An estimating function approach to inference for inhomogeneous Neyman-Scott processes

Waagepetersen, Rasmus Plenge

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

AALBORG UNIVERSITY

An estimating function approach to inference for inhomogeneous Neyman-Scott processes

by

Rasmus Plenge Waagepetersen

R-2005-30

September 2005

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: http://www.math.aau.dk

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

An estimating function approach to inference for inhomogeneous Neyman-Scott processes

Rasmus Plenge Waagepetersen
Institute of Mathematical Sciences, Aalborg University
Fredrik Bajersvej 7G, DK-9220 Aalborg
rw@math.aau.dk

Abstract

This paper is concerned with inference for a certain class of inhomogeneous Neyman-Scott point processes depending on spatial covariates. Regression parameter estimates obtained from a simple estimating function are shown to be asymptotically normal when the "mother" intensity for the Neyman-Scott process tends to infinity. Clustering parameter estimates are obtained using minimum contrast estimation based on the K-function. The approach is motivated and illustrated by applications to point pattern data from a tropical rain forest plot.

Keywords: asymptotic normality, clustering, estimating function, infill asymptotics, inhomogeneous point process, Neyman-Scott point process.

1 Introduction

This work is motivated by ecological studies of biodiversity in tropical rain forests. A question of particular interest is how the very high number of different tree species continue to coexist, see e.g. Burslem *et al.* (2001) and Hubbell (2001). One explanation is so-called niche assembly which hypothesizes that different species benefit from different habitats determined e.g. by topography or soil properties. In recent years huge amounts of data have been collected in tropical rain forest plots in order to investigate the niche assembly and other competing hypotheses (Losos and Leigh, 2004). The data sets consist of measurements of soil properties, digital terrain models, and individual locations of all trees growing in the plots.

A first attemp to study the niche assembly hypothesis might be to fit an inhomogeneous Poisson point process to the point pattern of a particular tree species where the intensity function might be log-linearly related to soil properties and topographical variables like elevation or gradient. However, the inhomogeneous Poisson point process assumes independent scattering of the trees. This is not realistic since the trees reproduce by seed dispersal. That is, in addition to large scale variation due to environmental variables, one may also expect clustering due to seed dispersal. The standard errors obtained assuming a Poisson point process then underestimate the uncertainty of the regression parameter estimates.

In this paper we model clustered point patterns of trees as realizations of certain inhomogeneous Neyman-Scott cluster point processes introduced in Section 2. Likelihood-based inference for such models can be carried out using Markov chain Monte Carlo methods, see Møller and Waagepetersen (2003) and Waagepetersen and Schweder (2005). However, the Markov chain Monte Carlo approach is computationally demanding and not yet amenable for routine analyses by non-specialists. We therefore in Section 3 consider another approach where estimates of the regression parameters are obtained from an estimating function given by the score of a Poisson likelihood function. This is similar to the approach in Schoenberg (2004) who considers consistent estimation of the intensity function of space-time point processes.

Given the number of "mother points", the clusters in the Neyman-Scott process provide *iid* random samples of the spatial covariates. Using this it is easy to demonstrate asymptotic normality of the score function under a kind of "infill" asymptotics where the intensity of mother points approaches infinity. Asymptotic normality of the regression parameters then follows from general results for estimating functions, see the Appendix. Asymptotics for inhomogeneous cluster processes seems to be a rather unexplored topic in statistics for spatial point processes. Heinrich (1992) and Guan (2005), for example, consider increasing domain asymptotics assuming stationarity.

The asymptotic variance depends on the Neyman-Scott clustering parameters which can be estimated using minimum-contrast methods, see Diggle (2003) or Møller and Waagepetersen (2003). Minimum-contrast estimates are in general not believed to be very efficient but may suffice in studies of the niche assembly hypothesis where the clustering parameters are essentially nuisance parameters.

The usefulness of the estimating function approach is demonstrated via applications and simulation studies in Sections 4 and 5.

2 Inhomogeneous Neyman-Scott cluster point processes

Let $S \subset \mathbb{R}^2$ denote the bounded plot where trees and environmental variables are observed. For $\xi \in \mathbb{R}^2$, $z_{1:p}(\xi)$ denotes the $1 \times p$, $p \geq 1$, vector of non-constant environmental variables. We assume that the point pattern of trees is a realization of a spatial point process $X \cap S$ where $X = X_{c \in C}$ is a superposition of clusters X_c of "offspring" associated with "mother" points c in a stationary Poisson point process of intensity $\kappa > 0$. Given C, the clusters X_c are independent Poisson processes with intensity functions

$$\lambda_c(\xi) = \alpha k(\xi - c; \omega) \exp(z_{1:p}(\xi)\beta_{1:p}^{\mathsf{T}}) \tag{1}$$

where $\alpha > 0$, $\beta_{1:p}$ is the $1 \times p$ vector of regression parameters, and k is a probability density depending on a parameter $\omega > 0$ determining the spread of offspring points around c. The parameter of main interest is the regression parameter $\beta_{1:p}$ while κ, α, and ω are regarded as nuisance parameters in this paper.

