KYBERNETIKA — VOLUME 47 (2011), NUMBER 6, PAGES 880–892

ESTIMATION OF SUMMARY CHARACTERISTICS FROM REPLICATED SPATIAL POINT PROCESSES

Zbyněk Pawlas

Summary characteristics play an important role in the analysis of spatial point processes. We discuss various approaches to estimating summary characteristics from replicated observations of a stationary point process. The estimators are compared with respect to their integrated squared error. Simulations for three basic types of point processes help to indicate the best way of pooling the subwindow estimators. The most appropriate way depends on the particular summary characteristic, edge-correction method and also on the type of point process. The methods are demonstrated on a replicated dataset from forestry.

Keywords: K-function, nearest-neighbour distance distribution function, non-parametric estimation, point process, replication

Classification: 60G55, 62G05, 62M30

1. INTRODUCTION

Data in the form of spatial point patterns occur in a wide range of disciplines, e. g. forestry, biology, medicine, seismology, epidemiology or materials science. Spatial point processes are used to model these kinds of data. The analysis of spatial point processes is one of basic objects in spatial statistics. For this analysis, summary characteristics have a fundamental role. Several types of summary characteristics for point processes have been defined and frequently used, for an overview see e. g. [9]. Much attention in the literature has been devoted to the estimation of summary characteristics from a single point pattern observed within a fixed bounded window. However, it is becoming more common in applications that replicated samples of a spatial point pattern are available (e. g. [2, 6]). The data may be often considered as independent replications of the same process. We may think of observation of several i.i.d. point processes (e. g. tissue sections in organs of different patients) or observation of a single process through several small windows that are sufficiently distant apart from each other (e. g. tissue sections in some organ at large distances). There are a number of papers dealing with the analysis of replicated spatial point patterns, e. g. [2, 4, 6, 7, 13]. In the present paper we focus mainly on the estimation of summary characteristics.

One of the most popular summary characteristics is the K-function. It is an example of a second-order characteristic. An important alternative to second-order

characteristics is given by the nearest-neighbour distance distribution function. Our main interest is the non-parametric estimation of both these characteristics. We will study an estimator of the K-function and three different estimators of the nearest-neighbour distance distribution function.

We assume that n independent replications $\Phi_1, \ldots, \Phi_n$ of a stationary point process Φ are observed in windows $W_1, \ldots, W_n$, respectively. Our aim is to estimate summary characteristics of Φ. The most natural approach is to aggregate the separate estimators for each of the windows W_i. Some of the refined methods are reviewed in [9], see also [1, 5]. We present a comparison of different approaches to the estimation of summary characteristics for the data in the form of replicated spatial point patterns. Since the processes Φ_i are assumed to be i.i.d. it is not difficult to obtain some results concerning asymptotic behaviour as $n \to \infty$, see [1]. However, for small n it is very complicated to derive any theoretical results and we have to investigate the properties of the estimators by a computer simulation. We study how the performance of individual methods depends on the type of point process, intensity and observation windows. We conclude with the application to real forestry dataset consisting of the positions of trees in several sampling regions.

2. SPATIAL POINT PROCESSES

Point processes are defined as locally finite random subsets of the d-dimensional Euclidean space $\mathbb{R}^d$. Denote the σ-algebra of Borel sets of $\mathbb{R}^d$ by $\mathcal{B}^d$ and the system of bounded Borel sets by $\mathcal{B}_0^d$. Let $\mathcal{N} = \{\varphi \subseteq \mathbb{R}^d : \varphi(B) < \infty \; \forall B \in \mathcal{B}_0^d\}$ be the family of locally finite subsets of $\mathbb{R}^d$. The symbol $\varphi(B)$ stands for the number of points of $\varphi \cap B$. We equip $\mathcal{N}$ with the σ-algebra $\mathfrak{N}$ generated by the sets $\{\varphi \in \mathcal{N} : \varphi(B) = m\}$, $m = 0, 1, 2, \ldots$, $B \in \mathcal{B}_0^d$. A spatial point process Φ is a measurable mapping from an abstract probability space $(\Omega, \mathcal{A}, \mathbb{P})$ into the measurable space $(\mathcal{N}, \mathfrak{N})$. We say that Φ is stationary if for all $y \in \mathbb{R}^d$ the translated point process $\Phi + y = \{X + y : X \in \Phi\}$ has the same distribution as Φ. Then the mean number of points falling in a set $B \in \mathcal{B}^d$ is proportional to the Lebesgue measure of B, i. e. $\mathbb{E}\Phi(B) = \lambda|B|$. The constant λ is called the intensity of the point process Φ. Given a stationary point process Φ with intensity λ, its reduced Palm distribution can be defined by

$$P_o^!(U) = \frac{1}{\lambda|A|}\mathbb{E}\sum_{X \in \Phi \cap A} \mathbf{1}_{[(\Phi \setminus \{X\}) - X \in U]}, \quad U \in \mathfrak{N},$$

where $A \in \mathcal{B}^d$ is an arbitrary set with $0 < |A| < \infty$. For more details on point processes we refer to [9] and references therein.

