A nonparametric measure of spatial interaction in point patterns

M. N. M. van Lieshout*

Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom

A. J. Baddeley

Department of Mathematics, University of Western Australia, Nedlands WA 6907, Australia, and Department of Mathematics and Computer Science, University of Leiden, The Netherlands

In memory of Philip Holgate

The strength and range of interpoint interactions in a spatial point process can be quantified by the function $J = (1 - G)/(1 - F)$, where G is the nearest-neighbour distance distribution function and F the empty space function of the process. $J(r)$ is identically equal to 1 for a Poisson process; values of $J(r)$ smaller or larger than 1 indicate clustering or regularity, respectively. We show that, for a large class of point processes, $J(r)$ is constant for distances r greater than the range of spatial interaction. Hence both the range and type of interaction can be inferred from J without parametric model assumptions. It is also possible to evaluate $J(r)$ explicitly for many point process models, so that J is also useful for parameter estimation. Various properties are derived, including the fact that the J function of the superposition of independent point processes is a weighted mean of the J functions of the individual processes. Estimators of J can be constructed from standard estimators of F and G. We compute estimates of J for several standard point pattern datasets and implement a Monte Carlo test for complete spatial randomness.

Key Words & Phrases: clustering, empty space function, J-statistic, Monte Carlo inference, nearest-neighbour distance distribution, Nguyen-Zessin formula, point process, spatial interaction, spatial statistics, regularity.

1 Introduction

The statistical analysis of a point pattern usually begins with the computation of estimates of the summary functions F (empty space function), G (nearest-neighbour distance distribution function) and K (reduced second moment function), defined e.g. in CRESSIE (1991), DIGGLE (1983), RIPLEY (1981), RIPLEY (1988). While these are useful descriptions of spatial pattern, and can easily be estimated from data, there

* mnl@stats.warwick.ac.uk

are very few stochastic models for which F, G or K is known analytically, so that parameter estimation and inference based on F, G, K are difficult.

Recall that, for a stationary point process, F is the distribution function of the distance from an arbitrary fixed point to the nearest random point of the process, and G of the distance from a point of the process to the nearest other point of the process. This paper advocates the use of

$$J(r) = \frac{1 - G(r)}{1 - F(r)}.$$

This is a nonparametric measure of the type of spatial interaction: the value 1 can be interpreted as indicating complete randomness or lack of interaction, while values less than 1 suggest "clustered" pattern and values greater than 1 suggest "regular" or "inhibitory" pattern.

We show that, for a very large class of point processes, the function J is constant for values of r larger than the effective range of spatial interaction. Hence J can be used to infer both the range and type of spatial interaction. Furthermore we are able to evaluate J explicitly for several stochastic models, so that it could be used directly for parameter estimation.

An appealing interpretation of J is that it compares the environment of a typical random point of the process with the environment of a fixed arbitrary point. $J(r)$ is the ratio of the probabilities, under these two situations, of the event that there are no points within a distance r of the given point. In terms of survival analysis, J is the ratio of the survival functions of the distance-to-nearest-point under these two situations; and our main result states that the hazard measures (GILL, 1994) of F and G are equal beyond the effective range of interaction r. The (signed) difference of hazard measures can be interpreted as a measure of spatial interaction.

Special cases of these results are implicit in the literature. The forms of F and G for a Neyman-Scott cluster process were derived by BARTLETT (1964); see AMBARTZUMIAN (1971), PALOHEIMO (1971a, b), BARTLETT (1975, pp. 8–9), and for detailed derivations DALEY and VERE-JONES (1988, §8.3, p. 243 ff), STOYAN et al. (1987, p. 143). For a Poisson cluster process

$$1 - G(r) = (1 - F(r))E(r)$$

where $E(r)$ is the probability that a randomly-chosen point in a typical cluster is more than r units distant from any other point belonging to the same cluster. Hence in particular if all offspring lie within a radius t of the parent point, $J(r)$ is constant for all $r \geqslant 2t$. Again, for a stationary, pairwise-interaction Gibbs process, STOYAN et al. (1987, p. 159) exhibit a relationship between $1 - F(r)$ and $1 - G(r)$ when r is exactly equal to the interaction distance R. In this paper we extend the relationship to all $r \geqslant R$.

Statistical inference based on comparisons between F and G has occasionally been suggested. DIGGLE (1979, (5.7)) proposed the statistic $D = \sup_r |\widehat{F(r)} - \widehat{G(r)}|$ as a measure of deviation from the Poisson process. Our proposal $(1 - G)/(1 - F)$ is in

some sense the canonical comparison of F and G as it arises directly from a fundamental point process formula.

The paper is organised as follows. In Section 2 we review the main techniques from spatial statistics that are used in the sequel. Section 3 introduces the J-function; the main theorem states that $J(r)$ is constant beyond the effective range of interaction. We also examine the behaviour of J under the basic operations of superposition and thinning, and show that the J-function of a superposition of independent processes is a convex combination of the J-functions of the components. The J-function is not invariant under thinning, in contrast to K.

In Section 4 we show that the J-function can be computed explicitly for a large class of point process models, including Poisson processes, Markov point processes, Neyman-Scott and Cox processes. For these examples at least, the classification of patterns as "clustered" or "regular" on the basis of their J-function values agrees with similar classifications on other grounds.

In Section 5 we discuss briefly how the J-statistic can be used for parameter estimation and statistical inference. Section 6 is a simple illustration on three standard data sets (see DIGGLE 1983), representing regular, random and clustered patterns.

