Published in final edited form as:

J Am Stat Assoc. 2009 December 1; 104(488): 1482. doi:10.1198/jasa.2009.tm08541.

On Nonparametric Variance Estimation for Second-Order Statistics of Inhomogeneous Spatial Point Processes With a Known Parametric Intensity Form

Yongtao Guan

Division of Biostatistics, Yale University, New Haven, CT 06520-8034 (yongtao.guan@yale.edu)

Abstract

We introduce new variance estimation procedures for second-order statistics that are computed from a single realization of intensity reweighted stationary spatial point processes. The statistics are defined either on a subset B of the observation window or on the whole window. For the former, we use subblocks that have the same size and shape as B as “replicates” of B in order to estimate the target variance. For the latter, we develop a subsampling estimator for a key component in the target variance and estimate its other components by method-of-moment methods. Under some suitable conditions, we prove that the proposed variance estimators are consistent for the target variances in both cases. Simulations and an application to a real data example are used to demonstrate the usefulness of the proposed methods. This article has supplemental material online.

Keywords

Intensity reweighted stationary processes; Second-order statistics; Variance estimation

1. INTRODUCTION

Let N be a spatial point process that is observed on a two-dimensional observation window $W \subset \mathbb{R}^2$. For a Borel set $B \subset \mathbb{R}^2$, let $|B|$ denote the Lebesgue measure of B , and let $N(B)$ be the number of events of N falling in B . Following Diggle (2003), we define the k th-order intensity function

$$\lambda_k(\mathbf{s}_1, \dots, \mathbf{s}_k) = \lim_{|\Delta \mathbf{s}_1|, \dots, |\Delta \mathbf{s}_k| \rightarrow 0} \frac{E[N(\Delta \mathbf{s}_1) \cdots N(\Delta \mathbf{s}_k)]}{|\Delta \mathbf{s}_1| \cdots |\Delta \mathbf{s}_k|},$$

where $\Delta \mathbf{s}$ denotes a small region containing $\mathbf{s}$. In what follows, we refer to the first-order intensity function as the intensity function, and assume that $\lambda_1(\mathbf{s}) = \lambda(\mathbf{s}; \beta)$, where $\lambda(\cdot; \beta)$ is a known parametric function and β is some unknown parameters.

© 2009 American Statistical Association

SUPPLEMENTAL MATERIALS

Additional algorithms and proofs: This document gives details of two additional algorithms to estimate of the covariance matrix of U_1 and the covariance between U_1 and $T(W)$. It also contains proofs for Theorems 1 & 2 and a small simulation study used to assess the effect of misspecifying the intensity model on the performance of the proposed variance estimation method for the empirical K -function. The equation numbers in this document follow those in the paper. (K variance supplement.pdf)

Supplementary materials for this article are available at <http://pubs.amstat.org/toc/jasa/104/488>.

Following Baddeley, Møller, and Waagepetersen (2000), we say that a spatial point process is k th-order intensity reweighted stationary if

$$\lambda_k(\mathbf{s}_1, \dots, \mathbf{s}_k) = \lambda(\mathbf{s}_1) \cdots \lambda(\mathbf{s}_k) g_k(\mathbf{s}_2 - \mathbf{s}_1, \dots, \mathbf{s}_k - \mathbf{s}_1), \quad (1)$$

where $g_k(\cdot)$ is a function that depends only on the interpoint lags $\mathbf{s}_2 - \mathbf{s}_1, \dots, \mathbf{s}_k - \mathbf{s}_1$. When $k = 2$, (1) implies the second-order intensity reweighted stationarity given in Baddeley, Møller, and Waagepetersen (2000). The process is further said to be isotropic if $g_2(\mathbf{s}_2 - \mathbf{s}_1) = g(\|\mathbf{s}_2 - \mathbf{s}_1\|)$ for some function $g(\cdot)$, where $\|\cdot\|$ is the Euclidean norm. This in turn yields the widely used reduced moment measure, or the K -function,

$$K(t) = 2\pi \int_0^t u g(u) du.$$

The K -function plays an important role in spatial point pattern analysis (Diggle 2003), due to its ability to quantify the clustering strength of a spatial point process and its ease of estimation. Typically $K(t) > \pi t^2$ indicates attraction at distance t and $K(t) < \pi t^2$ indicates repulsion at that distance. For the tropical rain forest example in Section 7, we want to compare the clustering strengths of two spatial point patterns, each comprised of the spatial locations of a given tree species in the study region. Such an analysis is useful because it can reveal biological insights on the difference in the seed dispersal modes of the species, which partly caused the clustering. The analysis can be done by comparing the empirical K -functions, which, if ignoring any edge effect, can be obtained as (Møller and Waagepetersen 2004)

$$\widehat{K}_{\text{inhom}}(t) = \frac{1}{|W|} \sum_{\mathbf{u}, \mathbf{v} \in (N \cap W)}^{\neq} \frac{I(\|\mathbf{u} - \mathbf{v}\| \leq t)}{\lambda(\mathbf{u}; \widehat{\beta}) \lambda(\mathbf{v}; \widehat{\beta})}. \quad (2)$$

In the above, $\neq$ means summation over distinct pairs, $I(\cdot)$ is an indicator function, and $\widehat{\beta}$ is an estimator for β . We set $\widehat{\beta}$ as the estimator obtained by maximizing a Poisson maximum likelihood (Schoenberg 2005), because the distributional properties of $\widehat{\beta}$ in this case are well developed (Guan and Loh 2007; Waagepetersen 2007).

For inference, the variance of $\widehat{K}_{\text{inhom}}(t)$ is needed. As given in Section 3, the variance depends on integrals involving the second to the forth-order intensity functions. These intensity functions are often not all available in closed forms and the calculation of the related integrals can be cumbersome even when they are. A common practice instead is to use parametric bootstrap (Diggle et al. 2007). However, it has been reported in literature (Cressie 1993; Møller, Syversveen, and Waagepetersen 1998) that commonly used spatial point process models often provide unsatisfactory fits to data. This in turn raises questions for the validity of parametric bootstrap. In this paper, we develop nonparametric variance estimators for second-order statistics taking the general form

$$S_0(B; \widehat{\beta}) = \sum_{\mathbf{u}, \mathbf{v} \in (N \cap B)}^{\neq} \frac{f(\mathbf{u}, \mathbf{v})}{\lambda(\mathbf{u}; \widehat{\beta}) \lambda(\mathbf{v}; \widehat{\beta})},$$

where $B \subseteq W$ and $f(\cdot)$ is a known function satisfying $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v}, \mathbf{u})$. In the case that $B = W$ and $f(\mathbf{u}, \mathbf{v}) = I(\|\mathbf{u} - \mathbf{v}\| \leq t)$, $S_0(B; \widehat{\beta})$ is $|W| \widehat{K}_{\text{inhom}}(t)$. Thus, the variance of $\widehat{K}_{\text{inhom}}(t)$ can be readily obtained from that of $S_0(B; \widehat{\beta})$. Statistics in the form of $S_0(B; \widehat{\beta})$ are also useful in residual

diagnostics and parametric estimation of the second-order intensity function; see Section 5 for details.