It is useful to describe point processes by simpler objects, such as numbers or functions. For this purpose, summary characteristics are frequently used in spatial statistics. We will introduce several well-established summary characteristics for stationary point processes. The most important numerical summary characteristic is the intensity λ. Modern point process statistics often works with functional summary characteristics. *Spherical contact distribution function* (also known as *empty space function*) is defined as

$$F(r) = \mathbb{P}(\Phi(b(o, r)) > 0), \quad r > 0,$$

where $b(x,r)$ denotes the ball of radius r with centre $x \in \mathbb{R}^d$. It can be interpreted as the distribution function of the random distance from the origin to the nearest point in Φ: $F(r) = \mathbb{P}(D \leq r)$, where $D = \inf_{X\in\Phi} \|X\|$. On the other hand, the *nearest-neighbour distance distribution function*

$$G(r) = P_o^!(\{\varphi \in \mathcal{N} : \varphi(b(o,r)) > 0\}), \quad r > 0,$$

is the distribution function of the random distance from the typical point of Φ to its nearest neighbour in Φ. Both F and G functions use only the nearest neighbour distances. Different type of characteristics are second-order characteristics. Probably the most popular second-order characteristic is the *K-function* which is defined by

$$K(r) = \frac{1}{\lambda}\mathbb{E}_o^!\Phi(b(o,r)), \quad r > 0,$$

where $\mathbb{E}_o^!$ is the expectation with respect to the reduced Palm distribution $P_o^!$. It means that $\lambda K(r)$ can be interpreted as the mean number of points in a ball of radius r centred at a typical point of the process which is not counted. Alternatively, we can write

$$K(r) = \frac{1}{\lambda^2|A|}\mathbb{E}\sum_{X,Y\in\Phi} \mathbf{1}_{[X\in A, \|X-Y\|\leq r]},$$

where A is an arbitrary set with positive and finite Lebesgue measure $|A|$.

The data are often formed by a single point pattern observed within a fixed bounded window $W \in \mathcal{B}_0^d$. However, natural estimators of the summary characteristics would need information from outside W. For example, a natural estimator of $G(r)$ would be

$$\hat{G}(r) = \frac{1}{\Phi(W)} \sum_{X\in\Phi\cap W} \mathbf{1}_{[e(X)\leq r]}, \tag{1}$$

where $e(X) = d(X, \Phi \setminus \{X\})$ is the distance from X to its nearest neighbour. This estimator is only applicable if we have additional information which enables us to determine $e(X)$. Mostly we have to deal with edge effects problems because only information on the points in W is available. For points X close to the boundary we can only observe $e_W(X) = d(X, (\Phi \setminus \{X\}) \cap W)$, the distance to the nearest neighbour in W, while the true nearest neighbour may lie outside W. Replacing $e(X)$ by $e_W(X) \geq e(X)$ in (1) introduces negative bias in the estimation. Therefore, different edge-correction methods have been developed, see [9] for a comprehensive review.

We mention three approaches for the estimation of $G(r)$. In the first approach, for a given distance r we consider only those points that have a distance larger than r from the boundary ∂W of the window W (so called minus sampling). This leads to the *border method estimator* (or *reduced sample estimator*), which has the form

$$\hat{G}_{\mathrm{b}}(r) = \frac{1}{\Phi(W_{\ominus r})} \sum_{X\in\Phi\cap W_{\ominus r}} \mathbf{1}_{[e(X)\leq r]}, \tag{2}$$

where $W_{\ominus r} = \{x \in W : d(x, \partial W) \geq r\}$. The second approach was suggested by Hanisch in [8], it uses only points that are closer to its nearest neighbour than to the boundary of the window. The *Hanisch estimator* is given by

$$\hat{G}_{\mathrm{H}}(r) = \frac{1}{\hat{\lambda}_{\mathrm{H}}} \sum_{X \in \Phi \cap W} \frac{\mathbf{1}_{[e(X) \leq d(X, \partial W)]}}{|W_{\ominus e(X)}|} \mathbf{1}_{[e(X) \leq r]}, \tag{3}$$

where

$$\hat{\lambda}_{\mathrm{H}} = \sum_{X \in \Phi \cap W} \frac{\mathbf{1}_{[e(X) \leq d(X, \partial W)]}}{|W_{\ominus e(X)}|}.$$

Finally, the third approach is the *Kaplan–Meier method*, motivated by the analogy with survival analysis (see [1]),

$$\hat{G}_{\mathrm{KM}}(r) = 1 - \prod_{s \leq r} \left(1 - \frac{\sum_{X \in \Phi \cap W} \mathbf{1}_{[e(X)=s, e(X) \leq d(X, \partial W)]}}{\sum_{X \in \Phi \cap W} \mathbf{1}_{[e(X) \geq s, d(X, \partial W) \geq s]}}\right). \tag{4}$$

The detailed comparison of the estimators (2), (3) and (4) can be found in [12].

