

Decomposition of Variance for Spatial Cox Processes

Abdollah Jalilian

Department of Statistics, Shahid Beheshti University,
Evin, Tehran, 19839-63113, Iran

Yongtao Guan

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

Rasmus Waagepetersen

Department of Mathematical Sciences, Aalborg University,
Fredrik Bajersvej 7G, DK-9220 Aalborg, Denmark

Abstract

Spatial Cox point processes is a natural framework for quantifying the various sources of variation governing the spatial distribution of rain forest trees. We introduce a general criterion for variance decomposition for spatial Cox processes and apply it to specific Cox process models with additive or log linear random intensity functions. We moreover consider a new and flexible class of pair correlation function models given in terms of normal variance mixture covariance functions. The proposed methodology is applied to point pattern data sets of locations of tropical rain forest trees.

Keywords: additive random intensity, composite likelihood, Cox process, Matérn covariance function, normal variance mixture, pair correlation function, variance component.

1 Introduction

The spatial distributions of tropical rain forest trees are influenced by many factors including e.g. spatially varying environmental conditions, seed dispersal, infectious diseases, and gap creation by hurricanes. One natural question is, loosely speaking, ‘how much of the variation’ in the spatial distribution of a tropical rain forest tree species can be attributed to each of these factors? In this paper we study methods for addressing this question with a particular focus on the contribution of the environment.

The most fundamental summaries of variation are the variances of counts of trees in bounded regions. A generalized linear mixed model (GLMM) for such counts, with a variance component for each of the sources of variation, might be a starting point for decomposing the variance. However, with this approach one loses the fine-scale information contained in extensive tropical rain forest data sets which include locations of individual trees and not just numbers of trees in certain regions (e.g. Condit, 1998). Moreover, conclusions obtained from a fitted GLMM may in general be strongly dependent on the sizes and shapes of the regions used to create the count data.

A more natural and flexible approach is to model the individual tree locations as a spatial point process (e.g. Møller and Waagepetersen, 2004) and then derive summaries of variation from a fitted spatial point process model. With this approach one can compute variances and covariances of counts in arbitrary regions and therefore recover all summaries that would be obtained from a GLMM approach. In particular, Cox processes (e.g. Møller and Waagepetersen, 2004) provide a useful framework for integrating different sources of variation.

This paper introduces a general criterion for decomposition of variance for Cox processes. The criterion is applied to specific Cox process models with either additive or log linear random intensity functions. The additive model is appealing in the context of variance decomposition but has not been well studied in the point process literature. For either model, the resulting variance decomposition depends on the assumed pair correlation function which characterizes the spatial correlation in the Cox process. It is therefore important to have a wide class of pair correlation

functions to choose from. To this end we consider a new class of shot-noise Cox processes with normal variance mixture pair correlation functions. These functions are closely related to widely used counterparts in geostatistics and hence serve as a bridge to further connect the theories of geostatistics and point processes.

1.1 A tropical rain forest data example

Waagepetersen and Guan (2009) fitted so-called inhomogeneous Thomas processes to locations of three tree species, namely *Acalypha diversifolia* (528 trees), *Lonchocarpus heptaphyllus* (836 trees), and *Capparis frondosa* (3299 trees) that were alive in 1995 in the 1000 m by 500 m Barro Colorado Island plot (Condit et al., 1996; Condit, 1998; Hubbell and Foster, 1983). The significant covariates in the fitted Thomas models were elevation and potassium for *Acalypha* and *Capparis* and nitrogen and phosphorous for *Lonchocarpus*. The point patterns of tree locations and the covariate potassium are shown in Figure 1.

[Figure 1 about here.]

In Section 7 we study decomposition of variance for these data sets. We moreover employ a much broader class of Cox processes than the inhomogeneous Thomas processes used in Waagepetersen and Guan (2009).

2 Background on spatial point processes

Let X be a point process on \mathbb{R}^2 and let $N(B)$, for any bounded $B \subseteq \mathbb{R}^2$, denote the number of points in $X \cap B$. The first- and second-order moments of the counts $N(B)$ are determined by the intensity function ρ and the second-order product density $\rho^{(2)}$ of X which are functions defined on \mathbb{R}^2 and $\mathbb{R}^2 \times \mathbb{R}^2$, respectively (see Møller and Waagepetersen, 2004). More precisely,

$$\mathbb{E}N(B) = \int_B \rho(u) du$$

and

$$\mathbb{E}N(A)N(B) = \int_{A \cap B} \rho(u) du + \int_A \int_B \rho^{(2)}(u, v) dudv$$

for bounded $A, B \subseteq \mathbb{R}^2$. The pair correlation function g is given by $g(u, v) = \rho^{(2)}(u, v) / [\rho(u)\rho(v)]$.

For a Poisson process, the counts in disjoint regions are independent and Poisson distributed. A Cox process is driven by a random intensity function $\Lambda = \{\Lambda(u) : u \in \mathbb{R}^2\}$. Conditional on $\Lambda = \lambda$, the Cox process becomes a Poisson process with intensity function λ . For a Cox process, $\rho(u) = \mathbb{E}\Lambda(u)$ and $\rho^{(2)}(u, v) = \mathbb{E}[\Lambda(u)\Lambda(v)]$.

3 Decomposition of variance

To quantify the various sources of variation in a spatial point pattern we consider the contribution of each source to the total variation of a count $N(B)$ for a region B . According to the previous section,

$$\text{Var}N(B) = \int_B \rho(u)du + \int_B \int_B \rho(u)\rho(v)[g(u, v) - 1]dudv. \quad (1)$$

The first term on the right-hand side of (1) is the variance of $N(B)$ for a Poisson process with intensity function $\rho(\cdot)$. The second term is the increase (or decrease) in variance due to possible attraction (or repulsion) between points corresponding to $g > 1$ (or $g < 1$).

For a Cox process X driven by a random intensity function Λ , we can decompose a count $N(B)$ as $N(B) = I_{|\Lambda}(B) + \hat{N}_{|\Lambda}(B)$ where

$$\hat{N}_{|\Lambda}(B) = \mathbb{E}[N(B)|\Lambda] = \int_B \Lambda(u)du$$

is the minimum mean square error predictor of $N(B)$ given Λ and $I_{|\Lambda}(B) = N(B) - \hat{N}_{|\Lambda}(B)$ is an innovation process in the terminology of Baddeley et al. (2005). Then $I_{|\Lambda}(B)$ and $\hat{N}_{|\Lambda}(B)$ are uncorrelated processes with

$$\mathbb{E}I_{|\Lambda}(B) = 0, \quad \text{Cov}[I_{|\Lambda}(A), I_{|\Lambda}(B)] = \int_{A \cap B} \rho(u)du$$

and

$$\text{Cov}[\hat{N}_{|\Lambda}(A), \hat{N}_{|\Lambda}(B)] = \int_A \int_B \rho(u)\rho(v)[g(u, v) - 1]dudv$$

since $\text{Cov}[\Lambda(u), \Lambda(v)] = \rho(u)\rho(v)[g(u, v) - 1]$. The process $I_{|\Lambda}$ may be viewed as a ‘nugget’ process in geostatistical terminology since $I_{|\Lambda}(A)$ and $I_{|\Lambda}(B)$ are uncorrelated when A and B are disjoint. The spatially correlated sources of variation causing the extra-Poisson variance in (1) enters via $\hat{N}_{|\Lambda}$.