2 Background

Throughout this paper we consider a stationary point process X in $\mathbb{R}^k$. For details of the theory of point processes see DALEY and VERE-JONES (1988), CRESSIE (1991) or STOYAN et al. (1987).

Define the empty space function F of X to be the distribution function

$$F(r) = \mathbb{P}\{ \rho(y, X) \leqslant r\}$$

of

$$\rho(y, X) = \min\{\|y - x\| : x \in X\},$$

the (Euclidean) distance from an arbitrary fixed point $y \in \mathbb{R}^k$ to the nearest point of the process. By stationarity, the definition of F does not depend on y.

Write $B(y, r) = \{x \in \mathbb{R}^k : \rho(x, y) \leqslant r\}$ for the closed ball of radius $r > 0$ centred at y in $\mathbb{R}^k$. Then $1 - F(r)$ is the probability that X puts no points in $B(y, r)$:

$$1 - F(r) = \mathbb{P}\{X \cap B(y, r) = \varnothing\}.$$

For example, for a Poisson process of intensity λ in $\mathbb{R}^2$ we obtain $1 - F(r) = \exp(-\lambda \pi r^2)$. The function F has been variously dubbed the "empty space", "point-event distance" and "spherical contact" distribution function.

To define G we need the Palm distribution $\mathbb{P}^y$ of X at $y \in \mathbb{R}^k$, which can be regarded as the conditional distribution of the entire process given that there is a point of X

at y (DALEY and VERE-JONES, 1988, chap. 12, CRESSIE, 1991, pp. 630–631, STOYAN et al. 1987, p. 110 ff.). Then define

$$G(r) = \mathbb{P}^y\{\rho(y, X\backslash\{y\}) \leqslant r\};$$

again this does not depend on y, by stationarity. Thus G is the distribution function of the distance from a point of the process to the nearest other point of the process, and is known variously as the "nearest-neighbour" or "event-event" function.

It is convenient to use the reduced Palm distribution $\mathbb{P}^{!y}$ defined as the distribution of $X\backslash\{y\}$ under $\mathbb{P}^y$, i.e. the conditional distribution of the rest of the process given that there is a point at y. Then the definition of G reads

$$G(r) = \mathbb{P}^{!y}\{\rho(y, X) \leqslant r\}$$

in harmony with the definition of F. For example, for a stationary Poisson process of intensity λ, the reduced Palm distribution $\mathbb{P}^{!y}$ is identical to $\mathbb{P}$, and $G \equiv F$.

Our main tool will be the following formula due to NGUYEN and ZESSIN (1979) which relates the reduced Palm distribution of X to its (ordinary) distribution:

$$\lambda \mathbb{E}^{!y} f(X) = \mathbb{E}[\lambda(y; X) f(X)] \tag{1}$$

holding (under suitable conditions on X) for any bounded nonnegative measurable function on the space of realisations of X. See also GLÖTZL (1980a, b), KOZLOV (1976), MATTHES et al. (1979), KALLENBERG (1983, 1984) and FIKSEL (1984, 1988), TAKACS (1983, 1986), SÄRKKÄ (1993) or DIGGLE et al. (1994), RIPLEY (1988, p. 54–55) for its application in parameter estimation. Here λ is the intensity of X and $\lambda(y; X)$ is the Papangelou conditional intensity of X at y. Roughly speaking, $\lambda(y; X)\,dy$ is the conditional probability of a point in the infinitesimal region dy centred at y given the configuration agrees with X outside this region (KALLENBERG, 1984). In statistical physics, $-\log \lambda(y; X)$ is interpreted as the energy needed to add y to the pattern X.

In other words, (1) states that $\mathbb{P}^{!y}$ is equivalent to the $\lambda(y; X)$-weighted distribution of X. In particular

$$\lambda = \mathbb{E}\lambda(0; X). \tag{2}$$

A necessary and sufficient condition (in the stationary case) for validity of (1) is that $\mathbb{P}^{!y}$ be absolutely continuous with respect to $\mathbb{P}$, whereupon $\lambda(y; X)$ is a.s. uniquely defined by (1). Formula (1) holds in particular for all stationary Gibbs point processes (PRESTON, 1976, RIPLEY, 1988) and for Poisson cluster processes when the cluster distribution is absolutely continuous. The corresponding expressions for $\lambda(y; X)$ are given in Section 4. Examples of processes which fail to satisfy (1) are randomly translated grids, and cluster processes consisting of pairs of points separated by a fixed distance.

3 The *J*-function

DEFINITION 1. *For a stationary point process X define*

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \tag{3}$$

for all $r \geq 0$ such that $F(r) < 1$.

For example, if X is a Poisson process then $F \equiv G$, so we obtain $J(r) \equiv 1$. Note that, even in a completely nonparametric context, the function J has an interpretation as the ratio of the survival functions of the distance to the nearest (other) point of X from (a) a point of the process, (b) a fixed arbitrary point. Values $J(r) < 1$ indicate that the survival function in situation (a) is smaller than that for (b), which may be interpreted as indicating "clustered" pattern; values $J(r) > 1$ indicate "regular" pattern. In the examples in Section 4 we will reconcile this with other definitions of "clustering" and "ordering".

Note that $J(0) = 1$. The denominator $1 - F$ is always absolutely continuous (BADDELEY and GILL, 1993) but the numerator $1-G$ need not be, so the discontinuity points of J are those of G. In general $1 - G(r)$ might be nonzero when $1 - F(r)$ is zero (e.g. for a randomly-translated unit square grid when $r = 1/\sqrt{2}$) but this does not occur for point processes of real interest (see Theorem 1 below).

