mathematics

Article

Cox Processes Associated with Spatial Copula Observed through Stratified Sampling

Walguen Oscar [1,2] and Jean Vaillant [1,*]

[1] Department of Mathematics and Computer Sciences, Université des Antilles, 97157 Pointe-à-Pitre, France; walguen.oscar@etu.univ-ag.fr
[2] Université d'Etat d'Haiti, HT6110 Port-au-Prince, Haiti
* Correspondence: jean.vaillant@univ-antilles.fr

Abstract: Cox processes, also called doubly stochastic Poisson processes, are used for describing phenomena for which overdispersion exists, as well as Poisson properties conditional on environmental effects. In this paper, we consider situations where spatial count data are not available for the whole study area but only for sampling units within identified strata. Moreover, we introduce a model of spatial dependency for environmental effects based on a Gaussian copula and gamma-distributed margins. The strength of dependency between spatial effects is related with the distance between stratum centers. Sampling properties are presented taking into account the spatial random field of covariates. Likelihood and Bayesian inference approaches are proposed to estimate the effect parameters and the covariate link function parameters. These techniques are illustrated using Black Leaf Streak Disease (BLSD) data collected in Martinique island.

Keywords: point process; Cox process; counting measure; overdispersion; stratified sampling; spatial copula; spatial sampling; likelihood; mixture distribution; negative multinomial

check for updates

Citation: Oscar, W.; Vaillant, J. Cox Processes Associated with Spatial Copula Observed through Stratified Sampling. *Mathematics* **2021**, *9*, 524. https://doi.org/10.3390/math9050524

Academic Editor: Mark Kelbert

Received: 31 January 2021
Accepted: 25 February 2021
Published: 3 March 2021

1. Introduction

Statistical Inference with spatial data requires to take into consideration a strong possibility of spatial dependency. In case of observations corresponding to the realization of a point process, the available data can be either point spatial locations or counts of points in spatial units [1,2]. These authors developed the mathematical theory of point processes which are used in many areas for describing event occurrences, like earthquakes, accidents, pest infestations, disease appearances, neuronal spikes, and many other situations [3–7]. Thus, a wide range of scientific applications can be cited, for example: ecology, epidemiology, geology, forestry, neurophysiology. De Oliveira [8] described a class of models for geostatistical count data generalizing the class proposed by Diggle et al., in 1998 [9]. Diggle [10] discussed about different spatial and spatio-temporal point process models, along with marked and multivariate counting processes. Wakefield [11] presented a critical review of spatial count data analysis methods for either disease mapping or spatial regression. He emphasizes the importance of the nature of variability in spatial risk and the fact that models must account for both spatial and non-spatial variability. Ickstadt and Wolpert [12] introduced a class of Bayesian hierarchical spatial models taking into account dependence on unobserved or unreported covariates. Examples and illustrations of state-of-the art developments in statistics for spatial data are presented in Cressie [13] with focus on lattice data, geostatistical data and point patterns. Sain and Cressie [14] emphasize that incorporating general forms of correlation is important in building spatial models for lattice data. Our goal is to consider Gamma-distributed effects which are spatially correlated. Given this constraint of marginal Gamma distributions, joint laws allowing such marginal properties, along with spatial dependence, are rare. This is the reason why we use the copula theory which allows both marginal laws and spatial dependence or spatial effects [15]. A spatial copula as defined in Durocher et al. [16] provide a full probabilistic model. Conversely, a covariance function describes only the covariance between marginal distributions except in the case of a multivariate Gaussian distribution which is fully specified by its expectation and covariance matrix.

In this paper, we consider count observations from a doubly stochastic Poisson process also called Cox process [8,17–20] on a measured space $(\mathcal{X}, \mathcal{B}, \nu)$ and driven by a random measure Λ such that there exists a positive random field $\lambda(.)$ verifying:

$$\forall B \in \mathcal{B}, \quad \Lambda(B) = \int_B \lambda(x)\nu(dx). \tag{1}$$

$\lambda(.)$ is the process intensity with respect to the reference measure ν. Its state space is $\mathcal{X}$ and $\Lambda(.)$ denotes its intensity measure. For example, ν can be the Lebesgue measure or a counting measure for individuals at risks.

The idea behind the model based on a Cox process is that environmental heterogeneity can generate overdispersion but also dependency. The counting measure N associated with such a Cox process verifies:

$$\forall B \in \mathcal{B}, N(B) \sim \mathcal{P}(\Lambda(B)) \mid \Lambda, \tag{2}$$

which means that, conditionally to Λ, the count $N(B)$ follows the Poisson distribution with parameter $\Lambda(B)$.

A well known result on doubly stochastic Poisson process is that stochasticity on Λ generates overdispersion: for any element B of $\mathcal{B}$ such that $E(\Lambda(B)) > 0$ and $Var(\Lambda(B)) > 0$, the dispersion index is

$$\frac{Var(N(B))}{E(N(B))} = \frac{E(\Lambda(B)) + Var(\Lambda(B))}{E(\Lambda(B))} = 1 + \frac{Var(\Lambda(B))}{E(\Lambda(B))} > 1.$$

If $\Lambda(B)$ is not a proper random variable or in other terms if $Var(\Lambda(B)) = 0$, then the dispersion index is equal to unity, what corresponds to a standard Poisson process. It is worth pointing out that modeling a Cox process is equivalent to modeling either its intensity $\lambda(.)$ or its driving measure $\Lambda(.)$. For example, Wolpert and Ickstadt [21] constructed λ as a linear combination of location-specific attributes and a kernel mixture measure associated with an independent-increment infinitely divisible random measure.

The general model studied in this paper is as follows. We assume that $\mathcal{X}$ has a finite n-partition $(\mathcal{X}_j)_{j \in J}$ and:

$$\lambda(x) = \sum_{j \in J} A_j f_\theta(Z(x)) \mathbb{1}_{\mathcal{X}_j}(x), \tag{3}$$

where A_j is a positive random variable, and $\mathbb{1}_{\mathcal{X}_j}$ stands for the indicator function of $\mathcal{X}_j$. We denote by $Z(x)$ the random vector of q covariates at point x and by f_θ the positive link function defined on $Z(\mathcal{X})$ and parameterized by θ. Therefore, Z is a random field taking values in a q-dimensional space and can not be observed entirely in practice. However we assume that for a given sampling unit, we have relevant summaries providing information on the Cox process. The unknown parameter θ belongs to $\mathbb{R}^{q'}$ with $q' \geq q$ and has to be estimated from observations of the Cox process and its q covariate summaries. Our aim is to present tools for statistical inference with data from the counting measure associated with a Cox process which intensity function given by equation (3), along with partial observations of its spatial covariates. This inference is on θ and also about the P_{A_j} parameters.

