Conditionally heteroscedastic intensity-dependent marking of log Gaussian Cox processes

Mari Myllymäki and Antti Penttinen
University of Jyväskylä, Finland

19th October 2007

Abstract

Spatial point processes are models for systems of points randomly distributed in space. If the points are heterogeneous with respect to size, environment or neighbourhood, for example, this variation is described by measured quantities called marks. A general framework for statistical analysis of such systems of random points and marks is based on marked point processes. This study deals with marking, methods of constructing marked point processes from unmarked ones. Two special cases, independent and geostatistical markings, are among the known simple examples of marking strategies and are often used in practice. However, these markings are not able to model density-dependence of marks, the case where the local point intensity affects the mark distribution. This study develops new marking models of such a generality that not only the mean of the mark distribution but also its variance is affected by the local intensity. The new models are employed for marking of the log Gaussian Cox process and their theoretical mean, variance and mark correlation properties are presented. The most important part is statistical inference for such heteroscedastic marked point processes. The performance of the suggested estimation methods are studied in a simulation experiment. A tropical rainforest data is modelled using the developed models and statistical methods.

Keywords: Marked point process, density-dependence, heteroscedastic, estimation, tropical rainforest.

1 Introduction

Models and statistical methods are developed for spatial data consisting of random locations of objects and their properties such as size, age, species or growth. The main focus is in *density-dependence* which means that the local variation in the intensity (or density) of objects affects their properties. A motivation has been the modelling of tropical rainforest structures but the new models are widely applicable to the analysis of marked spatial point pattern data in other areas as well.

Spatial marked point processes are models for collections of objects located randomly in the space $\mathbb{R}^d$ or in its bounded subset, denoted here by W, and the objects are provided with

random variables. The object locations are called 'points' and the associated variables 'marks'. In what follows, planar marked point processes with $d = 2$ are considered. The mathematical notation used here is $N_m = \{[x_i; m(x_i)]\}$ where x_i stands for a point location and $m(x_i)$ for the mark associated with x_i. The marks typically describe properties of the objects but, in addition to that, more general constructions are allowed. The marks can be 'environmental' describing heterogeneity of the environment of the points, or they can be so-called 'constructed' marks, derived from the point pattern alone such as the distance to the nearest neighbour of a point. The marks are qualitative if they obtain values in a finite set of labels. Such a marked point process is also called multivariate. If the marks are real-valued they are called quantitative, which case is the focus of this study.

The descriptive statistical analysis of marked point patters with quantitative marks relies most often on first- and second-order mark characteristics and their empirical counterparts derived from the data, called summaries. Especially the second-order summaries are central tools in modelling, both in the identification of models and in the evaluation of goodness of fit of the models fitted to data.

The first-order characteristics of marks are based on conditional distributions of a mark $m(x_i)$ given that 'there is a point at x_i', for which mathematical theory (as Palm probabilities) exists. Of interest are the mean and variance of the mark distributions. The second-order characteristics are various *mark correlation functions* and are based on the conditional (two-point Palm) distributions of two marks $m(x_i)$ and $m(x_j)$ on condition that 'there are points at the locations x_i and x_j'. The most important second-order characteristics are Stoyan's mark correlation function $k_{mm}(r)$, $E(r)$-function, $V(r)$-function and mark variogram $\gamma_m(r)$. These characteristics will be explained later. A presentation of mark characteristics can be found in Schlather (2001), see also Stoyan et al. (1995), Stoyan and Stoyan (1994) and Illian et al. (2008).

The focus of this study is to construct flexible marked point process models starting from an unmarked point process $N = \{x_n\}$ and providing each point $x_i \in N$ by a mark $m(x_i)$. This procedure is called 'marking' in the sequel. The two well-known examples are *independent marking*, where the marks are drawn independently from a probability distribution, and *geo-statistical marking* (Mase, 1996; Schlather et al., 2004). In the latter the marks are generated by a random field which is independent of the point locations. Both of these markings cannot model *intensity-dependent* marks, the case where the distribution of a mark varies along the point density.

A step forward in marking of stationary unmarked point processes is to allow the distribution of marks to be dependent on the local intensity as suggested by Menezes (2005), Ho and Stoyan (2007) and Myllymäki (2006), all for the stationary log Gaussian Cox process with random intensity $\{\Lambda(s)\}$, see Møller et al. (1998) and Møller and Waagepetersen (2004). This *intensity-dependent* marking assumes conditional independence of marks given the random intensity of the point process. In these models the mean of the conditional distribution of $m(x_i)$, on condition that there is a point at x_i and the intensity $\Lambda(x_i)$ is known, depends on $\Lambda(x_i)$ while the variance of this conditional distribution is constant. Heuristically, these models allow the marks to be large (small) in areas of low point intensity and small (large) in areas of high intensity. Intensity-dependent marking leads to correlation of marks which is affected by the second-order property of the unmarked Cox process.

The existing intensity-dependent markings are useful models but assume that the variance of

the conditional mark distribution does not depend on the point intensity. This can be a severe restriction as in the following example.

Figure 1 presents a marked point pattern of the species *T. tuberculata* (in 1990). The data originates from a tropical rainforest. It is a 200m×200m part of the 50 ha Forest Dynamics Plot at Barro Colorado Island, Panama, see Hubbell and Foster (1983), Condit et al. (1996) and Condit (1998). The marks of *T. tuberculata* are the diameters of trees at breast height (dbh). In Figure 1 the measured diameters are plotted against the estimated local tree density at the tree locations. It reveals that in regions with low intensity the marks are larger on an average but also the variance is larger. On the contrary, in areas densely occupied by trees the marks are small and variation around the average is small. The existing density-dependence models are inadequate to capture this feature. This calls for a need to introduce new *conditionally heteroscedastic* marking models.

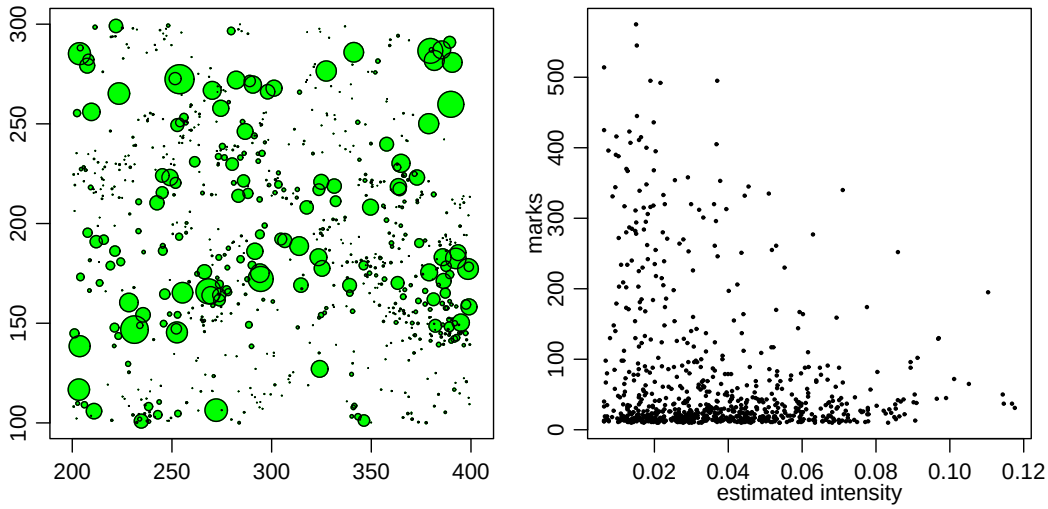

Figure 1: On the left: A 200m×200m -subplot of *T. tuberculata* trees of the BCI data in 1990. The size of a circle is proportional to the dbh of the tree. On the right: The diameters of trees plotted against the estimated tree densities at the tree locations.

Intensity-dependent marking presumes the existence of local variation in the intensity and that is why only clustered or heterogeneous point process models are relevant. This work uses the (stationary) log Gaussian Cox process as an unmarked point process model as do Menezes (2005) and Ho and Stoyan (2007). It is a flexible and very popular model, suitable for our purposes. The new marking schemes are extensions to the ones in Ho and Stoyan (2007) who suggest the conditional mark distribution

$$m(x_i)|\Lambda(x_i) \sim N(a + b\Lambda(x_i), d^2)$$

with constant variance d^2 independent of $\Lambda(x_i)$.

The first generalisation is to modify this model in a way that also the conditional mark variance depends on $\Lambda(x_i)$. The second model is intended for modelling negative dependence between the marks and the intensity and is based on normal distribution. These constructions lead to conditionally heteroscedastic normal models. The normal model is further replaced by the

gamma family of distributions including the exponential model

$$m(x_i)|\,\Lambda(x_i) \sim \mathrm{Exp}\left(\frac{1}{a + b/\Lambda(x_i)}\right).$$

Mark characteristics are derived analytically for each marked point process model considered and procedures for statistical inference are suggested. The parameters of the unmarked log Gaussian Cox process are estimated using the minimum contrast method recommended by Møller et al. (1998) and further with approximative maximum likelihood method suggested by Tanaka et al. (2007). Keeping the parameters of the log Gaussian Cox process fixed, estimates for the parameters related to the marks are obtained using minimum contrast method applied to the mark variogram. Two further estimation methods with weaker assumptions are suggested. The performance of the models and inference methods are evaluated by simulation experiments and the new models are applied to the marked point pattern of *T. tuberculata*.

The rest of the paper is arranged in the following way. Preliminaries of marked point processes and mark characteristics are outlined in Section 2. Previous marking strategies are recalled in Section 3 and new marking techniques for conditionally heteroscedastic intensity-dependent marks are introduced in Section 4. Statistical inference of the new models is considered in Section 5. Further, the new models are studied by simulation experiments in Section 6. An application to the marked point pattern of *T. tuberculata* in a tropical rainforest is given in Section 7. Section 8 is for discussion.

2 Marked point process preliminaries

Spatial point processes are models for random point patterns in (a subset of) $\mathbb{R}^d$. They are collections of point coordinates $x_i \in \mathbb{R}^d$ denoted by

$$N = \{x_i\}.$$

In this work we assume that N is a stationary and isotropic simple point process. Stationarity and isotropy mean that $N_s = \{x_i + s\}$ and $\mathbf{r}N = \{\mathbf{r}x_i\}$ have the same distribution as N for any $s \in \mathbb{R}^d$ and any (Euclidean) rotation $\mathbf{r}$ around the origin, respectively. 'Simple' means that multiple points are not allowed: $x_i \neq x_j$ for $i \neq j$.

A widely used point process class for aggregated or clustered point patterns is the Cox process, also so called 'doubly stochastic Poisson process', see e.g. Diggle (2003), Møller and Waagepetersen (2004) or Stoyan et al. (1995). Each realisation of a Cox process is an inhomogeneous Poisson process with an intensity function $\{\lambda(s)\}$, a realisation of the random intensity $\{\Lambda(s)\}$ of the process. In this work we use the (stationary and isotropic) *log Gaussian Cox process* $N = \{x_i\}$, which is a Cox process with a log Gaussian random intensity $\{\Lambda(s)\}$:

$$\Lambda(s) = \exp\{Z(s)\},$$

where $\{Z(s)\}$ is a (stationary and isotropic) Gaussian random field with mean μ_Z, variance σ_Z^2 and (positive definite) covariance function $C_Z(r)$, see Møller et al. (1998) and Møller and Waagepetersen (2004). The mean and the covariance function of $\{Z(s)\}$ characterize the distribution of $\{Z(s)\}$ and thus completely define the distribution of the log Gaussian Cox process N.

Alternatively, the process N is determined by the intensity and the pair correlation function, which are

$$\lambda = \exp\left\{\mu_Z + \frac{1}{2}\sigma_Z^2\right\}$$

and

$$g(r) = \exp\{C_Z(r)\},$$

respectively, see Møller et al. (1998). The intensity $\lambda = \mathbb{E}(\Lambda(s))$ tells the mean number of points per unit area or unit volume and, intuitively, $\lambda^2 g(\|o - r\|)\mathrm{d}o\,\mathrm{d}r$ is the probability that two infinitesimal disjoint regions of volumes $\mathrm{d}o$ and $\mathrm{d}r$ both contain exactly one point of N. For further details see e.g. Stoyan et al. (1995) or Stoyan and Stoyan (1994).