For the estimation of $K(r)$ translational edge-correction is often used. It can be shown that

$$\widehat{\lambda^2 K(r)} = \sum_{X, Y \in \Phi \cap W} \frac{\mathbf{1}_{[\|X - Y\| \leq r]}}{|W \cap (W + (X - Y))|} \tag{5}$$

is an unbiased estimator of $\lambda^2 K(r)$, see e. g. [9], p. 228. In order to get an estimator of $K(r)$ we divide by some estimator of λ^2. For example, we can use $\widehat{\lambda^2} = \Phi(W)(\Phi(W) - 1)/|W|^2$, which is unbiased in the case of Poisson process,

$$\hat{K}(r) = \frac{1}{\widehat{\lambda^2}} \sum_{X, Y \in \Phi \cap W} \frac{\mathbf{1}_{[\|X - Y\| \leq r]}}{|W \cap (W + (X - Y))|}. \tag{6}$$

3. REPLICATED PATTERNS

Frequently, the data can be regarded as replicate observations of the same point process through different windows (e. g. images taken at several different locations in the material or tissue). For analysis of replicated spatial data it is important to estimate summary characteristics of the underlying process.

3.1. Aggregation methods

We assume that n independent copies $\Phi_1, \ldots, \Phi_n$ of a stationary point process Φ are observed through observation windows $W_1, \ldots, W_n$. Our aim is to construct non-parametric estimators of a specific summary characteristic of our interest, say $S(r)$. There are two natural strategies how to estimate $S(r)$.

The first approach is to take separate estimators $\hat{S}_i(r)$ for each of the windows W_i and define a *weighted average*

$$\hat{S}_{\mathrm{w}}(r) = \sum_{i=1}^{n} \frac{C_i}{C} \hat{S}_i(r), \tag{7}$$

where $C_i \geq 0$ are the weights and $C = \sum_{i=1}^n C_i$. The C_i may be random and we also admit that they depend on r.

The second strategy is pooling information from all replicated observations. Often $\hat{S}_i(r)$ are ratio-unbiased estimators, i. e.

$$\hat{S}_i(r) = \frac{U_i(r)}{V_i}, \quad \text{where} \quad \frac{\mathbb{E}U_i(r)}{\mathbb{E}V_i} = S(r).$$

For example, (2) and (3) are ratio-unbiased estimators of $G(r)$. In [2] the authors use ratio regression approach to argue that preferable aggregated estimator is then the ratio of the sum of numerators to the sum of denominators,

$$\hat{S}_{\mathrm{p}}(r) = \frac{\sum_{i=1}^n U_i(r)}{\sum_{i=1}^n V_i}. \tag{8}$$

The justification of this intuitive approach can also be found in [9], p. 262. We will refer to (8) as the *pooled estimator*.

The pooled estimator (8) coincides with the weighted average (7) when we take $C_i = V_i$. In particular, $V_i = \Phi_i((W_i)_{\ominus r})$ for the case of the border estimator (2) and $V_i = (\hat{\lambda}_{\mathrm{H}})_i$ for the Hanisch estimator (3):

$$\hat{G}_{\mathrm{p,b}}(r) = \frac{\sum_{i=1}^n \sum_{X\in\Phi_i\cap(W_i)_{\ominus r}} \mathbf{1}_{[e_i(X)\leq r]}}{\sum_{i=1}^n \Phi_i((W_i)_{\ominus r})}, \tag{9}$$

$$\hat{G}_{\mathrm{p,H}}(r) = \frac{\sum_{i=1}^n \sum_{X\in\Phi_i\cap W_i} \mathbf{1}_{[e_i(X)\leq d(X,\partial W_i), e_i(X)\leq r]}/|(W_i)_{\ominus e_i(X)}|}{\sum_{i=1}^n \sum_{X\in\Phi_i\cap W_i} \mathbf{1}_{[e_i(X)\leq d(X,\partial W_i)]}/|(W_i)_{\ominus e_i(X)}|}, \tag{10}$$

