



HAL
open science

Point process models for complex spatio-temporal data

Morteza Raeisi

► **To cite this version:**

Morteza Raeisi. Point process models for complex spatio-temporal data. Statistics [math.ST]. Université d'Avignon, 2021. English. NNT : 2021AVIG0423 . tel-03675194

HAL Id: tel-03675194

<https://theses.hal.science/tel-03675194>

Submitted on 23 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT D'AVIGNON UNIVERSITÉ

École Doctorale N° 536
Agrosciences et Sciences

Mention de doctorat :
Mathématiques Appliquées

Laboratoire de Mathématiques d'Avignon

Présentée par
Morteza Raeisi

**MODÈLES DE PROCESSUS
PONCTUELS POUR DES DONNÉES
SPATIO-TEMPORELLES COMPLEXES**

Soutenue publiquement le 29/09/2021 devant le jury composé de :

Mme Giada Adelfio, Associate Professor, University of Palermo **Rapporteure**
M. Jean-François Coeurjolly, Professeur, Université Grenoble Alpes **Rapporteur**
M. Frédéric Lavancier, MCF/HDR, Université Nantes **Examineur**
M. Thomas Opitz, CRCN/HDR, INRAE **Examineur**
Mme Edith Gabriel, DR2, INRAE **Directrice de thèse**
M. Florent Bonneu, MCF, Avignon Université **Co-directeur de thèse**

Abstract

The theory of point processes is a branch of spatial statistics. A spatial (and spatio-temporal) point pattern, as a realization of a point process, is a collection of events for which locations (and times) of occurrence have been observed in a specified spatial region (and temporal period). Point patterns are often classified into three classes of single interaction structure: randomness, clustering, and inhibition that can be modeled for instance by Poisson process, Cox processes, and Gibbs processes, respectively. These single-structure point process models can be too simplistic to describe some complex phenomena in seismology, epidemiology, and forestry as they involve several structures at different spatial (and spatio-temporal) scales, thus requiring multi-structure point processes to describe them. The main concern of this Ph.D. thesis is the spatio-temporal modeling of such complex point patterns taking into account the spatio-temporal inhomogeneity driven by covariates and the complexity of the interaction structures.

In the spatial point processes literature, three general approaches are considered for constructing multi-structure point process models: thinning, superposition, and hybridization. The key contribution of the Ph.D. thesis is to introduce spatio-temporal hybrid point processes based on Gibbs and Cox models using hybridization approach and to develop their global and local statistical inference through composite likelihoods and Bayesian hierarchical approach and their simulation process through the birth–death Metropolis–Hastings algorithm.

Finally, we apply these new hybrid point processes to model the complex interaction structure observed on different datasets: forest fire occurrences in France and Spain and the hotspots of temperature in the United States.

Keywords: *Spatio-temporal Gibbs point processes, Hybridization, Composite likelihoods, Bayesian hierarchical approach, Forest fires, Hotspots of temperature.*

Résumé

La théorie des processus ponctuels est une branche de la statistique spatiale. Un modèle ponctuel spatial (spatio-temporel), en tant que réalisation d'un processus ponctuel, est un ensemble d'événements pour lesquels des positions (et des dates) d'occurrence ont été observées dans une région (et durant une période temporelle). Les semis de points sont souvent classés en trois classes de structure unique d'interaction : aléatoire, agrégée et répulsive qui peuvent être modélisées par exemple par les processus de Poisson, les processus de Cox et les processus de Gibbs, respectivement. Ces modèles de processus ponctuels à structure unique peuvent être trop simplistes pour décrire certains phénomènes complexes en sismologie, épidémiologie et foresterie car ils impliquent plusieurs structures à différentes échelles spatiales (et spatio-temporelles), nécessitant ainsi des processus ponctuels multi-structures pour les décrire. L'enjeu de cette thèse est la modélisation spatio-temporelle de semis de points en tenant compte de l'inhomogénéité spatio-temporelle induite par des covariables et la complexité des structures d'interaction.

Dans la littérature sur les processus ponctuels spatiaux, trois approches générales sont envisagées pour construire des modèles de processus ponctuels multi-structures : l'éclaircissement, la superposition et l'hybridation. La principale contribution de cette thèse est d'introduire de nouveaux processus ponctuels hybrides spatio-temporels basés sur des modèles de Gibbs et Cox en utilisant l'approche d'hybridation et de développer leur inférence statistique globale et locale à travers les vraisemblances composites et l'approche hiérarchique bayésienne et leur simulation à travers l'algorithme de naissance-mort de Metropolis-Hastings.

Enfin, nous appliquons ces nouveaux processus ponctuels hybrides pour modéliser la structure d'interaction complexe observée sur différents jeux de données : les occurrences de feux de forêt en France et en Espagne et les points de température extrême aux États-Unis.

Mots clés : *Processus ponctuels de Gibbs spatio-temporels, Hybridation, Vraisemblance composite, Approche hiérarchique bayésienne, Feux de forêt, Points de température extrême.*

Contents

Abstract	i
Résumé	ii
1 Introduction	1
1.1 Preliminaries	1
1.2 Point processes methods	3
1.2.1 Classical point process models	5
1.3 Multi-structure model review	8
1.3.1 Models based on Gibbs processes	8
1.3.2 Models based on Cox processes	11
1.4 Forest fires review	13
1.4.1 Exploratory analysis of forest fires	14
1.4.2 Point process models for forest fires	15
1.5 Outline of thesis	16
2 New models	19
2.1 Gibbs models	19
2.1.1 Spatio-temporal Gibbs models review	20
2.1.2 Spatio-temporal Geyer saturation model	22
2.1.3 Spatio-temporal Strauss hardcore model	23
2.2 Cox-Gibbs models	25
3 Inference	28
3.1 Global parameter estimation	28
3.1.1 Composite likelihoods	29
3.1.2 Bayesian approach	35
3.2 Local parameter estimation	38

3.2.1	Local pseudo-likelihood	38
3.2.2	Local logistic likelihood	40
3.2.3	Bandwidth selection	42
4	Simulations	45
4.1	Gibbs models' simulation	45
4.1.1	Birth-death Metropolis-Hastings algorithm	45
4.2	Cox-Gibbs models' simulation	46
4.3	Simulation study	47
4.3.1	Simulation study for Gibbs models	47
4.3.2	Simulation study for Cox-Gibbs models	52
4.3.3	First evaluation of the local logistic likelihood approach	53
5	Applications	55
5.1	Forest fires	55
5.1.1	Hybrid Geyer model for forest fires in France	56
5.1.2	Hybrid Strauss hardcore model for forest fires in Spain	62
5.2	Temperature anomalies	67
5.2.1	Hotspot extraction	68
5.2.2	Model fitting and validation	69
6	Conclusion and future work	71
6.1	Conclusion	71
6.2	Future work	72
	References	74
Appendices		
Appendix A	Review article published in <i>Annales of ISUP</i>	94
Appendix B	Original research published in <i>Spatial Statistics</i>	113
Appendix C	Original research submitted in <i>Stochastic Environmental Research and Risk Assessment</i>	133

Chapter 1

Introduction

Human activity is the source of environmental and climatic changes whose manifestations tend to become more and more frequent and extreme. Assessing the impact of weather and environmental change necessitates a better understanding of the random mechanisms governing the intensity of occurrences and the severity of risk events. The scientific reflection on risk prevention needs to consider a more realistic modeling of natural phenomena and their consequences, even at small spatial and temporal scales, in order to improve decision tools that remain simple and easy to use.

Single-structure point process models are too simplistic to describe such complex phenomena, e.g. in seismology (Siino et al. 2017, 2018b), epidemiology (Iftimi et al. 2017, 2018), and in forestry (Gabriel et al. 2017), as they involve several structures at different spatial (or spatio-temporal) scales. The main concern of this thesis is the spatio-temporal modeling of multi-structure point patterns taking into account the spatio-temporal inhomogeneity driven by covariates and the complexity of the interaction structures. Forest fires is one such risk events and motivate our work.

Note that this manuscript is a long summary of all published and drafted papers based on the thesis work. Published and submitted papers can be found in Appendix.

This chapter provides a review on construction methods and proposed point process models for multi-structure point patterns (based on Raeisi et al. (2019), see Appendix A) and about the use of point process theory to forest fire occurrences.

1.1 Preliminaries

Fundamental concepts of the theory of point processes emerged from life tables, renewal theory and counting problems (Daley and Vere-Jones 2003). The modern theory has mainly been developed between 1940's and 1970's (see e.g. the monographs by Palm (Palm 1943), Feller (Feller 1950), Bartlett (Bartlett 1954), Matérn (Matérn 1960) and Cox (Cox 1955, 1962)) and is linked to nonlinear techniques in stochastic process theory (Bartlett 1955, Bosq 1998). From 1980's spatial and spatio-temporal point processes have then become a subject on their own right. Today, they cover a plethora of applications in ecology, forestry, astronomy, epidemiology, seismology, fishery...