It should be noted that we consider count data derived from a stratified sampling procedure [22], whereas, in the literature, for each stratum $\mathcal{X}_j$, complete observations are available for the count measure and for the covariates [11], particularly for disease mapping [23]. We are, therefore, interested in a situation encountered in practice where the observations in each stratum are partial and carried out across sampling units. Furthermore, the spatial effects $(A_j)_{j \in J}$ are not necessarily mutually independent. Such a spatial correlation can be modeled by means of copula functions: Xue-Kun Song [24] presented a class of multivariate dispersion models based on multivariate Gaussian copula, Masarotto [25] identified Gaussian copula models for marginal regression analysis, Krupskii and Genton [26] proposed copula models for spatial data repeatedly observed in time, and Lee [27] developed a copula-based model for multisite precipitation occurrences.

In Section 2, distributional properties of the counting measure N are presented. A set of sampling units is considered in Section 3 with general features met in practice: the units are disjoint and entirely included in one of the partition elements $\mathcal{X}_j$. Firstly, the expressions of the likelihood conditional on the covariate field Z and the effect A_j are provided. Following Wakefield [11], spatially correlated log-normal effects are considered. This leads to observed counts in sampling units following the multivariate Poisson-log normal distribution [28]. Secondly, we show that the joint distribution of counts in the sampling units for each $\mathcal{X}_j$ is negative multinomial when the effects are gamma-distributed, and also that correlated gamma-distributed effects lead to dependent negative multinomial counts. Spatial correlations between effects are taken into account by means of Gaussian copula associated with Gamma distributed margins. In Section 4, the focus is on estimating the model parameters. In some situations, closed-form estimators may be obtained, but it is generally not the case. Likelihood inference can be performed when the effects are independent. Otherwise, a Markov chain Monte Carlo (MCMC) procedure is proposed in order to perform Bayesian inference [29]. An illustration is provided in Section 5 from BLSD data collected in Martinique Island within a network coordinated by the DAAL (Direction de l'Alimentation, de l'Agriculture et de la Forêt) and the FREDON Martinique (Fédération REgionale de Défense contre les Organismes Nuisibles de la Martinique).

2. Distributional Properties

In this section, we present the distributional properties associated with the counting measure N. These properties conditional on covariate vector Z give an idea of the dependency between the different spatial zones considered.

For any B in $\mathcal{B}$, let us write

$$S_\nu(B,\theta,Z) = \int_B f_\theta(Z(x))\nu(dx). \tag{4}$$

Here, the random measure $S_\nu(B,\theta,Z)$ depends on Z only through its restricted random field $Z_B = (Z(x))_{x\in B}$ on B.

The following proposition gives the expected number of points conditional on Z and $(A_j)_{j\in J}$:

Proposition 1. *Let N be the counting measure associated with the point process in which intensity is given by Equation (3); then,*

$$\forall B \in \mathcal{B}, \quad E(N(B)|(A_j)_{j\in J}, Z) = \sum_{j\in J} A_j S_\nu(B\cap\mathcal{X}_j,\theta,Z). \tag{5}$$

Proof of Proposition 1. For any element B of $\mathcal{B}$, we have

$$E(N(B)|(A_j)_{j\in J}, Z) = \Lambda(B) = \int_B \lambda(x)\nu(dx) = \int_B \sum_{j\in J} A_j f_\theta(Z(x))\mathbb{1}_{\mathcal{X}_j}(x)\nu(dx)$$

$$= \sum_{j\in J} A_j \int_{B\cap\mathcal{X}_j} f_\theta(Z(x))\nu(dx) = \sum_{j\in J} A_j S_\nu(B\cap\mathcal{X}_j,\theta,Z),$$

which shows (5). □

We have the following results about count correlation:

Proposition 2. *Consider the counting measure N associated with the point process defined by equation (3). For any (B_1, B_2) in $\mathcal{B}^2$, the following equality holds:*

$$Cov(N(B_1), N(B_2)|Z) = \sum_{(j,j')\in J^2} Cov(A_j, A_{j'}|Z) S_\nu(B_1 \cap \mathcal{X}_j, \theta, Z) S_\nu(B_2 \cap \mathcal{X}_{j'}, \theta, Z).$$

Proof of Proposition 2. From the covariance decomposition formula, we get

$$Cov(N(B_1), N(B_2)|Z) = Cov\Big(E(N(B_1)|(A_j)_{j\in J}, Z), E(N(B_2)|(A_j)_{j\in J}, Z)\Big) + E\Big(Cov(N(B_1), N(B_2)|(A_j)_{j\in J}, Z)\Big).$$

Since N is a Cox process, the second term of the right-hand side of the above equation is equal to zero so that Equation (5) leads to

$$Cov(N(B_1), N(B_2)|Z) = Cov\Big(\sum_{j\in J} A_j \int_{B_1\cap\mathcal{X}_j} f_\theta(Z(x)\nu(dx), \sum_{j'\in J} A_{j'} \int_{B_2\cap\mathcal{X}_{j'}} f_\theta(Z(x)\nu(dx)\Big),$$

and also to the final result. □

Corollary 1. *Consider the counting measure N associated with the point process defined by Equation (3) and assume the A_j are independent conditionally to Z. For any couple (B_1, B_2) in $\mathcal{B}^2$,*

$$Cov(N(B_1), N(B_2)|Z) = \sum_{j\in J} Var(A_j|Z) S_\nu(B_1 \cap \mathcal{X}_j, \theta, Z) S_\nu(B_2 \cap \mathcal{X}_j, \theta, Z).$$

Proof of Corollary 1. From Proposition 2 and conditional independence, whenever any distinct j, j' in J, we get $Cov(A_j, A_{j'}|Z) = 0$, and the final result follows. □