where $e_i(x) = d(x, \Phi_i\backslash\{x\})$. The Kaplan–Meier estimator (4) is not a ratio-unbiased estimator. However, it is natural to define the pooled estimator by analogue of (4) where both numerator and denominator are replaced by the corresponding sums over all replicates (see [1]),

$$\hat{G}_{\mathrm{p,KM}}(r) = 1 - \prod_{s\leq r}\left(1 - \frac{\sum_{i=1}^n \sum_{X\in\Phi_i\cap W_i} \mathbf{1}_{[e_i(X)=s, e_i(X)\leq\delta_i(X)]}}{\sum_{i=1}^n \sum_{X\in\Phi_i\cap W_i} \mathbf{1}_{[e_i(X)\geq s, \delta_i(X)\geq s]}}\right), \tag{11}$$

where $\delta_i(x) = d(x, \partial W_i)$.

For the estimation of $K(r)$ by the translational edge-correction, see (6), the pooled estimator (8) becomes

$$\hat{K}_{\mathrm{p}}(r) = \frac{\sum_{i=1}^n \sum_{X,Y\in\Phi_i\cap W_i} \mathbf{1}_{[\|X-Y\|\leq r]}/|W_i\cap(W_i+(X-Y))|}{\sum_{i=1}^n \Phi_i(W_i)(\Phi_i(W_i)-1)/|W_i|^2} \tag{12}$$

and it coincides with the weighted average estimator (7) if we take $C_i = \left(\widehat{\lambda^2}\right)_i = \Phi_i(W_i)(\Phi_i(W_i)-1)/|W_i|^2$. Another weighting is recommended in [5], p. 123, namely $C_i = \Phi_i(W_i)$. An alternative approach is to estimate $\lambda^2 K(r)$ by the mean of (5),

$$\widehat{\lambda^2 K(r)} = \frac{1}{n}\sum_{i=1}^n \sum_{X,Y\in\Phi_i\cap W_i} \frac{\mathbf{1}_{[\|X-Y\|\leq r]}}{|W_i\cap(W_i+(X-Y))|},$$

and divide by some estimator $\widehat{\lambda^2}$ of λ^2. This approach was used for example in [10]. It is easy to see that the resulting *alternative pooled estimator* has the form of weighted average,

$$\hat{K}_{\mathrm{a}}(r) = \sum_{i=1}^{n} \frac{\left(\widehat{\lambda^2}\right)_i}{n\widehat{\lambda^2}} \hat{K}_i(r). \tag{13}$$

However, the weights do not necessarily sum to one. In particular, we will consider $\left(\widehat{\lambda^2}\right)_i = \Phi_i(W_i)(\Phi_i(W_i) - 1)/|W_i|^2$ and

$$\widehat{\lambda^2} = \frac{\sum_{i=1}^{n} \Phi_i(W_i) \left(\sum_{i=1}^{n} \Phi_i(W_i) - 1\right)}{\left(\sum_{i=1}^{n} |W_i|\right)^2}.$$

3.2. Comparison of estimators

In [12] series of simulations were carried out for the border estimator of $G(r)$. Recall that in this case the pooled estimator $\hat{G}_{\mathrm{p,b}}(r)$ is equal to $\hat{G}_{\mathrm{w}}(r)$ with $C_i = \Phi_i((W_i)_{\ominus r})$. Other weights which were considered in [12] are $C_i = 1$, $C_i = |W_i|$, $C_i = \Phi_i(W_i)$ and $C_i = \Phi_i(W_i)^2$. Among these weights the weighting by the squared point numbers ($C_i = \Phi_i(W_i)^2$) was recommended. The comparison was based with respect to bias and mean squared error of the estimators.

Our aim is to compare different aggregation methods for the estimation of a summary characteristic $S(r)$. We measure the quality of the estimators by the mean weighted integrated squared error

$$\mathrm{MWISE}(a, w) = \mathbb{E} \int_0^a (\hat{S}(r) - S(r))^2 w(r) \, \mathrm{d}r,$$

where $w(r)$ is an appropriate weighting function and $a \geq 0$ is a chosen constant. In particular, for $S(r) = G(r)$ we choose $w(r) = 1$, while for $S(r) = K(r)$ we take $w(r) = 1/r^2$ in order to eliminate greater variability of estimators of $K(r)$ for large distances r. The same choice is usual when fitting a point process model by minimum contrast method, see [5].

