2000

Smoothing techniques and estimation methods for nonstationary Boolean models with applications to coverage processes

I. S. Molchanov
University of Glasgow, ilya@stats.gla.ac.uk

S. N. Chiu
Hong Kong Baptist University, snchiu@hkbu.edu.hk

This document is the authors' final version of the published article.
Link to published article: https://doi.org/10.1093/biomet/87.2.265

APA Citation
Smoothing techniques and estimation methods for nonstationary Boolean models with applications to coverage processes

BY I. S. MOLCHANOV

Department of Statistics, University of Glasgow, Glasgow G12 8QW, UK
ilya@stats.gla.ac.uk

S. N. CHIU

Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong
snciu@math.hkbu.edu.hk

Summary

Kernel smoothing methods are applied to nonparametric estimation for nonstationary Boolean models. In many applications only exposed tangent points of the models are observable rather than full realisations. Several methods are developed for estimating the distribution of the underlying Boolean model from observation of the exposed tangent points. In particular, estimation methods for coverage processes are studied in detail and applied to neurobiological data.

Some key words: Coverage; Johnson–Mehl model; Kernel smoothing; Nonstationary Boolean model
1 Introduction

Since Diggle’s (1985) pioneering paper smoothing methods have been used to estimate the intensity of a spatial point process. The key idea is to consider a realisation of a nonstationary point process to be a possibly infinite sample and to apply well-known kernel smoothing methods (Bowman & Azzalini, 1997; Silverman, 1986) to estimate the corresponding unnormalised density function. The latter provides an estimator for the intensity of the point process. Kernel estimators are also often used to estimate densities of higher-order moment measures of stationary point processes; see Heinrich (1988) and Jolivet (1984, 1991).

A relatively simple situation within this framework arises if the underlying point process is an inhomogeneous Poisson process. Its complete independence property ensures that the points can be interpreted as a sample of independent observations from an unnormalised density function. This allows us to use results from kernel smoothing in order to assess asymptotic properties of the corresponding kernel estimator of the intensity.

Points of a Poisson process in $\mathbb{R}^d$ can be interpreted as germs, and, if germs are marked by a sequence of random closed sets, this construction leads to the famous Boolean model; see Matheron (1975, § 3.3), Molchanov (1997) and Stoyan, Kendall & Mecke (1995, Ch. 3). More formally, let $\Pi$ be a Poisson point process in $\mathbb{R}^d$ with intensity measure $\Lambda$; the latter is assumed to be locally finite, so that the number of Poisson points in any bounded set is finite. Every point $x_i$ of $\Pi$ is marked by an independent and identically distributed realisation $\Xi_i$ of a random closed set $\Xi_0$ called the typical grain. The random set

$$\Xi = \bigcup_{x_i \in \Pi} (x_i + \Xi_i)$$  \hspace{1cm} (1)
is called a Boolean model. It is a non-trivial random closed set if

\[ E[\Lambda\{\Xi_0 \oplus B_r(x)\}] < \infty, \quad x \in \mathbb{R}^d, \quad r > 0, \quad (2) \]

where \( B_r(x) \) is the ball of radius \( r \) centred at \( x \in \mathbb{R}^d \) and \( \oplus \) denotes the Minkowski addition, so that \( A \oplus B = \{x + y : x \in A, y \in B\} \). Note that (2) will be assumed throughout the paper. For example, (2) holds if \( \Xi_0 \) is almost surely contained in a bounded deterministic set. As a result of overlapping, some germs and parts of the grains' boundaries are no longer observable from a realisation of \( \Xi \).

Assume that the typical grain \( \Xi_0 \) is almost surely convex. If a direction \( u \) in \( \mathbb{R}^d \) is fixed, then the tangent point of each grain \( \Xi_i \) is defined to be the lexicographical minimum among all points at which a hyperplane orthogonal to \( u \) and moving in the direction of \( u \) first touches \( \Xi_i \); see Molchanov & Stoyan (1994). Some of these tangent points are covered by other grains while other points are visible. These exposed, or observable, tangent points form a point process, which can be obtained by means of dependent thinning and possibly shifting of \( \Pi \). Figure 1 shows three grains and their tangent points, two of them being exposed and one covered by an overlapping grain.

[Figure 1 about here]

If the underlying Poisson process of germs is stationary, then \( \Lambda \) is proportional to Lebesgue measure and its density \( \lambda \), being a constant, is related to the intensity \( N^+ \) of the exposed tangent points process and the volume fraction \( p = \text{pr}(o \in \Xi) \), where \( o \) denotes the origin, by

\[ \lambda(1 - p) = N^+. \quad (3) \]

This leads to a simple method for estimating of \( \lambda \); see Molchanov & Stoyan (1994). Despite the loss of information caused by overlapping, a realisation of a stationary Boolean model provides enough information to estimate the density \( \lambda \) and the distribution of \( \Xi_0 \) (Molchanov, 1995).
This paper aims to study similar estimation problems for nonstationary and/or not fully observable Boolean models using the corresponding exposed tangent points process. The key idea is to use kernel smoothing methods in order to estimate the intensity of the exposed tangent points and to combine it with an estimator for the coverage probabilities in a formula that generalises (3). This method requires the full observation of a realisedation of the Boolean model. In many applications, however, only the exposed tangent points process is observable. Here we propose three different estimation methods, namely, the integral equation method, the reconstruction method and the maximum likelihood approach. For Boolean models which are stationary with respect to some coordinates, the integral equation method may be used even if only a projection of exposed tangent points is observed. Each of these three methods requires certain knowledge about or assumptions about the typical grain. This is not unreasonable since exposed tangent points alone are not sufficient to tell us anything about the typical grain.

A particularly important case in which our conditions on the typical grain are fulfilled is that of the Johnson–Mehl coverage processes, which are Boolean models defined on product spaces $\mathbb{R}^d \times [0, \infty)$, the first coordinate determining the spatial location and the second the time component. This model first arose from modelling the growth of crystal aggregates (Avrami, 1939, 1940, 1941; Johnson & Mehl, 1939; Kolmogorov, 1937) and surface film on metal (Evans, 1945), and has been studied in detail by Meijering (1953), Gilbert (1962) and Möller (1992); see Chiu (1995, 1997) and Okabe et al. (2000, Ch. 5) for subsequent developments and applications.

