European Society of Computational Methods in Sciences and Engineering (ESCMSE)

Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)
vol. 2, no. 1-2, 2007, pp. 79-102
ISSN 1790–8140

Computing Limiting Stochastic Processes for Spatial Structure Detection [1]

J. Mateu[2]

Department of Mathematics, Campus Riu Sec, University Jaume I,
E-12071 Castelln, Spain

Received 7 March, 2005; accepted in revised form 10 March, 2006

Abstract: Data showing spatial structure often arise in many applied scientific fields in form of points spatially distributed within a planar region. The basic methodology for analyzing spatial point pattern data is well established though most of the literature does not deal with point processes coming from the limit of lattice processes. However, spatial point processes that can be obtained from the limit of a suitable sequence of auto-Poisson stochastic lattice processes can be regarded as a useful tool to analyze spatial patterns exhibiting random and ordered structures. In this paper, we present a theoretical and computational framework and develop some practical issues in terms of a simulation study and real-data analysis.

Keywords: Cell nuclei, Computational stochastic methods, Gibbs point processes, Lattice processes, Dislocations of crystals, Pseudolikelihood, Spatial arrangement, Strauss process.

Mathematics Subject Classification: 60G12, 60G55.

1 Introduction

Spatial point processes define that branch of the statistical methodology which is focused on developing statistical tools that can handle spatial dependencies. Usually the basic elements this branch works with are spatial locations within a region together with measurable variables attached to each location. These elements are often represented by coordinates in a planar region. Then, a spatial point pattern is a set of locations, irregularly distributed within a region of interest, which have been generated by some unknown random mechanism.

Data in the form of a set of points distributed within a region of space, arise in many different scientific contexts; examples include the locations of trees in a forest, of nests in a breeding colony of birds, of cell nuclei in a microscopic section of tissue (see Figure 9), atoms in a disordered solid, arrays of obstacles (point defects or precipitates) for the motion of dislocations in crystals (see Figure 9) or the arrangement of the fibers in a fiber-reinforced material. In many scientific disciplines, such as biology, geology, medical engineering, materials sciences, technology, etc. there

[1] Published electronically April 14, 2007
[2] Corresponding author. E-mail:mateu@mat.uji.es. Fax: +34.964.728429

is a wealth of opportunities for collecting spatial point pattern data (for example, point pattern images can easily be obtained using standard microscopical equipment).

To understand how the underlying microstructure determines some of the very specific properties of todays high-performance materials, it is essential to be able to quantitatively describe relevant (spatial) features of the two and three-dimensional form of microstructures. The data required comes from detailed measurements of flat or thin sections of material and the transformation from the two-dimensional image to the three-dimensional structure exploits the models and methods developed in a particular area of mathematics known as stochastic geometry. However, it is often not necessary to describe the microstructure of a material in its whole geometrical complexity. In many situations it is sufficient to represent a phase by a point field. Such a point field of particle centers completely characterizes the particle system, and quantities describing this particle system can simply be estimated from a sample of the corresponding point field.

In biological or neuro-anatomical applications, where the points of interest are the centres of cells, say, the spatial arrangement of the points is likely to be very highly structured. We might, therefore, consider modelling the cell centre positions. At one extreme, the cell centres might be thought of as a hard-core process, with the hard-core parameter being equal to the cell diameter. A pairwise interaction process (Ripley [26]; Diggle *et al.* [10]) might be a more appropriate model for cell centre data; this class of point process is described at length in this paper.

The Poisson point process is the simplest spatial point process which is characterized by its strong (even ultimate) independence property: the number of points in disjoints regions are independent. Although the Poisson model is very simple and tractable, it can often serve only as a first approximation when studying natural phenomena represented by points in the space. In practice, point patterns exhibit interactions between points, which, for instance, can yield clustered or repulsive behaviour (see Figure 1).

Markov point processes were introduced in Ripley and Kelly [27] and are typically defined by their densities with respect to a Poisson process. Since the introduction of Markov point processes in spatial statistics attention has focused on the special case of pairwise interaction models in which each configuration of points interacts only via pairs of points from this configuration. These provide a large variety of complex patterns starting from simple potential functions which are easily interpretable as attractive and/or repulsive forces acting among points. Pairwise interaction models are simple exponential families whose sufficient statistics are often related to the popular K-function (Ripley [26], Diggle [8]) and they are very amenable to simulation and iterative statistical methods. However, pairwise interaction point processes (pipp) do not seem to be able to produce clustered patterns in sufficient variety (Strauss [32], Kelly and Ripley [15]). This led to the definition of area-interaction processes (Baddeley and Lieshout [2]) and shot-noise-weighted processes (Lieshout and Molchanov [16]). Thus, pipp models have been widely used as models for regular spatial patterns (Ripley [26]; Ogata and Tanemura [20], [21]; Diggle [8]; Penttinen [24]; Tomppo [34] and Stoyan, Kendall and Mecke [31]). However, the use of a pipp model as the limit of auto-Poisson lattice processes has been rarely used in literature despite the wide range of possible applications in such different fields as biology, material sciences, medicine, geology, economics or finances.

On the other hand, Ripley and Kelly [27], Besag [3] and Besag, Milne and Zachary [4] mentioned and further outlined possible connections between pipp and lattice models. Within this context, the pseudolikelihood estimation technique provides a computationally straightforward method of parameter estimation for such models. However, despite the applicability of this set-up, no practical grounds of such theoretical set-up were found in these and other related papers.

Motivated by biological applications and materials sciences applications, where the points of interest may be the centres of cells in microscopic tissue sections, in this paper we present a modified mathematical justification to prove that any purely inhibitory pipp can be obtained as the limit of a sequence of auto-Poisson lattice schemes and within this context we develop the pseudolikelihood

estimating equations, using a more friendly computational method. We carry out a Monte Carlo simulation study to analyze the behaviour of the parameters of a particular pipp model derived using this technique. We also stress that this methodology has a wide range of applications in many scientific fields, particularly in biology, as pointed out in our real-case studies. In particular, the stochastic models used in this paper will be motivated by simple considerations of possible underlying natural mechanisms which may be relevant in other disciplines. The local conditioning approach (our approach here) has also the advantage that it is quite statistically efficient, it is rather easy to correct for edge-effects and povides similar results than other likelihood-based methods.

The plan of the paper is as follows. Section 2 provides the set-up and definition of pipp models. Section 3 proves that a pipp model is the limit of a sequence of auto-Poisson lattice stochastic processes and derives the pseudolikelihood function in this situation. Section 4 is devoted to the Strauss example. Section 5 presents a short review of likelihood-based methods. A simulation study for the Strauss process is presented in Section 6. The paper ends with two applications and a Section of further research. The practical analysis presented in this paper are carried out using the author computing software, SPPA, which proofs to be very useful under the methodological and computational tools shown here.

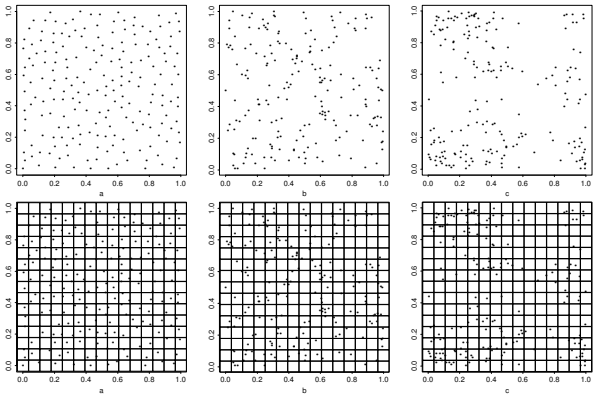

Figure 1: First row: Point patterns in the unit square showing three different spatial structures: *Left*, regular pattern; *Middle*, random pattern; *Right*, aggregated pattern. Second row: Spatial lattice of 15×15 contiguous quadrats superimposed over the point patterns

2 Set-up and definitions

A point process is a stochastic model governing the locations of events $\{s_i\}$ in some set X. Because our interest is in spatial point processes, we shall take X to be a bounded region in $\mathbf{R}^d$ or a torus, but more generaly X could be any locally compact Hausdorff space whose topology has a countable base (Cressie [6]). For example, events may be trees in a forest, towns in a geographic region, or epicenters of earthquakes. Here, it is assumed that the locations of events $\{s_i\}$ are the realization of some stochastic process with $s \in D$, where D is random.

Let $(X, \mathcal{F}, \tau)$ be a measure space such that $\tau(X) < \infty$, X is a bounded subset of $\mathbf{R}^d$ (the region within which the point process is defined), $\mathcal{F}$ its Borel σ-field and τ the Lebesgue measure on $(X, \mathcal{F})$. Let $(X_e, \mathcal{F}_e, \nu = \exp(\tau))$ be the corresponding *exponential space* (Carter and Prenter [5]). The measure ν can be identified as the distribution of a Poisson process of unit intensity. A typical point in X_e will be denoted by x. A measurable map from a probability space into $(X_e, \mathcal{F}_e)$ is a *point process* and the measure induced by such a map on $(X_e, \mathcal{F}_e)$ is the *distribution*

of the process. A point process is *simple* if its distribution is concentrated on the subset X_e^* of X_e, comprising all finite collections of distinct points from X. For any compact $K \subset \mathbf{R}^d$, $C(x \cap K)$ denotes the cardinality of $x \cap K$, i.e., number of points of the process inside K.