The main scope of this work is in marked point processes with quantitative marks. A stationary and isotropic simple marked point process is obtained when each point x_i of N is provided by a real-valued measurement $m(x_i)$ called mark. The marked point process is denoted by

$$N_m = \{[x_i; m(x_i)]\}.$$

Marking strategies are discussed in Sections 3 and 4. In what follows, we recall first- and second-order characteristics that are relevant for our purposes.

Various characteristics have been introduced to describe the properties of marked point processes, see Stoyan and Stoyan (1994), Stoyan et al. (1995), Schlather (2001) and Schlather et al. (2004). Mark characteristics are conditional quantities (in the Palm sense): let $\mathbb{E}_x$ and var_x stand for the conditional expectation and variance, respectively, given that there is a point of N at the location x. Further $\mathbb{E}_{xy}$ refers to the conditional expectation given there are two points of N at locations x and y. Because of stationarity and isotropy, it suffices to consider expectations $\mathbb{E}_o$ and $\mathbb{E}_{or}$ with $\|\mathbf{r}\| = r$.

The mean mark $\mu_m = \mathbb{E}_o(m(o))$, mark variance $\sigma_m^2 = \mathrm{var}_o(m(o))$ and mark distribution function $F_{\mathcal{M}}(m)$ are first-order characteristics of marks and conditional on 'there is a point of N at o'. Mathematical definition of these characteristics is the following. For any Borel sets B and L in $\mathbb{R}^d$ and $\mathbb{R}$, respectively, let $N_m(B \times L)$ stand for the number of points in B with mark in L. The corresponding mean number of points satisfies (Stoyan et al., 1995)

$$\mathbb{E}\{N_m(B \times L)\} = \lambda \nu_d(B)\mathcal{M}(L), \tag{1}$$

where ν_d denotes the d-dimensional Lebesque measure and $\mathcal{M}$ is the mark distribution. For real-valued marks $\mathcal{M}$ is described by the mark distribution function $F_{\mathcal{M}}(m)$, $F_{\mathcal{M}}(m) = \mathcal{M}((-\infty, m])$ for $-\infty < m < \infty$. The mean and variance of $F_{\mathcal{M}}(m)$ are μ_m and σ_m^2, respectively.

Several second-order summary characteristics describe the variability and correlations of marks. For stationary and isotropic processes these characteristics are functions of the distance r. The value of each characteristic at r is an expectation $\mathbb{E}_{or}\{f(m(o), m(\mathbf{r}))\}$ with $\|\mathbf{r}\| = r$ and some function f. Mathematical definitions can be found in Schlather (2001) and Stoyan and Stoyan (1994, p. 262), for example. Schlather (2001) extends the definitions for arbitrary real-valued marks, as they have been defined earlier only for positive marks.

The first measure of correlation of marks is Stoyan's $k_{mm}(r)$-function, also called 'mark correlation function' in literature. Stoyan's unscaled $\kappa_{mm}(r)$-function is heuristically the mean of the product of the marks at two points of N being distance r apart (Stoyan, 1984),

$$\kappa_{mm}(r) = \mathbb{E}_{or}\{m(o)m(\mathbf{r})\} \quad \text{for } r > 0.$$

Stoyan's $k_{mm}(r)$-function is a scaled version of $\kappa_{mm}(r)$:

$$k_{mm}(r) = \frac{\kappa_{mm}(r)}{\mu_m^2} \quad \text{for } r > 0,$$

where $\mu_m \neq 0$. Stoyan's $k_{mm}(r)$-function is not compatible with the classical definition of correlation, but if the $\kappa_{mm}(r)$-function is scaled differently, functions that correspond to the classical definitions of covariance and correlation are obtained.

The conditional expectation and variance of a mark given that there is a further point of the process a distance r apart are

$$E(r) = \mathbb{E}_{or}\{m(o)\} \quad \text{for } r > 0$$

and

$$V(r) = \mathbb{E}_{or}\{[m(o) - E(r)]^2\} \quad \text{for } r > 0,$$

respectively. Moreover, the *mark covariance* and *mark correlation* functions are

$$C_m(r) = \kappa_{mm}(r) - [E(r)]^2 \quad \text{for } r > 0$$

and

$$\rho_m(r) = \frac{\kappa_{mm}(r) - [E(r)]^2}{V(r)} \quad \text{for } r > 0,$$

respectively. The mark correlation and mark covariance functions carry slightly different information than Stoyan's $k_{mm}(r)$-function since $E(r)$ and $V(r)$, appearing in $C_m(r)$ and $\rho_m(r)$, are not necessarily constant for a marked point process. For example, Schlather (2001) shows that there exist two different marked point processes with the same $k_{mm}(r)$-function but different $\rho_m(r)$.

In geostatistics, the variogram is often used in estimation instead of the covariance function, see for example Cressie (1993). In point process statistics, a variogram can be defined for marks, but the *mark variogram*

$$\gamma_m(r) = \frac{1}{2}\mathbb{E}_{or}\{[m(o) - m(\mathbf{r})]^2\} \quad \text{for } r > 0$$

coincides with the geostatistical variogram only in specific situations such as independent marking and geostatistical marking, which is considered by Wälder and Stoyan (1996) and Stoyan and Wälder (2000). In fact $\gamma_m(r) = V(r) - C_m(r)$.

Schlather (2001) extends the above definitions to include the case $r = 0$ such that $k_{mm}(0) = \mathbb{E}_o[(m(o))^2]/\mu_m^2$, $E(0) = \mu_m$, $V(0) = \sigma_m^2$ and $\gamma_m(0) = 0$.

3 Marking of a point process

Traditional marking procedures, independent marking and geostatistical marking, assume that the marks are independent of the unmarked point process.

Independent marking

The marks are independently sampled from a mark distribution $F_{\mathcal{M}}(m)$ and are independent of the underlying point process. In many cases an independently marked point process is too simple and, in general, the model serves as a reference process. For the independently marked point process $k_{mm}(r) \equiv 1$, $E(r) \equiv \mu_m$, $V(r) \equiv \sigma_m^2$ and $\gamma_m(r) \equiv \sigma_m^2$ for $r > 0$. Especially, independent marking does not generate any dependence between the marks.

Geostatistical marking

An unmarked point process $N = \{x_i\}$ is marked by a stationary random field $\{U(s)\}$ being independent of N:

$$m(x_i) = U(x_i) \quad \text{for } x_i \in N.$$

The marks inherit the dependence structure from the random field. Independence between the random field and N implies that the mark variogram coincides with the geostatistical variogram of the random field $\{U(s)\}$ (Wälder and Stoyan, 1996; Stoyan and Wälder, 2000). The characteristics $E(r)$ and $V(r)$ are constants (Schlather et al., 2004), which rules are useful in testing whether the marking can be considered geostatistical.

In these two marking strategies N does not affect the distributional properties of the marks. An important situation, where independent or geostatistical marking are not valid strategies, is where the point density affects the marks, called density-dependence in plant ecology.

Point density is described by the intensity function $\{\lambda(s)\}$ for an inhomogeneous point process, or by the random intensity function $\{\Lambda(s)\}$ for a Cox process. In the latter case, stationarity implies that $\mathrm{E}(\Lambda(s))$ is constant but the point intensity in its realisations is defined by $\Lambda(s) = \lambda(s)$, a realisation of the random field.

It is assumed in the following that the marks are conditionally independent given the intensity function. A consequence is that the mark distribution of $m(x_i)$ depends on $\lambda(x_i)$ for an inhomogeneous point process and, conditional on $\{\Lambda(s)\}$, on $\Lambda(x_i)$ for a Cox process. Density-dependence of marks is considered in this generality.

Geostatistical model for preferential sampling

Menezes (2005) develops a model for geostatistical sampling where the sampling design depends on the observed random field. The sampling points and the values of the observed random field can be considered as a marked point pattern, but the interest is in fact in the random field, which can be observed wherever the investigator likes to study its properties.

The model for point locations is a log Gaussian Cox process $N = \{x_i\}$ with the intensity

$$\Lambda(s) = \exp\{\alpha + \beta Z(s)\},$$

where α and β are real parameters and $\{Z(s)\}$ is a Gaussian random field with mean μ_Z, variance σ_Z^2 and covariance function $C_Z(r)$. The intensity and the pair-correlation function of this log Gaussian Cox process are

$$\lambda = \exp\{\alpha + \beta\mu_Z + \beta^2\sigma_Z^2/2\}$$

and

$$g(r) = \exp\{\beta^2 C_Z(r)\} \quad \text{for } r \geq 0,$$

respectively (Ho and Stoyan, 2007). The marks of the model are defined by

$$m(x_i) = Z(x_i) + \epsilon(x_i) \quad \text{for } x_i \in N, \tag{2}$$

where the $\epsilon(x_i)$s are i.i.d. Gaussian random errors independent of $\{Z(s)\}$. The marks (2) can be considered as a noisy version of the underlying random field $\{Z(s)\}$ that is of interest. Clearly, the marks are not in general independent of the underlying point process and this is seen through the mark characteristics, which are studied in Ho and Stoyan (2007). For example $E(r)$ is not constant for this model.

Intensity-marked Cox process

Ho and Stoyan (2007) consider an intensity-marked model, in which point intensity and mark sizes are closely coupled. The log Gaussian Cox process is considered as a model for the points and, conditional on $\Lambda(x_i)$, the marks are provided by

$$m(x_i) = a + b\Lambda(x_i) + \epsilon(x_i) \quad \text{for } x_i \in N, \tag{3}$$

where a and b are model parameters and the $\epsilon(x_i)$s are i.i.d. Gaussian random errors with variance τ^2 being independent of $\{\Lambda(s)\}$. The marking (3) corresponds to

$$m(x_i)|\,\Lambda(x_i) \sim N\left(a + b\Lambda(x_i), \tau^2\right), \quad \text{for } x_i \in N,$$

and such is a special case of the Gaussian intensity-marked Cox process considered in Section 4.1. The dependence between marks and points affects mark characteristics, see Ho and Stoyan (2007) and Illian et al. (2008).

Log-intensity marked Cox process

A simple regression type model considered by us, independently of Menezes (2005) and Ho and Stoyan (2007), is the following. Let the log Gaussian Cox process serve as a model for the unmarked point pattern and provide the marks by

$$m(x_i) = a + bZ(x_i) + U(x_i) \quad \text{for } x_i \in N. \tag{4}$$

Here $\{U(s)\}$ is an additional zero-mean Gaussian random field which is independent of N and $\{Z(s)\}$. The 'error random field' $\{U(s)\}$ is allowed to be auto-correlated having covariance function $C_U(r)$. The use of correlated errors allows the mark structure to differ from the structure of the random field $\{Z(s)\}$. The random field $\{U(s)\}$ can be interpreted as the influence of unobserved variables affecting the marks but not the intensity.

This construction is related to independent marking in the residual sense. Indeed, the marking of $\{x_i\}$ by the random field $\{U(s)\}$ leads to the geostatistically marked point process $\{[x_i; U(x_i)]\}$. This mark is then modulated by the intensity through $\{Z(s)\}$ leading to intensity-dependence. An interesting feature of this marking is that it combines two sources of variation into the marking. Heuristically, if the marks are explained by the intensity-generated process, then the residuals $U(x_i) = m(x_i) - \alpha - \beta Z(x_i)$ may still contain autocorrelation. This is modelled using an external (and independent) random field.

Difficulties are met in fitting of the model with correlated error field $\{U(s)\}$ due to parameter confounding. Assuming that $\{U(s)\}$ is a noise field, the parameters of the model can be estimated in a similar way as for the models introduced in Section 4. If good estimates for the

intensities $\Lambda(x_i)$, $x_i \in N$, exist, then the estimation of the model with correlated error field may be conducted stepwise.

The mark distribution of the mark (4) is normal with mean

$$\mu_m = a + b(\mu_Z + \sigma_Z^2)$$

and variance

$$\sigma_m^2 = b\sigma_Z^2 + \sigma_U^2,$$

where $\sigma_U^2 = C_U(0)$. The characteristic $E(r)$ of the marking (4) is not constant,

$$E(r) = \mu_m + bC_Z(r) \quad \text{for } r > 0,$$

whilst $V(r) \equiv \sigma_m^2$. Further, Stoyan's unscaled $\kappa_{mm}(r)$-function and the mark variogram are

$$\kappa_{mm}(r) = b^2 C_Z(r) + C_U(r) + (\mu_m + bC_Z(r))^2 \quad \text{for } r > 0$$

and

$$\gamma_m(r) = b^2\sigma_Z^2 + \sigma_U^2 - b^2 C_Z(r) - C_U(r) \quad \text{for } r > 0.$$

4 Conditionally heteroscedastic modelling of marks

The intensity-marked point processes of Section 3 are reasonable models for marked point patterns, where the mark size depends on the point density. The conditional mark variance, however, is assumed not to depend on the point density. Thus, these models fail to model data where not only the mean but also the variance of the marks varies along the point density. This kind of phenomenon is seen in the marked point pattern of *T. tuberculata*, see Figure 1, for example. In this section models for conditionally heteroscedastic intensity-dependent marks are introduced.

