Aalborg Universitet

AALBORG UNIVERSITY

Likelihood inference for unions of interacting discs

Møller, Jesper; Helisová, Katarina

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

AALBORG UNIVERSITY

Likelihood inference for
unions of interacting discs

by

Jesper Møller and Kateřina Helisová

Likelihood inference for unions of interacting discs

Jesper Møller[1] and Kateřina Helisová[2,3]

[1] Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg, Denmark. Email: jm@math.aau.dk.

[2] Department of Probability and Mathematical Statistics, Charles University in Prague, Sokolovská 83, 18675 Praha 8, Czech Republic. Email: helisova@karlin.mff.cuni.cz.

[3] Department of Mathematics, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 16627 Praha 6, Czech Republic. Email: helisova@math.feld.cvut.cz.

Abstract

To the best of our knowledge, this is the first paper which discusses likelihood inference for a random set using a germ-grain model, where the individual grains are unobservable, edge effects occur, and other complications appear. We consider the case where the grains form a disc process modelled by a marked point process, where the germs are the centres and the marks are the associated radii of the discs. We propose to use a recent parametric class of interacting disc process models, where the minimal sufficient statistic depends on various geometric properties of the random set, and the density is specified with respect to a given marked Poisson model (i.e. a Boolean model). We show how edge effects and other complications can be handled by considering a certain conditional likelihood. Our methodology is illustrated by analyzing Peter Diggle's heather dataset, where we discuss the results of simulation-based maximum likelihood inference and the effect of specifying different reference Poisson models.

Keywords: Boolean model; connected component Markov process; disc process; edge effects; germ-grain model; interaction; quermass-interaction process; random closed set; simulation-based maximum likelihood; spatial Markov property; summary statistics.

1 Introduction

In spatial statistics and stochastic geometry, one often models a realization of a planar random set $\mathbf{Y}$ by a germ-grain model $\mathcal{U}_{\mathbf{X}}$, where

- $\mathbf{X}$ is a marked point process of points u_i called germs and associated marks K_i called primary grains,

- the germs u_i constitute a locally finite point process in $\mathbb{R}^2$,

- the primary grains K_i are random compact subsets of $\mathbb{R}^2$,

- $u_i + K_i = \{u_i + x : x \in K_i\}$, the translation of the set K_i by the vector u_i, is called a grain,

- $\mathcal{U}_{\mathbf{X}}$ is the union of all grains $u_i + K_i$,

see e.g. Hanisch (1981), Cressie (1993), and Stoyan, Kendall and Mecke (1995). It is well-known that any random closed set whose realizations are locally finite unions of compact convex sets is a germ-grain model with convex and compact grains (Weil and Wieacker, 1984, 1988). In practice, in order to make statistical inference, one uses a much smaller class of models, where random-disc Boolean models play the main role, see Stoyan *et al.* (1995) and the references therein, and so $\mathbf{Y} \approx \mathcal{U}_{\mathbf{X}}$ is an approximation. There is a lack of interaction

under the Boolean model, since $\mathbf{X}$ is an independently marked Poisson process. Many authors have mentioned the need of developing flexible germ-grain models with interaction between the grains (Diggle, 1981; Hall, 1988; Baddeley and Møller, 1989; Cressie, 1993; Stoyan *et al.*, 1995; Kendall, Van Lieshout and Baddeley, 1999; Møller and Helisová, 2008).

The present paper considers a new class of interacting disc processes from Møller and Helisová (2008), illustrated by analyzing Peter Diggle's heather dataset (Diggle, 1981), which previously have been modelled by a random-disc (or more general) Boolean model, cf. the references in Section 2. In particular, to the best of our knowledge, this is the first time that a paper discusses simulation-based likelihood inference for a germ-grain model. If we observe only $\mathbf{Y} \cap W$, where $W \subset \mathbb{R}^2$ denotes a bounded window, statistical inference is complicated by that

(i) as we observe only the union of the grains within W, the individual grains may be unobservable and edge effects may occur,

(ii) the model should account for the fact that the marked points (u_i, K_i) may interact,

(iii) the grains may only approximately be discs,

(iv) usually only a digital image is observed and the resolution makes it difficult to identify circular structures.

We discuss and to some extent try to solve these problems, in particular (i)-(ii). Briefly, we model the geometric properties of the connected components of $\mathbf{Y}$, where results from Møller and Helisová (2008) become useful, and our model extends the quermass-interaction disc model in Kendall *et al.* (1999) and forms a particular class of connected component

3

Markov processes (Baddeley and Møller, 1989; Baddeley, Van Lieshout and Møller, 1996; Møller, 1999; Chin and Baddeley, 2000).

Section 2 reviews previous statistical analyses of Diggle's heather dataset modelled by a random-disc Boolean model. Section 3 introduces our connected component Markov process model, where we specify the density with respect to an 'appropriate' reference Poisson disc process, and we have a first discussion on likelihood inference. The specification of the reference process is discussed further in Section 4 and will be based on the previous analyses reported in Section 2. Since there may be several reasonable reference processes, Section 5 compares the results of the likelihood analysis based on connected component Markov processes defined with respect to different reference processes. Section 6 discusses how well the estimated models fit the data.

2 Data example and previous analyses

As an illustrative application example, we consider the well known heather dataset first presented in Diggle (1981). The review in this section on previous statistical analysis of this dataset is relevant for the subsequent sections.

Figure 1 shows a binary image of the presence of heather (*Calluna vulgaris*, indicated by black) in a 10×20 m rectangular region W at Jädraås, Sweeden (henceforth units are meters). Assuming the heather plants grow from seedlings into roughly hemispherical bushes, a disc process model has a straightforward biological interpretation by identifying heather bushes and discs. The dataset in Diggle (1981) was coded as a 100×200 binary matrix, while Figure 1 consists of 250×508 pixels. Figure 1 was kindly provided by Adrian Baddeley
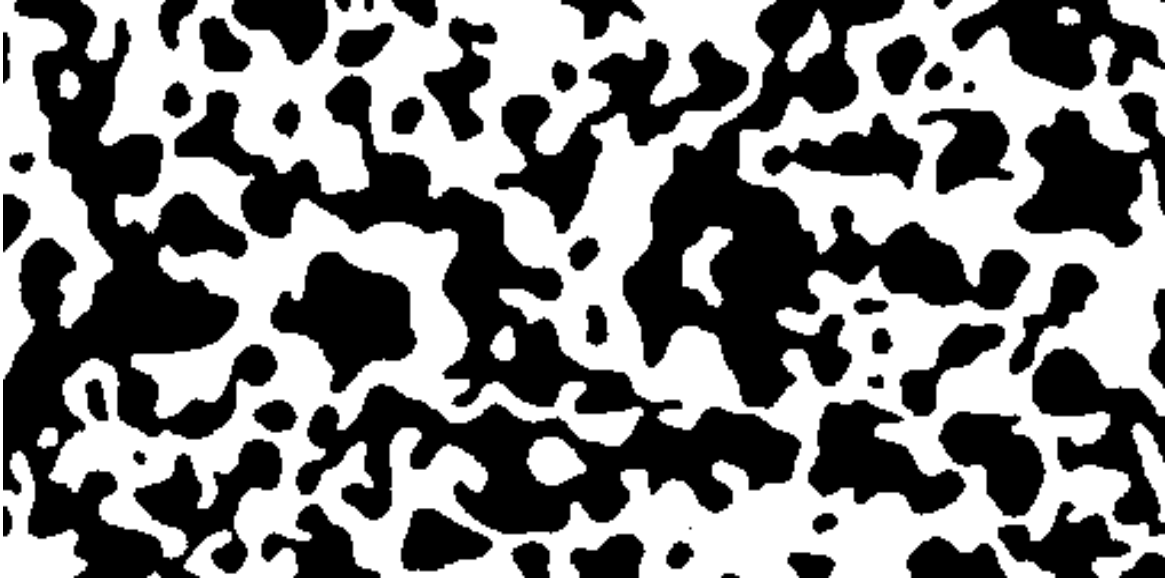
Figure 1: A redigitisation at higher resolution of the heather dataset.

after the original piece of paper, on which the map was hand-drawn by Peter Diggle, was
processed by Chris Jonker who scanned the paper into a grey-scale image, which was cleaned
up and converted into a binary image, which was smoothed using mathematical morphology
tools.