Spatial (and spatio-temporal) point process data are a collection of points for which locations (and times) of occurrence have been observed in a specified spatial region (and temporal period). Usually, the terms *points* and *events* are respectively used for arbitrary locations and for observations. The main goals in the analysis of point patterns concern the specification of intensity variations (first-order moment), interaction between events (second-order moment) and model identification for the underlying process. Processes are often classified into three classes of interaction structure (Diggle 1983):

- *randomness*: In the absence of any interaction between events, a point pattern is said Completely Spatially (or Spatio-Temporally) Random in the sense that the probability that an event occur at any point is equally likely to occur anywhere within a bounded region and that its location (and time) is independent of each any other event. This property provides the standard baseline against which point patterns are often compared. The simplest and most fundamental point process for modelling a complete random distribution of points is the Poisson point process (Kingman 1993, 2006). It is used as null hypothesis for statistical test of interaction (Diggle 2003, Illian et al. 2008).
- *clustering or aggregation*: In a clustered distribution, events tend to be closer than would be expected under complete randomness. Clustered patterns are mainly modelled 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 Kendall 2002, Gabriel 2014) and Shot-Noise Cox processes (Brix and Chadœuf 2000, Møller and Waagepetersen 2004, Møller and Diaz-Avalos 2010).
- *inhibition or regularity*: In a regular distribution, events are more evenly spaced than would be expected under complete randomness. This structure can be modelled by Strauss processes (Strauss 1975, Cronie and van Lieshout 2015), Matérn hard core processes (Matérn 1960, Gabriel et al. 2013) or determinantal point processes (Macchi 1975, Lavancier et al. 2015).

Gibbs processes (Ruelle 1969, Preston 1976, Dereudre 2019) offer a large class of models which allow any of the above interaction structure.

These single-structure point process models are too simplistic to describe phenomena with interactions at different spatial or spatio-temporal scales. That is for instance the case of seismic data as the different sources of earthquakes (faults, active tectonic plate and volcanoes) produce events with different displacements (Siino et al. 2017) and can be seen as the superposition of background earthquakes (which are distributed over a large spatio-temporal scale with low density) and clustered earthquakes (which are distributed over a small spatio-temporal scale with high density) (Pei et al. 2012).

Such multi-structure phenomena motivate statisticians to construct new spatial point process models, e.g. in ecology (Levin 1992, Wiegand et al. 2007, Picard et al. 2009), in epidemiology (Iftimi et al. 2017) and in seismology (Siino et al. 2017, 2018b), mainly based on Gibbs processes, but not only (Lavancier and Møller 2016). There are very few spatio-temporal models: Gabriel et al. (2017) modeled the multi-scale spatio-temporal structure of forest fires occurrences by log-Gaussian Cox processes (LGCP), Iftimi et al. (2018) developed a multi-scale area-interaction model for varicella cases and Illian et al. (2012b) modelled the locations of muskoxen herds by LGCP with a constructed covariate measuring local interactions.

In the spatial point processes literature, three general approaches are considered for constructing multi-structure point process models: thinning and superposition (Chiu et al. 2013), hybridization (Baddeley et al. 2013). Thinning consists in deleting points of a point process according to some probabilistic rule which is either independent or dependent of thinning other points (Chiu et al. 2013). This operation allows to get point processes with inhibition at small scales and attraction at large scales (Andersen and Hahn 2016, Lavancier and Møller 2016). Superposition of several processes is the union of the points of each process. It can be useful to model multi-scale clustered processes (Wiegand et al. 2007).

Hybridization consists in combining two or more point process models (Baddeley et al. 2015). Spatial hybrids of Gibbs models are defined in Baddeley et al. (2013) and hybrids of area-interaction potentials in Picard et al. (2009). Extension of the hybridization approach to the spatio-temporal framework has recently been considered in Iftimi et al. (2018).

The key contribution of this thesis is to develop new Gibbs and/or Cox model-based spatio-temporal multi-scale point processes by using hybridization. In the following, we review available models and methods for spatial (spatio-temporal) multi-structure point patterns and an overview of point process-based analyses and modeling of forest fire occurrences.

1.2 Point processes methods

We consider a finite spatial or spatio-temporal point process X observed in \mathcal{W} , where \mathcal{W} denotes either a spatial region $S \subset \mathbb{R}^d$ or a spatio-temporal region $S \times T \subset \mathbb{R}^d \times \mathbb{R}$. We denote \mathbf{x} a realization of the point process, i.e. a collection of events $\{\xi_i\}_{i=1,\dots,n}$ (or $\{(\xi_i, t_i)\}_{i=1,\dots,n}\} \subset \mathcal{W}$. Let η be any point in \mathcal{W} . We refer to Daley and Vere-Jones (2003), Chiu et al. (2013) (resp. Diggle and Gabriel (2010), Diggle (2013), Gonzalez et al. (2016)) for more formal definitions of spatial (resp. spatio-temporal) point processes. Without loss of generality, we set $d = 2$ throughout this thesis. The main characteristics driving the spatial (resp. spatio-temporal) distribution of points are the

intensity function, which governs the univariate distribution of the points of X , and the *pair correlation function*, which governs the bivariate distribution of the points of X , i.e. the interaction between events. In the following we remind some definitions and properties when X is a spatial or a spatio-temporal point process.

Campbell's theorem (Chiu et al. 2013) relates the expectation of a function, h assumed to be non-negative and measurable, summed over a point process X to an integral involving the mean measure of the point process :

$$\mathbb{E} \left[\sum_{\eta_1, \dots, \eta_k \in X}^{\neq} h(\eta_1, \dots, \eta_k) \right] = \int \dots \int h(\eta_1, \dots, \eta_k) \lambda^{(k)}(\eta_1, \dots, \eta_k) \prod_{i=1}^k d\eta_i,$$

where $\eta_i \in \mathcal{W}$ and $\lambda^{(k)}$, $k \geq 1$, are the product densities. For a simple point process, i.e. $\eta_i \neq \eta_j$ for $i \neq j$, if they exist, the product densities are related to the counting measure N in infinitesimal spatial or spatio-temporal regions $d\eta_1, \dots, d\eta_k \subset \mathcal{W}$, around η_1, \dots, η_k , with volumes $|d\eta_1|, \dots, |d\eta_k|$: $\mathbb{P} [N(d\eta_1) = 1, \dots, N(d\eta_k) = 1] = \lambda^{(k)}(\eta_1, \dots, \eta_k) \prod_{i=1}^k d\eta_i$. Thus, the intensity function is related to the expected number of points in infinitesimal regions

$$\lambda(\eta) = \lambda^{(1)}(\eta) = \lim_{|d\eta| \rightarrow 0} \frac{\mathbb{E}[N(d\eta)]}{|d\eta|}$$

and the pair correlation function is defined by

$$g(\eta_i, \eta_j) = \frac{\lambda^{(2)}(\eta_i, \eta_j)}{\lambda(\eta_i)\lambda(\eta_j)}. \quad (1.1)$$

A point process is *homogeneous* when its intensity is constant, $\lambda(\eta) = \lambda$, $\forall \eta$, *inhomogeneous* otherwise. In practice, the inhomogeneity is often driven by environmental covariates and we account for them by using parametric models for the intensity function (Baddeley et al. 2015). Under the assumption of *stationarity*, the properties of the point process are invariant under translation and the process is homogeneous. The *second-order stationarity* states that the second-order intensity only depends on the difference between points $\lambda^{(2)}(\eta_i, \eta_j) = \lambda^{(2)}(\eta_i - \eta_j)$. Because in practice most of processes are inhomogeneous, Baddeley and Turner (2000) and Gabriel and Diggle (2009) weakened it and defined the *second-order intensity-reweighted stationary* assumption for which the pair correlation function (1.1) is well-defined and a function of $\eta_i - \eta_j$. van Lieshout (2019) provides general concepts of factorial moment properties. The previous definition of inhomogeneous processes is not unique, Hahn et al. (2015) defined inhomogeneous model classes (including the class of reweighted second-order stationary processes) into the common general framework of hidden second-order stationary processes. The pair correlation function describes the structure of dependence/interaction

between points : $g(\eta_i, \eta_j) = 1, > 1$ and < 1 indicates that the pattern is, respectively, completely random, clustered and regular.

Assume that the distribution of the point process is defined by a *probability density* $f(\mathbf{x})$ with respect to the distribution of a unit rate Poisson process. The probability density can be used to study point processes. It can be viewed as the probability of getting the point pattern \mathbf{x} , divided by the same probability under Complete Randomness (Baddeley et al. 2015). The mathematical form of the probability density determines the structure of the point process, see Coeurjolly et al. (2017) and Coeurjolly and Lavancier (2019) about formulation of the density of point processes. A closely related concept is the Papangelou conditional intensity function (Papangelou 1974), which has been extended to the spatio-temporal framework by Cronie and van Lieshout (2015). It is defined by

$$\lambda(\eta|\mathbf{x}) = \frac{f(\mathbf{x} \cup \eta)}{f(\mathbf{x})}, \quad (1.2)$$

for $\eta \notin \mathbf{x}$ provided $f(\mathbf{x}) \neq 0$.

