Miscellanea

Fast block variance estimation procedures for inhomogeneous spatial point processes

By YONGTAO GUAN

Division of Biostatistics, Yale University, New Haven, Connecticut 06520-8034, U.S.A.

yongtao.guan@yale.edu

SUMMARY

We introduce two new variance estimation procedures that use non-overlapping and overlapping blocks, respectively. The non-overlapping blocks estimator can be viewed as the limit of the thinned block bootstrap estimator recently proposed in Guan & Loh (2007), by letting the number of thinned processes and bootstrap samples therein both increase to infinity. The non-overlapping blocks estimator can be obtained quickly since it does not require any thinning or bootstrap steps, and it is more stable. The overlapping blocks estimator further improves the performance of the non-overlapping blocks with a modest increase in computation time. A simulation study demonstrates the superiority of the proposed estimators over the thinned block bootstrap estimator.

Some key words: Block variance estimator; Inhomogeneous spatial point process; Thinning.

1. INTRODUCTION

Let N be a two-dimensional spatial point process that is observed on a domain of interest $D \subset \mathbb{R}^2$. For a Borel set $B \subset \mathbb{R}^2$, let $|B|$ denote the Lebesgue measure of B, and let $N(B)$ denote the number of events of N in B. Let $\lambda(s)$ and $\lambda(s_1, s_2)$ denote the first- and second-order intensity functions (Diggle, 2003, p. 43) of N, defined respectively as

$$\lambda(s) = \lim_{|ds| \to 0} \frac{E\{N(ds)\}}{|ds|}, \quad \lambda_2(s_1, s_2) = \lim_{|ds_1|, |ds_2| \to 0} \frac{E\{N(ds_1)N(ds_2)\}}{|ds_1||ds_2|},$$

where ds is an infinitesimal region containing s. We will focus on a flexible class of spatial point processes called second-order intensity reweighted stationary processes (Baddeley et al., 2000). We assume that $\lambda(s_1, s_2) = \lambda(s_1)\lambda(s_2)g(s_1 - s_2)$ for some function $g(\cdot)$, where $g(\cdot)$ is called the pair correlation function (Møller & Waagepetersen, 2004, p. 31). In the special case in which $\lambda(s) = \lambda$ for some constant $\lambda > 0$ for all $s \in \mathbb{R}^2$, the process is further said to be second-order stationary.

It is often of interest in practice to model the first-order intensity function of the process in terms of some observed covariates. For example, in the data example given in § 4, we are interested in modelling the first-order intensity function of the point process generating the locations of 3604 *Beilschmiedia pendula* trees by using two topographical variables, elevation and elevation gradient. Such a study can yield valuable biological insight into how these two variables affect the spatial distribution of the trees. For this, we assume that $\lambda(\cdot)$ can be written as a parametric function of the covariates, where the function is completely determined by a $p \times 1$ vector of unknown regression parameters, β. We thus rewrite $\lambda(\cdot)$ as $\lambda(\cdot; \beta)$. Our main goal is to estimate and make inference about β.

To estimate β, the following Poisson-based loglikelihood function (Schoenberg, 2005) is often used:

$$U(\beta) = \frac{1}{|D|} \sum_{x \in D \cap N} \log \lambda(x; \beta) - \frac{1}{|D|} \int_D \lambda(s; \beta) \, ds, \tag{1}$$

where $\sum_{x \in D \cap N}$ means summation over all events of N that are in D. Let $\hat{\beta}$ be the maximizer of (1). Schoenberg (2005) showed that $\hat{\beta}$ is consistent for β for a wide class of spatial-temporal point process models, even if the process is not Poisson. Waagepetersen (2007) and Guan & Loh (2007) established asymptotic normality for $\hat{\beta}$ for a class of inhomogeneous Neyman–Scott processes and a class of mixing point processes, respectively.