THEOREM 1. *Let X be a stationary point process with intensity λ whose Papangelou conditional intensity $\lambda(y; X)$ exists.*
(a) $G(r) < 1$ *implies* $F(r) < 1$ *and*

$$J(r) = \mathbb{E}\left[\frac{\lambda(0; X)}{\lambda}\bigg| X \cap B(0, r) = \emptyset\right] \tag{4}$$

$$= \left(\mathbb{E}^{!0}\left[\frac{\lambda}{\lambda(0; X)}\bigg| X \cap B(0, r) = \emptyset\right]\right)^{-1}. \tag{5}$$

(b) *Suppose X has "interactions of finite range R" in the sense that $\lambda(0; X)$ is constant (and thus equal to $\lambda(0; \emptyset)$) for all point patterns X which contain no points in $B(0, R)$. Then*

$$J(r) = \frac{\lambda(0; \emptyset)}{\lambda} \quad for \ r \geq R. \tag{6}$$

PROOF: Let A be the event $\{X \cap B(0, r) = \emptyset\}$, so that $1 - F(r) = \mathbb{P}(A)$ and $1 - G(r) = \mathbb{P}^{!0}(A)$. Apply formula (1) to $f(X) = \mathbb{1}_A$, the indicator of the event A. Then the left-hand side of (1) equals $\lambda(1 - G(r))$. Hence

$$1 - G(r) = \mathbb{E}\left[\frac{\lambda(0, X)}{\lambda}\mathbb{1}_A\right].$$

If $F(r) = 1$ we have $\mathbb{P}(A) = 0$ so that the expectation above will be zero, i.e. $G(r) = 1$. Hence if $G(r) < 1$ we have $F(r) < 1$ and $\mathbb{P}(A) > 0$; dividing the expression above by

$1 - F(r) = \mathbb{P}(A)$ yields (4). Next, apply (1) to

$$g(X) = \frac{\mathbb{1}_A}{\lambda(0; X)}.$$

Then, the right-hand side of (1) is $\mathbb{E}[\lambda(0; X)g(X)] = 1 - F(r)$ giving

$$1 - F(r) = \lambda \mathbb{E}^{!0}\left[\frac{\mathbb{1}_A}{\lambda(0; X)}\right]. \tag{7}$$

Moreover, dividing (7) by $1 - G(r) = \mathbb{P}^{!0}(A)$ gives the reciprocal of (5).

In case (b), if $\lambda(0; X) \equiv \lambda(0; \emptyset)$ on A then

$$g(X) \equiv \frac{\mathbb{1}_A}{\lambda(0; \emptyset)}$$

so that the left side of (1) is

$$\lambda \frac{1}{\lambda(0; \emptyset)} \mathbb{P}^{!0}(A) = \frac{\lambda}{\lambda(0; \emptyset)} (1 - G(r))$$

yielding (6). $\qquad\qquad\square$

COROLLARY 1. *Let X be a stationary point process with intensity λ whose Papangelou conditional intensity $\lambda(y; X)$ exists. Then the process is regular with respect to J, $J(r) \geqslant 1$ if and only if*

$$\mathrm{Cov}\,(\lambda(0; X), \mathbb{1}\{X \cap B(0, r) = \emptyset\}) \geqslant 0 \tag{8}$$

and clustered with respect to J if the reverse holds.

PROOF: Use expression (4) to see that $J(r) \geqslant 1$ if and only if

$$\mathbb{E}[\lambda(0, X)\mathbb{1}\{X \cap B(0, r) = \emptyset\}] \geqslant \lambda \mathbb{P}(X \cap B(0, r) = \emptyset).$$

By (2), (8) follows. $\qquad\qquad\square$

Next we examine the behaviour of J under the basic point process operations of superposition and thinning.

THEOREM 2. *Let X_1, X_2 be independent, stationary point processes with intensities λ_1, λ_2 and J-functions J_1, J_2 respectively. Then the J-function of the superposition $X = X_1 \cup X_2$ is a convex combination of the J-functions of the components:*

$$J(r) = \frac{\lambda_1}{\lambda_1 + \lambda_2} J_1(r) + \frac{\lambda_2}{\lambda_1 + \lambda_2} J_2(r). \tag{9}$$

PROOF: Write F_i ($i = 1, 2$) for the empty space function of X_i and similarly let G_i denote the nearest-neighbour distribution function corresponding to X_i. Then, by independence, the empty space function F of the superposition equals

$$1 - F(r) = (1 - F_1(r))(1 - F_2(r)).$$

Writing $P^0_{(i)}$ for the Palm distribution (on the entire probability space) with respect to X_i, $i = 1, 2$ and P^0 for the Palm distribution with respect to X, we have (e.g. STOYAN et al., 1987, p. 116)

$$P^0 = \frac{\lambda_1}{\lambda_1 + \lambda_2} P^0_{(1)} + \frac{\lambda_2}{\lambda_1 + \lambda_2} P^0_{(2)}$$

Under $P^0_{(1)}$, X_1 and X_2 are independent, with X_1 governed by its Palm distribution (the Palm distribution of its marginal distribution) and X_2 by its (ordinary) marginal distribution. Similarly for $P^0_{(2)}$. Hence, the G-function of the superposition is

$$1 - G(r) = \frac{\lambda_1}{\lambda_1 + \lambda_2} (1 - G_1(r))(1 - F_2(r)) + \frac{\lambda_2}{\lambda_1 + \lambda_2} (1 - F_1(r))(1 - G_2(r)).$$

Dividing this by the identity for F above gives (9).