In Section 4 we model Λ in terms of a random covariate process Z . To quantify how much of the spatially structured variation is due to Z we further decompose $\hat{N}_{|\Lambda}(B) = I_{|\Lambda, Z}(B) + \hat{N}_{|\Lambda, Z}$ into uncorrelated components

$$\hat{N}_{|\Lambda, Z} = \mathbb{E}[\hat{N}_{|\Lambda}(B)|Z] = \int_B \mathbb{E}[\Lambda(u)|Z]du$$

and

$$I_{|\Lambda, Z}(B) = \hat{N}_{|\Lambda}(B) - \hat{N}_{|\Lambda, Z}(B) = \int_B \{\Lambda(u) - \mathbb{E}[\Lambda(u)|Z]\}du.$$

If all random variation in Λ is due to Z then $\text{Var}I_{|\Lambda, Z}(B) = 0$. Further, $\text{Var}\hat{N}_{|\Lambda, Z}(B) = 0$ if Λ is independent of Z . The quantity

$$R^2(B) = \frac{\text{Var}\hat{N}_{|\Lambda, Z}(B)}{\text{Var}\hat{N}_{|\Lambda}(B)} = \frac{\text{Var}\int_B \mathbb{E}[\Lambda(u)|Z]du}{\text{Var}[\int_B \Lambda(u)du]} \quad (2)$$

thus describes the proportion of the variance of $\int_B \Lambda(u)du$ explained by Z . For sufficiently small B , $R^2(B)$ may be approximated by

$$R^2 = \frac{\text{Var}\mathbb{E}[\Lambda(u_B)|Z]}{\text{Var}\Lambda(u_B)}, \quad u_B \in B \quad (3)$$

which in the stationary case does not depend on B . The quantity R^2 can also be viewed as an analogue of the R^2 statistic for linear regression. To see this connection, we define the expected ‘sums of squares’

$$SSR = \mathbb{E} \int_S I_R^2(u)du, \quad SSE = \mathbb{E} \int_S I_E^2(u)du$$

and

$$SST = SSR + SSE$$

where $I_R(u) = \hat{\Lambda}_{|Z}(u) - \rho(u)$, $I_E(u) = \Lambda(u) - \hat{\Lambda}_{|Z}(u)$, and $\hat{\Lambda}_{|Z}(u) = \mathbb{E}[\Lambda(u)|Z]$. Then, in the case of a stationary random intensity function Λ , the ratio SSR/SST (known as R^2 for linear regression) coincides with (3). We discuss our R^2 criterion in relation to specific models in Section 4. Note that R^2 is not a goodness-of-fit criterion. In particular, $R^2 = 0$ whenever Λ is independent of Z .

4 Models for rain forest data

Let X be the process that generates the point pattern of locations for a particular tree species. We assume that X is a Cox process driven by a random intensity function Λ where Λ depends on a stationary process $Z = \{Z(u) : u \in S\}$, $Z(u) = [Z_1(u), \dots, Z_p(u)]$, of p observed covariates and a non-negative process $\Lambda_0 = \{\Lambda_0(u) : u \in S\}$ representing unobserved sources of variation. We assume that Λ depends on Z through $\tilde{Z} = \{\beta_0 + \beta_{1:p}Z(u)^\top\}_{u \in S}$ for some parameter $\beta = (\beta_0, \beta_{1:p}) = (\beta_0, \beta_1, \dots, \beta_p) \in \mathbb{R}^{p+1}$ and that \tilde{Z} and Λ_0 are independent second-order stationary processes.

We further let $\rho_0 = \mathbb{E}\Lambda_0(u)$ and $g_0(u-v) = \mathbb{E}[\Lambda_0(u)\Lambda_0(v)]/\rho_0^2$. Then ρ_0 and g_0 become respectively the intensity and the pair correlation function of a Cox process driven by the random intensity function Λ_0 . Note that

$$g_0(u-v) = c_0(u-v)/\rho_0^2 + 1 \quad (4)$$

where c_0 is the covariance function for Λ_0 . Moreover $\sigma_0^2 = \text{Var}[\Lambda_0(u)]$ is equal to $\rho_0^2[g_0(0) - 1]$. In Section 5 we discuss parametric models for g_0 or equivalently c_0 .

To quantify how much variation is due to Z , it is appropriate to model Z as a random field. However, since Z is observed, it is often convenient to base parameter estimation on $X|Z$ where Z is then treated as a fixed quantity, see Section 6.

4.1 An additive model

Variance decomposition is straightforward for the following additive model:

$$\Lambda(u) = \tilde{Z}(u) + \Lambda_0(u) = \beta_0 + \beta_{1:p}Z(u)^\top + \Lambda_0(u). \quad (5)$$

Thus, X can be considered as a superposition of two independent Cox processes with random intensity functions $\tilde{Z}(u)$ and Λ_0 . A drawback of this model is that Λ is not positive for all values of β and Z . This is probably why (5) has not attracted much interest in the point process literature; Best et al. (2000) is one notable exception. In Section 4.2 we discuss an alternative log linear model for which positivity of Λ is guaranteed.

Conditional on Z , the intensity function of X becomes

$$\rho(u|Z; \beta) = \rho_0 + \beta_0 + \beta_{1:p}Z(u)^\top.$$

In practice, a negative estimate of the intensity function may be obtained. This did, however, not happen in our data examples (Section 7). Possible problems with fitted negative intensity functions may be mitigated by enforcing linear constraints of the type $\rho_0 + \beta_0 + \beta_{1:p}Z(u_i)^\top \geq 0$ for a finite set of locations u_i e.g. including X . This is for instance possible with the R procedure `constrOptim()`.

Further, given Z , the second-order product density and the pair correlation function become

$$\rho^{(2)}(u, v|Z; \beta) = \rho(u|Z; \beta)\rho(v|Z; \beta) + c_0(u - v)$$

and $g(u, v|Z; \beta) = 1 + c_0(u - v)/[\rho(u|Z; \beta)\rho(v|Z; \beta)]$.

For (5) we obtain a straightforward decomposition of the variance of $\Lambda(u)$ into the sum of $\sigma_{\tilde{Z}}^2 = \text{Var}\tilde{Z}(u)$ and $\sigma_0^2 = \text{Var}\Lambda_0(u)$. Hence (3) becomes

$$R^2 = \frac{\sigma_{\tilde{Z}}^2}{\sigma_{\tilde{Z}}^2 + \sigma_0^2}.$$

4.2 A log linear model

The multiplicative log linear random intensity function

$$\Lambda(u) = \exp[\tilde{Z}(u) + \log \Lambda_0(u)] = \Lambda_0(u) \exp[\beta_0 + \beta_{1:p}Z(u)^\top] \quad (6)$$

is always non-negative. This model has an appealing interpretation in terms of location dependent thinning (Waagepetersen, 2007) where X is obtained by independent thinning of a Cox process driven by Λ_0 and the probability of retaining a point at u is proportional to $\exp[\tilde{Z}(u)]$. Hence in the tropical rain forest context, the covariates may be regarded as influencing the survival of plants in a stationary process of seedlings.

4.2.1 Intensity and pair correlation

Without loss of generality we assume in case of (6) that $\rho_0 = \mathbb{E}\Lambda_0(u) = 1$. Conditional on Z , the intensity function of X is

$$\rho(u|Z; \beta) = \mathbb{E}[\Lambda(u)|Z] = \exp[\tilde{Z}(u)],$$

the pair correlation function $g(u - v|Z) = g_0(u - v)$ coincides with g_0 , and

$$\rho^{(2)}(u, v|Z) = \rho(u|Z; \beta)\rho(v|Z; \beta) + \rho(u|Z; \beta)\rho(v|Z; \beta)c_0(u - v).$$

Unconditionally, the intensity and the pair correlation function become

$$\rho(u) = \rho_{\exp \tilde{Z}} = \mathbb{E} \exp[\tilde{Z}(u)] \text{ and } g(u - v) = g_0(u - v)g_{\exp \tilde{Z}}(u - v)$$

where $\rho_{\exp \tilde{Z}}$ and $g_{\exp \tilde{Z}}$ are the intensity and the pair correlation function of a Cox process with random intensity function $\exp[\tilde{Z}(\cdot)]$.

4.2.2 Decomposition of variance

For a Cox process with random intensity function (6),

$$\text{Var}[\Lambda(u)] = \sigma_{\exp \tilde{Z}}^2 + \sigma_0^2[\sigma_{\exp \tilde{Z}}^2 + \rho_{\exp \tilde{Z}}^2]$$

where

$$\sigma_{\exp \tilde{Z}}^2 = \text{Var} \mathbb{E}[\Lambda(u)|Z] = \text{Var} \exp[\tilde{Z}(u)].$$

According to (3), R^2 becomes

$$R^2 = \frac{\sigma_{\exp \tilde{Z}}^2}{\sigma_{\exp \tilde{Z}}^2 + \sigma_0^2[\sigma_{\exp \tilde{Z}}^2 + \rho_{\exp \tilde{Z}}^2]} = \frac{g_{\exp \tilde{Z}}(0) - 1}{g_{\exp \tilde{Z}}(0)g_0(0) - 1}.$$

A related approach is to consider the proportion of variance of $\log \Lambda$ explained by \tilde{Z} . If both \tilde{Z} and $\log \Lambda_0$ are Gaussian then (Møller et al., 1998)

$$\text{Var} \log \Lambda(u) = \text{Var} \tilde{Z}(u) + \text{Var} \log \Lambda_0(u) = \log g_0(0) + \log g_{\exp \tilde{Z}}(0).$$

The proportion of variance for $\log \Lambda$ explained by Z then becomes

$$\frac{\log g_{\exp \tilde{Z}}(0)}{\log g_0(0) + \log g_{\exp \tilde{Z}}(0)}$$

which is related to R^2 by the approximation $\log(x) \approx x - 1$.