We carry out a simulation study for three stationary planar point process models representing basic types of point processes: Poisson point process (complete spatial randomness), Thomas process (clustering) and Matérn hard-core process II (regularity), for definitions see [5] or [9]. Poisson process is determined by intensity λ, Thomas process has three parameters ($\bar{c}$ – mean number of points per cluster, λ_p – intensity of the Poisson process of parent points and σ – standard deviation of the displacement of a daughter point from its parent point), Matérn hard-core process II has two parameters (λ_b – intensity of the Poisson process of proposal points and r_0 – hard-core distance). The intensity of Thomas process is $\lambda = \lambda_p \bar{c}$ and the intensity of Matérn hard-core process II is $\lambda = (1 - \mathrm{e}^{-\lambda_b \pi r_0^2})/\pi r_0^2$.

For a given planar point process model we generate realizations in n windows, then we estimate summary characteristics and afterwards we determine the weighted integrated squared error. The simulations are independently repeated 10 000 times

and MWISE is approximated by the arithmetic mean over these 10 000 simulations. The analytical expression of $G(r)$ is only known for Poisson process ($G(r) = 1 - \mathrm{e}^{-\pi r^2}$), the formula for $K(r)$ is known in the case of Poisson process ($K(r) = \pi r^2$) and Thomas process ($K(r) = \pi r^2 + (1 - \mathrm{e}^{-r^2/4\sigma^2})/\lambda_p$). In other cases we have to approximate theoretical functions from Monte Carlo simulations. The statistical computing and simulation was conducted using R (see [11]) and its contributed package spatstat (see [3]).

We evaluate the weighted averages given by (7) with the following four choices of the weights C_i: identical ($C_i = 1$), window sizes ($C_i = |W_i|$), point numbers ($C_i = \Phi(W_i)$) and squared point numbers ($C_i = \Phi(W_i)^2$). These aggregated estimators are computed for each edge-correction method (2), (3), (4) and (6). Moreover, we determine the pooled estimators (9), (10), (11) and (12) and for the case of estimation of $K(r)$ we also include the alternative pooled estimator (13) into the comparison.

First we consider congruent windows. Let $n = 10$ and W_i be unit squares in $\mathbb{R}^2$. We put $a = 0.25$. For Poisson process two different intensities λ are considered (5 and 15). As an example of cluster process we use the Thomas process with parameters $\bar{c} = 5$, $\sigma = 0.2$ and as an example of regular process we consider the Matérn hard-core process II with hard-core distance $r_0 = 0.1$. In both these cases the intensity λ is 15. The resulting average errors obtained for the border estimator of $G(r)$ are summarized in Table 1, for the Hanisch estimator of $G(r)$ in Table 2, for the Kaplan–Meier estimator of $G(r)$ in Table 3 and for the translation-corrected estimator of $K(r)$ in Table 4. The smallest errors for each process are displayed in bold font. Since the windows are congruent, we have always only three different weighted averages. Furthermore, we consider the pooled estimators and in the case of estimation of $K(r)$ also the alternative pooled estimator (13).

estimator	Poisson(5)	Poisson(15)	cluster(15)	regular(15)
$C_i = 1 = \lvert W_i \rvert$	4.409	1.501	4.462	0.874
$C_i = \Phi_i(W_i)$	3.302	1.066	1.511	0.702
$C_i = \Phi_i(W_i)^2$	3.632	0.955	1.249	0.648
pooled $\hat{G}_{\mathrm{p,b}}$	**3.119**	**0.794**	**1.000**	**0.569**

Tab. 1. The values $1000 \cdot \mathrm{MWISE}(0.25, 1)$ obtained from 10 000 simulations in 10 unit square windows W_i. Different aggregated border method estimators of $G(r)$ are compared for four point processes.

The results for the border estimators (shown in Table 1) are in accordance with the conclusions in [12] where the same weights were considered and squared point numbers were found to be the best choice. However, the pooling approach (9), which was not involved in the simulation study of [12], outperforms all considered weighted averages. For the Kaplan–Meier estimator, the pooling approach (11) is the most acceptable because it has the smallest MWISE in all studied cases. When using the Hanisch estimator we recommend weighting by number of points. The Hanisch estimator uses only the points for which we observe the true nearest

estimator	Poisson(5)	Poisson(15)	cluster(15)	regular(15)
$C_i = 1 = \lvert W_i \rvert$	**3.217**	0.989	2.453	0.694
$C_i = \Phi_i(W_i)$	3.462	**0.820**	**0.934**	0.607
$C_i = \Phi_i(W_i)^2$	4.924	0.879	1.162	0.614
pooled $\hat{G}_{\mathrm{p,H}}$	4.592	0.835	1.034	**0.603**

Tab. 2. The values $1000 \cdot \mathrm{MWISE}(0.25, 1)$ obtained from 10 000 simulations in 10 unit square windows W_i. Different aggregated Hanisch estimators of $G(r)$ are compared for four point processes.