For comparison, the K-function of the superposition in the same situation is

$$K(r) = \frac{\lambda_1}{\lambda_1 + \lambda_2} \left[\frac{1}{\lambda_1 + \lambda_2} (\lambda_2 \pi r^2 + \lambda_1 K_1(r)) \right] + \frac{\lambda_2}{\lambda_1 + \lambda_2} \left[\frac{1}{\lambda_1 + \lambda_2} (\lambda_1 \pi r^2 + \lambda_2 K_2(r)) \right]$$

$$= \frac{1}{(\lambda_1 + \lambda_2)^2} [2\lambda_1 \lambda_2 \pi r^2 + \lambda_1^2 K_1(r) + \lambda_2^2 K_2(r)].$$

This equation can be verified by straightforward calculation using the identities in the proof above.

THEOREM 3. Let X_p be the process obtained from a stationary point process X by randomly deleting or retaining each point independently of other points, with retention probability $p > 0$. Then the J-function of X_p is

$$J_p(r) = \frac{Q^0_r(1 - p)}{Q_r(1 - p)} \tag{10}$$

where Q^0_r, Q_r are the generating functions of $n(X \cap B(0, r))$ under $\mathbb{P}^{!0}$ and $\mathbb{P}$ respectively. (The J-function of X itself is the case $p = 1$.)

PROOF: Let F_p, G_p be the F and G functions for X_p. Clearly $1 - F_p(r) = Q_r(1 - p)$. To prove $1 - G_p(r) = Q^0_r(1 - p)$ use the fact that the reduced Palm distribution of X_p coincides with the effect of random p-thinning on the reduced Palm distribution of X.   □

Thus while the K-function is invariant under random thinning (DIGGLE, 1983, ⁓7, STOYAN et al., 1987, p. 134), in general the J-function is not. There does not ⁓r to be a simple general relationship between J_p and J. This is not necessarily ⁓back, since a thinned process is generally different from the original process s of pattern or spatial interaction.

4 Examples

4.1 Poisson process

For a stationary Poisson process X of intensity λ we have $F \equiv G$ so that $J \equiv 1$. We could also derive this from Theorem 1 by observing that $\lambda(0; X) = \lambda$. It is interesting to note that J does not depend on the intensity parameter, a property also holding for the K-function, but not for F nor G.

4.2. Pairwise-interaction Markov point process

For a pairwise interaction point process (RIPLEY and KELLY, 1977, STOYAN et al., 1987, section 5.5) with activity constant β and interaction $\gamma(u, v)$ between points u, $v \in \mathbb{R}^k$,

$$\lambda(y; X) = \beta \prod_{x \in X} \gamma(x, y). \tag{11}$$

The process is *Markov* (in the sense of RIPLEY and KELLY, 1977) with interaction range R, if $\gamma(u, v) = 1$ when $\|u - v\| > R$. Examples include the *hard core process* defined by

$$\gamma(u, v) = \begin{cases} 0 & \text{if } \|u - v\| \leqslant R \\ 1 & \text{otherwise} \end{cases} \tag{12}$$

and the *Strauss process* defined by replacing 0 in (12) by a constant $0 < \gamma < 1$.

THEOREM 4. *For a Markov pairwise-interaction process with interaction range R,*
(a) $J(r)$ *is defined for all r;*
(b)

$$J(r) = \frac{\beta}{\lambda} \quad \text{for } r \geqslant R; \tag{13}$$

(c) *for "purely inhibitive" interactions,* $\gamma(u, v) \leqslant 1$, $\beta/\lambda \geqslant 1$;
(d) *for the hard core process* $J(r) = 1/(1 - F(r))$ *for* $r < R$, *and in particular J is continuous and monotone increasing for* $r < R$. *Furthermore* $J(r) = 1/(1 - \lambda m(B(0, r)))$ *for* $r < R/2$, *where m denotes Lebesgue measure.*

Thus, the hard-core and Strauss processes yield values outside the interaction radius indicating "regular" pattern in the sense of Corollary 1. Equation (13) was implicitly computed in STOYAN et al. (1987, (5.5.18), p. 159) for the value $r = R$ only.

PROOF: The product in (11) depends only on points $x \in X$ with $\|x - y\| \leqslant R$, so $\lambda(y; X)$ depends only on $X \cap B(y, R)$. Hence X has finite range interaction in the sense of Theorem 1(b) with $\lambda(0; \emptyset) = \beta$, and we get (13).

For a purely inhibitive process $\lambda(0; X) \leqslant \beta$ a.s. so that $\lambda \leqslant \beta$ using (2). This gives $J(r) \geqslant 1$ for $r \geqslant R$, proving (c).

For a hard core process (case (d)), clearly $G(r) = 0$ for $r < R$, so $J(r) = 1/(1 - F(r))$ for $r < R$. In particular J is monotone nondecreasing. Furthermore since

intensity measure λ. Note that by Jensen's inequality

$$F(r) = 1 - \mathbb{E}\exp[-A(B(0,r))] \leqslant 1 - \exp[-\mathbb{E}A(B(0,r))] = 1 - \exp[-\lambda\pi r^2]$$

and hence Cox processes are clustered with respect to F, i.e. they have larger empty spaces than a Poisson process with the same intensity.

THEOREM 6. *Let X be a Cox point process with driving random measure A which is stationary and a.s. nonatomic. The J-function of X is defined for all $r \geqslant 0$ and equals*

$$J(r) = \frac{\mathbb{E}^0\, e^{-A(B(0,r))}}{\mathbb{E}\, e^{-A(B(0,r))}} \tag{16}$$

where $\mathbb{E}^0$ denotes expectation with respect to the Palm distribution of A.

