PERGAMON

Pattern Recognition 32 (1999) 1631–1644

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

Size–biased random closed sets

M.N.M. van Lieshout*

CWI, Centre for Mathematics and Computer Science, P.O. Box 94079, 1090 GB Amsterdam, Netherlands

Received 18 September 1998; received in revised form 23 October 1998; accepted 23 October 1998

Abstract

We indicate how granulometries may be useful in the analysis of random sets. We define a suitable size distribution function as a tool in exploratory data analysis and give a new Hanisch style estimator for it. New Markov random sets are constructed which favour certain sizes above others. A size-biased random set model is fitted to a data set concerning the incidence of heather (Diggle, Biometrics 37 (1981) 531–539). © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Contact distribution function; Empty space function; Granulometry; Hanisch style estimator; Markov random set; Morphological opening and closing; Size distribution function

1. Introduction

One of the most basic properties of an object is its size. It is no wonder then that size measures have been used for a long time in the empirical sciences, and more recently in the analysis of (binary) images. For instance in classification problems, characteristics such as moments of the empirical size distribution may be used as features. See Vincent and Dougherty [1] for an overview as well as [2–4].

There are many other applications. To name a few, Serra [5] employs size distributions for shape and texture analysis, Maragos uses them for multiscale shape representation [6], Haralick et al. [7,8] apply size distributions to shape filtering and restoration problems, Sivakumar [9] gives applications in texture classification and morphological filtering, while texture synthesis and analysis are considered by Sivakumar and Goutsias [10].

In this paper we illustrate how size distribution functions may be used in stochastic geometry both as a descriptive tool in exploratory data analysis [11] and to build new random set models. Section 2 reviews the basic morphological operators [5,12,13] and shows how they can be used to define a granulometry to measure size, while Section 3 provides some background in stochastic geometry [14]. In Section 4 we introduce the size distribution function [13] of a stationary random closed set. Section 5 focuses on the estimation of the size distribution function, and proposes a new estimator in the spirit of Hanisch [15–17]. An application in exploratory data analysis is given in Section 6. New Markov random set models are constructed in Section 7 from a reference Boolean model by biasing towards certain sizes, generalising the discrete morphologically constrained Gibbs models of Sivakumar and Goutsias [10,18]. Finally, in Section 8 a size-biased random set model is fitted to a data set concerning the incidence of heather [19].

2. Morphological granulometries

Perhaps the oldest and most frequently used technique to quantify the size of solid particles in the empirical sciences is to use a series of sieves with varying mesh openings. Clearly, the particles that cannot pass through any given sieve are a subset of the total collection of

*Corresponding author. Tel. + 31-20-592-4008; fax: + 31-592-4199.

E-mail address: colette@cwi.nl (M.N.M. van Lieshout)

PII: S0031-3203(99)00025-4

particles; if the sieve is solid, no particle can pass through it, and if a larger number of particles are considered then the residual after sieving will be larger too. Moreover, if we sieve the particles successively with two different mesh sizes, the result will be the same as using only the one with the biggest mesh opening. These simple but essential features of sieving (as well as of other 'sizing' methods, see [5]) underly the following axiomatic definition of Matheron [13].

Definition 1. A family of operators $\psi_r : \mathscr{P}(\mathbb{R}^d) \to \mathscr{P}(\mathbb{R}^d)$ on the power set $\mathscr{P}(\mathbb{R}^d)$ of $\mathbb{R}^d$, indexed by $r \geqslant 0$ is a granulometry if for all $X \subseteq \mathbb{R}^d$

(G0) $\psi_0(X) = X$;

(G1) $\psi_r(X) \subseteq X$ for all $r \geqslant 0$;

(G2) if $Y \subseteq X$ then $\psi_r(Y) \subseteq \psi_r(X)$ for all $r \geqslant 0$;

(G3) $\psi_r(\psi_s(X)) = \psi_s(\psi_r(X)) = \psi_{\max(r,s)}(X)$ for all $r, s \geqslant 0$.

Condition (G3) is sometimes referred to as the *sieving condition*. Intuitively, $\psi_r(X)$ can be thought of as the subset of particles in X that remain after sieving with a mesh size $r \geqslant 0$.

In this paper, we are interested in a special class of granulometries based on Euclidean openings. Let B be a fixed subset of $\mathbb{R}^d$. Write $\check{B} = \{-b : b \in B\}$ for the reflection of B in the origin and use the subscript h for translation over the vector h. Then the *Minkowski addition* of a set $X \subseteq \mathbb{R}^d$ with *structuring element* B is defined by

$$X \oplus B = \{h \in \mathbb{R}^d : \check{B}_h \cap X \neq \emptyset\} \tag{2.1}$$

and similarly the *Minkowski subtraction* is given by

$$X \ominus B = \{h \in \mathbb{R}^d : B_h \subseteq X\}. \tag{2.2}$$

Seen as operators on $\mathscr{P}(\mathbb{R}^d)$, Eqs. (2.1) and (2.2) are referred to as *dilation* and *erosion*, respectively. The operators are dual, that is dilating the complement X^c of a set X amounts to eroding X itself: $(X^c \oplus B)^c = (X \ominus \check{B})$.

Compositions of Minkowski addition and subtraction define the *opening*

$$X \circ B = (X \ominus B) \oplus B = \bigcup \{B_h : h \in \mathbb{R}^d, B_h \subseteq X\} \tag{2.3}$$

and the *closing*

$$X \bullet B = (X \oplus B) \ominus B = (X^c \circ \check{B})^c. \tag{2.4}$$

For a comprehensive account on mathematical morphology, see [5,12,13].

From now on, we will assume that the structuring element B is non-empty, convex and compact. Let $rB = \{rb : b \in B\}$ $(r \geqslant 0)$ and set

$$\psi_r(X) = X \circ rB, \quad r \geqslant 0. \tag{2.5}$$

Then (2.5) is a granulometry [13]. Indeed, properties (G0)–(G2) follow directly from definition (2.3). The sieving condition is a consequence of the assumptions on the structuring element B, see Chap. 1-5 in [13] or p. 334 in [5].

In summary, openings with structuring elements of varying 'mesh' can be used to quantify the size of a set X (see Section 4). By duality, the size of the empty space X^c can be measured by the associated *anti-granulometry*

$$\psi_r(X^c)^c = (X^c \circ rB)^c = X \bullet r\check{B},$$

based on closings with structuring elements $r\check{B}$, $r \geqslant 0$.

3. Random sets and contact distributions

Stochastic geometry [14] is concerned with the study of random closed sets. Random sets arise in a variety of fields. Examples include samples of minerals in material science, microscopic sections of cells in cytology or vegetation maps such as Fig. 1 depicting the presence of heather [19].

Since a random closed set takes values in the family of all closed subsets of $\mathbb{R}^d$, its probability distribution is often untractable and lower-dimensional summary statistics are called for. Indeed, a statistical analysis usually starts with computing and plotting a few such statistics to 'get a feel for the data'. The plots may suggest a suitable parametric form for a probability model. Moreover, summary statistics may be used to estimate parameters and to perform a goodness-of-fit test. We will apply this approach to the data image of Fig. 1 in Sections 6 and 8. Details and more examples can be found in the textbooks [11,14].

Below we will assume that the random closed set X is stationary, i.e. its distribution is invariant under translations. In order to exclude pathological cases, we assume that for each $x \in \mathbb{R}^d$ the coverage probability $P(x \in X)$ is strictly between 0 and 1.

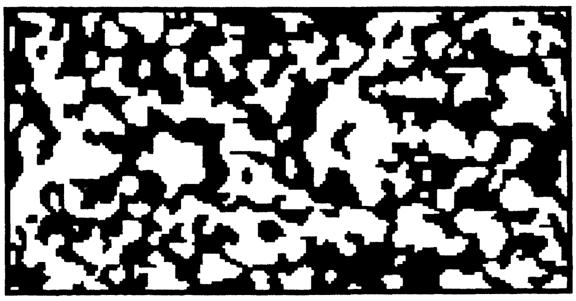


Fig. 1. Image heather [19].

A summary statistic commonly used to measure the 'size' of the pores in X^c is the *empty space function*

$$F_B(r) = P(x \in X \oplus r\check{B}) \tag{3.1}$$

or the related *contact distribution function*

$$H_B(r) = P(x \in X \oplus r\check{B} | x \notin X) = \frac{F_B(r) - F_B(0)}{1 - F_B(0)} \tag{3.2}$$

($r \geqslant 0$). Because of the stationarity of X, these definitions do not depend on the choice of x, and by the assumption on the coverage probabilities the conditioning in (3.2) is valid. Both (3.1) and (3.2) involve a structuring element B. Typical choices include balls and squares, although non-isotropic structuring elements may be preferred when investigating directional effects. If B is a compact convex set containing a neighbourhood of the origin, then $H_B(\cdot)$ is a distribution function but this is not the case for general B (see [14]). Note that $1 - H_B(r)$ can be interpreted as the conditional probability that a copy of rB placed at a test point 0 is entirely contained in the background given that the test point itself does not fall in X. A similar interpretation holds for the empty space function $F_B(r)$.