5 Models for g_0

Our estimation procedure in Section 6 requires the specification of parametric models for the intensity function and the second-order product density of $X|Z$. Hence it remains to specify a parametric model for the pair correlation function g_0 , or equivalently (cf. (4)), for the covariance function c_0 of Λ_0 . A given covariance function c_0 is not necessarily a covariance function of a non-negative random field and $g_0 = c_0 - 1$ hence might not be a pair correlation function of a Cox process. In the literature on Cox processes, pair correlation models have therefore been obtained from explicit constructions of non-negative random fields. The two most popular constructions are log Gaussian and shot-noise fields leading to log Gaussian Cox processes (Møller et al., 1998) and shot-noise Cox processes (Møller, 2003). On the other hand, the most common examples of Neyman-Scott and Poisson-cluster processes are special cases of shot-noise Cox processes.

5.1 Log Gaussian Cox processes

A log Gaussian random field is obtained by exponentiating a Gaussian random field Y and the pair correlation function for the corresponding Cox process is simply the exponentiated covariance function of Y , i.e.

$$g_0(u - v) = \exp[c(u - v)] \quad (7)$$

where $c(u - v) = \text{Cov}[Y(u), Y(v)]$ (Møller et al., 1998).

5.2 Shot-noise Cox and cluster processes

A shot-noise field is in the simplest form obtained by a sum of positive kernel functions $k(\cdot - u)$ scaled by a parameter $\xi > 0$ and centered around points u of a homogeneous Poisson point process with intensity $\kappa > 0$. A Cox process driven by a shot-noise field can equivalently be considered as a Poisson-cluster process given by a superposition of clusters where for each parent point u , a Poisson number of offspring is dispersed independently around u according to the density $k(\cdot - u)$. The

resulting pair correlation function becomes

$$g_0(r) = 1 + \kappa\xi^2 \int_{\mathbb{R}^2} k(u)k(u+r)du. \quad (8)$$

Suppose g_0 is a function so that $c_0(\cdot) = g_0(\cdot) - 1$ is an absolutely integrable covariance function. Further, let \tilde{k} denote the inverse Fourier transform of the square root of the spectral density for c_0 . Then g_0 has a representation of the form (8) (see e.g. page 65 and 489 in Chilès and Delfiner, 1999) with $k = \tilde{k}$. However, g_0 is not a pair correlation function of a shot-noise Cox process unless \tilde{k} is non-negative. In general for a given g_0 , this property of \tilde{k} is not easy to verify.

The usual approach in the literature on Poisson-cluster and shot-noise Cox processes (e.g. Stoyan et al., 1995; Diggle, 2003; Møller, 2003; Illian et al., 2008; Hellmund et al., 2008) is in fact to specify first k and then obtain g_0 rather than the other way around. It is, however, remarkable that in the literature on Poisson-cluster/shot-noise Cox processes essentially only two specific examples of k are considered. For the so-called modified Thomas process, k is a Gaussian density with standard deviation ω in which case c_0 is a Gaussian covariance function

$$c_0(r) = \sigma_0^2 \exp [- (\|r\|/\eta)^2] \quad (9)$$

where $\sigma_0^2 = \kappa\xi^2/(\pi\eta^2)$ and $\eta = 2\omega$. For the Matérn cluster process, k is the density of a uniform distribution on a disc of radius r_M , say. For a Matérn process, c_0 has a less simple expression (e.g. Stoyan and Stoyan, 1995) than for the modified Thomas process but an important feature is that $c_0(r) = 0$ for $r > r_M$. Hence, the modified Thomas process and the Matérn cluster model can only produce light tailed or extremely light tailed covariance functions and this is often not appropriate e.g. for tropical rain forest data. In Wiegand et al. (2007) and Tanaka et al. (2008) more general processes are constructed using modified Thomas processes as building blocks but again only light tailed correlation structures are obtained. Tanaka et al. (2008) also consider an inverse-power type kernel k but this does not admit a closed form expression for the pair correlation function.

In the next Section 5.3 we consider a more flexible class of normal variance mixture shot-noise fields where the kernel k and the pair correlation function is given

in terms of a normal variance mixture. These normal variance mixtures provide flexibility for modeling of both shape at the origin and tail behavior of the pair correlation function.

5.3 Normal variance mixtures

Our starting point for obtaining flexible kernel functions k for shot-noise Cox processes are normal variance mixtures

$$f(r) = \int_0^\infty \phi(r; s)h(s)ds \quad (10)$$

where $\phi(\cdot; s)$ is the density of a zero-mean two-dimensional Gaussian vector with covariance matrix sI and h is some probability density on \mathbb{R}_+ . Any function of the form (10) is a positive definite function on \mathbb{R}^2 (e.g. Schlather, 1999). If anisotropy is required, the identity matrix I may be replaced by a suitable non-diagonal correlation matrix.

One can easily verify that if h is a convolution $h = \tilde{h} * \tilde{h}$ then so is f , $f = k * k$ where

$$k(u) = \int_0^\infty \phi(u; s)\tilde{h}(s)ds \quad (11)$$

which is non-negative. Hence, for the shot-noise field with kernel k we obtain $c_0(r) = \kappa\xi^2 f(r)$. A wide class of covariance functions are obtained by choosing h in the class of generalized inverse Gaussian distributions (e.g. Jørgensen, 1982) which includes both gamma, inverse gamma and inverse Gaussian distributions as special cases. The resulting class of normal variance mixtures is the class of generalized hyperbolic distributions (Barndorff-Nielsen, 1977, 1978). By Barndorff-Nielsen and Halgreen (1977) any generalized inverse Gaussian distribution is infinitely divisible and hence any generalized hyperbolic density can be represented as a convolution. However, for a generalized inverse Gaussian density h it is not always easy to identify the convolution density \tilde{h} . We discuss below some special cases of generalized hyperbolic distributions where \tilde{h} and hence k is explicitly known. A concise overview of multivariate generalized hyperbolic distributions is given in Sections 2.1-2.2 of Hu (2005).

5.3.1 Variance gamma/Bessel

A gamma density

$$h(s; \alpha, \beta) = \beta^\alpha s^{\alpha-1} \exp(-\beta s) / \Gamma(\alpha), \quad s > 0,$$

with parameters $\alpha, \beta > 0$ is a convolution of $h(\cdot; \alpha/2, \beta)$ with itself. The densities f and k given by (10) and (11) become so-called variance gamma or Bessel densities (in the one-dimensional case an early reference is McKay, 1932). Specifically k is of the form

$$k(u) = \frac{1}{\pi 2^{\nu'+1} \eta^2 \Gamma(\nu' + 1)} (\|u\|/\eta)^{\nu'} K_{\nu'}(\|u\|/\eta) \quad (12)$$

where $\nu' = \alpha/2 - 1$, $\eta = (2\beta)^{-1/2}$, and K_ν is a modified Bessel function of the second kind. The covariance function for the corresponding shot-noise field becomes

$$c_0(r) = \sigma_0^2 \frac{(\|r\|/\eta)^\nu K_\nu(\|r\|/\eta)}{2^{\nu-1} \Gamma(\nu)} \quad (13)$$

where $\nu = \alpha - 1$ and $\sigma_0^2 = \kappa \xi^2 / (4\pi \eta^2 \nu)$. The resulting Matérn pair correlation function is

$$g_0(r) = 1 + \sigma_0^2 \frac{(\|r\|/\eta)^\nu K_\nu(\|r\|/\eta)}{2^{\nu-1} \Gamma(\nu) \rho_0^2}$$

where $\rho_0 = \kappa \xi$.

The smoothness parameter ν controls the shape of the pair correlation function and gives additional flexibility in the modeling. For instance, $\nu = 1/2$ yields the exponential model

$$g_0(r) = 1 + \sigma_0^2 \exp(-\|r\|/\eta) / \rho_0^2 \quad (14)$$

which offers more slowly decaying correlations than the Thomas pair correlation function obtained with (9). The Gaussian covariance function (9) may be viewed as a limiting case of a Matérn covariance function when $\nu \rightarrow \infty$ (Stein, 1999, page 50).