Diggle (1981) modelled the presence of heather by a stationary random-disc Boolean
model, i.e. the centres u_i are given by a stationary Poisson process with intensity $\beta > 0$,
and the radii r_i are independent and identically distributed and independent of the centres.
Let R denote a generic radius, where units are meters. Diggle modelled its distribution (the
mark distribution) by a left-truncated Weibull-density

$$\kappa\rho(r-\delta)^{\kappa-1}\exp\left(-\rho(r-\delta)^{\kappa}\right), \quad r \geq \delta \tag{1}$$

where δ, κ, ρ are positive parameters. He estimated the parameters by a minimum con-
trast method based on the covariance function for the indicator function of presence of

5

heather, splitting the 10×20 m window W into two 10×10 m non-overlapping squares in order to provide some cross validation of the results. His estimates of $(\beta, \delta, \kappa, \rho)$ were $(2.21, 0.281, 0.281, 20.6)$ for the left-hand square, and $(2.11, 0.226, 1.011, 12.5)$ for the right-hand square, corresponding to a mean and standard deviation of R of approximately 0.31 and 0.04 (left-hand square) and 0.31 and 0.08 (right-hand square), and an intensity of around two bushes per square meter (Diggle, 1981; Hall, 1988). This procedure passed a model control in Diggle (1981) based on the spherical contact distribution function (as defined in Section 6), however, as Diggle pointed out after comparing simulation of the fitted models with the data, the visual impression is not good.

Diggle (1981) observed that his fitted models generate more separate patches than in the data, and Hall (1985) noticed that Diggle's estimates of radius standard deviation may be too small. Ripley (1988) constructed diagnostic tools for the shape of the random set based on operations from mathematical morphology (as defined in Section 6), which also indicated that Diggle's random-disc Boolean model is not fitting well. Hall (1988) noticed that some of the differences between data and fitted models would diminish if the discs were given ragged edges.

Hall (1988) considered a stationary Boolean model with germ intensity β and isotropic convex grains, and assumed that the moments $\mu_1 = \mathrm{E}\left[L(K)\right]/(2\pi)$ and $\mu_2 = \mathrm{E}\left[A(K)\right]/\pi$ are finite, where A and L mean area and perimeter, respectively. Isotropy means that the distribution of a generic primary grain K is invariant under rotations about the origin in $\mathbb{R}^2$. Clearly, a stationary Boolean model with circular grains is isotropic with $\mu_i = \mathrm{E}(R^i)$, $i = 1, 2$, and we let $\mu = \mu_1$ denote the mean and $\sigma = \sqrt{\mu_2 - \mu^2}$ the standard deviation of R. Using that the number of grains intersecting a fixed convex set C is Poisson distributed

with mean $\beta(A(C) + L(C)\mu_1 + \pi\mu_2)$, and using a quadratic lattice with C corresponding to an arbitrary vertex, edge, or square, Hall (1988) derived moment equations from which estimates of β, μ_1, μ_2 were derived (this procedure is used in Section 4). In comparison with Diggle's minimum contrast estimation procedure, Hall's procedure is computationally much easier. Hall obtained the estimate 2.36 of β, and taking bushes to be discs, estimates of about 0.27 m and 0.12 m for μ and σ, respectively. Hall's estimates of β and μ are roughly similar to those of Diggle (1981), while Hall's estimate of σ is about twice as large.

Hall (1988) investigated the correlation structure and the hypothesis of isotropic convex grains. He provided some evidence of anisotropy (see also Renshaw and Ford, 1983), but concluded that anisotropy has no important effect on his techniques that assume isotropy. Comparing Figure 1 with Diggle's coarser digitization, Adrian Baddeley (personal communication) noticed a couple of discrepancies: the coarse data has some implausible horizontal 'spikes', presumably arising from errors in recording some of the run lengths. Baddeley also ran a quick spectral analysis using FFT on the new high-resolution data and did not find any evidence of anisotropy.

Further approaches of fitting Boolean models, in connection to random-discs as well as other grain structures, are discussed in Dupač (1980), Serra (1980), Cressie (1993), Stoyan *et al.* (1995), and Molchanov (1997). Cressie (1993, Section 9.5.3; based on Laslett, Cressie and Liow, 1985) considered a stationary Boolean model with intensity β and not necessarily convex grains. He used a method based on marker points (if grains are convex, the marker point of a grain is its most southwesterly point), or more precisely exposed marker points, i.e. the marker points on the boundary of $\mathbf{Y} \cap W$ (in the case of convex grains, exposed marker points are called lower convex tangent points in Stoyan *et al.*, 1995). Cressie's primary

7

aim was to estimate β using Laslett's transformation, which transforms the exposed marker points into a stationary Poisson process of intensity β. He obtained the estimate 1.16 of β, i.e. a much lower value than obtained by Diggle (1981) and Hall (1985, 1988). Cressie noticed that this discrepancy may be partly caused by missing exposed marker points due to the digitization of the image of heather data. Stoyan *et al.* (1995, page 95) referred to several simulation studies, which indicate that a certain method of intensities is the most precise method. Since this method (Stoyan *et al.*, 1995, page 89) requires to count the lower convex tangent points of $\mathbf{Y} \cap W$, for the same reason as for the method based on exposed marker points, it may not work well for the heather dataset. However, as Cressie (1993) remarked, exposed marker points could be sampled directly as part of the data collection process.

3 A parametric class of interacting disc process models

Section 3.1 specifies the parametric class of connected component Markov models considered in this paper, and Section 3.2 discusses various likelihood function used for the statistical analysis of the heather dataset in Section 5.

3.1 Density

The model is defined as follows, considering circular grains $u_i + K_i = b(u_i, r_i)$, where $b(u, r)$ denotes a closed disc with centre u and radius r. Since the heather plants also grow outside W, we let $S \supseteq W$ denote a bounded region and identify $\mathbf{X}$ by a finite marked point process with points $u_i \in S$ and marks $r_i \geq 0$. Note that the region S where the points live is unknown

8

to us. For any finite marked point configuration $\mathbf{x} = \{(u_1, r_1), \ldots, (u_n, r_n)\} \subset S \times [0, \infty)$, where n can be any non-negative integer (if $n = 0$ then $\mathbf{x}$ is empty), assume that $\mathbf{X}$ has density

$$f_\theta(\mathbf{x}) = \frac{1}{c_\theta} \exp \left(\theta_1 A(\mathcal{U}_\mathbf{x}) + \theta_2 L(\mathcal{U}_\mathbf{x}) + \theta_3 N_{\mathrm{cc}}(\mathcal{U}_\mathbf{x}) + \theta_4 N_{\mathrm{h}}(\mathcal{U}_\mathbf{x}) \right) \tag{2}$$

with respect to a reference Poisson process of discs, where A, L, N_{cc}, and N_{h} denote area, perimeter, number of connected components, and number of holes, respectively, $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ is an unknown four-dimensional real parameter, and c_θ is a normalizing constant. Note that c_θ is not expressible on closed form if $\theta \neq 0$.

Since the pattern of heather in Figure 1 may be assumed to be homogeneous, cf. Diggle (1981), we let the Poisson reference process be stationary with intensity $\beta > 0$. For technical reasons, we assume that the reference Poisson process of discs has a mark cumulative distribution function Q on $[0, \infty)$ such that

$$S \supseteq b(u, s) \setminus b(u, t) \text{ and } Q(t) > 0 \text{ for some } s, t, u \text{ with } \infty > s > t > 0 \text{ and } u \in \mathbb{R}^2. \tag{3}$$

For instance, (3) is satisfied for any bounded $S \supseteq W$ when using Diggle's mark distribution (1) and his parameter estimates given in Section 2. The specification of β and Q is important, since they will control the number of discs and their radii in the reference process, and hence effect the meaning of the parameter θ in (2). Since the discs are incompletely observed, we propose a pragmatic approach, where we first fit a Boolean disc model and second consider (2) as an extension of this Boolean model to an interacting disc process model. In Section 4 different reference Poisson disc process are suggested by using the results discussed in Section 2.

Combining (3) with Møller and Helisová (2008, Proposition 1), we obtain that (2) is a

regular exponential family with parameter space

$$\Theta = \{(\theta_1, \ldots, \theta_4) \in \mathbb{R}^4 : \int \exp\left(\pi\theta_1 r^2 + 2\pi\theta_2 r\right) Q(\mathrm{d}r) < \infty\} \tag{4}$$

and canonical minimal sufficient statistic

$$T(\mathbf{x}) = (A(\mathcal{U}_\mathbf{x}), L(\mathcal{U}_\mathbf{x}), N_{\mathrm{cc}}(\mathcal{U}_\mathbf{x}), N_{\mathrm{h}}(\mathcal{U}_\mathbf{x})). \tag{5}$$

Note that $\chi(\mathcal{U}_\mathbf{x}) = N_{\mathrm{cc}}(\mathcal{U}_\mathbf{x}) - N_{\mathrm{h}}(\mathcal{U}_\mathbf{x})$ is the Euler-Poincare characteristic, and if $\theta_3 + \theta_4 = 0$ then (2) specifies a quermass-interaction process (Kendall *et al.*, 1999) with parameter $(\theta_1, \theta_2, \theta_3)$ and canonical minimal sufficient statistic $(A(\mathcal{U}_\mathbf{x}), L(\mathcal{U}_\mathbf{x}), \chi(\mathcal{U}_\mathbf{x}))$. Further geometric characteristics than the four statistics in (5) are incorporated in the connected component Markov models studied in Møller and Helisová (2008), but they will be hard to determine from Figure 1 because of the problems (iii)-(iv) (Section 1).