Definitions (3.1) and (3.2) can be seen as the stochastic counterparts of the granulometries defined in Section 2 with particle size measured by

$$\tilde{\psi}_r(X) = (X^c \oplus r\check{B})^c = X \ominus rB.$$

However, although under our restrictions on $B, \tilde{\psi}_r(\cdot)$ satisfies (G0)–(G2), in general the sieving condition does not hold. Thus, in the remainder of this paper, we will consider replacing the dilation in (3.1) and (3.2) by a closing to obtain a proper granulometry.

4. Size distribution functions

In mathematical morphology (see eg. [15]), the *joint size distribution law* of a stationary random closed set X is defined using granulometry (2.5) described in Section 2 as

$$G(r) = \begin{cases} 1 - P(x \in X \circ rB), & r \geqslant 0, \\ 1 - P(x \in X \bullet |r|\check{B}), & r < 0. \end{cases} \tag{4.1}$$

By the stationarity of $X, G(r)$ does not depend on the choice of $x \in \mathbb{R}^d$. However, from a probabilistic point of view, $G(\cdot)$ is not a proper distribution function, since it is semi-continuous from the left rather than from the right. This will prove to be undesirable in Section 7 where we define size-biased random sets by integrating with respect to size. For this reason, the following definition is preferred.

Definition 2. Let X be a stationary random closed set and B a non-empty convex compact structuring element. Define the *size distribution function* of X by

$$P_B(r) = \begin{cases} P(x \in X \bullet r\check{B}), & r \geqslant 0, \\ P(x \in X \circ |r|B), & r < 0. \end{cases} \tag{4.2}$$

It is easily verified that $P_B(0) = P(0 \in X) = p$, the coverage fraction of the stationary random closed set X. Moreover, having excluded degenerate cases where X is either the whole space or empty almost surely, $P_B(r)$ approaches 0 at $-\infty$ and tends to 1 for $r \to \infty$.

At this point it is important to note that there is some ambiguous terminology in the literature. For instance, the size distribution function $P_B(\cdot)$ is called the *granulométrie bidimensionelle* in metallurgy [11,20], in image processing the granulometries of Definition 1 are sometimes called size distributions [21], and the phrase *pattern spectrum* is used to denote the normalised empirical size distribution of a single realisation of X, see Maragos [6]. Thus, the pattern spectrum can be seen as an estimator of (4.2).

Lemma 1. *Let X be a stationary random closed set and B a non-empty convex compact structuring element. Then $P_B(\cdot)$ is well-defined and does not depend on the choice of $x \in \mathbb{R}^d$. Seen as a function of r, $P_B(\cdot)$ takes values in $[0, 1]$, is increasing and semi-continuous from the right.*

Proof. First note that since $B \neq \emptyset$ is compact, $X \oplus rB$, $X \ominus rB$ hence $X \circ rB$, $X \bullet rB (r \geqslant 0)$ are closed sets (see p. 19 in [13]). Since X is stationary, so are $X \circ rB$ and $X \bullet r\check{B}$ implying that $P_B(r)$ is well-defined as the coverage fraction of the random closed set $X \bullet r\check{B}$ (for $r \geqslant 0$) or $X \circ |r|B$ (for $r < 0$). In particular, $P_B(r)$ does not depend on the choice of x in (4.2).

To check that $P_B(\cdot)$ is increasing, first consider $r \geqslant s \geqslant 0$. Then using well-known properties of the opening,

$$X \bullet r\check{B} = (X^c \circ rB)^c = ((X^c \circ sB) \circ rB)^c \supseteq (X^c \circ sB)^c$$
$$= X \bullet s\check{B},$$

hence $P_B(r) \geqslant P_B(s)$. Also, for $r \leqslant s < 0, \psi_{|r|}(X) = \psi_{|r|}(\psi_{|s|}(X)) \subseteq \psi_{|s|}(X)$ since ψ_t is a granulometry (using the notation of (2.5)). Finally, $X \circ |r|B \subseteq X$ implies $P_B(r) \leqslant P_B(0)$.

The mapping $(r, X) \to X \bullet r\check{B}$ is upper semi-continuous by properties 1-5-1 and 1-5-2 in [13] and increasing in r. Hence as $r_n \downarrow r \geqslant 0, X \bullet r_n\check{B} \downarrow X \bullet r\check{B}$. Thus $P_B(\cdot)$ is right-continuous on $[0, \infty)$. Similarly the mapping $(r, X) \to X \circ rB$ is upper semi-continuous and decreasing in $r > 0$, hence $0 < r_n \uparrow r$ implies $X \circ r_n B \downarrow X \circ rB$. Thus, if $0 > s_n \downarrow s$, then $0 < -s_n \uparrow -s$, and $X \circ |s_n| B \downarrow X \circ |s| B$. We conclude that $P_B(\cdot)$ is right-continuous.

Finally, since $P_B(\cdot)$ is defined in terms of probabilities, it takes its values in $[0,1]$. □

Summarising, $P_B(\cdot)$ is a proper distribution function. Note that the empty-space function (3.1) is absolutely continuous except for an atom in 0 (see [17] or [16]) but that the size distribution function may have countably many discontinuities.

Explicit expressions for $P_B(r)$ are hard to find with the notable exception of the linear size distribution functions [13,17]. The contact distribution function $H_B(r)$ is available in closed form for Boolean models [14] (but not for most other random set models!) and in this case, assuming the primary grain is convex, depends only on the moments of a few functionals of the grain. For instance in $\mathbb{R}^2$ the mean grain perimeter determines $H_B(r)$, which, as pointed out by Ripley [11], may result in poor distinguishing power as an exploratory data analysis tool. We will return to this point in Section 8.

5. Estimation and edge effects

In this section we discuss estimating the size distribution function $P_B(r)$ of a stationary random closed set X (cf. Definition 2). As a first step, note that for any $r \geqslant 0$ and any $A \subseteq \mathbb{R}^d$ of positive volume $|A| > 0$,

$$\frac{E|(X \bullet r\check{B}) \cap A|}{|A|} = \frac{1}{|A|}\int_A P(a \in X \bullet r\check{B})\,\mathrm{d}a = P_B(r) \tag{5.1}$$

using the fact that X is stationary. Thus the volume fraction of $X \bullet r\check{B}$ (or $X \circ |r|B$ for negative r) in any set A of positive volume yields a pointwise unbiased estimator of $P_B(r)$.

In practice, however, a random set is not observed over the whole space, but within some compact window W of positive volume $|W|$ (typically a square or rectangle). Thus, due to edge effects caused by parts of X outside W, $X \bullet r\check{B}$ and $X \circ |r|B$ are not completely observable and the volume fraction estimator with $A = W$ may be biased.

To overcome this problem, a minus sampling estimator

$$\hat{P}_B(r, X) = \begin{cases} \dfrac{|(X \bullet r\check{B}) \cap (W \ominus (rB \oplus r\check{B}))|}{|W \ominus (rB \oplus r\check{B})|}, & r \geqslant 0, \\[2ex] \dfrac{|(X \circ |r|B) \cap (W \ominus (|r|B \oplus |r|\check{B}))|}{|W \ominus (|r|B \oplus |r|\check{B})|}, & r < 0, \end{cases} \tag{5.2}$$

has been proposed. This estimator is based on the *local knowledge principle* [5] for openings and closings stating that if the random set X is observed in the compact window W then $X \circ rB$ and $X \bullet r\check{B}$ ($r \geqslant 0$) are observable within $W \ominus (rB \oplus r\check{B})$. More specifically,

$$(X \bullet r\check{B}) \cap (W \ominus (rB \oplus r\check{B})) = ((X \cap W) \bullet r\check{B}) \cap (W \ominus (rB \oplus r\check{B}))$$

with a similar formula for the opening.

From (5.1) with $A = W \ominus (rB \oplus r\check{B})$ it follows easily that the minus sampling estimator in (5.2) is unbiased whenever $|W \ominus (rB \oplus r\check{B})| > 0$. However, as both numerator and denominator in (5.2) depend on r, there is no guarantee that the minus sampling estimator is monotone in r. Nor is all available information used.

