

# 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 the so called niche assembly hypothesis 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 attempt 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 which are introduced in Section 2. Likelihood-based inference for such models can in principle be carried out using Markov chain Monte Carlo (MCMC) methods, see Møller and Waagepetersen (2003) and Waagepetersen and Schweder (2005). However, it is not straightforward to implement the MCMC approach for the inhomogeneous Neyman-Scott process, see Section 6 for a discussion of computational problems. 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. The estimating function approach is likely to be statistically less efficient than likelihood-based inference but is computationally very simple and fast.

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 Waagepetersen (2006). Asymptotics for inhomogeneous cluster processes seems to be a rather unexplored topic in statistics for spatial point processes. Heinrich (1992) and Guan (2006), 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 Stoyan (1992), 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 is demonstrated via applications and simulation studies in Sections 4 and 5. In Section 6 the estimating function is discussed in relation to maximum likelihood estimation and a second order estimating function.

## 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

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

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  $\kappa$ ,  $\alpha$ , and  $\omega$  are regarded as nuisance parameters in this paper.

Assume that  $\exp(z_{1:p}(\cdot) \beta_{1:p}^\top)$  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  $\exp(z_{1:p}(\cdot) \beta_{1:p}^\top)/M$ . 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  $X \cap S$ ,  $M = \max_{\xi \in S} \exp(z_{1:p}(\xi) \beta_{1:p}^\top)$  suffices.

The intensity function of  $X$  is

$$\lambda(\xi) = \kappa \alpha \exp(z_{1:p}(\xi) \beta_{1:p}^\top) = \exp(z(\xi) \beta^\top), \xi \in \mathbb{R}^2, \quad (2)$$

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  $\beta$  using an estimating function based on the intensity function (2). We therefore consider

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

which simply corresponds to the log likelihood function under the assumption that  $X$  is a Poisson process with intensity function (2). Our unbiased estimating function is the derivative

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

with sensitivity

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

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  $z(\xi)^\top z(\xi)$  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  $\exp(z(\cdot)\hat{\beta}^\top)$ . More specifically,

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

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  $\kappa$  and  $\omega$ . Minimum contrast estimates  $\hat{\kappa}$  and  $\hat{\omega}$  are then obtained by minimizing

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

with respect to  $(\kappa, \omega)$  for some user specified value of  $a$ . The choice of  $a$  introduces a certain level of arbitrariness in the estimation procedure, see Diggle (2003) who recommends that  $a$  should be considerably smaller than the dimensions of the observation plot. Finally  $\hat{\alpha} = \exp(\hat{\beta}_0)/\hat{\kappa}$ .

### 3.1 Approximate distribution of regression parameter estimates

Denote by  $\kappa^*$ ,  $\alpha^*$ ,  $\omega^*$ , and  $\beta_{1:p}^*$  the unknown parameter values for which the data is assumed to be generated. Suppose for a moment that  $\kappa^*$  is known in which case we obtain the estimate  $\hat{\beta}_0 - \log \kappa^*$  of  $\log \alpha$ . By Theorem 1 in Waagepetersen (2006), for large  $\kappa^*$ ,  $(\hat{\beta}_0 - \log \kappa^*, \hat{\beta}_{1:p})$  is approximately normal with mean  $(\log \alpha^*, \beta_{1:p}^*)$  and covariance matrix  $\Sigma^* = \Sigma(\kappa^*, \alpha^*, \omega^*, \beta_{1:p}^*)$  where

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

$$J(\beta_{1:p}) = \int_S z(\xi)^\top z(\xi) \exp(z_{1:p}(\xi) \beta_{1:p}^\top) d\xi,$$

$$G(\beta_{1:p}, \omega) = \int_{\mathbb{R}^2} H(\beta_{1:p}, \omega, c)^\top H(\beta_{1:p}, \omega, c) dc,$$

and

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

In practice we estimate the variance of  $\hat{\beta}_{1:p}$  using a plug-in approach where the unknown parameters in  $\Sigma^*$  are replaced by their estimates. Letting  $\hat{\text{sd}}_j$  denote the square root of the