3.2 Likelihood

Using Markov chain Monte Carlo (MCMC) methods, approximate maximum likelihood estimates of the parameter θ in (2) can be obtained and test statistics for reduced models (e.g. a quermass-interaction process, $\theta_3 + \theta_4 = 0$) can be evaluated as explained in Section 5. This section specifies various likelihood function used in Section 5. For later use, notice that under the Poisson reference disc process ($\theta = 0$), using a notation as in Section 2,

$$p = 1 - \exp\left[-\beta\pi(\sigma^2 + \mu^2)\right] \tag{6}$$

is the area fraction covered by $\mathbf{Y}$, and the empirical area fraction is $\hat{p} = A(\mathbf{Y} \cap W)/A(W) = 0.5014$.

In order to deal with edge effects, consider for the moment the quermass-interaction process ($\theta_3 + \theta_4 = 0$) and assume that the reference Poisson process is such that the radii are (almost surely) bounded by $r/2$ where $r < \infty$ is a constant. The quermass-interaction process is Markov in the sense of Ripley and Kelly (1977) with respect to the neighbour relation given by overlapping discs, and it satisfies an appealing spatial Markov property as follows. Let $W_{\ominus r} = \{u \in W : b(u,r) \subseteq W\}$ be the r-clipped window, and split $\mathbf{X}$ into $\mathbf{X}^{(1)}$, $\mathbf{X}^{(2)}$, $\mathbf{X}^{(3)}$ corresponding to discs with centres in $W_{\ominus r}$, $W \setminus W_{\ominus r}$, W^c, respectively. The spatial Markov property states that $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(3)}$ are conditionally independent given $\mathbf{X}^{(2)}$, and the conditional distribution $\mathbf{X}^{(1)}|\mathbf{X}^{(2)} = \mathbf{x}^{(2)}$ has density

$$f_{\theta_1,\theta_2,\theta_3}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)}) = \frac{1}{c_{\theta_1,\theta_2,\theta_3}(\mathbf{x}^{(2)})} \exp\left(\theta_1 A(\mathcal{U}_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}}) + \theta_2 L(\mathcal{U}_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}}) + \theta_3 \chi(\mathcal{U}_{\mathbf{x}^{(1)} \cup \mathbf{x}^{(2)}})\right)$$

with respect to the reference Poisson process restricted to discs with centres in the r-clipped window (Kendall *et al.*, 1999; Møller and Helisová, 2008). However, in practice it seems problematic to use this conditional density for likelihood inference, since $f_{\theta_1,\theta_2,\theta_3}(\mathbf{x}^{(1)}|\mathbf{x}^{(2)})$ depends on $\mathcal{U}_{\mathbf{x}^{(2)}} \setminus W$, which in general is not observable.

Instead the following spatial Markov property of a general connected component Markov processes becomes useful. Split $\mathbf{X}$ into $\mathbf{X}^{(a)}$, $\mathbf{X}^{(b)}$, $\mathbf{X}^{(c)}$ corresponding to discs belonging to connected components of $\mathcal{U}_{\mathbf{X}}$ which are respectively

(a) contained in the window W,

(b) intersecting both W and its complement W^c,

(c) contained in W^c,

see Figure 2. Let $\mathbf{x}^{(b)}$ denote any feasible realization of $\mathbf{X}^{(b)}$, i.e. $\mathbf{x}^{(b)}$ corresponds to a finite

11

configuration of discs such that for all $K \in \mathcal{K}(\mathcal{U}_{\mathbf{x}^{(b)}})$, K intersects both W and W^c. By Møller and Helisová (2008, Proposition 5), conditional on $\mathbf{X}^{(b)} = \mathbf{x}^{(b)}$, we have that $\mathbf{X}^{(a)}$ and $\mathbf{X}^{(c)}$ are independent, and the conditional distribution of $\mathbf{X}^{(a)}$ depends only on $\mathbf{x}^{(b)}$ through $V = W \cap \mathcal{U}_{\mathbf{x}^{(b)}}$ and has density

$$f_\theta(\mathbf{x}^{(a)}|V) = \frac{1}{c_\theta(V)} \mathbf{1}[\mathcal{U}_{\mathbf{x}^{(a)}} \subseteq W \setminus V] \exp\left(\theta \cdot T(\mathbf{x}^{(a)})\right) \tag{7}$$

with respect to the reference Poisson process of discs, where $\cdot$ is the usual inner product. Note that (7) depends only on what we can observe, and it does not depend on how we chose S. Therefore, we consider the conditional log likelihood

$$L_c(\theta) = \theta \cdot T(\mathbf{x}^{(a)}) - \log c_\theta(V) \tag{8}$$

corresponding to the conditional density (7), where $T(\mathbf{x}^{(a)})$ and V are determined as discussed below. Since $c_\theta(V)$ is not expressible on closed form, we approximate it by MCMC methods and find an approximate maximum likelihood estimate as discussed in Section 5.

For the heather data, Figure 3 shows the connected components intersecting the boundary of W (in grey), and we denote the union of these components by $\mathbf{Y}^{(b)}$, and set $\mathbf{Y}^{(a)} = \mathbf{Y} \cap W \setminus \mathbf{Y}^{(b)}$. In (8) we use the approximations $\mathbf{Y}^{(a)} \approx \mathcal{U}_{\mathbf{x}^{(a)}}$ and $Y^{(b)} \approx V$, cf. Sections 1-2, where

$$A(\mathbf{Y}^{(a)}) = 45.6, \quad L(\mathbf{Y}^{(a)}) = 190, \quad N_{\mathrm{cc}}(\mathbf{Y}^{(a)}) = 32, \quad N_{\mathrm{h}}(\mathbf{Y}^{(a)}) = 2. \tag{9}$$

Here the perimeter was calculated by a method based on intrinsic volume densities (Mrkvička and Rataj, 2008; for software, see `http://home.pf.jcu.cz/~mrkvicka/math/`), while the three other quantities were easily determined. Note that $A(V)/A(W) = 0.2734$ as compared to $\hat{p} = 0.5014$ from above, and there is of course a lack of information when inference is based on the conditional log likelihood (8).
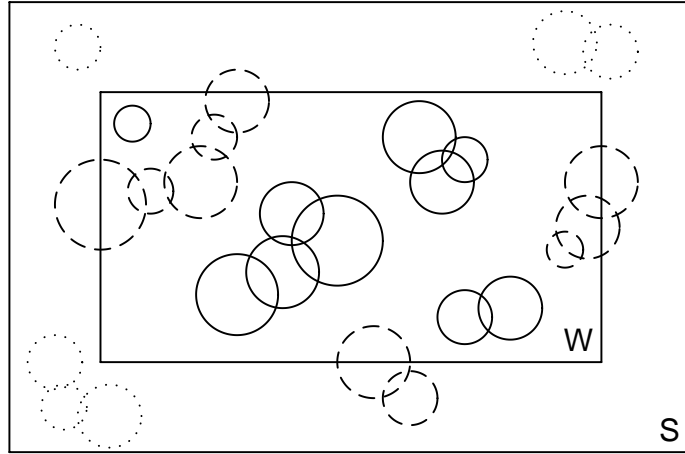
Figure 2: An example of possible realizations of $\mathbf{X}^{(a)}$ (the full circles), $\mathbf{X}^{(b)}$ (the dashed circles), and $\mathbf{X}^{(c)}$ (the dotted circles).
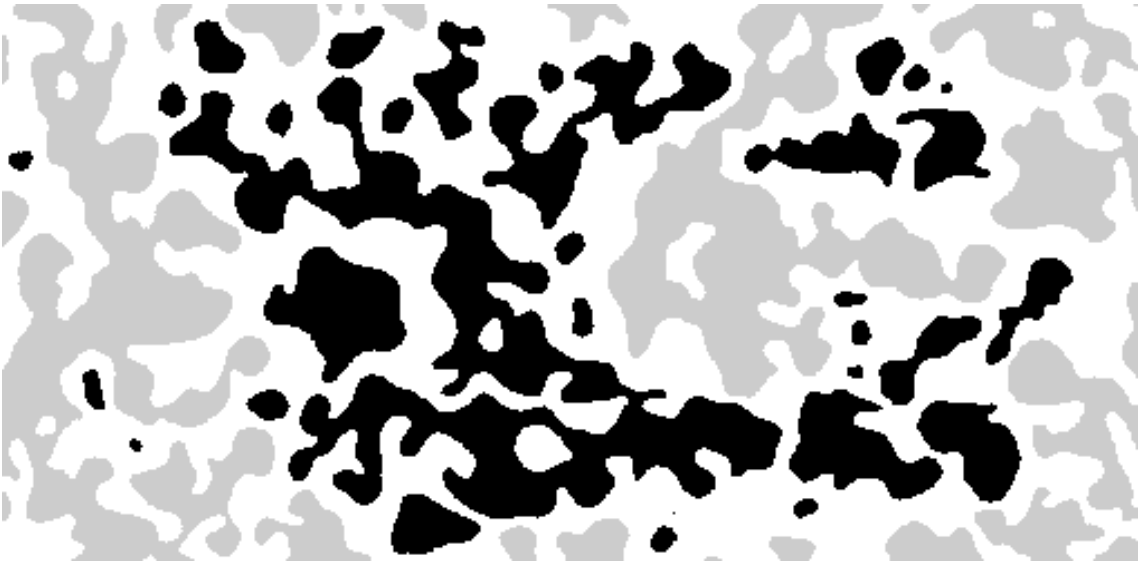