To make inference about β, the variance of $\hat{\beta}$ needs to be estimated. Let D_n be a sequence of domains that approach $\mathbb{R}^2$ in all directions as n increases, let $\hat{\beta}_n$ and β_0 be the estimated and the true parameter vectors, respectively and let $\lambda^{(i)}(\cdot; \beta)$ be the ith partial derivative of $\lambda(\cdot; \beta)$ with respect to β. Guan & Loh (2007) showed that, under suitable conditions,

$$\Sigma_n = |D_n| \mathrm{cov}(\hat{\beta}_n) \simeq |D_n|(A_n)^{-1} B_n (A_n)^{-1},$$

where

$$A_n = \int_{D_n} \frac{\lambda^{(1)}(s; \beta_0)\{\lambda^{(1)}(s; \beta_0)\}^{\mathrm{T}}}{\lambda(s; \beta_0)} ds, \qquad (2)$$

$$B_n = A_n + \iint_{D_n} \lambda^{(1)}(u; \beta_0)\{\lambda^{(1)}(v; \beta_0)\}^{\mathrm{T}} \{g(u-v)-1\} du\, dv. \qquad (3)$$

From (2), it can be seen that A_n depends only on the first-order intensity function, and thus can be calculated easily once it has been estimated. However, from (3) it can be seen that B_n depends also on the pair correlation function. Often a parametric model for the pair correlation function is first fitted by using, say, a minimum contrast estimation procedure (Møller & Waagepetersen, 2004, pp. 182–3), and the estimated pair correlation function is then inserted into (3) in order to estimate B_n. To avoid a parametric assumption about the pair correlation function, which can be restrictive in some applications, Guan & Loh (2007) proposed a thinned block bootstrap estimator for B_n. Their procedure involves repeatedly thinning the original point process and then bootstrapping each thinned realization of the process. The thinned block bootstrap procedure assumes that the process is second-order intensity reweighted stationary, but does not require any specific parametric form for the pair correlation function. A drawback of this procedure is that it can be very time-consuming because of the repeated thinning and bootstrap steps. The goal of this paper is to propose two alternative variance estimation procedures that can be performed much more quickly.

2. BACKGROUND ON THE THINNED BLOCK BOOTSTRAP PROCEDURE

The thinned block bootstrap procedure makes use of the fact that any second-order intensity reweighted stationary process can be thinned to be second-order stationary by application of proper thinning weights. For example, Guan & Loh (2007) considered the thinned process

$$\Psi_n = \left\{ x : x \in N, \mathrm{pr}(x \text{ is retained}) = \min_{s \in D_n} \lambda(s; \beta_0)/\lambda(x; \beta_0) \right\}. \qquad (4)$$

Clearly Ψ_n is second-order stationary on D_n since its first- and second-order intensity functions can be written respectively as

$$\lambda_n = \min_{s \in D_n} \lambda(s; \beta_0), \quad \lambda_{2,n}(s_1, s_2) = (\lambda_n)^2 g(s_1 - s_2).$$

For each thinned process, Ψ_n, Guan & Loh (2007) defined the statistic

$$S_n = \sum_{x \in \Psi_n \cap D_n} \lambda^{(1)}(x; \beta_0).$$

Since Ψ_n is second-order stationary, it follows that

$$\mathrm{cov}(S_n) = \lambda_n \int_{D_n} \lambda^{(1)}(s; \beta_0)\{\lambda^{(1)}(s; \beta_0)\}^{\mathrm{T}} ds + (\lambda_n)^2 (B_n - A_n), \qquad (5)$$

where A_n and B_n are given in (2) and (3), respectively. To estimate B_n, it is sufficient to estimate the covariance matrix of S_n. For this, Guan & Loh (2007) proposed the following block bootstrap algorithm.

Step 1. Obtain a thinned realization of the process as in (4) with β_0 replaced by $\hat{\beta}_n$.

Step 2. Divide D_n into k_n non-overlapping sub-blocks, $D_{l_n}^i$ ($i = 1, \ldots, k_n$), where l_n signifies the size of each sub-block. For each $D_{l_n}^i$, let c_i denote the 'centre' of the sub-block. For each thinned process, resample B times with replacement k_n sub-blocks from $D_{l_n}^i$ ($i = 1, \ldots, k_n$). For the bth collection of the resampled random sub-blocks, let J_b be the set of k_n random indices sampled from $\{1, \ldots, k_n\}$ that are associated with the selected sub-blocks. Define

$$S_n^b = \sum_{i=1}^{k_n} \sum_{x \in \Psi_n \cap D_{l(n)}^{J_b(i)}} \lambda^{(1)}\big(x - c_{J_b(i)} + c_i; \hat{\beta}_n\big),$$

where the second summation is over all events in sub-block $D_{l_n}^{J_b(i)}$ translated into sub-block $D_{l_n}^i$. Obtain the sample covariance matrix for S_n^b ($b = 1, \ldots, B$).