Corollary 2. *Consider the counting measure N associated with the point process defined by Equation (3) and assume the spatial effect variables $(A_j)_{j\in J}$ are mutually independent. Let B_1 and B_2 be two elements of $\mathcal{B}$. If $\exists (j, j') \in J^2$, with $j \neq j'$, and $B_1 \subset \mathcal{X}_j$, $B_2 \subset \mathcal{X}_{j'}$, then*

$$Cov(N(B_1), N(B_2)|Z) = 0.$$

Proof of Corollary 2. From corollary 1, the result is straightforward. □

Corollary 3. *Let B_1 and B_2 be two elements of $\mathcal{B}$. If there exists j in J, such that $B_1, B_2 \subset \mathcal{X}_j$, then*

$$Cov(N(B_1), N(B_2)|Z) = Var(A_j|Z) S_\nu(B_1 \cap \mathcal{X}_j) S_\nu(B_2 \cap \mathcal{X}_j).$$

Proof of Corollary 3. Both B_1 and B_2 are subsets of $\mathcal{X}_j$. Therefore, Proposition 2 give us the final result. □

It is worth noticing that, in the case where A_j is a non-degenerate random variable conditional on Z, then $Var(A_j|Z) > 0$, which implies that $Cov(N(B_1), N(B_2)|Z) > 0$.

3. Sampling Theory Results

In this section, we consider a set $\mathcal{S}$ of sampling units such that:
(C1) $\mathcal{S} = \{B_{ij} \in \mathcal{B}, i \in [\![1, n_j]\!], j \in J\}$,
(C2) $\forall j \in J, B_{ij} \subset \mathcal{X}_j$,
(C3) $\forall (i, i') \in [\![1, n_j]\!]^2, i \neq i' \Rightarrow B_{ij} \cap B_{i'j} = \emptyset$.
Condition (C3) means that there is no overlapping sampling units. When these three conditions are verified, we have the following results about the likelihood from the model defined by Equation (3).

Proposition 3. *Let $\mathcal{S}$ be a sample verifying conditions (C1), (C2), and (C3). For the model defined by Equation (3), if the joint distribution of $(A_j)_{j\in J}$ admits a probability density function g parameterized by β, then the conditional likelihood on $A_j = a_j$ for $\mathcal{S}$ with respect to the counting measure N and (θ, β) denoted by $\mathcal{L}\Big((\theta,\beta);(N(B))_{B\in\mathcal{S}},(a_j)_{j\in J}\Big)$ is equal to*

$$\left(\prod_{j\in J}\prod_{i=1}^{n_j}\frac{(a_jS_v(B_{ij},\theta,Z))^{N(B_{ij})}}{N(B_{ij})!}\exp(-a_jS_v(B_{ij},\theta,Z))\right)g\big((a_j)_{j\in J}\big). \tag{6}$$

It is worth noticing that this likelihood is conditional on Z.

Proof of Proposition 3. Since condition (C2) is verified, then $B_{ij} \subset \mathcal{X}_j$. Therefore, conditionally to A_j and Z, the count $N(B_{ij})$ is distributed according to the Poisson law $\mathcal{P}(A_jS_v(B_{ij},\theta,Z))$. Furthermore, the $N(B_{ij})$ are independent conditionally to the A_j. □

The likelihood given in Proposition 3 is useful for Bayesian approach when prior information on $(A_j)_{j\in J}$ are available, as well as sampling algorithms for augmented data. In practice, the A_j are not observed. Consequently, the likelihood techniques are based on count observations only. The following corollary provides such a likelihood:

Corollary 4. *Let $\mathcal{S}$ be a sample verifying conditions (C1), (C2), and (C3), if the joint distribution of $(A_j, j\in J)$ admits a probability density function g parameterized by β, then the unconditional likelihood for $\mathcal{S}$ with respect to the counting measure N and (θ,β) denoted by $\mathcal{L}\Big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\Big)$ is equal to*

$$\int_{\mathbb{R}_+^n}\left(\prod_{j\in J}\prod_{i=1}^{n_j}\frac{(a_jS_v(B_{ij},\theta,Z))^{N(B_{ij})}}{N(B_{ij})!}\exp(-a_jS_v(B_{ij},\theta,Z))\right)g\big((a_j)_{j\in J}\big)da_1\cdots da_n. \tag{7}$$

Proof of Corollary 4. The result is straightforward from Proposition 3. □

Many distributions can be used for the joint probability law of $(A_j, j \in J)$ which has to be considered as a mixing distribution in the framework of Poisson mixtures [30]. In the case where this joint probability law is a multivariate log-normal distribution $\mathcal{LN}(\mu,\Sigma)$ with $\mu\in\mathbb{R}^n$ and Σ a square matrix or order n, then the probability density function g is such that:

$$g\big((a_j)_{j\in J}\big) = \frac{\exp\left(-\frac{1}{2}\Big((\ln(a)-\mu)^t\Sigma^{-1}(\ln(a)-\mu)\Big)\right)}{(2\pi)^{n/2}(det\Sigma)^{1/2}\prod_{j\in J}a_j}, \tag{8}$$

where $\ln(a) = (\ln(a_j))_{j\in J}$ and n is the cardinality of J.

Following Wakefield [11], for any j in J, we can write $A_j = \exp(U_j + V_j)$, where the V_j are independent identically distributed according to $\mathcal{N}(0,\sigma_v^2)$ corresponding to non spatial contribution to overdispersion. On the other hand, the U_j are spatial contributions associated with a distance d on $\mathcal{X}$ such that $(U_j)_{j\in J}$ follows a multivariate normal distribution with $Cov(U_j,U_k) = \sigma_u^2\rho^{d(C_j,C_k)}$, where C_j and C_k are the centroids of $\mathcal{X}_j$ and $\mathcal{X}_k$, respectively. Consequently, $Cor(U_j,U_k) = \rho^{d(C_j,C_k)}$ so that parameter ρ stands for the correlation between U_j and U_k if $d(C_j,C_k) = 1$. We assume that $\rho \in]0,1[$. Therefore, in Equation (8), μ is the null vector of $\mathbb{R}^n$, and matrix Σ has each diagonal element equal to $\sigma_u^2+\sigma_v^2$, whereas element on line j and column k equal to $\sigma_u^2\rho^{d(C_j,C_k)}$. In the case where the mixture distribution is a multivariate log-normal distribution, Equations (7) and (8) provide tools for statistical analysis of counts in sampling units similarly to Aitchison and Ho [28].