Assume that $\exp(z_{1:p}(\cdot)\beta^{\mathsf{T}})$ is bounded by some constant M. A cluster X_c may then be regarded as an independent thinning of a cluster Y_c with intensity function $M\alpha k(\cdot - c; \omega)$ where the spatially varying thinning probability is $\exp(z_{1:p}(\cdot)\beta^{\mathsf{T}})/M$. From this point of view, the environmental variables control the survival of the offspring in Y_c. The thinning perspective is moreover useful for simulation purposes: it is straightforward to simulate the homogeneous Neyman-Scott process $Y = \cup_{c \in C} Y_c$ and secondly apply thinning to obtain a realization of X. For simulation of $X \cap S$, $M = \max_{\xi \in S} \exp(z_{1:p}(\xi)\beta^{\mathsf{T}})$ suffices.

The intensity function of X is

$$\lambda(\xi) = \kappa\alpha \exp(z_{1:p}(\xi)\beta_{1:p}^{\mathsf{T}}) = \exp(z(\xi)\beta^{\mathsf{T}}) \tag{2}$$

where $z(\xi) = (1, z_{1:p}(\xi))$ and $\beta = (\beta_0, \beta_{1:p}) = (\log(\kappa\alpha), \beta_{1:p})$. The so-called inhomogeneous K-function (Baddeley *et al.*, 2000) for X coincides with the K-function for the stationary process Y (letting $\lambda_Y = \kappa M\alpha$ denote the constant intensity of Y, $\lambda_Y K(t)$ is the expected number of points within distance t from a typical point of Y).

Note that the cluster model is a tractable but crude model for clustering due to seed dispersal. The clustering in reality results from an iteration of mother-offspring events over several generations.

3 Parameter estimation

Intuitively one may expect to obtain a useful estimate of the parameter β using an estimating function based on the intensity function (2). We therefore consider

$$l(\beta) = \sum_{\xi \in X \cap S} z(\xi)\beta^{\mathsf{T}} - \int_S \exp(z(\xi)\beta^{\mathsf{T}})\mathrm{d}\xi$$

which simply corresponds to the log likelihood function for a Poisson process with intensity function (2). Our unbiased estimating function is the derivative

$$u(\beta) = \frac{\mathrm{d}}{\mathrm{d}\beta} l(\beta) = \sum_{\xi \in X \cap S} z(\xi) - \int_S z(\xi) \exp(z(\xi)\beta^{\mathsf{T}})\mathrm{d}\xi \tag{3}$$

with sensitivity

$$j(\beta) = -\frac{\mathrm{d}}{\mathrm{d}\beta^{\mathsf{T}}} u(\beta) = \int_S z(\xi)^{\mathsf{T}} z(\xi) \exp(z(\xi)\beta^{\mathsf{T}})\mathrm{d}\xi.$$

The estimating equation $u(\beta) = 0$ has a unique solution $\hat{\beta}$ which maximizes $l(\beta)$ if the sensitivity j is positive definite. This is the case provided there exists a region $A \subseteq S$ of positive area $|A| > 0$ so that $z(\xi)^{\mathsf{T}} z(\xi)$ is positive definite for $\xi \in A$. The object function $l(\beta)$ can easily be maximized using the procedure `ppm` in the `R` package `spatstat` (Baddeley and Turner, 2005). Positive definiteness of j is moreover sufficient to establish asymptotic normality of the estimate $\hat{\beta}_{1:p}$ of $\beta_{1:p}$, see Section 3.1.

An estimate (Baddeley *et al.*, 2000; Møller and Waagepetersen, 2003) of the K-function for X can be obtained using the `spatstat` procedure `Kinhom` substituting the intensity function (2) by the estimate $\exp(z(\cdot)\hat{\beta}^{\mathsf{T}})$. More specifically,

$$\hat{K}(t) = \sum_{\xi, \eta \in X \cap S} \frac{1[0 < \|\xi - \eta\| < t]}{\exp\left((z(\xi) + z(\eta))\hat{\beta}^{\mathsf{T}}\right)} e_{\xi,\eta} \tag{4}$$

where $e_{\xi,\eta}$ is an edge correction.

In applications one typically uses a kernel k for which the K-function has a closed form expression depending on κ and ω. Minimum contrast estimates $\hat{\kappa}$ and $\hat{\omega}$ are then obtained by minimizing

$$\int_0^a (\hat{K}(t)^{1/4} - K(t; \kappa, \omega)^{1/4})^2 \mathrm{d}t \tag{5}$$

with respect to (κ, ω) for some user specified value of a, see Diggle (2003). Finally $\hat{\alpha} = \exp(\hat{\beta}_0)/\hat{\kappa}$.