Our work is motivated by Guan and Loh (2007) who developed nonparametric variance estimators for simpler first-order statistics calculated from intensity reweighted stationary spatial point processes. In contrast to their work, we consider the task of estimating the variance of more complicated second-order statistics. As can be seen from (5), the target variance involves not only the covariance matrix of a first-order statistic (i.e., U_1) but also other more complicated variance and covariance terms. Furthermore, our work nicely avoids the use of repeated thinnings and block bootstrap employed in previous work. This is significant because it leads to not only a more elegant presentation but also a far superior computational performance. Lastly, the essence of the proposed method when $B = W$ is based on subsampling (Politis 1999). To the best of our knowledge, this is the first time that a subsampling method has been developed for inhomogeneous spatial point processes.

The proposed methods are developed in two steps. First, we develop a consistent estimator for the variance of $S_0(B; \hat{\beta})$ when B is a small subset of W . To achieve this, we move across W a subblock that has the same size and shape as B and treat the obtained subblocks as “replicates” of B in order to estimate the target variance. Second, we apply the idea of subsampling to estimate a key component in the variance of $S_0(B; \hat{\beta})$ when $B = W$. Specifically, we define the estimator as that for the same component in the variance of $S_0(B; \hat{\beta})$ when B is a subset of W . Other components of the target variance are obtained by method-of-moment methods. For computational convenience, we standardize $S_0(B; \hat{\beta})$ as $S(B) = \{S_0(B; \hat{\beta}) - E[S_0(B; \beta_0)]\} / |B|^{1/2}$, where β_0 is the true value of β .

The remainder of the article is organized as follows. In Section 2, we give theoretical background on the variance of $S(B)$. In Sections 3 and 4, we propose our variance estimation methods for $S(B)$ when B is a small subset of W and when $B = W$, respectively. We discuss applications of the proposed methods in Section 5, present some simulation results in Section 6 and apply the proposed methods to a real data example in Section 7. A discussion is given in Section 8 and all theoretical proofs are in the supplemental material. In the subsequent development, we assume that the spatial point process is intensity reweighted stationary up to the fourth order.

2. THEORETICAL BACKGROUND

2.1 Poisson Maximum Likelihood

To calculate the second-order statistic of interest, it is necessary to estimate the unknown parameter vector β in the intensity function. To do so, Schoenberg (2005) proposed maximizing the (log) Poisson maximum likelihood criterion

$$L(\beta) = \frac{1}{|W|} \sum_{\mathbf{u} \in (N \cap W)} \log \lambda(\mathbf{u}; \beta) - \frac{1}{|W|} \int_W \lambda(\mathbf{u}; \beta) d\mathbf{u}. \quad (3)$$

The maximizer of (3) is called the Poisson maximum likelihood estimator (PMLE; denoted by $\hat{\beta}$). Note that the score function of $L(\beta)$ is an unbiased estimating equation and as such, consistency of $\hat{\beta}$ is expected following standard results for estimating equations (e.g., Crowder 1986; van der Vaart 1998). Schoenberg (2005) gave conditions for consistency of $\hat{\beta}$ for a class of possibly non-Poisson spatial-temporal point processes. For simplicity, we assume that $\hat{\beta}$ is consistent for β_0 .

Let $\mathbf{A}^T$ be the transpose of an arbitrary matrix $\mathbf{A}$ and let $\lambda^{(1)}(\cdot; \beta)$ be the vector of the first partial derivative of $\lambda(\cdot; \beta)$ with respect to β . Under suitable conditions, Guan and Loh (2007) showed that

$$|W|^{1/2}(\widehat{\beta} - \beta_0) \approx \mathbf{U}_2^{-1} \mathbf{U}_1, \quad (4)$$

where

$$\begin{aligned} \mathbf{U}_1 &= \frac{1}{\sqrt{|W|}} \left[\sum_{\mathbf{u} \in (N \cap W)} \frac{\lambda^{(1)}(\mathbf{u}; \beta_0)}{\lambda(\mathbf{u}; \beta_0)} - \int_W \lambda^{(1)}(\mathbf{u}; \beta_0) d\mathbf{u} \right], \\ \mathbf{U}_2 &= \frac{1}{|W|} \int_W \frac{\lambda^{(1)}(\mathbf{u}; \beta_0) \lambda^{(1)}(\mathbf{u}; \beta_0)^T}{\lambda(\mathbf{u}; \beta_0)} d\mathbf{u}. \end{aligned}$$

Note that $\mathbf{U}_2$ is fixed for a given observation window W . Furthermore, the covariance matrix of $\mathbf{U}_1$ is typically of order one under suitable conditions (Guan and Loh 2007), in the sense that the variance of $\mathbf{a}^T \mathbf{U}_1$ is of order one for any constant vector $\mathbf{a}$.

2.2 Effect of $\widehat{\beta}$ on the Variance of $S(B)$

The effect of using an estimated but not the true β on the variance of $S(B)$ is generally ignorable if B is small relative to W . Below we study this effect when $B = W$. To that end, let $\int_W^{\otimes k} = \int_W \cdots \int_W$ for k times. Then by Taylor series expansion,

$$S(W) \approx T(W) - 2\mathbf{U}_3^T [|W|^{1/2}(\widehat{\beta} - \beta_0)], \quad (5)$$

where

$$\begin{aligned} T(B) &= \frac{1}{\sqrt{|B|}} \left[\sum_{\mathbf{u}, \mathbf{v} \in (N \cap B)}^{\neq} \frac{f(\mathbf{u}, \mathbf{v})}{\lambda(\mathbf{u}; \beta_0) \lambda(\mathbf{v}; \beta_0)} - \int_{B^{\otimes 2}} f(\mathbf{u}, \mathbf{v}) g(\mathbf{u} - \mathbf{v}) d\mathbf{u} d\mathbf{v} \right], \\ \mathbf{U}_3 &= \frac{1}{|W|} \sum_{\mathbf{u}, \mathbf{v} \in (N \cap W)}^{\neq} \frac{f(\mathbf{u}, \mathbf{v}) \lambda^{(1)}(\mathbf{u}; \beta_0)}{\lambda(\mathbf{u}; \beta_0)^2 \lambda(\mathbf{v}; \beta_0)}. \end{aligned}$$

Note that the variance of $T(W)$ and $|W|^{1/2}(\widehat{\beta} - \beta_0)$ are typically both of order one. Furthermore, it can be shown that $\mathbf{U}_3$ is a consistent estimator for its expected value under mild conditions (Waagepetersen and Guan 2009). Thus, the variance of $S(W)$ is different to that of $T(W)$, where $T(W)$ can be regarded as $S(W)$ defined in terms of β_0 .

We construct $S(W)$ based on the PMLE because of the known distributional properties of the latter. Note that given (4), we can rewrite (5) as

$$S(W) \approx T(W) - 2\mathbf{U}_3^T \mathbf{U}_2^{-1} \mathbf{U}_1,$$

where $\mathbf{U}_1$ and $\mathbf{U}_2$ are given in the last section. Because $\mathbf{U}_2$ can be estimated directly based on a fitted intensity function and $\mathbf{U}_3$ is roughly a constant, it then follows that in order to estimate the variance of $S(W)$, we only need to estimate the variance of $T(W)$, the covariance matrix of $\mathbf{U}_1$ and the covariance between $T(W)$ and $\mathbf{U}_1$.