Step 3. Repeat Steps 1 and 2 M times and use the average of the resulting sample covariance matrices as the estimate for the covariance matrix given in (5).

Let $\text{côv}(S_n)$ be the estimator obtained from the above algorithm. Then (5) implies the following estimator for B_n:

$$\hat{B}_n = \text{côv}(S_n)/\hat{\lambda}_n^2 - \int_{D_n} \lambda^{(1)}(s; \hat{\beta}_n)\{\lambda^{(1)}(s; \hat{\beta}_n)\}^\mathrm{T} ds/\hat{\lambda}_n + \hat{A}_n, \qquad (6)$$

where

$$\hat{\lambda}_n = \min_{s \in D_n} \lambda(s; \hat{\beta}_n), \quad \hat{A}_n = \int_{D_n} \frac{\lambda^{(1)}(s; \hat{\beta}_n)\{\lambda^{(1)}(s; \hat{\beta}_n)\}^\mathrm{T}}{\lambda(s; \hat{\beta}_n)} ds.$$

3. The proposed variance estimation procedures
3·1. *The algorithms*

Throughout this section, let $\lambda(\cdot)$ and $\lambda^{(1)}(\cdot)$ denote $\lambda(\cdot; \hat{\beta}_n)$ and $\lambda^{(1)}(\cdot; \hat{\beta}_n)$, respectively. Let Ψ_n^m denote the mth thinned realization of the process. For a fixed M and if $B \to \infty$, some simple calculations show that $\text{côv}(S_n)$ converges to

$$\text{cov}\big(S_n^b \mid \Psi_n \cap D_n\big) = \frac{1}{k_n M} \sum_{m=1}^{M} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} \sum_{x \in D_{l_n}^j \cap \Psi_n^m} \sum_{y \in D_{l_n}^j \cap \Psi_n^m} \lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(y - c_j + c_i)\}^\mathrm{T}$$

$$-\frac{1}{k_n^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n} \sum_{j_2=1}^{k_n} \sum_{x \in D_{l_n}^{j_1} \cap \Psi_n^m} \sum_{y \in D_{l_n}^{j_2} \cap \Psi_n^m} \lambda^{(1)}(x - c_{j_1} + c_i)\{\lambda^{(1)}(y - c_{j_2} + c_i)\}^\mathrm{T}.$$

If $M \to \infty$, this further converges to

$$E\big\{\text{cov}\big(S_n^b \mid \Psi_n \cap D_n\big) \mid N\big\} = \lambda_n^2 \hat{V}_{n,1} - \lambda_n^2 \hat{V}_{n,2} + \lambda_n \hat{V}_{n,3} - \lambda_n \hat{V}_{n,3}/k_n, \qquad (7)$$

where

$$\hat{V}_{n,1} = \frac{1}{k_n} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} \sum_{\substack{x \in D_{l_n}^j \cap N, y \in D_{l_n}^j \cap N \\ x \neq y}} \frac{\lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(y - c_j + c_i)\}^\mathrm{T}}{\lambda(x)\lambda(y)}, \qquad (8)$$

$$\hat{V}_{n,2} = \frac{1}{k_n^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n} \sum_{j_2=1}^{k_n} \sum_{x \in D_{l_n}^{j_1} \cap N, y \in D_{l_n}^{j_2} \cap N \atop x \neq y} \frac{\lambda^{(1)}(x - c_{j_1} + c_i)\{\lambda^{(1)}(y - c_{j_2} + c_i)\}^{\mathrm{T}}}{\lambda(x)\lambda(y)}, \tag{9}$$

$$\hat{V}_{n,3} = \frac{1}{k_n} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} \sum_{x \in D_{l_n}^{j} \cap N} \frac{\lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(x - c_j + c_i)\}^{\mathrm{T}}}{\lambda(x)}. \tag{10}$$