Figure 3: Heather dataset with the components intersecting the boundary of the observation window coloured grey.

In some applications it may even happen that $\mathbf{Y}^{(a)} = \emptyset$, in which case the maximum likelihood estimate based on (8) does not exist. Partly for this reason and partly for comparison with the conditional maximum likelihood approach based on (8), we also consider the following unconditional approach. Let $S = W$ and $\tilde{\mathbf{Y}} = \mathbf{Y} \cap W$, and approximate the log likelihood by

$$L_u(\theta) = \theta_1 A(\tilde{\mathbf{Y}}) + \theta_2 L(\tilde{\mathbf{Y}}) + \theta_3 N_{\mathrm{cc}}(\tilde{\mathbf{Y}}) + \theta_4 N_{\mathrm{h}}(\tilde{\mathbf{Y}}) - \log c_\theta \tag{10}$$

where

$$A(\tilde{\mathbf{Y}}) = 100.28, \quad L(\tilde{\mathbf{Y}}) = 382.82, \quad N_{\mathrm{cc}}(\tilde{\mathbf{Y}}) = 56, \quad N_{\mathrm{h}}(\tilde{\mathbf{Y}}) = 6. \tag{11}$$

Here edge effects occur, since compared to (2) we have replaced $\mathcal{U}_{\mathbf{x}}$ by $\tilde{\mathbf{Y}}$, while $\mathcal{U}_{\mathbf{X}}$ may expand outside W. However, in fact S should be larger than W, in which case discs $b(u, r)$ with $(u, r) \in \mathbf{X}$, $u \in S \setminus W$, and $b(u, r) \cap W \neq \emptyset$ have been ignored. This may possibly to some extent reduce the problem of edge effects and hence possibly to some extent justify that in (10) we use the normalizing constant c_θ from (2) with $S = W$. However, for the present data where the values in (9) are about twice as large as those in (11), results in Section 5.3 will confirm that edge effects do play a crucial role.

4   Specifying the reference processes

We shall later compare results obtained by three rather different reference Boolean models which are specified by (R1)-(R3) below. Note that we suppress in the notation that the normalizing constants in the density (2), the conditional likelihood (8), and the unconditional likelihood (10) depend on the choice of reference process.

Using an R with unbounded support may imply undesirable restrictions on the parameter space $\boldsymbol{\Theta}$ given by (4). For instance, if we use Diggle's left-truncated Weibull-density (1) and his parameter estimate obtained from the left-hand square (see Section 2), we need to assume that either $\theta_1 < 0$ or both $\theta_1 = 0$ and $\theta_2 \leq 0$, and if we use his parameter estimate obtained from the right-hand square, we need to assume that $\theta_1 \leq 0$.

In the sequel we assume that R has bounded support, whereby (3) implies that $\boldsymbol{\Theta} = \mathbb{R}^4$ is as large as possible. Independent biological evidence suggests that R should be less than 0.5 m (Diggle, 1981), and spatial correlation analysis indicates that R is less than 0.4 m (Hall, 1985, 1988).

Below we refer to parameter estimates obtained by a method from Hall (1985) for an isotropic Boolean model. For the method, we use a square lattice, and after some experimentation we decided to work with the side length $c_1 = 0.48$ m. (Then a square roughly corresponds to 12 pixels in Figure 1 where 1 pixel corresponds to 4×3.94 cm.) Within W, the lattice has $n_0 = 903$ vertices, $n_1 = 1742$ edges, and $n_2 = 840$ squares. Denote $(N_0, N_1, N_2) = (450, 474, 71)$ the numbers of these vertices, edges, and squares which are not intersected by $\mathbf{Y}$. Let $A_i = \log(n_i/N_i)$, $i = 0, 1, 2$, and let $c_2 = c_1^2 = 0.2304$ m^2 be the area of a square. By Hall's method, the estimates are given by

$$\beta = \frac{1}{c_2}\left(A_0 - 2A_1 + A_2\right), \quad \mu_1 = \frac{1}{2c_1\beta}\left(A_1 - A_0\right), \quad \mu_2 = \frac{A_0}{\pi\beta}.$$

We obtain $(2.45, 0.26, 0.09)$ as the estimate of (β, μ_1, μ_2), and hence $(0.26, 0.16)$ as the estimate of (μ, σ).

Taking $\mu = 0.26$ and $\sigma = 0.16$ as estimated above, the three reference Poisson disc processes are specified by

(R1) $\beta = 2.45$ and R follows the restriction of $N(\mu, \sigma^2)$ to the interval $[0, 0.50]$;

(R2) $\beta = 2.45$ and R is uniformly distributed on $[0, 0.53]$;

(R3) $\beta = 1.16$ and R is uniformly distributed on $[0, 0.53]$.

A normal mark distribution has also been considered in Dupač (1980) and Hall (1988), and the restriction of $N(\mu, \sigma^2)$ to the interval $[0, 0.50]$ corresponds to 88 % of the probability mass. Under (R1), the mean of R is close to μ, and its standard deviation is 0.12. These values are close to the estimates of the mean and standard deviation obtained in Hall (1985). Under both (R2) and (R3), R has mean μ and standard deviation σ. By (6) the area fraction is $p = 0.46$ under (R1), $p = 0.51$ under (R2), and $p = 0.29$ under (R3), in comparison with the empirical area fraction $\hat{p} = 0.50$ from (6). The apparently too small area fraction under (R3) is caused by the small value of $\beta = 1.16$ taken from Laslett *et al.* (1985), cf. Section 2.

5   Results obtained by maximum likelihood

In general for any parametric random-disc Boolean model, maximum likelihood inference is very complicated because of the problems (i) and (iii)-(iv) (Section 1). To handle (i) by the EM-algorithm (Dempster, Laird and Rubin, 1977) seems impracticable, since we cannot do the expectation step even if (ii)-(iv) were not an issue. Alternatively, a missing data MCMC approach may be suggested (Geyer, 1999; Møller and Waagepetersen, 2004), but this would require us to make simulations from the conditional distribution of the missing data (the unobserved discs) given the observed data $\mathbf{Y} \cap W$. Constructing a well mixing MCMC algorithm for such conditional simulations seems very hard. For such an algorithm, in order

to fit a parametric random-disc Boolean model, the observed boundary of the random set may cause that too small discs are too often generated by the algorithm because of (iii)-(iv).

Section 5.1 discusses instead results obtained by simulation-based inference based on the conditional likelihood $L_c(\theta)$ in (8) where we compare the results obtained by using the different reference Poisson disc processes (R1)-(R3) from Section 4. Section 5.3 compare these results with other results based on the unconditional likelihood $L_u(\theta)$ in (10). Although our connected component Markov model may be viewed as a modification of the Boolean model obtained by allowing interaction, maximum likelihood based on MCMC methods becomes feasible, since $L_c(\theta)$ and $L_u(\theta)$ only depend on the heather data (Figure 1), and general MCMC importance sampling methods for finding an approximation of the normalizing constant in (8) and in (10) apply. Thereby we can obtain approximate maximum likelihood estimates (MLE's), confidence regions, and Wald test statistics, cf. Geyer (1999) and Møller and Waagepetersen (2004). For the simulations we use the Metropolis-Hastings birth-death algorithm discussed in Møller and Helisová (2008) with a burn-in of 10,000 iterations and samples for Monte Carlo estimates based on 1,000,000 iterations.