In what follows, we focus on the problem of estimating the variance of $T(W)$. Computational details for estimating the remaining terms are given in the supplemental materials. As an intermediate step, we first develop an estimation procedure for the variance of $T(B)$ where B is a small subset of W . The need for this will become clear in Section 4. We note that it is also of its own practical importance to estimate the variance of $T(B)$. See Section 5.1 for an example. For ease of presentation, we will abbreviate $\lambda(\cdot; \beta_0)$ as $\lambda(\cdot)$ from now on and assume that the true intensity $\lambda(\cdot)$ is known. We make this assumption only when estimating the variance of $T(W)$. In practice, the proposed methods can be directly applied with $\lambda(\cdot)$ being replaced by $\lambda(\cdot; \hat{\beta})$. Theoretical properties of the resulting estimators can be derived by repeatedly applying Taylor series expansion to $\lambda(\cdot; \hat{\beta})$ at $\hat{\beta} = \beta_0$ and by assessing the contribution of the higher-order terms based on (4) to each term of the proposed estimators. The derivation is rather tedious and we omit its presentation.

3. VARIANCE ESTIMATION WHEN B IS SMALL

3.1 The Proposed Method

Under the assumed intensity reweighted stationarity, it follows from a direct application of Campbell's Theorem (Daley and Vere-Jones 1988) that

$$V(B) = \text{var}[T(B)] = 2V_1(B) + 4V_2(B) + V_3(B), \quad (6)$$

where

$$\begin{aligned} V_1(B) &= \frac{1}{|B|} \int_{B^{\otimes 2}} \frac{f(\mathbf{u}, \mathbf{v})^2}{\lambda(\mathbf{u})\lambda(\mathbf{v})} g(\mathbf{v} - \mathbf{u}) d\mathbf{u} d\mathbf{v}, \\ V_2(B) &= \frac{1}{|B|} \int_{B^{\otimes 3}} \frac{f(\mathbf{u}, \mathbf{v})f(\mathbf{u}, \mathbf{w})}{\lambda(\mathbf{u})} \\ &\quad \times g_3(\mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{u}) d\mathbf{u} d\mathbf{v} d\mathbf{w}, \\ V_3(B) &= \frac{1}{|B|} \int_{B^{\otimes 4}} f(\mathbf{u}, \mathbf{v})f(\mathbf{w}, \mathbf{z}) \\ &\quad \times g_4(\mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{u}, \mathbf{z} - \mathbf{u}) d\mathbf{u} d\mathbf{v} d\mathbf{w} d\mathbf{z} \\ &\quad - |B| \left[\frac{1}{|B|} \int_{B^{\otimes 2}} f(\mathbf{u}, \mathbf{v})g(\mathbf{v} - \mathbf{u}) d\mathbf{u} d\mathbf{v} \right]^2 \\ &\equiv V_4(B) - |B| [\mu(B)]^2. \end{aligned} \quad (7)$$

Except for Poisson processes, in which case $V_3(B) = 0$, $V_3(B)$ is generally of the same order as $V_1(B)$ and $V_2(B)$ and thus is also important. To estimate these terms efficiently, it is desirable to use all observed data and not just those in B . The assumed intensity reweighted stationarity provides a nice framework to achieve this.

Let $B(\mathbf{s}) = B + \mathbf{s}$ be a shifted copy of B and let $W(B) = \{\mathbf{s}: B(\mathbf{s}) \subset W\}$. Then for some function $h(\mathbf{u}, \mathbf{v})$, $V_1(B)$ and $\mu(B)$ both take the form

$$\frac{1}{|B|} \int_{B^{\otimes 2}} h(\mathbf{u}, \mathbf{v})g(\mathbf{v} - \mathbf{u}) d\mathbf{u} d\mathbf{v},$$

which can be estimated unbiasedly based on a given $B(\mathbf{s})$ by

$$\frac{1}{|B|} \sum_{\mathbf{u}, \mathbf{v} \in [N \cap B(\mathbf{s})]}^{\neq} \frac{h(\mathbf{u} - \mathbf{s}, \mathbf{v} - \mathbf{s})}{\lambda(\mathbf{u})\lambda(\mathbf{v})}.$$

Let $\hat{V}_1(\mathbf{s}, B)$ and $\hat{\mu}(\mathbf{s}, B)$ be the resulting estimators for $V_1(B)$ and $\mu(B)$, respectively. Similarly, we can obtain the unbiased estimators for $V_2(B)$ and $V_4(B)$,

$$\begin{aligned}\hat{V}_2(\mathbf{s}, B) &= \frac{1}{|B|} \sum_{\mathbf{u}, \mathbf{v}, \mathbf{w} \in [N \cap B(\mathbf{s})]}^{\neq} \frac{f(\mathbf{u}-\mathbf{s}, \mathbf{v}-\mathbf{s})f(\mathbf{u}-\mathbf{s}, \mathbf{w}-\mathbf{s})}{\lambda(\mathbf{u}-\mathbf{s})\lambda(\mathbf{u})\lambda(\mathbf{v})\lambda(\mathbf{w})}, \\ \hat{V}_4(\mathbf{s}, B) &= \frac{1}{|B|} \sum_{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z} \in [N \cap B(\mathbf{s})]}^{\neq} \frac{f(\mathbf{u}-\mathbf{s}, \mathbf{v}-\mathbf{s})f(\mathbf{w}-\mathbf{s}, \mathbf{z}-\mathbf{s})}{\lambda(\mathbf{u})\lambda(\mathbf{v})\lambda(\mathbf{w})\lambda(\mathbf{z})},\end{aligned}$$

where the two $\neq$ signs in the above signify summations over any three and four distinct events of N , respectively. A “pooled” estimator for $V_i(B)$, $i = 1, 2, 3$, based on all $B(\mathbf{s})$, is then given by

$$\hat{V}_i(B) = \frac{1}{|W(B)|} \int_{W(B)} \hat{V}_i(\mathbf{s}, B) d\mathbf{s}, \quad i=1, 2, 3, \quad (8)$$

where

$$\begin{aligned}\hat{V}_3(\mathbf{s}, B) &= \hat{V}_4(\mathbf{s}, B) - |B|[\hat{\mu}(B)]^2, \\ \hat{\mu}(B) &= \frac{1}{|W(B)|} \int_{W(B)} \hat{\mu}(\mathbf{s}, B) d\mathbf{s}.\end{aligned}$$

The variance of $T(B)$ can then be estimated by plugging (8) back into (6). Let $\hat{V}(B)$ be the resulting estimator. We will establish consistency of $\hat{V}(B)$ in the next section. For ease of theoretical derivations, we assume that

$$\begin{aligned}|f(\mathbf{u}, \mathbf{v})| &< c \quad \text{and} \\ \int_{\mathbb{R}^2} |f(\mathbf{u}, \mathbf{v})| d\mathbf{v} &< c, \quad \text{where } 0 < c < \infty.\end{aligned} \quad (9)$$

3.2 Theoretical Justification

We study the large sample properties of the proposed variance estimator under an increasing-domain setup. To be specific, let W_n be a “convex averaging sequence” of subsets of $\mathbb{R}^2$, that is, W_n is bounded and convex, $W_n \subset W_{n+1}$ for any n and the radius of the largest circle contained in W_n goes to infinity as $n \rightarrow \infty$. Furthermore, let B_n be a sequence of convex subsets of W_n such that $|B_n|/|W_n| \rightarrow 0$ as $n \rightarrow \infty$, that is, B_n can change with n but should remain small relative to W_n .