The log Gaussian Cox process serves as a model for the unmarked point pattern and the marks are created such that they depend on the point density. The general setting is the following: Let $N = \{x_i\}$ be the log Gaussian Cox process in $W \subseteq \mathbb{R}^d$ and provide the (real-valued) marks by

$$m(x_i)|\,\Lambda(x_i) \sim F_{m|\Lambda}(\cdot|\Lambda(x_i)), \quad \text{for } x_i \in N,$$

where $F_{m|\Lambda}$ is the parametric conditional distribution of a mark $m(x_i)$ at (fixed) $x_i \in N$ given the intensity $\Lambda(x_i)$. The marks are assumed to be conditionally independent given the intensity.

Transformations that create negative and positive dependence between the marks and the intensity are considered, although the emphasis is in the negative dependence since its more interesting for our applications. In the following, we consider the models with marks provided by

$$m(x_i)|\,\Lambda(x_i) \sim N\left(a + b\Lambda(x_i), c^2\Lambda(x_i) + d^2\right) \quad \text{for } x_i \in N, \tag{5}$$

$$m(x_i)|\,\Lambda(x_i) \sim N\left(a + b/\Lambda(x_i), c^2/\Lambda(x_i) + d^2\right) \quad \text{for } x_i \in N, \tag{6}$$

$$m(x_i)|\,\Lambda(x_i) \sim \text{Gamma}\left(\alpha, \frac{1}{a + b/\Lambda(x_i)}\right) \quad \text{for } x_i \in N \tag{7}$$

and

$$m(x_i)|\,\Lambda(x_i) \sim \text{Exp}\left(\frac{1}{a + b/\Lambda(x_i)}\right) \quad \text{for } x_i \in N. \tag{8}$$

The exponential model (8) is a special case of the Gamma model (7) with shape parameter $\alpha = 1$. Here a, b, c^2 and d^2 are model parameters describing the relationship between the marks and the intensity, and α is a shape parameter of the gamma distribution.

Derivation of the mark characteristics, that are introduced here, can be found in Appendix.

4.1 Gaussian intensity-marked Cox process

Let $N = \{x_i\}$ be the log Gaussian Cox process in $W \subseteq \mathbb{R}^d$ and provide the marks by (5). The marking (5) with positive parameters a, b, c^2 and d^2 creates positive dependence between the marks and the intensity. This means that the marks are larger and there is more variation in marks in regions with high point density than in regions with low point density. The model is a modification of the intensity-marked Cox process considered by Ho and Stoyan (2007), the marking of which equals (5) with $c = 0$.

The mean mark μ_m, $E(r)$ and $k_{mm}(r)$, for $r > 0$, are not affected by the parameter c and are the same for the markings (3) and (5). Instead, there is an increase in the mark variance due to the term $c^2\Lambda(x_i)$ in (5):

$$\sigma_m^2 = b^2\lambda^2 e^{2\sigma_Z^2}\left(e^{\sigma_Z^2} - 1\right) + c^2\lambda e^{\sigma_Z^2} + d^2$$

and

$$\kappa_{mm}(0) = a^2 + 2ab\lambda e^{\sigma_Z^2} + b^2\lambda^2 e^{3\sigma_Z^2} + c^2\lambda e^{\sigma_Z^2} + d^2.$$

The second-order characteristics $V(r)$ and $\gamma_m(r)$ also change. For the marking (5) these are

$$V(r) = d^2 + c^2\lambda e^{\sigma_Z^2 + C_Z(r)} + b^2\lambda^2 e^{2\sigma_Z^2 + 2C_Z(r)}\left(e^{\sigma_Z^2} - 1\right) \quad \text{for } r > 0$$

and

$$\gamma_m(r) = d^2 + c^2\lambda e^{2\sigma_Z^2 + C_Z(r)} + b^2\lambda^2 e^{2\sigma_Z^2 + 2C_Z(r)}\left(e^{\sigma_Z^2} - e^{C_Z(r)}\right) \quad \text{for } r > 0.$$

Note that the second-order mark characteristics depend on the unmarked point process through the covariance function $C_Z(r)$ of the random field $\{Z(s)\}$ that generates the intensity.

From our point of view, more interesting is the model which creates negative dependence between the marks and the intensity. The first attempt is to assume that $b < 0$ in (3) or (5). However, if $C_Z(r)$ is a decreasing function, $V(r)$ of the models is decreasing regardless of the sign of b. In addition, having the log Gaussian Cox process as a model for point locations, the resulting mark distribution tends to be negatively skew (longer left tail) if $b < 0$. Instead, in many ecological applications, the variance of marks tends to be smaller in regions where the marks are smaller and mark distributions are often positively skew (longer right tail). For example, the empirical $E(r)$ and $V(r)$, estimated from the marked point pattern of *T. tuberculata*, shown in Figure 2, both increase along r.

Thus, further models are needed to create negative dependence between the marks and the intensity. An idea is to consider the conditional mean mark and mark variance linearly dependent on $1/\Lambda(x_i)$. The marking (6) with $a, b, c^2, d^2 > 0$ produces marks that are smaller in mean

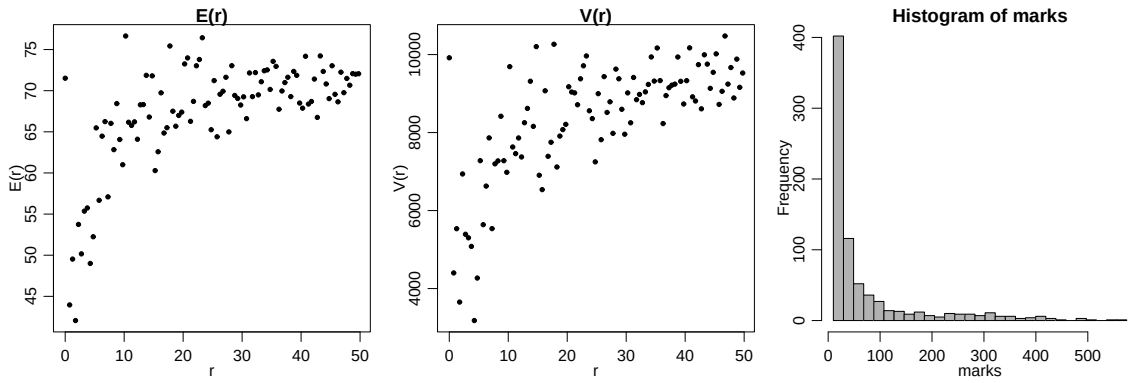

Figure 2: The $E(r)$ and $V(r)$ summaries calculated from the marked point pattern of *T. tuberculata*, and the histogram of the marks (dbh in mm) of *T. tuberculata*.

and vary less in regions with high point density. If $c = 0$ only the conditional mean of marks depends on the intensity and if $b = 0$ only the variance.

The mean mark and mark variance of the mark (6) are

$$\mu_m = a + \frac{b}{\lambda} \quad \text{and} \quad \sigma_m^2 = b^2 e^{-2\mu_Z}\left(1 - e^{-\sigma_Z^2}\right) + \frac{c^2}{\lambda} + d^2. \tag{9}$$

Assuming that $C_Z(r)$ is a decreasing function, the characteristics

$$E(r) = a + \frac{b}{\lambda} e^{-C_Z(r)} \quad \text{for } r > 0$$

and

$$V(r) = d^2 + \frac{c^2}{\lambda} e^{-C_Z(r)} + \frac{b^2}{\lambda^2}\left(e^{\sigma_Z^2} - 1\right) e^{-2C_Z(r)} \quad \text{for } r > 0$$

increase along r. The Stoyan's $k_{mm}(r)$-function

$$k_{mm}(r) = \begin{cases} 1 + \dfrac{b^2 e^{-2\mu_Z}\left(1 - e^{-\sigma_Z^2}\right) + \frac{c^2}{\lambda} + d^2}{\left(a + \frac{b}{\lambda}\right)^2}, & \text{if } r = 0, \\ \dfrac{a^2 + \left(\frac{2ab}{\lambda} + \frac{b^2}{\lambda^2}\right) e^{-C_Z(r)}}{\left(a + \frac{b}{\lambda}\right)^2}, & \text{if } r > 0, \end{cases}$$

indicates that at short inter-point distances the marks tend to be small and the mark variogram

$$\gamma_m(r) = d^2 + \left(\frac{c^2}{\lambda} - \frac{b^2}{\lambda^2}\right) e^{-C_Z(r)} + \frac{b^2}{\lambda^2} e^{\sigma_Z^2 - 2C_Z(r)} \quad \text{for } r > 0 \tag{10}$$

also implies positive association of the marks at short inter-point distances.

Although similar second-order characteristics as the second-order summaries of *T. tuberculata* can be obtained by marking (6) (with specific parameters), this model is not suitable for modelling the marked point pattern of *T. tuberculata* since the marks provided by (6) are not necessarily positive. The marginal mark distribution is not similar to the distribution of dbh.

The marking (5) adds intensity-dependent conditional mark variance to the intensity-marked Cox process with marking (3). Differently, conditionally heteroscedastic marks are obtained by the logarithmic transformation, for example: If $\log(m(x_i))$s are provided by (3) then both $\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)]$ and $\mathrm{var}_{x_i}[m(x_i)|\Lambda(x_i)]$ depend on $\Lambda(x_i)$.

4.2 Inverse gamma intensity-marked Cox process

In our applications the model assumption

$$\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)] = a + b/\Lambda(x_i),$$

where $x_i \in N$, seems reasonable. An idea is to construct the marks according to

$$m(x_i)|\,\Lambda(x_i) \sim \mathrm{Exp}\left(\frac{1}{a + b/\Lambda(x_i)}\right) \quad \text{for } x_i \in N.$$

Consequently, the conditional mark variance is

$$\mathrm{var}_{x_i}[m(x_i)|\Lambda(x_i)] = (a + b/\Lambda(x_i))^2$$

and thus it depends on the local intensity. Since the exponential distribution is a special case of the gamma distribution with the shape parameter $\alpha = 1$, we consider the gamma model.

Let again $N = \{x_i\}$ be the log Gaussian Cox process in $W \subseteq \mathbb{R}^d$. Marks of the *inverse gamma intensity-marked Cox process* are provided by (7):

$$m(x_i)|\,\Lambda(x_i) \sim \mathrm{Gamma}\,(\alpha, \beta) \quad \text{for } x_i \in N$$

with the shape parameter α and the rate parameter

$$\beta = \frac{1}{a + b/\Lambda(x_i)}.$$

Consequently, the conditional mean mark is

$$\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)] = \alpha/\beta$$

and the mark variance

$$\mathrm{var}_{x_i}[m(x_i)|\Lambda(x_i)] = \alpha/\beta^2.$$

Thus the construction with $a, b > 0$ creates negative dependence between the marks and the point density such that both the conditional mean and variance of the mark $m(x_i)$ depend on the intensity. Positive dependence can be studied by the model with $\beta = 1/(a + b\Lambda(x_i))$, for example. Here we restrict ourselves to the case with negative dependence. The *inverse exponential intensity-marked Cox process* is obtained as a special case with $\alpha = 1$.