An other choice of law for the effect A_j is the Gamma distribution with two parameters m_j and γ_j, where m_j is its expectation, and γ_j the shape parameter. Equation (9) gives its density function. It is worth noticing that several parameterizations can be met in the literature. We have the following theorem:

Theorem 1. *Let the random measure $S_\nu(B_{ij}, \theta, Z)$ be as in Equation (4). For any element j of J, and $(B_{ij})_{i \in [\![1,n_j]\!]}$ elements of $\mathcal{B}$ verifying conditions (C2) and (C3), if we assume that A_j follows a Gamma distribution $\Gamma(m_j, \gamma_j)$ with density function f_{A_j} given by:*

$$f_{A_j}(x) = \frac{x^{\gamma_j - 1}}{\Gamma(\gamma_j)} \left(\frac{\gamma_j}{m_j}\right)^{\gamma_j} \exp\left(-x\frac{\gamma_j}{m_j}\right), \tag{9}$$

then:

(i) $\Big(N(B_{ij})\Big)_{i \in [\![1,n_j]\!]}$ *follows a multinomial negative distribution,*

(ii) $\forall i \in [\![1, n_j]\!]$,

$$E(N(B_{ij})|Z) = m_j S_\nu(B_{ij}, \theta, Z)$$

and

$$Var(N(B_{ij})|Z) = E(N(B_{ij})|Z)\left(1 + \frac{E(N(B_{ij})|Z)}{\gamma_j}\right).$$

(iii) For two distinct elements i and i' of $[\![1, n_j]\!]$, the following equality holds:

$$Cov(N(B_{ij}, N(B_{i'j})|Z) = \frac{m_j^2}{\gamma_j} S_\nu(B_{ij}, \theta, Z) S_\nu(B_{i'j}, \theta, Z).$$

Proof of Theorem 1. (i) Using the probability generating function of $\Big(N(B_{ij})\Big)_{i \in [\![1,n_j]\!]}$, along with the moment generating function of A_j, leads us to the result.

In fact, for any j in J, if we write $N_j = \big(N(B_{ij})\big)_{i \in [\![1,n_j]\!]}$, then the probability generating function of N_j conditionally to A_j and Z at any point $(s_1, \ldots, s_{n_j})$ of $[0,1]^{n_j}$ is:

$$G_{N_j|A_j,Z}(s_1, \ldots, s_{n_j}) = E\left(\prod_{i=1}^{n_j} s_i^{N_{ij}} | A_j, Z\right) = \exp\left(-A_j \sum_{i=1}^{n_j} (1 - s_i) S_\nu(B_{ij}, \theta, Z)\right).$$

Therefore, the probability generating function of N_j conditional on Z is such that:

$$\begin{aligned} G_{N_j|Z}(s_1, \ldots, s_{n_j}) &= E\big(G_{N_j|A_j,Z}(s_1, \ldots, s_{n_j})\big) \\ &= E\left(\exp\big(-A_j \sum_{i=1}^{n_j} (1 - s_i) S_\nu(B_{ij}, \theta, Z)\big)\right) \\ &= M_{A_j}\left(-\sum_{i=1}^{n_j} (1 - s_i) S_\nu(B_{ij}, \theta, Z)\right), \end{aligned}$$

where M_{A_j} stands for the moment generating function of A_j.
Since $A_j \sim \Gamma(m_j, \gamma_j)$, then

$$G_{N_j|Z}(s_1, \ldots, s_{n_j}) = \left(1 + \frac{m_j}{\gamma_j} \sum_{i=1}^{n_j} (1 - s_i) S_\nu(B_{ij}, \theta, Z)\right)^{-\gamma_j},$$

so that N_j conditional on Z follows a multinomial negative distribution with parameters $\Big(m_j \big(S_\nu(B_{ij}, \theta, Z)\big)_{i \in [\![1,n_j]\!]}, \gamma_j\Big)$.

(ii) and (iii) The first derivative of $G_{N_j|Z}$ with respect to s_i at point 1 gives us $E(N(B_{ij})|Z)$. The second derivatives of $G_{N_j|Z}$ provide in a similar manner the second moments associated with N_j. $\square$

Proposition 4. *Let $\mathcal{S}$ be a sample verifying conditions (C1), (C2), and (C3). Assume the random variables $A_j, j \in J$, are mutually independent and distributed according to $\Gamma(m_j, \gamma_j)$ as defined by Equation (9). Writing $\beta = (m_j, \gamma_j)_{j\in J}$, then the likelihood $\mathcal{L}\Big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\Big)$ for $\mathcal{S}$ with respect to the counting measure N and (θ, β) is equal to*

$$\prod_{j=1}^{n}\frac{\Gamma(Y_{+j}+\gamma_j)}{\Gamma(\gamma_j)}\left(\frac{\gamma_j}{m_jS_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}\right)^{\gamma_j}\left(\frac{m_j}{m_jS_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}\right)^{Y_{+j}}\prod_{i=1}^{n_j}\frac{S_\nu(B_{ij},\theta,Z)^{Y_{ij}}}{Y_{ij}!}, \tag{10}$$

with $Y_{ij} = N(B_{ij})$, $Y_{+j} = \sum_{i=1}^{n_j} Y_{ij}$ and $S_{\nu+}(B_{\bullet j},\theta,Z) = \sum_{i=1}^{n_j} S_\nu(B_{ij},\theta,Z)$.