To implement (5.2), one typically does not compute the areas but rather uses a grid $T = \{t_i\}$ of points in W and simply counts the number of t_i's falling in $X \bullet r\check{B}$, $X \circ rB$ and $W \ominus (rB \oplus r\check{B})$. The resulting discretised estimator is defined for ranges r for which there are grid points in $W \ominus (|r|B \oplus |r|\check{B})$; it is still pointwise unbiased and in general suffers from the same lack of monotonicity as its area-based counterpart.

For the empty space function (3.1), more refined estimators have been proposed recently. Hansen et al. [17] used survival analysis ideas for deriving a Kaplan–Meier-type estimator [22] for $F_B(r)$. Chiu and Stoyan [16] showed that the ideas underlying this Kaplan–Meier approach are very similar to those involved in the Hanisch estimator [15], originally proposed for certain point pattern statistics. In the remainder of this section, we will derive a Hanisch style estimator for the size distribution function (Definition 2).

To do so, we need three local size measures: with respect to X, the background X^c and the boundary. As before, let B be a non-empty convex compact structuring element and X a stationary random closed set observed in a compact observation window W. Set

$$\rho(x, X) = \begin{cases} \sup\{r \geqslant 0 : \exists h \text{ such that } x \in (rB)_h \subseteq X\}, & x \in X, \\ 0 & x \notin X, \end{cases} \tag{5.3}$$

$$\eta(x, X) = \begin{cases} \inf\{r \geqslant 0 : x \in X \bullet r\check{B}\}, & x \notin X, \\ 0 & x \in X, \end{cases} \tag{5.4}$$

$$\zeta(t, W^c) = \begin{cases} \inf\{r \geqslant 0 : (rB \oplus r\check{B})_t \cap W^c \neq \emptyset\}, & t \in W, \\ 0 & t \notin W. \end{cases} \tag{5.5}$$

It is easy to see that $X \circ rB = \{x \in X : \rho(x, X) \geqslant r\}$ $X \bullet r\check{B} = \{x \in R^d : \eta(x, X) \leqslant r\}$ and $W \ominus (rB \oplus r\check{B}) = \{x \in W : \zeta(x, W^c) \geqslant r\}$. Thus, $\rho(x, A)$ measures the B-size of a point x in A. Note that for $r > 0$, the restriction to $x \in X$ may be omitted. Similarly, $\eta(x, A)$ measures the B-size of voids at x left by A. We already saw that observed distances are occluded by the edges of the sampling window. Hence our final function $\zeta(x, W^c)$ measures the 'distance' from any point x in W to the window boundary.

The discretised minus sampling estimator can be expressed in terms of ρ, η and ζ as follows:

$$\hat{P}_B(r,X)=\begin{cases}\dfrac{\#\{i:\eta(t_i,X)\leqslant r\leqslant\zeta(t_i,W^c)\}}{\#\{i:\zeta(t_i,W^c)\geqslant r\}}, & r\geqslant 0,\\ 1-\dfrac{\#\{i:\rho(t_i,X)<|r|;\zeta(t_i,W^c)\geqslant|r|\}}{\#\{i:\zeta(t_i,W^c)\geqslant|r|\}}, & r<0.\end{cases}\tag{5.6}$$

Estimator (5.6) is pointwise unbiased for those r for which the number of sampling points t_i at least a distance $|r|$ away from W^c is positive, that is $\#\{i:\zeta(t_i,W^c)\geqslant|r|\}>0$.

Note that (5.6) does not use all information contained in the data. In particular, if $t_i\notin W\ominus(rB\oplus r\check{B})$, but $\eta(t_i,X)\leqslant\zeta(t_i,W^c)$ the correct void size at t_i is measured. Using this observation, one can define a Hanisch style estimator for $P_B(r)$ $(r\geqslant 0)$ by replacing the condition $\zeta(t_i,W^c)\geqslant r$ by $\zeta(t_i,W^c)\geqslant\eta(t_i,X)$ with a similar adaptation for $r<0$.

Definition 3. Let X be a realisation of a stationary random closed set observed in a compact window W. Then for all $r\geqslant 0$ with $\#\{i:\zeta(t_i,W^c)\geqslant r\}>0$, define

$$\hat{P}^H_B(r,X)=\sum_{s\leqslant r}\frac{\#\{i:\eta(t_i,X)=s\leqslant\zeta(t_i,W^c)\}}{\#\{i:\zeta(t_i,W^c)\geqslant s\}}\tag{5.7}$$

and for $r<0$ with $\#\{i:\zeta(t_i,W^c)\geqslant|r|\}>0$, let

$$P^H_B(r,X)=1-\sum_{s<|r|}\frac{\#\{i:\rho(t_i,X)=s\leqslant\zeta(t_i,W^c)\}}{\#\{i:\zeta(t_i,W^c)\geqslant s\}}.\tag{5.8}$$

As we saw before, one of the disadvantages of the minus sampling estimator is that it is not necessarily increasing in r. The Hanisch style estimator does not suffer from this disadvantage and is pointwise unbiased too.

Theorem 1. *Let X be a stationary random closed set, observed in a compact window W. Let B be a non-empty convex compact structuring element. Then the Hanisch style estimator in Definition 3 is pointwise unbiased for $P_B(r)$, increasing and semi-continuous from the right.*

Proof. First of all, consider the case $r\geqslant 0$. Then by stationarity of the random closed set X,

$$E\hat{P}^H_B(r,X)=E\left[\sum_{s\leqslant r}\frac{\#\{i:(t_i,X)=s\leqslant\zeta(t_i,W^c)\}}{\#\{i:\zeta(t_i,W^c)\geqslant s\}}\right]$$

$$=E\left[\sum_{t_i}\frac{1\{\eta(t_i,X)\leqslant r\}1\{t_i\in W\ominus(\eta(t_i,X)B\oplus\eta(t_i,X)\check{B})\}}{\#\{j:t_j\in W\ominus(\eta(t_i,X)B\oplus\eta(t_i,X)\check{B})\}}\right]$$

$$=\sum_{t_i}E\left[\frac{1\{\eta(t_i,X)\leqslant r\}1\{t_i\in W\ominus(\eta(t_i,X)B\oplus\eta(t_i,X)\check{B})\}}{\#\{j:t_j\in W\ominus(\eta(t_i,X)B\ominus\eta(t_i,X)\check{B})\}}\right].$$

Since the expectation depends on the random closed set X only through the random variables $\eta(t_i,X)$, which have probability distribution function $P_B(\cdot)$, we obtain

$$E\hat{P}^H_B(r,X)=\sum_{t_i}\int_{[0,r]}\frac{1\{t_i\in W\ominus(sB\oplus s\check{B})\}}{\#\{j:t_j\in W\ominus(sB\oplus s\check{B})\}}\,\mathrm{d}P_B(s)$$

$$=P_B(r).$$

Similarly by duality, for $r<0$,

$$E\hat{P}^H_B(r,X)=1-\sum_{t_i}E\left[\frac{1\{\rho(t_i,X)<|r|\}1\{t_i\in W\ominus(\rho(t_i,X)B\oplus\rho(t_i,X)\check{B})\}}{\#\{j:t_j\in W\ominus(\rho(t_i,X)B\oplus\rho(t_i,X)\check{B})\}}\right]$$

$$=1-\sum_{t_i}\int_{[0,|r|)}\frac{1\{t_i\in W\ominus(sB\oplus s\check{B})\}}{\#j:t_j\in W\ominus(sB\oplus s\check{B})\}}\,\mathrm{d}Q_B(s)$$

$$=P(\rho(0,X)\geqslant|r|)=P(0\in X\circ|r|B)=P_B(r),$$

writing $Q_B(\cdot)$ for the probability distribution of $\rho(0,X)$.