The Johnson–Mehl model is in fact a random birth–growth coverage process. Germs are born and positioned in the space according to a spatial-temporal Poisson point process II. Once born, each germ $(x_i, t_i)$ generates a cell which grows radially. Locations occupied by growing cells are covered, and germs born in covered areas are not visible. In practice, it is often only...
possible to observe visible points of \( \Pi \). As will be explained later, the point process of such visible points is exactly the exposed tangent points process for a Boolean model in \( \mathbb{R}^d \times [0, \infty) \), where the typical grain represents the growth pattern of a typical cell and the vector \( u \) is oriented along the time axis. Hence, our proposed methods are appropriate in this situation.

For \( d = 1 \) and the space-homogeneous case, estimation of the constant growth rate \( v \) and the density of the time component of the underlying Poisson point process of a Johnson–Mehl model has been considered by Quine & Robinson (1992) and Chiu, Quine and Stewart (2000). Quine and Robinson’s method, which is confined to the time-homogeneous case in which the density of the time component is a positive finite constant, uses only the birth-times of the first visible germ and of the second visible germ, if there is one. Chiu et al. (2000) considered a time-inhomogeneous process and proposed a maximum likelihood estimator for \( v \). The estimated \( v \) enables them to reconstruct the unobservable Boolean model in order to estimate the intensity measure of the time component. Their estimator of the intensity measure makes use of both birth-times and birth-positions of all visible germs. The integral equation method proposed in this paper uses the birth-times only.

The paper is organised as follows. Section 2 summarises necessary probabilistic results concerning Boolean models that serve as a base for subsequent development of statistical techniques. Section 3 discusses intensity estimation for fully observable nonstationary Boolean models, using a kernel method to estimate the intensity of exposed tangent points. Section 4 discusses various cases in which estimation of the intensity, and sometimes the typical grain, is possible from observations of exposed tangent points alone. Section 5 adapts the general results to the special case of coverage processes. A number of simulation studies in § 5.2 confirm the applicability of the suggested estimation techniques. Finally, in § 6 we apply the integral equation method to neurobiological data. Further results of simulations related to the current
2 Tangent points in a nonstationary Boolean model

Let $\Xi$ be a Boolean model as defined in (1) with the germ process $\Pi$ of intensity $\Lambda$ and the typical grain $\Xi_0$ being a closed convex set. It is assumed throughout that $\Lambda$ is a locally finite measure and (2) holds. Fix a direction $u$ which will be used throughout to determine tangent points.

Since realisations of a Boolean model contain no information on how germs are positioned within grains, we avoid ambiguous translations of germs by assuming that $\Xi_0$ has its tangent point at the origin. The tangent points are then indistinguishable from the germs; some of them are exposed and others are covered by other grains as shown in Fig. 1. The process of exposed tangent points is denoted by $\Psi$, which as in the stationary case is obtained by means of a dependent thinning of $\Pi$. The capacity functional (Matheron, 1975, § 2.1, Molchanov, 1997, p. 11, Stoyan et al., 1995, § 6.1) of $\Xi$ is a functional defined on the family $\mathcal{K}$ of compact subsets of $\mathbb{R}^d$ by

$$T_{\Xi}(K) = \text{pr}(\Xi \cap K \neq \emptyset) = 1 - \exp \left[ -\int \text{pr}\{(K - z) \cap \Xi_0 \neq \emptyset\} \Lambda(dz) \right], \quad K \in \mathcal{K},$$

where $\tilde{\Xi}_0 = \{-x : x \in \Xi_0\}$ is the set centrally symmetric to $\Xi_0$. By Fubini’s theorem,

$$T_{\Xi}(K) = 1 - \exp \left[ -\int \text{pr}\{(K - z) \cap \Xi_0 \neq \emptyset\} \Lambda(dz) \right]. \quad (4)$$

In particular

$$p_{\Xi}(x) = \text{pr}(x \in \Xi) = 1 - \exp[-E\{\Lambda(x + \tilde{\Xi}_0)\}], \quad x \in \mathbb{R}^d, \quad (5)$$

is called the coverage function. Furthermore, define

$$q_{\Xi}(x, y) := \frac{1 - T_{\Xi}(\{x, y\})}{\{1 - p_{\Xi}(x)\}{1 - p_{\Xi}(y)}},$$

where
It follows from (4) that

\[
q_z(x, y) = \exp\{E[\Lambda\{(x + \Xi_0) \cap (y + \Xi_0)\}]\}
\]

\[
= \exp \left\{ \int \text{pr}(\{x - z, y - z \subseteq \Xi_0\}) \Lambda(dz) \right\}.
\]

The following results can be obtained by a simple modification of the arguments used by Molchanov & Stoyan (1994) for the stationary case. It refers to the concepts of intensity measure and the factorial second-order moment measure from the theory of point processes, see Daley & Vere-Jones (1988, § 5.4) and Stoyan et al. (1995, § 4.3).

**Theorem 1.** The intensity measure \(M\) and the factorial second-order moment measure \(M_2\) of \(\Psi\) are given by

\[
M(dx) = \Lambda(dx)\{1 - p_z(x)\},
\]

\[
M_2(dx \times dy) = \Lambda(dx)\Lambda(dy)q_z(x, y)\{1 - p_z(x)\}\{1 - p_z(y)\}
\]

\[
\times \text{pr}(x - y \notin \Xi_0)\text{pr}(y - x \notin \Xi_0).
\]

It might be useful to rewrite (6) using Fubini’s theorem as

\[
M(dx) = \Lambda(dx) \exp \left\{ - \int \text{pr}(x - z \in \Xi_0) \Lambda(dz) \right\}.
\]

Combining (6) and (7) yields

\[
M_2(dx \times dy) = M(dx)M(dy)q_z(x, y)\text{pr}(x - y \notin \Xi_0)\text{pr}(y - x \notin \Xi_0).
\]

Since the tangent point of \(\Xi_0\) is at the origin, convexity of \(\Xi_0\) implies that either \(x - y\) or \(y - x\) does not belong to \(\Xi_0\) with probability 1, provided that \(x \neq y\). Note that the tangent point of \(\Xi_0\) was defined as the lexicographical minimum of all points \(x \in \Xi_0\) that minimise the scalar product \(\langle x, u \rangle\). Define an order on \(\mathbb{R}^d\) by letting \(x \preceq y\) if \(\langle y - x, u \rangle \geq 0\). As \(\Xi_0\) has its tangent point at the origin, \(y - x\) does not belong to \(\Xi_0\) if \(x \neq y\), and so if \(\Lambda\) is absolutely continuous (9) can be rewritten as

\[
\mu_2(x, y) = \mu(x)\mu(y)q_z(x, y)\{1 - \text{pr}(y - x \in \Xi_0)\}, \quad x \preceq y,
\]
where $\mu$ and $\mu_2$ are the densities of $M$ and $M_2$, respectively.