The intensity-dependence of the marks is seen through the mark characteristics. The mean mark and the mark variance of the mark (7) are

$$\mu_m = \alpha\left(a + \frac{b}{\lambda}\right) \quad \text{and} \quad \sigma_m^2 = \alpha\left(a^2 + \frac{2ab}{\lambda} + b^2 e^{-2\mu_Z}\right) + \alpha^2 b^2 e^{-2\mu_Z}\left(1 - e^{-\sigma_Z^2}\right).$$

Further,

$$E(r) = \alpha\left(a + \frac{b}{\lambda}e^{-C_Z(r)}\right) \quad \text{for } r > 0,$$

$$V(r) = \alpha\left(a^2 + \frac{2ab}{\lambda}e^{-C_Z(r)} + \frac{b^2}{\lambda^2}e^{\sigma_Z^2 - 2C_Z(r)}\right) + \alpha^2\frac{b^2}{\lambda^2}e^{-2C_Z(r)}\left(e^{\sigma_Z^2} - 1\right) \quad \text{for } r > 0,$$

$$k_{mm}(r) = \begin{cases} 1 + \dfrac{\alpha\left(a^2 + \frac{2ab}{\lambda} + b^2\right) + \alpha^2 b^2\left(1 - \frac{1}{\lambda^2}\right)}{\alpha^2\left(a + \frac{b}{\lambda}\right)^2}, & \text{if } r = 0, \\ \dfrac{a^2 + \left(\frac{2ab}{\lambda} + \frac{b^2}{\lambda^2}\right)e^{-C_Z(r)}}{\left(a + \frac{b}{\lambda}\right)^2}, & \text{if } r > 0 \end{cases}$$

and

$$\gamma_m(r) = \alpha\left(a^2 + \frac{2ab}{\lambda}e^{-C_Z(r)} + \frac{b^2}{\lambda^2}e^{\sigma_Z^2 - 2C_Z(r)}\right) + \alpha^2\frac{b^2}{\lambda^2}e^{-C_Z(r)}\left(e^{\sigma_Z^2 - C_Z(r)} - 1\right) \quad \text{for } r > 0. \quad (11)$$

The characteristics of the marking (8) are obtained by setting $\alpha = 1$ in the equations above.

Theoretical characteristics above show that the inverse gamma intensity-marking, together with the log Gaussian Cox process, is able to generate marked point patterns with clusters of points having small marks, but simultaneously allow large variation for the marks.

5 Statistical inference

The likelihood is complex even for the unmarked log Gaussian Cox process. That is why practical methods which lead to tractable estimation equations are needed for the intensity-marked Cox processes. We suggest, that first the parameters of the log Gaussian Cox process are estimated with existing methods. Second, if the log Gaussian Cox process assumption is reasonable for the data, the parameters of the marking equation are estimated utilizing this information of the underlying point process model. Otherwise, the parameters related to the marks can be estimated weakening the assumptions. These methods can also be exploited in preliminary analysis investigating intensity-dependence of the marks.

5.1 Parameter estimation for the unmarked log Gaussian Cox process

Two parameter estimation methods for the unmarked log Gaussian Cox process are recalled. These are of use in the study of the assumption on the process of point locations. Further, the estimates are used in the estimation of the parameters of the marking equations.

Minimum contrast method

Møller et al. (1998) suggests to base the inference of the log Gaussian Cox process on first- and second-order summary statistics, which determine the distribution of the process completely. As non-parametric estimators of intensity λ and covariance function $C_Z(r)$ Møller et al. (1998) advises to use

$$\hat{\lambda} = n/\nu(W) \quad (12)$$

and

$$\hat{C}_Z(r) = \log(\hat{g}(r)), \tag{13}$$

respectively, where $\hat{g}(r)$ is an estimator of the pair-correlation, see for example Stoyan and Stoyan (1994, p. 284). Let $C_Z(r;\theta)$ be a parametric model for the covariance function, for example, the exponential covariance function

$$C_Z(r;\theta) = \sigma_Z^2 \exp\{-r/\phi_Z\}, \tag{14}$$

where σ_Z is the variance and ϕ_Z so-called range parameter. Estimates for the parameters $\theta = (\sigma_Z^2, \phi_Z)$ can be found by minimizing the integral

$$\int_\epsilon^{a_0} [(\hat{C}_Z(r))^\eta - (C_Z(r;\theta))^\eta]^2 \mathrm{d}r \tag{15}$$

with respect to θ. Here $0 \leq \epsilon < a_0$ and $\eta > 0$ are parameters specified by the user. Small values of η give more weight to small values of r. After obtaining estimates $\hat{\lambda}$ and $\hat{\sigma}_Z^2$ for λ and σ_Z^2 by (12) and (15), an estimate for mean of the random field Z can be obtained by

$$\hat{\mu}_Z = \log(\hat{\lambda}) - \frac{\hat{\sigma}_Z^2}{2}. \tag{16}$$

In computation, the integral (15) is approximated by a Riemann sum.

Approximative maximum likelihood method

Tanaka et al. (2007) puts forward an approximative likelihood method for stationary and isotropic Neyman-Scott and similar point processes. The method is based on analysing the point process N_D consisting of all difference points $x - y$ for $x, y \in N$, $x \neq y$. The difference pattern N_D is central symmetric, since it contains both $x - y$ and $y - x$ for all $x, y \in N$, and its intensity function can be given in the terms of the intensity λ and the pair correlation $g(r)$ of the original point process N. Let $\Lambda_D(r)$ be the mean number of points of N_D in the ball of radius r centered at o. The derivative of $\Lambda_D(r)$ is

$$\lambda_D(r) = \lambda^2 g(r) \bar{\gamma}_W(r),$$

where $\bar{\gamma}_W(r)$ is the set covariance of the window W, see Stoyan and Stoyan (1994, p. 123). The estimation method of Tanaka et al. (2007) is based on the assumption that N_D is approximated by an inhomogeneous Poisson process with intensity function

$$\lambda_D(r;\theta) = \lambda^2 g(r;\theta) \bar{\gamma}_W(r),$$

where θ are the parameters given by the parametric model $g(r;\theta)$ for the pair correlation function of N. If N_D indeed is an inhomogeneous Poisson process, then the log likelihood is given by

$$\log L(\theta) = \sum_{x,y \in N \cap W,\, x \neq y} \log(\lambda_D(\|x - y\|;\theta) - \int_0^R \lambda_D(r;\theta) d b_d r^{d-1} \mathrm{d}r, \tag{17}$$

where d is the dimension, b_d is the volume of the d-dimensional unit ball and $R = \min\{r : \bar{\gamma}_W(r) = 0\}$.

We set a parametric model for pair correlation function of the log Gaussian Cox process by $g(r;\theta) = \exp(C_Z(r;\theta))$, where $C_Z(r;\theta)$ is (14) for example. Maximization of the approximative log likelihood (17) gives estimates for $\theta = (\sigma_Z^2, \phi_Z)$. Further parameters λ and μ_Z are estimated through (12) and (16).

5.2 Estimation of the parameters of the marking equations

Assuming the log Gaussian Cox process as a model for the unmarked point pattern, the parameters of the marking equations are estimated using the minimum contrast method applied to the mark variogram. In addition to that, two straightforward two-step estimation methods are studied. These assume that first the intensity of the point pattern is possible to estimate reliably, e.g. for large data. Then the parameters which enter through the marking equation are considered to be a regression type problem after conditioning by the intensity function and by the point locations.

In the following estimation methods are discussed with emphasis on the inverse gamma and Gaussian intensity-marked Cox processes with the markings (7) and (6), respectively.

Minimum contrast method

In the present context the parameters are estimated by minimizing the integral

$$\int_{\epsilon}^{a_0} \left[(\hat{\gamma}_m(r))^{\eta} - (\gamma_m(r))^{\eta}\right]^2 \mathrm{d}r. \tag{18}$$

with respect to the model parameters. Here $0 \leq \epsilon < a_0$ and $\eta > 0$ are parameters specified by the user. The empirical variogram $\hat{\gamma}_m(r)$ is estimated from data and its theoretical counterpart $\gamma_m(r)$ is known up to the model parameters. The latter relies on the log Gaussian Cox process and conditionally independent marking.

Inverse gamma intensity-marked Cox process. For the marking (7) the equation to be minimized is

$$\int_{\epsilon}^{a_0} \left[(\hat{\gamma}_m(r))^{\eta} - \left(\alpha\left(a^2 + \frac{2ab}{\hat{\lambda}} e^{-\tilde{C}_Z(r)} + \frac{b^2}{\hat{\lambda}^2} e^{\hat{\sigma}_Z^2 - 2\tilde{C}_Z(r)}\right) + \alpha^2 \frac{b^2}{\hat{\lambda}^2} e^{-\tilde{C}_Z(r)}\left(e^{\hat{\sigma}_Z^2 - \tilde{C}_Z(r)} - 1\right)\right)^{\eta}\right]^2 \mathrm{d}r \tag{19}$$

Minimization with fixed shape parameter α gives estimates for a and b. Here $\hat{\sigma}_Z^2$, $\hat{\phi}_Z$ and $\hat{\lambda}$ are estimated parameters for the log Gaussian Cox process, and $\tilde{C}_Z(r)$ is either $\hat{C}_Z(r)$ or $C_Z(r;\hat{\theta})$, for example (14) with $\hat{\theta} = (\hat{\sigma}_Z^2, \hat{\phi}_Z)$ plugged in. If the log Gaussian Cox process fits well using $\hat{C}_Z(r)$ or $C_Z(r;\hat{\theta})$ is expected to give similar results. If there are more fluctuations in $\hat{C}_Z(r)$, one might want to base the estimation on the parametric form which is more stable. On the other hand, peculiarities that exist in $\hat{C}(r)$ may be transferred to the mark variogram as well, and using $\hat{C}(r)$ in (19) may cancel these out. We recommend to minimize (19) both with $\tilde{C}_Z(r) = \hat{C}_Z(r)$ and $\tilde{C}_Z(r) = C_Z(r;\hat{\theta})$ and compare the results.

Gaussian intensity-marked Cox process. For the marking (6) direct optimization of the integral (18) is problematic and, in addition, the model parameter a does not exist in the equation. More stable solution and also an estimate for a are achieved by the following two-step procedure:

The parameters a and d^2 tell the mean and variance of the marks in regions with very high intensity (then $b/\Lambda(x_i)$ is small), respectively. Consequently, these can be estimated from the marks at points $x_i \in N$ in regions with high intensity, which can be determined by a cusp-point method: Let us denote $H_h = \{[x_i; m(x_i)] \in N_m : \Lambda(x_i) \geq h\}$ and let $|H_h|$ stand for the number of points in H_h. Let h vary from $\min(\Lambda(x_i))$ to $\max(\Lambda(x_i))$, $x_i \in N$, and calculate the mean and variance of the marks in H_h. This gives the mean and variance of the marks as a function of h. A cusp-point is the value h_0 for which it holds that the mean and variance of the marks

are approximately constant for values $h \geq h_0$. The estimates for a and d^2 are obtained by

$$\hat{a} = \frac{1}{|H_{h_0}|} \sum_{[x_i;\, m(x_i)] \in H_{h_0}} m(x_i) \tag{20}$$

and

$$\hat{d}^2 = \frac{1}{|H_{h_0}| - 1} \sum_{[x_i;\, m(x_i)] \in H_{h_0}} (m(x_i) - \hat{a})^2, \tag{21}$$

respectively. Of course the $\Lambda(x_i)$ above must be replaced by the estimate $\hat{\Lambda}(x_i)$, $x_i \in N$. The kernel estimation of the intensity should be sufficient to detect the high intensity areas.

Considering $\hat{d}^2$ as a fixed value for d^2 (and $\hat{a}$ for a), we next estimate the parameters b and c by minimizing

$$\int_{\epsilon}^{a_0} \left[(\hat{\gamma}_m(r))^{\eta} - \left(\hat{d}^2 + \left(\frac{c^2}{\hat{\lambda}} - \frac{b^2}{\hat{\lambda}^2} \right) e^{-\tilde{C}_Z(r)} + \frac{b^2}{\hat{\lambda}^2} e^{\hat{\sigma}_Z^2 - 2\tilde{C}_Z(r)} \right)^{\eta} \right]^2 \mathrm{d}r. \tag{22}$$

Here $\tilde{C}_Z(r)$ is either $\hat{C}_Z(r)$ or $C_Z(r; \hat{\theta})$ as above.

Maximizing the conditional likelihood of the marks

If the marking is specified through a mark distribution which depends on the intensity $\Lambda(x_i)$ and the marking is assumed to be conditionally independent given $\Lambda(x_i)$, $x_i \in N$, and the x_is are fixed, then this leads to the log likelihood

$$\sum_{[x_i;\, m(x_i)] \in N_m} \log f_{m|\Lambda}(m(x_i)|\Lambda(x_i)), \tag{23}$$

where $f_{m|\Lambda}(m(x_i)|\Lambda(x_i))$ is the parametric density of the conditional mark distribution of the mark $m(x_i)$ at (fixed) $x_i \in N$ with known intensity $\Lambda(x_i)$. Maximization of (23) leads to parameter estimates. Because the intensity is in practise unknown, an estimate obtained e.g. through kernel method must be used instead of $\Lambda(x_i)$.