estimator	Poisson(5)	Poisson(15)	cluster(15)	regular(15)
$C_i = 1 = \lvert W_i \rvert$	3.580	1.064	3.274	0.693
$C_i = \Phi_i(W_i)$	2.445	0.750	0.997	0.559
$C_i = \Phi_i(W_i)^2$	2.680	0.703	0.980	0.527
pooled $\hat{G}_{\mathrm{p,KM}}$	**1.826**	**0.631**	**0.807**	**0.494**

Tab. 3. The values $1000 \cdot \mathrm{MWISE}(0.25, 1)$ obtained from 10 000 simulations in 10 unit square windows W_i. Different aggregated Kaplan–Meier estimators of $G(r)$ are compared for four point processes.

neighbour distance. It may happen that only a few points have this property. This is transparent in the case $\lambda = 5$, where individual Hanisch estimators are not very precise and eventually the best pooling turns out to be simple arithmetic mean. In our Monte Carlo simulations the mean number of points in a single window is rather small, this could explain lower quality of Hanisch estimators in comparison with Kaplan–Meier estimators. In fact, the quality of Hanisch estimators is comparable to the quality of the border method. Nevertheless, it is known that for larger intensities and larger r, the reduced sample estimator is less efficient. Of course, in all three cases the errors are smaller for larger intensity.

In the case of translation-corrected estimator (6) of the K-function, the pooled estimator (8) is the weighted average estimator with $C_i = \Phi_i(W_i)(\Phi_i(W_i) - 1)$. Hence, it is clear that this estimator yields almost the same errors as the weighted average with $C_i = \Phi_i(W_i)^2$. For Poisson and hard-core process these weightings are favourable. For cluster process averaging by point numbers behaves better, this approach is recommended in [5], p. 123 and also in [9], p. 263. However, the best results for cluster process are achieved by the alternative pooled estimator (13). This pooling method is intended to first estimate $\lambda^2 K(r)$ and then obtain the estimator of $K(r)$.

According to [9], p. 260, if the windows are congruent, then it is a good strategy to simply form arithmetic means of estimates from different windows. Our simulations reveal that this may not be the case if the point numbers are too varying. In the case of cluster process, the errors of arithmetic means are substantially higher than the errors of other aggregation methods.

estimator	Poisson(5)	Poisson(15)	cluster(15)	regular(15)
$C_i = 1 = \lvert W_i \rvert$	16.882	1.080	8.863	0.00969
$C_i = \Phi_i(W_i)$	9.118	0.865	4.618	0.00834
$C_i = \Phi_i(W_i)^2$	7.326	**0.814**	5.091	0.00789
pooled $\hat{K}_{\mathrm{p}}$	**7.296**	0.814	5.147	**0.00788**
alternative pooled $\hat{K}_{\mathrm{a}}$	7.660	0.850	**3.416**	0.00811

Tab. 4. The values $1000 \cdot \mathrm{MWISE}(0.25, 1/r^2)$ obtained from 10 000 simulations observed in 10 unit square windows W_i. Different aggregated translation-corrected estimators of $K(r)$ are compared for four point processes.

	Poisson		cluster		regular	
estimator	H	KM	H	KM	H	KM
$C_i = 1$	29.196	32.621	81.092	116.851	21.748	23.533
$C_i = \lvert W_i \rvert$	26.788	30.130	74.482	106.866	20.022	21.742
$C_i = \Phi_i(W_i)$	**23.823**	20.510	**24.498**	26.737	**17.727**	16.238
$C_i = \Phi_i(W_i)^2$	27.540	20.470	34.069	28.053	19.163	15.666
pooled $\hat{G}_{\mathrm{p}}$	24.618	**17.967**	28.074	**21.750**	18.804	**14.343**

Tab. 5. The average errors $1000 \cdot \mathrm{MWISE}(10, 1)$ obtained from 10 000 simulations of point processes observed in 15 windows W_i of different shape and size. Different aggregated Hanisch and Kaplan–Meier estimators of $G(r)$ are compared for three types of point processes with the same intensity.

The windows do not have to be congruent. Therefore, we consider different collection of sampling windows as well. We take $n = 15$ polygonal windows of different size and shape, they come from real dataset described in Section 4. The estimators are compared on three types of stationary point processes with intensity $0.0065\,\mathrm{m}^{-2}$: Poisson process, Thomas process with parameters $\bar{c} = 5$ and $\sigma = 8\,\mathrm{m}$ and Matérn hard-core process II with hard-core distance $r_0 = 4\,\mathrm{m}$. For each process 10 000 replicated samples are generated. Table 5 shows the results for Hanisch and Kaplan–Meier estimators of $G(r)$. Since the border method estimation has a bit poorer quality, we do not include it into the comparison this time. Four weighted averages are compared with the pooled estimators (10) and (11). The errors of the translation-corrected estimators of $K(r)$ are summarized in Table 6. Four weighted averages are compared with the pooled estimator (12) and the alternative pooled estimator (13). The results presented in Table 5 and Table 6 are in accordance with previous findings.