3 Intensity estimation for fully observable models

Let $\Xi$ be a nonstationary Boolean model and $W \subset \mathbb{R}^d$ an observation window. A full observation of $\Xi$ refers to the set $\Xi \cap W$. Assume that the intensity measure $\Lambda$ is absolutely continuous and has a density $\lambda$ with respect to the $d$-dimensional Lebesgue measure $\text{mes}_d$.

If the coverage function $p_\Xi(x)$ is known or can be estimated, then (6) yields an estimator of $\lambda$, since

$$\lambda(x) = \mu(x)/\{1 - p_\Xi(x)\}. \quad (11)$$

Note that the ratio of two zeros in the right-hand side of (11) must be interpreted as zero. As $\Psi$ is nonstationary, the density of its intensity measure can be estimated by kernel methods (Bowman & Azzalini, 1997; Simonoff, 1996). If $k$ is a kernel and $h$ is a bandwidth, then $\mu$ is estimated by

$$\hat{\mu}(x) = \sum_{x_i \in \Psi} k\{(x - x_i)/h\}. \quad (12)$$

In simulation studies we usually used the Gaussian kernel for $k$. The coverage function can be estimated by the Nadaraya-Watson nonparametric regression estimator as

$$\hat{p}_\Xi(x) = \frac{\int_{\Xi \cap W} k_1\{(y - x)/h_1\} dy}{\int_W k_1\{(y - x)/h_1\} dy}, \quad (13)$$

where $k_1$ is another kernel and $h_1$ is another bandwidth, which may be the same as the kernel $k$ and the bandwidth $h$ used in (12). When replacing (11) by its empirical counterpart, we have to avoid dividing by zeros, so that some oversmoothing of $p_\Xi$, by using a larger bandwidth, is preferable to undersmoothing. Alternatively, local linear regression or nonparametric logistic regression (Bowman & Azzalini, 1997, Ch. 3) may be used to smooth the indicator function of $\Xi$. 8
For these methods to provide reliable results, both \( \mu \) and \( p_\Xi \) must be sufficiently smooth. As may be seen from (5), this is guaranteed by smoothness assumptions on \( \lambda \), since

\[
E \{ \Lambda(x + \tilde{\Xi}_0) \} = \int \Pr(x - z \in \Xi_0) \lambda(z) dz
\]

shows that \( p_\Xi \) and \( \mu \) have the same degree of smoothness as \( \lambda \).

[Figure 2 about here]

Extensive simulation studies have been done concerning the estimation of the density of the germ points in the Boolean model sampled on the unit square. We report here the results of one study where the density was chosen as \( \lambda(x) = 4cx_1x_2 \), where \( x = (x_1, x_2) \), so that the two coordinates of germ points are independent, and \( c = 500 \), so that the distribution of the total number of germs is Poisson with mean 500. The typical grain is a fixed disk of radius \( r = 0.0234 \), i.e. equal to 6/256, as simulations have been done on a discrete 256 \( \times \) 256 grid with the typical grain of radius 6. Fig. 2(a) shows a typical realisation of this Boolean model. For smoothing purposes, we have used the Gaussian kernel for both the density of the exposed tangent points process and the coverage function, with the bandwidths \( h = 30/256 \approx 0.117 \) and \( h_1 = 15/256 \approx 0.058 \). For the tangent points, this choice corresponds to the mean of normal optimal bandwidths for 50 realisations. A rule of thumb for the choice of \( h_1 \) is to ensure that the value of the kernel estimate in the centre of the largest clump should be close to 1. For the Gaussian kernel, this amounts to choosing \( h_1 \) to be about one quarter of the size of the largest clump. In real applications, an approximately optimal bandwidth for the kernel density estimate (12) can be chosen by using the normal approximation as described in Bowman & Azzalini (1997, § 2.4). The choice of the bandwidth is considerably more complicated for the bivariate nonparametric regression estimator (13) of \( p_\Xi \); see Yang & Tschernig (1999). One can use
numerically intensive approaches like cross-validation (Bowman & Azzalini, 1997, § 4.5). It may be computationally preferable to use a triangular kernel $k_1(t) = (1 - \|t\|)$, for $\|t\| \leq 1$, for estimation of the coverage function. In the above set-up, acceptable results have been obtained for the bandwidth $h_1 = 30/256 \approx 0.117$, as the triangular kernel decreases faster than the Gaussian.

Fig. 2(b) provides an estimate for the density function $\lambda$ obtained from the realisation shown in Fig. 2(a). Figures 2(c),(d) are the corresponding contour plots of the estimated and the true density. The estimate from a single realisation is not very satisfactory, as there are few tangent points in the upper right-hand corner. However, the estimate catches the overall behaviour of the density function. The estimate improves considerably if the window of observation increases or if more independent realisations of $\Xi$ are used. The average of these estimates from fifty independent simulations is shown in Fig. 2(e) as a contour plot. Fig. 2(f) presents the difference between true density function $\lambda$ and the average value of the density estimates normalised by the observed standard deviation at every point. It has been observed that the average of estimates typically underestimates the true intensity. This can be reduced by decreasing the bandwidths at the cost of obtaining rougher individual estimates.

The smoothing technique may also be used to estimate the function $\mu_2$ by considering all pairs of points $\{(x, y) : x, y \in \Psi\}$ as a point process in the space $\mathbb{R}^d \times \mathbb{R}^d = \mathbb{R}^{2d}$. Together with equation (10) this leads to an estimator of the covering probabilities $\text{pr}(z \in \Xi_0)$ of the typical grain $\Xi_0$ provided that $q_{\Xi}(x, y)$ can be estimated, for example by using another kernel density estimator.