For any fixed ν (9.7.2. in Abramowitz and Stegun, 1965)

$$K_\nu(x) = \frac{\exp(-x)}{\sqrt{2x/\pi}} (1 + O(1/x)), \quad (15)$$

so in the tails $\|r\| \rightarrow \infty$, the Matérn covariance function behaves as $\|r\|^{\nu-1/2} \exp(-\|r\|)$. By Corollary 2 in Yu (2011) the Matérn covariance function is log concave for $\nu \geq 1/2$. Moreover, by the same corollary, $c_0(tr_0)$ is log convex as a function of $t > 0$ for any $r_0 \in \mathbb{R}^2$ when $\nu \leq 1/2$.

5.3.2 Cauchy

An inverse gamma density with parameters $\alpha, \beta > 0$ is of the form

$$h(s; \alpha, \beta) = \beta^\alpha s^{-\alpha-1} \exp(-\beta/s) / \Gamma(\alpha), \quad s > 0.$$

It is in general not easy to identify the corresponding convolution density \tilde{h} but the case $\alpha = 1/2$ is an exception. With $\alpha = 1/2$ the Laplace transform of h is simply $L(t) = \exp(-2\sqrt{-\beta t}), t \geq 0$, so that $h(\cdot; 1/2, \beta)$ is a convolution of $h(\cdot; 1/2, \beta/4)$ with itself. The corresponding kernel function becomes a Cauchy density

$$k(u) = \frac{1}{2\pi\omega^2} [1 + (\|u\|/\omega)^2]^{-3/2}$$

with $\omega = \sqrt{\beta/2}$ and the covariance function becomes

$$c_0(r) = \sigma_0^2 [1 + (\|r\|/\eta)^2]^{-3/2} \quad (16)$$

with $\sigma_0^2 = \kappa\xi^2/(2\pi\eta^2)$, and $\eta = 2\omega$. The Cauchy covariance function is polynomially decreasing and hence more suitable than the Matérn model for modeling of a slowly decaying covariance. The Cauchy covariance is log concave in a neighbourhood around the origin.

5.3.3 Normal inverse Gauss

The density of an inverse Gauss distribution is given by

$$h(s; \mu, \lambda) = \left(\frac{\lambda}{2\pi s^3}\right)^{1/2} \exp\left[-\frac{\lambda(s-\mu)^2}{2\mu^2 s}\right]$$

with $\mu, \lambda > 0$. The inverse Gauss density $h(\cdot; \mu, \lambda)$ is the convolution of $h(\cdot; \mu/2, \lambda/4)$ with itself. The resulting normal inverse Gauss kernel is

$$k(u) = \frac{(\gamma/4)^{1/2} \exp(\sqrt{\gamma/4}) K_{3/2}(\sqrt{\gamma/4 + (\|u\|/\eta)^2})}{\eta^2 \pi^{3/2} \sqrt{2} [\gamma/4 + (\|u\|/\eta)^2]^{3/4}}$$

where $\gamma = (\lambda/\mu)^2$ and $\eta = \mu/\sqrt{\lambda}$ and

$$c_0(r) = \sigma_0^2 \frac{K_{3/2}(\sqrt{\gamma + (\|r\|/\eta)^2}) \gamma^{3/4}}{K_{3/2}(\sqrt{\gamma}) [\gamma + (\|r\|/\eta)^2]^{3/4}}$$

where $\sigma_0^2 = \kappa \xi^2 \gamma^{1/2} \exp(\sqrt{\gamma}) K_{3/2}(\sqrt{\gamma}) / (\eta^2 \pi^{3/2} \sqrt{2} \gamma^{3/4})$. By the asymptotic formula (15), the normal inverse Gauss covariance behaves as $\exp(-\|r\|) / \|r\|^2$ in the tails. The parameter γ may be viewed as a kind of ‘offset’ parameter for the squared distance $\|r\|^2$.

Remark The convolution representations of the Matérn and Cauchy covariance functions above are not new and were found using Fourier transforms in Matérn (1986) (see Table 1, page 30 - a detailed derivation for the Matérn class is moreover provided in Jonsdottir et al., 2011). However, these representations and the shot-noise/Poisson-cluster processes with the corresponding kernels have not been considered previously in the point process literature. Moreover, the normal variance mixture perspective leads to straightforward simulation of offspring locations when the resulting shot-noise Cox process is viewed as a Poisson-cluster process. The Matérn and the normal inverse Gauss covariance functions both have three parameters. However, the Matérn class may seem preferable since the parameter ν allows for modeling of a range of kernel and covariance shapes varying from log convex to log concave.

6 Parameter estimation

In practice we consider a parametric model $c_0(\cdot; \psi)$ for the covariance function of Λ_0 where c_0 could for instance be the Matérn covariance function (13) with $\psi = (\sigma_0^2, \eta, \nu)$. Given (X, Z) observed within $W \subseteq \mathbb{R}^2$, we then obtain a plug-in estimate for R^2 in (3) after estimating the three parameters β , ψ , and $\sigma_{\tilde{Z}}^2$ (or $\sigma_{\exp \tilde{Z}}^2$).

6.1 Composite likelihood estimation

The inference about β and ψ is based on $X|Z$. The regression parameter β can be estimated by the first-order composite log likelihood function (CL_1) (Schoenberg, 2005; Waagepetersen, 2007)

$$\text{CL}_1(\beta) = \sum_{u \in X} \log \rho(u|Z; \beta) - \int_W \rho(u|Z; \beta) du. \quad (17)$$

This is the log likelihood function in the case where X is a Poisson process with intensity function $\rho(\cdot|Z; \beta)$. The Berman-Turner quadrature scheme can be used to approximate the integral in (17) both for the log linear and the additive model, see e.g. Baddeley and Turner (2000).

For a Cox process specified by the log linear random intensity function (6) one may subsequently estimate ψ using minimum contrast estimation based on the K -function, see Waagepetersen (2007) and Waagepetersen and Guan (2009). However, for the additive model the K -function is not well-defined since the pair correlation function is not translation invariant (Baddeley et al., 2000). As suggested by Guan (2006) and Waagepetersen (2007), ψ can instead be estimated using a second-order composite likelihood function

$$\begin{aligned} \text{CL}_2(\psi|\beta) &= \sum_{\substack{\neq \\ u,v \in X}} w(u,v) \log \rho^{(2)}(u,v|Z; \beta, \psi) \\ &\quad - \iint_{W^2} w(u,v) \rho^{(2)}(u,v|Z; \beta, \psi) dudv \end{aligned} \quad (18)$$

based on the second-order product density where w is a weight function, see e.g. (20) and (21) below. The integral term in (18) can also be approximated by a variant of the Berman-Turner scheme. The second-order composite likelihood can be evaluated for both the log linear model and the additive model and the maximized CL_2 can be used as a criterion for model selection, see Section 7.

Instead of maximizing the right-hand side of (18) with respect to both β and ψ , a computationally simpler approach is to obtain $\hat{\beta}$ by maximizing (17) and then $\hat{\psi}$ by maximizing $\text{CL}_2(\cdot|\hat{\beta})$. Moreover, an equivalent version of $\text{CL}_2(\psi|\hat{\beta})$ is given by

$$\begin{aligned} \text{CL}_2^*(\psi|\hat{\beta}) &= \sum_{\substack{\neq \\ u,v \in X}} w(u,v) \log g(u,v|Z; \hat{\beta}, \psi) \\ &\quad - \iint_{W^2} w(u,v) \text{Cov}[\Lambda(u), \Lambda(v)|Z; \hat{\beta}, \psi] dudv \end{aligned} \quad (19)$$

which is obtained from $\text{CL}_2(\psi|\hat{\beta})$ by subtracting the second-order composite likelihood of a Poisson process with intensity function $\rho(\cdot|Z; \hat{\beta})$. More stable convergence results were obtained using (19) instead of (18) when using a standard implemen-

tation of the Nelder-Mead algorithm for maximizing the second-order composite likelihood.

In simulation studies and applications to real data we discovered that the second-order composite likelihood estimates can be quite sensitive to the accuracy of the Berman-Turner quadrature scheme used to approximate the double integral in (19). This is especially the case when c_0 is steep at zero like for (13) with a small ν or small η . Regarding the weight function w ,

$$w(u, v) = 1[\|u - v\| \leq t]/[\pi t^2] \quad (20)$$

is the standard choice which ensures that only t -close pairs of points are used in the composite likelihood. In the data example in Section 7 we also considered

$$w(u, v) = 1[\|u - v\| \leq t]/[\pi t^2 \rho(u|Z, \hat{\beta})\rho(u|Z, \hat{\beta})]. \quad (21)$$

For log linear models, this w implies a simplification of the double integral since the intensity function is eliminated from the integrand.