Maximization of (23) is easily applicable for the marking (7) and a method to obtain an estimate for α. However, since the method is sensitive to the intensity estimation, one might use the conditional maximum likelihood to obtain an estimate for α and then minimize (19) using the log Gaussian Cox process assumption if reasonable. If α is known, then the log likelihood (23) can be maximized with respect to a and b with fixed α.

The conditional maximum likelihood method can be used as such to obtain estimates for a, b, c and d if good estimates for $\Lambda(x_i)$, $x_i \in N$, exist. However, it seems, according to limited experience, that, because of the sensitivity to the estimated intensity, better estimates are obtained, if a and d are first estimated as described above and then (23) is maximized with fixed a and d to obtain estimates for b and c.

Moment-based estimators

Instead of assuming a parametric conditional mark distribution one can proceed in terms of moments. Simple curve fitting can be done by minimizing the square sum

$$\sum_{[x_i;\, m(x_i)] \in N_m} \left(m(x_i) - \mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)] \right)^2 \tag{24}$$

or, alternatively, the weighted square sum

$$\sum_{[x_i;\, m(x_i)] \in N_m} \frac{1}{\Lambda(x_i)} \left(m(x_i) - \mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)]\right)^2 \tag{25}$$

with respect to parameters in $\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)]$. A parametric formula is assumed only for $\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)]$.

Assuming $\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)] = \alpha(a+b/\Lambda(x_i))$, α fixed, (24) and (25) can be minimized to obtain estimates for a and b in (7).

This moment-based estimation method is not especially suitable for the marking (6), since it assumes only $\mathbb{E}_{x_i}[m(x_i)|\Lambda(x_i)] = a + b/\Lambda(x_i)$ and gives estimates only for these a and b.

6 Simulation experiments

In this section simulation experiments of the heteroscedastic intensity-marked Cox processes are performed. Parameter estimation of these processes is demonstrated by applying estimation to the simulated processes.

A log Gaussian Cox process generated by a Gaussian random field with mean $\mu_Z = -4.0$ and exponential covariance function (14) with $\sigma_Z^2 = 1.5$ and $\phi_Z = 6.0$ is simulated in a window of size 200×200. A realisation of the Gaussian random field $\{Z(s) : s \in [0, 200] \times [0, 200]\}$, intensity function $\Lambda(s) = e^{Z(s)}$ and generated log Gaussian Cox process $N = \{x_i\}$ with 1523 points can be seen in Figure 3. We consider the marking of the point process by (8) and (6) such that

$$\begin{aligned} m^{(1)}(x_i)|\, \Lambda(x_i) &\sim \mathrm{Exp}(1/(30 + 0.8/\Lambda(x_i))) \\ m^{(2)}(x_i)|\, \Lambda(x_i) &\sim N(20 + 0.2/\Lambda(x_i), 0.5^2/\Lambda(x_i) + 2^2). \end{aligned}$$

An estimate with translation edge correction for $g(r)$ is calculated by means of the function `pcf` in R library `spatstat` (1.9-1), see Baddeley and Turner (2005), and $\hat{C}_Z(r)$ is obtained using (13). Empirical mark variograms of the marks $m^{(1)}(x_i)$ and $m^{(2)}(x_i)$ are calculated with translation edge correction. The estimates $\hat{\Lambda}(x_i)$ for $x_i \in N$ are obtained by kernel estimation with Epanechnikov kernel and bandwidth 10.

Log Gaussian Cox process. We fit the exponential covariance function and minimize (15) with $\epsilon = 0.25$ and $a_0 = 25$. We choose $\eta = 1/2$ as Møller et al. (1998) to give more weight to small values of r. The estimates obtained by the minimum contrast method and, additionally, by the approximative maximum likelihood method (for short AML) are shown in Table 1. An empirical covariance function $\hat{C}_Z(r)$ with fitted models is shown in Figure 4. The estimates given by these two methods are very similar. The approximative likelihood method seems to work well.

Marking 1. The parameters a and b of the inverse exponential intensity-marked Cox process are estimated by minimizing (19) with $\alpha = 1$, $\epsilon = 0.25$, $a_0 = 30$ and $\eta = 1$ assuming the log Gaussian Cox process (for short LGCP) with estimates obtained using both the minimum contrast method and the approximative maximum likelihood. Using $\eta = 1/2$ instead does not affect the estimates much in this case. Further estimates are obtained by minimizing (24) and

Table 1: Estimation results for the log Gaussian Cox process simulation.

	$\hat{\sigma}_Z^2$	$\hat{\phi}_Z$	$\hat{\mu}_Z$	$\hat{\lambda}$
Minimum contrast	1.35	7.20	-3.94	0.038
AML	1.32	7.48	-3.93	0.038
Values used in simulation	1.5	6.0	-4.0	0.039

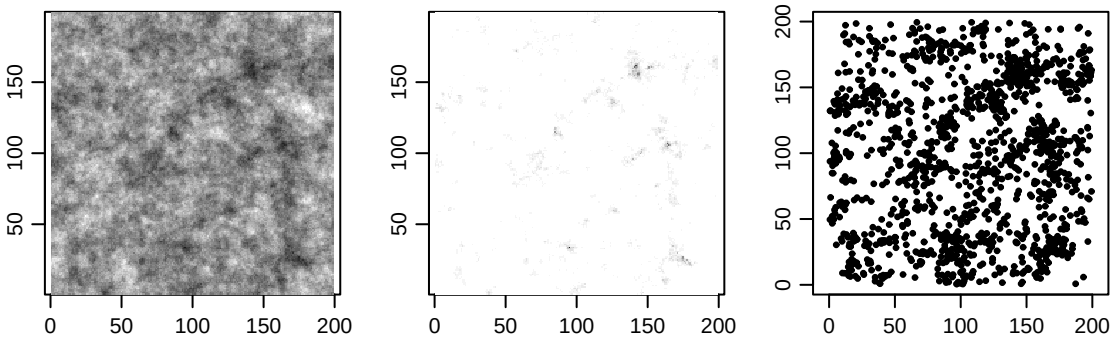

Figure 3: A realisation of the Gaussian random field $\{Z(s)\}$ (on the left) and the intensity $\Lambda(s) = e^{Z(s)}$ (in the middle). Dark color corresponds to a high value of a random field. A realisation of the log Gaussian Cox process with intensity $\{\Lambda(s)\}$ (on the right).

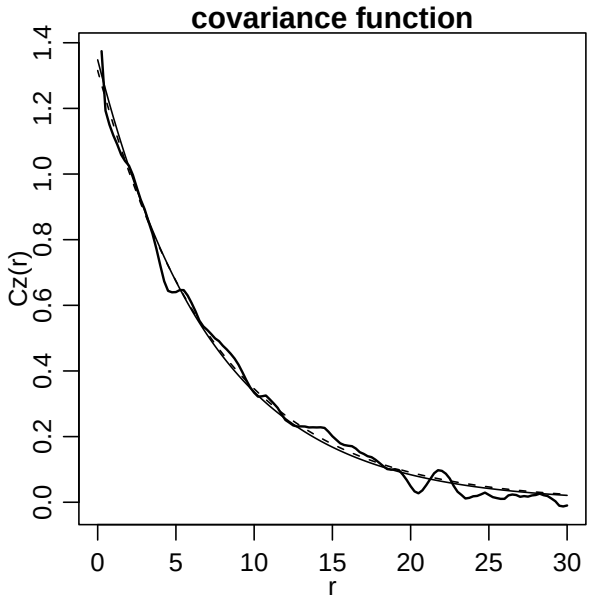

Figure 4: Estimated covariance function $\hat{C}_Z(r)$ (thick solid line) with fitted exponential covariance functions obtained through the minimum contrast method (solid line) and through the approximative maximum likelihood method (dotted line).

Table 2: Estimated mark parameters for the exponential marking.

	$\hat{a}$	$\hat{b}$	Model assumed for $\{x_i\}$
Method 1a: minimize (19) with $\hat{C}_Z(r)$	33.21	0.79	LGCP, minimum contrast
Method 1b: minimize (19) with $C_Z(r;\hat{\theta})$	32.33	0.80	LGCP, minimum contrast
Method 1c: minimize (19) with $\hat{C}_Z(r)$	33.18	0.80	LGCP, AML
Method 1d: minimize (19) with $C_Z(r;\hat{\theta})$	32.14	0.83	LGCP, AML
Method 2: conditional ML (23)	30.86	0.91	no assumption
Method 3a: minimize (24)	29.67	0.97	no assumption
Method 3b: minimize (25)	26.48	1.05	no assumption
Values used in simulation	30.00	0.80	

(25) and maximizing (23) with $\alpha = 1$. The results are shown in Table 2. In Figure 5 mark summaries are plotted with their theoretical counterparts with estimated parameters (a and b obtained by Methods 1a, 1b, 3a and 3b). The estimation based on the mark variogram (Method 1) seems to give better fit with respect to mark characteristics.

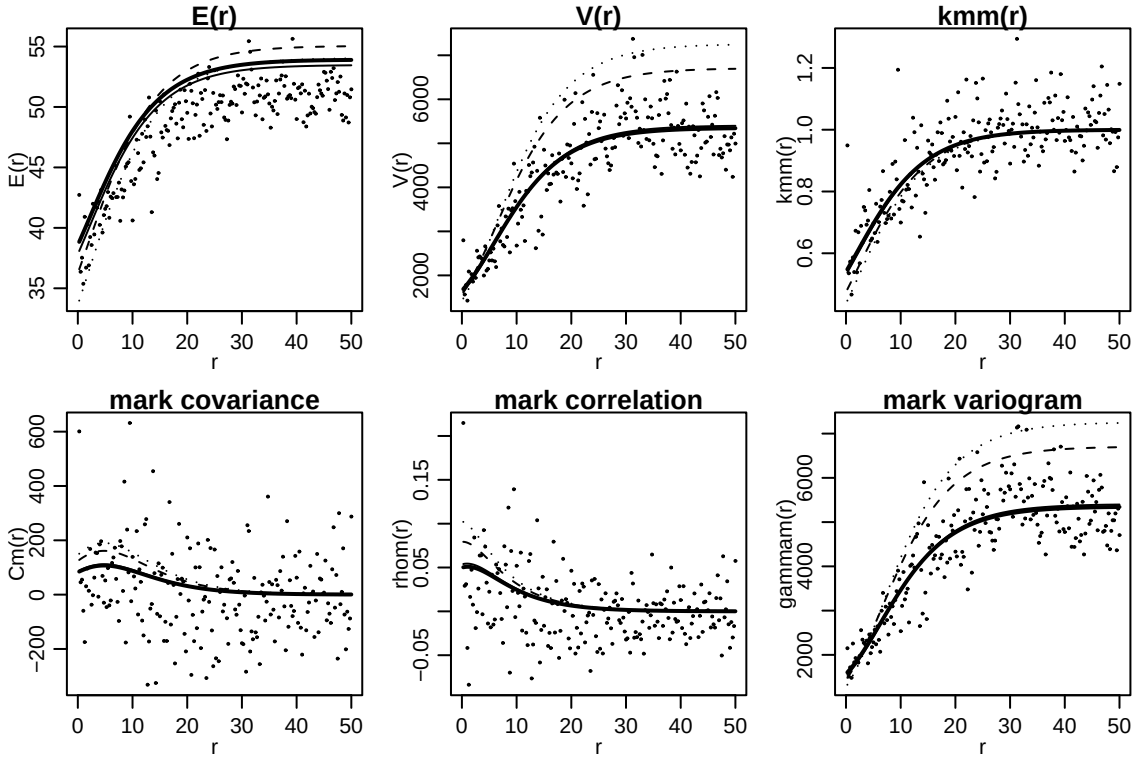

Figure 5: Summaries $E(r)$, $V(r)$, $k_{mm}(r)$, $C_m(r)$, $\rho_m(r)$ and $\gamma_m(r)$ calculated from the realisation of the inverse exponential intensity-marked Cox process (dots) and their theoretical counterparts with estimated parameters $\hat{a}$, $\hat{b}$ and $C_Z(r;\hat{\theta})$. The thick line corresponds to Method 1b, the solid thin line to Method 1a, the dashed line to Method 3a and the dotted line to Method 3b. The summaries are plotted for the values of r ranging from 0.25 to 50 by 0.25.