For applications and theoretical studies it is often possible to assume that $\Pi$ is partially stationary, i.e. stationary with respect to at least one coordinate. This leads to a family of identically distributed observations
of a nonstationary Boolean model, with respect to other coordinates, and significantly improves the quality of estimation. For example, if the Boolean model in the plane is stationary with respect to translations along the $x_1$-axis, then the coverage function depends on the second coordinate $x_2$ only, and as such may be estimated as the coverage fraction of the horizontal line drawn at a given value of $x_2$. In an example similar to one shown in Fig. 2 with $\lambda(x_1, x_2)$ depending on $x_2$ only, the projections of the exposed tangent points on the $x_2$-axis already suffice to estimate the bivariate density using a univariate kernel density estimate, and this results in an acceptable estimate on the basis of a single realisation. In the partially stationary case, estimators of the coverage function become statistically more reliable, and it is often possible to produce the corresponding limit theorems; see Mase (1982). One particularly important partially stationary model corresponding to coverage processes will be discussed in § 5.

If edge effects near the boundary of $W$ are taken into account, then we may simply reduce the window in order to subsample in a smaller window free from edge effects (Ripley, 1988, Ch. 3). Other sophisticated edge-corrected estimators such as the Kaplan–Meier estimator (Baddeley & Gill, 1997) and the Hanisch estimator (Chiu & Stoyan, 1998) may also be employed.

4 Estimation methods based on observations of exposed tangent points

4.1 A typical grain with known covering probabilities

In many cases the whole realisation of $\Xi \cap W$ is not observable, but instead we can observe only exposed tangent points of $\Xi \cap W$ with respect to a fixed direction $u$. Although this observation scheme drastically reduces the amount of information available, in some cases it is possible to estimate the intensity and sometimes even make inference about the typical grain. In this
situation we assume that the covering probabilities \( f(y) = \Pr(y \in \Xi_0) \) of the typical grain \( \Xi_0 \) are known, in which case integral equation (8) turns into

\[
\mu(x) = \lambda(x) \exp \left\{ - \int f(x - z) \lambda(z) dz \right\},
\]  
(14)

where \( \lambda \) and \( \mu \) are the densities of \( \Lambda \) and \( M \) respectively. It follows from (5) and (6) that equation (14) can be written as

\[
\mu(x) = \lambda(x) \exp \left\{ -E\{\Lambda(x + \hat{\Xi}_0)\} \right\}.
\]  
(15)

These integral equations can be solved by discretisation on a regular grid, which reduces (14) or (15) to a system of nonlinear equations with respect to the values of \( \lambda \) on the grid. Then a smoothed estimator of \( \mu \) can be plugged in to produce an estimator for \( \lambda \). However, for any fine grid this leads to considerable computational work.

As a simple illustration, suppose the grid is a regular lattice \( L \) with mesh size \( \delta \) and assume that \( \lambda(x) \) is a constant \( \lambda_i \) for \( x \in i + [0, \delta]^d \) for each lattice point \( i \). The value of \( \lambda_i \) can be approximated by the limit of the nondecreasing sequence \( \{\lambda_i^{(k)} : k \geq 0\} \) given by

\[
\log \lambda_i^{(k+1)} = \log \mu(i) + \sum_{j \in L} a_{ij} \lambda_j^{(k)},
\]

where \( a_{ij} = \int_{j+[0,\delta]^d} f(i - z) dz \) and \( \lambda_i^{(0)} = \mu(i) \).

Since the window \( W \) does not appear in the integral equation, this method is free from edge effects. In a sense, edge effects are being subsumed in the allowed nonstationarity for \( \lambda \).

As can be seen in § 5.2, the complexity of solving integral equation (14) will be reduced in partially stationary cases. For the stationary case, in which \( \lambda \) is a constant, equation (14) is simply a nonlinear equation of the form

\[
\mu = \lambda \exp(-c),
\]

where \( c = \int f(y) dy = E\{\text{mes}_d(\Xi_0)\} \) is the expected volume of the typical grain.
Suppose that the typical grain $\Xi_0$ is an unknown deterministic set. Then
\[ \Pr(x - y \notin \Xi_0)\Pr(y - x \notin \Xi_0) = \text{either zero or one}. \]
If we consider only the case $x \preceq y$, equation (9) turns into
\[
M_2(dx \times dy) = \begin{cases} 
M(dx)M(dy)q_\Xi(x, y), & y - x \notin \Xi_0, \\
0, & \text{otherwise},
\end{cases}
\]
for $x \preceq y$.

The function $q_\Xi$ cannot be estimated directly by observing only the exposed tangent points. However, $q_\Xi$ is always strictly positive, and so the pairs $(x, y)$ with $x \preceq y$ that provide zero values for $M_2(dx \times dy)$ can be used to estimate $\Xi_0$. In practice, this is reduced to calculating pairwise differences of exposed tangent points, smoothing their density by using a kernel estimator and then identifying a convex region where the kernel estimator vanishes. The estimated covering probabilities $f(y)$ can be used to solve integral equation (14).

### 4.3 A deterministic typical grain which is closed under the Minkowski addition

Suppose that the typical grain $\Xi_0$ is deterministic and closed under the Minkowski addition, i.e.
\[ \Xi_0 \oplus \Xi_0 \subset \Xi_0, \tag{16} \]
which implies that for every germ $y \in (x + \Xi_0)$ we have $(y + \Xi_0) \subset (x + \Xi_0)$.

Cones form the simplest class of sets which fulfil this condition. The convexity assumption on $\Xi_0$ can be waived if condition (16) is fulfilled. Let $\mathcal{M}$ denote a family of sets satisfying (16) and such that $\mathcal{M}$ is closed under taking finite intersections and closures of countable unions.

**Maximum likelihood estimation.** If $\Xi_0 = A$ is an unknown deterministic set from $\mathcal{M}$, we can use a maximum likelihood argument to estimate $\Xi_0$. The
likelihood \( L(\lambda, A) \) of observing a realisation \( \Psi \) of exposed tangent points inside a window \( W \) is proportional to the product of the likelihood that germ points appear at the observed points of \( \Psi \) and the probability that no other germ from \( \Pi \) apart from those appearing in \( \Psi \) is an exposed tangent point, i.e.

\[
L(\lambda, A) = \left\{ \prod_{x_i \in \Psi} \lambda(x_i) \right\} \exp\{-\Lambda\{W \setminus (\Psi \oplus A)\}\}.
\]  

(17)

Therefore, \( L(\lambda, A) \) is nondecreasing with respect to \( A \). However, \( A \) cannot be made arbitrarily large because all points of \( \Psi \) should remain exposed. Thus, in order to maximise the likelihood, the set \( A \) should be the largest, with respect to inclusion, set from \( \mathcal{M} \) that does not contain in its interior any of the pairwise differences \( x_i - x_j \), where \( x_i, x_j \in \Psi \). The conditions on \( \mathcal{M} \) ensure that such a set exists. Let \( \hat{A} \) denote this estimator of \( A \).