Proof of Proposition 4. Let Y_{ij} be a random variable on probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let y_{ij} be an element of $Y_{ij}(\Omega)$. Since the A_j are independent,

$$\mathbb{P}\Big(\bigcap_{j\in J}\bigcap_{i=1}^{n_j}(Y_{ij}=y_{ij})|Z\Big) = \prod_{j\in J}\mathbb{P}\Big(\bigcap_{i=1}^{n_j}(Y_{ij}=y_{ij})|Z\Big).$$

Moreover, each random variable A_j, $j \in J$, follows a Gamma distribution. From (i) in Theorem 1, $\Big(N(B_{ij})\Big)_{i\in[\![1,n_j]\!]}$ follows a multinomial negative distribution, and we get

$$\mathcal{L}\Big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\Big) =$$

$$\prod_{j\in J}\frac{\Gamma(y_{+j}+\gamma_j)}{\Gamma(\gamma_j)}\left(\frac{\gamma_j}{m_jS_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}\right)^{\gamma_j}\left(\frac{m_j}{m_jS_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}\right)^{y_{+j}}\prod_{i=1}^{n_j}\frac{S_\nu(B_{ij},\theta,Z)^{y_{ij}}}{y_{ij}!}$$

with $y_{+j} = \sum_i y_{ij}$. $\square$

In Proposition 4, we assume that the A_j are independent. Nevertheless, we can consider the case where $(A_j)_{j\in J}$ follows a multivariate Gamma distribution. We can refer to Kotz et al. [31] (Chapter 48) for an overview of multivariate Gamma distributions. We can also refer to Rahayu et al. [32] for statistical applications with multivariate Gamma distributions. In such a case of dependent univariate Gamma distributions, the result in Theorem 1 still holds so that we now have dependent negative multinomial counts. Thus, Equations (6) and (7) provide the conditional and unconditional likelihoods provided that we know the joint density function g of the A_j. The following proposition provides a joint gamma distribution for the effects $(A_j)_{j\in J}$ taking into account their spatial correlation.

Proposition 5. *Let $\mathcal{S}$ be a sample verifying conditions (C1), (C2), and (C3). For the model defined by Equation (3), if $(A_j)_{j\in J}$ follows a multivariate gamma distribution with joint density g parameterized by (β, ρ) and spatial correlation given by a copula density c, then the conditional likelihood on $A_j = a_j$, for $\mathcal{S}$ with respect to the counting measure N and (θ, β, ρ) is*

$$\begin{aligned}&\mathcal{L}\Big((\theta,\beta,\rho);(N(B))_{B\in\mathcal{S}},(a_j)_{j\in J}\Big) = \\ &\left(\prod_{j\in J}\prod_{i=1}^{n_j}\frac{(a_jS_\nu(B_{ij},\theta,Z))^{N(B_{ij})}}{N(B_{ij})!}\exp(-a_jS_\nu(B_{ij},\theta,Z))\right)c\Big(\big(F_{A_j}(a_j)\big)_{j\in J}\Big)\prod_{j\in J}f_{A_j}(a_j),\end{aligned} \tag{11}$$

where F_{A_j} is the cumulative distribution function of A_j, and f_{A_j} is given by Equation (9). ρ stands for the copula density parameter.

Proof of Proposition 5. The joint density g is the product of the marginal densities multiplied by the copula density c. Consequently, $g\Big((a_j)_{j\in J}\Big) = c\Big(\big(F_{A_j}(a_j)\big)_{j\in J}\Big)\prod_{j\in J} f_{A_j}(a_j)$. Proposition 3 leads to the final result. □

In the case where the copula density is constant equal to 1, the A_j are independent and expression (10) holds. In the sequel, we will use a multivariate gaussian copula [24] so that the correlation $cor(A_j, A_{j'})$ between two effects A_j and $A_{j'}$ depends on the distance $d(C_j, C_{j'})$ between centroids:

$$\forall u \in [0,1]^n, \quad c(u) = |W|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(u^*)^t(W^{-1} - I_n)u^*\right), \tag{12}$$

where
$u^* = (\phi^{-1}(u_1), \cdots, \phi^{-1}(u_n))$,
ϕ^{-1} is the inverse of the standard normal probability distribution function, and
$W = \left(\rho^{d(C_j,C_{j'})}\right)_{(j,j')\in J^2}$ with $\rho \in [0,1]$.

The greater ρ is, the higher the dependency between effects is. The greater the distance is between C_j and $C_{j'}$, the lower is the dependency between effects A_j and $A_{j'}$.

Note that Lee [27] has shown that, in the case of a bivariate Gaussian copula with gamma margins, $cor(A_j, A_{j'})$ cannot be resolved analytically but numerically.

4. Statistical Inference on the Model Parameters

In this section, we consider again a set $\mathcal{S}$ of sampling units verifying conditions C1, C2 and C3 as defined in Section 3, and effects distributed according to the gamma law. We first focus on the case where the effects $A_j, j \in J$, are independent and use likelihood techniques to estimate the model parameters. Then, the case of dependent A_j is considered within a Bayesian framework.

4.1. Case of Independent Effects

When the effects A_j are independent, the unconditional likelihood given by Equation (10) can be used to estimate the model parameters.

Proposition 6. *Let $\mathcal{S}$ be a sample verifying conditions (C1), (C2), and (C3). Assume the A_j are independent and distributed according to $\Gamma(m_j, \gamma_j)$ as defined by Equation (9). If θ is known, the maximum likelihood estimator for $\beta = (m_j, \gamma_j)_{j\in J}$ is such that:*

$$\widehat{m_j} = \frac{y_{+j}}{S_{v+}(B_{\bullet j}, \theta, Z)}$$

with $y_{+j} = \sum_i y_{ij}$ and $S_{v+}(B_{\bullet j}, \theta, Z) = \sum_{i=1}^{n_j} S_v(B_{ij}, \theta, Z)$.
The estimator $\widehat{\gamma}_j$ is obtained numerically from the following equation:

$$\sum_{k=0}^{y_{+j}-1} \frac{1}{\widehat{\gamma}_j + k} + \log\left(\frac{\widehat{\gamma}_j}{y_{+j} + \widehat{\gamma}_j}\right) = 0.$$

Proof of Proposition 6. The first order derivative of the log-likelihood $\log\Big(\mathcal{L}\big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\big)\Big)$ given by Equation (10) with respect to parameter m_j is as follows:

$$\frac{\partial \log \mathcal{L}\big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\big)}{\partial m_j} = \frac{\gamma_j y_{+j} - m_j\gamma_j S_{v+}(B_{\bullet j}, \theta, Z)}{m_j(m_j S_{v+}(B_{\bullet j}, \theta, Z) + \gamma_j)}.$$

Solving the equation $\dfrac{\partial \log \mathcal{L}\big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\big)}{\partial m_j} = 0$ for m_j gives:

$$\widehat{m_j} = \frac{y_{+j}}{S_{\nu+}(B_{\bullet j},\theta,Z)},$$

with $y_{+j} = \sum_i y_{ij}$ and $S_{\nu+}(B_{\bullet j},\theta,Z) = \sum_{i=1}^{n_j} S_\nu(B_{ij},\theta,Z)$.