Let c_1 and c_2 be some finite, positive constants such that $c_1 < c_2$. Define

$$\tilde{V}(\mathbf{s}, B_n) = \frac{1}{|B_n|} \left[\sum_{\mathbf{u}, \mathbf{v} \in [N \cap B_n(\mathbf{s})]}^{\neq} \frac{f(\mathbf{u}-\mathbf{s}, \mathbf{v}-\mathbf{s})}{\lambda(\mathbf{u})\lambda(\mathbf{v})} - \mu(B_n) \right]^2,$$

where $\mu(B_n)$ is $\mu(B)$ that is given in (7) but defined over B_n . We assume that

$$c_1 < \lambda(\cdot) < c_2 \quad \text{and} \quad g_i(\cdot) < c \quad \text{for } i=2, 3, 4, \quad (10)$$

$$\int_{\mathbb{R}^2} |g_4(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2+\mathbf{u}_3) - g(\mathbf{u}_1)g(\mathbf{u}_3)| d\mathbf{u}_2 < c, \quad (11)$$

$$\int_{\mathbb{R}^2} |g_6(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5) - g_3(\mathbf{u}_1, \mathbf{u}_2)g_3(\mathbf{u}_4, \mathbf{u}_5)| d\mathbf{u}_3 < c, \quad (12)$$

$$\text{var}[\tilde{V}(\mathbf{s}, B_n)] < c \quad \text{for all } \mathbf{s} \in \mathbb{R}^2, \quad (13)$$

$$\sup_{\substack{\mathbf{s}_1, \mathbf{s}_2: \|\mathbf{s}_1 - \mathbf{s}_2\| > d \\ \text{as } d \rightarrow \infty.}} \text{cov}[\tilde{V}(\mathbf{s}_1, B_n), \tilde{V}(\mathbf{s}_2, B_n)] \rightarrow 0 \quad (14)$$

Condition (10) is a mild condition. In particular, the condition on $\lambda(\cdot)$ is satisfied if the covariates are bounded and the intensity function is continuous; the condition on $g_i(\cdot)$ holds for a wide class of spatial point process models, for example, Neyman–Scott processes (Waagepetersen 2007) and log-Gaussian Cox processes (Møller, Syversveen, and Waagepetersen 1998).

Conditions (11) and (12) are standard assumptions of weak dependence for spatial point processes (Guan and Sherman 2007). They hold if the spatial point process is Brillinger-mixing (B -mixing; Heinrich 1988). The condition of B -mixing entails that all (reduced) cumulant densities (Daley and Vere-Jones 1988) of the point process are integrable. Examples of B -mixing processes include Poisson process, Poisson cluster process and a wide class of doubly stochastic Poisson processes (i.e., Cox processes; Heinrich 1985).

Condition (13) is also a mild condition. Following the proof of theorem 3 of Jolivet (1981), we see that (13) holds if the first eight (reduced) cumulant densities of the spatial point process are integrable. This is satisfied by all B -mixing processes. Condition (14) can be verified directly for a given type of spatial point processes, because the covariance can be expressed as integrals involving the first eight cumulant densities. Alternatively, it holds under suitable moment and strong mixing conditions (Doukhan 1994).

Theorem 1—If conditions (9)–(14) are true, then $\hat{V}(B_n) \rightarrow V(B_n)$ in probability.

4. VARIANCE ESTIMATION WHEN $B = W$

4.1 The Proposed Method

The variance estimation procedure in the last section uses moving subblocks $B(\mathbf{s})$, $\mathbf{s} \in W(B)$, as “replicates” of B to estimate the variance of $T(B)$. When $B = W$, such replicates are not available and the procedure is thus no longer applicable. However, the expression of the target variance remains the same as (6), except that the set B therein is now replaced by W . Thus, we only need to estimate the three terms involved, that is, $V_1(W)$, $V_2(W)$, and $V_3(W)$, in order to estimate the target variance.

To estimate $V_1(W)$ and $V_2(W)$, we can simply set

$$\widehat{V}_1(W) = \widehat{V}_1(\mathbf{s}, B) \quad \text{and} \quad \widehat{V}_2(W) = \widehat{V}_2(\mathbf{s}, B), \\ \text{with } \mathbf{s} = \mathbf{0} \text{ and } B = W,$$

where $\hat{V}_1(\mathbf{s}, B)$ and $\hat{V}_2(\mathbf{s}, B)$ are as given in Section 3.1. Under suitable conditions, we can show that $\hat{V}_i(W) \rightarrow V_i(W)$ in probability, $i = 1, 2$. We may wish to similarly estimate $V_3(W)$ by $\hat{V}_3(\mathbf{0}, W)$. However, this is not valid because it can be shown that $\hat{V}_3(\mathbf{0}, W) = -2\hat{V}_1(W) - 4\hat{V}_2(W)$, which always leads to a zero variance estimate.

We develop an asymptotically unbiased and consistent estimator for the term $V_3(W)$. To that end, we assume that for some constant η ,

$$V_3(B) \rightarrow \eta, \quad \text{if } |B| \rightarrow \infty. \quad (15)$$

To see when condition (15) holds, suppose that $f(\mathbf{u}, \mathbf{v}) = f(\mathbf{v} - \mathbf{u})$, that is, $f(\mathbf{u}, \mathbf{v})$ is a function depending only on the lag, and that $g_4(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2 + \mathbf{u}_3) - g(\mathbf{u}_1)g(\mathbf{u}_3) \geq 0$ for all $\mathbf{u}_1, \mathbf{u}_2$, and $\mathbf{u}_3 \in \mathbb{R}^2$. The latter implies $V_3(B') \geq V_3(B) > 0$ if $B \subset B'$. Then (15) holds if $f(\mathbf{u}, \mathbf{v}) \geq 0$ and condition (11) holds. For stationary quantitative spatial processes, (15) is assumed to develop subsampling variance estimation methods for statistics defined on such processes (Politis 1999). Here a quantitative spatial process means a spatial process with a quantitative (i.e., numerical) random variable defined at each location, for example, a spatial Gaussian process. Subsampling methods are also available for nonstationary quantitative spatial processes (Sherman 1997; Heagerty and Lumley 2000; Ekström and Luna 2004) and for stationary spatial point processes (Guan and Sherman 2007), all requiring a similar condition.

Condition (15) implies that if both B and W are large, then $V_3(B) \approx V_3(W)$. Thus, we can use $V_3(B)$ to approximate $V_3(W)$ if B is large. If B is also small relative to W , then we can estimate $V_3(B)$ by the proposed variance estimator given in Section 3.1. Below we study theoretical properties of the resulting estimator.