1.2.1 Classical point process models

We refer to Diggle (2003), Møller and Waagepetersen (2004), Illian et al. (2008), Chiu et al. (2013), Baddeley et al. (2015) and Cronie and van Lieshout (2015), Diggle and Gabriel (2010), Diggle (2013), Gabriel et al. (2013), Gonzalez et al. (2016) for a presentation of most of spatial and spatio-temporal point process models. Hereafter we only focus on the ones mentioned/used in Section 1.3 to construct multi-structure point process models, namely the Poisson, Cox and Gibbs processes.

Poisson process

The Poisson point process is the reference model for independence of the locations of events, i.e. for complete spatial (or spatio-temporal) randomness. It is also the simplest and most widely used inhomogeneous point process model. Poisson point processes with intensity function $\lambda(\eta)$ are defined by two postulates :

- The number of points in any region $B \subseteq \mathcal{W}$, $N(B)$, follows a Poisson distribution with parameter $\int_B \lambda(\eta) d\eta$,
- For all $B \subseteq \mathcal{W}$, given $N(B) = n$, the n events in B form an independent random sample from the distribution on B with probability density function $\lambda(\eta) / \int_B \lambda(\eta) d\eta$.

The probability density of a Poisson point process with respect to the unit rate Poisson process is

$$f(\mathbf{x}) = \exp\left(|\mathcal{W}| - \int_{\mathcal{W}} \lambda(\eta) d\eta\right) \prod_{\eta \in \mathbf{x}} \lambda(\eta).$$

Then, from Equation (1.2), the Papangelou conditional intensity is $\lambda(\eta|\mathbf{x}) = \lambda(\eta)$ and $\lambda^{(2)}(\eta_i, \eta_j) = \lambda(\eta_i)\lambda(\eta_j)$, so that $g(\eta_i, \eta_j) = 1$.

Cox process

Cox processes, so-called doubly stochastic point processes (Cox 1955), are considered as a generalization of inhomogeneous Poisson processes where the intensity is a realization of a random field $\Lambda = \{\Lambda(\eta)\}_{\eta \in \mathcal{W}}$. These models are particularly useful as soon as spatial variation in events density reflects both the environment and dependence between events. Moreover, their first- and second-order moments being tractable, they are very attractive. We have

$$\lambda(\eta) = \mathbb{E}[\Lambda(\eta)] \quad \text{and} \quad g(\eta_i, \eta_j) = \frac{\mathbb{E}[\Lambda(\eta_i)\Lambda(\eta_j)]}{\lambda(\eta_i)\lambda(\eta_j)} = 1 + \frac{\text{cov}(\Lambda(\eta_i), \Lambda(\eta_j))}{\lambda(\eta_i)\lambda(\eta_j)}. \quad (1.3)$$

The probability density $f(\mathbf{x}) = \mathbb{E}[\exp(|\mathcal{W}| - \int_{\mathcal{W}} \Lambda(\eta) d\eta) \prod_{\eta \in \mathbf{x}} \Lambda(\eta)]$ is intractable for these processes. Consequently, the Papangelou conditional intensity is not known. The second-order intensity function $\lambda^{(2)}(\eta_i, \eta_j) = \mathbb{E}[\Lambda(\eta_i)\Lambda(\eta_j)]$ is only tractable for two special cases of Cox processes, that we present below, the *Shot Noise Cox process* and the *log-Gaussian Cox process*.

Shot noise Cox processes (Møller 2003) (SNCP) are a wide class of Cox processes associated to

$$\Lambda(\eta) = \sum_{(c, \gamma) \in \Phi} \gamma k(c, \eta),$$

where Φ is a Poisson point process on $\mathcal{W} \times [0, \infty)$ with intensity measure ζ and $k(c, \cdot)$ is a density function on \mathcal{W} , $\forall c \in \mathcal{W}$. The intensity and pair correlation function are

$$\lambda(\eta) = \int \gamma k(c, \eta) d\zeta(c, \gamma) \quad \text{and} \quad g(\eta_i, \eta_j) = 1 + \frac{\int \gamma^2 k(c, \eta_i) k(c, \eta_j) d\zeta(c, \gamma)}{\lambda(\eta_i)\lambda(\eta_j)}.$$

SNCP include Poisson cluster processes, i.e. a Poisson process in which each point is replaced by a cluster of points, the original point is considered as the cluster center (Cox and Isham 1980). When the points in the cluster are independently and identically distributed about the cluster centre, the process is referred to as a Neyman-Scott process (Neyman and Scott 1958). Two mathematically tractable models of Neyman-Scott processes are the *Thomas process* (Thomas 1949), where k is a zero-mean normal density, and the *Matérn cluster process*, where k is a uniform density on a ball centered at the origin.

Log-Gaussian Cox processes (LGCP) have been introduced in Møller et al. (1998), considering that the intensity is a log-Gaussian process : $\Lambda(\eta) = \exp(Z(\eta))$, where Z is a real-valued Gaussian random field, with mean function $\mu(\eta)$ and covariance function

$C(\eta_i, \eta_j)$. In that case, from Equation (1.3) we have

$$\lambda(\eta) = \exp(\mu(\eta) + C(\eta, \eta)/2), \forall \eta \in \mathcal{W} \text{ and } g(\eta_i, \eta_j) = \exp(C(\eta_i, \eta_j)), \forall \eta_i, \eta_j \in \mathcal{W}.$$

The expression of the pair correlation function shows that the interaction is controlled by the second-order moment of Z . If $C(\eta_i, \eta_j) \geq 0$, we get $g(\eta_i, \eta_j) > 1$ and clustering. As they are based on a latent random field describing the intensity, LGCPs have a hierarchical structure making them particularly flexible (Illian et al. 2008). Note that the interaction is controlled through the second-order moment of the Gaussian random field, so that LGCPs do not describe the mechanistic process generating the points what is the case of most of Gibbs processes (see below) for which the dependence between points is controlled through local interaction between pairs of points.

Gibbs process

A finite Gibbs point process on \mathcal{W} admits a density

$$f(\mathbf{x}) = \exp(-\Psi(\mathbf{x})) \quad (1.4)$$

w.r.t. the Poisson process of unit intensity on \mathcal{W} . The potential function Ψ is often specified as the sum of pair potentials :

$$\Psi(\eta_1, \dots, \eta_n) = \alpha_0 + \sum_i \alpha_1(\eta_i) + \sum_{i < j} \alpha_2(\eta_i, \eta_j) + \dots + \alpha_n(\eta_1, \dots, \eta_n), \quad (1.5)$$

with α_0 a normalizing constant for the density and the pair potentials $\alpha_1, \alpha_2, \dots$ which determine the contribution to the potential from each δ -uple of points. Note that, if the $\alpha_\delta, \delta \geq 2$ are identically zero, the process is Poisson with intensity $\lambda(\eta) = \exp(-\alpha_1(\eta))$. Hence, α_1 can be viewed as controlling a spatial (or spatio-temporal) trend, while the $\alpha_\delta, \delta \geq 2$ control the interactions between events. The normalizing constant is generally intractable, so it is often impossible to compute the intensity and pair correlation function of Gibbs processes. However, the Papangelou conditional intensity can be computed (Coeurjolly and Lavancier 2019).

When the interaction between points is restricted to pairs, i.e. for

$$f(\mathbf{x}) = \alpha \prod_i \beta(\eta_i) \prod_{i < j} \gamma(\eta_i, \eta_j),$$

with $\alpha > 0, \beta$ an intensity function and γ a symmetric interaction function, the process is called *pairwise interaction process* (Diggle 1983, van Lieshout 2000). A well-known example of such processes is the *Strauss process* (Strauss 1975) for which

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \gamma^{s(\mathbf{x})},$$

where $\beta, \gamma > 0$, $n(\mathbf{x})$ is the number of points in \mathbf{x} and $s(\mathbf{x})$ the number of neighbour pairs of \mathbf{x} at distances less than a given distance R . When $\gamma = 0$, we get the *hardcore process*. Note that in the Strauss process, γ should be smaller than 1 otherwise the density is no integrable. Geyer (1999) modified the Strauss process and proposed the *Geyer saturation process* in which the overall contribution from each point is trimmed to never exceed a maximum value. We thus have

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{\eta \in \mathbf{x}} \gamma^{\min(s, t(\eta, r, \mathbf{x}))}, \quad (1.6)$$

where $\alpha, \beta, \gamma, r, s$ are parameters and $t(\eta, r, \mathbf{x})$ is the number of other events lying with a distance r of the point η .

1.3 Review on multi-structure point processes

Spatial and spatio-temporal single-structure point process models presented in Section 1.2 are generally used when only one type of interaction governs the structure of the point pattern. When there are indications that the spatial or spatio-temporal structure combines several structures or varies with ranges of distances, we need to consider multi-structure point process models.

We present in this section some of these models derived from the classes of Gibbs and Cox processes. By nature, few spatial point processes can exhibit directly several structures and/or scales of interaction and we recall some useful construction techniques to incorporate the multi-structure: hybridization, thinning, superposition or clustering.

