

# Bayesian Inference on a Cox Process Associated with a Dirichlet Process

Larissa Valmy

LAMIA (EA4540)

Université des Antilles-Guyane

BP 250 - 97157 Pointe-à-Pitre

Jean Vaillant

LAMIA (EA4540)

Université des Antilles-Guyane

BP 250 - 97157 Pointe-à-Pitre

## ABSTRACT

In ecology and epidemiology, spatio-temporal distributions of events can be described by Cox processes. Situations for which there exists a hidden process which contributes to random effects on the intensity of the observed Cox process are considered. The observed process is a generalized shot noise Cox process and the hidden process is a Poisson process associated with a Dirichlet process. The distributional properties of quadrat counts are presented and bayesian inference is proposed for estimating and predicting parameters of interest in the model. Illustrations are given from weed spatial count data and disease mortality data.

## General Terms:

Point process, Bayesian statistics, Mixture models

## Keywords:

point process, Cox process, bayesian inference, ecology, epidemiology

## 1. INTRODUCTION

In ecology and epidemiology, distributions of events like disease occurrences, predator arrivals or plant locations can be considered as realizations of a point process, of which each point represents a single event. The point process theory has been presented and discussed in [2], [3] and [8]. Statistical procedures for analysing point process realizations can be found in books ([6]; [22]; [12]; [19]) and a lot of papers deal with applications in special situations (e.g. [7]; [11]; [4]; [25]). Some studies are based on counts of events in sampling units ([5]), and some others on event spatial positions or occurrence dates ([25]), and also distance sampling ([16]). Perry et al. ([20]) discussed appropriate selection and use of method for analyzing spatial point patterns in plant ecology. One may refer to [19] for papers about statistical tools for spatial point processes. [5] discussed recently about spatial point process models for forest inventories exhibiting overdispersion. In applications in which overdispersion is assumed, Cox process modeling is a common choice since this class of point processes is wide enough to take into consideration many features. Thus, [26] presented various scientific fields in which the Cox process, also known as doubly stochastic Poisson process, occurs. The intensity process  $\lambda(\cdot)$  of

such a process on a measured space  $(X, \mathcal{B}, \nu)$  is a random field related to its driving random measure  $\Lambda(\cdot)$  as follows :

$$\Lambda(B) = \int_B \lambda(x) \nu(dx), \quad (1)$$

for any element  $B$  of  $\mathcal{B}$ . In expression (1),  $\Lambda(B)$  is a random variable which stands for the expected number of points in  $B$ . Thus, modeling a Cox process is equivalent to modeling either its intensity or its driving measure ([24]). For example, [4] described a specific class of Cox processes, namely shot-noise G Cox processes, by modeling  $\Lambda(\cdot)$  with a shot-noise G-measure.

In this paper, a generalized shot-noise Cox process  $N$  on a measured space ([18]) is considered. The intensity at any spatial location is expressed as a linear combination of kernels centered on contributing events. The model is extended to the case where there exist contributions to the intensity process associated with a Dirichlet process ([13]; [9]). This latter process describes hidden environmental effects on the observed intensity. The Dirichlet process is almost surely discrete so that the proposed model takes into account situations for which some contributions may be equal even if associated with different spatial locations or dates. The focus is on statistical procedures providing inference tools in a bayesian framework when observing count data. In section 2, the distributional properties of observed counts in spatial sampling units are presented with respect to bandwidth influence of the hidden spatial events and with respect to concentration parameter of hidden contributions to intensity. Section 3 focuses on performing posterior inference on the hidden process parameters by means of a hybrid Gibbs-Metropolis-Hastings algorithm. In sections 4 and 5, two case studies are carried out : the first one concerns spatial count data from *Ibicella lutea*, the second one is based on death numbers per county in Georgia, US, caused by chronic lower respiratory diseases. After performing dispersion and spatial autocorrelation tests, estimations of the expected number of hidden events, their expected contribution to intensity and other parameters of the proposed model are carried out. Section 6 concludes the paper with an overview of the main features of the proposed approach.

## 2. DISTRIBUTIONAL PROPERTIES OF COUNT MIXTURES

Let  $N$  be a Cox process with intensity process  $\lambda$  on space  $X$  defined by :

$$\forall x \in X, \quad \lambda(x) = \sum_{j=1}^M a_j K(x, y_j) \quad (2)$$

where  $K$  is a kernel function such that  $K(\cdot, y)$  is a probability density function on  $X$  for any  $y$  in  $X$  and  $M$  is the random number of contributions. The  $y_j$  are a realization of a point process  $L$  on  $X$  with positive real marks  $a_j$  identically distributed according to a probability law  $G$  on  $\mathbb{R}_+$ , and independent of the  $y_j$ . If  $L$  is a homogeneous Poisson process with scalar intensity  $\mu$  and the  $a_i$  are equal to the same value, then  $\lambda$  is a standard shot noise process ([17]), and then  $N$  is a standard shot noise Cox process. Moreover, if the bandwidth of  $K$  is a random variable, then  $N$  is a generalized shot noise Cox process ([18]). In expression (2),  $a_j$  is the contribution of event  $y_j$  to the intensity  $\lambda$ .