4.2 Theoretical Justification

Let W_n be a sequence of $n \times n$ square windows and let B_n be a sequence of $l_n \times l_n$ subsquare windows contained in W_n , where $l_n \rightarrow \infty$ but $l_n/n \rightarrow 0$ as $n \rightarrow \infty$. Let $\hat{V}_i(W_n)$ be the resulting estimators for $V_i(W_n)$, $i = 1, 2, 3$. Consistency of $\hat{V}_1(W_n)$ and $\hat{V}_2(W_n)$ follow from conditions (9)–(12). Consistency of $\hat{V}_3(W_n)$ for $V_3(W_n)$ is a direct result of Theorem 1. Thus, the proposed variance estimator is consistent for the target variance. Below we give further details on the rates for the bias and variance of $\hat{V}_3(B_n)$. To that end, we need the following three additional conditions:

$$|V_3(B_n) - \eta| < c |B_n|^{-1/2}, \quad (16)$$

$$\frac{1}{|B_n|^2} \text{var}\{[\mu(B_n) - \widehat{\mu}(B_n)]^2\} = o\left(\frac{|B_n|}{|W_n|}\right), \quad (17)$$

$$\text{cov}[\hat{V}_1(\mathbf{s}_1, B_n), \hat{V}_1(\mathbf{s}_2, B_n)] = O(\|\mathbf{s}_1 - \mathbf{s}_2\|^{-\alpha}) \quad \text{for some } \alpha > 2. \quad (18)$$

Condition (16) specifies the rate of the bias for the proposed estimator. A similar assumption is often used for subsampling methods in the quantitative process case (Sherman 1996). In the current context, (16) holds if the quantity $|g_4(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2 + \mathbf{u}_3) - g(\mathbf{u}_1)g(\mathbf{u}_3)|$ decays to zeros at a rate faster than $\|\mathbf{u}_2\|^{-3}$. Condition (17) is a mild condition. For B -mixing processes, it can be

shown that $E\{[\mu(B_n) - \hat{\mu}(B_n)]^4\}$ is of order $|B_n|^4/|W_n|^2$ (Jolivet 1981) so that (17) is true. Condition (17) ensures that the effect due to the use of $\hat{\mu}(B_n)$ but not $\mu(B_n)$ is ignorable. Condition (18) is only slightly stronger than condition (14) in that it specifies a rate on how fast the covariance between $\tilde{V}(s_1, B_n)$ and $\tilde{V}(s_2, B_n)$ decays to zero.

Theorem 2—If conditions (9)–(14), (16)–(18) are true, then

$$E\{[\tilde{V}_3(B_n) - V_3(W_n)]^2\} = O(1/|B_n| + |B_n|/|W_n|).$$

From Theorem 2, we see that the “optimal” rate for the sub-blocks is $|B_n| = c(B_n)|W_n|^{1/2}$, where the word “optimal” means minimizing the mean squared error and $c(B_n)$ may depend on B_n . If $c(B_n) \approx c$ for large enough B_n , which is a reasonable assumption when the covariate process is stationary, then the data-driven method proposed by Hall and Jing (1996) can be used to estimate the value of c . In a more general setting, we follow the empirical guideline given by Sherman (1996) to make the subblocks as large as possible such that about 10–30 nonoverlapping subblocks can be obtained.

5. APPLICATIONS

5.1 Residual Diagnostic

Baddeley et al. (2005) proposed novel residual diagnostic procedures for model assessment and Baddeley, Møller, and Pakes (2008) studied the theoretical properties of these residuals. Zhuang (2006) developed their second-order analogues for spatio-temporal point processes. Both Baddeley et al. and Zhuang’s residuals are defined in terms of some conditional intensity of the process but not in terms of a marginal quantity such as the second-order intensity function.

In practice, it is often of interest to fit a parametric model for the second-order intensity function. Let $\hat{\lambda}_2(\cdot)$ be such an estimate. To assess the goodness of fit of a fitted model, we propose examining the residual

$$R(B) = \sum_{\mathbf{u}, \mathbf{v} \in (N \cap B)} \sum_{\lambda(\mathbf{u}; \beta), \lambda(\mathbf{v}; \beta)}^{\neq} \frac{f(\mathbf{u}, \mathbf{v})}{\lambda(\mathbf{u}; \beta)\lambda(\mathbf{v}; \beta)} - \int_{B^{\otimes 2}} \frac{f(\mathbf{u}, \mathbf{v})}{\lambda(\mathbf{u}; \beta)\lambda(\mathbf{v}; \beta)} \hat{\lambda}_2(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v}.$$

If $\hat{\lambda}_2(\cdot)$ is a good model, then we expect $R(B) \approx 0$ and not so otherwise. In the special case that $f(\mathbf{u}, \mathbf{v}) = I(\|\mathbf{u} - \mathbf{v}\| \leq t)$, we can also consider

$$R(B) = \hat{K}_{\text{inhom}}(t; B) - \hat{K}_{\text{inhom}}(t),$$

where $\hat{K}_{\text{inhom}}(t; B)$ is the empirical K -function calculated on the subset B . If the spatial point process is second-order intensity reweighted stationary, then we again expect $R(B) \approx 0$. To formally compare the residuals with zero, we would need to estimate the variance of the residuals. We assume that the variance of $R(B)$ is dominated by that of the first term. The proposed variance estimation method for the small B case can then be used for this purpose.

5.2 Parametric Estimation of the K -Function

It is often of interest to fit a parametric model for the K -function, $K(t; \theta)$ say, where θ is some unknown parameter. To estimate θ , the least squared criterion

$$D(\theta) = \int_0^r [\widehat{K}_{\text{inhom}}(t)^c - K(t; \theta)^c]^2 dt$$

is often minimized, where r and c are two predefined tuning parameters. Waagepetersen and Guan (2009) showed that the variance of the resulting estimator $\hat{\theta}$ is closely related to the variance of the statistic $T(B)$ with $B = W$ and

$$f(\mathbf{u}, \mathbf{v}) = 2c^2 \int_{\|\mathbf{v}-\mathbf{u}\|}^r K(t; \theta)^{2c-2} K^{(1)}(t; \theta) dt,$$

where $K^{(1)}(t; \theta)$ is the first derivative of $K(t; \theta)$ with respect to θ . The proposed method can thus be applied to estimate the variance of $\hat{\theta}$. Waagepetersen and Guan (2009) used parametric bootstrap which can be computationally very intensive.

6. A SIMULATION STUDY

6.1 Simulation Design

We apply the proposed variance estimation methods to data generated on 1×1 and 2×2 square windows from inhomogeneous Neyman–Scott processes. The intensity of the process is set to be $\log \lambda(\mathbf{s}) = \alpha + \gamma X(\mathbf{s})$ for $\gamma = 0.5, 0.75$. Here $X(\mathbf{s})$ is a known covariate process and is fixed in all simulations. We simulate on average about 500 events on the 1×1 window and 2000 events on the 2×2 window each time. The number of events being simulated is directly regulated by the choice of α , where $\alpha \approx 425,360$ is used in our simulation for $\gamma = 0.5, 0.75$, respectively.

To simulate the data, we first generate a homogeneous Poisson process as the parent process and then simulate a Poisson number of offspring for each parent, where the position of each offspring relative to its parent is a radially symmetric Gaussian random variable (Diggle 2003). Let ρ and σ denote the intensity of the parent process and the standard deviation of the Gaussian dispersal variable, respectively. We set $\rho = 100, 400$ and $\sigma = 0.02, 0.04$, respectively. Note that a smaller ρ and σ lead to stronger clustering. For each simulated offspring event, we keep the event with a probability equal to the intensity at the offspring location divided by the maximum intensity in the study region. It's worth noting that the maximum intensity coincides with the intensity of the offspring process prior to the thinning.