1.3.1 Models based on Gibbs processes

Gibbs point processes are mainly used to model repulsion structure in point patterns, even if some examples exist for modelling low clustering (Chiu et al. 2013). Their definition through the potential function Ψ fit well in the statistical mechanics framework where the spatial modelling of particles needs often to consider their interaction. It is common in various domains (mechanics, biology. . .) to observe repulsion at short range and aggregation at medium-long range of entities, leading to define multi-structure point processes models.

For pairwise interaction processes, some parametric potential functions can be defined to take into account multiple scales of interaction, see e.g. Ruelle (1969), Ogata and Tanemura (1981), Penttinen (1984), Clyde and Strauss (1991), Habel et al. (2019). We consider in the sequel the homogeneous case, i.e. when α_1 is constant and the pair potential function $\alpha_2(\eta_i, \eta_j) = \alpha_2(\|\eta_i - \eta_j\|)$ in (1.5).

The Lennard-Jones pair potential function, well-known in statistical mechanics, is

given by

$$\alpha_2(r) = \epsilon_1 \left(\frac{\sigma}{r}\right)^{m_1} - \epsilon_2 \left(\frac{\sigma}{r}\right)^{m_2}, \quad \forall r > 0$$

where $m_1 > m_2$, $\epsilon_1, \sigma > 0$ and in the multi-structure case $\epsilon_2 > 0$. Another one is the step potential function given by

$$\alpha_2(r) = c_l \quad \text{if } R_{l-1} < r \leq R_l \quad \forall l = 1, \dots, m$$

where $R_0 = 0$, $R_m = \infty$, $c_1 = \infty$, $c_m = 0$ and $c_l \in \mathbb{R}$ for $l = 2, \dots, m-1$. The resulting model is an extension of the Strauss process to the multi-scale framework (Penttinen 1984). The square-well potential is obtained with $l = 2$. More recently, Goldstein et al. (2015) introduced a pair potential function varying smoothly over distance with scale interactions defined through a differential system of equations. Other pair potential functions can be found in the literature for modeling multi-structure phenomena, e.g. in Ogata and Tanemura (1981) and Chiu et al. (2013).

Some of these pair potential functions define multi-scale generalizations of single scale Gibbs processes. Indeed, the step potential functions of homogeneous pairwise interaction processes in Diggle (1983) and Penttinen (1984) represent multi-scale extensions of the Strauss process where the density is given by

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{l=1}^m \gamma_l^{s_l(\mathbf{x})},$$

where $s_l(\mathbf{x}) = \sum_{i < j} \mathbb{1}(R_{l-1} < \|\eta_i - \eta_j\| \leq R_l)$.

In the same way, the multi-scale generalization of the area-interaction model has been introduced in Ambler (2002) and Ambler and Silverman (2004, 2010) with a two-scale structure and in Picard et al. (2009) for multi-scale marked area-interaction processes. Its density function in a homogeneous multi-scale case is given by

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{l=1}^m \exp(-\kappa_l U(\mathbf{x}, r_l))$$

where $U(\mathbf{x}, r_l)$ is the d -dimensional volume of the set $\mathcal{W} \cap \bigcup_{\eta \in \mathbf{x}} b(\eta, r_l)$, with $b(\eta, r_l)$ the ball centered at η_i of radius $r_l > 0$. The sign of κ_l defines the l th structure : inhibition if negative, clustering otherwise. Nightingale et al. (2019) used area-interaction point processes for bivariate point patterns for modelling both attractive and inhibitive intra- and inter-specific interactions of two plant species.

Baddeley et al. (2013) defined a new class of multi-scale Gibbs point processes named hybrid models and including the two previous generalization examples. This

unified framework allows to define properly generalizations of single-scale Gibbs point processes by preserving Ruelle and local stability (van Lieshout 2000). This 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, so that

$$f(\mathbf{x}) = cf_1(\mathbf{x})\dots f_m(\mathbf{x})$$

where c is a normalization constant and f_l is a Gibbs density function for $l = 1, \dots, m$. The choice of the normalization constant allows to well define a probability density in the case where the product $f_1\dots f_m$ is integrable. The integrability condition is of course essential and induced by others conditions on the f_l (Ruelle stability, local stability or hereditary), see Baddeley et al. (2013) which play an important role in simulation algorithms and are established in general to demonstrate the model validity of the hybrid process.

Baddeley et al. (2013) introduced the spatial multi-scale Geyer saturation point process that was applied in epidemiology by Iftimi et al. (2017) and in seismology by Siino et al. (2017) and Siino et al. (2018b). Iftimi et al. (2018) extended the hybridization approach to the spatio-temporal framework and introduce the spatio-temporal multi-scale area-interaction point process (see Section 2.1.1). In Chapter 2 (resp. Chapter 3), we define (resp. provide an estimation procedure) the inhomogeneous spatio-temporal multi-scale Geyer saturation process. This work is published in Raeisi et al. (2021b), see Appendix B. Rajala et al. (2018) used a multitype generalization of Gibbs point processes with point-to-point interactions at different spatial scales in order to model a complex rainforest data of 83 species.

The definition of hybrid Gibbs models does not impose to consider the same m Gibbs models which is emphasized in Baddeley et al. (2015). In this way, Badreldin et al. (2015) applied a hybrid model with three model structures at different ranges of distance to the spatial pattern of halophytic species distribution in an arid coastal environment. They considered a hardcore process at very short distances, a Geyer process at short to medium distances and a Strauss process for the structure at large distances. Wang et al. (2020) fitted a spatial hybrid Geyer hardcore point process on the spatial distribution of trees. In Chapter 2, we extend this type of hybrids to the spatio-temporal context. This work is submitted for publication in Raeisi et al. (2021a), see Appendix C.

As a different approach to model the repulsion at short range and aggregation at medium-long range, Vihrs et al. (2021) embedded spatially structured Gaussian random effects in the Gibbs trend function and introduced, in particular, the log Gaussian Cox Strauss point process which we extend to the spatio-temporal context in Chapter 2.

1.3.2 Models based on Cox processes

Cox processes are mainly defined from additive or log-linear random intensity functions. Their hierarchical structure allows to quantify the various sources of variation governing the spatial or spatio-temporal distribution of the pattern of interest. They are widely used for modelling environmental and ecological patterns.

Cluster Cox processes and superposition

Some Cox processes are obtained by clustering of *offspring* points around *parent* points and correspond to specific cases of cluster processes. This two-step construction allows to consider easily different structures for the patterns of parents and offspring.

Møller and Torrisi (2005) introduced the class of Generalized Shot Noise Cox processes (GSNCP), extending the definition of SNCP, and allowing relevant multi-structure point processes for modelling regularity and clustering in many applications. This class has two advantages. Firstly, the parent process is not restricted to be Poisson, as in Neyman-Scott processes, and can be a repulsive Gibbs point process in order to add inhibition between the clusters. Secondly, in each cluster, the intensity and the bandwidth of the dispersion kernel can be random. By consequence, a GSNCP is a Cox process driven by a random field of the form

$$\Lambda(\eta) = \sum_{(c,\gamma,h) \in \Phi} \gamma k_h(c, \eta),$$

where Φ is a point process on $\mathcal{W} \times [0, \infty) \times [0, \infty)$ and h is a bandwidth for the kernel density $k_h(c, \cdot)$. So, given Φ , a GSNCP is distributed as the superposition $\cup_l X_l$ of independent Poisson processes with intensity functions $\gamma_l k_{h_l}(c_l, \cdot)$ where $\{\gamma_l\}_l$, $\{h_l\}_l$ are random and $\Phi_{cent} = \{c_l\}_l$ is the parent process. In population dynamics, with G_0 a Poisson process for the initial population and G_{n+1} a GSNCP where the cluster centers are given by G_n , the superposition of GSNCPs G_0, G_1, \dots is a spatial Hawkes process (Hawkes 1971). The GSNCP class contains the special cluster Cox process defined in Yau and Loh (2012), where the parents process is a Strauss process. This model coupling inhibition at medium/long range and aggregation in cluster is applied to tree locations in a rain-forest, in order to consider the competition and reproduction mechanisms. Albert-Green (2016) and Albert-Green et al. (2019) generalized the Neymann-Scott process by considering a log-Gaussian Cox process model for the parents, instead of a homogeneous Poisson process, leading to two scales of clustering, inter- and intra-clusters. This hierarchical model is applied to storm cell modelling in North Dakota.

Wiegand and co-authors' papers (Wiegand et al. 2007, 2009) consider several construction of Cox processes incorporating clustering at multiple scales. The nested

double-cluster process is an extension of the Thomas process in an multi-generation evolution of the population where the offspring become parents and generate offspring. They consider also the superposition of cluster processes, like the Thomas process.