The first order derivative of the log-likelihood $\log\Big(\mathcal{L}\big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\big)\Big)$ with respect to parameter γ_j is:

$$\frac{\partial \log \mathcal{L}\big((\theta,\beta);(N(B))_{B\in\mathcal{S}}\big)}{\partial \gamma_j} = \sum_{k=0}^{y_{+j}-1} \frac{1}{\gamma_j+k} + \log\Big(\frac{\gamma_j}{m_j S_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}\Big) + \frac{m_j S_{\nu+}(B_{\bullet j},\theta,Z) - y_{+j}}{m_j S_{\nu+}(B_{\bullet j},\theta,Z)+\gamma_j}.$$

By replacing m_j and γ_j with their maximum likelihood estimators, respectively, $\widehat{m_j}$ and $\widehat{\gamma_j}$, we get the final equation:

$$\sum_{k=0}^{y_{+j}-1} \frac{1}{\widehat{\gamma_j}+k} + \log\left(\frac{\widehat{\gamma_j}}{y_{+j}+\widehat{\gamma_j}}\right) = 0.$$

□

It is worth noticing that the maximum likelihood of γ_j does not depend on θ. The following corollary gives us the maximum likelihood estimators when θ is unknown.

Corollary 5. *Under conditions (C1), (C2), and (C3), if the A_j are independent and gamma-distributed, the maximum likelihood estimators for (θ,β), where $\beta = (m_j,\gamma_j)_{j\in J}$, are such that:*

$$\widehat{m_j} = \frac{y_{+j}}{S_{\nu+}(B_{\bullet j},\widehat{\theta},Z)},$$

with $S_{\nu+}(B_{\bullet j},\widehat{\theta},Z) = \sum_{i=1}^{n_j} S_\nu(B_{ij},\widehat{\theta},Z)$ and $\widehat{\theta}$ obtained from the following q equations:

$$\sum_{j\in J}\sum_{i=1}^{n_j}\left(\frac{y_{ij}}{S_\nu(B_{ij},\widehat{\theta},Z)} - \widehat{m_j}\right)\frac{\partial S_\nu(B_{ij},\widehat{\theta},Z)}{\partial \theta_l} = 0, \quad l = 1,\cdots,q.$$

$\widehat{\gamma_j}$ is obtained numerically from the same equation given in Proposition 6.

Proof of Corollary 5. This is a straightforward consequence of Proposition 6. □

Corollary 6. *Under conditions (C1), (C2), and (C3), and regularity conditions, if the A_j are independent and gamma-distributed, then the covariance matrix of $(\hat{\beta},\hat{\theta})$ is estimated by:*

$$M(\hat{\beta},\hat{\theta}) = \begin{pmatrix} H_{mm} & 0 & H_{m\theta} \\ 0 & H_{\gamma\gamma} & 0 \\ H_{\theta m} & 0 & H_{\theta\theta} \end{pmatrix}^{-1},$$

where

$$H_{mm} = diag\left[E\Big(\frac{\hat{\gamma}_j S_{\nu+}(B_{\bullet j},\hat{\theta},Z)}{y_{+j}+\hat{\gamma}_j}\Big)\right]_{j\in J}, \; H_{\gamma\gamma} = diag\left[E\Big(\sum_{k=0}^{y_{+j}-1}\frac{1}{(\hat{\gamma}_j+k)^2} - \frac{y_{+j}}{\hat{\gamma}_j(y_{+j}+\hat{\gamma}_j)^2}\Big)\right]_{j\in J},$$

$$H_{\theta\theta} = \left[E\Big(-\frac{\partial^2 \log \mathcal{L}\big((\hat{\beta},\hat{\theta});(N(B))_{B\in\mathcal{S}}\big)}{\partial\theta_l\partial\theta_s}\Big)\right]_{l,s} \text{ and } H_{m\theta} = \left[E\Big(-\frac{\partial^2 \log \mathcal{L}\big((\hat{\theta},\hat{\beta});(N(B))_{B\in\mathcal{S}}\big)}{\partial m_j\partial\theta_l}\Big)\right]_{j,l}.$$

Proof of Corollary 6. Differentiating twice the negative log-likelihood function with respect to the model parameters, we get the Fisher information as the expectation of this Hessian matrix. By replacing $\beta = (m_j, \gamma_j)_{j \in J}$ and θ with their maximum likelihood estimators, respectively, $\widehat{\beta}$ and $\widehat{\theta}$ in the second order derivatives, we get the final results. □

4.2. Case of Dependent Effects

When the A_j are dependent, the unconditional likelihood in Equation (7) is analytically intractable. Since the A_j are not observed, the conditional likelihood given by Equation (11) cannot be used directly for likelihood-based methods but rather within a Bayesian framework. For example, Lee [33] described implementations of spatial hierarchical models via simulation techniques. In fact, such MCMC methods provide tools for making inference about missing data and model parameters [29,34]. These MCMC techniques allow samples from the posterior distribution of $(\beta, \theta, \rho, (A_j)_{j \in J})$ to be drawn and then the calculation of any statistic based on the parameters and unobserved effects. This posterior distribution is proportional to

$$\mathcal{L}\Big((\theta, \beta, \rho); (N(B))_{B \in \mathcal{S}}, (a_j)_{j \in J}\Big) \times \pi(\theta)\pi(\beta)\pi(\rho), \tag{13}$$

where $\pi(\theta)$, $\pi(\beta)$, and $\pi(\rho)$ are the prior distributions for θ, β, and ρ.