6.2 Estimation of environmental variances

In practice Z is observed on a grid $G = \{u_i\}_{i=1, \dots, M}$ covering W . Since we assume that Z is a stationary process, simple estimators of $\sigma_{\bar{Z}}^2$ and $\sigma_{\exp \bar{Z}}^2$ are given by

$$\hat{\sigma}_{\bar{Z}}^2 = \frac{1}{M} \sum_{u \in G} [\hat{\tilde{Z}}(u) - \hat{\rho}_{\bar{Z}}]^2 \quad (22)$$

and

$$\hat{\sigma}_{\exp \bar{Z}}^2 = \frac{1}{M} \sum_{u \in G} \left\{ \exp [\hat{\tilde{Z}}(u)] - \hat{\rho}_{\exp \bar{Z}} \right\}^2 \quad (23)$$

where

$$\hat{\rho}_{\bar{Z}} = \frac{1}{M} \sum_{u \in G} \hat{\tilde{Z}}(u), \quad \hat{\rho}_{\exp \bar{Z}} = \frac{1}{M} \sum_{u \in G} \exp [\hat{\tilde{Z}}(u)], \quad (24)$$

and $\hat{\tilde{Z}}(u) = \hat{\beta}Z(u)^\top$.

6.3 Discussion of theoretical properties of estimates

One may be concerned about the joint properties of the composite likelihood estimates and (22)-(24). A fully detailed asymptotic analysis is outside the scope of this paper and we here only outline the essential arguments. We may view the estimate $\hat{\theta} = (\hat{\beta}, \hat{\psi})$ as a solution of an estimating equation $u(\theta) = 0$ where u is obtained by concatenating the gradients of the log composite likelihoods (17) and (18). Note that u is conditionally unbiased, $\mathbb{E}_\theta[u(\theta)|Z] = 0$. Moreover the estimates (22)-(24) are obtained in a simple manner from spatial averages of the form $A(\hat{\beta}) = \sum_{u \in M} f(Z(u), \hat{\beta})$ where $f(z, \beta)$ is either βz^\top , $(\beta z^\top)^2$, $\exp(\beta z^\top)$ or $\exp(2\beta z^\top)$.

Under appropriate mixing conditions for Λ_0 and Z one may verify using techniques as in Waagepetersen and Guan (2009) that asymptotically (increasing window size $|W|$), $|W|^{1/2}(\hat{\theta} - \theta)$ is distributed as $u(\theta)S^{-1}$ for a matrix $S = -\mathbb{E}du(\beta)/d\beta$ and that $u(\theta)$ is asymptotically normal. Hence $|W|^{1/2}(\hat{\theta} - \theta)$ is asymptotically zero-mean normal. Moreover,

$$|W|^{1/2}(A(\hat{\beta}) - \mathbb{E}A(\beta)) = |W|^{1/2}(A(\beta) - \mathbb{E}A(\beta)) + |W|^{1/2}(\hat{\beta} - \beta)\mathbb{E}\frac{d}{d\beta}A(\beta) + o_P(1).$$

By mixing the first term on the right hand side converges in distribution to a normal distribution. It thus follows that the spatial average $A(\hat{\beta})$ is a $|W|^{1/2}$ consistent estimate of $\mathbb{E}A(\beta)$ but the replacement of β with $\hat{\beta}$ introduces additional error in the estimation quantified by the second term in the right hand side of the above equation. This error is asymptotically independent of the error given by the first term since

$$\text{Cov}_\theta[A(\beta), u(\theta)] = \mathbb{E}_\theta \text{Cov}_\theta[A(\beta), u(\theta)|Z] + \text{Cov}_\theta\{\mathbb{E}_\theta[A(\beta)|Z], \mathbb{E}_\theta[u(\theta)|Z]\} = 0$$

by the unbiasedness of $u(\cdot)$ given Z and since $A(\beta)$ is constant given Z .

To summarize, under appropriate mixing conditions, the composite likelihood estimates and the estimates (22)-(24) all become consistent and asymptotically normal.

7 Decomposition of variance for tropical rain forest data example

In this section we return to the data example in Section 1.1. Waagepetersen and Guan (2009) fitted inhomogeneous Thomas models with random intensity functions of the form (6) and a Gaussian covariance function (9) for Λ_0 . Using the same covariates as in Waagepetersen and Guan (2009), we fit for each of the species the additive model (5) and the log linear model (6) with c_0 being the Gaussian covariance function (9), the Matérn covariance function (13), and the Cauchy covariance function (16). Note that these covariance functions have very distinct behaviors both at the origin and in the tails, see also Figure 3. As in Myllymäki and Penttinen (2009) we also consider a pair correlation function of the form (7) where Y is a Gaussian random field with a Matérn covariance function. We denote this covariance function LG-Matérn since it corresponds to the case of a log Gaussian random intensity function Λ_0 . In this case $\sigma_0^2 = \exp[\text{Var}Y(u)] - 1$. As measures of fit for these models, we use the maximal values of the composite likelihoods CL_1 and CL_2 .

Regarding the weight function w we tried both (20) and (21) with $t = 125$ in (18). The integrals in the composite likelihoods CL_1 and CL_2 were approximated using a Berman-Turner quadrature scheme consisting of data points and 200×100 dummy points over the observation window $W = [0, 1000] \times [0, 500]$. The ranking of the models according to CL_2 did not depend on the choice of w -function (except for a single swap of ranks between Cauchy and Matérn in case of the additive model for *Lonchocarpus*). However, for the log linear models we in general obtained somewhat smaller estimates of σ_0^2 with (20) than with (21) and vice versa for the additive models. According to a model check for the best fitting log linear models (see below) the estimates obtained with (21) gave the best fit. In the following we restrict attention to the results obtained with (21).

Table 1 shows the parameter estimates, the maximal composite likelihood values, and the estimated R^2 for each species and model.

[Table 1 about here.]

Considering first $\beta_{1:2}$ for each species, the regression parameters have similar

signs and relative magnitudes for the log linear model and the additive model. However, the maximal first-order composite likelihood CL_1 is always largest for the log linear model. The fitted \tilde{Z} are shown in Figure 2 for each species.

[Figure 2 about here.]

Considering CL_2 and comparing log linear and additive models for each species, the log linear models always yield the largest CL_2 values. Regarding the choice of covariance function, the best fit for *Acalypha* is obtained with the Cauchy covariance function while for *Lonchocarpus* and *Capparis*, the LG-Matérn covariance function performs better, followed by the Matérn covariance function. Both for the log linear and the additive models, the smallest CL_2 is obtained with the Gaussian covariance function. This suggests that the fast decaying pair correlation function obtained with the Gaussian covariance function is not suitable for the tropical rain forest data. For the additive model, the LG-Matérn and the Matérn covariance functions are almost identical because in this case $\hat{\sigma}_0^2$ is small and $\exp(x) - 1 \approx x$ when $x \approx 0$.

Regarding R^2 , the estimates vary considerably across models. However, the overall qualitative conclusion is stable: the contribution of the environment is smallest for *Acalypha* and largest for *Capparis*. This may be linked to the different modes of seed dispersal of the species, see Waagepetersen and Guan (2009). The R^2 obtained with the best fitting models are 0.01, 0.06, and 0.11 for *Acalypha*, *Lonchocarpus*, and *Capparis*.

The fitted covariance functions for the log linear models and for all species are shown in Figure 3. For all species, the Gaussian covariance function differ much from the other three covariance functions both at the origin and in the tail. For *Lonchocarpus* and *Capparis*, the fitted Matérn and LG-Matérn covariance functions appear rather similar.

[Figure 3 about here.]

We used a non-parametric estimate of the g_0 -function (see e.g. Chapter 4 in Møller and Waagepetersen, 2004) as a summary statistic for model assessment and computed pointwise 90% confidence bands for this statistic using simulations from the best fitting log linear models according to CL_2 . The non-parametric estimates

of $c_0(\cdot) = g_0(\cdot) - 1$ and 90% pointwise confidence bands are shown in Figure 4 and do not provide evidence against the fitted models.

[Figure 4 about here.]

8 Discussion

In this paper we introduced a method for decomposing variance for spatial Cox processes. As can be seen from our data analysis, the results can be sensitive to the choice of the pair correlation function. Fortunately, the flexible class of normal variance mixture pair correlation functions allow us to compare results from a wide range of pair correlation functions. We can then report results obtained with the best fitting model according to the second-order composite likelihood criterion.