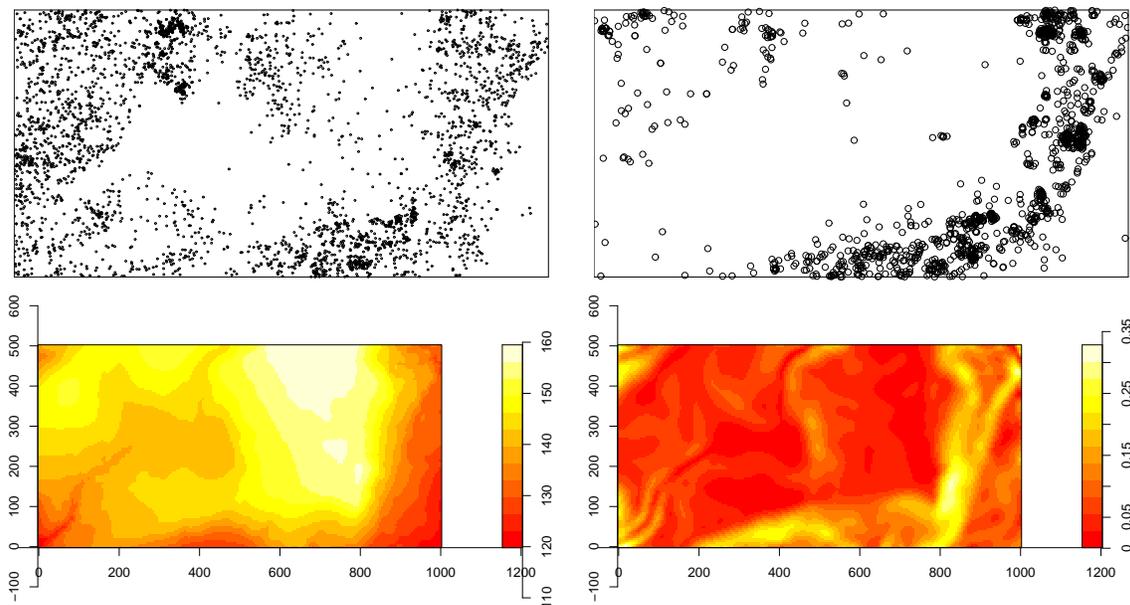


Figure 1: Upper plots: locations of *Beilschmiedia pendula* Lauraceae (left) and *Ocotea whitei* Lauraceae (right) trees. Lower plots: altitude (left) and norm of altitude gradient (right).

$j$ th diagonal element of  $\hat{\Sigma} = \Sigma(\hat{\kappa}, \hat{\alpha}, \hat{\omega}, \hat{\beta}_{1:p})$ ,  $[\hat{\beta}_j - 1.96\hat{s}d_j, \hat{\beta}_j + 1.96\hat{s}d_j]$  is an approximate 95% confidence interval for  $\beta_j$ ,  $j = 1, \dots, p$ . Our asymptotic result where  $\kappa$  tends to infinity does not justify the plug-in approach and the uncertainty of the plugged-in parameter estimates is not taken into account. 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 are generated under a Poisson process with intensity function (2), cf. Theorem 2 in Waagepetersen (2006).

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  $c$ . In Section 4  $k$  is a bivariate Gaussian density with standard deviation  $\omega$  and we use four times  $\omega$  as the radius for  $D_c$  (actually, using just three times  $\omega$  produces almost identical results).

## 4 Application to rain forest 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 positions in 1995 of the species *Beilschmiedia pendula* Lauraceae (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.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\text{Corr}(\hat{\beta}_1, \hat{\beta}_2)$	$\hat{\kappa}$	$\hat{\alpha}$	$\hat{\omega}$
Beilschmiedia	0.02 (-0.02;0.06)	5.84 (0.89;10.80)	0.39	8e-5	85.9	20.0
Ocotea	0.01 (-0.04;0.06)	14.87 (8.70;21.03)	0.55	1.2e-4	13.5	12.4

Table 1: Parameter estimates for the two species (approximate 95% confidence intervals for  $\beta_1$  and  $\beta_2$  in parentheses). The fourth column contains the estimated correlation between  $\hat{\beta}_1$  and  $\hat{\beta}_2$ .