5.1 Results based on the conditional likelihood

Approximate MLE's based on $L_c(\theta)$ under the full model $\theta \in \mathbb{R}^4$ and when using each of the reference processes (R1)-(R3) are shown in Table 1. To check our code, for each fitted value of θ, $\hat{\theta}$ say, we simulated a new Markov chain of length 1,000,000 (again after a burn-in of 10,000 iterations) with equilibrium density $f_{\hat{\theta}}$, and used the average of the 1,000,000 values of the canonical minimal sufficient statistic in (5) as new data, $\bar{T}$ say. By the law of large

17

	θ_1	θ_2	θ_3	θ_4
(R1)	-2.14	0.89	-1.78	-1.01
95%-CI	$[-3.68, -0.60]$	$[0.48, 1.31]$	$[-2.28, -1.28]$	$[-2.37, 0.34]$
Wald	7.45	17.89	48.96	2.14
(R2)	-4.81	1.17	-2.26	-0.69
95%-CI	$[-6.36, -3.26]$	$[0.75, 1.58]$	$[-2.74, -1.77]$	$[-2.04, 0.66]$
Wald	37.04	29.77	83.66	1.01
(R3)	-3.67	1.62	-2.25	-0.13
95%-CI	$[-5.42, -1.93]$	$[1.16, 2.09]$	$[-2.76, -1.73]$	$[-1.49, 1.23]$
Wald	17.01	46.67	73.01	0.04

Table 1: Under the full model (i.e. $(\theta_1, \theta_2, \theta_3, \theta_4) \in \mathbb{R}^4$), when using each of the reference processes (R1)-(R3), the table shows approximate MLE's and 95%-confidence intervals together with Wald statistics for testing each of the four hypotheses $\theta_i = 0$ with $i = 1, 2, 3, 4$.

numbers, $\bar{T}$ is expected to be close to the mean $\mathrm{E}_{\hat{\theta}}T$, and since the likelihood equation is given by $\mathrm{E}_\theta T = \bar{T}$, we expect the MLE based on $\bar{T}$ to be close to $\hat{\theta}$. In fact we did obtain approximate MLE's very close to those in Tables 1-2, e.g. the difference between the two estimates is $(0.00, 0.00, 0.00, 0.01)$ when (R1) is the reference process.

Assuming the MLE's are approximately normally distributed, Table 1 also shows approximate 95%-confidence intervals as well as the values of Wald statistics for testing each of the four hypotheses $\theta_i = 0$ with $i = 1, 2, 3, 4$. Compared to 3.842, the 95%-quantile in a χ^2-distribution with one degree of freedom, no matter which of the reference Poisson disc processes we use, the Wald statistic is highly significant except for the hypothesis $\theta_4 = 0$.

This indicates that $\theta_i \neq 0$, $i = 1, 2, 3$, and $\theta_4 = 0$, meaning that the characteristics of the connected components given by area, perimeter, and number of components seem all important, while the number of holes seems an irrelevant statistic. We have also considered the hypothesis of a quermass-interaction model (i.e. $\theta_3 + \theta_4 = 0$), where the values of the Wald statistic were given by 14.49, 16.22, 10.22 when using (R1), (R2), (R3), respectively, so these values seem highly significant too.

Columns 2-4 in Table 2 are similar to those in Table 1 but concern the reduced model with $(\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ and $\theta_4 = 0$. The results in the two tables are very similar. We refer to the reduced model as the (A, L, N_{cc})-interaction model, since (8) is the log likelihood for a regular exponential family model with canonical minimal sufficient statistic $T_r = (A(\mathcal{U}_{\mathbf{X}^{(a)}}), L(\mathcal{U}_{\mathbf{X}^{(a)}}), N_{cc}(\mathcal{U}_{\mathbf{X}^{(a)}}))$. Thus the likelihood equation is given by that the mean $\mathrm{E}_{(\theta_1, \theta_2, \theta_3)} T_r$ is equal to the observed value of T_r. For the three fitted models in Table 2, we obtained Monte Carlo estimates of $\mathrm{E}_{(\theta_1, \theta_2, \theta_3)} T_r$ given by $(46.45, 192.32, 31.34)$, $(45.76, 192.15, 32.24)$, $(45.85, 190.79, 31.79)$, respectively. These estimates are all rather close to the observation $T_r = (45.6, 190, 32)$, indicating that the approximate MLE's may be close to the exact MLE's. The values in Table 2 of the Wald statistic for testing $\theta_i = 0$ when $i = 1, 2$, or 3, seem again highly significant no matter which reference process is used. For $i = 1, 2, 3$, considering the confidence intervals in Table 2, the sign of θ_i seems effectively to be the same in all three fitted models. This may indicate that the characteristics of the connected components given by area, perimeter, and number of components have similar 'roles' in all three fitted models.

	L_c			L_u		
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
(R1)	-2.33	0.92	-1.77	-0.91	-0.02	-1.13
95%-CI	$[-3.80, -0.85]$	$[0.52, 1.31]$	$[-2.27, -1.26]$	$[-1.83, 0.00]$	$[-0.29, 0.25]$	$[-1.49, -0.78]$
Wald	9.54	21.01	46.89	3.80	0.02	39.25
(R2)	-4.91	1.18	-2.25	-1.75	1.02	-1.63
95%-CI	$[-6.48, -3.35]$	$[0.77, 1.59]$	$[-2.75, -1.75]$	$[-2.85, -0.65]$	$[0.70, 1.34]$	$[-2.01, -1.25]$
Wald	38.02	32.33	78.78	9.70	39.65	70.81
(R3)	-3.71	1.64	-2.25	-3.45	0.74	-1.63
95%-CI	$[-5.47, -1.95]$	$[1.17, 2.10]$	$[-2.77, -1.74]$	$[-4.41, -2.48]$	$[0.46, 1.01]$	$[-1.98, -1.28]$
Wald	17.04	47.11	73.89	49.35	26.82	82.86

Table 2: Under the reduced model (i.e. $(\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ and $\theta_4 = 0$), when using each of the reference processes (R1)-(R3), the table shows approximate MLE's and 95%-confidence intervals together with Wald statistics for testing each of the three hypotheses $\theta_i = 0$ with $i = 1, 2, 3$. Columns 2-4 are the results based on the conditional likelihood L_c. Columns 5-7 are the results based on the unconditional likelihood L_u.

	(R1)	(R2)	(R3)
mean of R	0.28	0.25	0.28
sd of R	0.12	0.13	0.13
intensity	2.57	2.36	1.76

Table 3: Under the reduced model (i.e. $(\theta_1, \theta_2, \theta_3) \in \mathbb{R}^3$ and $\theta_4 = 0$), when using each of the reference processes (R1)-(R3), the table shows MCMC estimates of the mean and standard deviation of the typical radius and intensity of bushes.

5.2 The typical radius and intensity of heather bushes

In Section 2, we reported on various results for the mean and variance of the typical radius R and the intensity of bushes as obtained by Diggle and Hall, and further such results were reported in Section 4. For comparison with these results, Table 3 summaries the results obtained by our estimated (A, L, N_{cc})-interaction models, and Figure 4 shows the estimated distribution of the typical radius R (as specified below) together with the densities of the typical radius under the corresponding reference processes. While our conclusions earlier have been rather insensitive to the choice of reference process, the distributions in Figure 4 are sensitive to this choice. In the case where (R1) is the reference process, there is a rather close agreement between the distributions of R under (R1) respective the fitted (A, L, N_{cc})-interaction model, while the disagreement is pronounced in the two other cases. The distributions of R under the fitted (A, L, N_{cc})-interaction models with reference processes (R1) respective (R3) look a bit similar and different from the distribution of R under the fitted (A, L, N_{cc})-interaction model with reference process (R2).
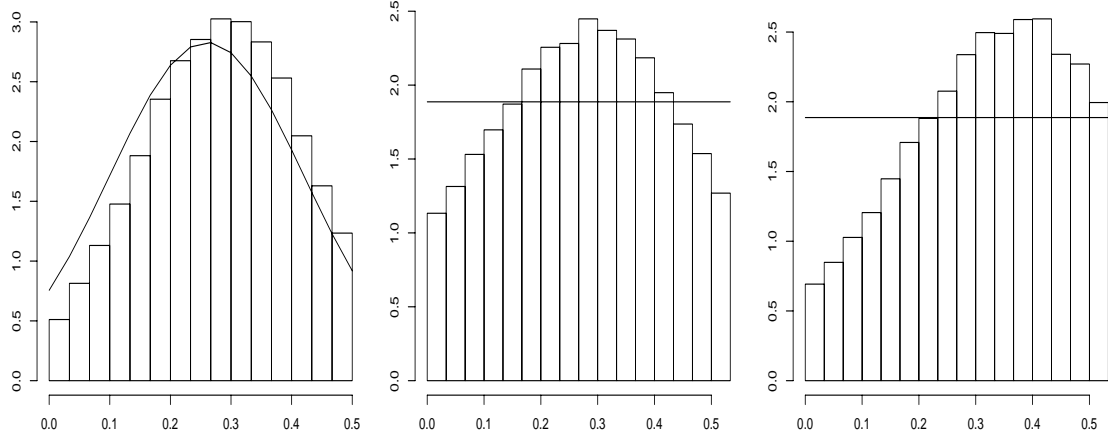
21

Figure 4: Estimated distribution of the typical radius under the fitted (A, L, N_{cc})-interaction models with parameters as estimated in Table 2 when using the the reference processes (R1)-(R3) (from the left to the right). The solid lines show the densities of the typical radius under the corresponding reference processes.