For each simulated realization, we first estimate the intensity parameters α and γ and then calculate the empirical K -function $\hat{K}_{\text{inhom}}(t)$ at lags 0.025, 0.05, 0.075, and 0.1 based on the estimated intensity function. Note that $\hat{K}_{\text{inhom}}(t)$ can be obtained both locally by using data on a subset of W and globally by using all data. The proposed variance estimation procedures in Sections 3 and 4 can then be applied to estimate the variance of the empirical K -function in each case. Due to space concern we focus only on the latter but the estimator for the former works quite well in our simulation.

To estimate the target variance in the global K -function case, we estimate $V_3(W)$ by $\hat{V}_3(B)$, where $\hat{V}_3(B)$ is as defined in (8) and B is 0.25×0.25 and 0.33×0.33 square subblocks for the 1×1 and 2×2 windows, respectively. Because a direct calculation of (8) is often difficult, we approximate $\hat{V}_3(B)$ by using both nonoverlapping and overlapping subblocks. To be specific, 16 nonoverlapping subblocks and 36 equally spaced overlapping subblocks are used for the 1×1 window; the numbers become 32 and 72 for the 2×2 window. For comparison, we also estimate the target variance by assuming that the process is Poisson. For ease of discussion, we will refer to the three estimators as the nonoverlapping, overlapping, and Poisson estimators, respectively. If an estimate from the nonoverlapping or overlapping estimators is smaller than

that from the Poisson estimator, we replace it with the latter. This rarely happens for the overlapping estimator but is more frequent for the nonoverlapping estimator.

6.2 Edge-Effect Correction

The empirical K -function given in (2) is slightly biased for the true K -function even if the true value of $\beta = (\alpha, \gamma)$ is known. To reduce the bias, we modify $\hat{K}_{\text{inhom}}(t)$ as

$$\tilde{K}_{\text{inhom}}(t) = \sum_{\mathbf{u}, \mathbf{v} \in (N \cap W)}^{\neq} \frac{I(\|\mathbf{u} - \mathbf{v}\| \leq t)}{\lambda(\mathbf{u}; \hat{\beta}) \lambda(\mathbf{v}; \hat{\beta}) |W \cap (W - \mathbf{u} + \mathbf{v})|},$$

where the term $|W \cap (W - \mathbf{u} + \mathbf{v})|$ is the common area between W and W shifted by $\mathbf{u} - \mathbf{v}$ (Møller and Waagepetersen 2004). For the variance of $\tilde{K}_{\text{inhom}}(t)$, we can analogously define $\tilde{V}_1(W)$, $\tilde{V}_2(W)$, and $\tilde{V}_3(W)$, corresponding to $V_1(W)$, $V_2(W)$, and $V_3(W)$ used in the variance of $\hat{K}_{\text{inhom}}(t)$. Variance estimators that are similar to $\tilde{V}_1(W)$ and $\tilde{V}_2(W)$ and account for the new edge-effect correction term can be easily developed for $\tilde{V}_1(W)$ and $\tilde{V}_2(W)$. Below we focus on estimating $\tilde{V}_3(W)$.

To proceed, we assume that $V_3(W) \approx \tilde{V}_3(W) \approx \eta$, where η is the limit of $V_3(W)$ as the size of W increases to infinity. This assumption is reasonable because both t and the dependence range here are small relative to the size of W . However, it is not reasonable to believe $V_3(B) \approx \eta$ for the subblocks B , because the values of t (e.g., $t = 0.1$) are quite large compared to the size of B . This in turn may result in severe edge effects. A similar phenomenon was observed and carefully discussed in Loh and Stein (2004) when estimating the variance of the empirical K -function in the stationary case. To correct for edge effects, we view the window as being wrapped on a torus so that the opposite boundaries of the square regions overlap. Define

$$\tilde{K}_{\text{inhom}}(t; B) = \frac{1}{|B|} \sum_{\mathbf{u} \in (N \cap B)}^{\neq} \sum_{\mathbf{v} \in (N \cap W)} \frac{I(\|\mathbf{u} - \mathbf{v}\| \leq t)}{\lambda(\mathbf{u}; \hat{\beta}) \lambda(\mathbf{v}; \hat{\beta})}.$$

We then estimate the term $V_3(B)$ associated with $\tilde{K}_{\text{inhom}}(t; B)$ in the above and use it as an estimate for $\tilde{V}_3(W)$.

6.3 Results

Table 1 shows the bias, interquartile range (IQR) and percentage of coverage at the 95% nominal level for $t = 0.1$. The results at other lags are similar and are thus omitted. The bias here is defined as the difference between the median and the target variance. Median and IQR are used because the proposed variances estimator can be highly skewed with relatively small sample sizes. From Table 1, we see that the bias is negative for all three estimators and in all cases, and that the magnitudes of the bias for the nonoverlapping and overlapping estimators are similar and are both much smaller than that of the Poisson estimator. When the window size is increased, both the bias and IQR for the proposed estimators decrease, which supports our consistency result. In almost all cases, the IQR for the overlapping estimator is smaller than that for the nonoverlapping estimator. This indicates the benefit of using overlapping blocks. Because of the large bias for the Poisson estimator, the resulting confidence intervals include the true K -function only 57.7%–83.6% of the times, which is in great contrast to 88.9%–95.2% and 90.1%–95.1% of the times for the nonoverlapping and overlapping estimators, respectively. The empirical confidence levels are especially close to the nominal level in the 2×2 window case.

The value of γ and ρ can have a large impact on the performance of the proposed method but the value of σ appears to affect it only mildly. Specifically, when γ is increased, that is, when the process becomes more inhomogeneous, both the bias and IQR become larger. Moreover, the resulting empirical confidence levels are further away from the nominal level. Thus, the performance of the proposed method becomes worse when the data are more inhomogeneous. When ρ is increased, the bias becomes smaller but the IQR remains similar. Thus, the variance estimator is more accurate. Reflected from the empirical confidence levels, better coverage is typically achieved when a larger ρ is used. Because a larger ρ indicates a weaker clustering, the proposed method is expected to perform better when the data are less clustered.

7. TROPICAL RAIN FOREST DATA

We apply our proposed variance estimation methods to a tropical forest dataset collected in a 1000 by 500 meter plot in the Barro Colorado Island, see Condit, Hubbell, and Foster (1996), Condit (1998), and Hubbell and Foster (1983). The top panel of Figure 1 shows the positions of alive trees of the species *Acalypha diversifolia* (526 trees) and *Capparis frondosa* (3299 trees), both recorded in a 1995 survey. We are interested in these two species because they are thought to have very distinct seed dispersal modes (Wright, Hernández, and Condit 2007). In particular, seeds are dispersed by exploding capsules for *Acalypha* and by birds and mammals for *Capparis*. Seidler and Plotkin (2006) hypothesized that the modes of seed dispersal can be reflected in the spatial distribution of tree locations with tight clusters for exploding capsules but loose clusters for bird and mammal dispersal. To investigate this, they fit homogeneous Neyman–Scott process models to the spatial locations of different species and used the dispersal parameter (i.e., σ used in the simulation) to quantify tightness of clustering. Waagepetersen and Guan (2009) argued that it's important to account for inhomogeneity due to the environment before assessing clustering. They instead fit inhomogeneous Neyman–Scott process models to the spatial locations of the species. Specifically, they used a log-linear intensity model with elevation and potassium content in the soil being the covariates.