Note that $E(\hat{V}_{n,3}) \simeq \int_{D_n} \lambda^{(1)}(s)\{\lambda^{(1)}(s)\}^{\mathrm{T}} ds$ if $\hat{\beta}_n \simeq \beta_0$. In view of (6)–(10), it is natural to consider the following estimator for B_n:

$$\hat{B}_n = \hat{V}_{n,1} - \hat{V}_{n,2} + \hat{A}_n, \tag{11}$$

where $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ are defined as in (8) and (9), respectively. The term $-\lambda_n \hat{V}_{n,3}/k_n$ in (7) is not included in (11) since it is ignorable because $k_n \to \infty$. The quantities $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ can be calculated directly without any thinning or bootstrapping steps, so that significant computational gains can be achieved; see Appendix 1 for the computational details. Furthermore, the new estimator can be regarded as the limiting version of the thinned block bootstrap estimator as both B and M increase to infinity. It is therefore reasonable to expect (11) to be more stable than any thinned block bootstrap estimator based on a fixed B and/or M.

From (8) and (9), it can be seen that $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ are both defined in terms of the point process N observed on the non-overlapping blocks $D_{l_n}^i$ ($i = 1, \ldots, k_n$). A direct extension of the above procedure is to use overlapping blocks. For this, let D_{l_n} be the sub-block centred at the origin and let $D_n^* = \{s : D_{l_n} + s \subset D_n\}$. Define $D_{l_n}^s = D_{l_n} + s$. The new versions of (8) and (9) based on overlapping blocks are given as follows:

$$\hat{V}_{n,1} = \frac{1}{|D_n^*|} \sum_{i=1}^{k_n} \int_{D_n^*} \sum_{x \in D_{l_n}^s \cap N, y \in D_{l_n}^s \cap N \atop x \neq y} \frac{\lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(y - c_j + c_i)\}^{\mathrm{T}}}{\lambda(x)\lambda(y)} ds, \tag{12}$$

$$\hat{V}_{n,2} = \frac{1}{|D_n^*|} \sum_{i=1}^{k_n} \int_{D_n^*} \int_{D_n^*} \sum_{x \in D_{l_n}^u \cap N, y \in D_{l_n}^v \cap N \atop x \neq y} \frac{\lambda^{(1)}(x - u + c_i)\{\lambda^{(1)}(y - v + c_i)\}^{\mathrm{T}}}{\lambda(x)\lambda(y)} du\, dv. \tag{13}$$

Compared to (8) and (9), (12) and (13) use more information from the data. Intuitively, we would expect them to be more stable than their counterparts, (8) and (9), which are both based on non-overlapping blocks. Indeed, overlapping blocks have been found to yield improved variance estimators for block bootstrap (Künsch, 1989). A similar result is anticipated in the current setting.

To calculate (12) and (13), we need to approximate the integral terms involved. One obvious approach is to 'tile' the region D_n^* by a grid consisting of k_n^* 'small' cells, and then approximate the integrals by the corresponding Riemann sums (Politis & Sherman, 2001). This leads to the following estimators for $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$:

$$\hat{V}_{n,1} = \frac{1}{k_n^*} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n^*} \sum_{x \in D_{l_n}^j \cap N, y \in D_{l_n}^j \cap N \atop x \neq y} \frac{\lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(y - c_j + c_i)\}^{\mathrm{T}}}{\lambda(x)\lambda(y)}, \tag{14}$$

$$\hat{V}_{n,2} = \frac{1}{(k_n^*)^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \sum_{x \in D_{l_n}^{j_1} \cap N, y \in D_{l_n}^{j_2} \cap N \atop x \neq y} \frac{\lambda^{(1)}(x - c_{j_1} + c_i)\{\lambda^{(1)}(y - c_{j_2} + c_i)\}^{\mathrm{T}}}{\lambda(x)\lambda(y)}. \tag{15}$$

Note that (8) and (9) can be viewed as special cases of (14) and (15) with $k_n^* = k_n$.