The distribution of R was estimated as follows. For each of the fitted (A, L, N_{cc})-interaction models from Table 2, let $n = 1,000,000$ be the number of MCMC iterations, k_i the number of discs and $r_j^{(i)}$ the radius of the j-th disc in the i-th iteration, and

$$F_i(r) = \frac{1}{k_i} \sum_{j=1}^{k_i} \mathbf{1}_{[r_j^{(i)} \leq r]}$$

the empirical distribution function of the radii obtained at the i-th iteration. The histogram in Figure 4 is obtained from the average of empirical distribution functions

$$\bar{F}(r) = \frac{1}{n} \sum_{i=1}^{n} F_i(r)$$

which may be interpreted as an estimate of the distribution of a typical radius R under the estimated (A, L, N_{cc})-interaction model, and which we refer to as the (estimated) mark

22

distribution. In fact we obtained a very similar distribution if we instead consider the average

$$\tilde{F}(r) = \frac{1}{\sum_{i=1}^{n} k_i} \sum_{i=1}^{n} \sum_{j=1}^{k_i} \mathbf{1}_{[r_j^{(i)} \leq r]}.$$

(not shown in this paper).

5.3 Results based on the unconditional likelihood

Rather different conclusions may be obtained when inference is based on the unconditional likelihood L_u in (10). This is exemplified in Table 2 when comparing the results in columns 2-4 (based on L_c) with those in columns 5-7 (based on L_u). Probably this difference is mainly due to edge effects.

6 Model control

In this section, we consider results when checking the fitted (A, L, N_{cc})-interaction models based on the conditional likelihood L_c, but similar conclusions are obtained when instead the unconditional likelihood L_u is considered.

Figure 5 shows simulated realizations of the random-disc Boolean models (R1)-(R3) and the fitted (A, L, N_{cc})-interaction models. As expected, because of the too low area fraction, the realization under (R3) looks very different from the others, and the many small connected components obtained under (R1)-(R3) seem less frequent under the fitted (A, L, N_{cc})-interaction models. Since it is hard by eye to check how well the estimated models fit the heather dataset, we consider in the sequel different summary statistics specified by various contact distribution functions and covariance functions, as usually considered in stochastic
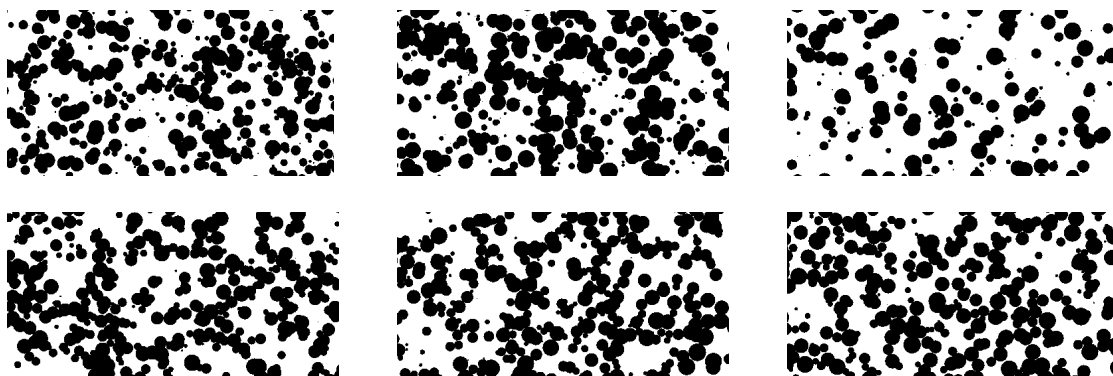
Figure 5: Simulations of the Boolean models (R1)-(R3) (first row from the left to the right) and of the fitted (A, L, N_{cc})-interaction models when using the the reference processes (R1)-(R3) (second row from the left to the right).

geometry (Stoyan *et al.*, 1995), and also shape characteristics obtained by certain operations from mathematical morphology (Ripley, 1988).

In Sections 6.1-6.2, we consider first a general planar random closed set **Z**. Assuming invariance properties such as stationarity or isotropy of **Z**, we naturally specify non-parametric estimates of the summary statistics with edge corrections, where we use a regular quadratic lattice with vertex set G given by the centroids of the 250×508 pixels in Figure 1. The non-parametric estimates can immediately be used for checking the Boolean models (R1)-(R3) which are stationary. They can also be used for checking our fitted (A, L, N_{cc})-interaction models from Table 2, which are finite (marked) point processes and hence non-stationary, since we supply the non-parametric estimates with 2.5% and 97.5% envelopes obtained from 39 simulations under a fitted model (see e.g. Møller and Waagepetersen, 2004, Section 4.3.4).

6.1 Contact distribution functions

Given a convex compact set $B \subset \mathbb{R}^2$ containing the origin, define

$$D = \inf\{r \geq 0 : \mathbf{Z} \cap rB \neq \emptyset\}$$

where $rB = \{(rx, ry) : (x, y) \in B\}$. Assuming $\mathrm{P}(D > 0) > 0$, the contact distribution with structuring element B is defined by

$$H_B(r) = \mathrm{P}(D \leq r | D > 0), \quad r \geq 0.$$

In the stationary case of Z, a natural non-parametric estimator is given by

$$\hat{H}_B(r) = \frac{\sum_{u \in G} \mathbf{1}[u \notin \mathbf{Z}, \, u + rB \subset W, \, (u + rB) \cap \mathbf{Z} \neq \emptyset]}{\sum_{u \in G} \mathbf{1}[u \notin \mathbf{Z}, \, u + rB \subset W]} \tag{12}$$

where we have used the border method (also known as minus sampling) to correct for edge effects and therefore only consider vertices u with $u + rB \subset W$. For the structuring element B, we use

- a unit line segment with endpoints $(-0.5 \cos \varphi, -0.5 \sin \varphi)$ and $(0.5 \cos \varphi, 0.5 \sin \varphi)$, where $0 \leq \varphi < \pi$, and we write H_φ for H_B (the linear contact distribution function),

- a unit disc $b(0, 1)$, and write H_s for H_B (the circular/spherical contact distribution function),

- a unit square $[-0.5, 0.5] \times [-0.5, 0.5]$, and write H_q for H_B (the quadratic contact distribution function).

For our random-disc Boolean models (recalling that $\mu_1 = \mathrm{E}R$),

$$H_B(r) = 1 - \exp\left(-\beta\left(L(B)\mu_1 r + A(B)r^2\right)\right) \tag{13}$$

25

where $L(B) = 2$ in the special case of the linear contact distribution function. Let

$$T_B(r) = -\frac{1}{r} \log \left(1 - H_B(r)\right), \quad r > 0,$$

and denote $\hat{T}_B(r)$ the non-parametric estimate obtained by replacing $H_B(r)$ by (12). Then (13) implies that

$$T_\varphi(r) = 2\beta\mu_1, \quad T_s(r) = 2\beta\pi\mu_1 + \beta\pi r, \quad T_q(r) = 4\beta\mu_1 + \beta r. \tag{14}$$

Figure 6 compares the theoretical functions $T_B(r)$ given by (14) and either (R1), (R2), or (R3) with $\hat{T}_B(r)$ based on the data and its simulated 2.5% and 97.5% envelopes obtained under respective (R1)-(R3) and the corresponding three (A, L, N_{cc})-interaction models. None of the Boolean models (R1)-(R3) provide a satisfactory fit, as $T_B(r)$ is systematically below $\hat{T}_B(r)$, with the upper 97.5% envelope close to $\hat{T}_B(r)$ in most cases of (R1)-(R2), while this envelope is much below $\hat{T}_B(r)$ in case of (R3). In contrast Figure 6 reveals no problems with any of the three (A, L, N_{cc})-interaction models.