3.1 Approximate distribution of regression parameter estimates

Denote by κ^*, α^*, ω^*, and $\beta^*_{1:p}$ the unknown parameter values for which the data is assumed to be generated. Suppose for a moment that κ^* is known in which case we obtain the estimate $\hat{\beta}_0 - \log\kappa^*$ of $\log\alpha$. By Theorem 1 in the Appendix, for large κ^*, $(\hat{\beta}_0 - \log\kappa^*, \hat{\beta}_{1:p})$ is approximately normal with mean $(\log\alpha^*, \beta^*_{1:p})$ and covariance matrix $\Sigma^* = \Sigma(\kappa^*, \alpha^*, \omega^*, \beta^*_{1:p})$ where

$$\Sigma(\kappa, \alpha, \omega, \beta_{1:p}) = (\kappa\alpha J(\beta_{1:p}))^{-1} + J^{-1}(\beta_{1:p})G(\beta_{1:p}, \omega)J^{-1}(\beta_{1:p})/\kappa, \tag{6}$$

$$J(\beta_{1:p}) = \int_S z(\xi)^\mathsf{T} z(\xi)\exp(z_{1:p}(\xi)\beta^\mathsf{T}_{1:p})\mathrm{d}\xi,$$

$$G(\beta_{1:p}, \omega) = \int_{\mathbb{R}^2} H(\beta_{1:p}, \omega, c)^\mathsf{T} H(\beta_{1:p}, \omega, c)\mathrm{d}c,$$

and

$$H(\beta_{1:p}, \omega, c) = \int_S z(\xi)\exp(z_{1:p}(\xi)\beta^\mathsf{T}_{1:p})k(\xi - c; \omega)\mathrm{d}\xi.$$

In practice we estimate the variance of $\hat{\beta}_{1:p}$ using a plug-in approach where the unknown parameters in Σ^* are replaced by their estimates. Letting $\hat{\mathrm{sd}}_j$ denote the square root of the jth diagonal element of $\hat{\Sigma} = \Sigma(\hat{\kappa}, \hat{\alpha}, \hat{\omega}, \hat{\beta}_{1:p})$, $[\hat{\beta}_j - 1.96\hat{\mathrm{sd}}_j, \hat{\beta}_j + 1.96\hat{\mathrm{sd}}_j]$ is an approximate 95 % confidence interval for β_j, $j = 1, \ldots, p$. Our asymptotic approach where κ tends to infinity does not justify the plug-in approach. We therefore assess the usefulness of standard errors and approximate confidence intervals obtained from $\hat{\Sigma}$ via simulation studies in Section 5.

Note that the first term in the right hand side of (6) is the asymptotic covariance matrix for the maximum likelihood estimate of $(\log\alpha, \beta_{1:p})$ when the data is generated under a Poisson process with intensity function (2), cf. Remark 1 in the Appendix.

The integrals J, G, and H are evaluated using Riemann sums where $k(\xi - c; \omega)$ is approximated by $1[\xi \in D_c]k(\xi - c; \omega)$ for a disc D_c around c.

4 Application to rainforest data

The tropical tree data sets considered in this section are extracted from a huge data set collected in the 500 by 1000 meter Barro Colorado Island plot, see Condit *et al.* (1996); Condit (1998); Hubbell and Foster (1983), and the Acknowledgment. The upper plots in Figure 1 show respectively all tree posititions in 1995 of the species *Beilschmiedia pendula Lauraceae*

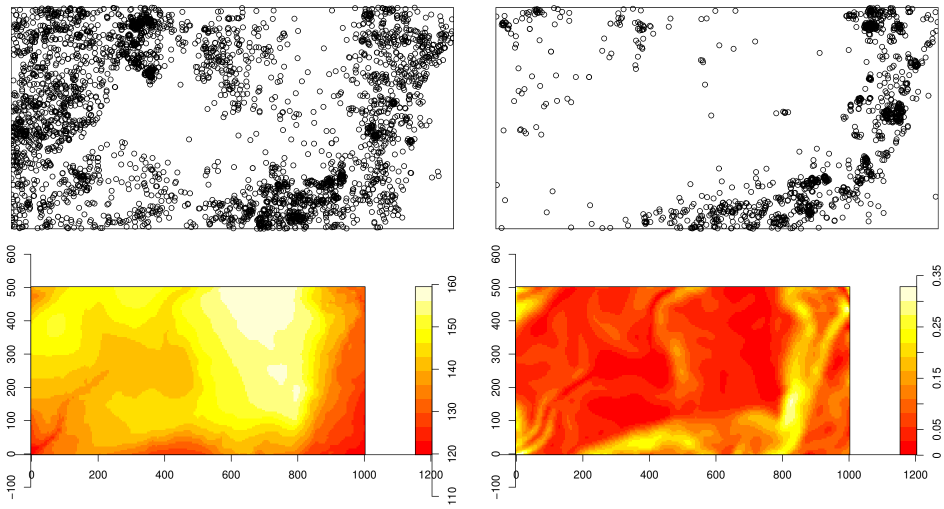

Figure 1: Upper plots: locations of *Beilschmiedia pendula Lauraceae* (left) and *Ocotea whitei Lauraceae* (right) trees. Lower plots: altitude (left) and norm of altitude gradient (right).

(3605 trees) and *Ocotea whitei Lauraceae* (1298 trees). The lower plots show covariates (altitude and norm of the altitude gradient) recorded on a 5 by 5 meter grid.

For both species, we let $z_{1:p}$ consist of the altitude and gradient covariates and k is assumed to be a bivariate isotropic normal density with standard deviation ω. The K-function is then

$$K(t; \kappa, \omega) = \pi t^2 + \left(1 - \exp(-t^2/(4\omega)^2)\right)/\kappa \tag{7}$$

and X can be viewed as an inhomogeneous version of the so-called Thomas process. The upper limit a in (5) is chosen to be 100 meter for both species (Diggle, 2003, recommends that a should be considerably smaller than the dimensions of the observation plot). We use 5 by 5 meter cells for the discretization in the Riemann approximation of the integrals in J, G, and H, and use four times the estimated ω for the radius in D_c, see Section 3.1.