Concerning the second-order composite likelihood estimation, further studies regarding numerical implementation and choice of w -function seem needed. For the Matérn model, the joint estimation of (σ_0^2, η, ν) is computationally demanding and simulation studies indicate that the statistical properties of the estimates can be poor. For routine use, a more feasible approach is to maximize only with respect to σ_0^2 and η for each ν in a moderate collection of ν values.

The Thomas process has enjoyed much popularity in the point process literature. However, at least for the tropical rain forest data considered in this paper, it seems inferior to models with a more slowly decaying pair correlation function. For all the data examples considered, the log linear model provided a better fit than the additive model. This is perhaps not so surprising since the log linear model has an appealing interpretation in terms of survival of seedlings. On the other hand, variance decomposition is more straightforward for the additive model than for the log linear model.

Our proposed variance decomposition procedure assumes a stationary random environment. This may e.g. not be tenable for *Acalypha* because the fitted \tilde{Z} in W shows a trend from right to left. Although this does not necessarily invalidate the stationarity assumption for \tilde{Z} , it at least implies that W may be too small to allow a precise estimate of the variance for \tilde{Z} . Further research is required to handle the

case where \tilde{Z} is not stationary but perhaps satisfies a weaker assumption of intrinsic stationarity.

Acknowledgements

We thank two anonymous referees for constructive comments which helped to improve the manuscript.

Abdollah Jalilian and Rasmus Waagepetersen's research was supported by the Danish Natural Science Research Council, grant 09-072331 'Point process modeling and statistical inference', and by Centre for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation. Yongtao Guan's research was supported by NSF grant DMS-0845368 and by NIH grant 1R01DA029081.

The BCI forest dynamics research project was made possible by National Science Foundation grants to Stephen P. Hubbell: DEB-0640386, DEB-0425651, DEB-0346488, DEB-0129874, DEB-00753102, DEB-9909347, DEB-9615226, DEB-9615226, DEB-9405933, DEB-9221033, DEB-9100058, DEB-8906869, DEB-8605042, DEB-8206992, DEB-7922197, support from the Center for Tropical Forest Science, the Smithsonian Tropical Research Institute, the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, the Celera Foundation, and numerous private individuals, and through the hard work of over 100 people from 10 countries over the past two decades. The plot project is part of the Center for Tropical Forest Science, a global network of large-scale demographic tree plots.

The BCI soils data set were collected and analyzed by J. Dalling, R. John and K. Harms with support from NSF DEB021104, 021115, 0212284, 0212818 and OISE 0314581 and OISE 031458, STRI and CTFS.

References

Abramowitz, M. and Stegun, I. (1965). *Handbook of Mathematical Functions*. Dover, New York, 9 edition.

- Baddeley, A. and Turner, R. (2000). Practical maximum pseudolikelihood for spatial point patterns. *Aust. N. Z. J. Stat.* **42**, 283–322.
- Baddeley, A. J., Møller, J., and Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat. Neerl.* **54**, 329–350.
- Baddeley, A. J., Turner, R., Møller, J., and Hazelton, M. (2005). Residual analysis for spatial point processes (with discussion). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**, 617–666.
- Barndorff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **353**, 401–419.
- Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Stat.* **5**, 151–157.
- Barndorff-Nielsen, O. E. and Halgreen, C. (1977). Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **38**, 309–312.
- Best, N. G., Ickstadt, K., and Wolpert, R. L. (2000). Spatial Poisson regression for health and exposure data measured at disparate resolutions. *J. Amer. Statist. Assoc.* **95**, 1076–1088.
- Chilès, J.-P. and Delfiner, P. (1999). *Geostatistics - modeling spatial uncertainty*. Probability and Statistics. Wiley.
- Condit, R. (1998). *Tropical Forest Census Plots*. Springer-Verlag and R. G. Landes Company, Berlin, Germany and Georgetown, Texas.
- Condit, R., Hubbell, S. P., and 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*. Arnold, second edition.
- Guan, Y. (2006). A composite likelihood approach in fitting spatial point process models. *J. Amer. Statist. Assoc.* **101**, 1502–1512.
- Hellmund, G., Prokešová, M., and Jensen, E. B. V. (2008). Lévy-based Cox point processes. *Adv. in Appl. Probab.* **40**, 603–629.
- Hu, W. (2005). *Calibration of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization and portfolio credit risk*. PhD thesis, Florida State University, Tallahassee, FL, USA.
- Hubbell, S. P. and Foster, R. B. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In Sutton, S. L., Whitmore, T. C., and Chadwick, A. C., editors, *Tropical Rain Forest: Ecology and Management*, pages 25–41. Blackwell Scientific Publications, Oxford.
- Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*. Statistics in Practice. Wiley.
- Jonsdottir, K. Y., Rønn-Nielsen, A., Mouridsen, K., and Jensen, E. B. V. (2011). Lévy-based modelling in brain imaging. Research Report 2, Centre for Stochastic Geometry and Advanced Bioimaging.
- Jørgensen, B. (1982). *Statistical properties of the generalized inverse Gaussian distribution*. Lecture Notes in Statistics. Springer.
- Matérn, B. (1986). *Spatial Variation*. Number 36 in Lecture Notes in Statistics. Springer-Verlag, Berlin.
- McKay, A. T. (1932). A Bessel function distribution. *Biometrika* **24**, 39–44.
- Møller, J. (2003). Shot noise Cox processes. *Adv. in Appl. Probab.* **35**, 614–640.

- Møller, J., Syversveen, A. R., and Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25**, 451–482.
- Møller, J. and Waagepetersen, R. P. (2004). *Statistical inference and simulation for spatial point processes*. Chapman and Hall/CRC, Boca Raton.
- Myllymäki, M. and Penttinen, A. (2009). Conditionally heteroscedastic intensity-dependent marking of log Gaussian Cox processes. *Stat. Neerl.* **63**, 450–473.
- Schlather, M. (1999). Introduction to positive definite functions and to unconditional simulation of random fields. ST 99-10, Lancaster University.
- Schoenberg, F. P. (2005). Consistent parametric estimation of the intensity of a spatial-temporal point process. *J. Statist. Plann. Inference* **128**, 79–93.
- Stein, M. L. (1999). *Spatial interpolation, some theory for kriging*. Springer Verlag, New York.
- Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and Its Applications*. Wiley, Chichester, second edition.
- Stoyan, D. and Stoyan, H. (1995). *Fractals, Random Shapes and Point Fields*. Wiley, Chichester.
- Tanaka, U., Ogata, Y., and Stoyan, D. (2008). Parameter estimation and model selection for Neyman-Scott point processes. *Biom. J.* **50**, 43–57.
- Waagepetersen, R. (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics* **63**, 252–258.
- Waagepetersen, R. and Guan, Y. (2009). Two-step estimation for inhomogeneous spatial point processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71**, 685–702.
- Wiegand, T., Gunatilleke, S., Gunatilleke, N., and Okuda, T. (2007). Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Ecology* **88**, 3088–3102.
- Yu, Y. (2011). On Normal Variance-Mean Mixtures. *ArXiv e-prints* .