We are concerned with point processes P on $(X_e, \mathcal{F}_e)$ which are absolutely continuous with respect to ν. Any such process can be identified by its Radon-Nikodym derivative $f = dP/d\nu$ with respect to the unit Poisson process. The derivative f is also considered to be *hereditary* (Preston [25]; Ripley and Kelly [27]): $f(x) > 0$ implies $f(y) > 0$ for all $y \subset x$.

In this paper we restrict our attention to *pairwise-interaction point processes* (pipp) (Ripley [26]), for which given $S \subset \mathbf{R}^d$ a bounded Borel set, the probability density f has the special form

$$f(x) \propto \prod_{\xi \in x} \chi(\xi) \prod_{\{\xi,\eta\} \subseteq x, \xi \neq \eta} \phi(\{\xi, \eta\}) \tag{1}$$

where $\chi(.)$ is usually a positive constant related to the intensity of the process, ϕ is an interaction function, i.e., a symmetric measurable function mapping $\mathbf{R}^d \times \mathbf{R}^d$ to $[0, \infty)$ such that $\phi(\{\xi, \xi\}) = 1$ for all $\xi \in X$. The normalizing constant in (1) can be written as $f(\emptyset)$.

Note that not every pair (χ, ϕ) defines a valid probability density f in that a suitable normalizing constant may fail to exist. Such processes for which f turns out to be integrable are called purely *inhibitory.*

For a process defined by a hereditary density f with respect to ν, the conditional intensity (related to the Papangelou intensity (Papangelou [23]; Daley and Vere-Jones [7])) at ξ given x on $X \backslash \xi$ is

$$\lambda(\xi|x) = \begin{cases} \frac{f(x \cup \{\xi\})}{f(x)} & f(x) > 0 \\ 0 & otherwise. \end{cases} \tag{2}$$

For any pipp of the form (1),

$$\begin{aligned} \lambda(\xi|x) &= \chi(\xi) \prod_{\eta \in x} \phi(\{\xi, \eta\}) \\ \lambda(\xi|x \setminus \{\xi\}) &= \chi(\xi) \prod_{\eta \in x \setminus \{\xi\}} \phi(\{\xi, \eta\}). \end{aligned} \tag{3}$$

An important and well known special class of processes is generated by interaction functions of the form $\phi(\{\xi, \eta\}) = \phi(d\{\xi, \eta\})$, where $d(.)$ denotes the usual Euclidean distance in $\mathbf{R}^d$. Conditions to ensure integrability of f are given by Ruelle [29], Preston [25] or Ripley [26].

The interaction function $\phi(.)$ usually depends on a set of parameters, say θ, which have to be estimated. In such cases, the parametric interaction function is given by $\phi(.;\theta)$. In this paper we focus on the *pseudolikelihood* estimation method as it can be used routinely in applications and do not place artificial restrictions on the parametric form of the interaction function. Other general methods such as approximations to maximum likelihood based on numerical or Monte Carlo approximations to the normalizing constant (Ogata and Tanemura [20], [21], [22]; Penttinen [24]) or recursive approximation methods (Moyeed and Baddeley [19]), the Takacs-Fiksel method (Takacs [33]; Fiksel [11], [12]) or even a non-parametric method (Diggle *et al.* [9]) have also been proposed in literature. For a more detailed review see (Diggle *et al.* [10]; Geyer [13] and Ripley [28]). In Section 5 we give a brief summary of some of these methods only for comparison purposes.

If χ and ϕ are bounded, then for any finite point configuration $x \subset S$ with $f(x) > 0$, the pseudolikelihood for pipp models is defined by (Besag [3]; Jensen and Mller [14])

$$PL(x;\theta) = \exp(-\int_S \lambda(\xi|x) d\xi) \prod_{\xi \in x} \lambda(\xi|x \setminus \{\xi\}). \tag{4}$$

Usually, (4) is re-cast in terms of its logarithm. Maximization of (4) with respect to the parameter set θ yields the maximum pseudolikelihood estimators.

In a more particular context, Diggle *et al.* [10] derived the following pseudolikelihood estimating equation

$$PL(x;\theta) = \sum_{i=1}^{n} \log\{\lambda_0(\xi_i|x \setminus \{\xi_i\})\} - n \log\left\{\int_S \lambda_0(\xi|x)d\xi\right\}, \tag{5}$$

where $\lambda_0(\xi|x)$ is defined as $\prod_{\eta\in x}\phi(\{\xi,\eta\})$.

3 Approximation of pipp by a lattice process and connection to pseudolikelihood

In this section we present the theoretical basis for which a pipp can be obtained as the limit of a lattice process. The pseudolikelihood of the corresponding lattice processes is proved to converge to the pseudolikelihood of the pipp model. The proposed theory also defines an alternative (compared to the classical point of view-see Section 5) and easier way to compute the pseudolikelihood.

Suppose that $S = \cup_{j=1}^{m_i} C_{ji},\ i = 1, 2, \ldots$ are nested subdivisions with $\Delta_{ji} = |C_{ji}| > 0$ for all (i, j), and that both $m_i \to \infty$ and $\max \operatorname{diam} C_{ji} \to 0$ as $i \to \infty$ (where 'diam C_{ji}' is the diameter of the minimal ball containing C_{ji}). For each (i, j) let $\xi_{ji} \in C_{ji}$ be a given 'center' point of the cell C_{ji}. Then for each $i = 1, 2, \ldots$ we define a spatial point process $X^{(i)}$ on S as follows. If $N_j^{(i)} = X^{(i)}(C_{ji})$, the number of points of $X^{(i)}$ in cell C_{ji}, then $N^{(i)} = (N_j^{(i)})_{j=1,\ldots,m_i}$ is an auto-Poisson model with probability density function given by

$$p^{(i)}(n) \propto \prod_{j=1}^{m_i} [\Delta_{ji}\chi(\xi_{ji})]^{n_j}/n_j! \prod_{1\le k\le l\le m_i} \phi(\{\xi_{ki}, \xi_{li}\})^{n_k n_l} \tag{6}$$

for $n = (n_j)_{j=1,\ldots,m_i} \in \{0, 1, \ldots\}^{m_i}$ and $p^{(i)}(n) = 0$ otherwise. Furthermore, the m_i point processes $X_j^{(i)} = X^{(i)} \cap C_{ji},\ j = 1, \ldots, m_i$, are independent and conditionally on $N_j^{(i)} = n_j$ the n_j points in $X_j^{(i)}$ form a binomial process on C_{ji}.

An example of nested subdivisions of region S with a grid of small square cells, each of the same area, is given in Figure 1.

Then, the conditional distribution of $N_j^{(i)}$ given the 'rest' $N_{-j}^{(i)} = n_{-j}$ (where $n_{-j} = (n_k)_{k\neq j,\, k=1,\ldots,m_i} \in \{0, 1, \ldots\}^{m_i-1}$) is a Poisson distribution with mean

$$\nu_{ji}(n_{-j}) = \Delta_{ji}\chi(\xi_{ji}) \prod_{k:k\neq j} \phi(\{\xi_{ji}, \xi_{ki}\})^{n_k}. \tag{7}$$

On the other hand, the conditional density (with respect to ν) of $X^{(i)}$ given $N^{(i)} = n$, is given by

$$f^{(i)}(x|n) = \prod_{j=1}^{m_i} \exp(-\Delta_{ji})(n_j!/\Delta_{ji}^{n_{ji}})\mathbf{1}[x(C_{ji}) = n_j]. \tag{8}$$

Then, $X^{(i)}$ has density $f^{(i)}$ with respect to ν given by

$$f^{(i)}(x) = f^{(i)}(x|n)p^{(i)}(n), \tag{9}$$

which combined with (6) gives

$$f^{(i)}(x) \propto \prod_{j=1}^{m_i} \chi(\xi_{ji})^{n_j} \prod_{1\le k<l\le m_i} \phi(\{\xi_{ki},\xi_{li}\})^{n_k n_l} \tag{10}$$

where $n_j = x(C_{ji})$.