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  $\omega$ . The  $K$ -function is then

$$K(t; \kappa, \omega) = \pi t^2 + (1 - \exp(-t^2/(4\omega)^2))/\kappa \quad (7)$$

and  $X$  can be viewed as an inhomogeneous version of the so-called Thomas process (Thomas, 1949). The upper limit  $a$  in (5) is chosen to be 100 meter for both species. 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  $\omega$  for the radius in  $D_c$ , see Section 3.1.

Table 1 shows parameter estimates for the two species. According to the approximate 95% confidence intervals (in parentheses) for  $\beta_1$  and  $\beta_2$  there is evidence that both species prefer to live on slopes but not that they favor low or high altitudes. The parameter estimates  $\beta_1$  and  $\beta_2$  seem to be rather strongly positively correlated according to the estimates of  $\text{Corr}(\hat{\beta}_1, \hat{\beta}_2)$  obtained from the approximate covariance matrix  $\hat{\Sigma}$ . The estimates of  $\kappa$  yield the expected numbers of mother points 40 and 60 within the plot for the two species.

Figure 2 shows for both species  $\hat{K}(t)$  given by (4),  $K(t, \hat{\kappa}, \hat{\omega})$ , and the  $K$ -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  $\beta_1$  and  $\beta_2$  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  $\Sigma^*$  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  $\omega^*$  equal to 10 or 20. The parameter  $\kappa^*$  is 5e-5, 1e-4, or 5e-4 corresponding to either 25, 50, or 250 expected numbers of mother points within the plot. For each value of  $\kappa^*$  we consider two values of  $\alpha^*$  so that the expected number  $\mu^*$  of simulated points is either 200 or 800 corresponding

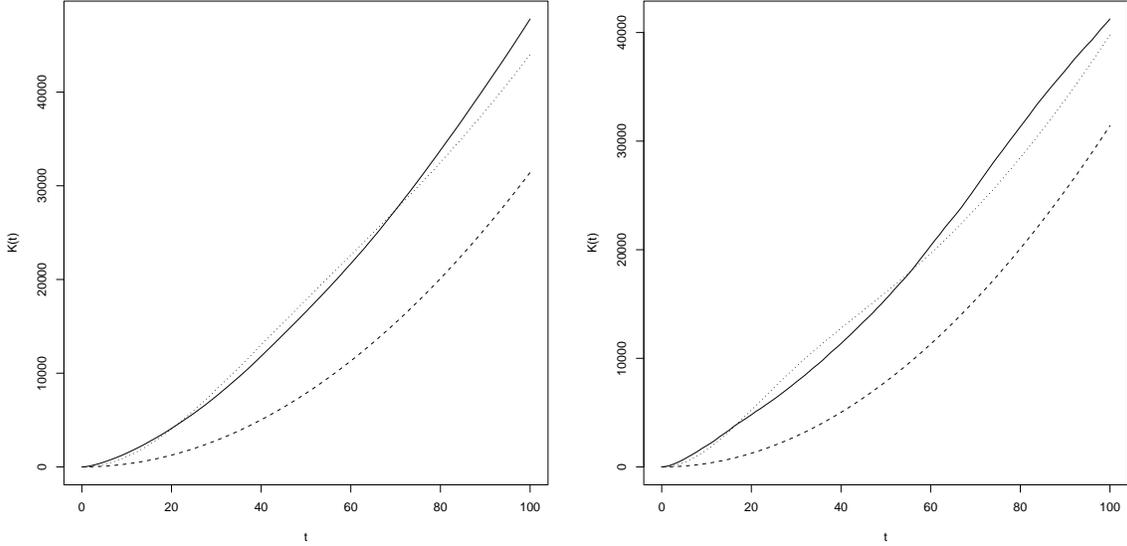


Figure 2: Solid lines:  $\hat{K}(t)$  (4) for *Beilschmiedia pendula Lauraceae* (left) and *Ocotea whitei Lauraceae* (right). Dotted lines:  $K(t, \hat{\kappa}, \hat{\omega})$  (7). Dashed lines:  $K$ -function  $K_{\text{pois}}(t) = \pi t^2$  for a Poisson process.

to “small” and “moderately large” point patterns. For each combination of  $\kappa^*$  and  $\mu^*$  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  $\omega^*$  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  $\kappa^* = 5e-5$  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  $\kappa^*$  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  $\beta_2$ : for  $\kappa^* = 5e-5$  the mean of  $\hat{\beta}_2$  is about 0.6 smaller than  $\beta_2^* = 5.84$ . For  $\kappa^* = 5e-4$  the bias is reduced to around 0.1. The estimate of  $\beta_2$  is essentially unbiased for all values of  $\kappa^*$ .