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

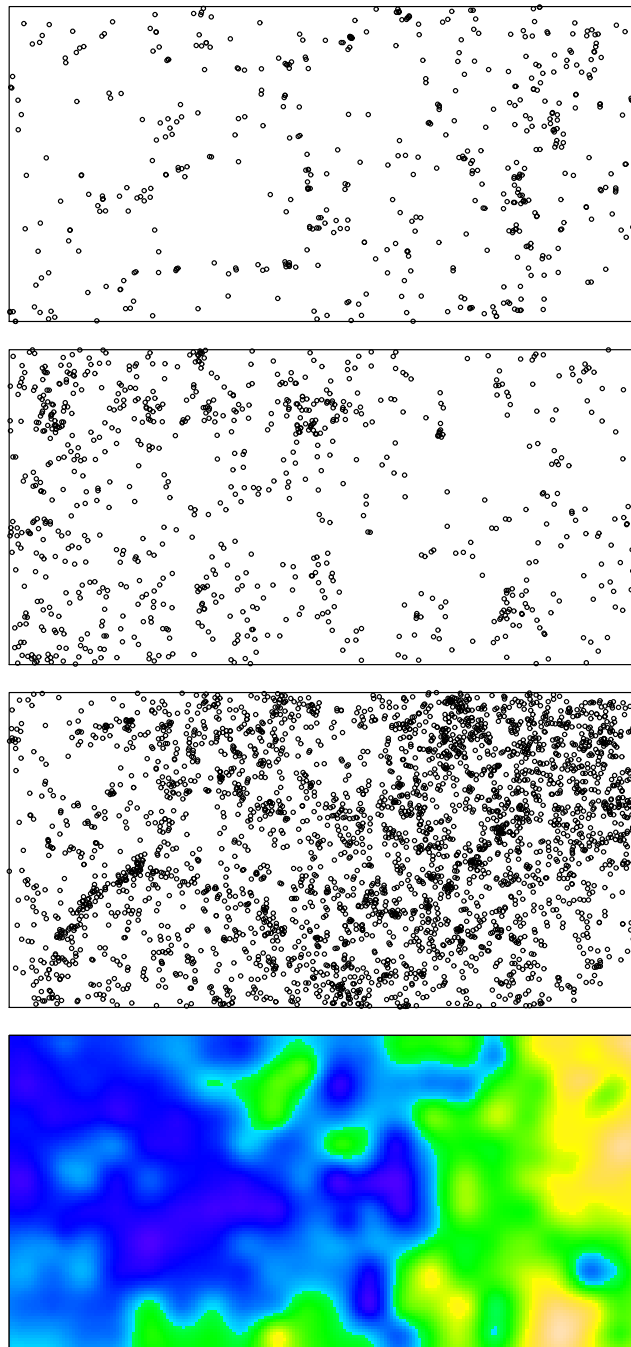


Figure 1: Locations of *Acalypha*, *Lonchocarpus*, and *Capparis* trees and image of interpolated potassium content in the surface soil (from top to bottom).

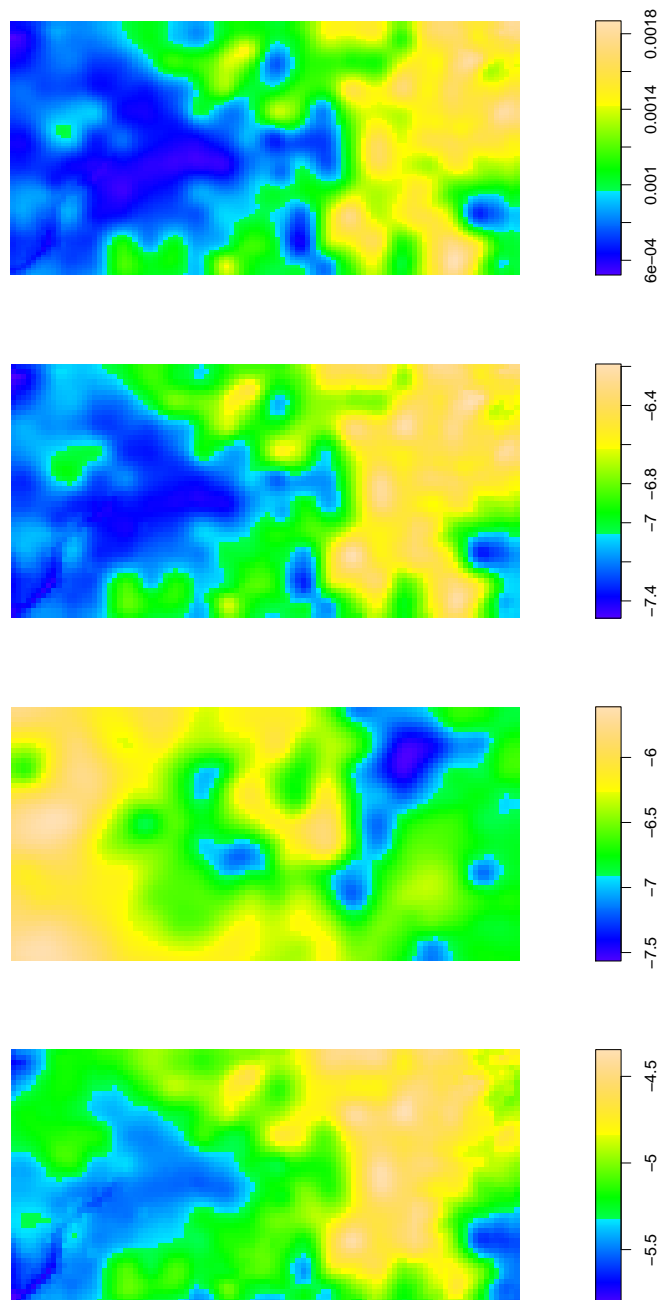


Figure 2: The fitted regression term, \hat{Z} , for *Acalypha* under the additive model and the log linear model, and for *Lonchocarpus* and *Capparis* under the log linear model (from top to bottom).

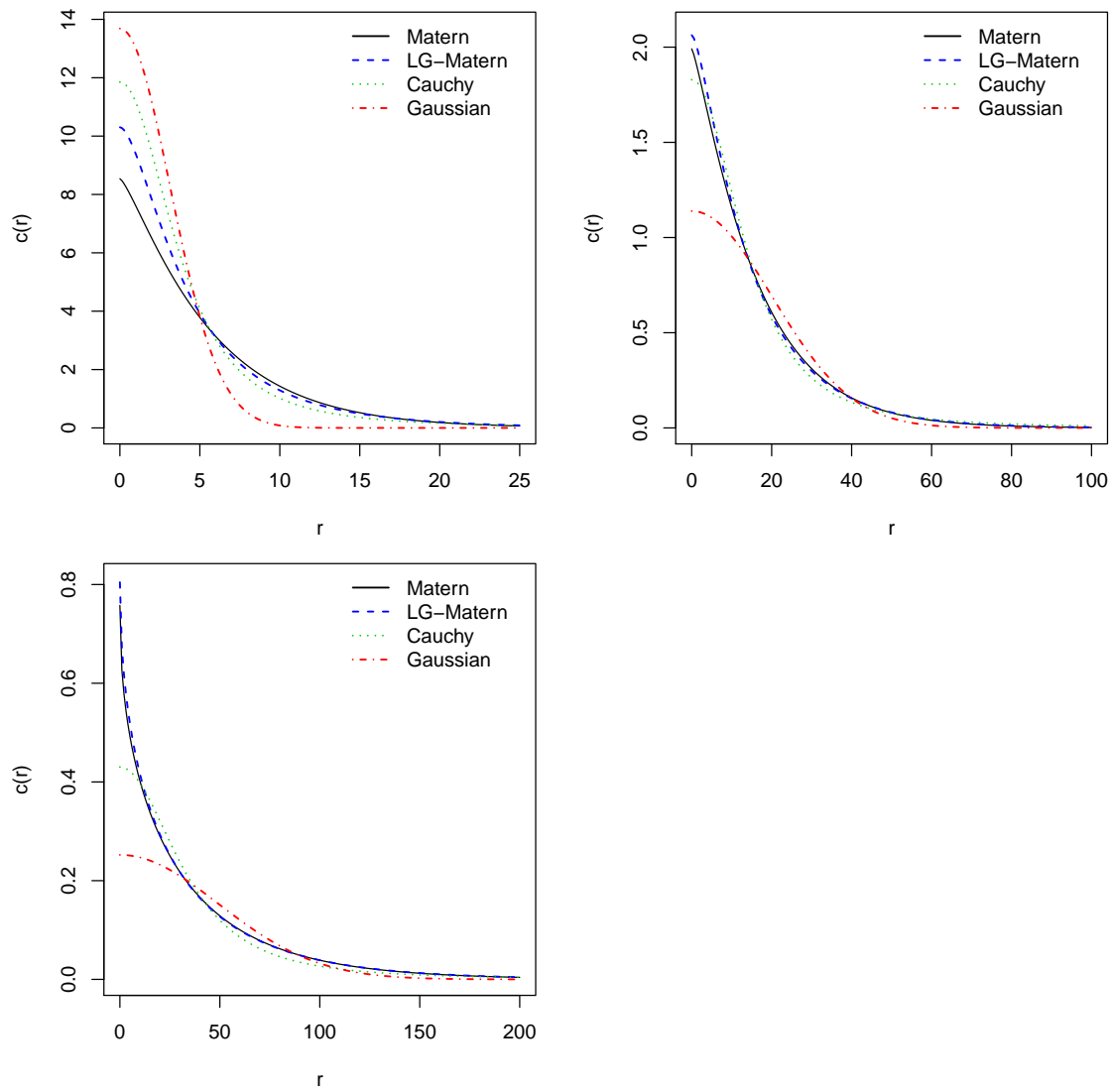


Figure 3: The various fitted covariance functions c_0 in case of the log linear model for *Acalypha* (top left), *Lonchocarpus* (top right), and *Capparis* (bottom left).