Let us now examine the limit behaviour of a sequence of processes $X^{(i)}$, $i = 1, 2, \cdots$. Using (10) we have that

$$\begin{aligned} f^{(i)}(x)/f^{(i)}(\emptyset) &= \prod_{j=1}^{m_i} \chi(\xi_{ji})^{n_j} \prod_{1\le k<l\le m_i} \phi(\{\xi_{ji},\xi_{ki}\})^{n_k n_l} \\ &\to \prod_{\xi\in x}^{m_i} \chi(\xi) \prod_{\{\xi,\eta\}\subseteq x,\ \xi\neq\eta} \phi(\{\xi,\eta\}) = f(x)/f(\emptyset) \end{aligned} \tag{11}$$

Then, for any finite point configuration $x \subset S$,

$$f^{(i)}(x)/f^{(i)}(\emptyset) \to f(x)/f(\emptyset) \quad \text{as} \quad i \to \infty. \tag{12}$$

Now, we define the pseudolikelihood based on data $X^{(i)} = x$ by

$$PL^{(i)}(x;\theta) = \prod_{j=1}^{m_i} p^{(i)}(n_j|n_{-j})/\Delta_{ji}^{n_j}. \tag{13}$$

Then using expression (7) and (13), the pseudolikelihood for the i-*th* point pattern is given by

$$PL^{(i)}(x;\theta) = \prod_{j=1}^{m_i} \{\exp(-\Delta_{ij}\chi(\xi_{ji}) \prod_{k:k\neq j} \phi(\{\xi_{ji},\xi_{ki}\})^{n_k}) \frac{\chi(\xi_{ji})^{n_j}}{n_j!} \prod_{k:k\neq j} \phi(\{\xi_{ji},\xi_{ki}\})^{n_j n_k})\}. \tag{14}$$

Finally, if we consider that χ and ϕ are bounded, then for any finite point configuration $x \subset S$ with $f(x) > 0$ and since n_j tends to 0 or 1 as $i \to \infty$,

$$\begin{aligned} PL^{(i)}(x;\theta) &\to \exp(-\int_S \chi(\xi) \prod_{\eta\in x} \phi(\{\xi,\eta\})d\xi) \prod_{\xi\in x} \chi(\xi) \prod_{\eta\in x\setminus\{\xi\}} \phi(\{\xi,\eta\}) \\ &= \exp(-\int_S \lambda(\xi|x)d\xi) \prod_{\xi\in x} \lambda(\xi|x\setminus\{\xi\}) = PL(x;\theta) \end{aligned} \tag{15}$$

where $\lambda(\xi|x) = f(x\cup\{\xi\})/f(x)$ is the conditional intensity given by (3) and $PL(x;\theta)$ is the general expression for the pseudolikelihood of a pipp model (4).

4 Example and interpretations

Attention in this Section is focused on the Strauss process, a well known and widely used parametric family of pipp models, having a specific interaction function depending on a scalar parameter. This model can precisely describe the spatial structure of a point pattern and minimize technical problems associated with the numerical optimisation. We develop the corresponding expressions of the density function $f(x)$ (1), the mean of the Poisson distribution (7) of the conditional distribution $p^{(i)}(n_j|n_{-j})$ and the pseudolikelihood (14).

Note that we can write (14) in terms of the mean value $\nu_{ji}(n_{-j})$,

$$PL^{(i)}(x;\theta) = \prod_{j=1}^{m_i}\{\exp(-\nu_{ji}(n_{-j}))\frac{\chi(\xi_{ji})^{n_j}}{n_j!}\prod_{k:k\neq j}\phi(\{\xi_{ji},\xi_{ki}\})^{n_j n_k})\}. \tag{16}$$

Taking logarithms on (16) we have

$$\begin{aligned}\log PL^{(i)}(x;\theta) &= \sum_{j=1}^{m_i} -\nu_{ji}(n_{-j}) + \sum_{j=1}^{m_i} n_j \log\chi(\xi_{ji}) - \sum_{j=1}^{m_i}\log n_j! \\ &\quad + \sum_{j=1}^{m_i} n_j \sum_{k:k\neq j}\log\phi(\{\xi_{ji},\xi_{ki}\})^{n_k}.\end{aligned} \tag{17}$$

4.1 A pipp model: the Strauss process

The Strauss process, due to Strauss [32], is characterized by having $\chi(\xi) = \lambda$ a constant, and an interaction function $\phi(\{\xi,\eta\}) = \rho^{\mathbf{1}_{[d(\xi,\eta)\leq R]}}$ which is only discontinuous at the Lebesgue null-set (in $\mathbf{R}^d \times \mathbf{R}^d$) where $d(\xi,\eta) = R$. The total number of neighbours of the pattern of $n(x)$ points is given by

$$s(x) = \frac{1}{2}\sum_{\xi,\eta}\mathbf{1}_{[d(\xi,\eta)\leq R]}, \tag{18}$$

and then, the density function (1) depending on $\theta = (\lambda,\rho)$ becomes

$$f(x) \propto \prod_{\xi\in x}\chi(\xi)\prod_{\{\xi,\eta\}\subseteq x,\xi\neq\eta}\phi(\{\xi,\eta\}) \propto \lambda^{n(x)}\rho^{s(x)}. \tag{19}$$

The case $\rho = 1$ defines a homogeneous planar Poisson process, whilst $\rho < 1$ defines a simple inhibition (regular) process. Strauss originally proposed (19) with $\rho > 1$ as a model for clustered patterns but, as subsequently pointed out by Kelly and Ripley [15], this violates the requirement of a finite normalising constant $f(\emptyset)$ in (19).

On the other hand, the conditional distribution $p^{(i)}(n_j|n_{-j})$ of $N_j^{(i)}$ given the 'rest' $N_{-j}^{(i)} = n_{-j}$, is given by

$$p^{(i)}(n_j|n_{-j}) = \exp(-\nu_{ji}(n_{-j}))\nu_{ji}(n_{-j})^{n_j}/n_j!, \tag{20}$$

where $\nu_{ji}(n_{-j})$, the mean of the Poisson distribution, is given by

$$\nu_{ji}(n_{-j}) = \Delta_{ji}\lambda\rho^{\sum_{k:k\in R(j)}n_k}. \tag{21}$$

and $k \in R(j)$ means that if celd $j = (r,s)$, then n_k is the number of points in the $k-th$ celd lying within the pre-defined range of j. If we write $t_j = \sum_{k:k\in R(j)} n_k$, then for a *first-order scheme* (see Figure 2) we have

$$t_j = n_{r-1,s} + n_{r+1,s} + n_{r,s-1} + n_{r,s+1} \tag{22}$$

and for a *second-order scheme* (see Figure 2) we have

$$t_j = t_{j(first-order)} + n_{r-1,s-1} + n_{r+1,s+1} + n_{r-1,s+1} + n_{r+1,s-1}. \tag{23}$$

The order schemes define the range of the pattern interaction structure and they represent the meaning of the interaction radius R in a more (not necessarily in a lattice system) general context. Note that Ripley [28] (Chapter 4) mentions that t_j is the total count for cells with at least part

within distance R of the centre of the cell j. This is equivalent to our definition of second-order scheme. So, our definition of t_j implicitely takes the value of the interaction radius R.

Finally, the local pseudolikelihood given by (17) becomes for the Strauss process

$$\begin{aligned}\log PL^{(i)}(x;\theta) &= -\lambda\sum_{j=1}^{m_i}\Delta_{ji}\rho^{t_j}+\sum_{j=1}^{m_i}n_j\log\lambda-\sum_{j=1}^{m_i}\log n_j! \\ &\quad +\sum_{j=1}^{m_i}n_j t_j\log\rho, \end{aligned} \tag{24}$$

which is proportional to

$$\log PL^{(i)}(x;\theta) = -\lambda\sum_{j=1}^{m_i}\Delta_{ji}\rho^{t_j}+\sum_{j=1}^{m_i}n_j[\log(\lambda\rho^{t_j})], \tag{25}$$

expression also derived by Ripley [28] (Chapter 4) considering $\Delta_{ji}=\Delta$, constant for all the grid cells. Note that expressions (24) and (25) depend on the interaction radius R through the definition, in terms of the order scheme, of t_j.

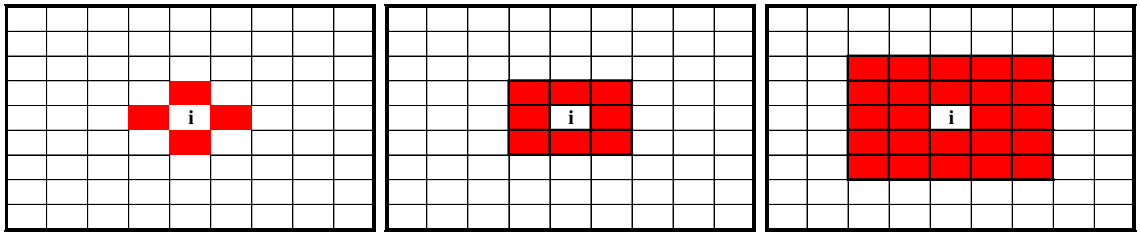

Figure 2: From left to right: first-, second- and third-order schemes for lattice processes

5 Likelihood-based methods for pipp models: a review

For most of pipp models (and more general Gibbs processes), the point process density (1) can be expressed in a general way as

$$f(x;\theta)=\frac{1}{c(\theta)}g(x;\theta), \tag{26}$$

where $g(x;\theta)$ is an unnormalized density belonging to the *exponential family*, and $c(.)$ is a normalizing constant which can be obtained through

$$c(\theta)=\int e^{<t(x),\theta>}\mu(d\mathbf{s}). \tag{27}$$

In (27), $t: X_e \to \Re^d$ is a map, and $<.,.>$ denotes the inner product. If $c(\theta)<\infty$ then $f(x;\theta)$ is the corresponding normalized density. For example, the Strauss process belongs to the exponential family and the unnormalized density (see (19)) takes the form

$$g(x;\theta)=e^{<t(x),\theta>}=\exp(n(x)\log(\lambda)+s(x)\log(\rho)). \tag{28}$$

Several methods have been proposed in literature for likelihood-based parametric estimation. Some of them focused on approximations to the normalizing constant and thus to the likelihood function, such as that of Penttinen [24]. Others focused on Monte Carlo likelihood approximation. Let us shortly review this method.