The estimates of β_1 (altitude) and β_2 (gradient) and associated approximate 95% confidence intervals are 0.02 (-0.02;0.06) and 5.84 (0.89;10.80) for Beilschmiedia and 0.01 (-0.04;0.06) and 14.87 (8.70;21.03) for Ocotea. Hence, there is evidence that both species prefer to live on slopes but not that they favour low or high altitudes. The estimates of (κ, α, ω) are (8e-5,85.9,20.0) for Beilschmiedia and (1.2e-4,13.5,12.4) for Ocotea (i.e. the expected numbers of motherpoints within the plot are 40 and 60, respectively).

Figure 2 shows for both species $\hat{K}(t)$ given by (4), $K(t, \hat{\kappa}, \hat{\omega})$, and the

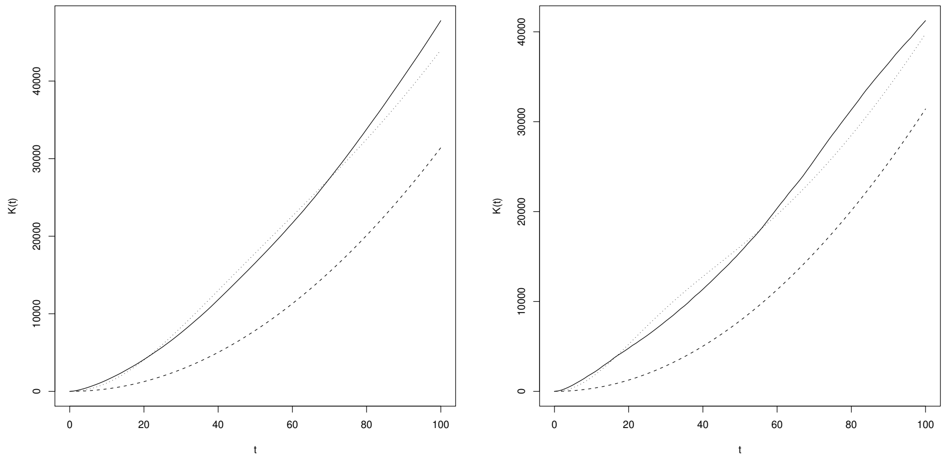

Figure 2: Solid lines: $\hat{K}(t)$ (4) for *Beilschmiedia pendula Lauraceae* (left) and *Ocotea whitei Lauraceae* (right). Dotted lines: $K(t, \hat{\kappa}, \hat{\omega})$ (7). Dashed lines: K-function $K_{\text{pois}}(t) = \pi t^2$ for a Poisson process.

K-function $K_{\text{pois}}(t) = \pi t^2$ for the Poisson process. The plots indicate clustering since the estimates $\hat{K}(t)$ are above $K_{\text{pois}}(t)$. Applying maximum likelihood estimation under the Poisson process assumption, we obtain the same estimates of β_1 and β_2 for the two types of trees but much too narrow confidence intervals (0.02;0.03) and (5.34;6.34) (Beilschmiedia) and (0.00;0.02) and (14.15;15.58) (Ocotea).

5 Simulation study

In the following simulation study we focus on the asymptotic normality of $\hat{\beta}_{1:p}$, the performance of the standard errors for $\hat{\beta}_{1:p}$ obtained from either Σ^* or $\hat{\Sigma}$, and the coverage properties of approximate confidence intervals, see Section 3.1.

We use the observation plot, covariates, and kernel k from the previous section, fix $\beta^*_{1:p}$ at the parameter estimates obtained for the Beilschmiedia trees and let ω^* equal to 10 or 20. The parameter κ^* is 5e-5, 1e-4, or 5e-4 corresponding to either 25, 50, or 250 expected numbers of motherpoints within the plot. For each value of κ^* we consider two values of α^* so that the expected number μ^* of simulated points is either 200 or 800 corresponding to "small" and "moderately large" point patterns. For each combination of κ^* and μ^* we generate 1000 synthetic data sets and obtain simulated parameter estimates by applying our estimation procedure to the synthetic data. The results obtained with the two ω^* values are qualitatively very similar, so below we only comment on the results for $\omega^* = 20$.

The qq-plots in Figure 3 and Figure 4 based on the simulated values of $\hat{\beta}_{1:p}$ indicate that the distribution of $\hat{\beta}_1$ is fairly close to normal already for κ^*=5e-5 while the convergence to normality is slower for $\hat{\beta}_2$ where the qq-plots reveal a heavy tail to the left for the smaller κ^* values. The different rates of convergence are probably due to the difference between the associated covariates. High values of the gradient covariate occur in rather narrow areas which are less likely to be sampled by a cluster of points. This induces a bias downwards for the estimates of β_2: for κ^*=5e-5 the mean of $\hat{\beta}_2$ is about 0.6 smaller than $\beta_2^* = 5.84$. For κ^*=5e-4 the bias is reduced to around 0.1. The estimate of β_2 is essentially unbiased for all values of κ^*.