In this paper, the contributions  $a_j$  are assumed to follow a Dirichlet process  $G$  ([13],[9]) with base measure  $G_0$  and concentration parameter  $\alpha$  denoted by  $DP(\alpha, G_0)$ . Therefore, there is no independence condition on the  $a_j$  as in many models but a weaker condition of exchangeability provided by De Finetti's theorem, see for example [10]. In fact the contributions  $a_j$  are conditionally independent given  $G$ . This provides a way of taking into account correlation between marks and mixed environmental effects. The  $y_j$  are assumed to be occurrence locations of a homogeneous Poisson process with parameter  $\mu$  denoted by  $HPP(\mu)$ . In other words,

$$\begin{array}{l} (a_j) \mid G \quad \overset{i.i.d}{\sim} \quad G \\ (y_j) \quad \quad \sim \quad HPP(\mu) \\ G \quad \quad \quad \sim \quad DP(\alpha, G_0). \end{array} \quad (3)$$

The following propositions give some distributional properties of the counting measure associated with  $N$  with respect to concentration parameter  $\alpha$ .

**PROPOSITION 1.** *Let  $N$  be a Cox process on the measured space  $(X, \mathcal{B}, \nu)$  with intensity defined by equations (2) and (3), then*

$$\forall A \in \mathcal{B}, \quad E(N(A)) = \mu E(a_1) \int_X \int_B K(x, y) \nu(dx) \nu(dy)$$

and

$$\begin{aligned} V(N(A)) = & E(N(A)) \left( 1 - \frac{1}{\alpha + 1} E(N(A)) \right) \\ & + \mu E(a_1^2) \int_X \left( \int_B K(x, y) \nu(dx) \right)^2 \nu(dy). \end{aligned}$$

**PROOF.** Taking into account (1) and (2), then for any element  $B$  of  $\mathcal{B}$ ,

$$\Lambda(B) = \sum_{j=1}^M a_j \int_B K(x, y_j) \nu(dx). \quad (4)$$

Since the  $y_j$  are uniformly distributed on  $X$  conditionally to  $M$ , the expectation of  $\Lambda(B)$  conditional on  $M$  is

$$ME(a_1) \int_X \frac{1}{\nu(X)} \int_B K(x, y) \nu(dx) \nu(dy)$$

and the final result for  $E(N(B))$  is obtained from  $E(M) = \mu\nu(X)$ .

Similarly, the second moment of  $\Lambda(B)$  conditional on  $M$  is

$$\begin{aligned} ME(a_1^2) \int_X \frac{1}{\nu(X)} \left( \int_B K(x, y) \nu(dx) \right)^2 \nu(dy) \\ + M(M-1)E(a_1 a_2) \left( \int_X \frac{1}{\nu(X)} \int_B K(x, y) \nu(dx) \nu(dy) \right)^2. \end{aligned}$$

$E(a_1 a_2) = (E(a_1))^2 \alpha / (\alpha + 1)$ ;  $E(M(M-1)) = (\mu\nu(X))^2$  and  $V(N(B)) = V(\Lambda(B)) + E(\Lambda(B))$  lead us to the final result.  $\square$

**PROPOSITION 2.** *With the same conditions as in proposition 1, let  $(B_1, B_2)$  be an element of  $\mathcal{B}^2$  with  $B_1 \cap B_2 = \emptyset$ . The covariance between  $N(B_1)$  and  $N(B_2)$  is*

$$\begin{aligned} Cov(N(B_1), N(B_2)) = \\ \mu E(a_1^2) \int_X \left( \int_{B_1} K(z, y) \nu(dz) \int_{B_2} K(x, y) \nu(dx) \right) \nu(dy) \\ - \frac{1}{\alpha + 1} E(N(B_1)) E(N(B_2)). \end{aligned}$$

**PROOF.** Since  $N$  is a Cox process, then  $Cov(N(B_1), N(B_2)) = Cov(\Lambda(B_1), \Lambda(B_2))$ . On the other hand,

$$\begin{aligned} E(\Lambda(B_1)\Lambda(B_2)) = \\ E \left( \sum_{j=1}^M a_j \int_{B_1} K(x, y_j) \nu(dx) \sum_{k=1}^M a_k \int_{B_2} K(x, y_k) \nu(dx) \right) \\ = E \left( ME(a_1^2) E \left( \int_{B_1} K(x, y_1) \nu(dx) \int_{B_2} K(x, y_1) \nu(dx) \right) \right) \\ + E \left( M(M-1) E(a_1 a_2) E \left( \int_{B_1} K(x, y_1) \nu(dx) \int_{B_2} K(x, y_2) \nu(dx) \right) \right) \end{aligned}$$