It is clear that both (5.7) and (5.8) are increasing and semi-continuous from the right. Furthermore,

$$\lim_{r\uparrow 0}\hat{P}^H_B(r,X)=1-\frac{\#\{i:\rho(t_i,X)=0\}}{\#\{i:t_i\in W\}}$$

$$=\frac{\#\{i:\rho(t_i,X)>0\}}{\#\{i:t_i\in W\}}\leqslant\frac{\#\{i:t_i\in X\}}{\#\{i:t_iW\}}$$

$$=\hat{P}^H_B(0,X).\quad\square$$

Thus the Hanisch style estimator preserves many properties of the size distribution function (cf. Lemma 1). However, in contrast to $P_B(r)$ itself, $\hat{P}^H_B(r,X)$ may be negative and exceed 1. If this is undesirable, one can take $R=\sup\{r>0:\#\{i:t_i\in W\ominus(rB\oplus r\check{B})\}>0\}$ and normalise the summands in (5.7) and (5.8) accordingly. The resulting estimator $\tilde{P}^H_B(r,X)$ is ratio-unbiased, for instance for $r\geqslant 0$

$$\frac{E\hat{P}^H_B(r,X)}{E\hat{P}^H_B(R,X)}=\frac{P_B(r)}{P_B(R)}.$$

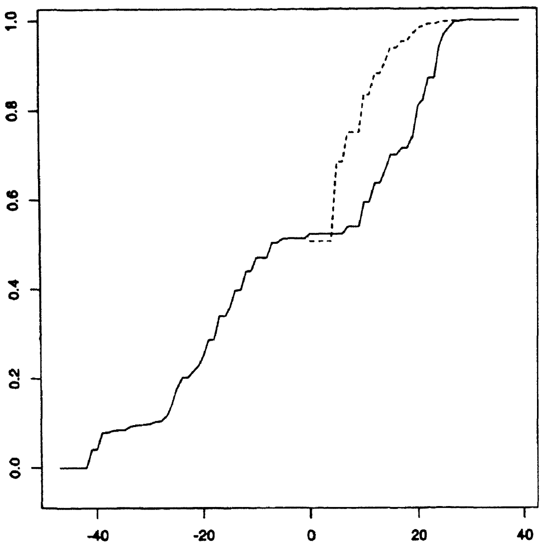
Similar techniques as in the proof of Theorem 1 can be employed to give an expression for the variance of $\hat{P}^H_B(r,X)$ $(r\geqslant 0)$ in terms of integrals with respect to the covariance measure

$$\gamma(t_i,t_k,s,t)=P(t_i\in X\bullet s\check{B};\,t_k\in X\bullet t\check{B})$$

(and similarly for $r<0$), but note that the variance depends on the choice of the grid T.

6. Exploratory data analysis

Usually, a statistical analysis of a binary image begins with plotting summary statistics such as the estimated



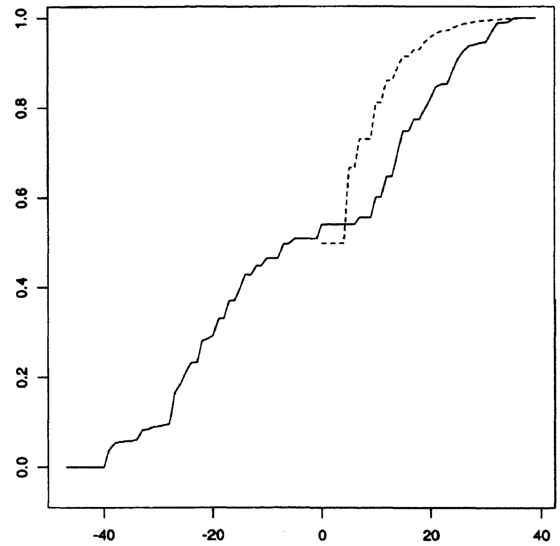


Fig. 2. Hanisch estimators of the size distribution function (solid line) and empty space function (broken line) for the left and right halves of heather.

empty space (3.1) or contact distribution function (3.2). Ripley [11] proposed to look at the plots of the normalised opening and closing transforms as well to get a better feel for the data. Here we present some further examples using the size distribution function (4.2).

Fig. 1 is a mapped pattern of heather (*Calluna vulgaris*), observed in a rectangle of 10 m × 20 m at Jädraås, Sweden. The data were collected by G. Ågren, T. Fagerström and P.J. Diggle and reproduced here with kind permission. The digitisation is on a 100 × 200 pixel grid, which is deemed a realistic reflection of the accuracy in the field. For further details see [19]. No apparent spatial inhomogeneity seems present, hence we may assume the random set X representing the area occupied by heather to be stationary.

In order to allow for cross validation, we will divide the data in to two equal square regions of side 10 m. Since bushes grow roughly spherical, a ball will be used as structuring element. To compute the distance measures ρ, η and ζ, the 5–7 chamfer metric-based algorithm of Nacken [21] was used to approximate the Euclidean metric, with distances calculated using the method of Borgefors [23]. The normalised Hanisch style estimators thus obtained are given in Fig. 2.

Note that the plots for the left and right halves of the field are similar, thus confirming the stationarity assumption. The coverage fraction is close to 1/2 and empty spaces measured by dilation are mostly smaller than 40 cm. The graph of the estimated empty space function lies above the graph of the size distribution function, reflecting the fact that for a centred ball B the inclusion relationship $X \bullet r\check{B} \subseteq X \oplus r\check{B}$ holds. The sizes of heather bushes and patches of background range up to about 80 cm.

7. Size-biased Markov random sets

Perhaps the best-known example of a random closed set is the *Boolean model* [13,14]. In this model, an ensemble random set is build from basic building blocks (the so-called *grains*) positioned at random locations (the *germs*) that are independently and uniformly scattered in space. More precisely, let W be a compact region in $\mathbb{R}^d$ of positive volume. We require that

- the number N of germs in W follow a Poisson distribution, that is

$$P(N = n) = e^{-\lambda|W|}\frac{(\lambda|W|)^n}{n!}, \quad n = 0,1,\ldots,$$

  the constant $\lambda > 0$ is called the intensity;
- given $N = n$, the germs $X_1, \ldots X_n$, are independent and uniformly distributed on W;
- at each germ $X_i = x_i$ a (random) grain K_i is placed independently of other grains according to some probability distribution $\mu(\cdot)$ on the family of non-empty compact sets.

Then the union $\cup_i(x_i \oplus K_i)$ is called a Boolean model. Note that the finite union of compact grains is indeed closed, as desired.

Other random closed set models may be obtained by weighting of a reference Boolean model (see [24] for technical details). In particular, a size distribution weight function can be used to bias towards certain sizes. The same idea underlies the *morphologically constrained random field models* on a discrete pixel grid, defined by Sivakumar and Goutsias [18] as

$$P(X) = \frac{1}{Z}\exp\left[- \sum_{i=0}^{I} \beta_i |X \circ iB \backslash X \circ (i+1)B| - \sum_{j=1}^{J} \gamma_j |X \bullet j\check{B} \backslash X \bullet (j-1)\check{B}| \right], \quad (7.1)$$

where X is the set of foreground pixels in a binary image W and for any $A \subseteq W$, $|A|$ denotes the number of pixels in A. The special case $\beta_i, \gamma_j \equiv 1$ had already been studied by Chen and Kelly [25]. It is worth noting that the exponent in (7.1) is based on volume fractions, thus failing to account for edge effects (cf. Section 5). Various extensions of morphologically constrained random fields have been considered; see [9,10] for details.

In the remainder of this section we will extend (7.1) to continuous random set models observed in a compact window W of positive volume $|W| > 0$.

Definition 4. A size-biased Markov random set is any random closed set whose density with respect to a Boolean model exists and is of the form

$$p(X) = \frac{1}{Z}\exp\left[- \int f(s)\, \mathrm{d}\hat{P}_B(s, X) \right], \quad (7.2)$$

where $f: \mathbb{R} \to \mathbb{R}$ is a bounded (measurable) function and $\hat{P}_B(\cdot, X)$ is an estimator of the size distribution function (4.2) based on X.

Some care has to be taken to ensure that (7.2) is well defined. In particular, the lack of monotonicity of the minus sampling estimator causes problems in defining the integral in the exponent of (7.2). However, by Theorem 1, for every X the Hanisch style estimator $\hat{P}_B^H(\cdot, X)$ can be normalised into a probability distribution function and hence $\int f(s)\, \mathrm{d}\hat{P}_B^H(s, X)$ is well defined for measurable functions $f(\cdot)$. By similar arguments it can be shown that if the naive volume fraction estimator based on (5.1) is taken for $\hat{P}_B(\cdot, X)$, the exponent in (7.2) is well defined. A sufficient condition to avoid explosion, that is to ensure that the normalising constant Z is finite, is *Ruelle stability*:

$$\int f(s)\, \mathrm{d}\hat{P}_B(s, X) > -a - bn$$

for some positive constants a and b (writing n for the number of grains in X). It follows that if $f(\cdot)$ is bounded in absolute value, $|f(s)| \leqslant F$, its integral with respect to the normalised Hanisch style estimator is bounded as well and hence (7.2) is well defined. Finally, the unnormalised Hanisch style estimator (5.7)–(5.8) gives rise to a signed measure, bounded in absolute value by $|T|$, the number of sampling points in the grid T. Hence, under the same condition $|f(s)| \leqslant F$, (7.2) is well defined.

Before turning to Markov properties of (7.2), we present some examples.