By maximising (17) with respect to \( A \) we obtain the profile likelihood for \( \lambda \)

\[
L_{\mathcal{M}}(\lambda) = \sup_{A \in \mathcal{M}} \left\{ \prod_{x_i \in \Psi} \lambda(x_i) \right\} \exp\{-\Lambda\{W \setminus (\Psi \oplus A)\}\}
\]

\[
= \left\{ \prod_{x_i \in \Psi} \lambda(x_i) \right\} \exp\{-\Lambda\{W \setminus (\Psi \oplus \hat{A})\}\}.
\]

The profile loglikelihood is given by

\[
\ell_{\mathcal{M}}(\lambda) = \log L_{\mathcal{M}}(\lambda) = \sum_{x_i \in \Psi} \log \lambda(x_i) - \Lambda\{W \setminus (\Psi \oplus \hat{A})\}.
\]  

(18)

If we have a parametric family of densities \( \lambda \), then (18) can be used to obtain parameter estimates by maximising profile likelihood and eventually the corresponding confidence intervals based on likelihood.

The following example shows how the maximum likelihood approach yields well-known results elegantly.
Example 1. Let \( \lambda(x) = \lambda \) be a constant density function on \( W \). The profile loglikelihood

\[
l_M(\lambda) = \#(\Psi \cap W) \log \lambda - \lambda \text{mes}_d\{W \setminus (\Psi \oplus \hat{A})\},
\]

where \( \#(\cdot) \) denotes the number of points in the corresponding set, leads to the maximum likelihood estimator for \( \lambda \) as

\[
\hat{\lambda} = \frac{\#(\Psi \cap W)}{\text{mes}_d\{W \setminus (\Psi \oplus \hat{A})\}}.
\] (19)

Since \( \hat{A} \) satisfies (16), the denominator of the estimator (19) equals \( \text{mes}_d(W)(1 - \hat{p}) \), where \( \hat{p} \) is the estimated volume fraction of the Boolean model with typical grain \( \hat{A} \). Thus, the estimator given in (19) is identical to the well-known exposed tangent points estimator for the intensity of a stationary Boolean model with convex grains based on (3); see Molchanov (1997, § 5.3), Molchanov & Stoyan (1994) and Stoyan et al. (1995, p. 89). The asymptotic properties of the likelihood yield formulae for confidence intervals for \( \lambda \) of the Boolean model with deterministic typical grain from the family \( \mathcal{M} \). Consider

\[
\frac{\partial l_M(\lambda)}{\partial \lambda} = \frac{\#(\Psi \cap W)}{\lambda} - \text{mes}_d(W)(1 - \hat{p}),
\]

\[
\frac{\partial^2 l_M(\lambda)}{\partial \lambda^2} = -\frac{\#(\Psi \cap W)}{\lambda^2}.
\]

The quadratic approximation to the loglikelihood yields the following \( 100(1 - \alpha) \) % confidence interval for \( \lambda \), which coincides with a result of Molchanov & Stoyan (1994):

\[
\hat{\lambda} \pm z_{\alpha/2} \left\{ \frac{\hat{\lambda}}{\text{mes}_d(W)(1 - \hat{p})} \right\}^{1/2},
\]

where \( z_{\alpha/2} \) is the \( 100(1 - \alpha/2) \) th-percentile of the standard normal distribution.

In contrast to the general derivation of the exposed tangent points estimator and its asymptotic properties (see Molchanov & Stoyan, 1994), the
above argument presumes that the grain is deterministic and belongs to the predetermined family $\mathcal{M}$. On the other hand, the maximum likelihood estimator is derived from observing only the exposed tangent points instead of a full observation of the Boolean model.

*Integral equation method.* The method described in § 4.1 is applicable here, because the deterministic typical grain $\Xi_0$ is either known or can be estimated by the maximum likelihood argument and so the covering probabilities $f(y)$ are simply equal to 1 if $y \in \Xi_0$ and to 0 otherwise.

*Reconstruction method.* Since $\Xi_0$ satisfies (16), the whole realisation of the Boolean model can be reconstructed using the exposed tangent points $\Psi$ and the known or estimated $\Xi_0$ as

$$\Xi = \bigcup_{x_i \in \Psi} (x_i + \Xi_0) = \Psi \oplus \Xi_0. \tag{20}$$

This provides a full realisation of the Boolean model, so that the methods described in § 3 are applicable and can be used to estimate the density $\lambda$.

Suppose the exposed tangent points are observed in a window $W$. If the support of $\lambda$ is a subset of $W$, then the reconstruction does not create edge effects and what we reconstruct is the same as what we should have observed. However, if the support of $\lambda$ is a superset of $W$, the reconstruction creates unwanted edge effects and the reconstructed model is not the corresponding full realisation of the Boolean model in $W$. Nevertheless, we may consider a smaller window which is free from edge effects.