Given the normalized densities (26), the log likelihood for an observation x is given by $l(x;\theta) = \log g(x;\theta) - \log c(\theta)$. Let $X_1, X_2, ...$ be simulations from the known distribution given by $\psi \in \Theta$, a known parameter belonging to the neighbourhood of $\widehat{\theta}$. The log likelihood ratio against ψ can be estimated by

$$l(x;\theta,n) = \log \frac{g(x;\theta)}{g(x;\psi)} - \log \left(\frac{1}{n} \sum_{i=1}^{n} \frac{g(X_i;\theta)}{g(X_i;\psi)} \right) \tag{29}$$

Maximizing (29) gives a Monte Carlo approximation $\widehat{\theta}_n$ to the maximum likelihood estimator $\widehat{\theta}$.

Setting the normalized importance weights as (Geyer [13])

$$w_{n,\theta,\psi}(x) = \frac{e^{<t(x),\theta-\psi>}}{\sum_{i=1}^{n} e^{<t(X_i),\theta-\psi>}}, \tag{30}$$

we obtain the estimates of the score and Fisher information

$$\begin{aligned} \nabla l(x;\theta,n) &= t(x) - E_{n,\theta,\psi} t(X) \\ -\nabla^2 l(x;\theta,n) &= Var_{n,\theta,\psi} t(X), \end{aligned} \tag{31}$$

where, for any function g,

$$\begin{aligned} E_{n,\theta,\psi} g(X) &= \sum_{i=1}^{n} g(X_i) w_{n,\theta,\psi}(X_i) \\ Var_{n,\theta,\psi} g(X) &= E_{n,\theta,\psi} g(X) g(X)^T - \left[E_{n,\theta,\psi} g(X)\right] \left[E_{n,\theta,\psi} g(X)\right]^T. \end{aligned} \tag{32}$$

A Monte Carlo Newton-Raphson algorithm is a particular case when using formulas (31) and (32) in the special case when $\theta = \psi$ and putting all the importance weights equal to $\frac{1}{n}$. Then, the Newton-Raphson update of the Monte Carlo approximation to the maximum likelihood estimator consists of

$$\psi_{k+1} = \psi_k - \left[\nabla^2 l(x;\psi_k,n)\right]^{-1} \nabla l(x;\psi_k,n).$$

6 A simulation study

For this Section we have chosen the Strauss pipp model as a parametric family of point process models. This model has been widely used in spatial statistics, it is quite simple in terms of computational grounds and its parameters give a clear information about the spatial structure of the pattern. Consequently, conclusions drawn from this model are easily generalized and applied to other pipp models.

The goal of this simulation study is to evaluate if the pseudolikelihood estimation method applied to the lattice system by the local conditioning technique outlined in Section 3 works in practice and if it is useful to detect spatial point structures. In this sense we analyze the behaviour of the estimated model parameters, λ and ρ obtained through the pseudolikelihood estimation

under a variety of practical situations. Recall that basically there are three qualitatively different point structures (Diggle [8]): random, regular and aggregated (clustered) patterns. The values parameter ρ can take give hints about the spatial structure of the point process. Basically, $\rho = 1$ defines a random pattern, $\rho < 1$ indicates a regular inhibitory structure and values of $\rho > 1$ are indicative of clustered patterns. Recall that pipp models are not good models for fitting clustered patterns though can be used as descriptive statistics to detect them. If we finally end up with a clustered pattern we should think of models which are not just pipp models (Baddeley and Lieshout [2]).

6.1 Design of the simulation study

We considered three qualitatively different spatial point structures: random, aggregated and regular patterns, which should be identified by our methodology. Then we make use of the local conditioning method which approximates the pseudolikelihood function of a Strauss process through a lattice system. For each type of pattern structure we simulated 2000 realisations, each with an average of $n = 1000$ points in the unit square $(0, 1) \times (0, 1)$, except for regular structures for which we generated an average of $n = 800$ points per pattern (due to restrictions of the software SPPA [30] used).

Random patterns were generated according to a homogeneous Poisson process. Regular or inhibitory point patterns were defined through a hard-core distance, δ, using a sequential inhibition process based on the following facts: (a) x_1 is uniformly distributed in the region A; (b) Given $\{x_j, j = 1, ..., i-1\}$, x_i is uniformly distributed on the intersection of A with $\{y : d(y, x_j) \geq \delta\}$, for $j = 1, ..., i-1$. Finally, aggregated patterns were defined in terms of Poisson clustered processes as defined by Cressie [6]. These processes incorporate an explicit form of spatial clustering, and as such provide a more satisfactory basis for the modelling of aggregated spatial point patterns. They are defined based on the following three postulates: (a) Parent events form a Poisson process with a fixed intensity; (b) Each parent produces a random number of offsprings, realized independently and identically for each parent; (c) The positions of the offsprings relative to their parents are independently and identically distributed according to a bivariate normal density function.

For each pattern, and to apply the lattice approximation system, we used several grid sizes and order schemes as shown in Table 1. Each line of this table gives us information about the type of simulated pattern, the average number of points in each pattern (*Points per pattern*), the number of replications for each experiment (*Replications*) and the grid order values. All grid orders are magnitudes to the square, i.e., $(10 \times 10, 15 \times 15, 20 \times 20$, etc).

Table 1: Design of the simulation study for first and second-order schemes

Points per pattern	Replications	Grid order
Random pattern		
1000	2000	10, 15,...,35, 40, 50,..., 220, 230
Cluster pattern		
1000	2000	10, 20, 30, 40, 50, 60,..., 180, 190,200
Regular pattern		
800	2000	10, 20, 30, 40, 50, 60

For each combination of pattern, replication, grid size and order scheme we evaluated the estimate of parameters λ and ρ using the pseudolikelihood estimation method as described in preceeding Sections. Each combination yielded 2000 estimates of both parameters which are summarised in box-plot format. For each sample we tested the null hypothesis of $\rho = 1$ based on

the likelihood ratio test, the Wald test and the Lagrange multipliers test. As they require the Gaussian distribution, we carried out Kolmogorov-Smirnov tests to evaluate this assumption using the empirical parameters as maximum likelihood approximations to the theoretical ones.

6.2 Results

The simulation results of parameters λ and ρ for each combination are shown in Figures 3 to 8. The figures show the parameters box-plots with vertical axes measuring parameters magnitude, and horizontal axes the corresponding grid order.

6.2.1 Random structures

(Figures 3, 6). The convergence of λ parameter to the number of points per unit area (the intensity) for any order scheme is clear as the number of grids increase. In general standard deviations of λ parameter decrease with the grid size.

On the other hand, the value of parameter ρ always oscillates around the unit for any order scheme. In fact, the three statistical tests can not reject the null hypothesis $H_0 : \rho = 1$. The standard deviations of this parameter keep on growing up as the number of grids increases though this increment is not significantly important to reject the null hypothesis. Moreover, note that a bigger order scheme reduces the standard deviation.

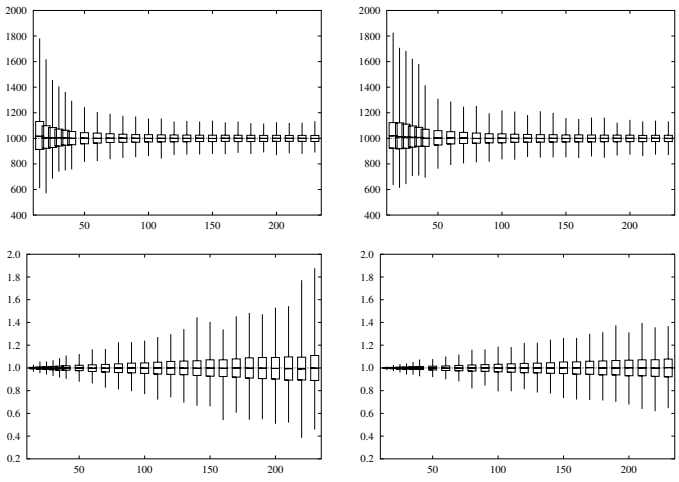

Figure 3: Boxplots of parameters λ (first row) and ρ (second row) for random spatial point patterns with an average of 1000 points in the unit square. First (left column) and second-order (right column) schemes and grid sizes from 10×10 to 230×230

6.2.2 Aggregated structures

(Figures 4, 7). The value of parameter λ clearly increases with the grid size until convergence is reached at around 900 indicating the importance of the grid size in a correct specification of this parameter under a clustered structure. In any case this method underestimates the true value of the pattern intensity. The corresponding standard deviation keeps within the same range for all grid size and order scheme.

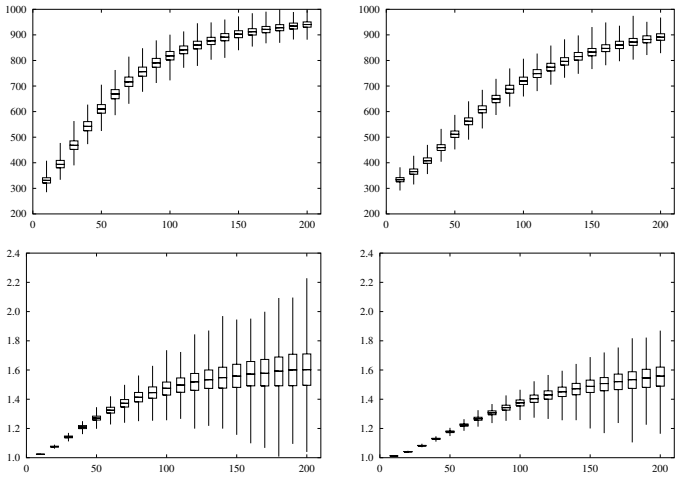

Figure 4: As in Figure 3 but for clustered spatial point patterns with an average of 1000 points in the unit square, and grid sizes from 10×10 to 200×200

The value of parameter ρ takes values above the unit and increases with the number of grids until reaching a comfortable value significantly different from 1 which is indicative of aggregated structures. The magnitude of the order scheme, i.e. the definition of the interaction structure, plays an important role in the value shown by this parameter. Given a fixed grid size, this value decreases as the order scheme increases. On the other hand, the larger the order scheme, the lower standard deviations parameter ρ shows. Recall that the results given in this Section can only be used as descriptive and no pipp model should be used to fit these aggregated structures.