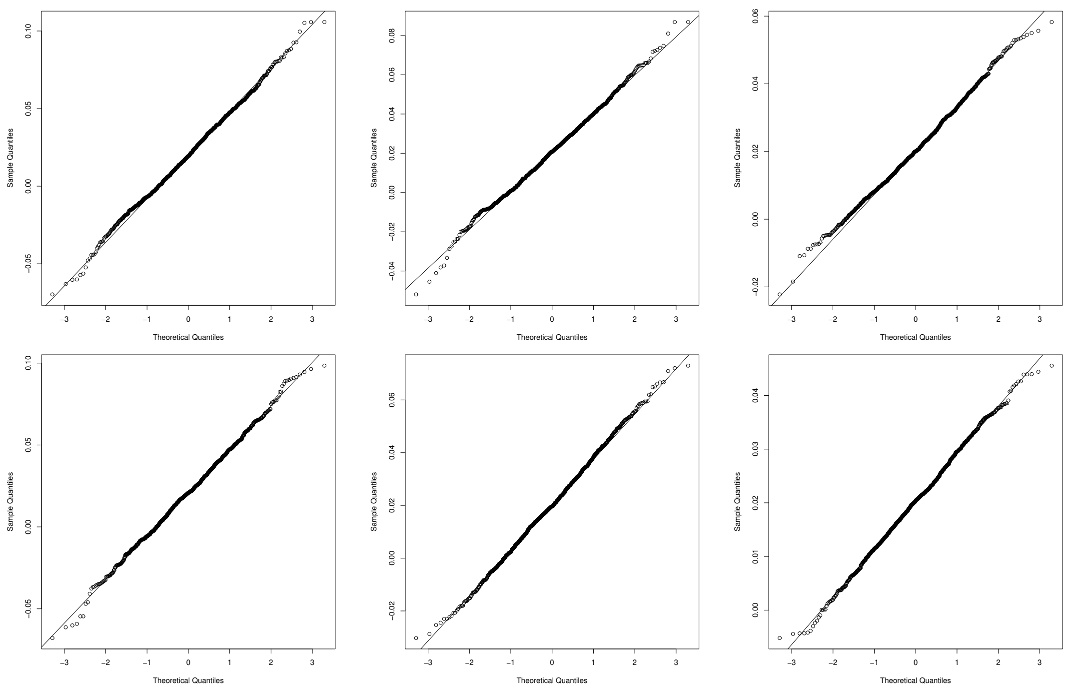

Figure 3: Quantiles obtained from 1000 simulated parameter estimates of β_1 against quantiles of a normal distribution. Left to right κ^*=5e-5,1e-4,5e-4 and top to bottom $\mu^* = 200, 800$.

The first column in Table 1 shows for each combination of κ^* and μ^*, a Monte Carlo estimate of the standard deviation for $\hat{\beta}_1$ obtained from the 1000 simulated parameter estimates. The second column contains the standard deviations obtained from Σ^* while Monte Carlo estimates of the medians of the standard deviations obtained from $\hat{\Sigma}$ are given in the third column. The

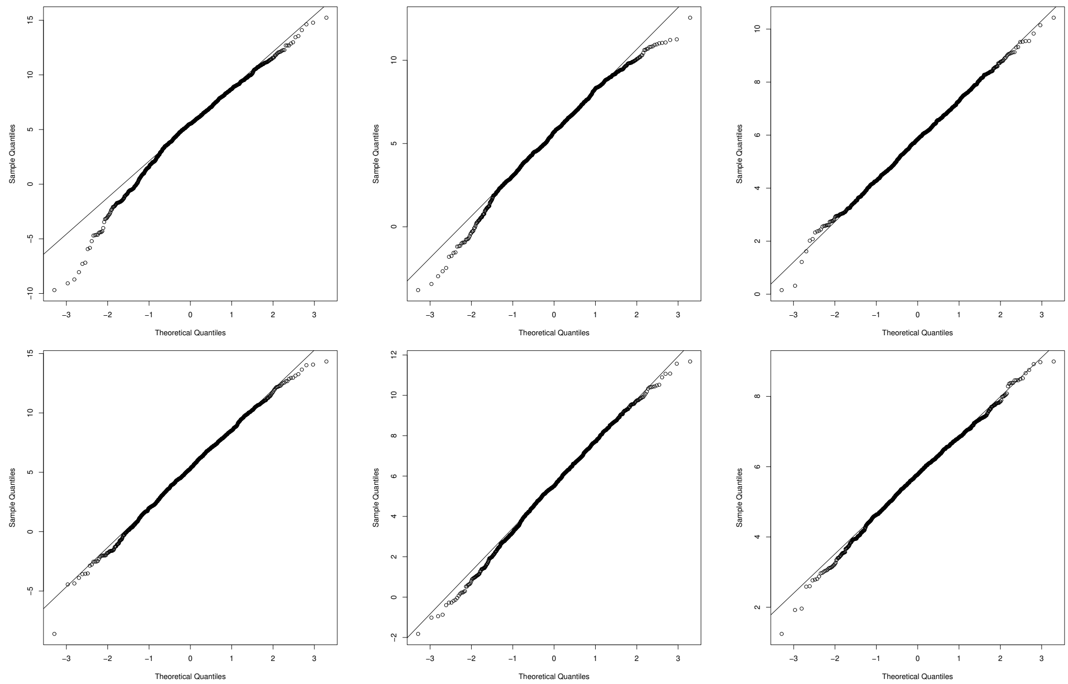

Figure 4: Quantiles obtained from 1000 simulated parameter estimates of β_2 against quantiles of a normal distribution. Left to right κ^*=5e-5,1e-4,5e-4 and top to bottom $\mu^* = 200, 800$.

fourth column contains the estimated coverage percentages for the approximate 95 % confidence intervals for β_1. The last four columns are as the four first but for β_2.