Note that the reconstruction procedure based on (20) cannot be extended to the case when the typical grain $\Xi_0$ is a random set from $\mathcal{M}$. The reason for this is that the grains $\Xi_i$ at the exposed tangent points $\Psi$ are not independent. Thus, some exposed tangent points may not be observed in the reconstructed Boolean model.
A random typical grain with unknown distribution

Assume that $\Xi_0 = VA$, where $A$ is a known deterministic convex set having its tangent point at the origin and $V$ is a random variable that determines the grain's size and has an unknown distribution with density $\theta(v)$. For every point $x_i \in \Psi$ we may find the largest possible $v$ such that $x_i + vA$ does not have any other point of $\Psi$ in its interior. This produces a sample of values $\tilde{v}_i$ that correspond to the points of $\Psi$. However, for every $x_i \in \Psi$, the value $\tilde{v}_i$ is not the true value $v_i$ that corresponds to the grain implanted at $x_i$. The maximum likelihood argument above cannot be applied here because (16) does not hold and so the likelihood is not the one given by (17). We only can say that $v_i \in [0, \tilde{v}_i]$, so that the sample $\tilde{v}_i$ provides a right censored variant of the sample of the true values of $v_i$. Nevertheless, integral equations (6) and (9) yield

$$
\mu(x) = \lambda(x) \exp \left\{ - \int \lambda(z) dz \int_{0}^{\infty} 1_{[x-z] \cap \nu A \neq \emptyset} \theta(v) dv \right\},
$$

$$
\frac{\mu_2(x, y)}{\mu(x)\mu(y)} = \left\{ \int_{0}^{\infty} 1_{[x-y, y-x] \cap \nu A \neq \emptyset} \theta(v) dv \right\} \times \exp \left\{ \int \lambda(z) dz \int_{0}^{\infty} 1_{[x-z, y-z] \cap \nu A \neq \emptyset} \theta(v) dv \right\},
$$

and, in principle, provide a way to estimate both $\lambda$ and $\theta$ by solving a system of two integral equations. Here $1_{E}$ equals 1 if $E$ holds and 0 otherwise.

5 Application to coverage processes

5.1 Introduction

Consider a Poisson point process $\Pi = \{(x_i, t_i)\}$ in $D \times [0, \infty)$, where $D \subseteq \mathbb{R}^d$, with the intensity measure $\Lambda$; $D$ may or may not be the observation window $W$. Consider a random closed set

$$
\Xi_0 = \{(x, t) : x \in Z(t), \ t \geq 0\},
$$
where the growth function \( Z(t), t \geq 0 \), is a possibly random increasing set-valued function defined on \([0, \infty)\) with values being compact convex subsets of \( \mathbb{R}^d \) so that \( Z(0) = \{o\} \) is the origin. If \( \Xi \) is the Boolean model with the typical grain \( \Xi_0 \) and germs \( \Pi \), then its lower border

\[
\zeta(x) = \inf\{t \geq 0 : (x, t) \in \Xi\}
\]

is called a coverage process; see Fig. 3. One particularly important example, called steady circular growth, arises if \( Z(t) = B_{tV}(o) \) is the ball of radius \( tV \) centred at the origin, where a positive random variable \( V \) determines the growth speed of the typical grain.

[Figure 3 about here]

In studies of coverage processes the germs from \( \Pi \) are usually called seeds, and \( \Lambda \) is often the product of two measures \( N \times H \), so that coordinates of seeds are independent. Here \( N \) is a measure on \( W \) that governs positions of the seeds and \( H \) is a measure on \([0, \infty)\) that controls their temporal components. A coverage process is said to be space-homogeneous if \( N \) is the Lebesgue measure \( \text{mes}_d \) in \( \mathbb{R}^d \).

Many germs, i.e. seeds, coming from \( \Pi \) will not be visible as they appear inside covered areas. The visible seeds \( \Psi \) are obtained by a thinning of \( \Pi \), so that a point \((x_i, t_i)\) is retained if \((x_i, t_i)\) is not covered by any \((x_j, t_j) + \Xi_j, j \neq i\). If we fix the vector \( u \) pointing upwards along the time axis, then \( \Xi_0 \) has its tangent point at the origin, and the visible seeds become the exposed tangent points of \( \Xi \).

The basic aim in the statistics of coverage processes is to estimate the intensity measure \( \Lambda \) and the growth distribution. We always assume that \( \Lambda = N \times H \), where \( \nu \) and \( \eta \) are densities of \( N \) and \( H \) respectively. Clearly, the density of \( \Lambda \) is \( \lambda(x, t) = \nu(x)\eta(t) \).
5.2 Estimation of intensity

**Full observation.** Suppose that the coverage process is fully observable. The kernel smoothing technique for the estimation of \( \lambda \) and the coverage function described in § 3 may be employed. A typical realisation of the coverage process is shown in Fig. 4(a) as the lines that connect neighbouring visible seeds.

**Exposed tangent points.** In applications, however, full observations are usually not available and instead of the full growth pattern only visible seeds are observable. As all seeds have two coordinates we will speak about their birth-positions and birth-times. As before, \( M \) denotes the intensity measure of the point process of the visible seeds and \( \mu(x, t) \) is the density of \( M \).

Assume that the growth function \( Z(t) \) has a known distribution, which is equivalent to saying that the distribution of \( \Xi_0 \) is known. Then the covering probabilities of \( \Xi_0 \) can be found as \( f(x, t) = \text{pr}\{x \in Z(t)\} \). The density of \( \Lambda \) satisfies integral equation (15), which in the current framework simplifies to

\[
\mu(x, t) = \nu(x) \eta(t) \exp \left\{ - \int_0^t E[N \{ D \cap (x + \tilde{Z}(t - s)) \}] \eta(s) ds \right\}, \quad x \in D.
\]

(23)

For simplicity, assume that the coverage process is space-homogeneous, i.e. \( N = \text{mes}_d \) and that the growth is steady circular with the speed determined by a random variable \( V \) with a known distribution. Then \( Z(t - s) = B_{(t-s)V(o)} \) and

\[
\mu(x, t) = \eta(t) \exp \left\{ - \int_0^t E[\text{mes}_d \{ D \cap B_{(t-s)V} (x) \}] \eta(s) ds \right\}, \quad x \in D.
\]

(24)

Finally, if \( D = \mathbb{R}^d \), \( \mu(x, t) \) does not depend on \( x \) and is given by

\[
\mu(x, t) = \eta(t) \exp \left\{ - \int_0^t \omega_d E(V^d)(t - s)^d \eta(s) ds \right\},
\]

19
where \( \omega_d = \pi^{d/2}/\Gamma(1 + d/2) \) is the volume of a unit ball in \( \mathbb{R}^d \). The integral of \( \mu \) over the observation window \( W \) is then given by

\[
\tilde{\mu}(t) = \int_W \mu(x, t)dx = \text{mes}_d(W)\eta(t) \exp \left\{ -\int_0^t \omega_d E(V^d)(t - s)^d \eta(s)ds \right\}.
\]

(25)

Note that the function \( \tilde{\mu}(t) \) on the left-hand side represents the density of birth-times of visible seeds if their positions within \( W \) are ignored. That is, the birth-times of visible seeds contain sufficient information to estimate \( \eta \), provided that the distribution of the growth speed is known.