PROOF: This follows from the fact that the reduced Palm distribution of X is the distribution of a Cox process with driving measure distributed as the Palm distribution of A, cf. STOYAN et al. (1987, p. 141).   □

As an example, let X be a *mixed Poisson process* where $A = \alpha m(\cdot)$ for any nonnegative random variable α with finite positive expectation. Then the Palm distribution of A is simply the α-weighted distribution, and (16) reduces to

$$J(r) = \frac{\mathbb{E}[\alpha\, e^{-\alpha\pi r^2}]}{\mathbb{E}\alpha\, \mathbb{E}\, e^{-\alpha\pi r^2}}$$

It can be shown that J is monotonically decreasing with limit $(\text{ess inf}\,\alpha)/\mathbb{E}\alpha$. Hence a mixed Poisson process is clustered with respect to J, that is $J(r) \leqslant 1$ for all r. The inequality is strict unless α is constant almost surely.

5 Statistical aspects

5.1 Nonparametric estimation of J

Edge-corrected estimators for F and G based on observations of X within a bounded window $W \subseteq \mathbb{R}^k$ are reviewed in CRESSIE (1991, chap. 8), RIPLEY (1988, chap. 3) and STOYAN et al. (1987), pp. 122–131). For recent variations see BADDELEY and GILL (1993), BADDELEY et al. (1993), BARENDREGT and ROTTSCHÄFER (1991), DOGUWA (1989, 1990, 1992), DOGUWA and CHOJI (1991), DOGUWA and UPTON (1989, 1990), FIKSEL (1988) or STEIN (1990).

We propose estimating J by plugging into (3) estimates of F and G obtained by methods that are comparable to one another. For example one may estimate F by the standard "border correction" estimator (RIPLEY, 1988, chap. 3) and G by Hanisch's border correction estimator $\hat{G}_4$ (HANISCH, 1984). See also STOYAN et al. (1987, p. 128) where G is called D. These are Horvitz-Thompson type ratio estimators with comparable denominators, and are pointwise unbiased for F and pointwise approximately unbiased for G, respectively. Alternatively the Kaplan-Meier style estimators of F and G proposed in BADDELEY and GILL (1993) could be used. These have the advantage of being proper distribution functions (possibly defective), and

correspond to unbiased and approximately unbiased estimators of the hazard measures of F and G, respectively. Furthermore the estimator of F in BADDELEY and GILL (1993) has the same continuity properties as F itself.

We know little about the sampling properties of either estimator of J. Clearly $\widehat{J}(0) = 1$ always. It seems plausible that the relative error of J will increase with r, since this is true of standard estimators of $F(r)$ and $G(r)$ (BADDELEY and GILL, 1993), DOGUWA, 1992, DOGUWA and UPTON, 1989, 1990). Central limit theorems have been proved for $\hat{F}$ and $\hat{G}$ of both the Horvitz-Thompson and Kaplan-Meier types under various regimes (BADDELEY, 1980, BADDELEY and GILL, 1993, HEINRICH, 1988, JOLIVET, 1980, STEIN, 1991, CRESSIE, 1991, p. 480); a joint central limit theorem for $(\hat{F}, \hat{G})$, and hence for $\hat{J}$, has also been established.

Edge effects have a far greater influence on $\hat{G}$ than on $\hat{F}$ (RIPLEY, 1988, chap. 3, BADDELEY and GILL, 1993). The sampling properties of $\hat{G}$ and therefore of $\hat{J}$ may be particular cause for concern when the sampling window W is irregular, or in dimensions higher than two (BADDELEY and GILL, 1993, BADDELEY et al. 1993).

5.2 Estimation and inference based on J

Following are some speculative remarks on techniques for statistical inference based on J.

Model parameters θ can be estimated by matching the empirical value of a summary statistic to the theoretical one, e.g. by minimising

$$\int_0^{t_0} \{J(t, \theta)^c - \widehat{J(t)}^c\}^2 \, dt$$

where c is a suitably chosen constant, cf. DIGGLE (1983). If necessary, the theoretical value can be replaced by a Monte Carlo estimate. An advantage of using J instead of F or G separately may be its better mathematical tractability (see the results in Section 4).

Another possibility is the Takacs-Fiksel estimation method (FIKSEL, 1984, 1986, TAKACS 1983, 1986, RIPLEY, 1988, p. 54–55, DIGGLE et al., 1994, §2.4, SÄRKKÄ, 1993) since the basic equations (5)–(6) are special cases of (1) with the choice of $f = f_r$ given in the proof of Theorem 1.

For a Markov pairwise-interaction process, (13) gives the constant value of $J(r)$ for all $r > R$ in terms of the parameter β and the intensity λ. The intensity is determined by β and by the interaction function $\gamma(\,\cdot\,,\,\cdot\,)$ in a complex way. However λ may be estimated directly from the data, as $\hat{\lambda} = n(X \cap W)/m(W)$ in the usual way. If R is assumed known then β can be estimated via (13). This is semi-parametric estimation, since γ is unknown apart from the constraint that $\gamma(u, v) = 1$ for $\|u - v\| > R$.

Similarly, for an area-interaction process, (15) allows us to estimate the parameters β and $\eta = \gamma^{m(B(0,t))}$ given the interaction radius $R = 2t$.

Finally, note that estimation of the interaction distance R, in any of the models studied, amounts to estimating the largest interval $[R, \infty)$ on which J is constant.

6   Applications

We have taken three standard point pattern datasets discussed at length by DIGGLE (1983), dubbed *pines* ("Japanese pine saplings"), *redwood* ("Californian redwood seedlings") and *cells* ("biological cells"). These were exhibited as typical examples of random, clustered, and regular patterns respectively.