The bottom panel of Figure 1 shows two typical realizations from the fitted inhomogeneous Neyman–Scott process models. These realizations look somehow visually different to the observed patterns, thus raising questions over the validity of the fitted models and of any subsequent analysis based on these models. Waagepetersen and Guan (2009) acknowledged an unsatisfactory fit for *Acalypha* on the basis of examining the J -function (van Lieshout and Baddeley 1996). In light of these facts, it seems necessary to compare the clustering strengths of the two spatial point patterns in a model-free way, for example, by comparing the empirical K -functions. To that end, we need to obtain the variance estimates associated with the empirical K -functions. The proposed variance estimation method in Section 4 can then be used for this purpose. Furthermore, we may wish to assess whether the clustering strength of each species, as reflected by the empirical K -functions, varies over region. If the empirical K -function remains roughly constant across the study region, then the assumption of second-order intensity reweighted stationarity seems reasonable. For this we will assess the difference between locally and globally obtained empirical K -functions. For a formal comparison we will need to estimate the variance of the local empirical K -function. The proposed variance estimation procedure in Section 3 can then be used for this purpose.

Figure 2 shows the empirical K -functions for *Acalypha* and *Capparis*, the associated 95% confidence bands and the theoretical K -function in the Poisson process case. From the plot, we see that both species are clustered because their lower confidence bands are larger than the Poisson K -function. Furthermore, the K -function for *Acalypha* is significantly larger than that of *Capparis*, suggesting that *Acalypha* is more clustered than *Capparis*. The difference in the

clustering strengths may be caused by the difference in the seed dispersal modes, as postulated by Seidler and Plotkin (2006). Our analysis provides support for this hypothesis.

Figure 3 shows plots of standardized differences (residuals) between the local and global empirical K -functions at 20 different lags. To obtain the local empirical K -functions, we divide the study region into 18 and 32 subregions for *Acalypha* and *Capparis*, respectively. Variance estimates for the local empirical K -functions are obtained by using the proposed variance estimation procedure in Section 3 and are subsequently used to create the standardized residuals. These plots reveal one highly clustered subregion (highlighted in Figure 1) for each species. The extreme clustering can be due to underestimation of either the intensity or the K -function in the subregion or both. For *Acalypha*, standardized residual plots suggest that the fitted log-linear model fits the data well. Thus, spatial locations of *Acalypha* in the identified subregion appear to be more clustered than in other regions. The biological cause behind is unclear to us. For *Capparis*, we have found that the fitted intensity function model severely underestimates the intensity in the identified subregion. This in turn leads to a strong clustering effect. We have also tried to include the quadratic terms of the covariates in the model. However, such a model does not seem to improve the fit in the given subregion. Other than in the two identified subregions, the standardized residuals are all reasonably small. Thus, the second-order intensity reweighted stationarity appears to be reasonable for both datasets. Note that we reach this conclusion without assuming any specific models for the pair correlation function.

8. DISCUSSION

We have developed two nonparametric variance estimation procedures for second-order statistics calculated from inhomogeneous spatial point processes. To apply the proposed procedures, we need to assume a parametric model for the intensity function of the spatial point process. Such an assumption is not overly restrictive, because the calculation of most commonly used second-order statistics (e.g., the empirical K -function) requires a (parametric) estimate of the intensity function. Nevertheless, it excludes the use of the proposed methods for a certain class of Gibbs point process models with a parametric first-order (i.e., trend) component (e.g., Stoyan and Stoyan 1998), because the intensity function of such processes often can not be written in closed parametric forms.

It is important to properly model the intensity function, as a misspecified model may result in misleading estimates for the K -function. To assess the goodness of fit of a fitted intensity model, we can generalize the residual diagnostic procedures proposed by Baddeley et al. (2005) to the current setting. In our real data analysis, we have also compared estimates of the intensity obtained from a fitted parametric model with those obtained from a consistent nonparametric estimator (e.g., Guan 2008), so as to evaluate the validity of the fitted parametric model. Our results (not included) show that the log-linear model with only linear terms seems to fit the data well for *Capparis*, but there seems to be a slight curvature trend for *Acalypha*. We only report the results from the simpler model here so as to be consistent with the analysis done by Waagepetersen and Guan (2009). We have nevertheless also fitted a log-linear model with quadratic terms for *Acalypha* but the main finding remains unchanged. To investigate the effect of using an incorrect model on the proposed method, we have conducted a small simulation study in a setting similar to our real data for *Acalypha*. These results suggest that the proposed variance estimation procedure still works well (see the supplemental materials), although the interpretation of the empirical K -function becomes awkward. In general, we believe that second-order analysis should be done only if the intensity function has been modeled satisfactorily, for example, as suggested by results obtained from the various diagnostic procedures.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Yongtao Guan's research was supported by NSF grants DMS-0706806 and DMS-0845368 and by NIH/NIDCR grant UL1 DE019586. The author thanks the Editor, an Associate Editor and two referees for their helpful comments that have greatly improved this manuscript.

REFERENCES

- Baddeley AJ, Møller J, Pakes AG. Properties of Residuals for Spatial Point Processes. *Annals of the Institute of Statistical Mathematics* 2008;60:627–649.
- Baddeley AJ, Møller J, Waagepetersen R. Non- and Semi-Parametric Estimation of Interaction in Inhomogeneous Point Patterns. *Statistica Neerlandica* 2000;54:329–350.
- Baddeley AJ, Turner R, Møller J, Hazelton M. Residual Analysis for Spatial Point Processes. *Journal of the Royal Statistical Society, Ser. B* 2005;67:617–666. (with discussion).
- Condit, R. Tropical Forest Census Plots. Berlin: Springer-Verlag and R. G. Landes Company; 1998.
- Condit R, Hubbell SP, Foster RB. Changes in Tree Species Abundance in a Neotropical Forest: Impact of Climate Change. *Journal of Tropical Ecology* 1996;12:231–256.
- Cressie, NAC. *Statistics for Spatial Data*. New York: Wiley; 1993.
- Crowder MJ. On Consistency and Inconsistency of Estimating Equations. *Econometric Theory* 1986;2:305–330.
- Daley, DJ.; Vere-Jones, D. *An Introduction to the Theory of Point Processes*. New York: Springer-Verlag; 1988.
- Diggle, PJ. *Statistical Analysis of Spatial Point Patterns*. New York: Oxford University Press; 2003.
- Diggle PJ, Gomez-Rubio V, Brown PE, Chetwynd AG, Gooding S. Second-Order Analysis of Inhomogeneous Spatial Point Processes Using Case-Control Data. *Biometrics* 2007;63:550–557. [PubMed: 17688507]
- Doukhan, P. *Mixing: Properties and Examples*. New York: Springer-Verlag; 1994.
- Ekström M, Luna SS. Subsampling Methods to Estimate the Variance of Sample Means Based on Nonstationary Spatial Data With Varying Expected Values. *Journal of the American Statistical Association* 2004;99:82–95.
- Guan Y. On Consistent Nonparametric Intensity Estimation for Inhomogeneous Spatial Point Processes. *Journal of the American Statistical Association* 2008;103:1238–1247.
- Guan Y, Loh JM. A Thinned Block Bootstrap Variance Estimation Procedure for Inhomogeneous Spatial Point Patterns. *Journal of the American Statistical Association* 2007;102:1377–1386.
- Guan Y, Sherman M. On Least Squares Fitting for Stationary Spatial Point Processes. *Journal of the Royal Statistical Society, Ser. B* 2007;69(1):31–49.
- Hall P, Jing B. On Sample Reuse Methods for Dependent Data. *Journal of the Royal Statistical Society, Ser. B* 1996;58:727–737.
- Heagerty PJ, Lumley T. Window Subsampling of Estimating Functions With Application to Regression Models. *Journal of the American Statistical Association* 2000;95:197–211.
- Heinrich L. Normal Convergence of Multidimensional Shot Noise and Rates of This Convergence. *Advances in Applied Probability* 1985;17:709–730.
- Heinrich L. Asymptotic Gaussianity of Some Estimators for Reduced Factorial Moment Measures and Product Densities of Stationary Poisson Cluster Processes. *Statistics* 1988;19:87–106.
- Hubbell, SP.; Foster, RB. Diversity of Canopy Trees in a Neotropical Forest and Implications for Conservation. In: Sutton, SL.; Whitmore, TC.; Chadwick, AC., editors. *Tropical Rain Forest: Ecology and Management*. Oxford: Blackwell Scientific Publications; 1983. p. 25-41.
- Jolivet, E. Central Limit Theorem and Convergence of Empirical Processes for Stationary Point Processes. In: Bartfai, P.; Tomko, J., editors. *Point Processes and Queuing Problems*. New York: Elsevier; 1981. p. 117-161.