7.1. Morphologically smooth random sets

Let $f(s) = |W| \mathbf{1}_{(-1,0]}(s) \log \gamma$. Then, using the volume fraction estimator yields density

$$p(X) = \frac{1}{Z}\gamma^{-|X \backslash (X \circ B)|}, \quad (7.3)$$

generalising the Chen–Kelly model [25] for binary random fields. Note that for $\gamma > 1$, the most likely realisations X are open with respect to the structuring element B. Thus sets built of approximately convex components are favoured over those with thin or elongated pieces, sharp edges or small isolated clutter.

By duality, taking $f(s) = |W| \mathbf{1}_{(0,1]}(s) \log \gamma$ yields

$$p(X) = \frac{1}{Z}\gamma^{-|(X \bullet \check{B}) \backslash X|}, \quad (7.4)$$

favouring for $\gamma > 1$ sets that are approximately closed with respect to $\check{B}$ and discouraging small holes or rough edges.

Both models are well defined for $\gamma \leqslant 1$ too, for $\gamma < 1$ encouraging morphological roughness.

Note that in (7.3) and (7.4), $|\cdot|$ denotes Lebesgue measure restricted to W and hence $p(\cdot)$ is susceptible to edge effects. This can be alleviated by using the Hanisch style estimator. The resulting model influences the morphological smoothness of its realisations as described above.

7.2. Morphological area-interaction random sets

Let $f(s) = |W| \mathbf{1}\{s \leqslant -1\} \log \gamma$ and $\hat{P}_B(\cdot, X)$ the volume fraction estimator, yielding

$$p(X) = \frac{1}{Z}\gamma^{-|X \circ B|}.$$

If B were the empty set, $p(\cdot)$ would define an area-interaction random set [24,26]. Thus, for general B, $p(\cdot)$ can be seen as an opening-smoothed area-interaction model. By duality, $f(s) = -\mathbf{1}\{s > 1\} \log \gamma$,

$$p(X) = \frac{1}{Z}\gamma^{1 - |X \bullet \check{B}|/|W|},$$

defines a closing-smoothed area-interaction random set. Again, Hanisch style estimators may be employed to better account for edge effects.

Similar ideas may be used if the area measure in the exponent of γ is replaced by the Euler characteristic or other quermass integral [27], but note that some care has to be taken to ensure that the model does not explode. However, since the closing operator removes small holes, the closing-smoothed Euler-interaction model may be integrable when the non-smoothed version is not.

7.3. Size-symmetric random sets

Let f be the indicator function of $(-g, g]$, hence

$$p(X) = \frac{1}{Z}\exp[-\gamma\int_{-g}^{g} \mathrm{d}\hat{P}_B(s, X)]. \tag{7.5}$$

For $\gamma > 0$, particle and pore sizes exceeding g will be favoured, while for $\gamma < 0$ the sizes tend to be smaller than g.

The probabilistic model (7.2) (as well as (7.1)) involves a constant Z ensuring $p(\cdot)$ integrates to 1. Due to the high dimensionality of the model, a closed-form expression for Z is usually not available. Therefore from a practical point of view, it would be particularly convenient if small changes to the set X would affect $p(X)$ only locally. In that case, iterative algorithms can be designed that avoid Z and involve local computations only.

First, consider the case where both the germs and the grains of which the random closed set X is composed are fully observable and write $Y = \{(x_i, K_i) : i = 1, \dots, N\}$ for the collection of germ–grain pairs. Referring back to Fig. 1 this means that individual heather bushes would be observed instead of the area occupied by the heather. Writing $X = \cup_i(x_i \oplus K_i)$ for the union of all germs, (7.2) can serve as a probability density for Y with respect to a Poisson 'grain-marked germ' process [14].

Suppose that addition of a grain K at u is considered. Then the likelihood ratio

$$\frac{p(X \cup K_u)}{p(X)} = \exp\left[-\int f(s)\,\mathrm{d}\hat{P}_B(s, X \cup K_u) + \int f(s)\,\mathrm{d}\hat{P}_B(s,X)\right] \tag{7.6}$$

does not depend on the intractable constant Z (writing as before K_u for the translation of K over the vector u). If moreover $f(\cdot)$ is supported on $[-g, g]$, the log-likelihood ratio reduces to

$$-\int_{-g}^{g} f(s)\,\mathrm{d}\hat{P}_B(s, X \cup K_u) + \int_{-g}^{g} f(s)\,\mathrm{d}\hat{P}_B(s, X) \tag{7.7}$$

which involves only germ–grain pairs close' to (u, K).

More precisely, define a *neighbourhood relation* by

$$(u, K) \sim (v, L) \Leftrightarrow L_v \oplus (gB \oplus g\check{B}) \cap K_u \oplus (gB \oplus g\check{B}) \neq \emptyset, \tag{7.8}$$

where $u, v \in W$ and K, L are non-empty compact sets. In the following theorem we will show that (7.7) only depends on those $(x_i, K_i) \in Y$ that are $\sim$-neigbours of (u, K), i.e. $(x_i, K_i) \sim (u, K)$. In mathematical parlance, this local dependence property means that in the fully observable case Y is *Markov* [28,29] with respect to the neighbourhood relation $\sim$.

Theorem 2. *Let Y be a germ–grain model defined by its density* (7.2) *for some bounded function $f(\cdot)$ that is supported on $[-g, g](g > 0)$. Then Y is Markov with respect to $\sim$ for $\hat{P}_B(\cdot, X)$ either the naive estimator* (5.1) *or the Hanisch style estimator* (5.7) *and* (5.8).

Proof. We start with proving that if $x \notin K_u \oplus (gB \oplus g\check{B})$, then for all $s \leq g$, $x \in (X \cup K_u) \circ sB \Leftrightarrow x \in X \circ sB$ and $x \in (X \cup K_u) \bullet s\check{B} \Leftrightarrow x \in X \bullet s\check{B}$.

To see this, let $x \in (X \cup K_u) \circ sB$. Then $\exists h$ such that $x \in (sB)_h \subseteq X \cup K_u$, and, in particular, we can write $x = h + sb$ for some $b \in B$. Now if $(sB)_h \cap K_u \neq \emptyset$, then $h + sb' \in K_u$ for some $b' \in B$ and hence $x = h + sb = h + sb' + (sb - sb') \in K_u \oplus (sB \oplus s\check{B}) \subseteq K_u \oplus (gB \oplus g\check{B})$ using the convexity of the structuring element B. This contradicts the assumption that $x \notin K_u \oplus (gB \oplus g\check{B})$ and hence $x \in (sB)_h \subseteq X$, that is $x \in X \circ sB$.

Similarly for the closing, let $x \in (X \cup K_u) \bullet s\check{B}$. Then by duality $x \notin (X \cup K_u)^c \circ sB$ and hence for any h such that $x \in (sB)_h$ the intersection $(sB)_h \cap (X \cup K_u)$ must be non-empty. By the previous argument, K_u cannot be intersected and hence $(sB)_h \cap X \neq \emptyset$. Thus $x \in X \bullet s\check{B}$.

Secondly, for $x \notin K_u \oplus (gB \oplus g\check{B})$, suppose that $\eta(x, X \cup K_u) \leq g$. Then $x \in (X \cup K_u) \bullet g\check{B}$, hence by the above $x \in X \bullet g\check{B}$ or equivalently $\eta(x, X) \leq g$. Also $\eta(x, X \cup K_u) = \inf\{s \geq 0 : x \in (X \cup K_u) \bullet s\check{B} = \inf\{0 \leq s \leq g : x \in (X \cup K_u) \bullet s\check{B}\} = \inf\{0 \leq s \leq g : x \in X \bullet s\check{B} = \eta(x, X)$. Dually, suppose that $\rho(x, X) < g$. Then by $x \notin X \circ gB$ hence by the above $x \notin (X \cup K_u) \circ gB$ or $\rho(x, X) < g$. Also $\rho(x, X) = \sup\{s \geq 0 : x \in X \circ sB\} = \sup\{0 \leq s \leq g : x \in X \circ sB\} = \sup\{0 \leq s \leq g : x \in (X \cup K_u) \circ sB\} = \rho(x, X \cup K_u)$.