Figures 1 to 3 show the data and corresponding estimates $\hat{J}$ obtained using the Kaplan-Meier estimators (BADDELEY and GILL, 1993) of F and G. For pines the value of J is close to 1 for almost the entire range of r values except at high r values; for
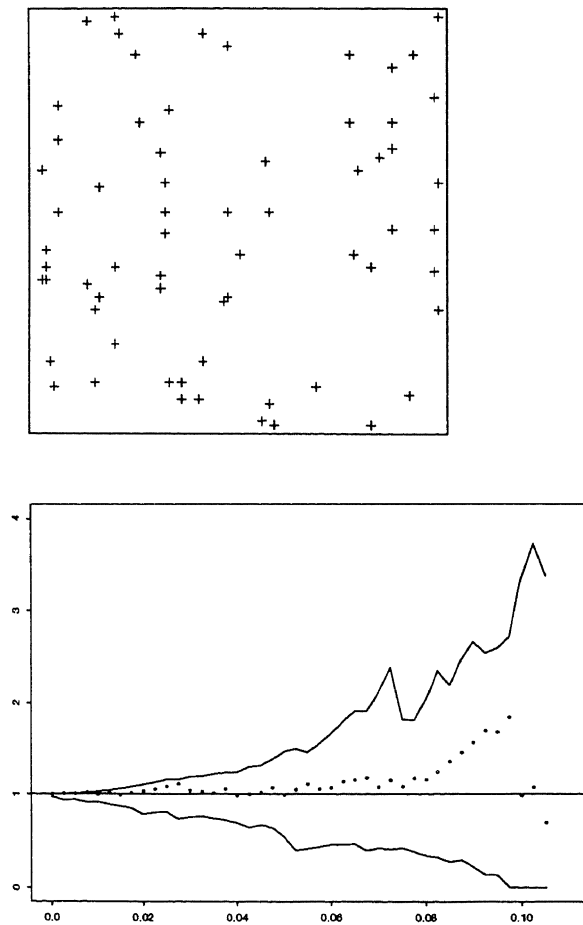
Fig. 1.   Above: pines data. Below: empirical J-function (dotted line) and envelope of 99 simulations of a binomial process with the same intensity (solid lines).
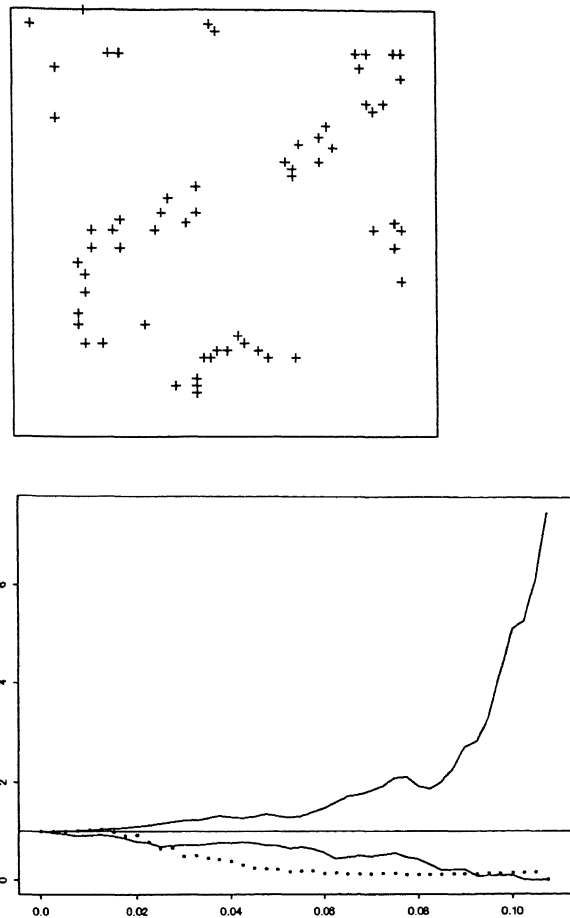
Fig. 2. Above: redwood data. Below: empirical J-function (dotted line) and envelope of 99 simulations of a binomial process with the same intensity (solid lines).

redwood it is below 1 and monotonically decreasing except for small fluctuations; and for cells it is above 1 for the entire range and is monotonically increasing.

We also computed the upper and lower envelopes from 99 simulations of a binomial process (independent uniform random points in the sampling window). Note that as r increases, the envelopes tend to span a wider range. For pines, the estimate $\hat{J}$ lies between the upper and lower envelopes over the entire range, in support of a Poisson model. For redwood $\hat{J}$ drifts below the lower envelope, suggesting aggregation, while for cells the estimated J-statistic exceeds the upper envelope at larger distances, suggesting regular pattern.