Cox processes with constructed covariate

Another way to incorporate both small and large spatial scale structure in Cox processes is to define a constructed covariate measuring the local structure of a point pattern associated to an additional spatial effect at medium-long range. This methodology developed in Illian et al. (2012a) and applied to koala data is used again in Illian et al. (2012b, 2013) for other spatial ecological data. They consider a log-Gaussian Cox process in a Bayesian framework in order to apply the INLA approach for speeding up the estimation of parameters in comparison to MCMC approaches that are very time-consuming. Gabriel et al. (2017) used also this approach in the context of wildfire modelling in Mediterranean France. In the case of a spatial LGCP model, the method consists in estimating the random field Λ on grid cells s_i as follow

$$\Lambda(s_i) = \exp \left(\beta_0 + f(z_c(s_i)) + \sum_{k=1}^p f_k(z_k(s_i)) + Y(s_i) \right)$$

where β_0 is the intercept, $f(z_c(\cdot))$ is a function of the constructed covariate z_c , f_k , $k = 1, \dots, p$ are functions of the observed covariates z_k and Y is a Gaussian random field taking into account the spatial autocorrelation not explained by the covariates. This intensity is estimated for each cell s_i of a grid partitioning the observation window.

In Illian et al. (2012a), the constructed covariate at each center point c of the grid cell s is the distance from c to the nearest point in the pattern outside the grid cell, i.e $z_c(s) = \min_{\eta \in \mathbf{x} \setminus s} (\|c - \eta\|)$. This constructed covariate describes small scale inter-individual behavior whereas the random field Y captures the spatial autocorrelation at a large spatial scale. The space-time and space-mark extensions of the constructed covariate definition are respectively introduced in Illian et al. (2012b) and Illian et al. (2013). In Gabriel et al. (2017) the constructed covariate corresponds to a temporal intensity index given by the ratio between the number of wildfires observed spatially close to an other in a specified period and the total number of closed wildfires observed outside this given period. This covariate measures the temporal wildfire inhibition at close spatial distances induced by the local burn of vegetation after a wildfire occurrence. Sørbye et al. (2019) fitted a LGCP to rainforest tree species by adding to the combination of covariates in the log-intensity a spatial random field and error field. The first random field captures the spatial autocorrelation in point counts among neighboring grid cells and the second one the clustering within grid cells, as a nugget effect in geostatistics.

The intensity in $s \in \mathcal{W}$ is thus given by

$$\Lambda(s) = \exp \left(\beta_0 + \sum_{k=1}^p \beta_k z_k(s) + \frac{1}{\sqrt{\tau}} \left\{ \sqrt{\rho} \times Y(s) + \sqrt{1 - \rho} \times \epsilon(s) \right\} \right)$$

where β_k are linear effects of observed covariates z_k , Y is a spatial random field with autocorrelation between grid cells and ϵ the error field driving the aggregation structure within grid cells.

Thinned point processes

Thinning is an operation allowing to delete points in a point process in order to obtain a new one with different characteristics. Each point of a point process has a probability $1 - \pi$ of deletion, where the retention probability π can be constant or not, independent of the location point or depending on one to several points. For Cox processes, this technique is generally applied to create random local regularity. For example, Andersen and Hahn (2016) applied a Matérn hard core dependent thinning to a Shot Noise Cox process to obtain short range repulsion with medium range clustering. For a given point pattern and a specified distance h , Matérn hard core thinning acts by first attaching random positive marks (arrival times) to each point. Subsequently a point is removed if it has a neighbour within distance h and with a smaller mark (i.e. the neighbour arrived earlier). In that way, for a given location η , the retention probability $\pi(\eta)$ is the ratio between the intensities of the thinned process and the original process at η . Lavancier and Møller (2016) extended the definition of interrupted point processes in Stoyan (1979) and Chiu et al. (2013) and considered a spatial point process X obtained by an independent thinning driven by a random process Z on a regular point process Y . An example is given with Y a Matérn hard core process and Z the transformation by a characteristic function of a Boolean disc model (Chiu et al. 2013).

1.4 Point process-based analysis and modeling of forest fire occurrences

Statistical modeling of forest fires appeared in the late 1970s with the works of Wilkins (1977) and Dayananda (1977). More recently, the theory of point processes has been considered as statistical tools to analyze spatial (spatio-temporal) structures of forest fire occurrences. Podur et al. (2003) analyzed the occurrences of lightning-caused forest fires in Canada. Raeisi (2018), Pereira and Turkman (2018), and Xi et al. (2019) reviewed briefly the literature in forest fire occurrences modeling with (spatial) spatio-temporal point processes. The studies are classified into two types: those related to exploratory analysis, based on summary statistics and those related to model analysis which might include non-Poisson point process models.

1.4.1 Exploratory analysis of forest fires

In exploratory analysis of point patterns, aims are to map local spatial/spatio-temporal density variations and to test interactions between points. The former is achieved by estimating the first-order intensity, the later by using summary statistics as the pair correlation function or any related function. One can further investigate separability in space, time and/or marks, see Gabriel and Diggle (2009), Gonzalez et al. (2016).

The intensity function reflects the average density of points in the point process and can be used to identify areas with a high or low expected number of points. It can be estimated from parametric and non-parametric methods. Estimating kernel density of fire occurrences patterns indicate heterogeneous spatial distributions with hotspots (Yin et al. 2007, del Hoyo et al. 2011, Gralewicz et al. 2011, Gonzalez-Olabarria et al. 2015, Fuentes-Santos et al. 2016, Yin et al. 2019, Li and Banerjee 2020) and can be used to create maps of “fire occurrence zones” (Koutsias et al. 2004, 2014, 2015). Parametric models further quantify the influence of various covariates that drive local density variations of fire ignitions (Yang et al. 2008, Mundo et al. 2013). The inclusion and test of spatially (or spatio-temporally) varying covariates in intensity function has been of particular interest in Diaz-Avalos et al. (2014), Borrajo et al. (2017, 2018, 2020a,b), and Myllymäki et al. (2021).

Fuentes-Santos et al. (2017) and Fuentes-Santos et al. (2021) proposed non-parametric tests to compare the spatial distribution of two observed forest fire patterns based on comparison of their first-order intensities.

For spatio-temporal point patterns, one can preliminary test for first-order separability. Schoenberg (2004), Diaz-Avalos et al. (2013), and Fuentes-Santos et al. (2018) show that the intensity of forest fire occurrences varies in space and time in a non-separable way.

Dependencies between points can be described through the analysis of second-order characteristics. From spatial (Genton et al. 2006) or spatio-temporal (Wang and Anderson 2010, Comas et al. 2014, Costafreda-Aumedes et al. 2016, Tonini et al. 2017) K or L -functions, forest fire occurrences show clustering, i.e. a local over-density in space and/or in time. Some analyses also identified multi-structure properties. Forest fire occurrences can be spatially clustered at a relatively small scale but regularly spaced at a larger scale (Podur et al. 2003), significantly clustered at different spatial scales (Gonzalez-Olabarria et al. 2011, Juan et al. 2012). In the spatio-temporal framework, Tonini et al. (2017) carried out a spatio-temporal clustering analysis for forest fires in Portugal and found a complex structure associated with different behavior according to the size of the fire ”*medium fires tend to aggregate around small fires, while large fires aggregate at a larger distance and longer times, indicating that the return time*

Table 1.1 Overview of spatial point process-based studies for modelling forest fires

Studies	Model	Estimation	Study area
Podur et al. (2003)	Poisson	Composite likelihood	Canada
Yang et al. (2007)	Poisson	Composite likelihood	USA
Yang et al. (2008)	Poisson	Composite likelihood	USA
Hering et al. (2009)	Poisson and Geyer	Composite likelihood	USA
Turner (2009)	Poisson and Geyer and Strauss hardcore	Composite likelihood	Canada
Juan et al. (2012)	Poisson, Thomas and area-interaction	Composite likelihood	Spain
Liu et al. (2012)	Poisson	Composite likelihood	China
Miranda et al. (2012)	Poisson	Composite likelihood	USA
Serra et al. (2013)	Poisson, Thomas and area-interaction	Composite likelihood	Spain
Trilles et al. (2013)	Poisson and area-interaction	Composite likelihood	Spain
Yang et al. (2015)	Poisson	Composite likelihood	USA
Arago et al. (2016)	Poisson and area-interaction	Composite likelihood	Spain
Woo et al. (2017)	Poisson and area-interaction	Composite likelihood	Korea
Juan (2019)	LGCP	INLA	Spain
Gomez-Rubio (2020)	LGCP	INLA	Spain
Baile et al. (2021)	Multifractal Cox	Composite likelihood	France

following these events is longer than for small and medium fires". Defining a normalized empirical intensity ratio index, Gabriel et al. (2017) showed that inhibitive patterns between neighboring events can span several years.

Some other studies considered first-order and second-order characteristics for identifying drivers and spatial distribution of wildfires (Kwak et al. 2009, 2010, Gua et al. 2011, Pereira and Turkman 2012, Fuentes-Santos et al. 2013, Zhang et al. 2013, Gua et al. 2015, 2016, 2017, Bates et al. 2018, Rihan et al. 2019).

Further information can be added to fire locations, as the burned areas or the cause of ignition, and can be treated as marks. Zhang and Zhuang (2014) proposed a local odds ratio approach to estimate the localized dependence structure between burned areas and fire locations. Schoenberg (2004) found evidence of a lack of separability between fire occurrences and sizes due to small-scale clustering. Hence, the marks may not be separable from the points. Pereira and Turkman (2012) also rejected the separability assumption between space, time and marks.