The estimated coverage percentages in general differ less from the nominal 95% than twice the Monte Carlo standard error which is around 0.007. The approximate confidence intervals seem to be slightly too conservative for β_1 and slightly too restrictive for β_2. There is in general good agreement between the first three columns regarding β_1. Larger discrepancies can be observed between the columns 5 to 7. In particular, the standard errors obtained from $\hat{\Sigma}$ (column 7) seem to underestimate somewhat the sampling standard deviation of $\hat{\beta}_2$ (column 5). Perhaps the underestimation of the standard deviation is counterbalanced by the bias of $\hat{\beta}_2$ so that reasonable coverage percentages are still obtained.

Considering the generally decent coverage properties of the approximate confidence intervals, basing inference on standard errors obtained from $\hat{\Sigma}$ does not seem unreasonable even when the expected number of mother points in the observation plot is as low as 25. However, for covariates with peaks and

κ^*	μ^*	sd_1	sd_1^*	$\hat{\text{sd}}_1$	cvrg_1	sd_2	sd_2^*	$\hat{\text{sd}}_2$	cvrg_2
5e-5	200	0.027	0.027	0.027	0.96	3.67	3.35	3.34	0.94
	800	0.026	0.026	0.025	0.97	3.36	3.21	3.16	0.94
1e-4	200	0.020	0.020	0.020	0.96	2.57	2.49	2.41	0.94
	800	0.018	0.018	0.018	0.96	2.23	2.3	2.23	0.95
5e-4	200	0.013	0.013	0.012	0.94	1.50	1.48	1.42	0.94
	800	0.009	0.009	0.009	0.95	1.14	1.14	1.10	0.94

Table 1: First four columns: Monte Carlo estimate of the standard deviation for $\hat{\beta}_1$, standard deviation obtained from Σ^*, median of standard deviation obtained from $\hat{\Sigma}$, and coverage of approximate confidence interval. Last four columns: as first four columns but for β_2.

narrow ridges, one should be careful with possible bias of the estimates of the associated parameter and standard error.

6 Estimating functions based on second order properties

Our estimating function (3) does not provide estimates of the clustering parameters. Inspired by Guan (2005) one might therefore consider

$$l_2(\beta_{1:p}, \kappa, \alpha, \omega) = \sum_{\substack{\xi,\eta \in X \cap S:\\ \xi \neq \eta}} \log \lambda^{(2)}(\xi, \eta; \beta_{1:p}, \kappa, \alpha, \omega) - \int_S \int_S \lambda^{(2)}(\xi, \eta; \beta_{1:p}, \kappa, \alpha, \omega) \mathrm{d}\xi \mathrm{d}\eta$$

where

$$\lambda^{(2)}(\xi, \eta; \beta_{1:p}, \kappa, \alpha, \omega) = \exp(z(\xi)\beta^{\mathsf{T}}) \exp(z(\eta)\beta^{\mathsf{T}}) g(\|\xi - \eta\|; \kappa, \omega)$$

is the second order product density and the pair correlation function $g(t; \kappa, \omega)$ is the derivative of the K-function divided by $2\pi t$. The function l_2 may be viewed as a limit of log composite likelihoods

$$\sum_{i \neq j} \big(1[n_i > 0 \text{ and } n_j > 0] \log P(n_i > 0 \text{ and } n_j > 0) + 1[n_i = 0 \text{ or } n_j = 0] \log P(n_i = 0 \text{ or } n_j = 0) \big)$$

where n_i is the number of points in $X \cap A_i$ for a disjoint partitioning $\{A_i\}$ of S and the sizes of the A_i tend to zero. In the stationary case, Guan (2005) considers an object function obtained by replacing the last term in l_2 by $\sum_{\substack{\xi,\eta\in X\cap S\\ \xi\neq\eta}} \log \int_S \int_S \lambda^{(2)}(\xi,\eta;\beta_{1:p},\kappa,\alpha,\omega)\mathrm{d}\xi\mathrm{d}\eta$.

Differentiating l_2 an unbiased estimating function u_2 is obtained. Disadvantages of u_2 compared to (3) are that standard software is not applicable and that the asymptotics are more complicated.

Acknowledgments This work resulted from a joint NERC Centre for Population Biology and UK Population Biology Network working group 'Spatial analysis of tropical forest biodiversity' which is funded by the Natural Environment Research Council and English Nature. The data sets used in the paper were collected in the forest dynamics plot of Barro Colorado Island which has been made possible through the generous support of the U.S. National Science Foundation, The John D. and Catherine T. MacArthur Foundation, and the Smithsonian Tropical Research Institute and through the hard work of over 100 people from 10 countries over the past two decades. The Barro Colorado Island Forest Dynamics Plot is part of the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

References

Baddeley, A. J. & Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**, 1–42.

Baddeley, A. J., Møller, J. & Waagepetersen, R. (2000). Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica* **54**, 329–350.