3·2. *Theoretical justification*

Let $\hat{V}_n = \hat{V}_{n,1} - \hat{V}_{n,2}$, where $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ are given as in (8) and (9) in the non-overlapping case or (12) and (13) in the overlapping case. We would like to show that $\hat{V}_n$ converges to

$$V_n = B_n - A_n = \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} \int_{D_{l_n}^i} \int_{D_{l_n}^j} \lambda^{(1)}(u; \beta_0) \lambda^{(1)}(v; \beta_0) \{g(u - v) - 1\} du\, dv.$$

To show this, assume that, for all β in a small neighbourhood of β_0,

$$\lambda(s; \beta) > 0, \quad \lambda^{(i)}(s; \beta) < \infty \quad (i = 1, 2). \tag{16}$$

In addition, conditions on the cumulant density function of the process are needed. Define the kth-order cumulant density functions of N as

$$Q_k(s_1, \ldots, s_k) = \lim_{|ds_i| \to 0} \frac{\text{cum}\{N(ds_1), \ldots, N(ds_k)\}}{|ds_1| \cdots |ds_k|} \quad (i = 1, \ldots, k),$$

where $\text{cum}(Y_1, \ldots, Y_k)$ is the coefficient of $i^k t_1 \cdots t_k$ in the Taylor series expansion of $\log[E\{\exp(i \sum_{j=1}^k Y_j t_j)\}]$ about the origin (Brillinger, 1975, p. 19) and Y_i $(i = 1, \ldots, k)$ are random variables. The cumulant density functions are useful tools for describing the dependence between events of the process, where close-to-zero values of the cumulant density functions often indicate near independence. In the extreme case of complete independence, i.e. when N is Poisson, $Q_k(s_1, \ldots, s_k) = 0$ if at least two of $s_1, \ldots, s_k$ are different. In terms of the cumulant density functions, assume that

$$\sup_{s_1} \int \cdots \int |Q_k(s_1, \ldots, s_k)| ds_2 \ldots ds_k < \infty \quad (k = 2, 3, 4). \tag{17}$$

Condition (17) is a fairly weak condition. It holds for a broad class of inhomogeneous models including, but not limited to, the log-Gaussian Cox process (Møller et al., 1998), the inhomogeneous Neyman–Scott process (Waagepetersen, 2007) and any inhomogeneous process that is obtained by thinning a homogeneous process satisfying this condition. The following theorem, proved in Appendix 2, establishes the consistency of $\hat{V}_n$.

THEOREM 1. *Assume that* (16) *and* (17) *hold. If* $|D_{l_n}| = o(|D_n|^{1/2})$ *and* $|D_n|^{1/2}(\hat{\beta}_n - \beta_0) = O_p(1)$, *then* $(\hat{V}_n - V_n)/|D_n| \to 0$ *in mean square.*

The condition on the convergence rate of $\hat{\beta}_n$ to β_0 is a standard assumption and holds under conditions given in Guan & Loh (2007). The condition on the sub-block size is undesirable. However, a further relaxation of this condition appears to be difficult. It may not be reasonable after all to set the sub-block size to be of order larger than or equal to $|D_n|^{1/2}$ given that $|D_n|^{1/2}(\hat{\beta}_n - \beta_0) = O_p(1)$. Finally, although we are focusing on estimating the variance of $\hat{\beta}_n$ in this paper, the proposed methods can be applied to estimate the variance of any statistic taking the form $\sum_{x \in N \cap D_n} Z(x; \beta_0)$. This can be done simply by replacing $\lambda^{(1)}(\cdot)$ with $Z(\cdot)\lambda(\cdot)$ in the definitions of $\hat{V}_n$.

4. NUMERICAL STUDIES

4·1. *A simulation study*

To illustrate the superior performance of the proposed methods over the thinned block bootstrap approach, we simulated realizations from an inhomogeneous Neyman–Scott process model on a unit square. The first-order intensity function of the process was $\lambda(s) = \alpha + \beta X(s)$, where $\alpha = 7\cdot02$, $\beta = 2$ and $X(s)$ was the same covariate process as used in Guan & Loh (2007). For each simulation, we first simulated a homogeneous Poisson process as the parent process, where the intensity of the process was $\kappa = 50$. For each parent, we then generated a Poisson number of offspring. The position of each offspring relative to its parent was determined by a radially symmetric Gaussian random variable (Diggle, 2003, p. 66) with standard deviation $\omega = 0\cdot02$ or $0\cdot04$, which represent relatively strong and weak clustering, respectively.