The first column in Table 2 shows for each combination of  $\kappa^*$  and  $\mu^*$ , 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  $\Sigma^*$  while Monte Carlo estimates of the medians of the standard deviations obtained from  $\hat{\Sigma}$  are given in the third column. The fourth column contains the estimated coverage percentages for the approximate 95 % confidence intervals for  $\beta_1$ . The last four columns are as the four first but for  $\beta_2$ .

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  $\beta_1$  and slightly too restrictive for  $\beta_2$ .

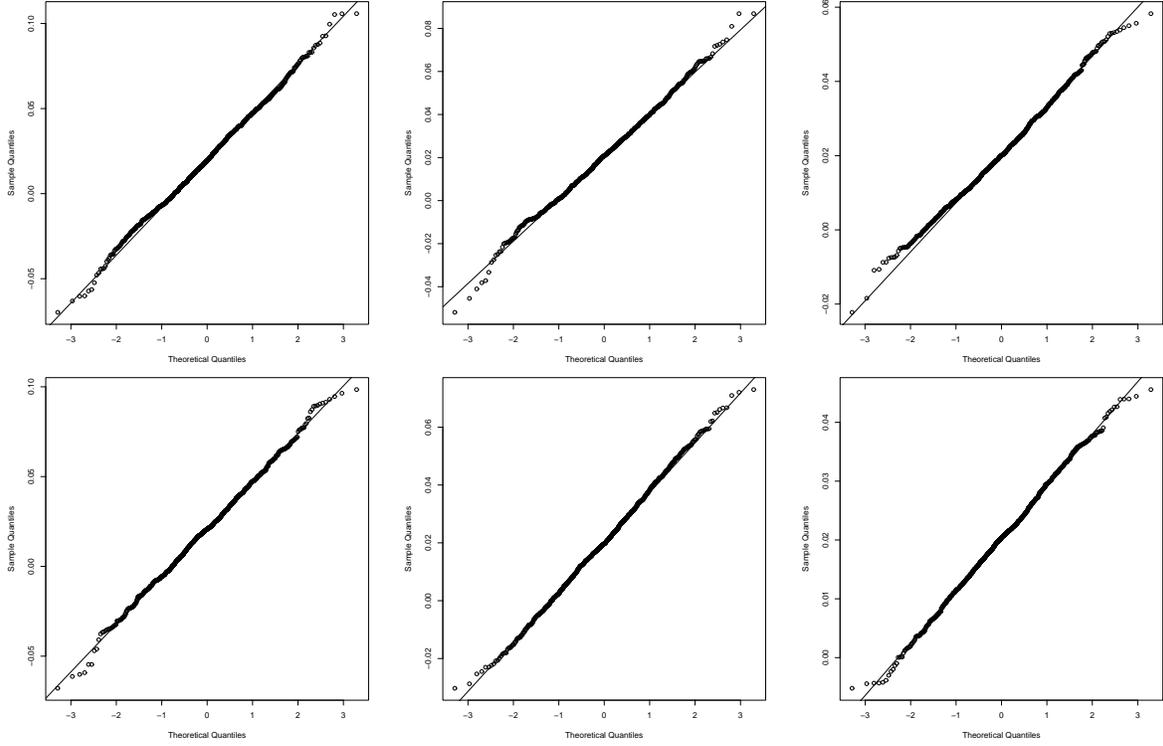


Figure 3: Quantiles obtained from 1000 simulated parameter estimates of  $\beta_1$  against quantiles of a normal distribution. Left to right  $\kappa^*=5e-5, 1e-4, 5e-4$  and top to bottom  $\mu^* = 200, 800$ .

There is in general good agreement between the first three columns regarding  $\beta_1$ . 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 narrow ridges, one should be careful with possible bias of the estimates of the associated parameter and standard error.

## 6 Discussion

We conclude by discussing merits and disadvantages of the proposed estimating function approach in relation to the alternatives of maximum likelihood estimation and a second order estimating function.