6.2.3 Regular structures

(Figures 5, 8). The value of parameter λ increases when with the number of grids and clearly varies among the two order schemes. There is no clear structure between the values of standard deviations and number of grids with order schemes.

On the other hand, ρ values start to be significantly different from 1 (detecting a lack of randomness in favour to a regular structure) as the grid size increases, but it strongly depends on the magnitude of the order scheme. To detect the same strength of the repulsion between points we certainly need to specify a combination between the order scheme (equivalent to the interaction radius) and the grid size. This idea agrees with the one explored by Mller [18]. The standard deviations are kept within the same range for all grid sizes and order schemes. Analyzing the behaviour of the value of ρ, we suggest that there may be an optimum value for the grid size, given an order scheme.

7 Analysis of real data

In this Section we analyze two real case studies which show common interest in the fields of neuroanatomy and/or biology and the field of materials science engineering. Both data sets show spatial locations, scaled to the unit square $(0, 1) \times (0, 1)$ (Figure 9). The first one shows cells nuclei in a microscopic section of a tissue. The second data set shows arrays of obstacles (point defects or precipitates) for the motion of dislocations in crystals. The main goal here is to detect if the

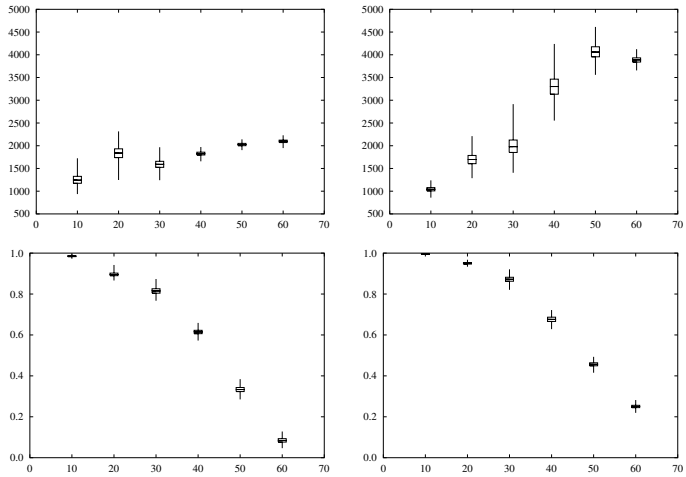

Figure 5: As in Figure 3 but for regular spatial point patterns with an average of 800 points in the unit square, and grid sizes from 10×10 to 60×60

points in both spatial patterns locate according to an underlying (natural, environmental factor) spatial structure based on repulsion, aggregation or randomness.

Two descriptive functions, the $G(\cdot)$ function (a nearest-neighbour method based on point-to-point nearest distances) and the $K(\cdot)$ function (usually called Ripley's K-function, a second-order characteristic), will be used to help us to distinguish among possible spatial structures (Diggle [8]; Ripley [28]; Cressie [6]; Stoyan, Kendall and Mecke [31]).

For comparison purposes, in Figure 10 we have shown these two functions acting on the three possible spatial structures. A simple test of randomness consists of calculating simulation envelopes of both G and K functions under this assumption and plotting the empirical data-based function together with the maximum and minimum envelopes. If the empirical function falls outside this envelope either at short distances or at long distances, there is evidence against the hypothesis assessed. The results based on both functions clearly reject the randomness hypothesis for both spatial patterns (Figure 11). Then we should learn more about the underlying spatial structure.

Table 2 shows the results of the pseudolikelihood estimation method based on the local conditioning approach for the cells data set. We report the results of parameter ρ for five different grid sizes and the two order schemes. The p-value of the Lagrange multipliers test is also shown. For any order scheme, the 10×10 grid size is not big enough to detect the regular pattern underlying the cells nuclei locations. The other grids do. However, in order to select the correct parameter value (and so the correct combination of grid size and order scheme) we can use goodness-of-fit statistical tests based on G and K functions. For comparison purposes we also show the results obtained with Penttinen approximate likelihood method, Newton-Raphson and MCMC likelihood methods (see Section 5). Both techniques, goodness-of-fit statistical tests and likelihood-based methods are based on simulating Strauss processes with the corresponding parameter values. This simulation has been carried out by Metropolis-Hastings algorithms with a random number of points as suggested by Geyer [13] and Mller [17].

We simulated the G and K functions from 99 realizations of a Strauss process with a given parameter set on the unit square region. Then, the maximum and minimum provide a simulation envelope. The best fit was obtained using a second-order scheme for grid size 40×40 which provided a value of $\rho = 0.588$ (Table 2). For shortness, it is only shown the G and K-based goodness-of-fit

study with a second-order scheme for grid sizes 30×30, 40×40 and 50×50 (Figures 12, 13). The only empirical K and G functions totally lying within the envelopes are that based on a 40×40 grid size. A second-order scheme within a grid size of 40×40 corresponds to an approximately interaction radius of $R = 0.05$. This value was confirmed by an additional exercise consisting of evaluating the profile pseudolikelihood 4 as a function of R for a given set of parameters and looking for the value of R that maximizes this profile function (Ripley [28]; Baddeley and Lieshout [2]). Finally, to reinforce our results, we evaluated the three likelihood-based methods with the following results: $\rho = 0.661$ (Penttinen), $\rho = 0.576$ (Newton-Raphson) and $\rho = 0.565$ (MCMC method). Except the approximate method of Penttinen, the other two more robust methods (Geyer [13]) show very close results to our.

Consequently, the spatial point pattern of cells nuclei follow a Strauss process with parameter $\rho = 0.588$ and interaction radius $R = 0.05$, indicating a clear underlying regular or repulsive spatial structure.

Table 3 shows the results of the pseudolikelihood estimation method based on the local conditioning approach for the arrays of obstacles (point defects or precipitates) for the motion of dislocations in crystals. We again report the results of parameter ρ for five different grid sizes and the two order schemes together with the p-value of the Lagrange multipliers test.

This time ρ parameter values are quite close to 1 though for certain combinations of grid size and order scheme this value is significantly different from 1. Again, to select the correct parameter value we use goodness-of-fit statistical tests based on G and K functions. Comments concerning the simulation of the Strauss process remain valid.

We simulated the G and K functions from 99 realizations of a Strauss process with a given parameter set on the unit square region and calculated the envelopes. The best fit was obtained using a second-order scheme for grid size 30×30 which provided a value of $\rho = 0.915$ and grid size 50×50 with a value of $\rho = 0.908$ (Table 3). The G and K-based goodness-of-fit study with a second-order scheme for grid sizes 30×30, 40×40 and 50×50 is shown in Figures 14 and 15. Note that both values of ρ parameter are similar and such a small difference is not enough to detect a change in the pattern structure. Moreover, empirical K and G functions for these two grid sizes seem to agree with the envelopes. A second-order scheme within a grid size of 30×30 corresponds to an approximately interaction radius of $R = 0.04$, and the corresponding value for grid size 50×50 is $R = 0.035$. These values were confirmed by the evaluation of the profile pseudolikelihood (4) as a function of R. The results shown by the likelihood-based methods were: $\rho = 0.858$ (Penttinen), $\rho = 0.960$ (Newton-Raphson) and $\rho = 0.885$ (MCMC method). Again, Newton-Raphson and MCMC methods show very close results to ours.

Table 2: Results for cells nuclei locations. sig_{LM}= p-value of the Lagrange multipliers test

Grid	ρ_1	sig_{LM}	ρ_2	sig_{LM}
10	0.985	0.316	0.987	0.082
20	0.913	0.002	0.931	0.001
30	0.724	0.000	0.734	0.000
40	0.598	0.000	0.588	0.000
50	0.741	0.000	0.672	0.000

Consequently, the arrays of obstacles for the motion of dislocations in crystals follow a Strauss process with parameter ρ around 0.90 and interaction radius $R = 0.04$, indicating an underlying weak regular or repulsive spatial structure.

Let us note that the resistance to the dislocation motion depends on distribution properties of the spatial point field of the obstacle's centers. The planar point field formed by the centers of the

Table 3: Results for arrays of obstacles (point defects or precipitates) for the motion of dislocation in crystals. sig_{LM}= p-value of the Lagrange multipliers test

Grid	ρ_1	sig_{LM}	ρ_2	sig_{LM}
10	0.981	0.117	0.994	0.422
20	0.971	0.277	0.995	0.791
30	0.862	0.000	0.915	0.002
40	0.815	0.000	0.809	0.000
50	0.897	0.086	0.908	0.036

dislocation etch pits in a single crystal carries decisive information about the dislocation arrangement. In fact, the the spatial distribution depends on the crystal growth conditions regarding the anisotropy and orientation of the crystal lattice.