Table 1. *Bias, standard deviation and computation time, in hours of cpu time, for* 1000 *simulations for the various variance estimators using non-overlapping blocks, overlapping blocks and the thinned block bootstrap. The target parameter is the standard deviation of $\hat{\beta}$, which is* 0·8708 *for $\omega = 0·02$ and* 0·8025 *for $\omega = 0·04$*

| | | | OB | | TBB ($B = 499$) | | TBB ($B = 999$) | |
	ω	NOB	$k_n^* = 64$	$k_n^* = 144$	$M = 5$	$M = 20$	$M = 5$	$M = 20$
Bias	0·02	−0·1233	−0·1193	−0·1209	−0·1271	−0·1264	−0·1286	−0·1273
	0·04	−0·1648	−0·1652	−0·1671	−0·1711	−0·1699	−0·1709	−0·1700
SD	0·02	0·1310	0·1197	0·1193	0·1340	0·1330	0·1339	0·1326
	0·04	0·1216	0·1114	0·1124	0·1256	0·1229	0·1241	0·1228
Time	0·02	0·0117	0·2076	0·8745	0·8041	2·9513	1·4951	5·7367
	0·04	0·0118	0·2237	0·9487	0·8203	2·9655	1·5096	5·7571

OB, overlapping blocks method; NOB, non-overlapping blocks method; TBB, thinned block bootstrap method; SD, standard deviation.

One thousand realizations of the process were simulated for each ω value. For each realization, the thinned block bootstrap estimator and the proposed methods were all applied. For the thinned block bootstrap estimator, we took $B = 499$, 999 and $M = 5$, 20. For all estimators, the sub-block size was $0·25 \times 0·25$, which led to $k_n = 16$ non-overlapping blocks. For the estimator based on overlapping blocks, $k_n^* = 64$, 144, where k_n^* was the number of small cells used to calculate (14) and (15).

Table 1 shows the bias, standard deviation and computation time for each estimator. For all estimators, there is a negative bias. The bias for $\omega = 0·04$ is larger, probably because the range of dependence is larger in this case. The non-overlapping blocks and overlapping blocks methods appear to be less biased than the thinned block bootstrap method, because of the removal of the term $-\lambda_n \hat{V}_{n,3}/k_n$ in (11). For the thinned block bootstrap method, the standard deviation generally decreases with M, or B, for a fixed B, or M. The effect of B is not very significant in this example, probably because both B values being considered here are much larger than the number of blocks to be resampled from. As expected, all the standard deviations for the thinned block bootstrap method are larger than their counterparts for the non-overlapping blocks method. For the overlapping blocks method, the standard deviations are significantly smaller than those for the non-overlapping blocks method. However, the number of overlapping blocks does not need to be very big to achieve most of the benefit. In terms of computation time, the non-overlapping blocks method is the most computationally efficient. Its computation time for 1000 simulations is only about 1/18 of that for the overlapping blocks method when $k_n^* = 64$, and about 1/70 of the smallest computation time for the thinned block bootstrap method, which occurs when $M = 5$ and $B = 499$. The computation time for the overlapping blocks method quickly increases for $k_n^* = 144$. Even in this case, it is only slightly larger than the smallest computation time for the thinned block bootstrap method. The latter increases roughly in proportion to M and B.

4·2. *Beilschmiedia pendula data*

These data contain locations of 3604 *Beilschmiedia pendula* trees in a 1000 metre by 500 metre plot in the Barro Colorado Island. Measurements on elevation and elevation gradient in the plot are also available. Waagepetersen (2007) and Guan & Loh (2007) analyzed the same data and both used the first-order intensity function model

$$\lambda(s) = \exp\{\beta_0 + \beta_1 E(s) + \beta_2 G(s)\},$$

where $E(s)$ and $G(s)$ are the elevation and elevation gradient at s, respectively. By maximizing (1), Waagepetersen (2007) obtained the estimates $\hat{\beta}_1 = 0·02$ and $\hat{\beta}_2 = 5·84$. To estimate the standard errors of $\hat{\beta}_1$ and $\hat{\beta}_2$, he used a plug-in method that assumes a parametric model for the pair correlation function. Guan & Loh (2007) estimated the standard errors by using the thinned block bootstrap method without requiring a parametric model for the pair correlation function.