The final result is obtained from

$$\begin{aligned} E(M) = \mu\nu(X), \quad E(a_1 a_2) = (E(a_1))^2 \alpha / (\alpha + 1) \quad \text{and} \\ E(M(M-1)) = (\mu\nu(X))^2. \end{aligned}$$

$\square$

Let us consider count data in  $r$  disjoint subsets of  $X$ . When  $G_0$  belongs to a probability distribution family parameterized by  $b$  and  $K$  a kernel family parameterized by  $\sigma$ , we have the following result :

**PROPOSITION 3.** *Let  $N$  be a Cox process on the measured space  $(X, \mathcal{B}, \nu)$  with intensity defined by equations (2) and (3) and let  $B_1, \dots, B_r$  be  $r$  disjoint elements of  $\mathcal{B}$ . Consider  $G_0$  parameterized by  $b$  and  $K$  parameterized by  $\sigma$ . Under independent priors for  $\alpha, b, \mu$  and  $\sigma$ , the posterior distribution of  $(\alpha, b, \mu, \sigma, (a_j)_{j=1, \dots, M}, M)$ , conditional on counts in  $B_1, \dots, B_r$ , is proportional to :*

$$\begin{aligned} p(\alpha) p(b) p(\sigma) p(\mu) p(a_1, \dots, a_M \mid \alpha, b, \mu) \frac{(\mu\nu(X))^M}{M!} e^{-\mu\nu(X)} \\ \times \prod_{i=1}^r \Lambda(B_i)^{N(B_i)} e^{-\Lambda(B_i)} \end{aligned} \quad (5)$$

with the conditional joint distribution of the  $a_j$  obtained from the following expression :

$$a_j | a_1, \dots, a_{j-1} \sim \frac{\alpha}{\alpha + j - 1} G_0 + \frac{1}{\alpha + j - 1} \sum_{k=1}^{j-1} \delta_{a_k}. \quad (6)$$

The  $\Lambda(B_i)$  depend on  $(\sigma, (a_j)_{j=1, \dots, M}, M)$  as described in (4).

PROOF. Conditionally to process  $\lambda$ , the  $N(B_i)$  are independent counts following respectively a Poisson distribution with parameter  $\Lambda(B_i)$ . The associated conditional likelihood is then multiplied by the joint prior distribution of the parameters  $\alpha, b, \mu, \sigma, a_1, \dots, a_M$  and  $M$ . The Bayes theorem is then applied.

Equation (5) is obtained from the predictive law representation presented by [1].  $\square$

In the sequel,  $K$  is assumed to be an isotropic gaussian kernel with bandwidth  $\sigma$ . Another aspect of the model is the limits when the spatial influence parameter  $\sigma$  tends to zero or when the concentration parameter  $\alpha$  tends to zero or infinity.

PROPOSITION 4. With the same conditions as in proposition 1, consider  $B$  and  $B'$  two disjoint elements of  $\mathcal{B}$  and let  $K$  be an isotropic gaussian kernel with bandwidth  $\sigma$ . Then

- 1)  $\lim_{\sigma \rightarrow 0} E(N(B)) = \mu E(a_1) \nu(B)$ ,
- 2)  $\lim_{\sigma \rightarrow 0} \frac{V(N(B))}{E(N(B))} = 1 - \frac{\mu E(a_1)}{\alpha + 1} \nu(B) + \frac{E(a_1^2)}{E(a_1)}$ ,
- 3)  $\lim_{\sigma \rightarrow 0} Cov(N(B), N(B')) = \mu^2 (E(a_1))^2 \frac{\alpha}{\alpha + 1} \nu(B) \nu(B')$ ,
- 4) The variance-to-mean ratio  $\frac{V(N(B))}{E(N(B))}$  is an increasing function with respect to  $\alpha$ .

PROOF. 1)  $\Lambda(B) = \sum_{j=1}^M a_j \int_B K(x, y_j) \nu(dx)$  and for any  $(x, j) \in X \times \{1, \dots, M\}$ ,  $K(x, y_j)$  converges to  $\delta_{y_j}(\{x\})$  as  $\sigma$  converges to 0. This gives us

$$\lim_{\sigma \rightarrow 0} \Lambda(B) = \sum_{j=1}^M a_j \delta_{y_j}(B) \quad (7)$$

$E(N(B)) = E(\Lambda(B)) = E(M E(a_1 | M) E(\delta_{y_j}(B) | M))$ . Independence of  $a_j$  and  $y_j$  conditional on  $M$  leads to  $\lim_{\sigma \rightarrow 0} E(N(B)) = \mu E(a_1) \nu(B)$ .