1.4.2 Point process models for forest fires

In the last two decades, forest fire occurrences have been modeled with point process models. Spatial point process models mainly include Poisson and area interaction models, while spatio-temporal models are Cox models (Log-Gaussian and Shot Noise Cox processes). Tables 1.1 and 1.2 report spatial and spatio-temporal models used in different articles, as well as the inference method used for model fitting and the study area. These Tables show that models tend to be more complex along years, as an attempt to better include the relatively large number of covariates (e.g. land use and the meteorological covariates such as temperature and precipitation that are measured in time), but also the interaction structures highlighted in the previous section.

Hawkes point process models have also been considered for forest fire occurrences

Table 1.2 Overview of spatio-temporal point process-based studies for modelling forest fires

Studies	Model	Estimation	Study area
Møller and Diaz-Avalos (2010)	SNCP	Composite likelihood	USA
Serra et al. (2012)	LGCP	INLA	Spain
Pereira et al. (2013)	LGCP	INLA	Portugal
Serra et al. (2014a)	Poisson hurdle model	INLA	Spain
Serra et al. (2014b)	LGCP	INLA	Spain
Najafabadi et al. (2015)	SNCP	Composite likelihood	Iran
Gabriel et al. (2017)	LGCP	INLA	France
Opitz et al. (2020)	LGCP	INLA	France
Juan (2020)	LGCP	INLA	Spain

(Brillinger et al. 2003, Peng et al. 2005, Schoenberg et al. 2007, 2009, Xu and Schoenberg 2011, Taylor et al. 2013). Due to the occurrence of human-caused forest fires on network of roads, paths and trails, (Uppala and Handcock 2014, Comas et al. 2019) model forest fires on linear networks. Marked point process models allowed to consider fire sizes (Diaz-Avalos et al. 2016). Pimont et al. (2021) developed a two-components Bayesian hierarchically-structured probabilistic model for daily fire activity which contains a spatio-temporal Poisson model for fire occurrence and piecewise-estimated Pareto and Generalized-Pareto distributions for fire sizes. As an extension, since the heavy tails of burned areas lead to a dominant influence of the most extreme wildfires, Koh et al. (2021) focus on accurate modeling of the distribution of extreme wildfires, and its spatio-temporal variation. As an alternative application of the marked point processes in forest fires modelling, Quinlan et al. (2021) considered the duration of fires as marks.

Different approaches can be used to fit these models, e.g. moment-, likelihood- and Bayesian-based methods. For most of point process models, the likelihood has no closed form expression (and thus is intractable). To address this issue, a simple and quick inference procedure is using the composite likelihood-based inference, as the pseudo-likelihood (Baddeley and Turner 2000) or logistic likelihood (Baddeley et al. 2014) for Gibbs models. Cox models involve an unobserved (Gaussian) random field in which composite likelihood estimation methods would involve complex integrals. These models are hierarchical and are therefore naturally suited to a Bayesian hierarchical approach for inference based on integrated nested Laplace approximation (INLA) (Rue et al. 2009) or on the Markov chain Monte Carlo (MCMC) method (Taylor et al. 2015). See Taylor and Diggle (2014) for a comparison.

1.5 Outline of thesis

As discussed in the previous section, the spatio-temporal distribution of forest fire occurrences is very complex in nature with non-separability in space and time and multiple structures (repulsion and aggregation) at different spatial and/or temporal scales. Spatio-temporal variations of fire occurrences further depend on the spatial distribu-

tion of current land use and meteorological conditions, but also depends on past events (changes in vegetation due to fires affect the probability of fire occurrence during a regeneration period). In this thesis we develop new models for such complex phenomena, as well as some methods for their inference and simulation.

In Chapter 2, based on the extension of hybridization approach (Baddeley et al. 2013) to the spatio-temporal framework (Iftimi et al. 2018), we propose the spatio-temporal hybrid Geyer saturation point process for multi-scale point patterns and spatio-temporal hybrid Strauss hardcore point process that combines both multi-scaling by hybridization and hardcore distances.

A different approach, leading to more flexibility in the model but also to more challenging inference, consists of Gibbs models that contain both random and fixed effects as in Illian and Hendrichsen (2010) to take into account complex patterns of heterogeneity. We propose a new modeling approach for this case, and we embed spatio-temporally structured Gaussian random effects in the Gibbs trend function. Therefore, our approach focuses on models derived from the multi-scale classes of combinations of Gibbs and log-Gaussian Cox point processes, to which we refer as Cox-Gibbs models in the following.

In Chapter 3, we aim to extend and implement available inference methods for these new models in the spatio-temporal framework. We classify the inference procedure into two approaches: global and local estimation methods. We tailor common methods for a global statistical inference in Gibbs models: the pseudo-likelihood and logistic likelihood approaches. However, the calculation of the likelihoods variants (composite likelihoods: pseudo-likelihood and logistic likelihood) used in point process inference would involve complex, high-dimensional integrals for Cox-Gibbs models, and we would need estimation methods that allow handling the latent (i.e., unobserved) Gaussian variables. However, due to the hierarchical structure of such models, they can be naturally formulated and estimated within a Bayesian hierarchical approach, using techniques as the INLA (Rue et al. 2009).

The models based on global parameter estimates can not take into account different local interaction structures. Thus, we extend the local likelihood approach (Baddeley 2017) to the spatio-temporal context as an alternative method for modeling multi-structure point patterns with spatially and/or temporally varying parameters in Gibbs point process models.

In Chapter 4, we implement a birth-death Metropolis-Hasting algorithm for simulating the hybrid Gibbs models. We propose a two-step procedure for simulating the hybrid Cox-Gibbs model by simulating, firstly, a realisation of a Gaussian random field

and then simulating a realisation of hybrid Cox-Gibbs model given that Gaussian realisation using the birth-death Metropolis-Hasting algorithm.

The models, estimation and simulation methods proposed in this thesis have been carried out using R together with the `spatstat` (Baddeley and Turner 2005), `stpp` (Gabriel et al. 2013), `splanCS` (Rowlingson and Diggle 1993), `fields` (Nychka et al. 2017), `sparr` (Davies et al. 2011), `raster` (Hijmans 2020), `INLA` (Lindgren and Rue 2015) and `GET` (Myllymäki and Mrkvička 2019) packages.

In Chapter 5, we investigate the hybrid Gibbs models proposed in Chapter 2 for fitting forest fire occurrences in South of France and in central Spain. We also develop an innovative application of spatio-temporal modeling of temperature hotspots, and in particular of temperature anomalies, in the United States using Cox-Gibbs models.

Chapter 2

New models

In this chapter, we extend two spatial Gibbs models to the spatio-temporal and multi-scale contexts and then propose a model which is a combination of Cox and Gibbs models.

2.1 Models based on Gibbs point processes

Gibbs models are flexible point processes that allow the specification of point interactions via a probability density defined with respect to the unit rate Poisson point process. These models allow to characterize a form of local or Markovian dependence amongst events. Gibbs point processes contain a large class of flexible and natural models that can be applied for:

- Postulating the interaction mechanisms between pairs of points,
- Taking into account clustering, randomness or inhibition structures,
- Combining several structures at different scales with the hybridization approach.

Let $\mathbf{x} = \{\eta_1, \dots, \eta_n\} = \{(\xi_1, t_1), \dots, (\xi_n, t_n)\}$ be a spatio-temporal point pattern where $(\xi_i, t_i) \in \mathcal{W} = S \times T \subset \mathbb{R}^2 \times \mathbb{R}$. We consider $(\mathcal{W}, d(\cdot, \cdot))$ where $d((u, v), (u', v')) := \max\{|u - u'|, |v - v'|\}$ for $(u, v), (u', v') \in \mathcal{W}$ is a complete, separable metric space. The cylindrical neighbourhood $C_r^q(u, v)$ centred at $(u, v) \in \mathcal{W}$ is defined by

$$C_r^q(u, v) = \{(a, b) \in \mathcal{W} : \|u - a\| \leq r, |v - b| \leq q\}, \quad (2.1)$$

where $r, q > 0$ are spatial and temporal radius and $\|\cdot\|$ 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 centre (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).

A finite Gibbs point process is a finite simple point process defined with a density $f(\mathbf{x})$ that satisfies the hereditary condition, i.e. $f(\mathbf{x}) > 0 \Rightarrow f(\mathbf{y}) > 0$ for all $\mathbf{y} \subset \mathbf{x}$.

A closely related concept to density functions is Papangelou conditional intensity function (Papangelou, 1974) which is a tool for simulating Gibbs models and inferring

its parameters. The Papangelou conditional intensity of a spatio-temporal point process on \mathcal{W} with density f for $(u, v) \in \mathcal{W}$ is defined by

$$\lambda((u, v)|\mathbf{x}) = \frac{f(\mathbf{x} \cup (u, v))}{f(\mathbf{x} \setminus (u, v))}, \quad (2.2)$$

with $a/0 := 0$ for all $a \geq 0$ (Cronie and van Lieshout, 2015).