Maximum likelihood estimation is likely to be statistically more efficient but the practi-

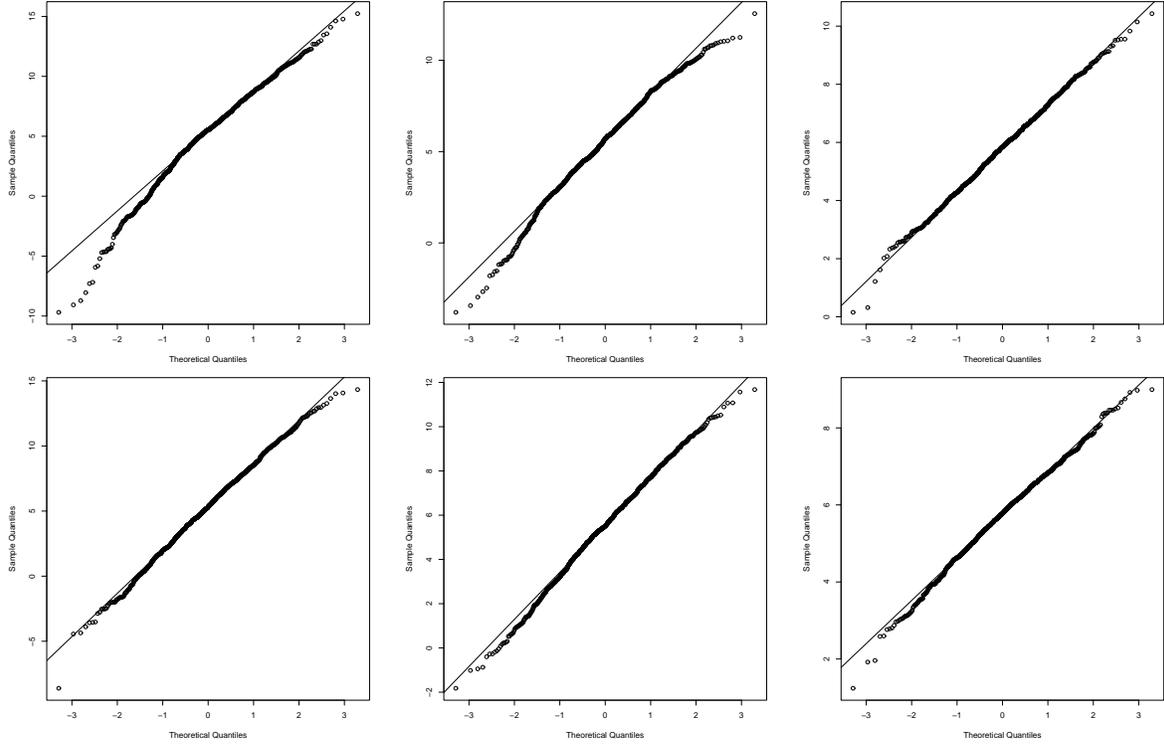


Figure 4: Quantiles obtained from 1000 simulated parameter estimates of  $\beta_2$  against quantiles of a normal distribution. Left to right  $\kappa^* = 5e-5, 1e-4, 5e-4$  and top to bottom  $\mu^* = 200, 800$ .

cal implementation using MCMC is not easy in the case of an inhomogeneous Neyman-Scott process. Let

$$\Lambda(\xi|C, \beta_{1:p}, \alpha, \omega) = \alpha \exp(z_{1:p}(\xi)\beta_{1:p}^\top) \sum_{\eta \in C} k(\xi - \eta; \omega), \quad \xi \in \mathbb{R}^2,$$

denote the random intensity function of  $X$  viewed as a Cox process. The basic computational obstacle is the integral  $\int_S \Lambda(\xi|c, \beta_{1:p}, \alpha, \omega) d\xi$  which appears in the conditional density of the data  $X \cap S$  given  $C$  and which can only be evaluated using a Riemann sum approximation. This becomes computationally very time consuming since the integral must be evaluated millions of times in connection with birth-death updates in an MCMC algorithm for simulation of  $C$  given  $X \cap S$  and with computation of importance sampling weights, see Møller and Waagepetersen (2003) Section 7.1.2, 8.6, and 10.3.1 for details. Evaluating the integral in the birth-death MCMC updates can be avoided using a data augmentation technique in Waagepetersen and Schweder (2005). However, in experiments for the Beilschmiedia data with an MCMC sampler based on data augmentation, very low acceptance rates for births or deaths of mother points are obtained. This is partly due to the large value of  $\alpha$  for this data set which implies that adding a point to  $C$  or removing a point to  $C$  induces a major change in the intensity function  $\Lambda(\xi|C, \beta_{1:p}, \alpha, \omega)$ ,  $\xi \in S$ , for