- 2)  $\int_B K(x, y) \nu(dx)$  converges to  $\mathbb{1}_B(y)$  as  $\sigma$  converges to 0. Therefore  $\lim_{\sigma \rightarrow 0} \int_X \left( \int_B K(x, y) \nu(dx) \right)^2 \nu(dy) = \nu(B)$  which leads us to the final result.
- 3) Equation (7) leads to  $E(\Lambda(B) \Lambda(B')) = \mu^2 \nu(B) \nu(B') E(a_1 a_2)$ . On the other hand,  $Cov(N(B), N(B')) = Cov(\Lambda(B), \Lambda(B'))$ . Therefore,  $Cov(N(B), N(B')) = \mu^2 \nu(B) \nu(B') Cov(a_1, a_2)$ .

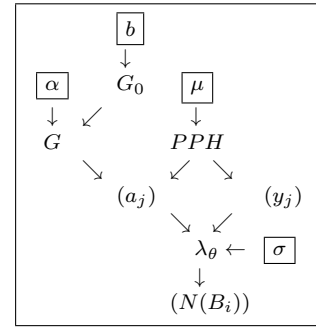


Fig. 1. Acyclic directed graph of the hierarchical model defined by (3) with  $G_0 = \text{Gamma}(1, b)$

4) From proposition 1, it results that

$$\frac{V(N(B))}{E(N(B))} = 1 - \frac{1}{\alpha + 1} E(N(B)) + \frac{E(a_1^2)}{E(a_1)} \frac{\int_X \left( \int_B K(x, y) \nu(dx) \right)^2 \nu(dy)}{\int_X \int_B K(x, y) \nu(dx) \nu(dy)}.$$

Moreover, in the above equation, the third term of the right member and  $E(N(B))$  do not depend on  $\alpha$ . Consequently, the variance-to-mean ratio is increasing with  $\alpha$ .

$\square$

Proposition 4 shows that even if the spatial influence of hidden events is negligible ( $\sigma$  close to zero), count correlation may be high according to the value taken by parameter  $\alpha$ . It is worth pointing out that when  $\alpha$  equals zero, the  $a_j$  are all equal to each other whereas when  $\alpha$  converges to infinity, the  $a_j$  are independent identically distributed according to  $G_0$ . Thus property 4 in proposition 4 indicates that  $\alpha$  is a dispersion parameter. The lower  $\alpha$  is, the higher the correlation between  $a_j$  is and the lower the overdispersion is.

In the sequel, the Gamma distribution with scale parameter  $a$  and shape parameter  $b$  is denoted by  $\text{Gamma}(a, b)$ .

### 3. ESTIMATION PROCEDURE

Posterior inference methods can be performed on spatial models ([14]) by implementing MCMC sampling. The model defined by expressions (2) and (3) and described by figure 1 involves two sets of unknowns, first  $\alpha, b, \mu, \sigma$  which are the parameters of interest, secondly, the positions  $(y_j)$  and effects  $(a_j)$  of the hidden events. Let us denote by  $q(\cdot)$  the joint distribution of the unknowns and count data up to a constant. Then :

$$q((N(B_i)), (\alpha, b, \mu, \sigma, M, (a_j, y_j)_{j=1, \dots, M})) \propto p(a_1, \dots, a_M | \alpha, b, \mu, \sigma, M) \frac{\mu^M}{M!} \frac{p(\alpha, b, \mu, \sigma)}{e^{\mu \nu(X)}} \prod_{i=1}^r \Lambda(B_i)^{N(B_i)} e^{-\Lambda(B_i)} \quad (8)$$

As expressed in (4), the positions  $(y_j)$  intervene in the posterior distribution of  $(\alpha, b, \mu, \sigma)$  in the following way :

$$\Lambda(B_i) = \sum_{j=1}^M a_j \int_{B_i} K(x, y_j) \nu(dx).$$

For  $t = 1, \dots, N$

- Generate  $\log(\alpha^{(t+1)})$  according to  $\mathcal{N}(\log(\alpha^{(t)}), \xi_1)$

- Accept  $\alpha^{(t+1)}$  with probability

$$\min \left( \frac{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t)}, \mu^{(t)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}{q \left( (N(B_i)), \left( \alpha^{(t)}, b^{(t)}, \mu^{(t)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}, 1 \right)$$

- Generate  $\log(b^{(t+1)})$  according to  $\mathcal{N}(\log(b^{(t)}), \xi_2)$

- Accept  $b^{(t+1)}$  with probability

$$\min \left( \frac{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t+1)}, \mu^{(t)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t)}, \mu^{(t)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}, 1 \right)$$

- Generate  $\log(\mu^{(t+1)})$  according to  $\mathcal{N}(\log(\mu^{(t)}), \xi_3)$

- Accept  $\mu^{(t+1)}$  with probability

$$\min \left( \frac{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t+1)}, \mu^{(t)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}, 1 \right)$$

- Generate  $\sigma^{(t+1)}$  according to  $\mathcal{N}(\log(\sigma^{(t)}), \xi_4)$