The Papangelou conditional intensity is also very useful to describe the local interactions in a point pattern, and leads to the notion of a Markov point process which is base of implementation MCMC algorithms for simulation of Gibbs models. We say that the point process has "interactions of range R at (ξ, t) " if points further than R away from (ξ, t) do not contribute to the conditional intensity at (ξ, t) .

Definition 2.1.1. *A spatio-temporal Gibbs point process X has a finite interaction range R if the Papangelou conditional intensity satisfies*

$$\lambda((u, v)|\mathbf{x}) = \lambda((u, v)|\mathbf{x} \cap C_R^R(u, v)) \quad (2.3)$$

for all configurations \mathbf{x} of X and all $(u, v) \in \mathcal{W}$, where $C_R^R(u, v)$ denotes the cylinder of radius $R > 0$ and height $2R > 0$ centred at (u, v) .

Spatio-temporal Gibbs models usually have finite interaction range property (spatio-temporal Markov property) and are called in this case Markov point processes (van Lieshout 2000). The finite range property of a spatio-temporal Gibbs model implies that the probability to insert a point (u, v) into \mathbf{x} depends only on some cylindrical neighborhoods of (u, v) . We further refer to Dereudre (2019) for more formal introduction of Gibbs point processes.

Here, we first review spatio-temporal Gibbs models and then extend the spatial Geyer and Strauss hardcore models to the spatio-temporal single- and multi-scale context.

2.1.1 Spatio-temporal Gibbs models review

In the literature, several spatio-temporal Gibbs point process models have been proposed such as the hardcore (Cronie and van Lieshout 2015), Strauss (Gonzalez et al. 2016) and area-interaction (Iftimi et al. 2018) point processes.

A Gibbs point process model explicitly postulates that interactions traduce dependencies between the points of the pattern. The hardcore interaction is one of the simplest type of interactions, which forbids points being too close to each other. The homogeneous spatio-temporal hardcore point process is defined by the density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})} \mathbb{1}\{ \|\xi - \xi'\| > hc_s \text{ or } |t - t'| > hc_t; \forall (\xi, t) \neq (\xi', t') \in \mathbf{x} \}, \quad (2.4)$$

with respect to a unit rate Poisson point process on \mathcal{W} , where $c > 0$ is a normalizing constant, $\lambda > 0$ is an activity parameter, $hc_s, hc_t > 0$ are, respectively, the spatial and the temporal hardcore distances and $n(\mathbf{x})$ is the number of points in \mathbf{x} . The Papangelou conditional intensity of a homogeneous spatio-temporal hardcore point process for $(u, v) \notin \mathbf{x}$ is obtained

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda \mathbb{1}\{|\xi - u| > hc_s \text{ or } |t - v| > hc_t; \forall (\xi, t) \in \mathbf{x}\} \\ &= \lambda \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{|\xi - u| > hc_s \text{ or } |t - v| > hc_t\} \\ &= \lambda \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{(\xi, t) \notin C_{hc_s}^{hc_t}(u, v)\}. \end{aligned} \quad (2.5)$$

The homogeneous spatio-temporal Strauss point process is defined by density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})}\gamma^{S_r^q(\mathbf{x})}, \quad (2.6)$$

with respect to a unit rate Poisson point process on \mathcal{W} , where $S_r^q(\mathbf{x}) = \sum_{(\xi, t) \neq (\xi', t') \in \mathbf{x}} \mathbb{1}\{|\xi - \xi'| \leq r, |t - t'| \leq q\}$ and the Papangelou conditional intensity of the model for $(u, v) \notin \mathbf{x}$ is

$$\lambda((u, v)|\mathbf{x}) = \lambda\gamma^{n[C_r^q(u, v); \mathbf{x}]}, \quad (2.7)$$

and for $(\xi, t) \in \mathbf{x}$

$$\lambda((\xi, t)|\mathbf{x}) = \lambda\gamma^{n[C_r^q(\xi, t); \mathbf{x} \setminus (\xi, t)]}, \quad (2.8)$$

where $n[C_r^q(y, z); \mathbf{x}] = \sum_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{|\xi - y| \leq r, |t - z| \leq q\}$ is the number of points in \mathbf{x} which are in a cylinder $C_r^q(y, z)$. Although the Strauss point process was originally intended as a model of clustering, it can only be used to model inhibition, because the parameter γ cannot be greater than 1. If we take $\gamma > 1$, the density function of Strauss model is not integrable, so it does not define a valid probability density.

Iftimi et al. (2018) defined the homogeneous spatio-temporal area-interaction point process by density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})} \prod_{(\xi, t) \in \mathbf{x}} \gamma^{-\ell(\cup_{(\xi, t) \in \mathbf{x}} C_r^q(\xi, t))}, \quad (2.9)$$

with respect to a unit rate Poisson point process on \mathcal{W} , where ℓ is the Lebesgue measure restricted to \mathcal{W} . Iftimi et al. (2018) extended the hybridization approach for an inhomogeneous area-interaction model in spatio-temporal framework where the density is given by

$$f(\mathbf{x}) = c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{-\ell(\cup_{(\xi, t) \in \mathbf{x}} C_{r_j}^{q_j}(\xi, t))}, \quad (2.10)$$

with respect to a unit rate Poisson process on \mathcal{W} , where (r_j, q_j) are pairs of irregular parameters of the model and γ_j are interaction parameters, $j = 1, \dots, m$. New hybrid

Gibbs models can also be defined from the hardcore process (2.4) and the Strauss process (2.6) 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.

As mentioned, Strauss point process model only achieves the inhibition structure. In spatial framework, two ways are introduced to overcome this problem that we extend to spatio-temporal framework hence defining two new spatio-temporal Gibbs point process models.

2.1.2 Spatio-temporal Geyer saturation model

A first way to propose the Gibbs models based on Strauss model which intend for clustering structures is to consider an upper bound for the number of neighboring points that interact. Indeed, we extend the spatial Geyer saturation point process (1.6) to the spatio-temporal framework replacing the Euclidean balls by spatio-temporal cylindrical neighborhoods (Gonzalez et al. 2016).

Definition 2.1.2. *We define the spatio-temporal Geyer saturation point process as the point process with density*

$$f(\mathbf{x}) = c \prod_{(\xi,t) \in \mathbf{x}} \lambda(\xi, t) \gamma^{\min\{s, n(C_r^q(\xi,t); \mathbf{x})\}}, \quad (2.11)$$

with respect to a unit rate Poisson process on \mathcal{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 $n(C_r^q(\xi, t); \mathbf{x}) = \sum_{(u,v) \in \mathbf{x} \setminus (\xi,t)} \mathbb{1}(\|u - \xi\| \leq r, |v - t| \leq q)$ 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 (2.11).

Definition 2.1.3. *We define the spatio-temporal multi-scale Geyer saturation point process as the point process with density*

$$f(\mathbf{x}) = c \prod_{(\xi,t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(\xi,t); \mathbf{x})\}}, \quad (2.12)$$

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

For any $j \in \{1, \dots, m\}$, 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 $j \in \{1, \dots, m\}$, the density (2.12) corresponds to the density of an inhomogeneous Poisson process. Equation (2.12) 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 2.1.1. *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 (2.12) is measurable and integrable for all $\gamma_j, j = 1, \dots, m$ 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 (2.11) 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 (2.12) is also a Markov point process in Ripley-Kelly sense at interaction range $\max_{1 \leq j \leq m} \{2 \max\{r_j, q_j\}\} = 2 \max\{r_m, q_m\}$ (Baddeley et al. 2013). \square

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

$$\begin{aligned} \lambda((u, v) | \mathbf{x}) &= \lambda(u, v) \prod_{j=1}^m \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(u, v); \mathbf{x})\}} \\ &\times \prod_{(\xi, t) \in \mathbf{x} \setminus (u, v)} \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(\xi, t); \mathbf{x} \cup (u, v))\} - \min\{s_j, n(C_{r_j}^{q_j}(\xi, t); \mathbf{x} \setminus (u, v))\}}, \end{aligned} \quad (2.13)$$

The Markovian property (Lemma 2.1.1) 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 Chapter 4.

2.1.3 Spatio-temporal Strauss hardcore model

A second way to propose the Gibbs models based on Strauss model which intend for clustering structures is to introduce a hardcore condition to the Strauss density (2.6). Hence, we can define a Strauss hardcore model in the spatio-temporal context.

Definition 2.1.4. *We define the spatio-temporal Strauss hardcore point process as the*

point process with density

$$f(\mathbf{x}) = c \prod_{i=1}^n \lambda(\xi_i, t_i) \gamma^{S_r^q(\mathbf{x})} \mathbb{1}\{\|\xi - \xi'\| > hc_s \text{ or } |t - t'| > hc_t; \forall (\xi, t) \neq (\xi', t') \in \mathbf{x}\}, \quad (2.14)$$

where $0 < hc_s < r$ and $0 < hc_t < q$.