So, the preferred configuration of dislocations in the above analyzed crystal is closer to the homogeneous Poisson field, indicating that the material is highly resistent to the dislocation motion, and consequently it is of sufficient high quality to be further used for practical industrial purposes.

8 Conclusions and further research

We have presented a methodology based on stochastic point processes obtained as the limit of lattice-based point processes and an estimation procedure to obtain the parameters of pair-potentials of Gibbs processes. Thus, we consider an important statistical problem. A simulation study has been conducted to analyze the behaviour of this methodology under different practical situations.

Using this local conditioning-based methodology, we can successfully approach a pipp model by means of a colection of lattice processes. Particularly, we have demostrated either theoretically and by simulation that this method can detect all the three possible pattern structures, even under aggregation, for which traditionally pipp models have been rejected. Moreover, this methodology can be largely used (in an efficient and easy way) in many applied fields and by non-expert researchers in the spatial biological context. It is worth pointing out here that there are lots of applications coming from different applied scientific fields that could benefit from applying this methodology. However, if the researcher is also looking for an adequate parameter value, there exists a balance between grid size and order scheme, both playing an important role. This balance can be detected by means of goodness-of-fit tests as we have used in the applications or we can borrow some help from other likelihood-based methods.

In this paper we have not considered marked point processes. These kind of processes can be used to work with, for example, locations of cells together with biological measures defined upon each cell. The author is actually working on this kind of approximation for marked models. Also, it might be interesting to explore how this approximation behaves with other models which are not pipp, such as nearest-neighbour or area-interaction processes.

Finally, the SPPA [30] computer program has been a useful tool to work with this kind of spatial statistical techniques, since it allows us to reach the appropriate theoretical limits in terms of grid sizes. This computational software is available from the author upon request.

Acknowledgement

The author wishes to thank the anonymous referees for their careful reading of the manuscript and their fruitful comments and suggestions.

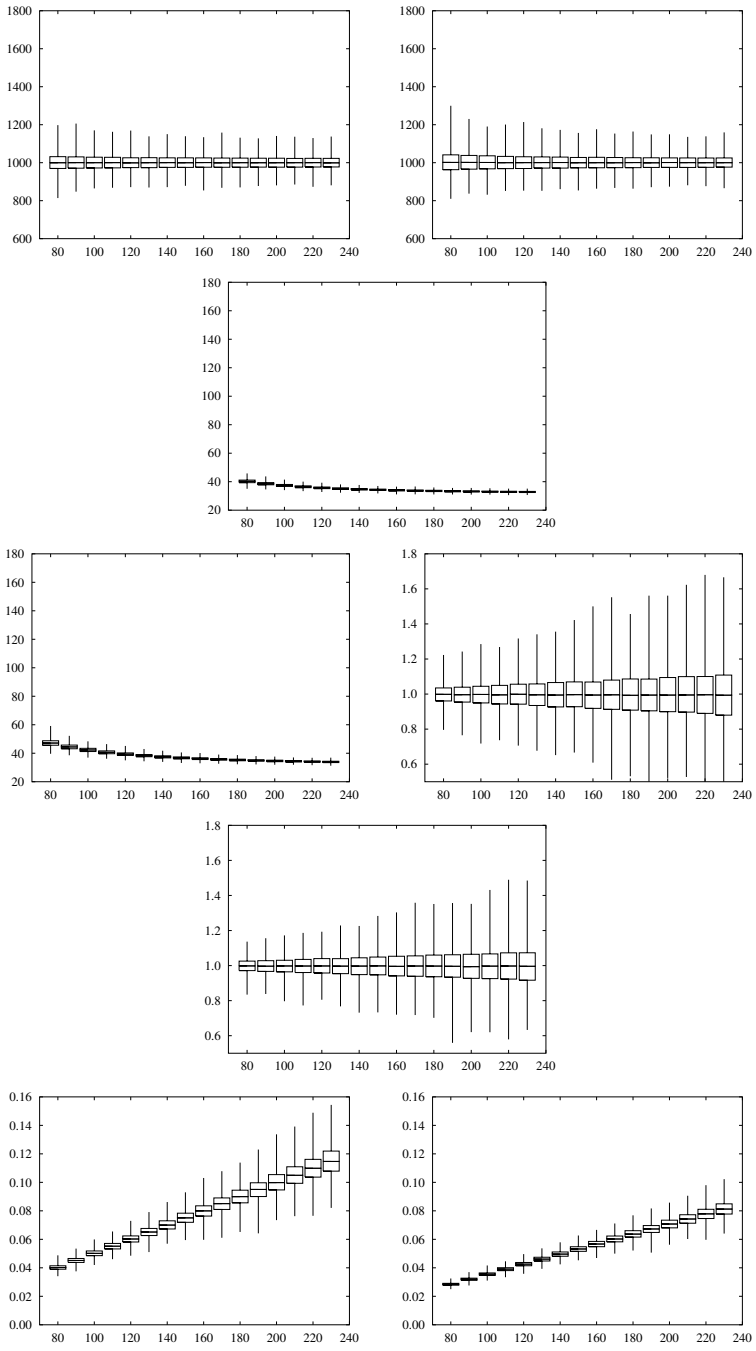

Figure 6: Boxplots of means and standard deviations of parameters λ and ρ for random spatial point patterns with an average of 1000 points in the unit square. Comparison for first to second-order schemes for grid sizes from 80×80 to 230×230. From left to right and up and down: $\lambda_1, \lambda_2, s_{\lambda_1}, s_{\lambda_2}, \rho_1, \rho_2, s_{\rho_1}, s_{\rho_2}$

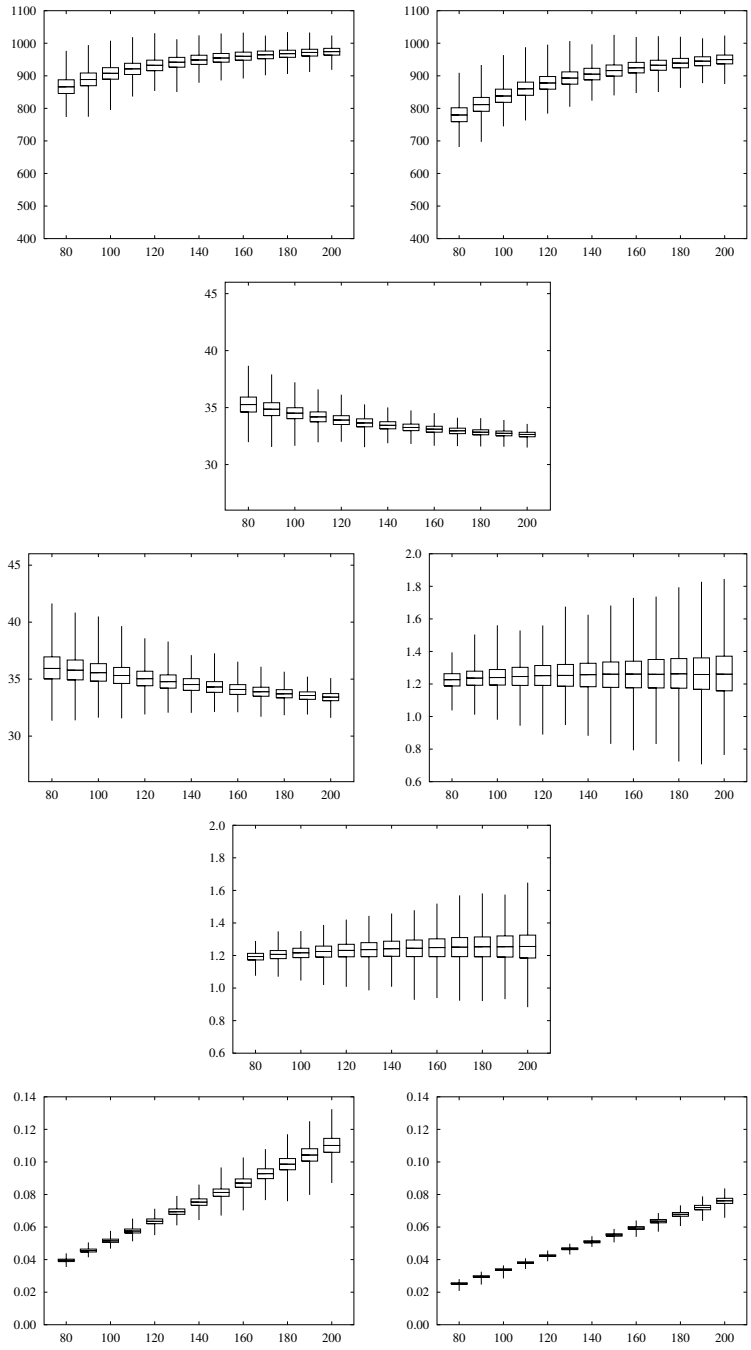

Figure 7: As in Figure 6 but for clustered spatial point patterns with an average of 1000 points in the unit square, and grid sizes from 80×80 to 200×200. From left to right and up and down: $\lambda_1, \lambda_2, s_{\lambda_1}, s_{\lambda_2}, \rho_1, \rho_2, s_{\rho_1}, s_{\rho_2}$