- Accept  $\sigma^{(t+1)}$  with probability

$$\min \left( \frac{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma^{(t+1)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}{q \left( (N(B_i)), \left( \alpha^{(t+1)}, b^{(t+1)}, \mu^{(t+1)}, \sigma^{(t)}, M^{(t)}, (a_j, y_j)_{j=1, \dots, M^{(t)}}^{(t)} \right) \right)}, 1 \right)$$

- Generate  $M^{(t+1)}$  according to  $\frac{1}{2} (\delta_{M^{(t)}+1} + \delta_{M^{(t)}-1})$

- Generate  $(a_1^{(t+1)}, \dots, a_{M^{(t+1)}}^{(t+1)})$  in the following way

- if  $M^{(t+1)} = M^{(t)} - 1$  by uniform random thinning on  $\{a_1^{(t)}, \dots, a_{M^{(t)}}^{(t)}\}$

- if  $M^{(t+1)} = M^{(t)} + 1$  by a random adding according to

$$\frac{\alpha^{(t+1)}}{\alpha^{(t+1)} + M^{(t)}} \text{Gamma}(1, b^{(t+1)}) + \frac{1}{\alpha^{(t+1)} + M^{(t)}} \sum_{k=1}^{M^{(t)}} \delta_{\alpha_k^{(t)}}$$

- Generate  $((y_1)^{(t+1)}, \dots, (y_{M^{(t+1)}})^{(t+1)})$  in the following way

- if  $M^{(t+1)} = M^{(t)} - 1$  by suppressing  $(y_j)$  corresponding to  $a_j$  previously suppressed

- if  $M^{(t+1)} = M^{(t)} + 1$  by adding  $(y_{M^{(t+1)}})^{(t+1)}$  according to  $\mathcal{U}(X)$ .

Fig. 2. Hybrid Gibbs Metropolis-Hastings algorithm

Expression (8) enables convenient use of the MCMC algorithm proposed in figure 2.

It is worth noticing that independent priors were chosen for  $\alpha$ ,  $b$ ,  $\mu$  and  $\sigma$ . Consequently, the Metropolis-Hastings ratios at the different steps of the algorithm may be simple to calculate, particularly for updating parameters  $\alpha$ ,  $b$  and  $\mu$ .

For each individual parameter, the MCMC sampling enables inferences made via posterior marginal distribution. In fact, at the completion of the MCMC run, we have a posterior joint distribution sample for the parameters of interest which provides a good approximation of their posterior probability law.

[23] offered a set of suggestions for choosing among models by examining the posterior distribution of the log-likelihood under each model. They introduced the deviance information criterion in aiming to combine measure of fit and complexity (effective number of parameters). The resulting criterion is the difference between the

posterior mean of the deviance and the deviance at the posterior mean which can be easily calculated from the MCMC posterior sample.

#### 4. ANALYSIS OF *IBICELLA LUTEA* COUNT DATA

Data provided by [16] are considered. They consist of 742 weeds (*Ibicella lutea*) in an australian farming paddock of area  $1200 \times 1200$  square meter. This weed dispersal mechanism is through seeds falling to the ground from ripen fruits. Additionally, fruits are carried away from plants by mammals. These latter unobserved events can be mathematically considered as the realization of an hidden process.

Figure 3 displays *Ibicella lutea* count data by means of a regular grid of size  $32 \times 32$ . In order to measure overdispersion and spatial autocorrelation at different scales, counts were aggregated with

Table 1. Dispersion and spatial autocorrelation test results for *Ibicella lutea* with respect to grid size

grid size	Mean	Variance	Dispersion index	Dispersion pvalue $\times 10^{-5}$	Moran index $\times 10^{-1}$	Moran pvalue $\times 10^{-2}$
$4^2$	46.25	142.33	3.08	5	3.32	3.9
$8^2$	11.56	18.60	1.61	157	1.74	4.16
$16^2$	2.89	3.98	1.38	6	1.50	0.075
$32^2$	0.72	0.87	1.20	1	0.64	0.37

Table 2. Bayesian inference results for *Ibicella lutea* count data

	Mean	Median	Standard error	95% HPD
$\alpha$	22.516	21.974	6.367	(11.920, 36.761)
$b$	5.157	5.880	1.750	(2.150, 7.456)
$\mu$	20.139	19.145	6.598	(9.808, 34.285)
$\sigma$	0.048	0.033	0.024	(0.030, 0.097)

different grid sizes and the dispersion and Moran autocorrelation indices were calculated and tested for each grid size (table 1). A significant overdispersion was found for each of the spatial scales which were considered. The spatial autocorrelation was significantly positive for grid sizes  $2^k \times 2^k, k \geq 2$  (table 1). This led us to apply the Cox process model defined from expressions (2) and (3). Then  $\mu$  stands for the expected number of dispersal events whereas  $\sigma$  is a seed dispersal parameter : the closer  $\sigma$  is to zero, the shorter are the dispersal distances. We assume that the contributions  $a_j$  to weed intensity follow a Dirichlet process centered on a Gamma distribution with scale parameter equal to unity and shape parameter  $b$ , denoted by  $Gamma(1, b)$ . In fact the Gamma distribution is frequently used in environmental effect modeling ([24]). Here parameter  $b$  is the expected contribution to intensity for a dispersal event since  $Gamma(1, b)$  is the marginal distribution of any  $a_j$ . When concentration parameter  $\alpha$  tends to zero, the contributions  $a_j$  are strongly and positively correlated. When  $\alpha$  tends to infinity, these contributions are independent.