Marking 2. Consider the inference of the Gaussian intensity-marked Cox process. The mean and the variance of the marks in $H_h = \{[x_i; m(x_i)] \in N_m : \hat{\Lambda}(x_i) \geq h\}$ as functions of h are shown in Figure 6. The cusp-point $h_0 = 0.25$ is chosen. The variation with large values of h is due to the small number of observations in H_h. The estimates obtained by (20) and (21) with $h_0 = 0.25$ are the following (values used in simulation are given in parentheses): $\hat{a} = 20.32$ (20.00) and $\hat{d} = 1.87$ (2.00).

Minimization of (22) with $\epsilon = 0.25$, $a_0 = 30$ and $\eta = 1/2$ and maximization of (23) with respect to b and c yield the estimates shown in Table 3. Mark summaries with their theoretical counterparts (b and c obtained by Methods 1a, 1b and 2) are shown in Figure 7. With respect to mark characteristics minimum contrast method applied to the mark variogram seems to give best estimates for b and c, but estimates that are closest to the ones used in simulation are obtained by the conditional maximum likelihood method.

Table 3: Estimated parameters b and c for the Gaussian marking.

	$\hat{b}$	$\hat{c}$	Model assumed for $\{x_i\}$
Method 1a: minimize (22) with $\hat{C}_Z(r)$	0.20	0.76	LGCP, minimum contrast
Method 1b: minimize (22) with $C_Z(r;\hat{\theta})$	0.21	0.69	LGCP, minimum contrast
Method 1c: minimize (22) with $\hat{C}_Z(r)$	0.21	0.78	LGCP, AML
Method 1d: minimize (22) with $C_Z(r;\hat{\theta})$	0.22	0.70	LGCP, AML
Method 2: conditional ML (23)	0.19	0.51	no assumption
Values used in simulation	0.20	0.50	

7 Modelling the dbh of trees in a tropical rainforest data

Assume that the marked point pattern of *T. tuberculata* in Figure 1 origins from an inverse exponential intensity-marked Cox process. The model could be appropriate for the data, since the point pattern is clustered and the relationship between the marks and the intensity of the marked point patterns seems to be similar to that of the inverse exponential intensity-marked Cox process. In addition, maximization of (23) for the inverse gamma intensity-marked Cox process gives $\hat{\alpha} = 0.97$ using kernel estimates for $\Lambda(x_i)$, $x_i \in N$. Thus, assume $\alpha = 1$.

The minimum contrast method and the approximative maximum likelihood method described in Section 5 are used to fit a log Gaussian Cox process to the unmarked point pattern. The goodness of fit of the log Gaussian Cox process as a model for the point pattern is evaluated not only by means of the covariance function of the random field Z (or pair-correlation), which generates the intensity, but also by means of two other functions, the nearest neighbour distance distribution function and the spherical contact distribution function, which are defined by

$$D(r) = 1 - \mathbb{P}_o(N(b(o,r)) = 1), \quad \text{for } r \geq 0,$$

and

$$H_s(r) = 1 - \mathbb{P}(N(b(o,r)) = 0), \quad \text{for } r \geq 0,$$

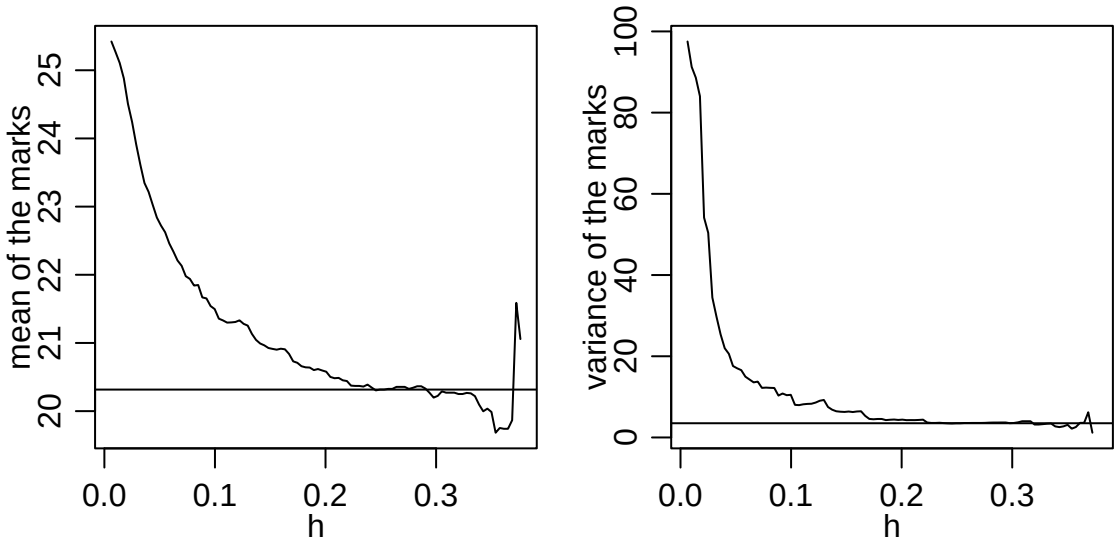

Figure 6: The cusp-point method for the realisation of the Gaussian intensity-marked Cox process. The horizontal lines correspond to the mean and the variance of marks $m(x_i)$ for which $\hat{\Lambda}(x_i) > 0.25$.

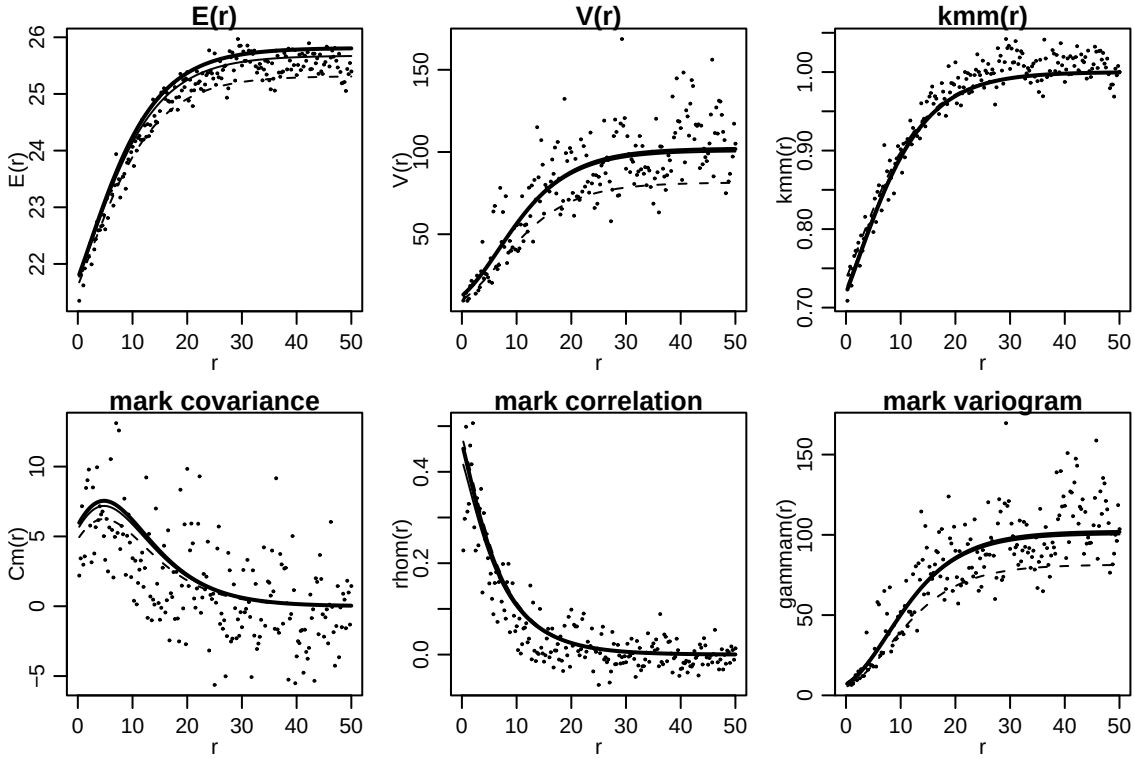

Figure 7: Summaries $E(r)$, $V(r)$, $k_{mm}(r)$, $C_m(r)$, $\rho_m(r)$ and $\gamma_m(r)$ calculated from the realisation of the Gaussian intensity-marked Cox process (dots) and their theoretical counterparts with estimated parameters $\hat{a}$, $\hat{b}$, $\hat{c}$, $\hat{d}$ and $C_Z(r;\hat{\theta})$. The thick solid line corresponds to Method 1b, the thin solid line to Method 1a and the dashed line to Method 2.

respectively, see for example Stoyan et al. (1995) or Diggle (2003). Here $\mathbb{P}_o$ is conditional probability (Palm probability) given that there is a point of N at o and $b(o,r)$ is a circle of radius r centred at o. Other parameters of the inverse exponential intensity-marked Cox process are estimated by means of various methods described in Section 5. Confidence envelopes for the summary characteristics $E(r)$, $V(r)$, $k_{mm}(r)$ and $\gamma_m(r)$ are constructed from 99 independent simulations from the fitted model for testing the goodness of fit.

The empirical covariance function $\hat{C}_Z(r)$ is calculated using the translation edge correction method and the exponential covariance function (14) is fitted to $\hat{C}_Z(r)$ by minimizing (15) with $\epsilon = 0.25$, $a_0 = 30$ and $\eta = 1/2$. The estimates for the log Gaussian Cox process (for short LGCP) obtained by the minimum contrast method and approximative maximum likelihood (AML) can be found in Table 4. The estimate of the covariance function and the fitted exponential

Table 4: Estimated LGCP parameters for the tropical rainforest data.

	$\hat{\sigma}_Z^2$	$\hat{\phi}_Z$	$\hat{\mu}_Z$	$\hat{\lambda}$
Minimum contrast	0.70	12.63	-4.29	0.019
AML	0.51	19.69	-4.20	0.019

covariance functions are plotted in Figure 8 with envelopes calculated from 99 independent simulations of the log Gaussian Cox process. The envelopes from 99 simulations are calculated also for the $D(r)$- and $H_s(r)$-functions, see Figure 9. All of these functions show that the fit is satisfactory. The minimum contrast method gives better estimates with respect to the functions $C_Z(r)$ and $D(r)$. Thus we can consider the log Gaussian Cox process (estimated by the minimum contrast method) as a suitable model for the unmarked point pattern of the data.

We assume that the marks are provided by (8): $m(x_i)|\,\Lambda(x_i) \sim \mathrm{Exp}(1/(a + b/\Lambda(x_i)))$. Minimization of (19), (24) and (25) and maximization of (23), all with $\alpha = 1$, yield the estimates for a and b gathered in Table 5. Estimates $\hat{\Lambda}(x_i)$, $x_i \in N$, are obtained by the kernel method. The values $\epsilon = 0.25$, $a_0 = 30$ and $\eta = 1$ are used in parameter estimation through the mark variogram. Minimization of (19) using the parameters of the log Gaussian Cox process obtained by the minimum contrast method yields similar results with $\hat{C}_Z(r)$ and $C_Z(r;\hat{\theta})$. However using the estimates obtained by the AML method, the results differ remarkable, which is due to large difference between $\hat{C}_Z(r)$ and $C_Z(r;\hat{\theta})$ with $r < 10$m. In this case we rely more on the minimum contrast method. Mark summaries calculated from the data and their theoretical counterparts of the inverse exponential intensity-marked Cox process (obtained by Methods 1a, 1b, 3a and 3b) are plotted in Figure 10.

The estimates based on the mark variogram seem to fit better to the data with respect to mark characteristics $E(r)$, $V(r)$, $k_{mm}(r)$, $C_m(r)$, $\rho_m(r)$ and $\gamma_m(r)$ than the ones obtained by methods based on (non-parametric) intensity estimation, which is probably due to goodness of the log Gaussian Cox process as a model for the point pattern.

The estimates obtained by Method 1b are used to calculate envelopes for the mark characteristics from simulations of the inverse exponential intensity-marked Cox process. See Figure 11 for results. According to these mark characteristics and the functions $C_Z(r)$, $D(r)$ and $H_s(r)$ in Figures 8 and 9, the inverse exponential intensity-marked Cox process seems to describe the marked point pattern of *T. tuberculata* well.