Graphical plots such as Figures 1 to 3 are useful as a first exploratory step in the analysis, indicating the type and range of spatial interaction and are helpful in
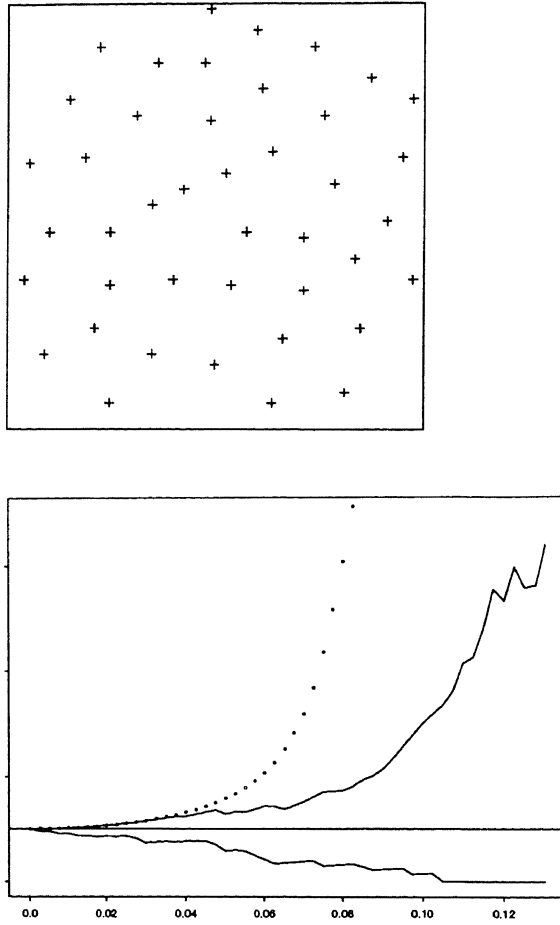
Fig. 3.   Above: cells data. Below: empirical J-function (dotted line) and envelope of 99 simulations of a binomial process with the same intensity (solid lines).

formulating plausible models. Furthermore, $\widehat{J(r_0)}$ for a specific value r_0,

$$\int_0^{r_0} \{\widehat{J(r)} - 1\}^2 \, dr$$

or

$$\sup_{r \leqslant r_0} |\widehat{J(r)} - 1|$$

can be used as Monte Carlo test statistics for departures from a Poisson model.

To investigate whether the empirical J-function can indicate the range of interaction as well, we simulated a Strauss process, conditional on 50 points with
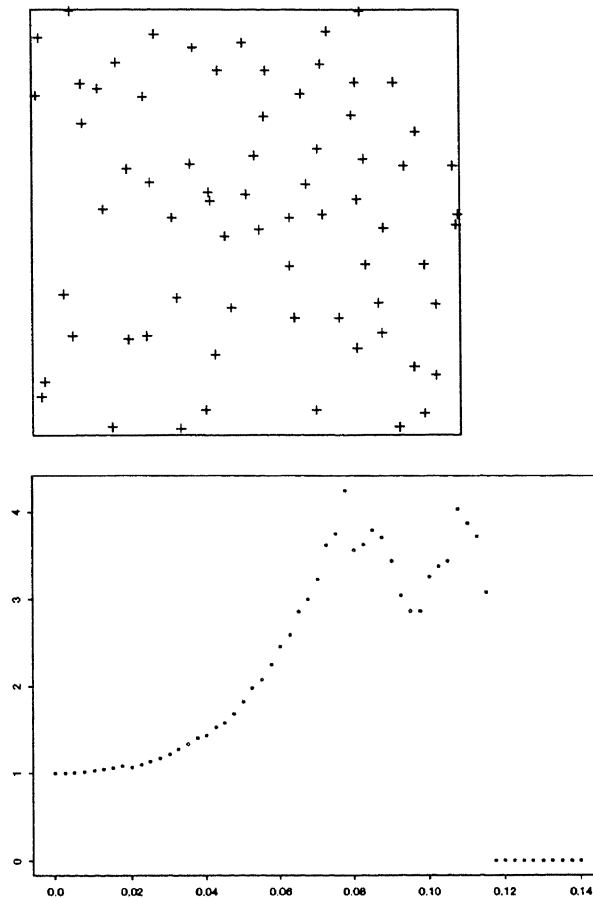
Fig. 4. Realisation from a Strauss process with 50 points, $r \approx 0.04$ and $\gamma = 0.1$ (above) and estimate of J (below).

$\gamma = 0.1$ and $r \approx 0.04$ (Fig. 4). Examining the estimated J-graph, there appears to be a positive bias and experience from change point techniques may prove helpful.

Acknowledgements

References

AMBARTZUMIAN, R. V. (1972), Palm distributions and superpositions of independent point processes in R^n P.A.W. Lewis, *Stochastic point processes: statistical analysis, theory, and applications*, 626–645. John Wiley and Sons, New York.

BADDELEY, A. J. (1980), A limit theorem for statistics of spatial data, *Advances in Applied Probability* **12**, 447–461.

BADDELEY, A. J. and R. D. GILL (1993), Kaplan-Meier estimators for interpoint distance distributions of spatial point processes, Research Report BS-R9315, Centrum voor Wiskunde en Informatica.

BADDELEY, A. J. and M. N. M. VAN LIESHOUT (1995), Area-interaction point processes, *Annals of the Institute of Statistical Mathematics* **47**, 601–619.

BADDELEY, A. J. and J. MØLLER (1989), Nearest-neighbour Markov point processes and random sets, *International Statistical Review* **57**, 89–121.

BADDELEY, A. J., M. N. M. VAN LIESHOUT, and J. MØLLER (1994), Markov properties of cluster processes, Technical Report 278, Department of theoretical statistics, University of Aarhus.

BADDELEY, A. J., R. A. MOYEED, C. V. HOWARD, and A. BOYDE (1993) Analysis of a three-dimensional point pattern with replication, *Applied Statistics* **42**, 641–668.

BARENDREGT, L. G. and M. J. ROTTSCHÄFER (1991), A statistical analysis of spatial point patterns. A case study, *Statistica Neerlandica* **45**, 345–363.

BARTLETT, M. S. (1964), The spectral analysis of two-dimensional point processes, *Biometrika* **51**, 299–311.

BARTLETT, M. S. (1975), *The statistical analysis of spatial pattern*, Chapman and Hall, London.

CRESSIE, N. A. C. (1991), *Statistics for spatial data*, John Wiley and Sons, New York.