The bayesian inference results are summarized in table 2 using posterior means, medians, standard deviations and 95% highest posterior density (HPD) intervals. The bayesian mean number of hidden events is around 20. Posterior mean contribution is 5.16. The influence parameter bayesian mean is 0.048. The concentration parameter estimates are around 22. The Bayesian deviance calculated at the posterior mean was chosen as criterion of goodness-of-fit ([23]). In table 5, the bayesian deviance result indicates that the hypothesis of equality of some contributions is more likely than the one of independent contributions.

In addition to information on structure of weed communities in crops such as the one provided by [21], these results may give some relevant information about unobserved events in the framework of weed control management. The proposed approach provides estimates of hidden process parameters influencing the weed dispersal and can be integrated into weed management programs.

## 5. ANALYSIS OF CHRONIC LOWER RESPIRATORY DISEASES DEATH NUMBERS

A set of data from Georgia (USA) was analyzed. These data available at the Web site <http://www.georgiastats.uga.edu/> consist of counts of death cases in each of the 159 counties of Georgia, for chronic lower respiratory diseases (CLRD) in 2007. For these diseases, death occurrences are sometimes the consequence of unobserved events due to environmental factor effects which generate

Table 3. Dispersion and spatial autocorrelation test results for death number per county in Georgia, US in 2007 for CLRD.

Mean	Variance	Dispersion index	Dispersion pvalue $\times 10^{-5}$	Moran index	Moran pvalue $\times 10^{-5}$
46.25	789.04	37.36	0	0.405	0

Table 4. Bayesian inference results for death number caused by CLRD per county in Georgia, US in 2007.

	Mean	Median	Standard error	95% HPD
$\alpha$	8.189	8.048	1.868	(4.928, 12.285)
$b$	1.498	1.489	0.241	(1.057, 1.993)
$\mu$	26.844	26.493	5.130	(17.643, 37.796)
$\sigma$	0.075	0.075	0.002	(0.070, 0.079)

Table 5. Bayesian deviance calculation for *Ibicella Lutea* count data (grid  $32 \times 32$ ) and Georgia death number per county for CLRD.

	Bayesian Deviance ( $\alpha = 0$ )	Bayesian Deviance ( $\alpha = +\infty$ )
Ibicella Lutea	1027.537	1048.074
CLRD	10245.690	10252.780

clustering. However, it is possible to formulate a bayesian cluster model taking into account such unobserved process of cluster centers ([15]).

Figure 4 shows the case counts for each county for CLRD along with the neighbourhood links between counties. The neighbourhood weight is based on whether or not two counties share borderlines with each other. The dispersion index and the Moran spatial index based on this neighbourhood were calculated (table 3). The corresponding p-values indicate a very strong overdispersion and positive spatial autocorrelation for CLRD.

Tables 4 and 5 show the results of bayesian inference for these diseases. The bayesian mean number of hidden events is 26.84 for chronic lower respiratory disease. The bayesian deviance is weaker under the hypothesis of equal contributions compared the one under the hypothesis of independent contributions. In other words, the equality of contributions is more likely than their independence.

## 6. CONCLUSION

In this article, a model of Cox process associated with a Dirichlet process was proposed with an emphasis on modeling spatial distributions of events generated by hidden occurrences. Event contributions are distributed according to a Dirichlet process centered on a Gamma distribution. A hybrid Gibbs-Metropolis-Hastings algorithm was developed. It provides the posterior distribution for the following parameters : expected number of hidden events per spatial unit, expected contribution per hidden event, spatial influence of hidden events and dispersion index due to hidden events. The hypotheses of contribution equality and contribution independence are compared by means of bayesian deviance calculation. Application to real data shows the potential of the method considered.

## 7. REFERENCES

- [1] D. Blackwell and J.B. Macqueen. Ferguson distributions via poly urn schemes. *Annals of Statistics*, 1(2):353–355, 1973.
- [2] P. Bremaud. *Point Processes and Queues: Martingale Dynamics*. 1981.