- Loh JM, Stein M. Bootstrapping a Spatial Point Process. *Statistica Sinica* 2004;14:69–101.
- Møller, J.; Waagepetersen, RP. *Statistical Inference and Simulation for Spatial Point Processes*. Boca Raton: Chapman & Hall/CRC; 2004.
- Møller J, Syversveen AR, Waagepetersen RP. Log Gaussian Cox Processes. *Scandinavian Journal of Statistics* 1998;25:451–482.
- Politis, DN. *Subsampling*. New York: Springer-Verlag; 1999.
- Schoenberg FP. Consistent Parametric Estimation of the Intensity of a Spatial–Temporal Point Process. *Journal of Statistical Planning and Inference* 2005;128(1):79–93.
- Seidler TG, Plotkin JB. Seed Dispersal and Spatial Pattern in Tropical Trees. *PLoS Biology* 2006;4:2132–2137.
- Sherman M. Variance Estimation for Statistics Computed From Spatial Lattice Data. *Journal of the Royal Statistical Society, Ser. B* 1996;58:509–523.
- Sherman M. Subseries Methods in Regression. *Journal of the American Statistical Association* 1997;92:1041–1048.
- Stoyan D, Stoyan H. Non-Homogeneous Gibbs Process Models for Forestry—a Case Study. *Biometrical Journal* 1998;40:521–531.
- van der Vaart, AW. *Asymptotic Statistics*. New York: Cambridge University Press; 1998.
- van Lieshout MNM, Baddeley AJ. A Nonparametric Measure of Spatial Interaction in Point Patterns. *Statistica Neerlandica* 1996;50:344–361.
- Waagepetersen RP. An Estimating Function Approach to Inference for Inhomogeneous Neyman–Scott Processes. *Biometrics* 2007;62:252–258. [PubMed: 17447951]
- Waagepetersen RP, Guan Y. Two-Step Estimation for Inhomogeneous Spatial Point Processes. *Journal of the Royal Statistical Society, Ser. B* 2009;71:685–702.
- Wright SJ, Hernández A, Condit R. The Bushmeat Harvest Alters Seedling Banks by Favoring Lianas, Large Seeds, and Seeds Dispersed by Bats, Birds, and Wind. *Biotropica* 2007;39:363–371.
- Zhuang J. Second-Order Residual Analysis of Spatiotemporal Point Processes and Applications in Model Evaluation. *Journal of the Royal Statistical Society, Ser. B* 2006;68:635–653.

Figure 1. Locations of *Acalypha diversifolia* and *Capparis frondosa* trees. Top panel: observed locations; bottom panel: simulated locations. The subsquares are the outlying subregions revealed by the residual plots (Figure 3).

Figure 2. The empirical K -functions (solid lines) of a Poisson process (bottom), *Acalypha diversifolia* (top), and *Capparis frondosa* (middle). The dashed and dotted lines are the 95% confidence bands for *Acalypha diversifolia* and *Capparis frondosa*, respectively.

Figure 3. Standardized residual plots for *Acalypha diversifolia* (left) and *Capparis frondosa* (right).

Table 1

Bias, interquartile range (IQR), and coverage percent at the 95% nominal level for (1) the proposed estimator using nonoverlapping blocks, (2) using overlapping blocks, and (3) the Poisson plug-in estimator. Bias and IQR are both divided by the target variance parameter

γ	ρ	σ	W	Bias			IQR			Coverage		
				(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
0.5	100	0.02	1 × 1	-0.270	-0.305	-0.792	0.634	0.518	0.144	0.929	0.934	0.628
			2 × 2	-0.095	-0.121	-0.713	0.416	0.382	0.110	0.949	0.942	0.642
	0.04	0.04	1 × 1	-0.290	-0.299	-0.721	0.550	0.527	0.194	0.911	0.901	0.639
			2 × 2	-0.074	-0.096	-0.622	0.432	0.398	0.146	0.935	0.934	0.695
0.75	400	0.02	1 × 1	-0.092	-0.100	-0.541	0.627	0.503	0.295	0.935	0.940	0.735
			2 × 2	0.040	0.036	-0.381	0.418	0.370	0.216	0.947	0.951	0.808
	0.04	0.04	1 × 1	-0.104	-0.122	-0.484	0.607	0.485	0.330	0.927	0.943	0.787
			2 × 2	-0.043	-0.049	-0.374	0.369	0.334	0.230	0.952	0.949	0.828
	100	0.02	1 × 1	-0.489	-0.477	-0.828	0.617	0.623	0.152	0.914	0.916	0.577
			2 × 2	-0.181	-0.187	-0.728	0.584	0.562	0.172	0.937	0.939	0.668
	0.04	0.04	1 × 1	-0.420	-0.417	-0.731	0.621	0.593	0.256	0.889	0.901	0.669
			2 × 2	-0.225	-0.233	-0.646	0.485	0.471	0.214	0.932	0.927	0.718
	400	0.02	1 × 1	-0.354	-0.307	-0.597	0.621	0.586	0.369	0.916	0.922	0.753
			2 × 2	-0.176	-0.163	-0.484	0.466	0.440	0.268	0.945	0.946	0.795
	0.04	0.04	1 × 1	-0.381	-0.319	-0.551	0.593	0.554	0.396	0.913	0.918	0.796
			2 × 2	-0.122	-0.100	-0.360	0.434	0.425	0.322	0.946	0.948	0.836