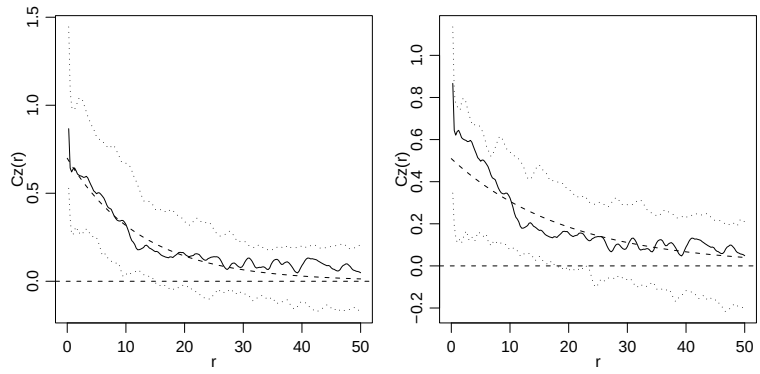

Figure 8: Estimated covariance function $\hat{C}_Z(r)$ calculated with translation edge correction from the point pattern of *T. tuberculata* (solid line), fitted exponential covariance function (dashed line) and limits from 99 simulations of the log Gaussian Cox process with estimated parameters (dotted lines). On the left: fitting by the minimum contrast method. On the right: fitting by the approximative maximum likelihood method.

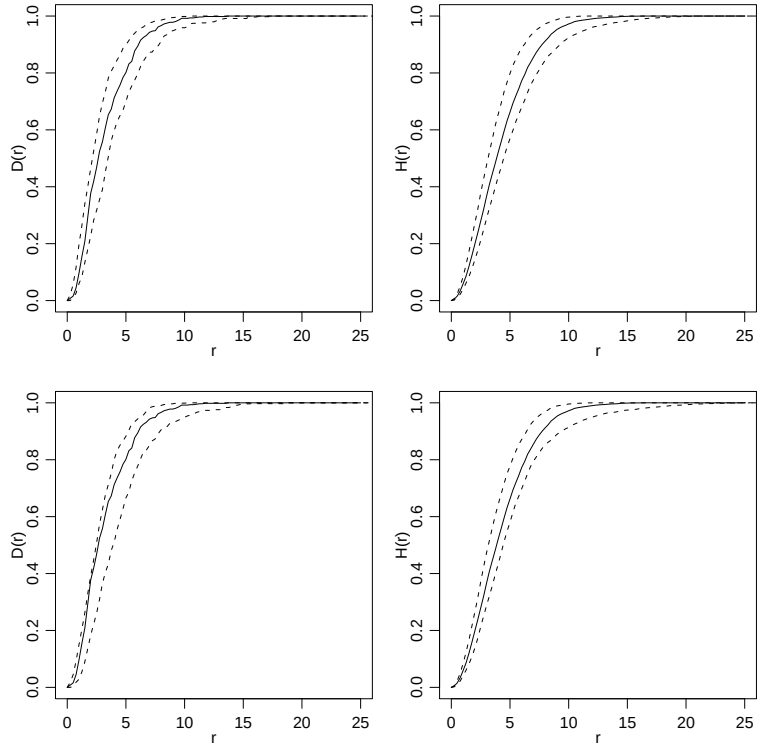

Figure 9: Estimated nearest neighbour distance distribution function (on the left) and spherical contact distribution function (on the right) of *T. tuberculata* (solid lines) with the limits calculated from 99 simulations of the log Gaussian Cox process (dashed lines). The functions have been calculated with border edge correction with functions `Fest` and `Gest` in R library `spatstat`. Upper row: fitting by the minimum contrast method. Lower row: fitting by the approximative maximum likelihood method.

Table 5: Estimated mark parameters for the tropical rainforest data.

	$\hat{a}$	$\hat{b}$	Model assumed for $\{x_i\}$
Method 1a: minimize (19) with $\hat{C}_Z(r)$	22.06	1.00	LGCP, minimum contrast
Method 1b: minimize (19) with $C_Z(r;\hat{\theta})$	23.38	0.98	LGCP, minimum contrast
Method 1c: minimize (19) with $\hat{C}_Z(r)$	23.54	1.10	LGCP, AML
Method 1d: minimize (19) with $C_Z(r;\hat{\theta})$	0.22	1.43	LGCP, AML
Method 2: conditional ML (23)	32.10	1.02	no assumption
Method 3a: minimize (24)	39.23	0.83	no assumption
Method 3b: minimize (25)	53.36	0.59	no assumption

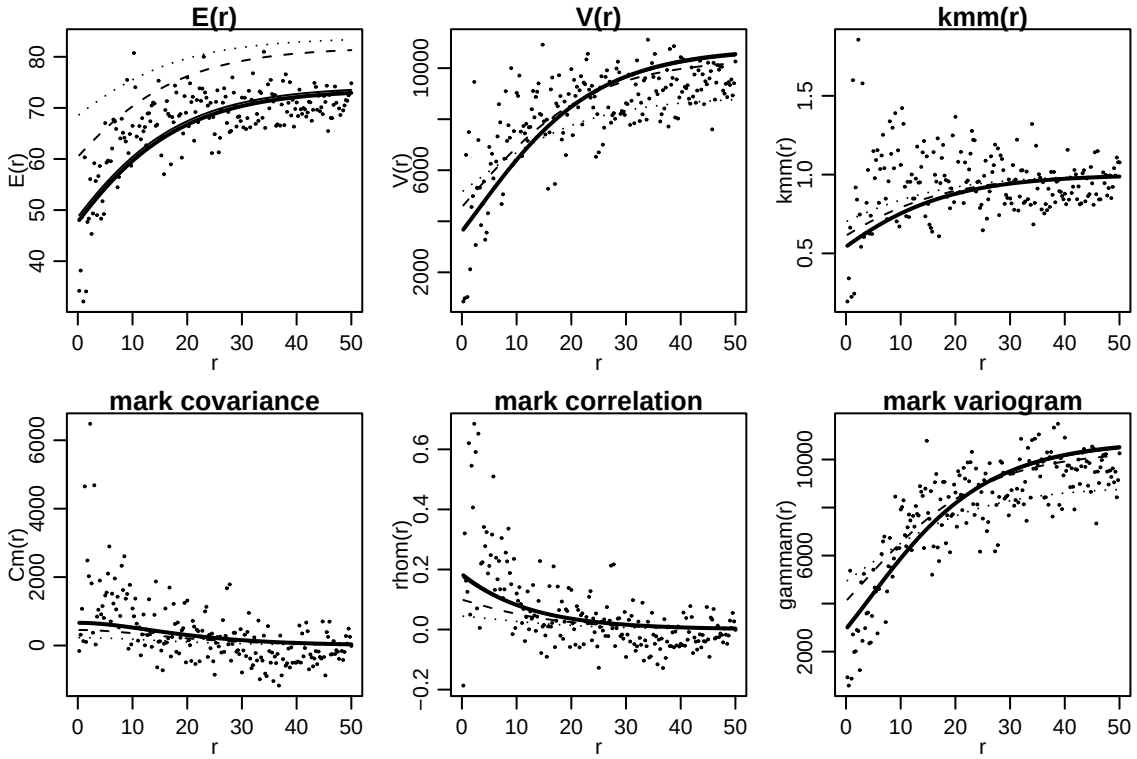

Figure 10: Summaries $E(r)$, $V(r)$, $k_{mm}(r)$, $C_m(r)$, $\rho_m(r)$ and $\gamma_m(r)$ calculated from *T. tuberculata* (dots) and their theoretical counterparts with estimated parameters (lines). The thick line corresponds to Method 1b, the solid thin line to Method 1a, the dashed line to Method 3a and the dotted line to Method 3b.

8 Discussion

The new marking models extend the independent and geostatistical marking strategies in two ways. First the marks are allowed to depend on the point intensity, leading to "density-dependence" of marks. Second, not only the mean but also variation around the mean of the marks is modelled. This is an advantage e.g. in forestry applications where in tree clusters trees tend to be small on average, but at the same time, show large variation in the form of

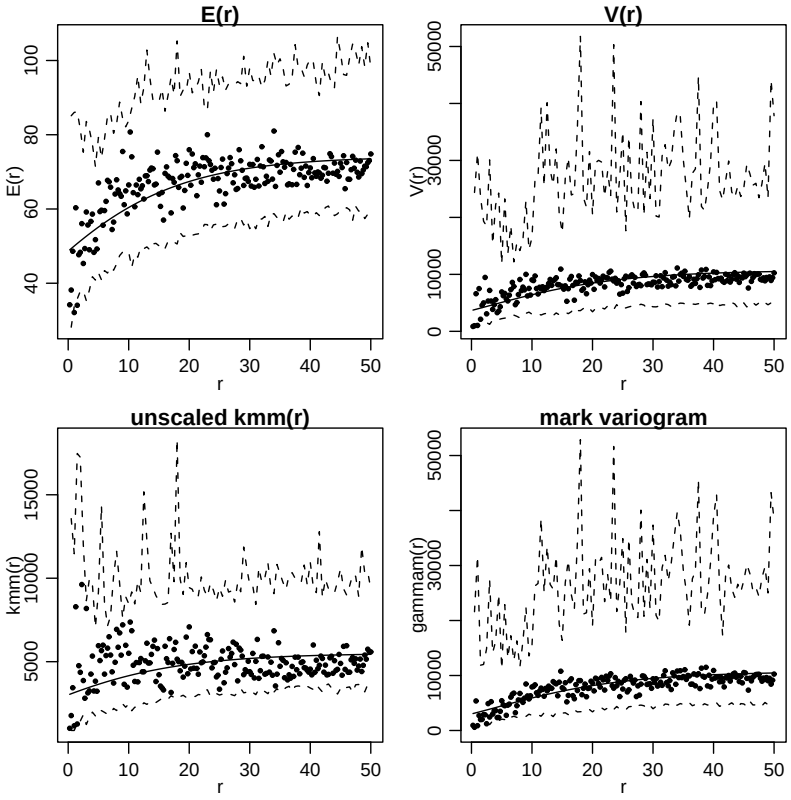

Figure 11: Summaries $E(r)$, $V(r)$, $\kappa_{mm}(r)$ and $\gamma_m(r)$ calculated from *T. tuberculata* (dots) with their theoretical counterparts with parameters obtained by Method 1b (solid lines) and upper and lower envelopes from 99 simulations from the fitted inverse gamma intensity-marked Cox process (dashed lines).

long tail of the size (or dbh) distribution.

The unmarked point process is chosen to the log Gaussian Cox process (LGCP) which generates clustered patterns as do all Cox processes. Intensity-dependent marking makes sense only if the point process realisations are clustered. The LGCP assumption can be relaxed to other clustering models, e.g. to inhomogeneous Poisson process, cluster process or Poisson/gamma random field generated Cox process (Wolpert and Ickstadt, 1998). However, it should be noted that, some of the estimation methods are based on the LGCP assumption. Our choice to base the reasoning on LGCP is a good choice for many point pattern data and a reasonable starting point for developing new marking strategies.

In the marking strategies, conditional indepence of the marking of the points given the intensity is assumed. This simplifies the theoretical calculations and seems not to be very restrictive. The marking models are ingenious in the sense, that analytical mark characteristics are obtainable.

As the estimation method two-stage procedures are suggested. At the first stage the intensity (or pair-correlation) of the unmarked point process are estimated, with or without a model assumption. This leads to a regression-type second stage, where the mark distribution is estimated conditional on the intensity (or pair-correlation). One may wonder that simultaneous estimation of the intensity and marking parameters would be more efficient. For large data, the loss of efficiency is probably unessential. Density-dependent marking is useful for large marked

point patterns.

Various methods are suggested for the estimation of the parameters of the mark distribution. The method based on moments may often be used in preliminary analysis investigating the relationship between the marks and the intensity. The minimum contrast method applied to the mark variogram is recommended, when the log Gaussian Cox process is a good model for the point locations. Limited experience suggests that this method, which relies more on the model, works better than the moment-based method, when the log Gaussian Cox process is a reasonable model. The minimum contrast method is applicable with other point process models as well, if one is able to obtain the theoretical mark variogram for the applied model. The conditional maximum likelihood assuming a parametric model for the mark distribution might be preferred if good estimates for intensity exist.