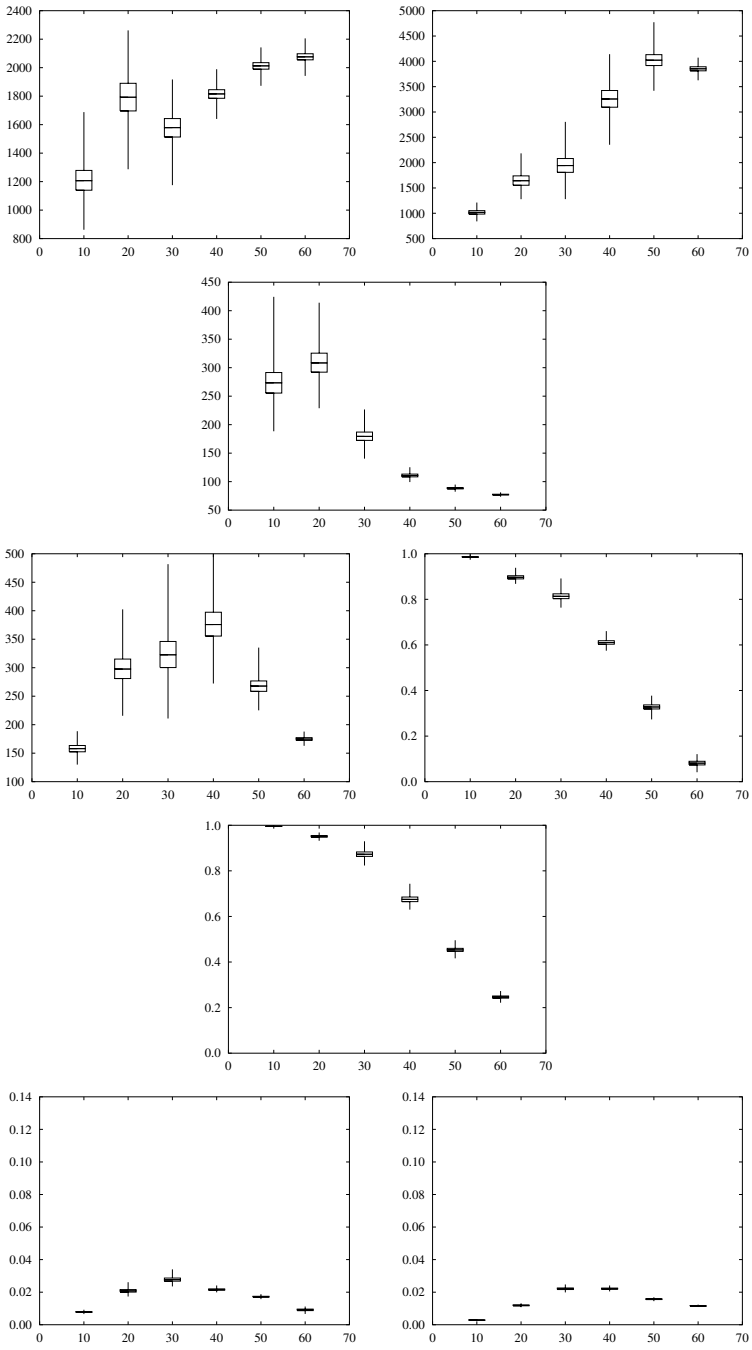

Figure 8: As in Figure 6 but for regular spatial point patterns with an average of 800 points in the unit square, and grid sizes from 10×10 to 60×60. From left to right and up and down: $\lambda_1, \lambda_2, s_{\lambda_1}, s_{\lambda_2}, \rho_1, \rho_2, s_{\rho_1}, s_{\rho_2}$

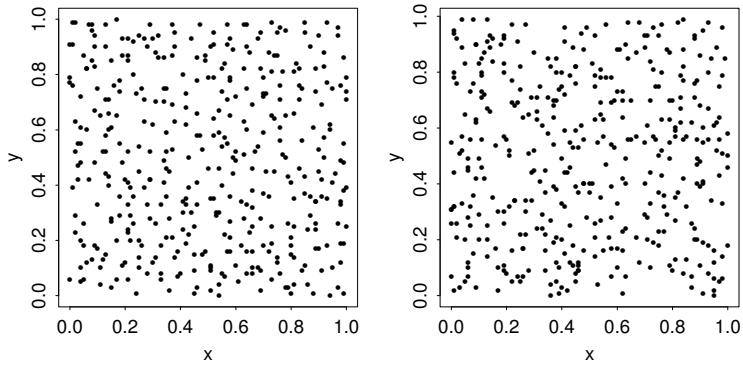

Figure 9: Data sets. *Left*: Locations of cells nuclei scaled to $(0,1) \times (0,1)$ in a microscopic section of tissue; *Right*: Locations of arrays of obstacles for the motion of dislocations in crystals scaled to $(0,1) \times (0,1)$

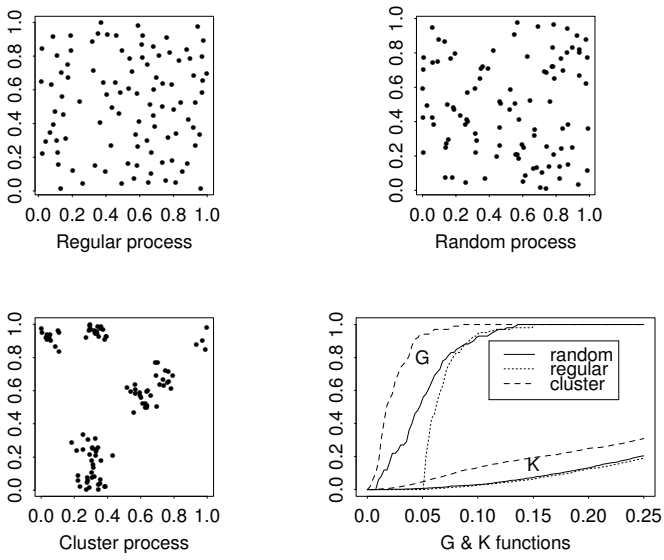

Figure 10: Simulated point models from three qualitatively different spatial structures and their corresponding G and K-functions

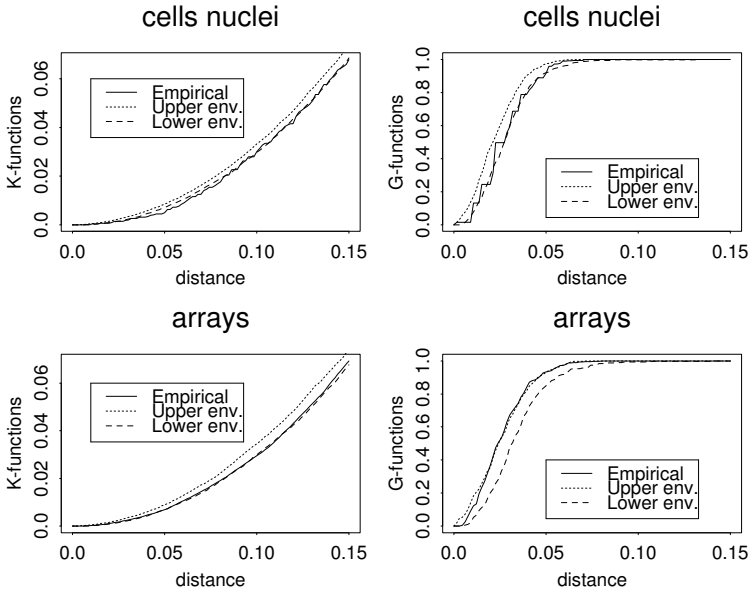

Figure 11: K and G-functions for the empirical patterns of cells nuclei and arrays of obstacles for the motion of dislocations in crystals. None of them corresponds to a random spatial pattern

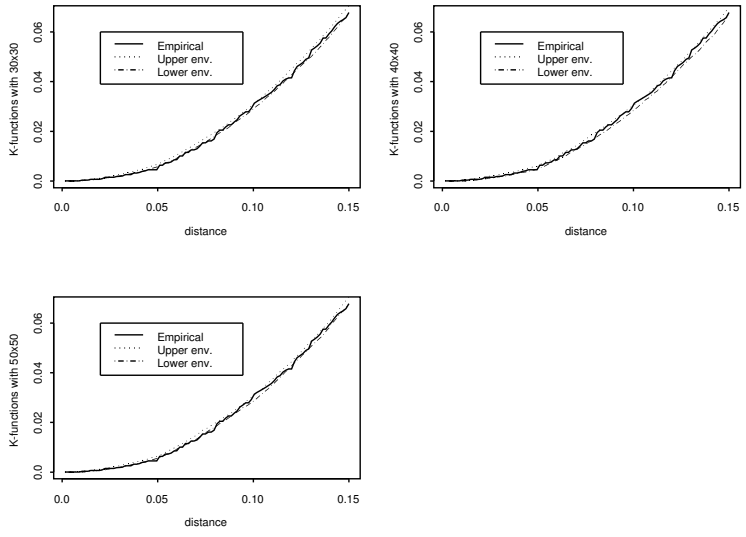

Figure 12: Empirical and simulated envelopes of the second-order K-function for the pattern of cells nuclei calculated using a Strauss process with those parameters obtained with three different grid sizes