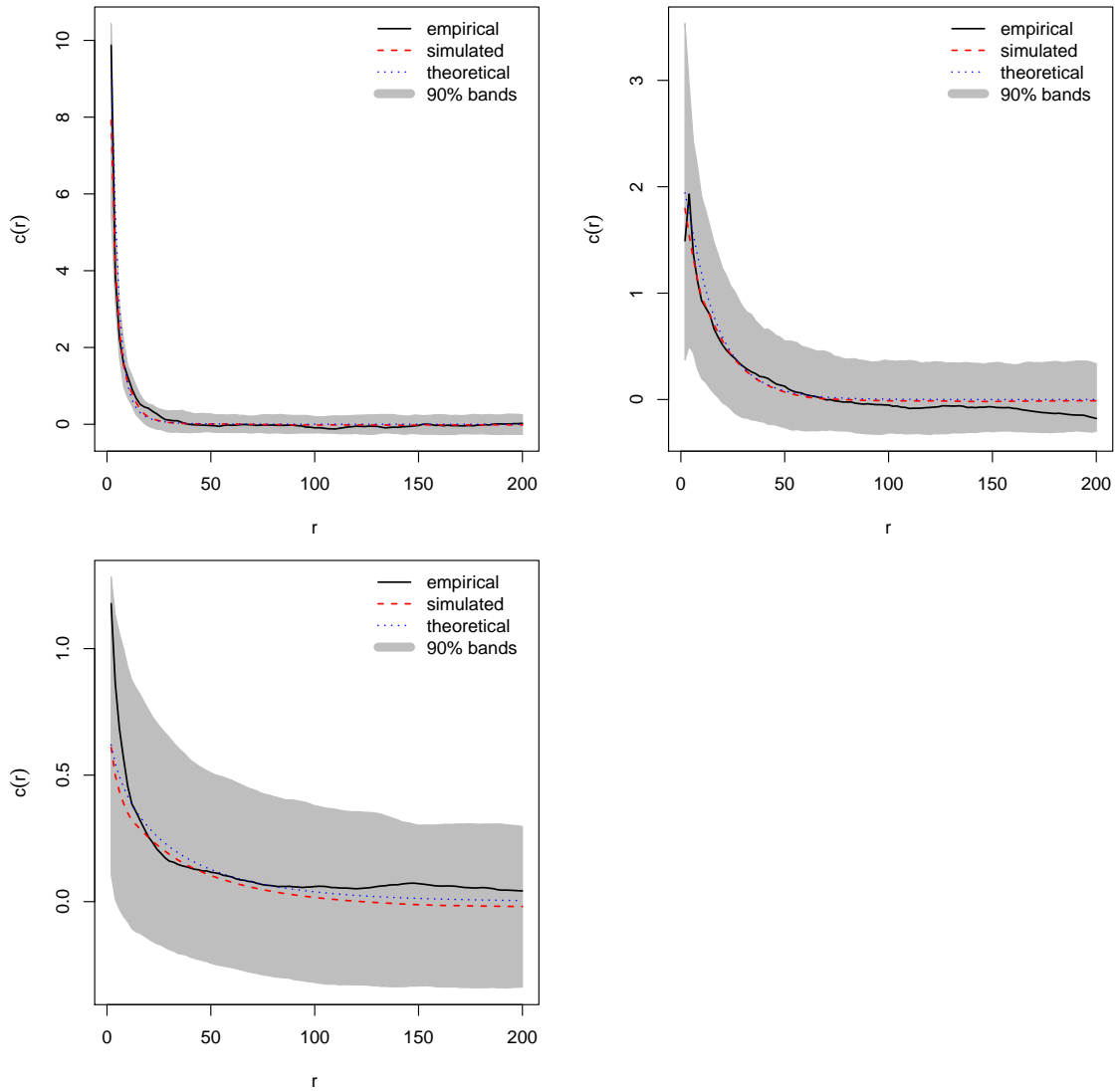


Figure 4: Non-parametric estimates of c_0 (solid line) and 90% pointwise confidence bands (gray area) obtained from 100 simulations under best fitting models. The red dashed line shows mean of simulated non-parametric estimates and the blue dotted line shows the parametric estimate of c_0 . Top left: Acalypha, top right: Lonchocarpus, and bottom left: Capparis.

Table 1: Estimates of β , ψ and R^2 and maximal values of the composite likelihoods CL_1 and CL_2 . For each species, the first four rows in the last three columns correspond to the log linear model with Gaussian, Cauchy, Matérn and LG-Matérn covariance functions. The next four rows are for the additive model with the same covariance models. For LG-Matérn, abusing notation, $\hat{\psi}$ denotes the parameter estimate for the covariance function of the Gaussian field Y . In the first column, the estimate $\hat{\rho}$ is given by $N(W)/|W|$. The second column also shows estimates of $\hat{\sigma}_{\exp\tilde{Z}}^2$ and $\hat{\sigma}_{\tilde{Z}}^2$.

Species	$\hat{\beta}_{1:2}$ / $CL_1(\hat{\beta})$	$\hat{\psi} = (\hat{\sigma}_0^2, \hat{\eta}, \hat{\nu})$	$CL_2(\hat{\psi} \hat{\beta}) - C$	R^2	
Acalypha	(0.02, 0.005)	(13.7, 4.4, ∞)	27438.39	0.01	
	-4117.7	(11.9, 4.9, -)	28744.97	0.01	
	$\hat{\sigma}_{\exp\tilde{Z}}^2 = 0.13 \times 10^{-6}$	(8.5, 4.7, 0.69)	28507.05	0.01	
		(2.4, 5.6, 1.02)	28675.13	0.01	
	$(15.5, 4.9) \times 10^{-6}$	$(11.6 \times 10^{-6}, 4.1, \infty)$	0	0.01	
		$(8.1 \times 10^{-6}, 5.2, -)$	1724.32	0.01	
	$\hat{\rho} = 1056 \times 10^{-6}$	-4119.7	$(6.1 \times 10^{-6}, 5.8, 0.56)$	1128.50	0.02
	$C = -6291053.0$	$\hat{\sigma}_{\tilde{Z}}^2 = 0.11 \times 10^{-6}$	$(6.1 \times 10^{-6}, 5.8, 0.56)$	1128.50	0.02
Lonchocarpus	(-0.03, -0.16)	(1.1, 28.4, ∞)	82006.98	0.11	
	-6117.6	(1.8, 18.4, -)	82174.76	0.07	
	$\hat{\sigma}_{\exp\tilde{Z}}^2 = 0.45 \times 10^{-6}$	(2.0, 14.0, 0.65)	82326.85	0.06	
		(1.1, 14.8, 0.86)	82344.08	0.06	
	$(-38.2, -193.3) \times 10^{-6}$	$(1.5 \times 10^{-6}, 36.6, \infty)$	0	0.17	
		$(2.1 \times 10^{-6}, 27.4, -)$	934.78	0.12	
	$\hat{\rho} = 1672 \times 10^{-6}$	-6121.9	$(2.8 \times 10^{-6}, 23.1, 0.41)$	702.26	0.09
	$C = -6168628.5$	$\hat{\sigma}_{\tilde{Z}}^2 = 0.29 \times 10^{-6}$	$(2.8 \times 10^{-6}, 23.1, 0.41)$	702.26	0.09
Capparis	(0.03, 0.004)	(0.25, 69.8, ∞)	5012.70	0.28	
	-19693.0	(0.43, 43.1, -)	5223.48	0.18	
	$\hat{\sigma}_{\exp\tilde{Z}}^2 = 4.84 \times 10^{-6}$	(0.76, 48.2, 0.22)	5342.78	0.11	
		(0.59, 49.7, 0.26)	5361.99	0.11	
	$(193.2, 24.8) \times 10^{-6}$	$(10.0 \times 10^{-6}, 70.2, \infty)$	0	0.29	
		$(15.0 \times 10^{-6}, 48.7, -)$	285.51	0.21	
	$\hat{\rho} = 6598 \times 10^{-6}$	-19700.1	$(28.8 \times 10^{-6}, 51.6, 0.21)$	466.02	0.12
	$C = -5089810.54$	$\hat{\sigma}_{\tilde{Z}}^2 = 4.06 \times 10^{-6}$	$(28.8 \times 10^{-6}, 51.6, 0.21)$	466.02	0.12