Burslem, D. F. R. P., Garwood, N. C. & Thomas, S. C. (2001). Tropical forest diversity: the plot thickens. *Science* **291**, 606–607.

Condit, R. (1998). *Tropical Forest Census Plots*. Springer-Verlag and R. G. Landes Company, Berlin, Germany and Georgetown, Texas.

Condit, R., Hubbell, S. P. & Foster, R. B. (1996). Changes in tree species abundance in a neotropical forest: impact of climate change. *Journal of Tropical Ecology* **12**, 231–256.

Diggle, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*. Oxford University Press, 2nd edition.

Guan, Y. (2005). A composite likelihood approach in fitting spatial point process models. Preprint.

Heinrich, L. (1992). Minimum contrast estimates for parameters of spatial ergodic point processes. In: *Transactions of the 11th Prague Conference on Random Processes, Information Theory and Statistical Decision Functions*, Academic Publishing House, Prague, 479–492.

Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Number 32 in Monographs in Population Biology, Princeton University Press.

Hubbell, S. P. & Foster, R. B. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In: *Tropical Rain Forest: Ecology and Management* (eds. S. L. Sutton, T. C. Whitmore and A. C. Chadwick), Blackwell Scientific Publications, 25–41.

Losos, E. C. & Leigh, Jr., E. G., eds. (2004). *Tropical Forest Diversity and Dynanism. Findings from a Large-Scale Plot Network*. The University of Chicago Press, Chicago.

Møller, J. & Waagepetersen, R. P. (2003). *Statistical inference and simulation for spatial point processes*. Chapman and Hall/CRC, Boca Raton.

Schoenberg, F. P. (2004). Consistent parametric estimation of the intensity of a spatial-temporal point process. *Journal of Statistical Planning and Inference* **128**, 79–93.

Sørensen, M. (1999). On asymptotics of estimating functions. *Brazilian Journal of Probability and Statistics* **13**, 111–136.

Waagepetersen, R. & Schweder, T. (2005). Likelihood-based inference for clustered line transect data. (submitted).

A Asymptotic normality of regression parameter estimates

Please recall the notation introduced in Section 2 and 3. In this appendix we derive asymptotic normality of the estimate of the interest parameter $\beta_{1:p}$ when the mother intensity tends to infinity, i.e. we consider an increasing sequence $(\kappa_n)_{n\geq 1}$ of κ values where $\kappa_n = n\tilde{\kappa}$ for some $\tilde{\kappa} > 0$ and $n \to \infty$. The constant $\tilde{\kappa}$ is introduced to allow for non-integer values of κ.

Let $\tilde{\beta}_0 = \log(\alpha)$ and let $u_n(\tilde{\beta}_0, \beta_{1:p}) = u(\log(\kappa_n) + \tilde{\beta}_0, \beta_{1:p})$ be the estimating function for $(\tilde{\beta}_0, \beta_{1:p})$ when κ is known and given by κ_n. Denote by

$(\tilde{\beta}_0^n, \hat{\beta}_{1:p}^n)$ the estimate obtained by solving $u_n(\tilde{\beta}_0, \beta_{1:p}) = 0$. The following Theorem 1 is concerned with asymptotic normality of $\sqrt{\kappa_n}(\tilde{\beta}_0^n - \beta_0^*, \hat{\beta}_{1:p}^n - \beta_{1:p}^*)$.

Theorem 1. *Suppose $J(\beta_{1:p}^*)$ is positive definite. Then $\sqrt{\kappa_n}(\tilde{\beta}_0^n - \beta_0^*, \hat{\beta}_{1:p}^n - \beta_{1:p}^*)$ is asymptotically zero mean normal with covariance matrix*

$$(\alpha^* J(\beta_{1:p}^*))^{-1} + J^{-1}(\beta_{1:p}^*) G(\beta_{1:p}^*, \omega^*) J^{-1}(\beta_{1:p}^*). \tag{8}$$

Proof. Below we show that $u_n(\tilde{\beta}_0^*, \beta_{1:p}^*)/\sqrt{n}$ is asymptotically normal. Asymptotic normality of $\sqrt{\kappa_n}(\tilde{\beta}_0^n - \beta_0^*, \hat{\beta}_{1:p}^n - \beta_{1:p}^*)$ then follows directly from Theorem 2.8 in Sørensen (1999).

Let $S_n = [-g(n), g(n)]^2$ where $g(\cdot)$ is an increasing function chosen so that

$$\lim_{n\to\infty} \sqrt{n} \int_S \int_{\mathbb{R}^2 \setminus S_n} k(\xi - c; \omega^*) \mathrm{d}c \mathrm{d}\xi = 0. \tag{9}$$

The increasing regions S_n are introduced to handle cases where $k(\cdot; \omega^*)$ is a probability density with unbounded support. Note that the bounded observation plot S is contained in S_n for sufficiently large n.

We now identify $C \cap S_n$ with $\cup_{i=1}^n \{C_{i1}^n, \ldots, C_{iN_i^n}^n\}$ where for each n, the N_i^n are independent Poisson variables with mean $\tilde{\kappa}|S_n|$ and the C_{ij}^n are *iid* uniform on S_n given the N_i^n. For each C_{ij}^n let $Z_{C_{ij}^n} = \sum_{\xi \in X_{C_{ij}^n} \cap S} z(\xi)$ where given C_{ij}^n, $X_{C_{ij}^n}$ is a Poisson process of intensity $\lambda_{C_{ij}^n}$, cf. (1). Similarly, to any point c in $C \setminus S_n$ we associate $Z_c = \sum_{\xi \in X_c \cap S} z(\xi)$.