For an application of this method, \( \tilde{\mu} \) can be estimated using a kernel estimator. It should be noted that birth-times of visible seeds usually have a skewed distribution, since seeds that are born earlier have considerably more chances to be visible. This calls for accurate tuning of the smoothing procedures. We have found that the logarithmic transformation, typically used to handle positive data, does not work well for the type of data coming from the coverage processes. However, the square-root transformation works well and is used throughout, prior to smoothing with a Gaussian kernel and the bandwidth determined by using normal optimal smoothing. After smoothing, the inverse transformation is applied, the estimate of \( \tilde{\mu}(t) \) for negative \( t \) is discarded and its positive part is rescaled to keep the total integral unchanged. Note that (25) can be usefully rewritten as

\[
\eta(t) = \frac{\tilde{\mu}(t)}{\text{mes}_d(W)} \exp \left\{ \int_0^t \omega_d E(V^d)(t - s)^d \eta(s)ds \right\}, \quad t > 0.
\]

(26)

Therefore, \( \eta(t) \) can be expressed by using an integral of \( \eta(s) \) for \( s \leq t \), which makes it possible to solve the integral equation in a stepwise manner, by extending a solution from small to large \( t \)’s and using the fact that \( \eta(t) \approx \tilde{\mu}(t)/\text{mes}_d(W) \) for small \( t \).

Consider a constant \( \delta > 0 \) and let \( t_i = i\delta \) for \( i = 0, 1, 2, \ldots \). If \( \delta \) is small, we can approximate the right-hand side of (26) by considering \( \eta(t) \) to be a
constant $\eta_i$ in each $[t_i, t_{i+1})$. Then, $\eta(t)$ can be approximated by the solutions of
\[
\eta_i = \frac{\tilde{\mu}(t_i)}{\text{mes}_d(W)} \exp \left[ \sum_{k=0}^{i-1} \eta_k \omega_d E(V^d) \left\{ \frac{(t_i - t_k)^{d+1}}{d + 1} - \frac{(t_i - t_{k+1})^{d+1}}{d + 1} \right\} \right], \quad \text{for } i \geq 1.
\]
The initial value $\eta_0$ can be chosen to be $\tilde{\mu}(\delta)/\text{mes}_d(W)$. Note that the integral equation method can also handle an inhomogeneous density $\nu$ of birth-positions. This requires a discretisation of the integral equation (23). Similarly, if the positions are homogeneous and $D$ is a bounded subset of $\mathbb{R}^d$, this calls for a solution of (24).

Note that the typical grain $\Xi_0$ is closed under the Minkowski addition as required by (16). This makes it tempting to have recourse to the reconstruction method, which does not depend on the space-homogeneity assumption. However, if the growth is random, the reconstruction method does not work, because the growth speeds at different visible seeds are not independent. In other words, it is impossible to ensure that all visible seeds remain visible if we assign independent realisations of $\Xi_0$ to visible seeds.

However, if the growth function is deterministic, for example assuming deterministic steady growth, the reconstruction method is perfectly applicable. The advantages of the integral equation method over the reconstruction method are as follows: the integral equation method does not require birth-positions if the underlying Poisson process is space-homogeneous; it is free from edge effects when the coverage process is on $\mathbb{R}^d$; and it works for random speeds with known distribution. On the other hand, the reconstruction method works without any changes for an arbitrary distribution of seed positions, and so naturally handles edge effects if they are present originally. In the following we mostly report estimation results for densities based on application of the integral equation method. The results for cumulative functions are generally far better and more stable.
Simulation studies. We report simulation results from the case in which

\[ \eta(t) = \frac{a \gamma^k}{\Gamma(k)} t^{k-1} e^{-\gamma t}. \]

Fig. 4(a) shows a realisation of the coverage process on the line \([0, L]\) with \(a = 6, \gamma = 1, k = 2, L = 50\) and \(v = 0.2\). Fig. 4(b) provides the estimates of the densities \(\tilde{\mu}/L\) and \(\eta\) on the time-axis of the exposed tangent points and the underlying Poisson process, respectively, based on this realisation, in which 137 out of 250 tangent points are exposed.

[Figure 4 about here]

Fig. 5(a) contains the estimates of \(\eta\), based on the integral equation method, from 10 independent realisations of such a coverage process. The mean, a pointwise 95% confidence bound and the true \(\eta\) are given in Fig. 5(b). In agreement with the findings of Chiu et al. (2000), the maximum likelihood estimates of the constant speed \(v\), discussed in \$5.3\), are sufficiently reliable. For the current simulations, the estimates of \(v\) range from 0.2003 to 0.2123 with the median 0.2006.

[Figure 5 about here]

As Fig. 4(b) shows, the integral equation method and the reconstruction method produce very similar results. This fact has been confirmed by numerous simulation studies for various density functions. In a sense, the reconstruction method is a Monte-Carlo solution to the integral equation. Suppose that equation (25) is to be solved. The exponential part can be estimated by the volume fraction of a simulation of the corresponding Boolean model and then a Monte-Carlo solution of \(\eta\) can be obtained. Similarly, in the reconstruction method, the reconstructed full realisation is regarded as the result of a simulation.
About half the tangent points are exposed in the above simulation study. If the speed $v$ increases from 0.2 to 1.0, say, while the other parameters remain unchanged, then fewer, about one quarter, of the tangent points are exposed. Fig. 6(a) presents individual estimates and the mean with a pointwise 95% confidence bound of the estimated $\eta$ from 10 realisations in comparison with the true density. Clearly, increase in speed results in a much smaller number of visible seeds and poorer quality of tail estimation for the density function. Moving from one dimension, $[0, L]$, to two dimensions, $[0, L]^2$, does not create any difficulty, and the results with the same parameter values apart from $L = 10$ and $v = 1$ are given in Fig. 6(b). In this case fewer than one quarter out of 600 tangent points are exposed.

[Figure 6 about here]

If $k = 1$, then $\eta$ has a pole at zero, since it is actually a constant times the density of an exponential distribution. The results are given in Fig. 7(a) with all other values kept the same as in Fig. 5. Finally, consider that $\eta$ is simply a constant times the density of a uniform distribution on $(0, 1)$. Fig. 7(b) shows the results. As usual, the tail of the estimated $\eta$ is not so satisfactory because not many exposed tangent points with large birth-times are available.