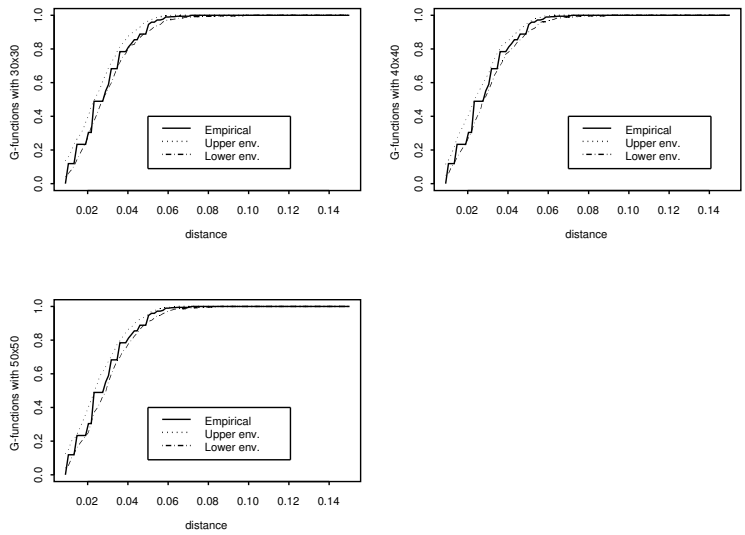

Figure 13: Empirical and simulated envelopes of the nearest-neighbour distribution function for the pattern of cells nuclei calculated using a Strauss process with those parameters obtained with three different grid sizes

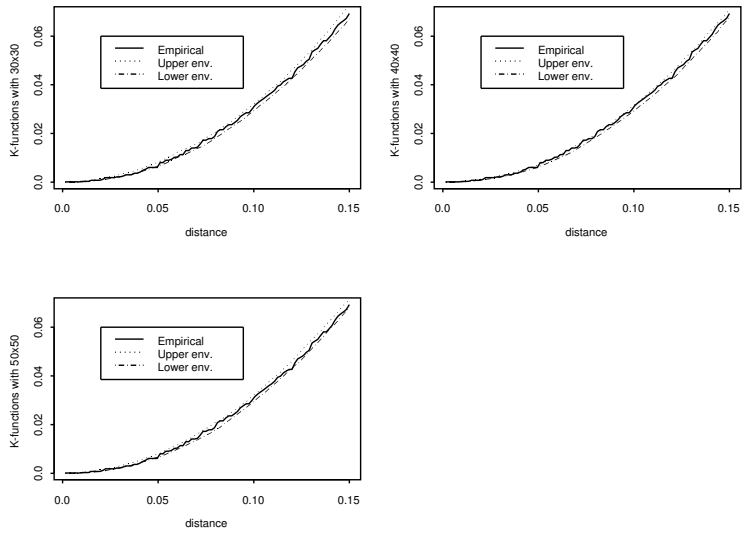

Figure 14: Empirical and simulated envelopes of the second-order K-function for the pattern of of arrays of obstacles for the motion of dislocations in crystals calculated using a Strauss process with those parameters obtained with three different grid sizes

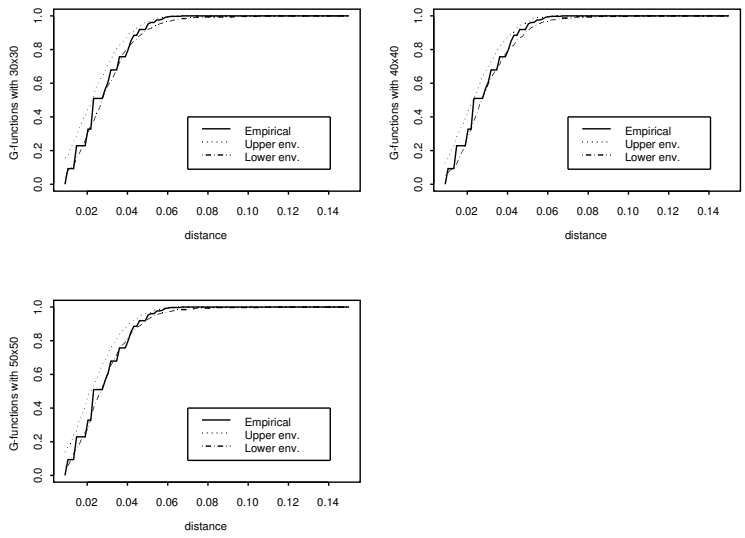

Figure 15: Empirical and simulated envelopes of the nearest-neighbour distribution function for the pattern of arrays of obstacles for the motion of dislocations in crystals calculated using a Strauss process with those parameters obtained with three different grid sizes

References

[1] T.E. Simos, Atomic Structure Computations in Chemical Modelling: Applications and Theory (Editor: A. Hinchliffe, UMIST), *The Royal Society of Chemistry* 38-142(2000).

[2] A.J. Baddeley and M.N.M. Lieshout, Area-interaction point processes, *Annals of Institute of Statistical Mathematics* **46** 601-619(1995).

[3] J. Besag, Some methods of statistical analysis for spatial data, *Bulletin of the International Statistical Institute* **47** 77-92(1977).

[4] J. Besag, R. Milne and S. Zachary, Point process limits of lattice processes, *Journal of Applied Probability* **19** 210-216(1982).

[5] D.S. Carter and P.M. Prenter, Exponential spaces and counting processes, *Z. Wahrscheinlichkeitsth* **21** 1-19(1972).

[6] N. Cressie, *Statistics for Spatial Data*, John Wiley and Sons, New York, 1993.

[7] D.J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer, New York, 1988.

[8] P.J. Diggle, *Statistical Analysis of Spatial Point Patterns*, Academic Press, London, 1983.

[9] P.J. Diggle, D.J. Gates and A. Stibbard, A nonparametric estimator for pairwise interaction point processes, *Biometrika* **74** 763-770(1987).

[10] P.J. Diggle, T. Fiksel, P. Grabarnik, Y. Ogata, D. Stoyan and M. Tanemura, On parameter estimation for pairwise interaction point processes, *International Statistical Review* **62** 99-117(1994).

[11] T. Fiksel, Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes, *Elektron. Inform. Kybernet.* **20** 270-278(1984).

[12] T. Fiksel, Estimation of interaction potentials of gibbsian point processes, *Math. Operationsf. Statist. Ser. Statist.* **19** 77-86(1988).

[13] C. Geyer, Likelihood inference for spatial point processes, In: *Stochastic Geometry: Likelihood and Computation* (Editors: O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout), Chapman and Hall/CRC London, 79-140(1999).

[14] J.L. Jensen and J. Møller, Pseudo-likelihood for exponential family models of spatial point processes, *Annals of Applied Probability* **1** 445-461(1991).

[15] F.P. Kelly and B.D. Ripley, On Strauss model for clustering, *Biometrika* **63** 357-360(1976).

[16] M.N.M. Lieshout and I.S. Molchanov, Shot-noise-weighted processes: a new family of spatial point processes, *Stochastic Models* **14** 35-48(1997).

[17] J. Møller, Markov Chain Monte Carlo and spatial point processes. In: *Stochastic Geometry: Likelihood and Computation* (Editors: O.E. Barndorff-Nielsen, W.S. Kendall and M.N.M. van Lieshout), Chapman and Hall/CRC London, 141-172(1999).

[18] J. Møller and K. Schladitz, Extensions of Fill's algorithm for perfect simulation, *Journal of the Royal Statistical Society* B **61** 955-969(1999).

[19] R.A. Moyeed and A.J. Baddeley, Stochastic approximation of the MLE for a spatial point pattern, *Scandinavian Journal of Statistics*, **18** 39-50(1991).

[20] Y. Ogata and M. Tanemura, Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure, *Annals of the Institute of Statistical Mathematics* **33** 315-338(1981).

[21] Y. Ogata and M. Tanemura, Likelihood analysis of spatial point patterns, *Journal of the Royal Statistical Society* B **46** 496-518(1984).

[22] Y. Ogata and M. Tanemura, Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns, *Annals of the Institute of Statistical Mathematics* **41** 583-600(1989).

[23] F. Papangelou, The conditional intensity of general point processes and an application to line processes, *Z. Wahrscheinlichkeitsth* **28** 207-26(1974).

[24] A. Penttinen, Modelling interaction in spatial point patterns: parameter estimation by the maximum likelihood method, *Jyvaskyla Studies in Computer Science, Economics and Statistics* **7**, 1984.

[25] C.J. Preston, *Random fields.* Lecture Notes in Mathematics **534**, Springer-Verlag, Berlin, 1976.

[26] B.D. Ripley, Modelling spatial patterns (with discussion), *Journal of the Royal Statistical Society* B **39** 172-212(1977).

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

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

[29] D. Ruelle, *Statistical Mechanics*, Wiley, New York, 1969.

[30] SPPA, *Spatial Point Pattern Analysis*, Computer Software developed by Albert, Albert, Mateu, Pernías, Universitat Jaume I, Castellón, 1997.

[31] D. Stoyan, W. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, Akademie-Verlag, Berlin, 1995.

[32] D.J. Strauss, A model for clustering, *Biometrika* **63** 467-475(1975).

[33] R. Takacs, Estimator for the pair-potential of a Gibbsian point process, *Mathematische Operationsforschung und Statistik, series Statistik* **17** 429-433(1986).

[34] E. Tomppo, Models and methods for analysing spatial patterns of trees, *Communicationes Instituti Forestalis Fenniae* **138** 1-65(1986).