We applied our proposed procedures to estimate the standard errors of $\hat{\beta}_1$ and $\hat{\beta}_2$. We used 200 metre by 100 metre sub-blocks for both the non-overlapping and overlapping methods. The non-overlapping method yielded the estimates 0·0173 for $\hat{\beta}_1$ and 2·1659 for $\hat{\beta}_2$, and the overlapping blocks method with $k_n^* = 225$ yielded the estimates 0·0164 for $\hat{\beta}_1$ and 2·1544 for $\hat{\beta}_2$. Similar results were obtained for the latter with other k_n^* values, and therefore we do not report them here. Our estimates are very similar to those of Guan & Loh (2007) and are slightly smaller than those of Waagepetersen (2007). On the basis of a 5% level of significance, we thus conclude that β_2 is significant but β_1 is not. Biologically, this implies that *Beilschmiedia pendula* prefers to live on slopes but does not favour either high or low elevations. The same conclusion was obtained in the previous analyses. The strengths of our methods are their relative ease of implementation, as compared to the computationally intensive thinned block bootstrap method, and their flexibility in not assuming a parametric model for the pair correlation function, as required by the plug-in approach.

APPENDIX 1

Computational details

For $\hat{V}_{n,1}$ given in (14), $\hat{V}_{n,1} = \hat{V}_{n,1}^a - \hat{V}_{n,1}^b$, where

$$\hat{V}_{n,1}^a = \frac{1}{k_n^*} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n^*} \left\{ \sum_{x \in D_{l_n}^j \cap N} \frac{\lambda^{(1)}(x - c_j + c_i)}{\lambda(x)} \right\} \left\{ \sum_{x \in D_{l_n}^j \cap N} \frac{\lambda^{(1)}(x - c_j + c_i)}{\lambda(x)} \right\}^{\mathrm{T}},$$

$$\hat{V}_{n,1}^b = \frac{1}{k_n^*} \sum_{i=1}^{k_n} \sum_{j=1}^{k_n^*} \sum_{x \in D_{l_n}^j \cap N} \frac{\lambda^{(1)}(x - c_j + c_i)\{\lambda^{(1)}(x - c_j + c_i)\}^{\mathrm{T}}}{\lambda(x)^2}.$$

For $\hat{V}_{n,2}$ given in (15), $\hat{V}_{n,2} = \hat{V}_{n,2}^a - \hat{V}_{n,2}^b$, where

$$\hat{V}_{n,2}^a = \frac{1}{(k_n^*)^2} \sum_{i=1}^{k_n} \left\{ \sum_{j=1}^{k_n^*} \sum_{x \in D_{l_n}^j \cap N} \frac{\lambda^{(1)}(x - c_{j_1} + c_i)}{\lambda(x)} \right\} \left\{ \sum_{j=1}^{k_n^*} \sum_{x \in D_{l_n}^j \cap N} \frac{\lambda^{(1)}(x - c_{j_1} + c_i)}{\lambda(x)} \right\}^{\mathrm{T}},$$

$$\hat{V}_{n,2}^b = \frac{1}{(k_n^*)^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \sum_{x \in D_{l_n}^{j_1} \cap D_{l_n}^{j_2} \cap N} \frac{\lambda^{(1)}(x - c_{j_1} + c_i)\{\lambda^{(1)}(x - c_{j_2} + c_i)\}^{\mathrm{T}}}{\lambda(x)^2}.$$

Formulae for $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ in the non-overlapping case correspond to $k_n^* = k_n$.

APPENDIX 2

Proof of Theorem 1

Direct application of Taylor series expansion shows that it is sufficient to prove the theorem for $\hat{\beta}_n = \beta_0$. Here we outline the proof only for the overlapping blocks estimator. The proof in the non-overlapping case follows trivially. We therefore consider the case in which $\hat{V}_{n,1}$ and $\hat{V}_{n,2}$ are given by (14) and (15), respectively, since they, not (12) and (13), are used in practice. First note that

$$E(\hat{V}_n) = \sum_{i=1}^{k_n} \iint_{D_{l_n}^i} \lambda^{(1)}(u)\lambda^{(1)}(v)\{g(u - v) - 1\}du\,dv$$

$$- \frac{1}{(k_n^*)^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \iint_{D_{l_n}^i} \lambda^{(1)}(u)\lambda^{(1)}(v)\{g(u - v + c_{j_1} - c_{j_2}) - 1\}du\,dv. \qquad (\text{A1})$$