6.2 Covariance function

Assuming $\mathbf{Z}$ is both stationary and isotropic, its non-centred covariance function is defined by

$$C(r) = \mathrm{P}(u \in \mathbf{Z}, v \in \mathbf{Z})$$

for any two points $u, v \in \mathbb{R}^2$ with distance $\|u - v\| = r$. An unbiased and edge corrected non-parametric estimator based on the border method is given by

$$\hat{C}(r) = \frac{\sum_{u,v \in G} \mathbf{1}[\|u - v\| = r, \{u, v\} \subset \mathbf{Z}]}{\sum_{u,v \in G} \mathbf{1}[\|u - v\| = r]} \tag{15}$$
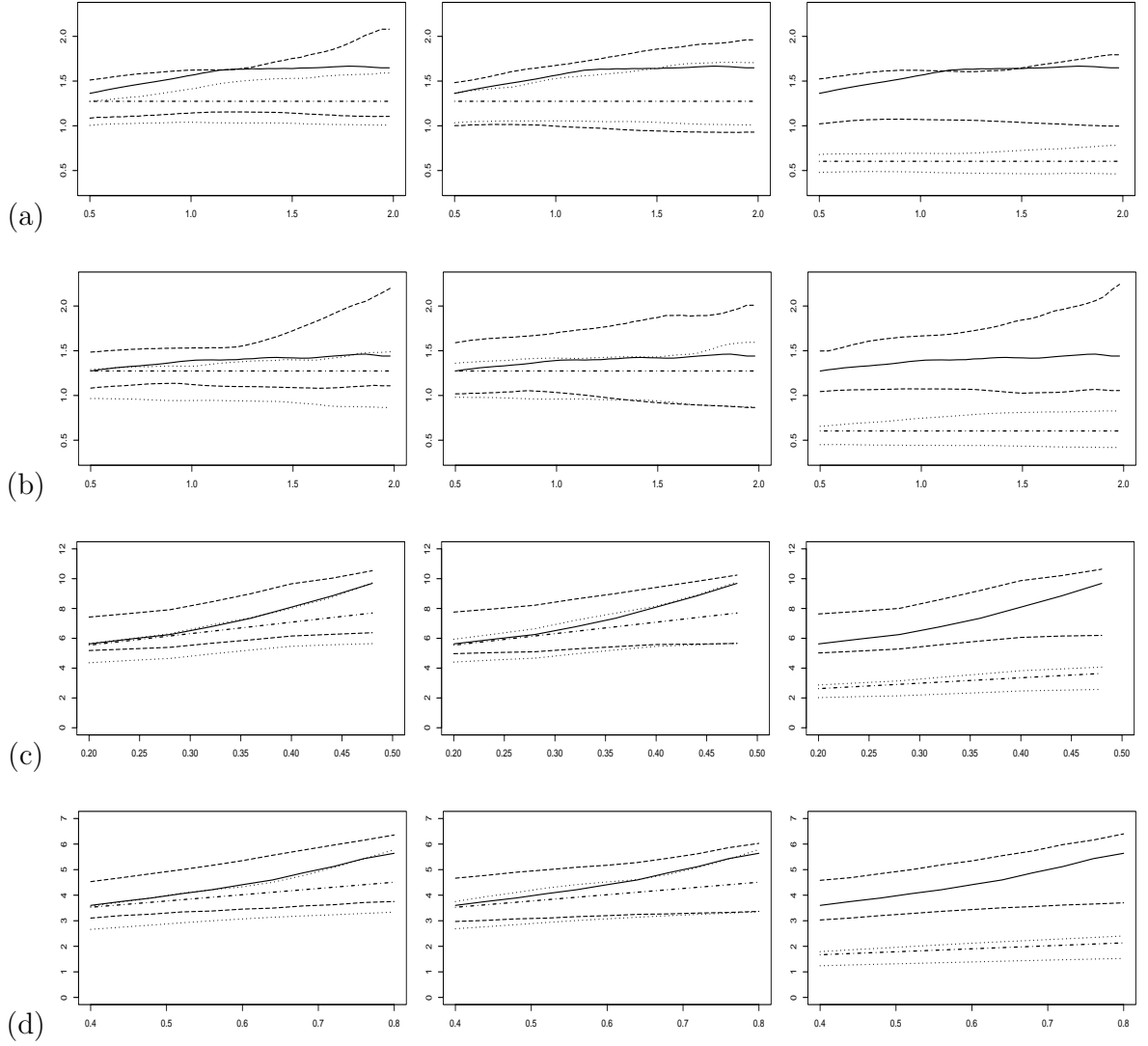
Figure 6: Comparing the theoretical functions $T_B(r)$ (dot-dashed lines) with $\hat{T}_B(r)$ (solid lines) based on the data and its simulated 2.5% and 97.5% envelopes obtained under the Boolean model (R1), (R2), or (R3) (dotted lines) and the corresponding (A, L, N_{cc})-interaction model (dashed lines). The three columns correspond from the left to the right to results when (R1), (R2), or (R3) is used either as a fitted model or as a reference process. For the rows, different structuring elements are specified by (a) the line segment with $\varphi = 0$, (b) the line segment with $\varphi = \pi/2$, (c) the unit disc $b(0, 1)$, and (d) the unit square $[-0.5, 0.5] \times [-0.5, 0.5]$.
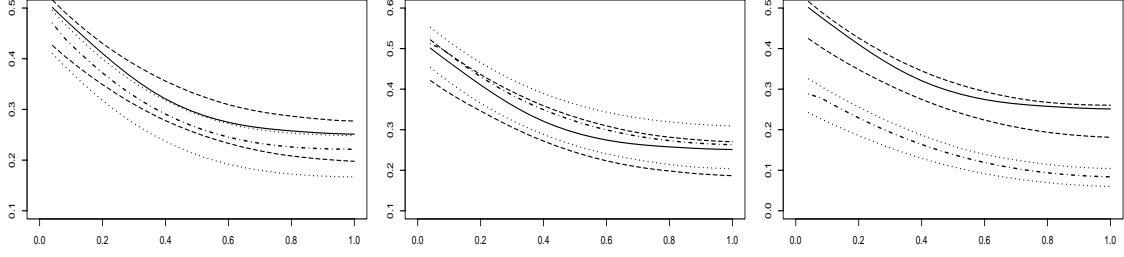
Figure 7: Comparing the theoretical functions $C(r)$ (dot-dashed lines) with $\hat{C}(r)$ (solid lines) based on the data and its simulated 2.5% and 97.5% envelopes obtained under the Boolean model (R1), (R2), or (R3) (dotted lines) and the corresponding (A, L, N_{cc})-interaction model (dashed lines). The three columns correspond from the left to the right to results when (R1), (R2), or (R3) is used either as a fitted model or as a reference process.

provided the denominator is non-zero.

For a random-disc Boolean model,

$$C(r) = 2p - 1 + (1 - p)^2 \exp\left(\beta \mathrm{E}\left[2R^2 \arccos \frac{R}{2r} - \frac{r}{2}\sqrt{4R^2 - r^2}\right]\right) \tag{16}$$

where the expectation may be evaluated by numerical methods.

Figure 7 compares the theoretical function $C(r)$ given by (16) and either (R1), (R2), or (R3) with $\hat{C}(r)$ in (15) based on the data and its simulated 2.5% and 97.5% envelopes obtained under respective the Boolean models (R1)-(R3) and the corresponding three (A, L, N_{cc})-interaction models. The figure reveals no misfit for neither (R2) or any of the three (A, L, N_{cc})-interaction models. However, $\hat{C}(r)$ is very close to the 97.5%-envelope obtained for (R1), and there is a clear misfit in case of (R3).