The models presented here are meant for clustered point patterns with heteroscedastic intensity-dependent marks with the property that the marks and points have similar ranges of correlation. The mark characteristics depend on the correlation function of the Gaussian random field $\{Z(s)\}$, that generates the intensity of the log Gaussian Cox process. From the point of view of applications we have been successful in capturing the density-dependence which exists in the rainforest data. However, as Ho and Stoyan (2007) states, there is also a need for intensity-marked processes which allow different ranges of correlation for marks and points. A marking similar to (4), allowing additional source of variation for marks, might be a way to proceed.

Acknowledgements This work has been stimulated by participation in the joint NERC Centre for Population Biology and UK Population Biology Network working group 'Spatial analysis of tropical forest biodiversity' funded by the Natural Environment Research Council and English Nature. The research has been funded by the Academy of Finland (Project number 111156). We thank professor Dietrich Stoyan for giving the manuscript Ho and Stoyan (2007) at our disposal.

References

Baddeley, A. and Turner, R. (2005). Spatstat: An R package for analysing spatial point patterns. *Journal of Statistical Software* **12**, 1–42.

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

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

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley & Sons, New York. Revised Edition.

Diggle, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*. Arnold, London.

Ho, L. P. and Stoyan, D. (2007). Modelling Marked Point Patterns by Intensity-marked Cox Processes. *Statistics & Probability Letters* (to appear).

Hubbell, S. and Foster, R. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. Pages 24-41 in S. L. Sutton, T. C. Whitmore, and A. C. Chadwick, eds. *Tropical Rain Forest: Ecology and Management*. Blackwell Scientific Publications, Oxford.

Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). Statistical Analysis and Modelling of Spatial Point Patterns. Wiley (to appear).

Mase, S. (1996). The treshold method for estimating total rainfall. *Annals of the Institute of Statistical Mathematics* **48**, 201–213.

Menezes, R. (2005). *Assessing spatial dependency under non-standard sampling*. Ph.D. thesis (ISBN 84-9750-595-6). Universidad de Santiago de Compostela, Santiago de Compostela, Spain.

Møller, J., Syversveen, A. R., and Waagepetersen, R. P. (1998). Log Gaussian Cox Processes. *Scandinavian Journal of Statistics* **25**, 451–482.

Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Prosecces*. Chapman & Hall/CRC, Boca Raton.

Myllymäki, M. (2006). On Intensity-dependent Marking of Log Gaussian Cox Processes. M.Sc. thesis, University of Jyväskylä, Finland.

Schlather, M. (2001). On the second-order characteristics of marked point processes. *Bernoulli* **7**, 99–117.

Schlather, M., Ribeiro, P., and Diggle, P. (2004). Detecting Dependence between Marks and Locations of Marked Point Processes. *Journal of the Royal Statistical Society, B* **66**, 79–93.

Stoyan, D. (1984). On correlations of marked point processes. *Mathematische Nachrichten* **116**, 197–207.

Stoyan, D., Kendall, W. S., and Mecke, J. (1995). *Stochastic Geometry and its Applications*. John Wiley & Sons, Chichester. 2nd ed.

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*. John Wiley & Sons, Chichester.

Stoyan, D. and Wälder, O. (2000). On Variograms in Point Process Statistics, II: Models of Markings and Ecological Interpretation. *Biometrical Journal* **42**, 171–187.

Tanaka, U., Ogata, Y., and Stoyan, D. (2007). Parameter estimation and model selection for Neyman-Scott point processes. *Biometrical journal* (to appear).

Wälder, O. and Stoyan, D. (1996). On Variograms in Point Process Statistics. *Biometrical Journal* **38**, 895–905.

Wolpert, R. and Ickstadt, K. (1998). Poisson/gamma random field models for spatial statistics. *Biometrika* **85**, 251–267.

Appendix: Calculation of the mark characteristics

Assume that $N = \{x_i\}$ is a stationary and isotropic log Gaussian Cox process with an intensity function $\{\Lambda(s)\}$. The calculation of moments of the marks for the intensity-marked Cox process in Ho and Stoyan (2007) yields that

$$\mathbb{E}_o[\Lambda(o)] = \lambda e^{\sigma_Z^2}$$

and

$$\mathrm{var}_o[\Lambda(o)] = \lambda^2 e^{2\sigma_Z^2}(e^{\sigma_Z^2} - 1).$$

Thus, using formulas $\mathbb{E}_o(m(o)) = \mathbb{E}_o\mathbb{E}_o[m(o)|\Lambda(o)]$ and $\mathrm{var}_o(m(o)) = \mathrm{var}_o[\mathbb{E}_o(m(o)|\Lambda(o))] + \mathbb{E}_o[\mathrm{var}_o(m(o)|\Lambda(o))]$, the mean mark and mark variance are obtained for the marks (5).

In a similar way we derive $\mathbb{E}_o[1/\Lambda(o)]$ and $\mathrm{var}_o[1/\Lambda(o)]$: Consider the marks $m(x_i) = 1/\Lambda(x_i)$ and denote the lognormal distribution of $\Lambda(o)$ by F. Then the corresponding mark distribution function is obtained in the following way using Campbell's and Fubini's theorems and stationarity of the process:

$$\begin{aligned}
\mathbb{E}\{N_m(B \times (-\infty, m])\} &= \mathbb{E}\left\{\mathbb{E}\left[\sum_{x\in N} \mathbf{1}_B(x)\mathbf{1}_{(-\infty,m]}\left(\frac{1}{\Lambda(x)}\right) \mid \Lambda\right]\right\} \\
&= \mathbb{E}\left[\int \mathbf{1}_B(x)\mathbf{1}_{(-\infty,m]}\left(\frac{1}{\Lambda(x)}\right)\Lambda(x)\mathrm{d}x\right] \\
&= \int \mathbf{1}_B(x)\mathbb{E}\left[\mathbf{1}_{(-\infty,m]}\left(\frac{1}{\Lambda(x)}\right)\Lambda(x)\right]\mathrm{d}x \\
&= \int \mathbf{1}_B(x)\mathbb{E}\left[\mathbf{1}_{(-\infty,m]}\left(\frac{1}{\Lambda(o)}\right)\Lambda(o)\right]\mathrm{d}x \\
&= \mathbb{E}\left[\mathbf{1}_{(-\infty,m]}\left(\frac{1}{\Lambda(o)}\right)\Lambda(o)\right]\int \mathbf{1}_B(x)\mathrm{d}x \\
&= \mathbb{E}\left[\mathbf{1}_{[1/m,\infty)}(\Lambda(o))\,\Lambda(o)\right]\int \mathbf{1}_B(x)\mathrm{d}x \\
&= \nu_d(B)\int_{1/m}^{\infty} x\mathrm{d}F(x),
\end{aligned}$$

where B is a Borel set in $\mathbb{R}^d$ and $\nu_d(B)$ its volume. Formula (1) yields that the distribution of the mark $m(x_i) = 1/\Lambda(x_i)$ is

$$F_{\mathcal{M}}(m) = \frac{1}{\lambda}\int_{1/m}^{\infty} x\mathrm{d}F(x).$$

The corresponding mean, second moment and variance are

$$\mu = \frac{1}{\lambda}, \quad \mu_2 = e^{-2\mu_Z} \quad \text{and} \quad \sigma^2 = e^{-2\mu_Z}\left(1 - e^{-\sigma_Z^2}\right),$$

respectively. For the marks (6) it holds

$$\mu_m = \mathbb{E}_o\left[a + b\frac{1}{\Lambda(o)}\right] \quad \text{and} \quad \sigma_m^2 = \mathrm{var}_o\left[a + b\frac{1}{\Lambda(o)}\right] + \mathbb{E}_o\left[c^2\frac{1}{\Lambda(o)} + d^2\right]$$

and thus (9) follows. The mean mark and mark variance of the marking (7) are obtained similarly.

The mark characteristics $E(r)$, $V(r)$, $\kappa_{mm}(r)$ and $\gamma_m(r)$ can be derived by using the equation (Ho and Stoyan, 2007)

$$\mathbb{E}_{or}\{f(m(o), m(\mathbf{r}))\} = \frac{\mathbb{E}[f(m(o), m(\mathbf{r}))\Lambda(o)\Lambda(\mathbf{r})]}{\mathbb{E}[\Lambda(o)\Lambda(\mathbf{r})]} = \frac{\mathbb{E}\{\Lambda(o)\Lambda(\mathbf{r})\mathbb{E}[f(m(o), m(\mathbf{r}))|\Lambda(o), \Lambda(\mathbf{r})]\}}{\mathbb{E}[\Lambda(o)\Lambda(\mathbf{r})]}.$$

The characteristics of the marking (3) are derived in Ho and Stoyan (2007) and these of the marking (5) are obtained by small modifications.

Consider the marking (6). Due to conditional independence of marks conditional on the intensity,

$$\mathbb{E}[m(o)m(\mathbf{r})\Lambda(o)\Lambda(\mathbf{r})] = \mathbb{E}\left[\Lambda(o)\Lambda(\mathbf{r})\left(a + \frac{b}{\Lambda(o)}\right)\left(a + \frac{b}{\Lambda(\mathbf{r})}\right)\right].$$

Consequently,

$$\begin{aligned}\mathbb{E}[m(o)m(\mathbf{r})\Lambda(o)\Lambda(\mathbf{r})] &= a^2\mathbb{E}e^{Z(o)+Z(\mathbf{r})} + ab\mathbb{E}e^{Z(o)} + ab\mathbb{E}e^{Z(\mathbf{r})} + b^2 \\ &= a^2e^{2\mu_Z+\sigma_Z^2+C_Z(r)} + 2abe^{\mu_Z+\sigma_Z^2/2} + b^2 \\ &= a^2\lambda^2e^{C_Z(r)} + 2ab\lambda + b^2.\end{aligned}$$

Since $\mathbb{E}[\Lambda(o)\Lambda(\mathbf{r})] = \lambda^2e^{C_Z(r)}$,

$$\mathbb{E}_{or}[m(o)m(\mathbf{r})] = a^2 + \frac{2ab}{\lambda}e^{-C_Z(r)} + \frac{b^2}{\lambda^2}e^{-C_Z(r)},$$

from which $k_{mm}(r)$ is obtained by dividing by μ_m^2. The value of k_{mm} at zero is $k_{mm}(0) = (\sigma_m^2 + \mu_m^2)/\mu_m^2$. Further,

$$\mathbb{E}[m(o)\Lambda(o)\Lambda(\mathbf{r})] = \mathbb{E}\left[\Lambda(o)\Lambda(\mathbf{r})\left(a + \frac{b}{\Lambda(o)}\right)\right] = a\lambda^2e^{C_Z(r)} + b\lambda.$$

and

$$\begin{aligned}\mathbb{E}[m(o)^2\Lambda(o)\Lambda(\mathbf{r})] &= \mathbb{E}\left[\Lambda(o)\Lambda(\mathbf{r})\left\{\frac{c^2}{\Lambda(o)} + d^2 + \left(a + \frac{b}{\Lambda(o)}\right)^2\right\}\right] \\ &= c^2\lambda + (a^2 + d^2)\lambda^2e^{C_Z(r)} + 2ab\lambda + b^2e^{\sigma_Z^2-C_Z(r)}.\end{aligned}$$

Thus,

$$E(r) = a + \frac{b}{\lambda}e^{-C_Z(r)}$$

and

$$\mathbb{E}_{or}[(m(o))^2] = a^2 + d^2 + \frac{2ab + c^2}{\lambda}e^{-C_Z(r)} + \frac{b^2}{\lambda^2}e^{\sigma_Z^2-2C_Z(r)},$$

and, consequently,

$$V(r) = \mathbb{E}_{or}[(m(o))^2] - [E(r)]^2 = d^2 + \frac{c^2}{\lambda}e^{-C_Z(r)} + \frac{b^2}{\lambda^2}(e^{\sigma_Z^2} - 1)e^{-2C_Z(r)}.$$

The mark variogram is obtained in a similar way:

$$\mathbb{E}[(m(o) - m(\mathbf{r}))^2\Lambda(o)\Lambda(\mathbf{r})] = c^2\lambda + d^2\lambda^2e^{C_Z(r)} + b^2e^{\sigma_Z^2-C_Z(r)} - b^2.$$

The mark characteristics of the inverse gamma intensity-marked Cox process are obtained by similar calculations.