Therefore, we propose a hybrid Gibbs-Metropolis-Hasting algorithm based on an acyclic directed graph (Figure 1) which provides posterior distribution samples for θ, β, and ρ. An example of such MCMC applications was presented in Valmy and Vaillant [35].

$$\begin{tikzcd}
\Big(d(C_j, C_{j'})\Big), \boxed{\rho} \arrow[d] & & \boxed{\beta = (m_j, \gamma_j)_{j \in J}} \arrow[d] \\
\textit{Gaussian copula} \arrow[dr] & & \textit{Gamma margins} \arrow[dl] \\
& (a_j) \arrow[d] & \\
& \Lambda_{\rho,\beta,\theta} \arrow[d] & \boxed{\theta}, Z \arrow[l] \\
& (N(B_{ij})) &
\end{tikzcd}$$

Figure 1. Acyclic directed graph of the hierarchical model defined by Proposition 5.

5. Analysis of BLSD Data

We performed our analysis on a dataset from Martinique consisting of spatial counts of the BLSD, a major foliar disease of bananas. Preliminary results show that four strata could be considered on the basis of the average annual rainfall (Figure 2).

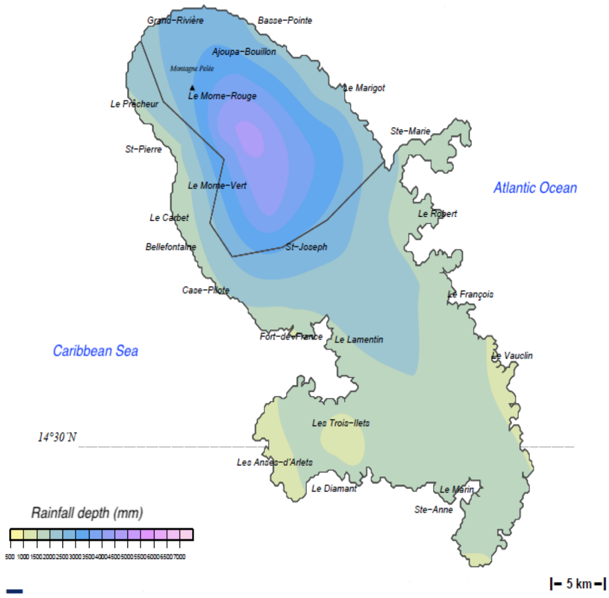

Figure 2. Average annual precipitation in Martinique (from Météo France).

On the other hand, because of sampling cost and effort, data could not be collected over the whole island (Figure 3). The proportion of Cavendish plots was the only covariate considered in this analysis. In fact, Cavendish banana is the most susceptible cultivar to BLSD and is mainly grown for export.

Figure 3. Available set of sampling units: black units are observed, grey units are not observed, white units are uncolonizable: high mountain and volcano (from Landry et al. (2021) [34]).

Table 1 summarizes some information on the number of positive cases with BLSD observed in each sampling unit. The overdispersion is very significant in each stratum.

Samples from the parameter's posterior distribution obtained by performing the MCMC procedure provide the estimates given in Table 2. From the 95% confidence intervals, we can deduce that the parameters ρ, $(m_j, \gamma_j)_{j\in\{1,2,3,4\}}$ are significantly different from zero at the 5% level. Only θ is not significantly different from zero which means that the proportion of Cavendish plot is not correlated to the number of positive cases. The m_j distributions are significantly different from each other (p-value $< 2.2 \times 10^{-16}$) and this is also the case for the γ_j distributions. The posterior distribution of ρ spreads over [0.03, 0.96] with a mode at 0.49 which implies a significant spatial dependency between strata. Note that the lower bound of the 95% posterior confidence interval for γ_4 is equal to $\exp(0.07) = 1.0725$. Therefore, parameter γ_4 is most likely greater than 1: the spatial effect in stratum 4 is significant.

Table 1. Test of dispersion for the number of positive cases. $n_j, \bar{x}_j, s_j^2, \widehat{Id}_j$ are, respectively, the number of sampling units, the mean number of positive cases, the variance of the positive case number, the estimated index of dispersion in stratum j. The pvalue corresponds to the test of overdispersion.

Stratum j	n_j	$\bar{x}_j$	s_j^2	$\widehat{Id}_j$	p-Value
1	150	1.21	2.80	2.32	0.0×10^0
2	60	1.90	4.50	2.37	1.7×10^{-8}
3	70	1.13	4.72	4.18	0.0×10^0
4	60	0.32	0.90	2.84	2.6×10^{-12}

Table 2. Markov chain Monte Carlo (MCMC) posterior distribution results.

Parameter	Posterior Mean	Posterior Standard Deviation	95% Posterior Confidence Interval
ρ	0.486	0.28	[0.03, 0.96]
θ	5.21	5.21	[−3.70, 11.32]
$\log(m_1)$	1.94	1.19	[−0.26, 3.63]
$\log(m_2)$	1.57	0.61	[0.26, 2.70]
$\log(m_3)$	1.29	0.53	[−0.07, 2.08]
$\log(m_4)$	4.29	1.47	[0.59, 5.69]
$\log(\gamma_1)$	−0.93	0.92	[−2.56, 0.70]
$\log(\gamma_2)$	−0.67	0.74	[−1.97, 0.65]
$\log(\gamma_3)$	1.67	1.56	[−0.55, 4.48]
$\log(\gamma_4)$	5.86	3.40	[0.07, 11.14]

6. Conclusions

We propose a model for spatial count data obtained from a stratified sampling in presence of overdispersion with spatial dependency built from a gaussian copula. Stratum random effects and spatial covariates can be considered and statistical inference can be performed within a Bayesian framework. Posterior samples of the model parameters are obtained by means of a hybrid Gibbs-Metropolis-Hasting algorithm. There are potential applications in various scientific fields, such as epidemiology and ecology. An illustration is given from data of Black Leaf Streak Disease (BLSD) on banana in Martinique. Differences from one stratum to another are significant, as well as their spatial dependency. A non-significant influence of proportion of Cavendish plots on the number of positive cases is found.

Author Contributions: W.O. and J.V. contributed equally to this work in conceptualization, methodology, software and original draft preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the French Embassy in Haiti (grant 973355H).

Acknowledgments: We would like to thank DAAF and FREDON Martinique for providing the BLSD spatio-temporal data, Météo France for providing Martinique's rainfall data, and the French Embassy in Haiti for partly funding Walguen Oscar's PhD work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brémaud, P. *Point Processes and Queues: Martingale Dynamics*; Springer: Berlin/Heidelberg, Germnay, 1981; Volume 50.
2. Daley, D.J.; Vere-Jones, D. *An Introduction to the Theory of Point Processes: Volume 1: Elementary Theory and Methods*; Springer: New York, NY, USA; Berlin/Heidelberg, Germnay, 2003.
3. Lánskỳ, P.; Vaillant, J. Stochastic model of the overdispersion in the place cell discharge. *Biosystems* **2000**, *58*, 27–32. [CrossRef]
4. Møller, J.; Ghorbani, M.; Rubak, E. Mechanistic spatio-temporal point process models for marked point processes, with a view to forest stand data. *Biometrics* **2016**, *72*, 687–696. [CrossRef]

5. Peng, R.D.; Schoenberg, F.P.; Woods, J.A. A space–time conditional intensity model for evaluating a wildfire hazard index. *J. Am. Stat. Assoc.* **2005**, *100*, 26–35. [CrossRef]
6. Vaillant, J.; Puggioni, G.; Waller, L.A.; Daugrois, J. A spatio-temporal analysis of the spread of sugarcane yellow leaf virus. *J. Time Ser. Anal.* **2011**, *32*, 396–406. [CrossRef]
7. Li, Z.; Cui, L.; Chen, J. Traffic accident modelling via self-exciting point processes. *Reliab. Eng. Syst. Saf.* **2018**, *180*, 312–320. [CrossRef]
8. De Oliveira, V. Hierarchical Poisson models for spatial count data. *J. Multivar. Anal.* **2013**, *122*, 393–408. [CrossRef]
9. Diggle, P.J.; Tawn, J.A.; Moyeed, R.A. Model-based geostatistics. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1998**, *47*, 299–350. [CrossRef]
10. Diggle, P.J. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*; CRC Press: Boca Raton, FL, USA, 2014.
11. Wakefield, J. Disease mapping and spatial regression with count data. *Biostatistics* **2007**, *8*, 158–183. [CrossRef] [PubMed]
12. Ickstadt, K.; Wolpert, R.L. Spatial regression for marked point processes. *Bayesian Stat.* **1998**, *6*, 323–341.
13. Cressie, N. *Statistics for Spatial Data*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
14. Sain, S.R.; Cressie, N. A spatial model for multivariate lattice data. *J. Econom.* **2007**, *140*, 226–259. [CrossRef]
15. Ma, X.; Luan, S.; Du, B.; Yu, B. Spatial copula model for imputing traffic flow data from remote microwave sensors. *Sensors* **2017**, *17*, 2160. [CrossRef] [PubMed]
16. Durocher, M.; Chebana, F.; Ouarda, T.B. On the prediction of extreme flood quantiles at ungauged locations with spatial copula. *J. Hydrol.* **2016**, *533*, 523–532. [CrossRef]
17. Cox, D.R. Some statistical methods connected with series of events. *J. R. Stat. Soc. Ser. B (Methodol.)* **1955**, *17*, 129–157. [CrossRef]
18. Møller, J.; Syversveen, A.R.; Waagepetersen, R.P. Log gaussian cox processes. *Scand. J. Stat.* **1998**, *25*, 451–482. [CrossRef]
19. Moller, J.; Waagepetersen, R.P. *Statistical Inference and Simulation for Spatial Point Processes*; CRC Press: Boca Raton, FL, USA, 2004.
20. Neyens, T.; Faes, C.; Molenberghs, G. A generalized Poisson-gamma model for spatially overdispersed data. *Spat. Spatio-Temporal Epidemiol.* **2012**, *3*, 185–194. [CrossRef]
21. Wolpert, R.L.; Ickstadt, K. Poisson/gamma random field models for spatial statistics. *Biometrika* **1998**, *85*, 251–267. [CrossRef]
22. Sharma, G. Pros and cons of different sampling techniques. *Int. J. Appl. Res.* **2017**, *3*, 749–752.
23. Pascutto, C.; Wakefield, J.; Best, N.; Richardson, S.; Bernardinelli, L.; Staines, A.; Elliott, P. Statistical issues in the analysis of disease mapping data. *Stat. Med.* **2000**, *19*, 2493–2519. [CrossRef]
24. Xue-Kun Song, P. Multivariate dispersion models generated from Gaussian copula. *Scand. J. Stat.* **2000**, *27*, 305–320. [CrossRef]
25. Masarotto, G.; Varin, C. Gaussian copula marginal regression. *Electron. J. Stat.* **2012**, *6*, 1517–1549. [CrossRef]
26. Krupskii, P.; Genton, M.G. Factor copula models for data with spatio-temporal dependence. *Spat. Stat.* **2017**, *22*, 180–195. [CrossRef]
27. Lee, T. Multisite stochastic simulation of daily precipitation from copula modeling with a gamma marginal distribution. *Theor. Appl. Climatol.* **2018**, *132*, 1089–1098. [CrossRef]
28. Aitchison, J.; Ho, C. The multivariate Poisson-log normal distribution. *Biometrika* **1989**, *76*, 643–653. [CrossRef]
29. Smith, A.F.; Roberts, G.O. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B (Methodol.)* **1993**, *55*, 3–23. [CrossRef]
30. Karlis, D.; Xekalaki, E. Mixed poisson distributions. *Int. Stat. Rev./Rev. Int. Stat.* **2005**, *73*, 35–58. [CrossRef]
31. Kotz, S.; Balakrishnan, N.; Johnson, N.L. *Continuous Multivariate Distributions, Volume 1: Models and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
32. Rahayu, A.; Prastyo, D.D. Multivariate Gamma Regression: Parameter Estimation, Hypothesis Testing, and Its Application. *Symmetry* **2020**, *12*, 813. [CrossRef]
33. Lee, D. CARBayes: An R package for Bayesian spatial modeling with conditional autoregressive priors. *J. Stat. Softw.* **2013**, *55*, 1–24. [CrossRef]
34. Landry, C.; Abadie, C.; Bonnot, F.; Vaillant, J. A Spatio-temporal Stochastic Model for an Emerging Plant Disease Spread in a Heterogeneous Landscape. *Int. J. Comput. Appl.* **2021**, *975*, 1–7.
35. Valmy, L.; Vaillant, J. Bayesian Inference on a Cox Process Associated with a Dirichlet Process. *Int. J. Comput. Appl.* **2014**, *95*, 1–7. [CrossRef]