By the Slivnyak-Mecke theorem (see e.g. Theorem 3.1 in Møller and Waagepetersen, 2003),

$$\mathbb{E}[Z_{C_{ij}^n}] = \frac{\alpha^*}{|S_n|} \int_S z(\xi) \exp(z_{1:p}(\xi)(\beta_{1:p}^*)^\mathsf{T}) \int_{S_n} k(\xi - c; \omega^*) \mathrm{d}c \mathrm{d}\xi.$$

Then $u_n(\tilde{\beta}_0^*, \beta_{1:p}^*)/\sqrt{n}$ is distributed as $U_n/\sqrt{n} + Z_n/\sqrt{n}$ where

$$U_n = \sum_{c \in C \cap S_n} Z_c - n\tilde{\kappa}|S_n| \mathbb{E} Z_{C_{11}^n} = \sum_{i=1}^n \Big(\sum_{j=1}^{N_i^n} Z_{C_{ij}} - \tilde{\kappa}|S_n| \mathbb{E} Z_{C_{11}^n}\Big)$$

and

$$Z_n = \sum_{c \in C \setminus S_n} Z_c - n\tilde{\kappa}\alpha^* \int_S z(\xi) \exp(z_{1:p}(\xi)(\beta_{1:p}^*)^\mathsf{T}) \int_{\mathbb{R}^2 \setminus S_n} k(\xi - c; \omega^*) \mathrm{d}c \mathrm{d}\xi.$$

The term $Z_n/\sqrt{n}$ converges to zero in probability since

$$\mathbb{E}\|Z_n/\sqrt{n}\| \le 2\sqrt{n}\tilde{\kappa}\alpha^* \int_{\mathbb{R}^2 \setminus S_n} k(\xi - c; \omega^*) \int_S \|z(\xi)\| \exp(z_{1:p}(\xi)(\beta_{1:p}^*)^\mathsf{T}) \mathrm{d}\xi \mathrm{d}c$$

where the right hand side tends to zero due to (9).

Regarding $U_n/\sqrt{n}$, note that $\mathbb{E}\sum_{j=1}^{N_i^n} Z_{C_{ij}^n} = \tilde{\kappa}|S_n|\mathbb{E}Z_{C_{11}^n}$ and $\mathbb{V}\mathrm{ar}\sum_{j=1}^{N_i^n} Z_{C_{ij}^n} = \tilde{\kappa}|S_n|\mathbb{E}(Z_{C_{11}^n})^2$ where by the extended Slivnyak-Mecke theorem (see e.g. Theorem 3.2 in Møller and Waagepetersen, 2003),

$$\begin{aligned}\mathbb{E}[(Z_{C_{11}^n})^2] =& \frac{\alpha^*}{|S_n|}\int_S z(\xi)^\mathsf{T} z(\xi)\exp(z_{1:p}(\xi)(\beta_{1:p}^*)^\mathsf{T})\int_{S_n} k(\xi-c;\omega^*)\mathrm{d}c\mathrm{d}\xi \\ &+\frac{(\alpha^*)^2}{|S_n|}\int_{S_n} H(\beta_{1:p}^*,\omega^*,c)^\mathsf{T} H(\beta_{1:p}^*,\omega^*,c)\mathrm{d}c.\end{aligned}$$

Thus, $\mathbb{V}\mathrm{ar}\sum_{j=1}^{N_i^n} Z_{C_{ij}^n}$ converges to $V=\tilde{\kappa}\alpha^* J(\beta_{1:p}^*)+\tilde{\kappa}(\alpha^*)^2 G(\beta_{1:p}^*,\omega^*)$. By the multivariate central limit theorem, $U_n/\sqrt{n}$ and hence $u_n(\tilde{\beta}_0^*,\beta_{1:p}^*)/\sqrt{n}$ converges to a multivariate normal distribution with mean zero and covariance matrix V.

It follows from Theorem 2.8 in Sørensen (1999) (under condition 2.1 and 2.4 with $G_n(\bar{\theta})=u_n(\tilde{\beta}_0^*,\beta_{1:p}^*)$ and $W(\bar{\theta})=\tilde{\kappa}\alpha^* J(\beta_{1:p}^*)$) that $\sqrt{\tilde{\kappa}n}(\tilde{\beta}_0^n-\beta_{1:p}^*,\hat{\beta}_{1:p}^n-\beta_{1:p}^*)$ is asymptotically zero-mean normal with covariance matrix (8). □

Remark 1. Suppose we consider a sequence of Poisson processes on S with intensity functions $\kappa_n\alpha\exp(z_{1:p}(\xi)\beta_{1:p}^\mathsf{T})$ and obtain estimates of $(\log\alpha,\beta_{1:p})$ by solving $u(\log(\kappa_n)+\log(\alpha),\beta_{1:p})=0$. Then, by the same type of argument as in the proof of Theorem 1, the asymptotic covariance is given by the first term in (8).