estimator	Poisson	cluster	regular
$C_i = 1$	72.110	716.71	1162.22
$C_i = \|W_i\|$	66.475	665.65	1073.26
$C_i = \Phi_i(W_i)$	51.510	361.04	889.53
$C_i = \Phi_i(W_i)^2$	**50.265**	447.69	**872.46**
pooled $\hat{K}_{\mathrm{p}}$	50.340	490.08	877.41
alternative pooled $\hat{K}_{\mathrm{a}}$	54.719	**276.98**	966.44

Tab. 6. The average errors MWISE$(10, 1/r^2)$ obtained from 10 000 simulations of point processes observed in 15 windows W_i of different shape and size. Different aggregated translation-corrected estimators of $K(r)$ are compared for three types of point processes with the same intensity.

4. FORESTRY DATA

We illustrate our study on the data provided by the Forest Management Institute, Brandýs nad Labem, Czech Republic. The data give the locations of trees in Vysočina region (Czech Republic) collected in 15 sampling windows that are sufficiently distant apart from each other to assume independence. Altogether 1 777 trees were located. The windows are polygons of different shape and size ranging from 1 679 to 4 031 square metres. They were chosen in spruce forests of approximately the same age (about 100 years). Therefore, we may regard the point patterns as samples of i.i.d. planar point processes. From the visual inspection of the data it seems that the stationarity assumption holds.

Basic summary characteristics were estimated separately from each window, see Figure 1 for Hanisch estimates of $G(r)$. Except windows 10 and 14 there are no apparent substantial differences between windows. All summary statistics suggest a regular point pattern.

We aggregate the information from all subwindows and calculate the pooled Kaplan–Meier estimator (11), see Figure 2. For comparison we plot also the theoretical function for the Poisson process with the same intensity as the overall intensity estimated from the data.

To facilitate visual inspection of estimates of $K(r)$ we prefer to plot the difference between the estimate and the theoretical value for the Poisson process, which is πr^2. Superimposed translation-corrected estimators $\hat{K}_i(r)$, $i = 1, \ldots, 15$, and the resulting pooled estimator are shown in Figure 3.

The pooled estimators deviate from theoretical functions under complete spatial randomness and reveal slight regularity at small distances (up to 3 metres). It is not surprising because trees can not grow too close together due to their physical size. Based on the pooled estimators of summary characteristics we can perform Monte-Carlo tests of the complete spatial randomness hypothesis. The L-test (see [9], p. 95) rejects this null hypothesis.

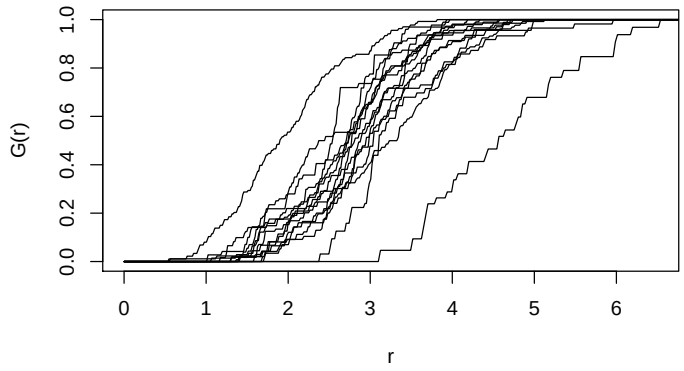

Fig. 1. Superimposed Hanisch estimates of $G(r)$ from all 15 windows.

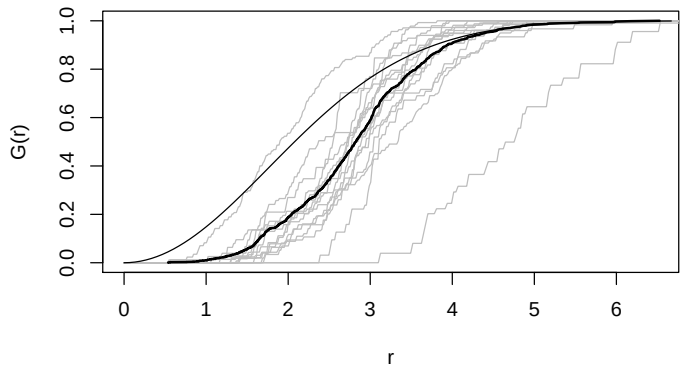

Fig. 2. Pooled Kaplan–Meier estimate of $G(r)$ is shown by thick black line, theoretical function for the Poisson process is thin black line and individual Kaplan–Meier estimates for separate subwindows are in grey.