Hence for $s \in [-g, g]$, $\hat{P}(s, X \cup K_u) - \hat{P}(s, X)$ depends only on grid points $t_i \in K_u \oplus (gB \oplus g\check{B})$. By the local knowledge principle [5]

$$|(X \bullet g\check{B}) \cap (A \ominus (gB \oplus g\check{B}))| = |((X \cap A) \bullet g\check{B}) \cap (A \ominus (gB \oplus g\check{B}))|$$

(similarly for openings) with $A = K_u \oplus (gB \oplus g\check{B}) \oplus (gB \oplus g\check{B})$ and noting that $A \ominus (gB \oplus g\check{B}) \supseteq K_u \oplus (gB \oplus g\check{B})$ only knowledge of Y in so far as it intersects A is needed. The result follows. □

If grains are not individually observable, note that if $X = X_1 \cup \cdots \cup X_k$ is partitioned into its connected components $X_1, \cdots, X_k$, the opening $X \circ B = \cup\{B_h \subseteq X\}$

also partitions, as the convexity of B implies that B_h must fall entirely in one of the X_i. Thus, both the volume fraction (5.1) and Hanisch style (5.8) estimators satisfy $\hat{P}_B(s, X) = \sum_{i=1}^{k}[\hat{P}_B(s, X_i)]$ for $s < 0$. Similarly, $X \bullet \check{B}$ partitions over the connected components of $W \backslash X$ and hence

$$p(X) = \prod_{i=1}^{k} \phi(X_i) \prod_{i=1}^{l} \phi(X_i^c).$$

Thus, altering X will only affect the connected components that are modified, a state-dependent Markov property as introduced by Baddeley and Møller [29]. See also [24,30].

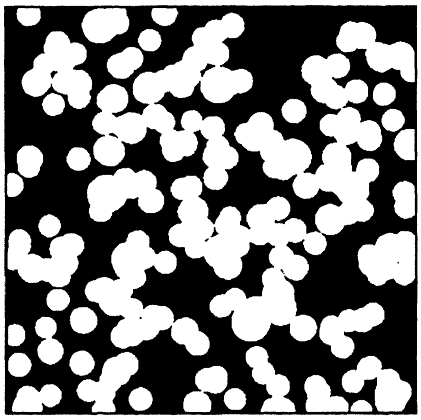


Fig. 3. Realisation of Boolean model with circular grains.

8. Application

In Section 6, a binary image of a species of heather was considered (cf. Fig. 1). It is well known to biologists that Calluna vulgaris grows from seedlings into roughly circular bushes, reaching a maximum radius of about 50 cm in some 20–25 years. As they grow, the branches intermingle and overlapping areas of the field are occupied. Thus it is that in maps such as Fig. 1 no individual bushes can be observed. The particular field depicted in Fig. 1 is 25 years old.

The above considerations motivated Diggle [19] to fit a Boolean model (see Section 7) with circular grains. The radius distribution was taken to be of the form $c + W$ where c is a fixed constant and W a Weibull distributed random variable. Although the model is plausible on biological grounds and fits well using a test based on the empty space distribution function (3.1), it was found that realisations of the fitted model contained more isolated patches than the data (cf. Figs. 1–3). Further discussion can be found in Ripley [11], who compared the graphs of normalised opening and closing transforms for data and fitted model. Hall [31] discusses counting methods, Cressie and Laslett [32] estimate the mean number of heather bushes per unit area using marker points and Baddeley and Gill [33] consider a statistic derived from the empty space distribution.

Below we will fit a size-biased Markov random set model (Definition 4). In order to suggest a weight function $f(\cdot)$, the plots of the estimated size distribution function $\tilde{P}_B^H(r, X)$ for the data are compared to that of a realisation of the Boolean model fitted by Diggle [19]. Identifying the left half-field with the unit square [19], found the following estimates: the intensity is $\lambda = 221$, the minimal radius 0.0281 and the Weibull parameters are (0.8471, 355.2). A typical realisation is given in Fig. 3. The normalised Hanisch style estimators of the size distribution function are given in Fig. 4 (broken lines), using again a ball for the structuring element B. In order to assess the variability, upper and lower envelopes based on 19 independent realisations are plotted as well (corresponding to a significance level of 5%).

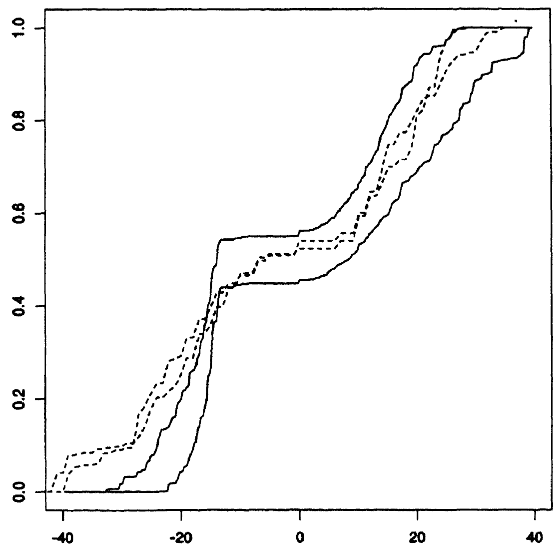


Fig. 4. Upper and lower envelopes of the estimated size distribution function based on 19 simulations of a Boolean model (solid lines) compared to those for the left and right halves of heather (broken lines).

The minimal grain size shows up clearly in the flat pieces of the envelopes. Moreover, the data curves for r less than about -40 cm lie above the upper simulation envelope, indicating that the fitted model has too few large grains. Thus, a goodness-of-fit test based on Euclidean openings would reject the Boolean model. A possible explanation of the greater discriminating power of size distribution functions compared to the empty space function (3.1) is that for Boolean models with circular grains, $F_B(r)$ is a function of the mean grain area and perimeter only (see Section 3).

For larger positive r, the empirical curve of the left half of Fig. 1 is near the upper envelope, suggesting that the pores are somewhat too large as well (although since parameters were estimated from the left half of the data, to determine the fit the estimated size distribution function of the right half should be considered). Hence, to favour larger particles and smaller pores, consider the size-biased model

$$p(X) = \frac{1}{Z}\exp[-\gamma_1 + \gamma_1 \hat{P}_B^H(-g, X) + \gamma_2 \hat{P}_B^H(g, X)] \tag{8.1}$$

parametrised by $\gamma_1, \gamma_2 \geqslant 0$. The Boolean model corresponds to $\gamma_1 = \gamma_2 = 0$.

In contrast to [19] which used a least-squared-error approach, we will estimate the parameter vector $\gamma = (\gamma_1, \gamma_2)$ by the maximum likelihood approach outlined by Geyer [34]. Writing $h_\gamma(X) = Zp(X)$, the log likelihood $l(\gamma)$ with respect to a fixed reference value $\psi \geqslant 0$ equals

$$l(\gamma) = \log\frac{h_\gamma(X)}{h_\psi(X)} - \log\frac{Z(\gamma)}{Z(\psi)}. \tag{8.2}$$

Here, the fixed ψ terms do not affect the modes of $l(\gamma)$ and a maximum likelihood estimator can, in principle, be found by optimising (8.2) over γ. A complication is that the term $\log Z(\gamma)/Z(\psi)$ is not known in closed form. However, since Z is a normalising constant, $Z(\gamma)/Z(\psi) = E_\psi h_\gamma(X)/h_\psi(X)$ and therefore the log likelihood $l(\gamma)$ can be approximated by

$$l_n(\gamma) = \log\frac{h_\gamma(X)}{h_\psi(X)} - \log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{h_\gamma(X_i)}{h_\psi(X_i)}\right), \tag{8.3}$$

where $X_1, \ldots X_n$ are samples from (8.1) under parameter value ψ. Thus, given a data image $X = x$ and samples $x_1, \ldots, x_n$, $\hat{\gamma}$ is obtained by optimising (8.3) with respect to γ. Provided the covariance matrix of $(\hat{P}_B^H(-g, X) - 1, \hat{P}_B^H(g,X))$ is non-singular, any solution to the maximum likelihood equations is necessarily unique. For details see [34].

The simulations needed in (8.3) are performed using the Metropolis–Hastings sampler of Geyer and Møller [35], a special case of Green's reversible jump technique [36]. This is an iterative procedure based on successive additions and deletions of a grain (see the discussion preceding Theorem 2). Given an initial set of germ-grain pairs $Y_0 = \{(x_i, K_i)\}$ with associated set $X_0 = \bigcup_i (x_i \oplus K_i)$, with probability 1/2 propose adding a grain ('birth'); with probability 1/2 propose deleting one of the grains in Y_0 if any ('death'). A new germ is proposed uniformly with a grain drawn from $\mu(\cdot)$. To ensure the correct distribution in the long run, the proposal (u, K) is accepted with probability

$$\min\left\{1, \frac{p(X_0 \cup K_u)}{p(X_0)}\frac{\lambda|W|}{n(Y_0)+1}\right\},$$

where $n(Y_0)$ denotes the number of elements of Y_0. If the new grain K at u is accepted, set $Y_1 = Y_0 \cup \{(u, K)\}$, $X_1 = X_0 \cup K_u$; otherwise $Y_1 = Y_0$ and $X_1 = X_0$. Similarly, select grain K_i at x_i for deletion from Y_0 with probability $1/n(Y_0)$. The proposal is accepted with probability

$$\min\left\{1, \frac{p(X_0 \setminus (K_i)_{x_i})}{p(X_0)}\frac{n(Y_0)}{\lambda|W|}\right\}$$

and if so, $Y_1 = Y_0 \setminus \{(x_i, K_i)\}$, $X_1 = X_0 \setminus (K_i)_{x_i}$. Otherwise $Y_1 = Y_0$ and $X_1 = X_0$. Continuing in this fashion, we obtain a sequence $X_k, k \in \mathbb{N}_0$, whose distribution converges in total variation to $p(\cdot)$ as $k \to \infty$ (from almost all initial states). For details see [38,39].