DALEY, D. J. and D. VERE-JONES (1988), *An introduction to the theory of point processes*, Springer Verlag New York.

DIGGLE, P. J. (1979), On parameter estimation and goodness-of-fit testing for spatial point patterns, *Biometrics* **35**, 87–101.

DIGGLE, P. J. (1983), *Statistical analysis of spatial point patterns*, Academic Press, London.

DIGGLE, P. J., T. FIKSEL, P. GRABARNIK, Y. OGATA, D. STOYAN, and M. TANEMURA (1994), On parameter estimation for pairwise interaction processes, *International Statistical Review* **62**, 99–117.

DOGUWA, S. I. (1989), A comparative study of the edge-corrected kernel-based nearest neighbour density estimators for point processes, *Journal of Statistical Computation and Simulation* **33**, 83–100.

DOGUWA, S. I. (1990), On edge-corrected kernel-based pair correlation function estimators for point processes, *Biometrical Journal* **32**, 95–106.

DOGUWA, S. I. (1992), On the estimation of the point-object nearest neighbour distribution $F(y)$ for point processes, *Journal of Statistical Computation and Simulation* **41**, 95–107.

DOGUWA, S. I. and D. N. CHOJI (1991), On edge-corrected probability density function estimators for point processes, *Biometrical Journal* **33**, 623–637.

DOGUWA, S. I. and G. J. G. UPTON (1989), Edge-corrected estimators for the reduced second moment measure of point processes, *Biometrical Journal* **31**, 563–675.

DOGUWA, S. I. and G. J. G. UPTON (1990), On the estimation of the nearest neighbour distribution, $G(t)$, for point processes, *Biometrical Journal* **32**, 863–876.

FIKSEL, T. (1984), Estimation of parametrized pair potentials of marked and non-marked Gibbsian point processes, *Elektronische Informationsverarbeitung und Kybernetik* **20**, 270–278.

FIKSEL, T. (1988), Edge-corrected density estimators for point processes, *Statistics* **19**, 67–75.

FIKSEL, T. (1988), Estimation of interaction potentials of Gibbsian point processes, *Statistics* **19**, 77–86.

GILL, R. D. (1994), Lectures on survival analysis, P. Bernard, *22e Ecole d'Eté de Probabilités de Saint-Flour 1992*, Lecture Notes in Statistics, Vol. **1581**, Springer, New York, 115–241.

GLÖTZL, E. (1980), Bemerkungen zu einer Arbeit von O.K. Kozlov, *Mathematische Nachrichten* **94**, 277–289.

GLÖTZL, E. (1980), Lokale Energien und Potentiale für Punktprozesse, *Mathematische Nachrichten* **96**, 195–206.

HANISCH, K.-H. (1984), Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point patterns, *Mathematische Operationsforschung und Statistik, series Statistics* **15**, 409–412.

HEINRICH, L. (1988), Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes, *Statistics* **19**, 87–106.

JOLIVET, E. (1980), Central limit theorem and convergence of empirical processes for stationary point processes in: P. BASTFAI and J. TOMKO (eds.), *Point processes and queueing problems*, 117–161, North-Holland, Amsterdam.

KALLENBERG, O. (1983), *Random measures*, Akademie Verlag/Academic Press, Berlin/New York, third edition.

KALLENBERG, O. (1984), An informal guide to the theory of conditioning in point processes, *International Statistical Review* **52**, 151–164.

KOZLOV, O. K. (1976), Gibbsian description of point random fields, *Theory of Probability and its Applications* **21**, 339–355.

MATTHES, K., J. WARMUTH and J. MECKE (1979), Bemerkungen zu einer Arbeit von X.X. Nguyen und H. Zessin, *Mathematische Nachrichten* **88**, 117–127.

NGUYEN, X. X. and H. ZESSIN (1979), Integral and differential characterization of the Gibbs process, *Mathematische Nachrichten* **88**, 105–115.

PALOHEIMO, J. E. (1971), Discussion contribution to the paper by Matérn, in: G. P. PATIL, E. C. PIELOU, and W. E. WATERS (eds.), *Statistical ecology, Volume 1: Spatial patterns and statistical distributions*, 210–212, Pennsylvania State University Press, University Park.

PALOHEIMO, J. E. (1971), On a theory of search, *Biometrika* **58**, 61–75.

PRESTON, C. J. (1976), *Random fields*, Springer Verlag, Berlin-Heidelberg-New York.

RIPLEY, B. D. (1981), *Spatial statistics*, John Wiley and Sons, New York.

RIPLEY, B. D. (1988), *Statistical inference for spatial processes*, Cambridge University Press, Cambridge.

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

SÄRKKÄ, A. (1993), Pseudo-likelihood approach for pair potential estimation of Gibbs processes, PhD thesis, University of Jÿvaskÿla.

STEIN, M. L. (1990), A new class of estimators for the reduced second moment measure of point processes, Technical Report 278, Department of Statistics, University of Chicago, Chicago, Illinois, USA.

STEIN, M. L. (1991), Asymptotically optimal estimation for the reduced second moment measure of point processes, Technical Report 300, Department of Statistics, University of Chicago, Chicago, Illinois, USA.

STOYAN, D., W. S. KENDALL, and J. MECKE (1987), *Stochastic geometry and its applications*, Springer-Verlag, Berlin.

TAKACS, R. (1983), Estimator for the pair-potential of a Gibbsian point process, Institutsbericht 238, Institut für Mathematik, Johannes Kepler Universität Linz, Austria.

TAKACS, R. (1986), Estimator for the pair potential of a Gibbsian point process, *Statistics* **17**, 429–433.