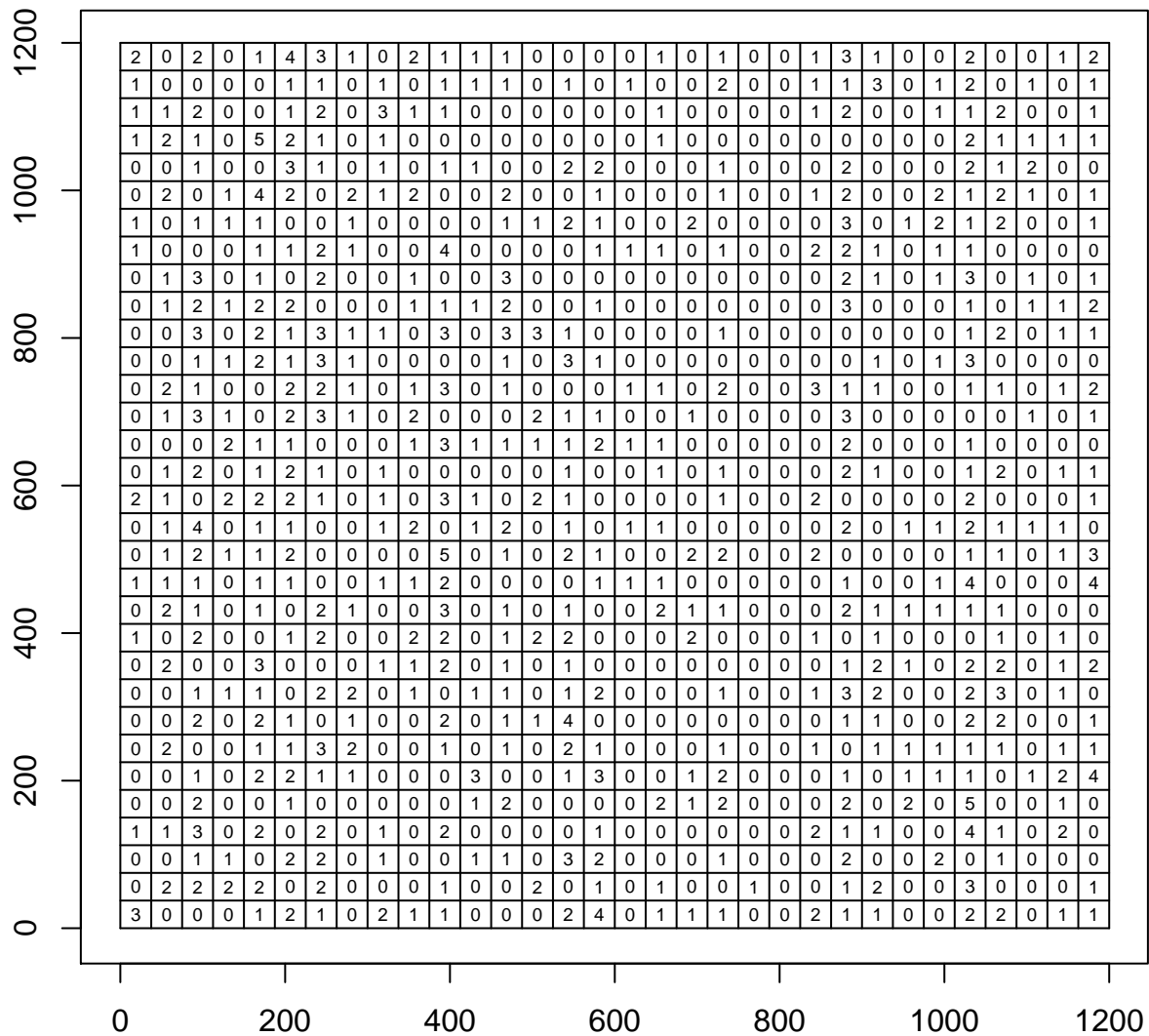


Fig. 3. 32 × 32 regular grid of *Ibicella lutea* spatial count data

[3] P. Bremaud. *Point Processes and Their Statistical Inference*. 1991.

[4] A. Brix and J. Chadœuf. Spatio-temporal modeling of weeds and shotnoise cox processes. *Biometrical Journal*, 44:83–99, 2002.

[5] Comas C., Mateu J., and Delicado P. On tree intensity estimation for forest inventories: Some statistical issues. *Biometrical Journal*, 23(6):994–1010, 2011.

[6] A.D. Cliff and K. Ord. *Spatial Processes: Models & Applications*. 1981.

[7] J. Cuzick and R. Edwards. Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society: Series B*, 52:73–104, 1990.

[8] D.J. Daley and D. Vere-Jones. *Introduction to the Theory of Point Processes. Vol 1: Elementary Theory and Methods (second edition)*. 2003.