We applied this strategy to the model (8.1). As observed above, for the fitted Boolean model, $\hat{P}_B^H(r, X)$ is significantly too small for r less than -40 cm. Recalling that a pixel represents a square of side 10 cm, to favour larger particles, we choose $g = 5$ pixels. The reference vector ψ was taken to be (100, 100). In order to check whether the Metropolis–Hastings sampler has converged, time series of the sufficient statistics $\hat{P}_B^H(g, X)$ and $\hat{P}_B^H(-g, X) - 1$ are plotted in Fig. 5 over 200 000 iterations. The sampler appears to be mixing, thus there seems no reason to doubt convergence.

As for the parameters, the maximum likelihood approach outlined above was used, deleting the first 100 000 samples to allow for burn-in of the Metropolis–Hastings chain. The estimates are $\hat{\gamma}_1 = 117$ and $\hat{\gamma}_2 = 84$. Thus, the sizes of heather bushes are influenced more strongly than those of the background spaces. Typical samples (with the same resolution as the data) are given in Fig. 6. Note that the similarity to Fig. 1 seems greater than for the Boolean model, and that, in particular, the tendency of bushes to intermingle has increased.

To assess the goodness of fit, Fig. 7 plots the normalised Hanisch estimator $\tilde{P}_B^H(r, X)$ for the right half of the heather data (broken line) and the envelopes based on 19 (dependent) samples of the fitted size-biased random set model (8.1) taken every 5000 steps after a burn-in time of 100 000 steps. In comparison to Fig. 4, note that for $r > 0$ (as for the reference Boolean model) the estimated size distribution of the right half of the data lies within the simulation envelopes, but that for $r < 0$ the fit is much better than for a Boolean model.

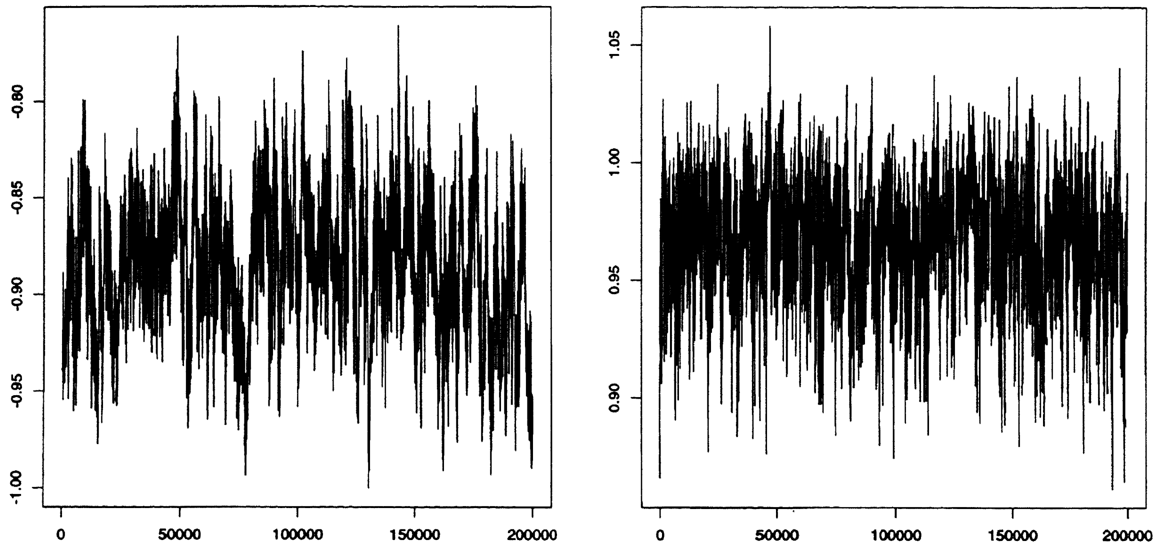


Fig. 5. Time series of the sufficient statistics $\hat{P}_B^H(-g, X) - 1$ (left) and $\hat{P}_B^H(g, X)$ (right) over 200 000 Metropolis-Hastings steps of (8.1) with $g = 5$ pixels and $\gamma_1 = \gamma_2 = 100$.

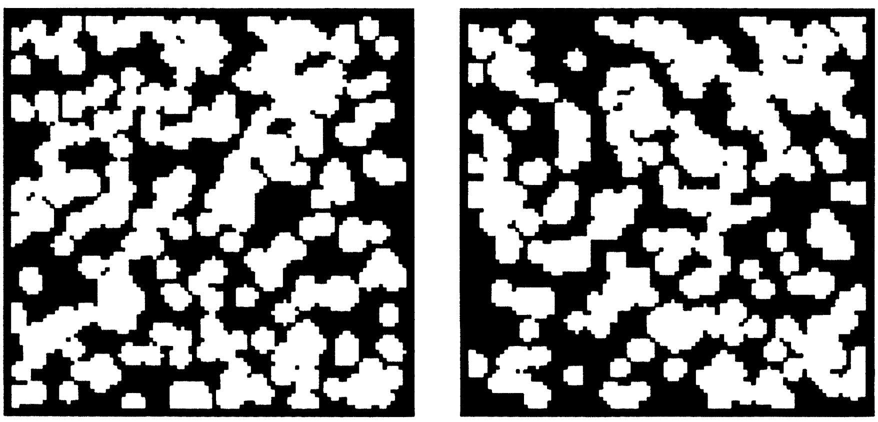


Fig. 6. Realisations of (8.1) with $g = 5$ pixels and $\gamma_1 = 117$, $\gamma_2 = 842$ after 100 000 (left) and 130 000 (right) Metropolis–Hastings steps.

Nevertheless, further improvements may be obtained by relaxing the minimal grain size assumption and finer discretisation.

9. Conclusions

In stochastic geometry, one of the trends in recent years has been the development of Markov models for simple point and object patterns [28,29,37] and the development of the computational tools to deal with them [35,36,38]. Some models can be adapted to random closed sets [24], but nevertheless explicit random set models defined in likelihood terms are scarce [39].

As a step in the direction of the development of more flexible and genuinely set-based models, this paper introduced size-biased Markov random set models, generalising the discrete morphologically constrained random fields [18]. These were obtained by biasing a Boolean

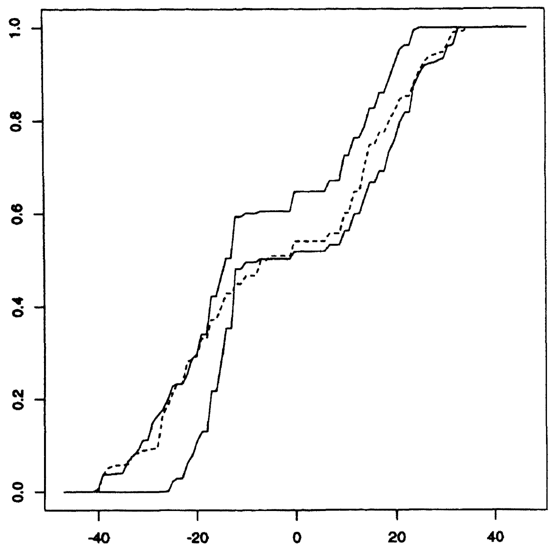


Fig. 7. Upper and lower envelopes of the estimated size distribution function based on 19 simulations of (8.1) with $g = 5$ pixels and $\gamma_1 = 117, \gamma_2 = 84$ (solid line) compared to that of the right half of heather (broken line).

model towards certain sizes, and include smoothed versions of quermass-interaction random sets [24,27]. It was shown that the models are well defined and under mild regularity conditions satisfy a Markov property.

From a statistical point of view, inference for random sets is predominantly non-parametric in nature, based on summary statistics such as the empty space function and on least squares or method of moments techniques (see [14] for an overview). In [34], Geyer advocates likelihood-based inference in the context of spatial point processes. In this paper, we show that Monte Carlo maximum likelihood techniques are also feasible for random set models.