$\kappa^*$	$\mu^*$	$sd_1$	$sd_1^*$	$\hat{sd}_1$	$cvg_1$	$sd_2$	$sd_2^*$	$\hat{sd}_2$	$cvg_2$
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 2: First four columns: Monte Carlo estimate of the standard deviation for  $\hat{\beta}_1$ , standard deviation obtained from  $\Sigma^*$ , median of standard deviation obtained from  $\hat{\Sigma}$ , and coverage of approximate confidence interval. Last four columns: as first four columns but for  $\beta_2$ .

$X \cap S$  given  $C$ .

Inspired by Guan (2006) another alternative is the second order object function

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

where

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

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  $\sum_{i \neq j} [N_{ij} \log P(N_{ij} = 1) + (1 - N_{ij}) \log P(N_{ij} = 0)]$  where  $N_{ij} = 1[X \cap A_i \neq \emptyset \text{ and } X \cap A_j \neq \emptyset]$  is the indicator for simultaneous occurrence of points within  $A_i$  and  $A_j$  where  $\{A_i\}$  is a disjoint partitioning of  $S$  and the sizes of the  $A_i$  tend to zero. Differentiating  $l_2$  an unbiased second order estimating function  $u_2$  is obtained. No choice of tuning parameters is needed for  $u_2$ .

In preliminary experiments the numerical solution of  $u_2(\beta_{1:p}, \kappa, \alpha, \omega) = 0$  turned out to be rather time consuming due to the need for numerical integration to evaluate  $u_2$  and its derivative. Using a grid search for  $\omega$  combined with Newton-Raphson updates for the remaining parameters and about one hour of computing time, the estimate (0.02, 5.73, 7e-5, 95, 30) is obtained for  $(\beta_1, \beta_2, \kappa, \alpha, \omega)$  in the case of the Beilschmiedia data. The estimate for  $\omega$  differs substantially from the estimate obtained using the minimum contrast method whereas the other parameter estimates are rather similar to the ones obtained in Section 4. In a simulation study under parameter settings as in the fourth row of Table 2 we obtained Monte Carlo estimates 0.018 and 2.26 of the standard deviations for the  $u_2$  estimates of  $\beta_1$  and  $\beta_2$ . These values are very similar to the ones obtained using the first order estimating function (3), see Table 2. Hence, it is not obvious that  $u_2$  provides more efficient regression parameter estimates than (3).

In conclusion, the approach for parameter estimation suggested in Section 3 suffers from two disadvantages: the need for choosing the constant  $a$  in (5) and the potential loss

of efficiency in comparison with maximum likelihood estimation. From a computational point of view, however, the method is a very simple and fast alternative to maximum likelihood estimation and to the second order estimating function  $u_2$ . An advantage from a mathematical point of view is the easy characterization of the asymptotic properties of the regression parameter estimates obtained using (3), see Waagepetersen (2006).

**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. & Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software* **12**(6), 1–42, URL: [www.jstatsoft.org](http://www.jstatsoft.org), ISSN: 1548-7660.
- Baddeley, A. J., Møller, J. & Waagepetersen, R. (2000). Non- and semi-parametric 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. (2006). A composite likelihood approach in fitting spatial point process models. *J. Amer. Statist. Assoc.* To appear.
- 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 Dynamism. 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.
- Stoyan, D. (1992). Statistical estimation of model parameters of planar Neyman-Scott cluster processes. *Metrika* **39**, 67–74.
- Thomas, M. (1949). A generalization of Poisson's binomial limit for use in ecology. *Biometrika* **36**, 18–25.
- Waagepetersen, R. & Schweder, T. (2005). Likelihood-based inference for clustered line transect data. (submitted).
- Waagepetersen, R. P. (2006). Web-based appendix. Available at <http://www.tibs.org/biometrics>.