5. CONCLUSIONS

When dealing with replicated point patterns there are several strategies how to estimate summary characteristics such as nearest-neighbour distance distribution function or K-function. The appropriate strategy depends on the specific summary characteristic, edge-correction method and point process under consideration. There are three main estimation approaches for $G(r)$, the border method estimator per-

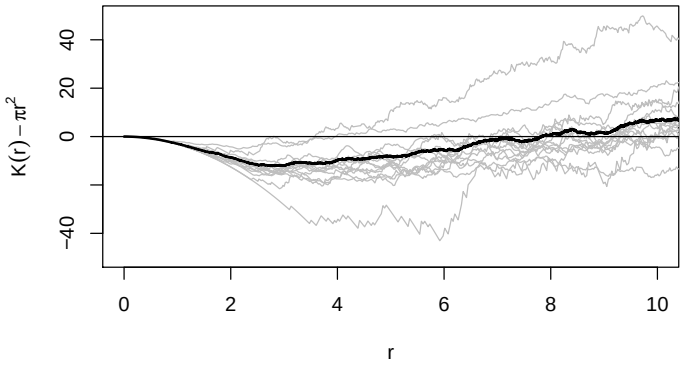

Fig. 3. Translation-corrected estimates of $K(r)$: estimates from individual subwindows (grey) and pooled estimate (thick black). Identical zero function (thin black line) corresponds to the theoretical value for the Poisson process.

forms slightly worse than the Hanisch estimator or the Kaplan–Meier estimator. For both the border and the Kaplan–Meier estimators the pooling strategy is advisable while for the Hanisch estimator weighting by the number of points can be recommended. The translational edge-correction provides a standard way to estimate $K(r)$. The pooling method and weighting by the squared point numbers are very similar and yield aggregated estimates of comparable quality. They turn to be the most convenient in the case of Poisson and regular processes. For cluster processes the alternative pooling method behaves particularly well.

ACKNOWLEDGEMENT

The author thanks the Forest Management Institute in Brandýs nad Labem for providing the data. This work was partially supported by the Czech Ministry of Education, research project MSM 0021620839 and by the Grant Agency of the Czech Republic, grant P201/10/0472. It has been presented at Prague Stochastics 2010.

(Received October 29, 2010)

REFERENCES

[1] A. J. Baddeley and R. Gill: Kaplan–Meier estimators of distance distributions for spatial point processes. Ann. Statist. *25* (1997), 263–292.

[2] A. J. Baddeley, R. A. Moyeed, C. V. Howard, and A. Boyde: Analysis of a three-dimensional point pattern with replication. J. Roy. Statist. Soc. Ser. C *42* (1993), 641–668.

[3] A. J. Baddeley and R. Turner: Spatstat: an R package for analyzing spatial point patterns. J. Statist. Software *12* (2005), 1–42.

[4] M. L. Bell and G. K. Grunwald: Mixed models for the analysis of replicated spatial point patterns. Biostatistics *5* (2004), 633–648.

[5] P. J. Diggle: Statistical Analysis of Spatial Point Patterns. 2nd edition. Arnold, London 2003.

[6] P. J. Diggle, N. Lange, and F. M. Beneš: Analysis of variance for replicated spatial point patterns in clinical neuroanatomy. J. Amer. Statist. Assoc. *86* (1991), 618–625.

[7] P. J. Diggle, J. Mateu, and H. E. Clough: A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns. Adv. in Appl. Probab. (SGSA) *32* (2000), 331–343.

[8] K.-H. Hanisch: Some remarks on estimators of the distribution function of nearest neighbour distance in stationary spatial point patterns. Statistics *15* (1984), 409–412.

[9] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan: Statistical Analysis and Modeling of Spatial Point Patterns. John Wiley & Sons, Chichester 2008.

[10] A. A. Philimonenko, J. Janáček, and P. Hozák: Statistical evaluation of colocalization patterns in immunogold labeling experiments. J. Struct. Biol. *132* (2000), 201–210.

[11] R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna 2010. URL: `http://www.R-project.org`.

[12] D. Stoyan: On estimators of the nearest neighbour distance distribution function for stationary point processes. Metrika *64* (2006), 139–150.

[13] C. G. Wager, B. A. Coull, and N. Lange: Modelling spatial intensity for replicated inhomogeneous point patterns in brain imaging. J. R. Statist. Soc. B *66* (2004), 429–446.

Zbyněk Pawlas, Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8. Czech Republic.
e-mail: pawlas@karlin.mff.cuni.cz