It should be noted that exact simulation of size-biased Markov random set models based on the Propp–Wilson coupling from the past idea [40] is theoretically possible [38]. However, since tight upper and lower bounds on the likelihood ratio based on the current state of the algorithm would have to be computed at every iteration, we preferred to use the computationally easier (but only asymptotically exact) Metropolis–Hastings method in fitting a size-biased random set to a binary image.

Finally, it is important to recognise that most problems involving spatial data are hampered by edge effects due to parts of the image extending beyond the observation window. Thus, a Hanisch style estimator for the size distribution function of stationary random closed sets was developed and shown to be unbiased.

10. Summary

In this paper, size distribution functions were considered, both as an exploratory tool and as an ingredient for modelling random closed sets.

After a brief review of basic notions from mathematical morphology, a size distribution function $P_B(r)$, $r \in \mathbb{R}$, was defined for stationary random closed sets and compared to other summary statistics such as the empty space function and the contact distribution function. A Hanisch style estimator for $P_B(r)$ based on a partial observation of the random set within a bounded sampling window was derived. We proved that this estimator is pointwise unbiased, monotonically increasing and semi-continuous from the right implying that it can be normalised into a probability distribution function, in contrast to the commonly used minus sampling estimator.

Random set models were constructed by biasing a reference Boolean model towards certain sizes. Examples include morphologically smooth random sets and morphologically constrained area-interaction models. It was shown that under mild conditions the resulting models are well defined and possess a Markov property. From a practical point of view, Markov properties are useful in considerably reducing the computational burden inherent in working with random sets, and allow for likelihood-based inference using Markov chain Monte Carlo techniques.

As an illustration, a mapped pattern of heather was analysed. Fitting a Boolean model (representing spatial independence between the individual bushes) resulted in an underrepresentation of larger bushes. Thus, a size-biased model was suggested which gave a better fit.

Acknowledgements

The author is grateful to Peter Diggle for providing the heather image, to Adri Steenbeek for programming assistance and to Henk Heijmans and anonymous referees for valuable comments. This research was partially carried out while the author was at the University of Warwick.

References

[1] L. Vincent, E.R. Dougherty, Morphological segmentation for textures and particles, in: E.R. Dougherty (Ed.), Digital Image Processing Methods, Marcel Dekker, New York, 1994.

[2] Y. Chen, E.R. Dougherty, Gray-scale morphological granulometric texture classification, Opt. Eng. 33 (1994) 2713–2732.

[3] E.R. Dougherty, J.T. Newell, J.B. Pelz, Morphological texture-based maximum likelihood pixel classification based on local granulometric moments, Pattern Recognition 25 (1992) 1181–1198.

[4] E.R. Dougherty, F. Sand, Representations of linear granulometric moments for deterministic and random binary Euclidean images, J. Visual Commun. Image Representation 6 (1995) 69–79.

[5] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, London, 1982.

[6] P. Maragos, Pattern spectrum and multiscale shape representation, IEEE Trans. Pattern Ana. Mach. Intell. 11 (1989) 701–716.

[7] R.M. Haralick, P.J. Katz, E.R. Dougherty, Model based morphology: the opening spectrum, Graphical Models Image Process. 57 (1995) 1–12.

[8] D. Schonfeld, J. Goutsias, Optimal morphological pattern restoration from noisy binary images, IEE Trans. Pattern Anal. Mach. Intell. 13 (1991) 14–29.

[9] K. Sivakumar, Morphological analysis of random fields: theory and applications. Ph.D. Thesis, Johns Hopkins University, 1997.

[10] K. Sivakumar, J. Goutsias, Morphologically constrained Gibbs random fields: applications to texture synthesis and analysis, Report JHU/ECE 97-11, Johns Hopkins University, 1997.

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

[12] H.J.A.M. Heijmans, Morphological Image Operators, Academic Press, Boston, 1994.

[13] G. Matheron, Random Sets and Integral Geometry, Wiley and Sons, New York, 1975.

[14] D. Stoyan, W.S. Kendall, J. Mecke, Stochastic Geometry and its Applications, second ed., Akademie-Verlag, Berlin, 1995.

[15] K.H. Hanisch, Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point processes, Math. Operationsforsch. Statist. Ser. Statistik 15 (1984) 409–412.

[16] S.N. Chiu, D. Stoyan, Estimators of distance distributions for spatial patterns, Statist. Neerlandica 52 (1998) pp 239–246.

[17] M.B. Hansen, R.D. Gill, A.J. Baddeley, Kaplan-Meier type estimators for linear contact distributions, Scand. J. Statist. 23 (1996) 129–155.

[18] K. Sivakumar, J. Goutsias, Morphologically constrained discrete random sets, in: D. Jeulin (Ed.), Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets, World Scientific Publishing, Singapore, 1997, pp. 49–66.

[19] P.J. Diggle, Binary mosaics and the spatial pattern of heather, Biometrics 37 (1981) 531–539.

[20] A.G. Fabbri, Image Processing of Geological Data, Van Nostrand Reinhold, New York, 1984.

[21] P.F.M. Nacken, Image analysis methods based on hierarchies of graphs and multi-scale mathematical morphology, Ph.D. Thesis, University of Amsterdam, 1994.

[22] A.J. Baddeley, R.D. Gill, Kaplan-Meier estimators for interpoint distance distributions of spatial point processes, Ann. of Statist. 25 (1997) 263–292.

[23] G. Borgefors, Distance transformations in digital images, Comput. Vision Graphics Image Process. 34 (1986) 344–371.

[24] M.N.M. van Lieshout, On likelihoods for Markov random sets and Boolean models, in: D. Jeulin (Ed.), Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets, World Scientific Publishing, Singapore, 1997, pp. 121–135.

[25] F. Chen, P.A. Kelly, Algorithms for generating and segmenting morphologically smooth binary images, Proceedings of the Twenty Six Conference on Information Sciences, 1992, pp. 902.

[26] A.J. Baddeley, M.N.M. van Lieshout, Area-interaction point processes, Ann. Inst. of Statis. Math. 47 (1995) 601–619.

[27] W.S. Kendall, M.N.M. van Lieshout, A.J. Baddeley, Quermass-interaction models: conditions for stability, Adv. Appl. Probab. (SGSA) 31 (1999) in press.

[28] B.D. Ripley, F.P. Kelly, Markov point processes, J. London Math. Soc. 15 (1977) 188–192.

[29] A.J. Baddeley, J. Møller, Nearest-neighbour Markov point processes and random sets, Int. Statist. Rev. 57 (1989) 89–121.

[30] J. Møller, R. Waagepetersen, Markov connected component fields, Research Report 96-2009, Department of Mathematics and Computer Science, Aalborg University, 1996.

[31] P. Hall, Introduction to the Theory of Coverage Processes, Wiley, New York, 1988.

[32] N.A.C. Cressie, Statistics for Spatial Data, Wiley, New York, 1991.

[33] A.J. Baddeley, R.D Gill, The empty space hazard of a spatial pattern, Research Report 1994/3, Department of Mathematics, University of Western Australia, 1994.

[34] C.J. Geyer, Likelihood inference for spatial point processes, in: O. Barndorff-Nielsen, W.S. Kendall, M.N.M. van Lieshout (Eds.), Proceedings Seminaire Européen de Statistique, Stochastic Geometry, Likelihood, and Computation, Chapman & Hall, London, 1998.

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

[36] P.J. Green, Reversible jump MCMC computation and Bayesian model determination, Biometrika 82 (1995) 711–732.

[37] M.N.M. van Lieshout, Markov Point Processes and their Applications, Imperial College Press, London, 1999, in press.

[38] W.S. Kendall, J. Møller, Perfect Metropolis–Hastings simulation of locally stable spatial point processes, in preparation.

[39] I.S. Molchanov, Random closed sets: results and problems, O. Barndorff-Nielsen, W.S. Kendall, M.N.M. van Lieshout (Eds.), Proceedings Seminaire Européen de Statistique, Stochastic Geometry, Likelihood, and Computation, Chapman and Hall, London, 1998.

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

About the Author — MARIE-COLETTE VAN LIESHOUT received her M.Sc. and Ph.D. degrees in mathematics from Free University Amsterdam, the Netherlands, in 1990 and 1994, respectively. She is currently a senior researcher at the Centre for Mathematics and Computer Science (CWI) in Amsterdam. Previously she was employed as a junior researcher at Free University Amsterdam and at CWI, and as a lecturer in the Department of Statistics of the University of Warwick, United Kingdom. Her current research interests include stochastic geometry, image analysis, spatial statistics and Markov chain Monte Carlo techniques.