[Figure 7 about here]

5.3 Estimation of growth pattern

Assume first that the growth speed $v$ is a deterministic constant. This case was considered in Chiu et al. (2000), where it was shown that the maximum likelihood estimator of $v$ is in fact the maximum possible speed. Since $v$ is a deterministic constant, we need not consider all pairwise differences to calculate the maximum possible speed but only consider nearest visible neighbours of each visible seed.
Suppose that the growth rate is deterministic but not constant. The maximum likelihood argument in § 4 is applicable only when (16) is satisfied, i.e. only when the growth speed \( v \) itself is also a nondecreasing function of \( t \). The maximum possible speed can be obtained by considering the difference between each seed and its nearest neighbours and then determining the minimum of all possible functions with non-increasing slopes as shown in Fig. 8. The resulting maximum likelihood estimate of the speed is obtained as the pointwise maximum of the functions found as described for every visible seed.

[Figure 8 about here]

If the growth is steady and circular such that the speed follows an unknown distribution with density \( \theta(v) \), then the maximum likelihood argument is not applicable because (16) is not satisfied. We can estimate \( \lambda \) and \( \theta \) using integral equations (21) and (22), which lead to

\[
\mu(x, t) = \lambda(x, t) \exp \left\{ -\int_{|z-s| \leq |x-t|} \lambda(z, s) dz ds \int_0^{|x-z|/|t-s|} \theta(v) dv \right\},
\]

\[
\frac{\mu_2((x, t), (y, u))}{\mu(x, t) \mu(y, u)} = \int_0^{\|x-y\|/\|t-u\|} \theta(v) dv \times \exp \left\{ \int_{|z-s| \leq |x-t|} \lambda(z, s) dz ds \int_0^{\max\{|x-z|/|t-s|, |y-z|/|u-s|\}} \theta(v) dv \right\},
\]

where \( \| \cdot \| \) denotes Euclidean distance in \( \mathbb{R}^d \).

6 A case study

Bennett & Robinson (1990) proposed the following model of autoinhibited release of neurotransmitters at a synapse. The terminal of a neuronal axon at the neuromuscular junction has branches consisting of strands containing many randomly scattered sites. At a synapse an action potential triggers
the release of neurotransmitter at these sites. Each quantum released causes release of an inhibitory substance which diffuses along the terminal at a constant rate preventing further releases in the inhibited region. Measurements of time and a surrogate for location of release of each transmitter are possible.

This is in fact a coverage process in \([0, L] \times [0, \infty)\) with intensity measure \(\text{mes}_1 \times \eta\) and constant growth speed. Chiu et al. (2000) used the maximum likelihood method to estimate the speed \(v\), and then reconstructed the process. They estimated the cumulative function \(H([0, t])\) of \(\eta\) by using observation of the reconstructed coverage process inside \([0, L] \times [0, t]\). Clearly, kernel smoothing can be avoided when estimating the cumulative density by using the reconstruction method. Here we use the smoothing technique and the integral equation method to estimate the density \(\eta\) of \(H\).

The data, provided by Professor M.R. Bennett of the Neurobiology Research Centre at the University of Sydney, consist of the results of 800 experiments. The number of exposed tangent points in each experiment ranges from 0 to 4. If we use the same scale as in Chiu et al. (2000), the maximum likelihood estimate of the speed \(v\) was 0.018 and \(L\) was set to be 1. The estimated density \(\eta(t)\) and a pointwise 95\% confidence bound from these 800 experiments, divided into 10 groups of 80, are given in Fig. 9.

[Figure 9 about here]

ACKNOWLEDGEMENT

We thank Adrian Bowman for helpful discussions concerning smoothing methods. The smoothing algorithms for coverage processes have been implemented using \texttt{sm} library of S-plus functions described in Bowman & Azzalini (1997). This research is supported by the UK/Hong Kong Joint Research Scheme. We are grateful to the editor and the referees for stimulating remarks.
References


**Molchanov, I. S.** (1997). *Statistics of the Boolean Model for Practitioners*
and Mathematicians. Chichester: Wiley.


Figure captions

Figure 1.
Overlapping grains, tangent lines and tangent points.

Figure 2
Simulation results for estimates of the density function of a fully observable nonstationary Boolean model. All density estimates are shown on a reduced window $[0.2, 0.8]$ to avoid edge effects and scaled by $500^{-1}$. (a) A realisation of the Boolean model with $\lambda(x_1, x_2) = 2000x_1x_2$. (b) An estimate of the density function by one realisation. (c) Contour plot of the estimate of the density function shown in (b). (d) Contour plot of the true density function. (e) Contour plot of the average estimate of the density function. (f) Normalised deviance plot. Maximum absolute deviance is 2.17.

Figure 3
Visible and invisible seeds in a coverage process.

Figure 4
(a) A realisation of the coverage process. The germs or seeds are represented by solid dots. Dots joined by lines are visible seeds and the connected lines provide a realisation of the coverage process. Dots above the lines and the lines themselves are unobservable. (b) The corresponding density estimates of $\eta$. Solid line — true density; dotted line — integral equation estimate; short dashed line — reconstruction estimate, dashed line — density estimate for birth times of seeds.
Figure 5
Simulation results for the integral equation method. (a) Ten individual estimates for the density. (b) The mean of the ten individual estimates and pointwise 95% confidence bounds; the true density is shown as solid line.

Figure 6
Effects of changing speed (a) and dimension (b). Solid lines — true density; dotted lines — averages of ten individual estimates; dashed lines — pointwise 95% confidence bounds.

Figure 7
Estimates for the exponential (a) and uniform (b) densities. Solid lines — true densities; dotted lines — averages of ten individual estimates; dashed lines — pointwise 95% confidence bounds.

Figure 8
The maximum likelihood estimation of a nondecreasing growth speed at a visible seed. The solid dots (●) represent pairwise differences, the solid lines (—) the maximum possible speed and the dash-dotted lines (−·) the maximum likelihood estimate of the speed at the visible seed.

Figure 9
Density estimate for neurotransmitter data. The mean is shown as a solid line. Dotted lines provide pointwise 95% confidence bounds.
Figure 1:
Figure 2:
Figure 3:
Figure 4:
Figure 5:
Figure 6:
Figure 7:
Figure 8: