Spatial Statistics 41 (2021) 100492

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A spatio-temporal multi-scale model for Geyer saturation point process: Application to forest fire occurrences

Morteza Raeisi [a,*], Florent Bonneu [a,b], Edith Gabriel [b]

[a] LMA EA2151, Avignon University, F-84000 Avignon, France
[b] INRAE, BioSP, F-84914 Avignon, France

ARTICLE INFO

Article history:
Received 21 October 2020
Received in revised form 22 December 2020
Accepted 13 January 2021
Available online 27 January 2021

Keywords:
Spatio-temporal Gibbs point processes
Hybridization
Pseudo-likelihood
Logistic likelihood
Forest fires

ABSTRACT

Because most natural phenomena exhibit dependence at multiple scales like locations of earthquakes or forest fire occurrences, spatio-temporal single-scale point process models are unrealistic in many applications. This motivates us to construct generalizations of classical Gibbs models. In this paper, we extend the Geyer saturation point process model to the spatio-temporal multi-scale framework. The simulation process is carried out through a birth–death Metropolis–Hastings algorithm. In a simulation study, we compare two common methods for statistical inference in Gibbs models: the pseudo-likelihood and logistic likelihood approaches that we tailor to this model. Finally, we illustrate this new model on forest fire occurrences modeling in Southern France.

1. Introduction

Nowadays point process models are widely used to highlight trends and interactions in the spatial or spatio-temporal distribution of events. Most of them are single-structure in the sense that they exhibit either spatial randomness (e.g. modeled by the Poisson process Kingman, 1993, 2006) or clustering (mostly modeled by Cox processes (Cox, 1972), in particular log-Gaussian Cox processes (Møller et al., 1998; Brix and Møller, 2001; Brix and Diggle, 2001; Diggle et al., 2013), Poisson Cluster processes (Neyman and Scott, 1958; Brix and Kendal, 2002; Gabriel, 2014) and

* Corresponding author.
E-mail address: morteza.raeisi@univ-avignon.fr (M. Raeisi).

https://doi.org/10.1016/j.spasta.2021.100492

Shot-Noise Cox processes (Brix and Chadœuf, 2000; Møller and Waagepetersen, 2004; Møller and Diaz-Avalos, 2010)) or inhibition (modeled by Strauss processes Strauss, 1975; Cronie and van Lieshout, 2015, Matérn hard core processes Matérn, 1960; Gabriel et al., 2013 and determinantal point processes Macchi, 1975; Lavancier et al., 2015). However, lot of phenomena present interactions at different scales what motivate statisticians to develop new models, mainly spatial models in ecology (Levin, 1992; Wiegand et al., 2007; Picard et al., 2009), epidemiology (Iftimi et al., 2017) or seismology (Siino et al., 2017, 2018b), but very few spatio-temporal models in environment (Gabriel et al., 2017) or epidemiology (Iftimi et al., 2018) as lately reviewed in Raeisi et al. (2019). Multi-scale models are mostly based on Gibbs models (see Dereudre, 2019 for a recent review on Gibbs models) as they offer a large class of models which allow any of the above mentioned interaction structure. Multi-structure models can then be obtained by hybridization (Baddeley et al., 2013).

Gibbs point processes are studied by their probability density, defined with respect to the unit rate Poisson point process. Well-known inhibitive Gibbs models include the hardcore model (events are forbidden to come too close together) and the Strauss model (Strauss, 1975) (pairs of close events are not impossible but are unlikely to occur). Generalizing the Strauss process, the Geyer saturation process (Geyer, 1999) intends to model both inhibition and clustering. It is able to take into account the clustering nature of a pattern due to interactions between points in absence of covariate information (Anwar and Stein, 2015).

Baddeley et al. (2013) defined a new class of multi-scale Gibbs point processes, so-called hybrid models. The hybridization technique consists in defining the density function of a multi-scale point process model as the product of several densities of Gibbs point processes, f_l for $l = 1, \dots, m$, so that $f = cf_1 \times \cdots \times f_m$ where c is a normalization constant. The choice of the normalization constant allows to well define a probability density in the case where the product of densities is integrable. In particular, Baddeley et al. (2013) introduced the spatial multi-scale Geyer saturation point process that has then been applied in epidemiology (Iftimi et al., 2017) and in seismology (Siino et al., 2017, 2018b). Iftimi et al. (2018) extended the hybridization approach to the spatio-temporal framework and introduced the spatio-temporal multi-scale area-interaction process. New hybrid Gibbs models can also be defined from the hardcore process (Cronie and van Lieshout, 2015) and the Strauss process (Gonzalez et al., 2016) introduced in the spatio-temporal framework, but much more hybrid Gibbs models remain to be developed to better describe spatio-temporal complex phenomena in practice.

Forest fire occurrences present multi-scale structures which are related to spatial or spatio-temporal inhomogeneities of environmental and climate covariates as well as influence of past events. Their complex interaction structure has been modeled by a spatio-temporal log-Gaussian Cox process in Opitz et al. (2020) and with an inhibitive effect as covariate in Gabriel et al. (2017). Gibbs point process models have also been considered in the spatial context for modeling wildfires like the area-interaction point process (Juan et al., 2012; Serra et al., 2013; Trilles et al., 2013; Arago et al., 2016; Woo et al., 2017) or the Geyer point process (Turner, 2009). In this paper, we aim to extend the spatial Geyer saturation point process to the spatio-temporal framework replacing the Euclidean balls by spatio-temporal cylindrical neighborhoods (Gonzalez et al., 2016). We also introduce its multi-scale version by extending the hybridization approach (Baddeley et al., 2013) to space and time. We then model forest fire occurrences using our spatio-temporal multi-scale Geyer saturation point process. Our data, available from the Prométhée database[1], concern forest fire occurrences in the Bouches-du-Rhône department (South of France) between 2001 and 2015.

The spatio-temporal multi-scale Geyer saturation point process model is introduced in Section 2. In Section 3, we extend the pseudo-likelihood and logistic likelihood approaches for statistical inference of Gibbs models to the spatio-temporal framework. Then in Section 4 we implement the model simulation using a birth–death Metropolis–Hastings algorithm and present a simulation study to compare the performance of the two estimation methods. Finally, in Section 5, we apply our model to forest fire occurrences in Southern France.

[1] https://www.promethee.com/en.

2. Spatio-temporal Geyer saturation point process

A spatio-temporal point process can be viewed as a random locally finite subset of a Borel set $W = S \times T \subset \mathbb{R}^2 \times \mathbb{R}$. We consider a complete, separable metric space $(W, d(\cdot,\cdot))$ where $d((u,v),(u',v')) := \max\{\|u-u'\|, |v-v'|\}$ for $(u,v),(u',v') \in W$. For $\mathcal{N}$ the state space of points configurations of W, $\mathbf{x} \in \mathcal{N}$ denotes a point pattern, i.e. $\mathbf{x} = \{(\xi_1,t_1),\ldots,(\xi_n,t_n)\}$ where (ξ_i,t_i) describes the location and time, respectively, associated with the ith event.

The cylindrical neighborhood $C_r^q(u,v)$ centered at $(u,v) \in W = S \times T$ is defined as

$$C_r^q(u,v) = \{(a,b) \in W = S \times T : \|u-a\| \le r, |v-b| \le q\}, \tag{1}$$

where $r, q > 0$ are spatial and temporal radii, $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $|\cdot|$ denotes the usual distance in $\mathbb{R}$. Note that $C_r^q(u,v)$ is a cylinder with center (u,v), radius r, and height $2q$ that represents a natural neighborhood for extending spatial Gibbs models to the spatio-temporal context (Gonzalez et al., 2016).

The Papangelou conditional intensity (Papangelou, 1974) of a spatio-temporal point process on W with density f is defined by

$$\lambda((u,v)|\mathbf{x}) = \frac{f(\mathbf{x} \bigcup (u,v))}{f(\mathbf{x} \backslash (u,v))}, \tag{2}$$

with $a/0 := 0$ for $a \ge 0$ and $(u,v) \in W$ (Cronie and van Lieshout, 2015). Hence, we have $\lambda((u,v)|\mathbf{x}) = \frac{f(\mathbf{x} \bigcup (u,v))}{f(\mathbf{x})}$ if $(u,v) \notin \mathbf{x}$ and $\lambda((u,v)|\mathbf{x}) = \frac{f(\mathbf{x})}{f(\mathbf{x} \backslash (u,v))}$ if $(u,v) \in \mathbf{x}$.

Gonzalez et al. (2016) introduced a spatio-temporal Strauss process with conditional intensity for $(u,v) \notin \mathbf{x}$

$$\lambda((u,v)|\mathbf{x}) = \lambda \gamma^{\tilde{n}(C_r^q(u,v);\mathbf{x})}, \tag{3}$$

where $\tilde{n}(C_r^q(u,v);\mathbf{x}) = \sum_{(\xi,t)\in\mathbf{x}} \mathbb{1}\{\|u-\xi\| \le r, |v-t| \le q\}$ is the number of points of $\mathbf{x}$ lying in $C_r^q(u,v)$.

The density function of Strauss model is not integrable for $\gamma > 1$, it thus does not define a valid probability density and the Strauss process cannot be intended for clustering structures. To avoid this issue, Geyer (1999) considers an upper bound (saturation parameter) for the number of neighboring points that interact and define the (spatial) Geyer saturation point process.

Definition 1. We define the *spatio-temporal Geyer saturation point process* as the point process with density

$$f(\mathbf{x}) = c \prod_{(\xi,t)\in\mathbf{x}} \lambda(\xi,t) \gamma^{\min\{s, n(C_r^q(\xi,t);\mathbf{x})\}}, \tag{4}$$

with respect to a unit rate Poisson process on W, where $c > 0$ is a normalizing constant, λ is a non-negative, measurable and bounded function, $\gamma > 0$ is the interaction parameter, s is the saturation parameter, and $n(C_r^q(\xi,t);\mathbf{x}) = \sum_{(u,v)\in\mathbf{x}\backslash(\xi,t)} \mathbb{1}(\|u-\xi\| \le r, |v-t| \le q)$ is the number of points of $\mathbf{x}$ lying in $C_r^q(\xi,t)$ and different from (ξ,t).

The function λ describes some spatio-temporal trend in point pattern that can be estimated using covariates. The scalars γ, r, q and s are the parameters of the model. The saturation parameter s is an upper bound of the number of points in the cylinder C_r^q. By using hybridization approach (Baddeley et al., 2013; Iftimi et al., 2018), we define a multi-scale version of (4).

Definition 2. We define the *spatio-temporal multi-scale Geyer saturation point process* as the point process with density

$$f(\mathbf{x}) = c \prod_{(\xi,t)\in\mathbf{x}} \lambda(\xi,t) \prod_{j=1}^{m} \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(\xi,t);\mathbf{x})\}}, \tag{5}$$

with respect to a unit rate Poisson process on W, where $\gamma_j > 0, j = 1,\ldots,m$, are the interaction parameters, and $r_1 < \cdots < r_m$, $q_1 < \cdots < q_m$ are spatial and temporal interaction ranges.

For any $j \in \{1, \ldots, m\}$, the interaction parameters $0 < \gamma_j < 1$ reflect inhibition, while $\gamma_j > 1$ reflect clustering between points at some spatio-temporal scales. When $s_j = 0$ or $\gamma_j = 1$ for all $j \in \{1, \ldots, m\}$, the density (5) corresponds to the density of an inhomogeneous Poisson process. Eq. (5) indicates that the structure of the process changes with the spatial and temporal distances r_j, q_j. Covariates can be added to the model by assuming that the spatio-temporal trend λ is function of a covariate vector $Z(\xi, t)$, i.e. $\lambda(\xi, t) = \Psi(Z(\xi, t))$.

Lemma 1. *The spatio-temporal multi-scale Geyer point process is a Markov point process in the sense of Ripley–Kelly (Ripley and Kelly, 1977) and its density (5) is measurable and integrable for all γ_j, $j = 1, \ldots, m$ with $m \in \mathbb{N}$.*

Proof. A Geyer model is hereditary, locally and Ruelle stable and hence integrable (Geyer, 1999). Baddeley et al. (2013) showed these properties for hybrids. As in Iftimi et al. (2018), we can show that the spatio-temporal Geyer saturation point process (4) is a Markov point process in Ripley–Kelly's sense at interaction range $2\max\{r, q\}$ and that the spatio-temporal multi-scale Geyer saturation process (5) is also a Markov point process in Ripley–Kelly sense at interaction range $\max_{1\le j\le m}\{2\max\{r_j, q_j\}\} = 2\max\{r_m, q_m\}$ (Baddeley et al., 2013). □

For any $(u, v) \in W$, the Papangelou conditional intensity function of the spatio-temporal multi-scale Geyer saturation process is

$$\lambda((u,v)|\mathbf{x}) = \lambda(u,v)\prod_{j=1}^{m}\gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(u,v);\mathbf{x})\}} \times \prod_{(\xi,t)\in\mathbf{x}\setminus(u,v)} \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(\xi,t);\mathbf{x}\cup(u,v))\} - \min\{s_j, n(C_{r_j}^{q_j}(\xi,t);\mathbf{x}\setminus(u,v))\}}. \tag{6}$$

The Markovian property (Lemma 1) ensures that this conditional intensity only depends on (u, v) and its neighbors in $\mathbf{x}$. Hence, we can design simulation algorithms for generating realizations of the model, see Section 4.

3. Inference

Geyer saturation point process model (4) involves two types of parameters: regular parameters and irregular parameters. A parameter is called *regular* if the log likelihood is a linear function of that parameter, *irregular* otherwise. Regular parameters like trend λ and interaction γ can be estimated using the pseudo-likelihood method (Baddeley and Turner, 2000) or the logistic likelihood method (Baddeley et al., 2014) rather than the maximum likelihood method (Ogata and Tanemura, 1981). Indeed, they are based on the conditional intensity which is tractable for most Gibbs models and is free from the normalization constant c (whose estimation is computationally very expensive, even for a small number of regular parameters). Here we tailor these two methods to estimate regular parameters of our spatio-temporal model and we compare their performance in the next section.

Irregular parameters, like saturation threshold s and distances r and q, are difficult to estimate using the maximum likelihood method because the likelihood function is not differentiable with respect to them. These parameters can be estimated using the profile pseudo-likelihood approach (Baddeley and Turner, 2000) or predetermined by the user using some summary statistics, like the pair correlation and the auto-correlation functions (Iftimi et al., 2018), in order to determine the interaction ranges. Baddeley and Turner (2006) presented the methods that are used for irregular parameter estimation in the spatial framework.

In this paper, we combine the advantages of the two previous methodologies. By computing some statistics summarizing the range of interactions in space and time, we consider a set of feasible irregular parameter values and we choose the combination of them providing the best Akaike's Information Criterion (AIC) for the fitted model.

3.1. Pseudo-likelihood approach

Let $\boldsymbol{\theta}$ be the vector of regular parameters that we aim to estimate. Besag (1977) defined the pseudo-likelihood for spatial point processes in order to avoid computational problems with point process likelihoods. One can easily extend it for a spatio-temporal point process with conditional intensity $\lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x})$ over W as follows

$$PL(\mathbf{x};\boldsymbol{\theta}) = \exp\left(-\int_S\int_T \lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x})dvdu\right)\prod_{(\xi,t)\in\mathbf{x}}\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}). \tag{7}$$

The pseudo score is defined by

$$U(\mathbf{x};\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}}\log PL(\mathbf{x};\boldsymbol{\theta}), \tag{8}$$

that is an unbiased estimating function. The maximum pseudo-likelihood normal equations are then given by

$$\frac{\partial}{\partial\boldsymbol{\theta}}\log PL(\mathbf{x};\boldsymbol{\theta}) = 0, \tag{9}$$

where

$$\log PL(\mathbf{x};\boldsymbol{\theta}) = \sum_{(\xi,t)\in\mathbf{x}}\log\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}) - \int_S\int_T\lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x})dvdu, \tag{10}$$

and $\lambda_{\boldsymbol{\theta}}(\cdot|\mathbf{x})$ is defined by (6) for hybrid Geyer model (5).

For sake of clarity, we now assume that $\boldsymbol{\theta} = [\log\gamma_1,\ldots,\log\gamma_m]^\top$ the logarithm of interaction parameters in model (5). To estimate $\boldsymbol{\theta}$, we use the pseudo-likelihood approach. Eq. (6) can be rewritten as $\lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x}) = \lambda(u,v)\prod_{j=1}^m\exp(\theta_j S_j((u,v),\mathbf{x}))$ where

$$\begin{aligned} S_j((u,v),\mathbf{x}) = {} & \min\{s_j, n(C_{r_j}^{q_j}(u,v);\mathbf{x})\} \\ & + \sum_{(\xi,t)\in\mathbf{x}\setminus(u,v)}[\min\{s_j, n(C_{r_j}^{q_j}(\xi,t);\mathbf{x}\cup(u,v))\} \\ & - \min\{s_j, n(C_{r_j}^{q_j}(\xi,t);\mathbf{x}\setminus(u,v))\}], \end{aligned} \tag{11}$$

is a sufficient statistics. Then, for $\mathbf{S}((u,v),\mathbf{x}) = [S_1((u,v),\mathbf{x}),\ldots,S_m((u,v),\mathbf{x})]^\top$

$$\log\lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x}) = \log\lambda(u,v) + \boldsymbol{\theta}^\top\mathbf{S}((u,v),\mathbf{x}) \tag{12}$$

is a linear model in $\boldsymbol{\theta}$ with offset $\log\lambda(u,v)$. Thus, Eq. (9) gives us the pseudo-likelihood equations

$$\frac{\partial}{\partial\boldsymbol{\theta}}\left[\sum_{(\xi,t)\in\mathbf{x}}[\log\lambda(\xi,t) + \sum_{j=1}^m\theta_j S_j((\xi,t),\mathbf{x})] - \int_S\int_T\lambda(u,v)\prod_{j=1}^m e^{\theta_j S_j((u,v),\mathbf{x})}dvdu\right] = 0, \tag{13}$$

For each parameter θ_i, $i = 1,\ldots,m$, Eqs. (13) can be rewritten

$$\sum_{(\xi,t)\in\mathbf{x}} S_i((\xi,t),\mathbf{x}) = \int_S\int_T\lambda(u,v)S_i((u,v),\mathbf{x})\prod_{j=1}^m e^{\theta_j S_j((u,v),\mathbf{x})}dvdu, \tag{14}$$

The major difficulty is to estimate the integrals on the right hand side of Eqs. (14). The pseudo-likelihood cannot be computed exactly but must be approximated numerically.

For a point process model, the approximation of likelihood is converted into a regression model. In the following, we refer to generalized log-linear Poisson regression approach as approximation of integrals in (14). In the next subsection, we also investigate an alternative, the logistic regression.

Berman and Turner (1992) developed a numerical quadrature method to approximate maximum likelihood estimation for an inhomogeneous Poisson point process. Berman–Turner method has then been extended to Gibbs point processes by Baddeley and Turner (2000), approximating the integral in (10) by a Riemann sum

$$\int_S \int_T \lambda_{\boldsymbol{\theta}}((u,v)|\mathbf{x}) dv du \approx \sum_{k=1}^{n+p} w_k \lambda_{\boldsymbol{\theta}}((\xi_k, t_k)|\mathbf{x}), \tag{15}$$

where (ξ_k, t_k) are points in $\{(\xi_1, t_1), \ldots, (\xi_n, t_n), (\xi_{n+1}, t_{n+1}), \ldots, (\xi_{n+p}, t_{n+p})\} \in W$ consisting of the n events of $\mathbf{x}$ and p dummy points, and w_k are quadrature weights such that $\sum_{k=1}^{n+p} w_k = \ell(S \times T)$ where ℓ is Lebesgue measure. This yields an approximation for the log pseudo-likelihood of the form

$$\log PL(\mathbf{x}; \boldsymbol{\theta}) \approx \sum_{(\xi,t)\in\mathbf{x}} \log \lambda_{\boldsymbol{\theta}}((\xi, t)|\mathbf{x}) - \sum_{k=1}^{n+p} w_k \lambda_{\boldsymbol{\theta}}((\xi_k, t_k)|\mathbf{x}). \tag{16}$$

Note that if the set of points $\{(\xi_k, t_k), k = 1, \ldots, n+p\}$ includes all the points of $\mathbf{x} = \{(\xi_1, t_1), \ldots, (\xi_n, t_n)\}$, we can rewrite (16) as

$$\log PL(\mathbf{x}; \boldsymbol{\theta}) \approx \sum_{k=1}^{n+p} w_k \left(y_k \log \lambda_{\boldsymbol{\theta}}((\xi_k, t_k)|\mathbf{x}) - \lambda_{\boldsymbol{\theta}}((\xi_k, t_k)|\mathbf{x}) \right), \tag{17}$$

where

$$y_k = \begin{cases} 1/w_k, & \text{if } (\xi_k, t_k) \in \mathbf{x} \text{ is an event,} \\ 0, & \text{if } (\xi_k, t_k) \notin \mathbf{x} \text{ is a dummy point.} \end{cases} \tag{18}$$

The right hand side of (17), for fixed $\mathbf{x}$, is formally equivalent to the log-likelihood of independent Poisson variables $Y_k \sim Poisson(\lambda_{\boldsymbol{\theta}}((\xi_k, t_k)|\mathbf{x}))$ taken with weights w_k. Therefore, by using the `glm` function in `R` (R Core Team, 2016), we can perform the maximum likelihood-based parameter estimation of this Poisson generalized linear model and obtain the maximum value for (17).

Note that in hybrid Geyer model (5), we consider $\lambda(\xi, t) = \lambda_\beta(\xi, t) = \beta\mu(\xi, t)$ where $\mu(\xi, t)$ is known or estimated beforehand and β is a parameter to estimate. In summary, the method is as follows.

Algorithm 1

- Generate a set of p uniform dummy points in W and merge them with all the data points in $\mathbf{x}$ to construct the set of quadrature points $(\xi_k, t_k) \in W$ with $k = 1, \ldots, n+p$.
- Compute the quadrature weights w_k and the indicators y_k defined in (18),
- Compute the sufficient statistics $\boldsymbol{S}((\xi_k, t_k), \mathbf{x})$ at each quadrature point,
- Fit a log-linear Poisson regression with explanatory variables $\boldsymbol{S}((\xi_k, t_k), \mathbf{x})$, and offset $\log \lambda(\xi_k, t_k)$ on the responses y_k with weights w_k to obtain estimates $\hat{\boldsymbol{\theta}}$ for the $\boldsymbol{S}$-vector and intercept $\hat{\theta}_0$,
- Return the maximum pseudo-likelihood-based parameter estimates $\hat{\gamma}_j = \exp(\hat{\theta}_j)$ for $j = 1, \ldots, m$ and $\hat{\beta} = \exp(\hat{\theta}_0)$.

We define the quadrature scheme by defining a spatio-temporal partition of W into cubes C_k of equal volumes ν and by using the counting weights proposed in Baddeley and Turner (2000). We then assign to each dummy or data point (ξ_k, t_k) a weight $w_k = \nu/n_k$ where n_k is the number of dummy and data points that lie in the same cube as (ξ_k, t_k). The number of dummy points should be sufficient for an accurate estimate of the pseudo-likelihood. We follow Baddeley and Turner (2000) and start with $p \approx 4n(\mathbf{x})$. Then, we increase it until $\sum_k w_k = \ell(W)$, what can lead to high computational costs.

An alternative way to define the quadrature scheme for *Algorithm 1* is based on Dirichlet tessellation (Baddeley and Turner, 2000) and the weight of each point is equal to the volume of the corresponding Dirichlet 3D cell. In this paper, we consider cubes because it is less time consuming and provides similar results (see Opitz, 2009 for quadrature schemes comparison of 3D Gibbs point processes).

3.2. Logistic likelihood approach

The logistic likelihood method (Baddeley et al., 2014) is an alternative for estimating the regular parameters of Gibbs models that is closely related to the pseudo-likelihood method. The Berman–Turner approximation often requires a quite large number of dummy points. Hence, fitting such GLM can be computationally intensive, especially when dealing with a large dataset. Baddeley et al. (2014) formulated the pseudo-likelihood estimation equation as a logistic regression using auxiliary dummy point configurations and proposed a computational technique for fitting Gibbs point process models to spatial point patterns. Iftimi et al. (2018) extended the logistic likelihood approach for spatio-temporal Gibbs point processes and we tailored it to our model.

Let $\mathbf{x}$ be a realization of a spatio-temporal point process defined on W having a density $f_{\boldsymbol{\theta}}$ with respect to the unit rate Poisson process and with conditional intensity function $\lambda_{\boldsymbol{\theta}}(\cdot|\mathbf{x})$. We consider an independent Poisson process for dummy points, with intensity function ρ, and we denote by $\boldsymbol{d}$ a set of dummy points. We follow Baddeley et al. (2014) (resp. Iftimi et al., 2018) for choosing ρ of a homogeneous (resp. inhomogeneous) Poisson process in simulation study (resp. application). See Baddeley et al. (2014), for a data-driven determination of ρ and its effect on efficiency and practicability of the method.

By defining $Y(\xi,t) = \mathbb{1}_{\{(\xi,t)\in\mathbf{x}\}}$ for $(\xi,t)\in \mathbf{x}\cup\boldsymbol{d}$, we obtain independent Bernoulli variables taking one for data points and zero for dummy points. We have

$$Pr(Y(\xi,t)=1) = \frac{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))}{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t)) + \rho(\xi,t)}, \tag{19}$$

By considering the log linearity assumption for the conditional intensity $\lambda_{\boldsymbol{\theta}}(\cdot|\mathbf{x})$ in (12), the logit of $Pr(Y(\xi,t)=1)$ is

$$\log\frac{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))}{\rho(\xi,t)} = \log\frac{\lambda(\xi,t)}{\rho(\xi,t)} + \sum_{j=1}^{m}\theta_j S_j((\xi,t),\mathbf{x}\backslash(\xi,t)), \tag{20}$$

which is a linear model in $\boldsymbol{\theta}$ with offset $\log\frac{\lambda(\xi,t)}{\rho(\xi,t)}$.

Since $\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}) = \lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))$ for $(\xi,t)\in\boldsymbol{d}$, the log logistic likelihood is defined by

$$\begin{aligned}\log LL(\mathbf{x},\boldsymbol{d};\boldsymbol{\theta}) &= \sum_{(\xi,t)\in\mathbf{x}\cup\boldsymbol{d}} Y((\xi,t))\log\frac{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))}{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))+\rho(\xi,t)}\\ &\quad+\sum_{(\xi,t)\in\mathbf{x}\cup\boldsymbol{d}}[1-Y((\xi,t))]\log\frac{\rho(\xi,t)}{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x})+\rho(\xi,t)}\\ &=\sum_{(\xi,t)\in\mathbf{x}}\log\frac{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))}{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x}\backslash(\xi,t))+\rho(\xi,t)}\\ &\quad+\sum_{(\xi,t)\in\boldsymbol{d}}\log\frac{\rho(\xi,t)}{\lambda_{\boldsymbol{\theta}}((\xi,t)|\mathbf{x})+\rho(\xi,t)}.\end{aligned} \tag{21}$$

The maximum of the log-logistic likelihood exists and under regularity condition (Baddeley et al., 2019) is unique. Hence, estimation can be implemented in `R` by using the `glm` function.

As in *Algorithm 1*, we consider $\lambda(\xi,t) = \lambda_\beta(\xi,t) = \beta\mu(\xi,t)$ and we estimate the regular parameters form the following algorithm.

Algorithm 2

- Generate dummy points $\boldsymbol{d}$ from a Poisson process with intensity function ρ and merge them with all the data points in $\mathbf{x}$ to construct the set of quadrature points $(\xi_k,t_k)\in W$,
- Obtain the response variables y_k (1 for data points, 0 for dummy points),
- Compute the sufficient statistics $\boldsymbol{S}((\xi_k,t_k),\mathbf{x}\backslash(\xi_k,t_k))$ at each quadrature point,
- Fit a logistic regression model with explanatory variables $\boldsymbol{S}((\xi_k,t_k),\mathbf{x}\backslash(\xi_k,t_k))$, and offset $\log(\mu(\xi_k,t_k)/\rho(\xi_k,t_k))$ on the responses y_k to obtain estimates $\hat{\boldsymbol{\theta}}$ for the $\boldsymbol{S}$-vector and intercept $\hat{\theta}_0$,

- Return the parameter estimator $\hat{\boldsymbol{\gamma}} = \exp(\hat{\boldsymbol{\theta}})$ and $\hat{\beta} = \exp(\hat{\theta}_0)$ and in the case where $\mu(\xi_k, t_k)/\rho(\xi_k, t_k)$ is a constant c we have $\hat{\beta} = c^{-1}\exp(\hat{\theta}_0)$.

4. Simulation

The simulation algorithms of Gibbs point process models require only computation of the Papangelou conditional intensity which avoids to consider the difficult estimation of the unknown normalizing constant in the density function. Gibbs point process models can be simulated by using Markov chain Monte Carlo (MCMC) algorithms like the birth–death Metropolis–Hastings algorithm (Møller and Waagepetersen, 2004) that belongs to the large class of Metropolis–Hastings algorithms (Geyer and Møller, 1994). In this section, we first present the birth–death Metropolis–Hastings algorithm and secondly we investigate the goodness of parameter estimation of the two approaches introduced before.

4.1. Birth–death Metropolis–Hastings algorithm

For $\mathbf{x}$ a spatio-temporal point pattern in W, we can propose either a birth with probability $q(\mathbf{x})$ or a death with probability $1 - q(\mathbf{x})$. For a birth, a new point $(u, v) \in W$ is sampled from a probability density $b(\mathbf{x}, \cdot)$ and the new point configuration $\mathbf{x} \cup (u, v)$ is accepted with probability $A(\mathbf{x}, \mathbf{x} \cup (u, v))$, otherwise the state remains unchanged. For a death, the point $(\xi, t) \in \mathbf{x}$ chosen to be removed is selected according to a discrete probability distribution $d(\mathbf{x}, .)$ on $\mathbf{x}$, and the proposal $\mathbf{x} \setminus (\xi, t)$ is accepted with probability $A(\mathbf{x}, \mathbf{x} \setminus (\xi, t))$, otherwise the state remains unchanged. For simplicity, we consider $q(\mathbf{x}) = \frac{1}{2}$, $b(\mathbf{x}, \cdot) = 1/\ell(W)$ and $d(\mathbf{x}, \cdot) = 1/n(\mathbf{x})$. By setting $A(\mathbf{x}, \mathbf{x} \cup (u, v)) = \min\{1, r((u, v); \mathbf{x})\}$, and $A(\mathbf{x}, \mathbf{x} \setminus (\xi, t)) = \min\{1, 1/r((\xi, t); \mathbf{x} \setminus (\xi, t))\}$ where $r((u, v); \mathbf{x}) = \frac{\ell(W)}{n(\mathbf{x})+1} \times \lambda((u, v)|\mathbf{x})$ is the Hastings ratio (Iftimi et al., 2018), we obtain the following birth–death Metropolis–Hastings algorithm.

Algorithm 3

For $n = 0, 1, \ldots$, given $X_n = \mathbf{x}$ (e.g. a Poisson process for $n = 0$), generate X_{n+1}:

- Generate two uniform numbers y_1, y_2 in $[0, 1]$,
- If $y_1 \leq \frac{1}{2}$ then
  - A new point (u, v) is uniformly sampled from a probability density $1/\ell(W)$,
  - Compute $r((u, v); \mathbf{x}) = \frac{\ell(W)}{n(\mathbf{x})+1}\lambda((u, v)|\mathbf{x})$, $(u, v) \notin \mathbf{x}$.
    If $y_2 < r((u, v); \mathbf{x})$ then $X_{n+1} = \mathbf{x} \cup (u, v)$ else $X_{n+1} = \mathbf{x}$
- If $y_1 > \frac{1}{2}$ then
  - Uniformly select a point (ξ, t) in $\mathbf{x}$ according to a discrete probability density $1/n(\mathbf{x})$,
  - Compute $r((\xi, t); \mathbf{x}\backslash(\xi, t)) = \frac{\ell(W)}{n(\mathbf{x})}\lambda((\xi, t)|\mathbf{x} \setminus (\xi, t))$, $(\xi, t) \in \mathbf{x}$.
    If $y_2 < 1/r((\xi, t); \mathbf{x}\backslash(\xi, t))$ then $X_{n+1} = \mathbf{x}\backslash(\xi, t)$ else $X_{n+1} = \mathbf{x}$.
  - Note that if $\mathbf{x} = \emptyset$ then $X_{n+1} = \mathbf{x}$.

This simulation process is repeated a large number of time in order to ensure the convergence of the algorithm to the expected distribution. This number of iterations is unknown a priori and must be determined by the user from practical knowledge and/or diagnostic tools. We choose 20,000 iteration steps in simulation study (70,000 iteration steps in the application study). To investigate the convergence of the algorithm, we use a "trace plot" which shows the evolution of the number of points at each iteration of *Algorithm 3*. Thus, we check that the number of points in the simulated point pattern is stabilized (see Møller and Waagepetersen, 2004; Illian et al., 2008 for more details).

4.2. Simulation study

We compare the performance of the pseudo-likelihood and logistic likelihood approaches on the spatio-temporal multi-scale Geyer point process. We generate 100 simulated realizations in

Table 1
Parameters of the three multi-scale Geyer point process models used in simulation study.

Model	Values of parameter				
	Regular parameters		Irregular parameters		
	λ	γ	r	q	s
Model 1	70	(1.5,1.5)	(0.05,0.1)	(0.05,0.1)	(2,2)
Model 2	100	(0.5,1.5)	(0.05,0.1)	(0.05,0.1)	(1,3)
Model 3	200	(0.8,0.8)	(0.05,0.1)	(0.05,0.1)	(1,1)

Table 2
RMSE of parameter estimates from 100 simulated realizations of the multi-scale Geyer point process model.

Method	*Model 1*			*Model 2*			*Model 3*		
	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
pseudo	62.09	0.59	0.25	103.74	0.09	0.27	**22.13**	0.45	0.29
logistic	**12.07**	**0.18**	**0.16**	**17.30**	**0.08**	**0.08**	27.48	**0.20**	**0.12**

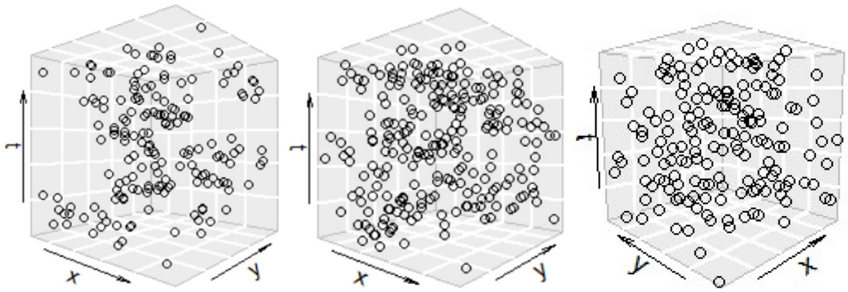

Fig. 1. Realizations of *Model 1* (left); *Model 2* (middle); *Model 3* (right).

the unit cube from three models. The first one exhibits strong clustering (*Model 1*), the second one exhibits small scale inhibition and large scale clustering (*Model 2*) and the third one exhibits inhibition (*Model 3*). Model parameters are reported in Table 1. We consider a burn-in period of 20,000 steps in *Algorithm 3*. Fig. 1 shows one realization of each model.

According to Baddeley et al. (2014), we generate a spatio-temporal Poisson process with intensity $\rho = 4n(\mathbf{x})$ (resp. $4n(\mathbf{x})/\ell(W)$) as dummy points in *Algorithm 1* (resp. *Algorithm 2*). For each model, we compute the root mean square error (RMSE) of each set of estimated parameters (Table 2) and plot the related boxplots (Fig. 2). In Table 2 the lowest RMSE value is in bold and in Fig. 2 the true values are represented by horizontal red lines. Both RMSE and boxplots show that the logistic likelihood approach performs better than the pseudo-likelihood approach for any model.

Note that in the spatial framework, Baddeley et al. (2014) showed that for large datasets the logistic likelihood method is preferable than the pseudo-likelihood method as it requires a smaller number of dummy points and performs quickly and efficiently. Daniel et al. (2018) and Choiruddin et al. (2018) investigated a similar comparison when these methods are regularized (i.e. using an approach with a simultaneous parameter estimation and variable selection by maximizing a penalized likelihood functions). Iftimi et al. (2018) found the advantage of the logistic likelihood approach for the spatio-temporal multi-scale area-interaction point process model. We here confirm this advantage for the spatio-temporal multi-scale Geyer point process model.

5. Application to forest fire occurrences

Economic and ecological disasters caused by wildfires in the world have led the scientific community to develop many novel statistical analysis and modeling wildfire occurrences to better

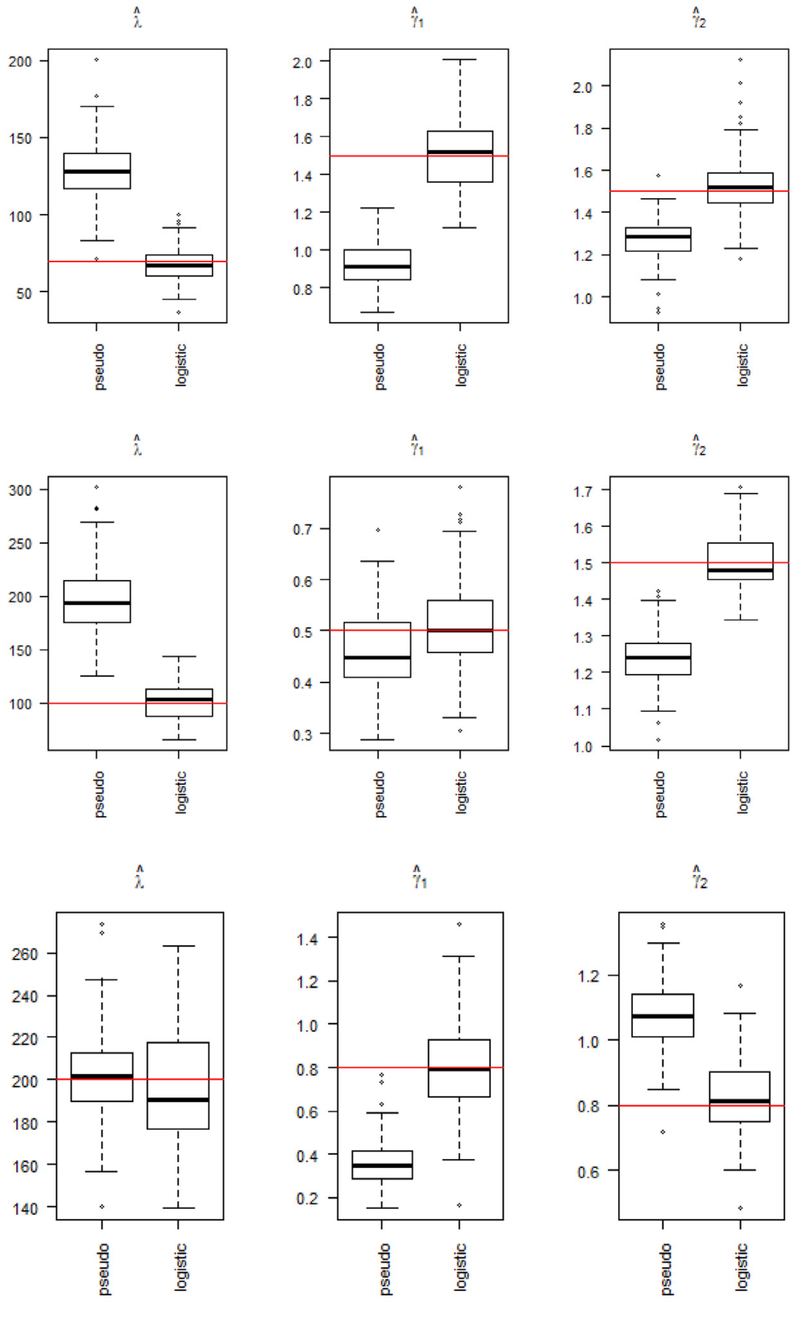

Fig. 2. Boxplots of regular parameters estimated from the pseudo-likelihood and logistic likelihood approaches for *Model 1* (first row), *Model 2* (second row) and *Model 3* (third row). True values are represented by horizontal red lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

understand their behaviors. In this section, we focus on the modeling of forest fire occurrences in the Bouches-du-Rhône county (Southern France) between 2001 and 2015.

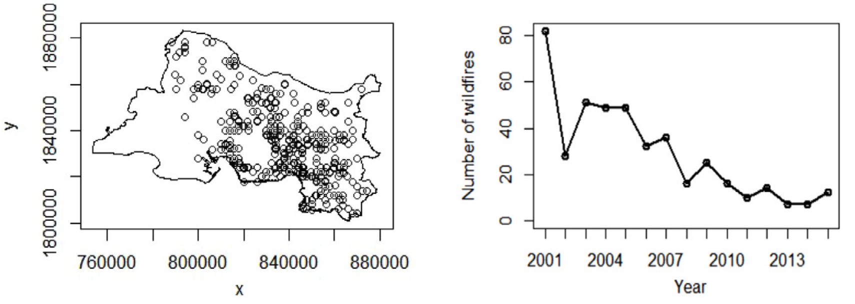

Fig. 3. (Left) Forest fire locations in UTM coordinate system (distance in meters), with more than 1 *hectare* of burnt area, recorded during the years 2001 to 2015 in the Bouches-du-Rhône county in France. (Right) Number of recorded forest fires per year.

Several statistical studies have shown the influence of environmental and meteorological factors on forest fire occurrences. In the French Mediterranean basin, Opitz et al. (2020) fit a spatio-temporal log-Gaussian Cox process model for forest fire occurrences with a log-linear intensity depending on spatio-temporal land use and weather covariates. Ganteaume and Jappiot (2013) investigated the impact of the different covariates on the number of fires using multivariate analysis and Gabriel et al. (2017) explored the influence of land cover covariates, temperature and precipitation on the probability of event occurrence. In addition to the spatio-temporal clustering of events induced by some covariates, Gabriel et al. (2017) detected spatio-temporal interaction structures at different scales and notably an inhibitive effect that arises locally in time and space after wildfires as we expect lesser occurrences at these locations during a vegetation regeneration period.

We propose to fit the spatio-temporal hybrid Geyer point process model (5) on wildfire occurrences to take into account both the inhomogeneities induced by covariates and the multi-scale structure of interactions.

5.1. Data

Our dataset is of the form (ξ_i, t_i), $i = 1, \ldots, 434$, where (ξ_i, t_i) corresponds to a wildfire with more than 1 *hectare* of burnt surface spatially indexed by a DFCI[2] cell center ξ_i in the Lambert 93 projection system and year $t_i \in \{2001, \ldots, 2015\}$. To avoid duplicated points we uniformly jittered ξ_i in its DFCI cell. We refer the reader to Gabriel et al. (2017) and Opitz et al. (2020) for further information on the data. Whilst forest fires are daily reported, we consider here the yearly scale, as done in many works (see e.g. Serra et al., 2012, 2014a,b), because of the small number of reports and to optimize computation time in model fitting and validation steps. Fig. 3 plots locations of forest fires (left panel) and yearly number of occurrences (right panel). It shows some clustering at short and medium spatial distances. Note that there exist two particular areas without any fire occurrences as they correspond to a lake (center) and marshlands (South-West). The number of fires slightly exponentially decreases in time over the 15 years, mainly due to improvements of fire-fighting resources.

We consider the same framework as in Gabriel et al. (2017) and restrict our attention to the following covariates: water coverage, elevation, coniferous cover and building cover as spatial covariates and temperature average, precipitation as spatio-temporal covariates. Hence, we can consider these covariates as good proxies of the main environmental, climatic and human factors. Maps of covariates are shown in Fig. 4 in 2001.

2 District units for fire management strategies, see Opitz et al. (2020).

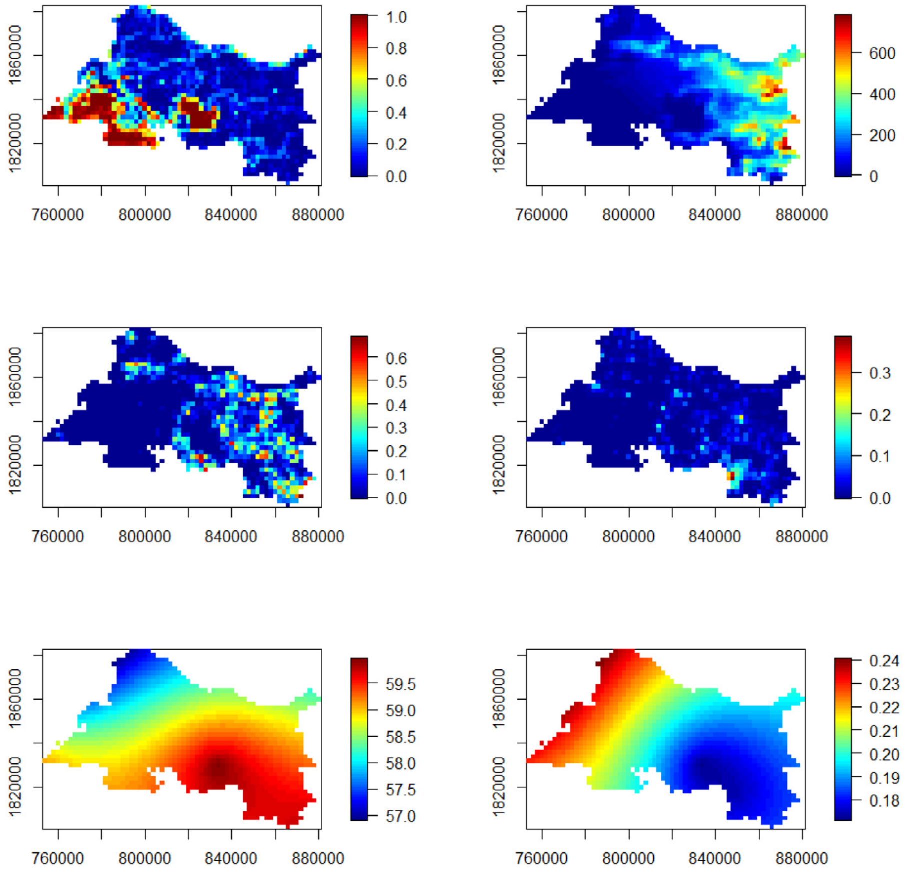

Fig. 4. Maps of covariates: water coverage (top left), elevation (top right), coniferous cover (middle left), building cover (middle right), temperature average (bottom left) and square root of precipitation (bottom right) in 2001.

5.2. Model fitting

Here we first estimate the spatio-temporal trend and then fit the spatio-temporal multi-scale Geyer model to forest fire occurrences. This two-step model fitting procedure follows our assumption that most forest fire occurrences are firstly due to environmental and meteorological conditions and secondly due to unobserved pairwise interactions. This technique will allow to see the benefits of the multi-scale interaction structure in our hybrid model compared to an inhomogeneous Poisson model with the same spatio-temporal trend.

5.2.1. Spatio-temporal trend estimation

We express the spatio-temporal trend (5) as $\lambda(\xi, t) = \beta\mu(\xi, t)$ where $\log\mu(\xi, t)$ is assumed to linearly depend on covariates:

$$\log\mu(\xi, t) = \beta_0 + \sum_{k=1}^{4} \beta_k^S Z_k^S(\xi) + \sum_{l=1}^{2} \beta_l^{ST} Z_l^{ST}(\xi, t) + \alpha t \tag{22}$$

with $Z_k^S(\xi)$, $k = 1, \ldots, 4$, the spatial covariates, $Z_l^{ST}(\xi, t)$, $l = 1, 2$, the spatio-temporal covariates and αt a decreasing trend of fire counts over time. Because the covariates are known at a fixed

Table 3
Estimated coefficients, standard errors and *p*-values based on two-tailed Student's t-tests of significant differences from zero.

Covariates	Coefficients	Estimates	Standard error	*p*-value
Intercept	β_0	262	26	$< 2 \times 10^{-16}$ ***
Water	β_1^S	−1.88	0.29	5.89×10^{-11} ***
Elevation	β_2^S	−0.001	0.0004	0.0008 ***
Coniferous	β_3^S	0.77	0.36	0.031 *
Building	β_4^S	4	0.89	8.08×10^{-6} ***
Temperature	β_1^{ST}	0.37	0.06	1.13×10^{-10} ***
Precipitation	β_2^{ST}	−11.3	1.48	1.75×10^{-14} ***
Time	α	−0.14	0.001	$< 2 \times 10^{-16}$ ***

discretization scale, $\mu(\xi, t)$ does not vary for points ξ inside the same DFCI unit with a time t corresponding to the same year. By consequence, we can restrict our attention on DFCI grid cell centers $\xi_i, i = 1, \ldots, 1320$ and years $t_j = 2001, \ldots, 2015$ for $j = 1, \ldots, 15$, and we consider a Poisson response for our model $N_{ij}|\mu(\xi_i, t_j) \sim Poisson(\mu(\xi_i, t_j))$, where N_{ij} is the number of forest fires in ith DFCI cell at year t_j. The coefficient β will be estimated simultaneously with the other regular parameters by the logistic likelihood approach. Table 3 reports the coefficients β_0, β_k^S, β_l^{ST} and α estimated as in Gabriel et al. (2017) and Opitz et al. (2020). The sign indicates if covariates favor (if positive, like coniferous, building and temperature) or prevent (if negative, like water, elevation, precipitation and time) fire occurrences. All covariates are globally significant and results are consistent with previous works (Ganteaume and Jappiot, 2013; Gabriel et al., 2017; Opitz et al., 2020) for this county. Note that p-values have been computed during the trend fitting under a Poisson model and not for the overall fitting of forest fire occurrences under our spatio-temporal hybrid Geyer saturation process. Thus, we might have obtained more significance of the covariates than under our hybrid Geyer saturation model.

5.2.2. Parameters estimation

There is no common method for estimating irregular parameters in spatial or spatio-temporal Gibbs point process models. Here we considered several combinations of ad-hoc values within a reasonable range and select the optimal irregular parameters according to the Akaike's Information Criterion (AIC) of the fitted model.

Baddeley and Turner (2006) suggest that the spatial interaction radius r of the Geyer saturation point process should be between 0 and the maximum nearest neighbor distance, about 8000 *meters* for our dataset. For the temporal radius q, we consider small values to be in accordance with the natural phenomena of forest fire occurrences. Finally, for the saturation parameter s, we have $n(C_r^q(\xi_i, t_i); \mathbf{x}) \leq s$ for all $(\xi_i, t_i) \in \mathbf{x}$. Hence, for any pair (r, q), we set $s = \max_{1 \leq i \leq n} n(C_r^q(\xi_i, t_i); \mathbf{x})$.

According to the former section, we use the logistic likelihood method and *Algorithm 2* to estimate the regular parameters. We simulate dummy points from an inhomogeneous Poisson point process with intensity $\rho(\xi, t) = C\mu(\xi, t)/\nu$ where $C = 4$ by a classical rule of thumb in the logistic likelihood approach and $\nu = 2000 \times 2000 \times 1$ (area of a DFCI cell multiplied by 1 year).

We fitted the spatio-temporal multi-scale Geyer point process model for a range of ad-hoc values $(r_j, q_j) \in [0, 8000] \times \{1, 2, 3, 4, 5\}$, and their corresponding values of s_j, $j = 1, \ldots, m$, with varying m in $\{1, 2, 3, 4, 5\}$. The minimum AIC is obtained for the combination given in Table 4. Estimated regular parameters γ_j associated with their 95% bootstrap confidence intervals show strong clustering at very short distances, weak repulsion (resp. clustering) at small (resp. medium) scale, and randomness at large scale. Another methodology for testing the significance of γ_j parameters from 1 could be to extend the pseudo-likelihood or composite likelihood ratio test introduced in Baddeley et al. (2016) to the spatio-temporal case.

Table 4
Parameter estimates for $m = 4$.

Irregular parameters				
r	500	2000	5000	7500
q	1	2	3	4
s	4	7	27	57
Estimated regular parameters and 95% confidence intervals				
$\hat{\beta} = 0.66$	$\hat{\gamma}_1 = 2.73$	$\hat{\gamma}_2 = 0.93$	$\hat{\gamma}_3 = 1.07$	$\hat{\gamma}_4 = 0.98$
[0.442, 0.968]	[1.818, 3.405]	[0.820, 0.994]	[1.020, 1.120]	[0.962, 1.011]

5.3. Model validation

We validate our fitted model from several Monte Carlo tests using statistics based on the spatio-temporal inhomogeneous K-function (Gabriel and Diggle, 2009). First, we generate $n_{sim} = 99$ simulations from our fitted hybrid Geyer model (5) by *Algorithm 3* with a burn-in period of 70,000 steps, representing realizations from our null hypothesis. Then, we compute the spatio-temporal inhomogeneous K-function for the observed and simulated point patterns, denoted respectively by $\hat{K}^{inh}_{obs}(h_s, h_t)$ and $\hat{K}^{inh}_i(h_s, h_t), i \in \{1, \ldots, n_{sim}\}$, with an estimated separable intensity function obtained by kernel smoothing. For each value of the spatio-temporal distance (h_s, h_t), lower (L) and upper (U) critical envelopes of the summary statistics are computed locally

$$L(h_s, h_t) = \min_{1 \le i \le n_{sim}} \hat{K}^{inh}_i(h_s, h_t), \quad U(h_s, h_t) = \max_{1 \le i \le n_{sim}} \hat{K}^{inh}_i(h_s, h_t). \tag{23}$$

In addition to these local envelopes, we compute local and global p-values as in Tamayo-Uria et al. (2014), Siino et al. (2018a) in order to respectively detect spatio-temporal distances where the departure from the null hypothesis is the most significant and the overall adequacy of our model. Let $E(h_s, h_t)$ and $V(h_s, h_t)$ denote the mean and variance of $\left\{\hat{K}^{inh}_1(h_s, h_t), \ldots, \hat{K}^{inh}_{n_{sim}}(h_s, h_t), \hat{K}^{inh}_{obs}(h_s, h_t)\right\}$. We define the local p-value for each pair (h_s, h_t) by

$$p(h_s, h_t) = \frac{1 + \sum_{i=1}^{n_{sim}} \mathbb{1}\{T_i(h_s, h_t) > T_{obs}(h_s, h_t)\}}{n_{sim} + 1}, \tag{24}$$

where $T_i(h_s, h_t)$ (resp. $T_{obs}(h_s, h_t)$) denotes the local statistic T computed from the ith simulation (resp. the data) at (h_s, h_t). The local statistic is defined by

$$T(h_s, h_t) = \sqrt{\frac{(\hat{K}^{inh}(h_s, h_t) - E(h_s, h_t))^2}{V(h_s, h_t)}}. \tag{25}$$

The global test combines the information for all spatial and temporal distances. We define the test statistic

$$\tilde{T} = \int_0^{h_{t,max}} \int_0^{h_{s,max}} T(h_s, h_t) dh_s dh_t, \tag{26}$$

where $h_{s,max}$ and $h_{t,max}$ are user-specific maximum spatial and temporal distances which are preferable to choose close to the (expected) range of interaction of the underlying point process. Illian et al. (2008) recommend to compare the results for several values of $h_{s,max}$ and $h_{t,max}$. The p-value of the global test is then given by

$$p_{global} = \frac{1 + \sum_{i=1}^{n_{sim}} \mathbb{1}\{\tilde{T}_i > \tilde{T}_{obs}\}}{n_{sim} + 1}.$$

Fig. 5.(a) shows the spatio-temporal inhomogeneous K function computed on our dataset (dark gray) and the envelopes obtained from our hybrid Geyer model (light gray); $\hat{K}^{inh}_{obs}(h_s, h_t)$ lies inside the envelopes, meaning that the fitted model seems to describe properly the spatio-temporal structure of the data. This is confirmed by local p-values at any distances (Fig. 5.(b)). Global p-values

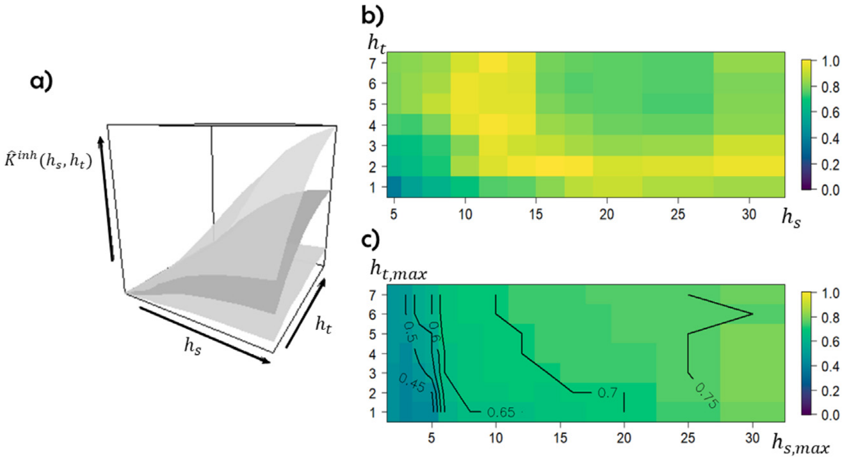

Fig. 5. Temporal separations h_t are in *year* and spatial distances h_s are in *kilometer*. (a) Envelopes of the spatio-temporal inhomogeneous K-function for the simulated spatio-temporal multi-scale Geyer point process according to the estimated parameters. (b) Image plot of the local p-value. (c) Image plot of the global p-value for any pairs of $(h_{s,max}, h_{t,max})$.

are given in Fig. 5.(c) for any combination of $h_{s,max}$ and $h_{t,max}$. Again, it shows that our fitted model is validated.

In addition, we also compute global envelopes and p-value of the spatio-temporal $\hat{K}^{inh}$ functions based on the Extreme Rank Length (ERL) measure defined in Myllymäki et al. (2017) and implemented in the R package GET (Myllymäki and Mrkvička, 2019). The main advantage is that the resulting p-value will not depend on a priori parameters as in the definition of p_{global} with the $h_{s,max}$ and $h_{t,max}$ values. For each point pattern, we consider the long vector T_i, $i = 1, \ldots, n_{sim}$ (resp. T_{obs}) merging the $K_i^{inh}(\cdot, h_t)$ (resp. $K_{obs}^{inh}(\cdot, h_t)$) estimates for all considered values h_t. The ERL measure of vector T_i (resp. T_{obs}) of length n_{st} is defined as

$$E_i = \frac{1}{n_{ns}} \sum_{j=1}^{n_{st}} \mathbb{1}\{R_j \prec R_i\},$$

where R_i is the vector of pointwise ordered ranks and $\prec$ is an ordering operator (Myllymäki et al., 2017; Myllymäki and Mrkvička, 2019). The final p-value is obtained by

$$p_{erl} = \frac{1 + \sum_{i=1}^{n_{sim}} \mathbb{1}\{E_i \geq E_{obs}\}}{n_{sim} + 1}.$$

The global p-value p_{erl} is equal to 0.34 consolidating previous results and validating our hybrid Geyer model.

Note that we did the same tests for 99 simulations of an inhomogeneous Poisson process with intensity $\mu(\xi, t)/(2000 \times 2000 \times 1)$ (22). This model has been rejected at the level 5%, with a median global p-value equals to 0.04. The p_{erl} value is equal to 0.04 under the Poisson assumption rejecting also this baseline model.

Conclusion

Due to the capability of Gibbs point processes to cover prevalent structures (inhibition, randomness and clustering), the hybridization approach allows to introduce new Gibbs models combining several structures at different scales. In this paper, we defined the spatio-temporal multi-scale Geyer saturation point process model and detailed the classical statistical inference methods and MCMC simulation techniques that we have extended to the spatio-temporal framework and implemented in R code[3] that will be added to the stpp package (Gabriel et al., 2013). Our simulation study highlighted a better goodness-of-fit of parameters for the logistic likelihood approach compared to the pseudo-likelihood approach. Finally, we illustrated the interest of using this model on a spatio-temporal dataset of forest fire locations associated with environment covariates. The model validation shows that our model captures the multi-scale interaction structure inherent to forest fire occurrences.

In this paper, we focused our attention on the definition of a new hybrid Gibbs model, the inference methods and MCMC simulation algorithms that we needed to adapt to the spatio-temporal context. Some of our choices can be discussed and eventually improved in future works, notably in our application to forest fire occurrences which is not presented as an in-depth study but as an illustration of the model fitting on real data.

In our application study, we considered a log-linear form for the trend depending on covariate information. We chose a two-step procedure for estimating, at first, the trend coefficients and then the regular parameters of the interaction function. Our knowledge on forest fire mechanisms guided this choice because the main driver of occurrence locations is the environmental heterogeneity and the secondary one is the interaction phenomena. The trend is estimated at the spatial DFCI scale and at the yearly one, corresponding to our covariate resolution. In that way, we estimated a global trend at a medium scale whereas the interaction parameters are estimated at the point locations and represent a local interaction behavior at a fine scale. This procedure could be improved by incorporating variable selection methods, e.g. via regularization (Choiruddin et al., 2018; Daniel et al., 2018).

Our two-step estimation procedure allows us to provide confidence intervals for both the trend coefficients and the regular parameters. We notice that some parameters γ_j are closed to one. Here we consider a bootstrap estimate of the confidence interval for each γ_j. We could further test departure from one by extending the adjusted composite likelihood ratio test (Baddeley et al., 2016) to the spatio-temporal framework. Indeed, Baddeley et al. (2016) proposed a likelihood ratio test for spatial Gibbs point process models fitted by maximum pseudo-likelihood. They discussed that implementing other composite likelihood as the logistic likelihood would provide a better composite likelihood ratio test. Estimating diagnostics related to the logistic likelihood requires to estimate the variance–covariance matrix of the logistic score and the sensitivity matrix. Baddeley et al. (2014) provide consistent estimators of these quantities. The extension to the spatio-temporal framework is a full-blown work that also involves efficient implementation.

For the choice of irregular parameters, because the likelihood is not differentiable with respect to them, we used a maximum profile likelihood approach based on the logistic likelihood estimation procedure and AIC values for model selection. Introduced for the pseudo-likelihood estimates in Anwar and Stein (2015) and applied to the logistic likelihood approach by us using the results in Baddeley et al. (2014), this method consists in fixing irregular parameters and maximizing the composite likelihood with respect to the regular ones. This technique is a computationally-intensive method. Thanks to a preliminary spatio-temporal exploratory analysis of the interaction ranges done with the inhomogeneous pair correlation function g, the maximum nearest neighbor distance and the temporal autocorrelation function, we chose few configurations of feasible values for the nuisance parameters m, r_j, q_j and s_j, $j = 1, \ldots, m$. Considering more values would be very time-consuming and developing a new estimation method would be a subject in its own right. During the model validation procedure, we could use the global envelope tests based on the ERL measure to assess the goodness-of-fit of submodels with fewer irregular parameters to be parsimonious.

Our model can be used in many fields, like seismology and epidemiology for example, because several mechanisms exhibit interaction between points at multiple scales in space and time. Relying on this work, we can also develop hybrid models with different density structures. Indeed, although it was not necessarily highlighted here, we know that forest fires with large burnt areas avoid future fire occurrences during a vegetation regeneration period. Such cases of strong inhibition may be modeled by hybrid Gibbs point processes with a hardcore component like the hybrid Geyer hardcore point process. We recently extended our work to this model.

3 http://edith.gabriel.pagesperso-orange.fr/software.html.

References

Anwar, S., Stein, A., 2015. Spatial pattern development of selective logging over several years. Spat. Statist. 13, 90–105. http://dx.doi.org/10.1016/j.spasta.2015.03.001.

Arago, P., Juan, P., Diaz-Avalos, C., Salvador, P., 2016. Spatial point process modeling applied to the assessment of risk factors associated with forest wildfires incidence in Castellon, Spain. Eur. J. Forest Res. 135 (3), 451–464. http://dx.doi.org/10.1007/s10342-016-0945-z.

Baddeley, A., Coeurjolly, J., Rubak, E., Waagepetersen, R., 2014. Logistic regression for spatial Gibbs point processes. Biometrika 101 (2), 377–392. http://dx.doi.org/10.1093/biomet/ast060.

Baddeley, A., Rubak, E., Turner, R., 2019. Leverage and influence diagnostics for Gibbs spatial point processes. Spat. Statist. 29, 15–48. http://dx.doi.org/10.1016/j.spasta.2018.09.004.

Baddeley, A., Turner, R., 2000. Practical maximum pseudolikelihood for spatial point patterns (with discussion). Aust. N. Z. J. Stat. 42, 283–322. http://dx.doi.org/10.1111/1467-842X.00128.

Baddeley, A., Turner, R., 2006. Modelling spatial point patterns in R. In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (Eds.), Case Studies in Spatial Point Process Modeling. In: Lecture Notes in Statistics, vol. 185, Springer, New York, pp. 23–74. http://dx.doi.org/10.1007/0-387-31144-0_2.

Baddeley, A., Turner, R., Mateu, J., Bevan, A., 2013. Hybrids of Gibbs point process models and their implementation. J. Stat. Softw. 55 (11), 1–43. http://dx.doi.org/10.18637/jss.v055.i11.

Baddeley, A., Turner, R., Rubak, E., 2016. Adjusted composite likelihood ratio test for spatial Gibbs point processes. J. Stat. Comput. Simul. 86 (5), 922–941. http://dx.doi.org/10.1080/00949655.2015.1044530.

Berman, M., Turner, R., 1992. Approximating point process likelihoods with GLIM. Appl. Stat. 41 (1), 31–38. http://dx.doi.org/10.1016/0167-6687(93)90845-G.

Besag, J., 1977. Some methods of statistical analysis for spatial data. Bull. Int. Stat. Inst. 47, 77–92.

Brix, A., Chadœuf, J., 2000. Spatio-temporal modeling of weeds and shot-noise G Cox processes. Biom. J. 44, 83–99. http://dx.doi.org/10.1002/1521-4036(200201)44:1<83::AID-BIMJ83>3.0.CO;2-W.

Brix, A., Diggle, P., 2001. Spatiotemporal prediction for log-Gaussian Cox processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (4), 823–841. http://dx.doi.org/10.1111/1467-9868.00315.

Brix, A., Kendal, W., 2002. Simulation of cluster point processes without edge effects. Adv. Appl. Probab. 34, 267–280. http://dx.doi.org/10.1239/aap/1025131217.

Brix, A., Møller, J., 2001. Space-time multitype log Gaussian Cox processes with a view to modelling weed data. J. Scand. J. Stat. 28, 471–488. http://dx.doi.org/10.1111/1467-9469.00249.

Choiruddin, A., Coeurjolly, J., Letué, F., 2018. Convex and non-convex regularization methods for spatial point processes intensity estimation. Electron. J. Stat. 12 (1), 1210–1255. http://dx.doi.org/10.1214/18-EJS1408.

Cox, D., 1972. The statistical analysis of dependencies in point processes. In: Lewis, P. (Ed.), Stochastic Point Processes. Wiley, New York, pp. 55–66.

Cronie, O., van Lieshout, M., 2015. A J-function for inhomogeneous spatio-temporal point processes. Scand. J. Stat. 42 (2), 562–579. http://dx.doi.org/10.1111/sjos.12123.

Daniel, J., Horrocks, J., Umphrey, G., 2018. Penalized composite likelihoods for inhomogeneous Gibbs point process models. Comput. Statist. Data Anal. 124, 104–116. http://dx.doi.org/10.1016/j.csda.2018.02.005.

Dereudre, D., 2019. Introduction to the theory of Gibbs point processes. In: Coupier, D. (Ed.), Stochastic Geometry. Springer, pp. 181–229. http://dx.doi.org/10.1007/978-3-030-13547-8_5.

Diggle, P., Moraga, P., Rowlingson, B., Taylor, B., 2013. Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. Statist. Sci. 28 (4), 542–563. http://dx.doi.org/10.1214/13-STS441.

Gabriel, E., 2014. Estimating second-order characteristics of inhomogeneous spatio-temporal point processes. Methodol. Comput. Appl. Probab. 16 (2), 411–431. http://dx.doi.org/10.1007/s11009-013-9358-3.

Gabriel, E., Diggle, P., 2009. Second-order analysis of inhomogeneous spatio-temporal point process data. Stat. Neerl. 63 (1), 43–51. http://dx.doi.org/10.1111/j.1467-9574.2008.00407.x.

Gabriel, E., Opitz, T., Bonneu, F., 2017. Detecting and modeling multi-scale space-time structures: the case of wildfire occurrences. J. French Stat. Soc. 158 (3), 86–105, URL: http://journal-sfds.fr/article/view/629.

Gabriel, E., Rowlingson, B., Diggle, P., 2013. `stpp`: a R package for plotting, simulating and analyzing spatio-temporal Point Patterns. J. Stat. Softw. 53 (2), 1–29. http://dx.doi.org/10.18637/jss.v053.i02.

Ganteaume, A., Jappiot, M., 2013. What causes large fires in Southern France. Forest Ecol. Manag. 294, 76–85. http://dx.doi.org/10.1016/j.foreco.2012.06.055.

Geyer, C., 1999. Likelihood inference for spatial point processes. In: Barndorff-Nielsen, O., Kendall, W., van Lieshout, M. (Eds.), Stochastic Geometry: Likelihood and Computation, Vol. 80. Chapman and Hall, London, pp. 79–140. http://dx.doi.org/10.1201/9780203738276-3.

Geyer, C., Møller, J., 1994. Simulation procedures and likelihood inference for spatial point processes. Scand. J. Stat. 18, 505–544, URL: http://www.jstor.org/stable/4616323.

Gonzalez, J., Rodriguez-Cortes, F., Cronie, O., Mateu, J., 2016. Spatio-temporal point process statistics: A review. Spat. Statist. 18, 505–544. http://dx.doi.org/10.1016/j.spasta.2016.10.002.

Iftimi, A., van Lieshout, M., Montes, F., 2018. A multi-scale area-interaction model for spatio-temporal point patterns. Spat. Statist. 26, 38–55. http://dx.doi.org/10.1016/j.spasta.2018.06.001.

Iftimi, A., Montes, F., Mateu, J., Ayyad, C., 2017. Measuring spatial inhomogeneity at different spatial scales using hybrids of Gibbs point process models. Stoch. Environ. Res. Risk Assess. 31 (6), 1455–1469. http://dx.doi.org/10.1007/s00477-016-1264-0.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. John Wiley & Sons, Chichester, http://dx.doi.org/10.1002/9780470725160.

Juan, P., Mateu, J., Saez, M., 2012. Pinpointing spatio-temporal interactions in wildfire patterns. Stoch. Environ. Res. Risk Assess. 26 (8), 1131–1150. http://dx.doi.org/10.1007/s00477-012-0568-y.

Kingman, J., 1993. Poisson Processes, Vol. 3. Oxford University Press, Oxford.

Kingman, J., 2006. Poisson processes revisited. Probab. Math. Statist. 26 (1), 77–95.

Lavancier, F., Møller, J., Rubak, E., 2015. Determinantal point process models and statistical inference. J. R. Stat. Soc. Ser. B Stat. Methodol. 77, 853–877. http://dx.doi.org/10.1111/rssb.12096.

Levin, C., 1992. The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. Ecology 73, 1943–1967. http://dx.doi.org/10.2307/1941447.

Macchi, O., 1975. The coincidence approach to stochastic point processes. Adv. Appl. Probab. 7, 83–122. http://dx.doi.org/10.2307/1425855.

Matérn, B., 1960. Spatial Variation. In: Lectures Notes in Statistics, Springer-Verlag, Chichester, http://dx.doi.org/10.1007/978-1-4615-7892-5.

Møller, J., Diaz-Avalos, C., 2010. Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. Scand. J. Stat. 37 (1), 2–25. http://dx.doi.org/10.1111/j.1467-9469.2009.00670.x.

Møller, J., Syversveen, A., Waagepetersen, R., 1998. Log Gaussian Cox processes. Scand. J. Stat. 25 (3), 451–482. http://dx.doi.org/10.1111/1467-9469.00115.

Møller, J., Waagepetersen, R., 2004. Statistical Inference and Simulation for Spatial Point Processes. Chapman and Hall/CRC, Boca Raton, http://dx.doi.org/10.1201/9780203496930.

Myllymäki, M., Mrkvička, T., 2019. GET: Global envelopes in R. URL: https://arxiv.org/abs/1911.06583.

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H., Hahn, U., 2017. Global envelope tests for spatial processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 79, 381–404. http://dx.doi.org/10.1111/rssb.12172.

Neyman, J., Scott, E., 1958. Statistical approach to problems of cosmology. J. R. Stat. Soc. Ser. B Stat. Methodol. 20, 1–29. http://dx.doi.org/10.1111/j.2517-6161.1958.tb00272.x.

Ogata, Y., Tanemura, M., 1981. Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. Ann. Inst. Statist. Math. 33, 315–338. http://dx.doi.org/10.1007/BF02480944.

Opitz, T., 2009. Simulating and Fitting of Gibbs Processes in 3D–Models, Algorithms and Their Implementation (Master's thesis). Ulm University, Germany, URL: https://scholar.google.com/scholar?cluster=18315270875961905679&hl=en&oi=scholarr.

Opitz, T., Bonneu, F., Gabriel, E., 2020. Point-process based Bayesian modeling of space-time structures of forest fire occurrences in Mediterranean France. Spat. Statist. 100429. http://dx.doi.org/10.1016/j.spasta.2020.100429.

Papangelou, F., 1974. The conditional intensity of general point processes and an application to line processes. Probab. Theory Related Fields 28 (3), 207–226. http://dx.doi.org/10.1007/BF00533242.

Picard, N., Bar-Hen, A., Mortier, F., Chadoeuf, J., 2009. The multi-scale marked area-interaction point process: a model for the spatial pattern of trees. Scand. J. Stat. 36, 23–41. http://dx.doi.org/10.1111/j.1467-9469.2008.00612.x.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL: https://www.R-project.org.

Raeisi, M., Bonneu, F., Gabriel, E., 2019. On spatial and spatio-temporal multi-structure point process models. Les Ann. l'ISUP 63 (2–3), URL: https://arxiv.org/abs/2003.01962.

Ripley, B., Kelly, F., 1977. Markov point processes. J. Lond. Math. Soc. 15 (1), 188–192. http://dx.doi.org/10.1112/jlms/s2-15.1.188.

Serra, L., Juan, P., Varga, D., Mateu, J., Saez, M., 2013. Spatial pattern modelling of wildfires in Catalonia, Spain 2004–2008. Environ. Modell. Softw. 40, 235–244. http://dx.doi.org/10.1016/j.envsoft.2012.09.014.

Serra, L., Saez, M., Juan, P., Varga, D., Mateu, J., 2014a. A spatio-temporal Poisson hurdle point process to model forest fires. Stoch. Environ. Res. Risk Assess. 28 (7), 1671–1684. http://dx.doi.org/10.1007/s00477-013-0823-x.

Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Diaz-Avalos, C., 2014b. Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of Catalonia, 1994–2008. Environ. Ecol. Stat. 21 (3), 531–563. http://dx.doi.org/10.1007/s10651-013-0267-y.

Serra, L., Saez, M., Varga, D., Tobias, A., Mateu, J., 2012. Spatio-temporal modelling of wildfires in Catalonia, Spain, 1994–2008, through log Gaussian Cox processes. WIT Trans. Ecol. Environ. 158, 34–49. http://dx.doi.org/10.2495/FIVA120041.

Siino, M., Adelfio, G., Mateu, J., 2018a. Joint second-order parameter estimation for spatio-temporal log-Gaussian Cox processes. Stoch. Environ. Res. Risk Assess. 32, 3525–3539. http://dx.doi.org/10.1007/s00477-018-1579-0.

Siino, M., Adelfio, G., Mateu, J., Chiodi, M., D'Alessandro, A., 2017. Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity. Stoch. Environ. Res. Risk Assess. 31 (7), 1633–1648. http://dx.doi.org/10.1007/s00477-016-1294-7.

Siino, M., D'Alessandro, A., Adelfio, G., Scudero, S., Chiodi, M., 2018b. Multiscale processes to describe the eastern sicily seismic sequences. Ann. Geophys. 61 (2), 1–17. http://dx.doi.org/10.4401/ag-7688.

Strauss, D., 1975. A model for clustering. Biometrika 62 (2), 467–475. http://dx.doi.org/10.1093/biomet/62.2.467.

Tamayo-Uria, I., Mateu, J., Diggle, P., 2014. Modelling of the spatiotemporal distribution of rat sightings in an urban environment. Spat. Statist. 9, 192–206. http://dx.doi.org/10.1016/j.spasta.2014.03.005.

Trilles, S., Juan, P., Diaz, L., Arago, P., Huerta, J., 2013. Integration of environmental models in spatial data infrastructures: a use case in wildfire risk prediction. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6 (1), 128–138. http://dx.doi.org/10.1109/JSTARS.2012.2236538.

Turner, R., 2009. Point patterns of forest fire locations. Environ. Ecol. Stat. 16, 197–223. http://dx.doi.org/10.1007/s10651-007-0085-1.

Wiegand, T., Gunatillekem, N., Okudam, T., 2007. Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. Ecology 88, 3088–3102. http://dx.doi.org/10.1890/06-1350.1.

Woo, H., Chung, W., Graham, J., Lee, B., 2017. Forest fire risk assessment using point process modelling of fire occurrence and Monte Carlo fire simulation. Int. J. Wildland Fire 26, 789–805. http://dx.doi.org/10.1071/WF17021.