6.3 Shape characteristics

Following Ripley (1988) in using operations from mathematical morphology, we describe the shape of $\tilde{\mathbf{Y}} = \mathbf{Y} \cap W$ as follows. For any set $B \subset \mathbb{R}^2$ and number $r > 0$, let $|B| = A(B)$, $B_{\ominus r} = \{u \in \mathbb{R}^2 : b(u,r) \subseteq B\}$, and $B_{\oplus r} = \cup_{u \in B} b(u,r)$. Erosion e_r, dilation d_r, opening o_r, and closing c_r of $\tilde{\mathbf{Y}}$ by the disc $b(0,r)$ are defined by

$$e_r = \frac{|\tilde{\mathbf{Y}}_{\ominus r}|}{|W_{\ominus r}|}, \quad d_r = \frac{|\tilde{\mathbf{Y}}_{\oplus r} \cap W_{\ominus r}|}{|W_{\ominus r}|}, \quad o_r = \frac{|(\tilde{\mathbf{Y}}_{\ominus r})_{\oplus r} \cap W_{\ominus 2r}|}{|W_{\ominus 2r}|}, \quad c_r = \frac{|(\tilde{\mathbf{Y}}_{\oplus r})_{\ominus r} \cap W_{\ominus 2r}|}{|W_{\ominus 2r}|}.$$

Figure 8 compares these shape-characteristics based on the data with simulated 2.5% and 97.5% envelopes obtained under respective the Boolean models (R1)-(R3) and the corresponding three (A, L, N_{cc})-interaction models. In case of dilation and closing, the figure reveals no clear misfit for neither (R2) or any of the three (A, L, N_{cc})-interaction models, while (R1) and particularly (R3) are not fitting well. Furthermore, in case of erosion and opening, the figure indicates that any of the six models is not fitting well, possibly since the heather data are rather smooth while a disc process is naturally more rugged.
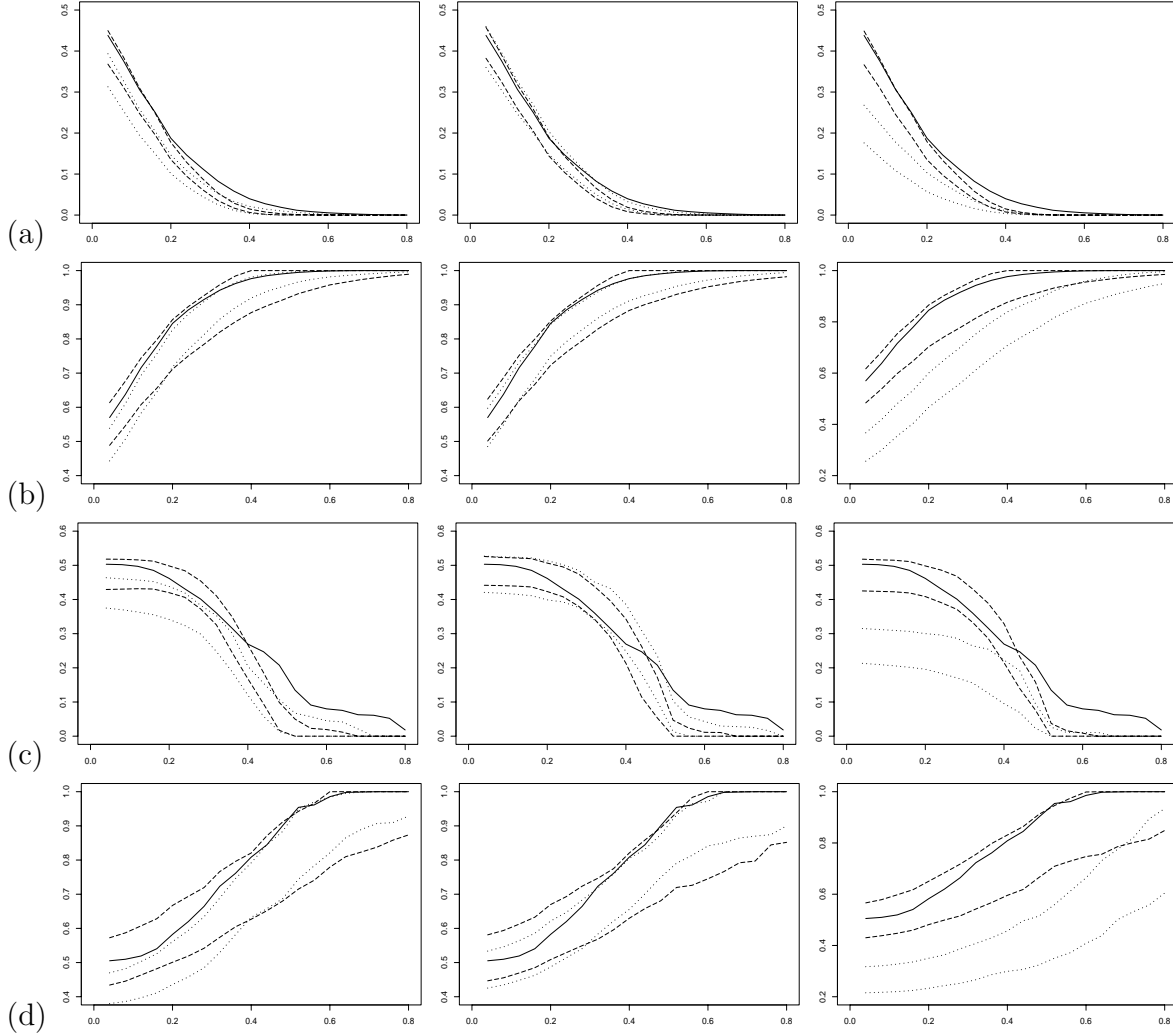
Figure 8: Comparing (a) erosion e_r, (b) dilatation d_r, (c) opening o_r, and (d) closing c_r of the heather data set (solid lines) with simulated 2.5% and 97.5% envelopes obtained under the Boolean model (R1), (R2), or (R3) (dotted lines) and the corresponding (A, L, N_{cc})-interaction model (dashed lines). The three columns correspond from the left to the right to results when (R1), (R2), or (R3) is used either as a fitted model or as a reference process.

References

Baddeley, A. and Møller, J. (1989). Nearest-neighbour Markov point processes and random sets. *International Statistical Review* **2**, 89–121.

Baddeley, A. J., van Lieshout, M. N. M. and Møller, J. (1996). Markov properties of cluster processes. *Advances in Applied Probability* **28**, 346–355.

Chin, Y. C. and Baddeley, A. J. (2000). Markov interacting component processes. *Advances in Applied Probability* **32**, 597–619.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New York, 2nd edition.

Dempster, A. P., Laird, N. M. and Rubin, D. B. (2004). Detecting dependence between marks and locations of marked point processes. *Journal of Royal Statistical Society Series B* **66**, 79–93.

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

Dupač, V. (1980). Parameter estimation in the Poisson field of discs. *Biometrika* **67**, 187–190.

Geyer, C. J. (1999). Likelihood inference for spatial point processes. In: *Stochastic Geometry: Likelihood and Computation* (eds. O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout), Chapman & Hall/CRC, Boca Raton, Florida, 79–140.

Hall, P. (1985). Counting methods for inference in binary mosaics. *Biometrics* **41**, 1049–1052.

Hall, P. (1988). *Introduction to the Theory of Covarage Processes*. Wiley, New York.

Hanisch, K.-H. (1981). On classes of random sets and point processes. *Serdica* **7**, 160–167.

Kendall, W., van Lieshout, M. and Baddeley, A. (1999). Quermass-interaction processes: conditions for stability. *Advances in Applied Probability* **31**, 315–342.

Laslett, G. M., Cressie, N. and Liow, S. (1985). Intensity estimation in a spatial model of overlapping particles. Upublished manuscript, Division of Mathematics and Statistics, CSIRO, Melbourne.

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

Møller, J. (1999). Markov chain Monte Carlo and spatial point processes. In: *Stochastic Geometry: Likelihood and Computation* (eds. O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout), Monographs on Statistics and Applied Probability 80, Chapman and Hall/CRC, Boca Raton, 141–172.

Møller, J. and Helisová, K. (2008). Power diagrams and interaction processes for unions of discs. *Advances in Applied Probability* **40**, 321–347.

Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.

Mrkvička, T. and Rataj, J. (2008). On estimation of intrinsic volume densities of stationary random closed sets. *Stochastic Processes and their Applications* **118**, 213–231.

Renshaw, E. and Ford, E. D. (1983). The interpretation of process from pattern using two-dimensional spectral analysis: Methods and problems of interpretation. *Applied Statistics* **32**, 51–63.

Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.

Ripley, B. D. and Kelly, F. P. (1977). Markov point processes. *Journal of the London Mathematical Society* **15**, 188–192.

Serra, J. (1980). The Boolean model and random sets. *Computer Graphics and Image Processing* **12**, 99–126.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. Wiley, Chichester, 2nd edition.

Weil, W. and Wieacker, J. (1984). Densities for stationary random sets and point processes. *Advances in Applied Probability* **16**, 324–346.

Weil, W. and Wieacker, J. (1988). A representation theorem for random sets. *Probability and Mathematical Statistics* **9**, 147–151.