Therefore,

$$E(\hat{V}_n) - V_n = -\sum_{i=1}^{k_n} \sum_{j \neq i} \int_{D_{l_n}^i} \int_{D_{l_n}^j} \lambda^{(1)}(u)\lambda^{(1)}(v)\{g(u-v)-1\}du\,dv$$

$$-\frac{1}{(k_n^*)^2} \sum_{i=1}^{k_n} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \iint_{D_{l_n}^i} \lambda^{(1)}(u)\lambda^{(1)}(v)\{g(u-v+c_{j_1}-c_{j_2})-1\}du\,dv. \quad \text{(A2)}$$

Equation (A1) is of order $o(|D_n|)$, from the proof of Theorem 2 of Guan & Loh (2007). Equation (A2) is also of order $o(|D_n|)$, because of (16) and (17) and the fact that, for each fixed j_1, the number of overlapping blocks that are within a distance l_n of $D_{l_n}^{j_1}$ is of order $o(k_n^*/k_n)$.

For the variance of $\hat{V}_n$, some elementary algebra shows that the variance is bounded by the following terms:

$$\frac{Ck_n^2}{(k_n^*)^2} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \int_{D_{l_n}^{j_1} \cap D_{l_n}^{j_2}} \int_{D_{l_n}^{j_1} \cap D_{l_n}^{j_2}} g(x_1-x_2)dx_1dx_2,$$

$$\frac{Ck_n^2}{(k_n^*)^2} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \int_{D_{l_n}^{j_1} \cap D_{l_n}^{j_2}} \int_{D_{l_n}^{j_1}} \int_{D_{l_n}^{j_2}} |Q_3(x_1,x_2,x_3)+2Q_2(x_1,x_2)+Q_2(x_2,x_3)|\,dx_1dx_2dx_3,$$

$$\frac{Ck_n^2}{(k_n^*)^2} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \int_{D_{l_n}^{j_1}} \int_{D_{l_n}^{j_1}} \int_{D_{l_n}^{j_2}} \int_{D_{l_n}^{j_2}} |Q_3(x_1,x_2,x_3)|\,dx_1dx_2dx_3dx_4,$$

$$\frac{Ck_n^2}{(k_n^*)^2} \sum_{j_1=1}^{k_n^*} \sum_{j_2=1}^{k_n^*} \left\{ \int_{D_{l_n}^{j_1}} \int_{D_{l_n}^{j_2}} |Q_2(x_1,x_2)|\,dx_1dx_2 \right\}^2.$$

All the above terms are of order $|D_n|^2/k_n$ because of (16) and (17) and the fact that, for each fixed j_1, the number of overlapping blocks that are within a distance l_n of $D_{l_n}^{j_1}$ is of order $o(k_n^*/k_n)$. $\qquad\square$

REFERENCES

BADDELEY, A. J., MØLLER, J. & WAAGEPETERSEN, R. P. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statist. Neer.* **54**, 329–50.

BRILLINGER, D. R. (1975). *Time Series*: *Data Analysis and Theory*. New York: Holt, Rinehart & Winston.

DIGGLE, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*. New York: Oxford University Press.

GUAN, Y. & LOH, J. M. (2007). A thinned block bootstrap variance estimation procedure for inhomogeneous spatial point patterns. *J. Am. Statist. Assoc.* **102**, 1377–86.

KÜNSCH, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17**, 1217–41.

MØLLER, J., SYVERSVEEN, A. R. & WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25**, 451–82.

MØLLER, J. & WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. New York: Chapman and Hall.

POLITIS, D. N. & SHERMAN, M. (2001). Moment estimation for statistics from marked point processes. *J. R. Statist. Soc.* B **63**, 261–75.

SCHOENBERG, F. P. (2005). Consistent parametric estimation of the intensity of a spatial-temporal point process. *J. Statist. Plan. Infer.* **128**, 79–93.

WAAGEPETERSEN, R. P. (2007). An estimating function approach to inference for inhomogeneous Neyman–Scott processes. *Biometrics* **62**, 252–8.