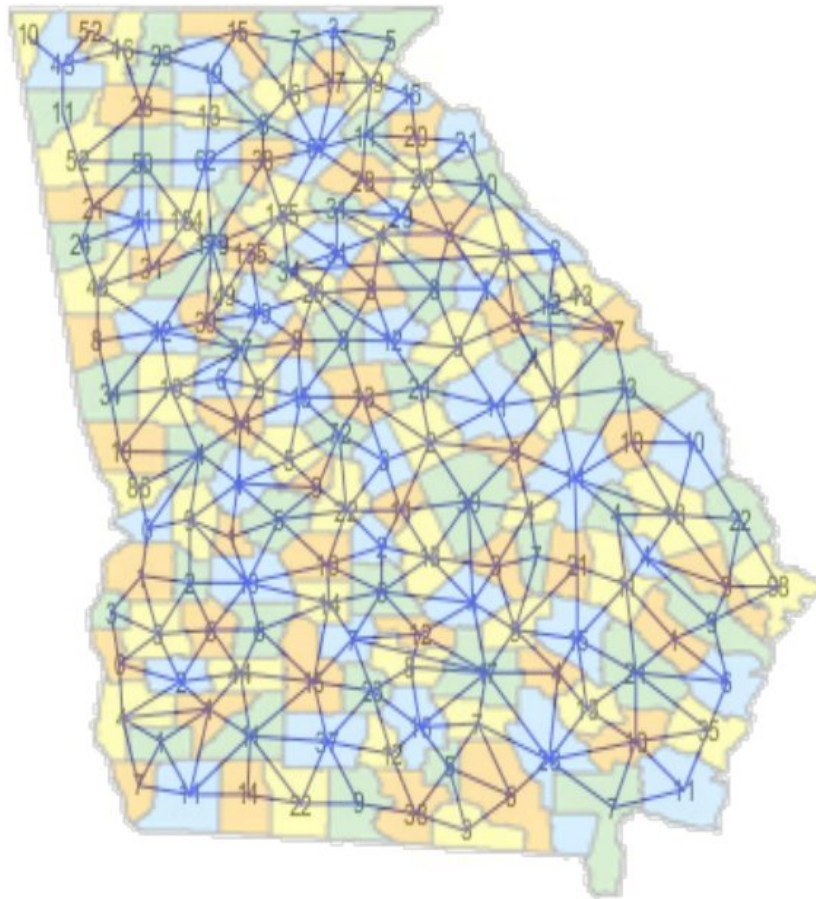


Fig. 4. Neighbourhood link and death number per county for Georgia state, US in 2007 for CLRD.

- [9] P. Diaconis and D. Freedman. Prior distributions on spaces of probability measures. *The Annals of Statistics*, 2:615–629, 1974.
- [10] P. Diaconis and D. Freedman. De finetti’s theorem for markov chains. *The Annals of Probability*, 8(1):115–130, 1980.
- [11] P. Diggle, B. Rowlingson, and T. Su. Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics*, 16:423–434, 2005.
- [12] P.J. Diggle. *Statistical Analysis of Spatial Point Patterns*. 1983.
- [13] T.S. Ferguson. A bayesian analysis of some nonparametric problems. *The Annals of Statistics*, 1:209–230, 1973.
- [14] A. Kottas, J. A. Duan, and A. E. Gelfand. Modeling disease incidence data with spatial and spatio-temporal dirichlet process mixtures. *Biometrical Journal*, 49:1–14, 2007.
- [15] A.B. Lawson. *Bayesian disease mapping: hierarchical modeling in spatial epidemiology*. 2009.
- [16] G.J. Melville and A.H. Welsh. Line transect sampling in small regions. *Biometrics*, 4(5):1130–1137, 2001.
- [17] J. Møller. Shot noise cox processes. *Advanced in Applied Probability*, 35:614–640, 2003.
- [18] J. Møller and G.L. Torrisi. Generalised shot noise cox processes. *Advances in Applied Probability*, 37:48–74, 2005.
- [19] J. Møller and R. Waagepetersen. Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34:643–684, 2007.
- [20] G.L.W Perry, B.P. Miller, and N.J. Enright. A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology*, 187(1):59–82, 2006.
- [21] S.L Poggio, E.H. Satorre, and E.B. de la Fuente. Structure of weed communities occurring in pea and wheat crops in the rolling pampa (argentina). *Agriculture, Ecosystems and Environment*, 103:225–235, 2004.
- [22] B.D. Ripley. *Spatial Statistics*. 1981.
- [23] S.D. Spiegelhalter, N.G. Best, B.P. Carlin, and A. van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, 4:583–639, 2002.
- [24] J. Vaillant. Negative binomial distributions of individuals and spatio-temporal cox processes. *Scandinavian Journal of Statistics*, 18:235–248, 1991.
- [25] J. Vaillant, Puggioni G., Waller L., and Daugrois J.-H. A spatio-temporal analysis of the spread of sugar cane yellow leaf virus. *Journal of Time Series Analysis*, 32:396–406, 2011.
- [26] T. Zhang and S.C. Kou. Non parametric inference of doubly stochastic poisson process data via the kernel method. *The Annals of Applied Statistics*, 4(4):1913–1941, 2010.