The model could be used to model clustering patterns with a softer attraction between the points like a pattern with a combination of interaction terms that show repulsion between the points at a small scale and attraction between the points at a larger scale. The Papangelou conditional intensity of a spatio-temporal Strauss hardcore point process for $(u, v) \notin \mathbf{x}$ is obtained

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda(u, v) \gamma^{n[C_r^q(u, v); \mathbf{x}]} \mathbb{1}\{\|\xi - u\| > hc_s \text{ or } |t - v| > hc_t; \forall (\xi, t) \in \mathbf{x}\} \\ &= \lambda(u, v) \gamma^{n[C_r^q(u, v); \mathbf{x}]} \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{(\xi, t) \notin C_{hc_s}^{hc_t}(u, v)\}. \end{aligned} \quad (2.15)$$

A hybrid version of spatio-temporal Strauss hardcore model can be defined by hybridization approach.

Definition 2.1.5. We define the spatio-temporal hybrid Strauss hardcore point process with density

$$\begin{aligned} f(\mathbf{x}) &= c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{S_{r_j}^{q_j}(\mathbf{x})} \\ &\times \mathbb{1}\{\|\xi' - \xi''\| > hc_s \text{ or } |t' - t''| > hc_t; \forall (\xi', t') \neq (\xi'', t'') \in \mathbf{x}\}, \end{aligned} \quad (2.16)$$

where $0 < hc_s < r_1 < \dots < r_m$ and $0 < hc_t < q_1 < \dots < q_m$.

In the same way, Papangelou conditional intensity of an inhomogeneous spatio-temporal hybrid Strauss hardcore process for $(u, v) \notin \mathbf{x}$ is obtained

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda(u, v) \prod_{j=1}^m \gamma_j^{n[C_{r_j}^{q_j}(u, v); \mathbf{x}]} \mathbb{1}\{\|\xi - u\| > hc_s \text{ or } |t - v| > hc_t; \forall (\xi, t) \in \mathbf{x}\} \\ &= \lambda(u, v) \prod_{j=1}^m \gamma_j^{n[C_{r_j}^{q_j}(u, v); \mathbf{x}]} \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{(\xi, t) \notin C_{hc_s}^{hc_t}(u, v)\}. \end{aligned} \quad (2.17)$$

The Papangelou conditional intensity of the Gibbs point process models including a hardcore interaction term takes the value zero at some locations. We can thus write that

for all parameters of the model

$$\lambda((u, v)|\mathbf{x}) = m((u, v)|\mathbf{x})\lambda^+((u, v)|\mathbf{x}), \quad (2.18)$$

where $m((u, v)|\mathbf{x})$ takes only the values 0 and 1, and $\lambda^+((u, v)|\mathbf{x}) > 0$ everywhere.

In the same way as Lemma 2.1.1, the spatio-temporal hybrid Strauss hardcore point process (2.16) is a Markov point process in Ripley and Kelly (1977) sense at interaction range $\max\{r_m, q_m\}$.

2.2 Models based on Cox and Gibbs processes

Gibbs models represent a flexible class of processes for setting direct interaction between points. The spatio-temporal heterogeneity in the expected number of points observed per unit of space and time can be captured by estimating a non constant trend term estimation of a Gibbs models. In the literature, this trend is typically considered as a function of the covariates, whose influence is expressed through a small number of parameters, for instance by estimating fixed effects in a generalized linear model as in Iftimi et al. (2018). In this section, we introduce the models derived from the multi-scale classes of combinations of Gibbs and log-Gaussian Cox point processes, to which we refer as Cox-Gibbs models in the following.

We consider the popular class of pairwise interaction point processes with density

$$f(\mathbf{x}) = c \prod_{i=1}^n \lambda(\xi_i, t_i) \prod_{i<j} \gamma((\xi_i, t_i), (\xi_j, t_j)), \quad (2.19)$$

with respect to a unit rate Poisson process on \mathcal{W} for all point patterns \mathbf{x} , where $c > 0$ is a normalizing constant, $\lambda : \mathcal{W} \rightarrow \mathbb{R}^+$ is a first-order interaction function which models systematic aggregation of points and $\gamma : \mathcal{W} \times \mathcal{W} \rightarrow \mathbb{R}^+$ is a second-order interaction function which models repulsion between the points with form $\gamma((\xi_i, t_i), (\xi_j, t_j)) = \gamma(\|\xi_i - \xi_j\|, |t_i - t_j|)$.

The simplest nontrivial pairwise interaction process is the Strauss process with $\gamma((\xi_i, t_i), (\xi_j, t_j)) = \gamma^{\mathbb{1}\{\|\xi_i - \xi_j\| \leq r, |t_i - t_j| \leq q\}}$ in (2.19) where r and q are spatial and temporal radii, respectively. Pairwise interaction processes are mainly models for a repulsive behaviour. However, a Strauss hardcore process with density (2.14) rewritten as

$$f(\mathbf{x}) = c \prod_{i=1}^n \lambda(\xi_i, t_i) \prod_{i<j} \gamma^{\mathbb{1}\{\|\xi_i - \xi_j\| \leq r, |t_i - t_j| \leq q\}} \mathbb{1}\{\|\xi - \xi'\| > hc_s \text{ or } |t - t'| > hc_t; \forall (\xi, t) \neq (\xi', t') \in \mathbf{x}\}, \quad (2.20)$$

where hc_s, hc_t are spatial and temporal hardcore distances, is a model for both repul-

sive and attractive behaviours. Interesting Gibbs point process models are usually introduced with infinite order of interaction such as Geyer saturation point process with density (2.11).

In point process literature, modeling the small-scale interactions is an important and difficult challenge. To overcome that, we consider firstly a spatio-temporal pairwise interaction point process X with density (2.19) and then consider a doubly stochastic construction by replacing λ with a random function Λ in order to introduce random aggregation to the model which is an extension of a Cox process (when $\gamma = 1$). When Λ is the random intensity function of a log Gaussian Cox process, we have a log Gaussian Cox pairwise interaction process. Specifically, we consider for $(u, v) \in \mathcal{W}$

$$\Lambda(u, v) = \exp(Z(u, v)), \quad (2.21)$$

where $\mathbf{Z} := \{Z(u, v)\}_{(u,v) \in \mathcal{W}}$ is a Gaussian random field (GRF). Indeed, we suggest a model for regularity at small-scale and aggregation at larger-scale. Due to different values for parameters of the model, we have some well-known special cases (e.g. homogeneous Poisson, log Gaussian Cox process, and pairwise interaction process) in spatio-temporal context.

The log Gaussian Cox pairwise interaction process has a density (with respect to the unit rate Poisson process) with the form

$$f(\mathbf{x}) = \mathbb{E} \left[\frac{1}{c(\mathbf{Z})} \prod_{i=1}^n \exp(Z(\xi_i, t_i)) \prod_{i < j} \gamma(\|\xi_i - \xi_j\|, |t_i - t_j|) \right], \quad (2.22)$$

where the expectation is with respect to the GRF \mathbf{Z} and $c(\mathbf{Z})$ is the normalising constant obtained by conditioning on \mathbf{Z} .

Hence, a log Gaussian Cox Strauss process (LGCSP) has density

$$f(\mathbf{x}) = \mathbb{E} \left[\frac{1}{c(\mathbf{Z})} \prod_{i=1}^n \exp(Z(\xi_i, t_i)) \prod_{i < j} \gamma^{\mathbb{1}\{\|\xi_i - \xi_j\| \leq r, |t_i - t_j| \leq q\}} \right]. \quad (2.23)$$

In the same way, we can introduce the log Gaussian Cox Strauss hardcore process (LGCSHP) with density

$$f(\mathbf{x}) = \mathbb{E} \left[\frac{1}{c(\mathbf{Z})} \prod_{i=1}^n \exp(Z(\xi_i, t_i)) \prod_{i < j} \gamma^{\mathbb{1}\{\|\xi_i - \xi_j\| \leq r, |t_i - t_j| \leq q\}} \right. \\ \left. \times \mathbb{1}\{\|\xi - \xi'\| > hc_s \text{ or } |t - t'| > hc_t; \forall (\xi, t) \neq (\xi', t') \in \mathbf{x}\} \right]. \quad (2.24)$$

We can also introduce the log Gaussian Cox Geyer process (LGCGP) with density

$$f(\mathbf{x}) = \mathbb{E} \left[\frac{1}{c(\mathbf{Z})} \prod_{i=1}^n \exp(Z(\xi_i, t_i)) \gamma^{\min\{s, n(C_r^q(\xi_i, t_i); \mathbf{x})\}} \right]. \quad (2.25)$$

Due to existence the GRF \mathbf{Z} in density of hybrid Gibbs-Cox models, the Papangelou conditional intensity can have a general form for $(u, v) \notin \mathbf{x}$

$$\lambda((u, v) | \mathbf{x}, \mathbf{Z}) = \Gamma((u, v) | \mathbf{x}) \zeta((u, v) | \mathbf{Z}), \quad (2.26)$$

where Γ is related to second-order interaction of the model which is free of complex integrals of GRF while

$$\zeta((u, v) | \mathbf{Z}) = \frac{\mathbb{E} [\exp(Z(u, v)) \prod_{i=1}^n \exp(Z(\xi_i, t_i))]}{\mathbb{E} [\prod_{i=1}^n \exp(Z(\xi_i, t_i))]}.$$

Hence, for a LGCSHP we have in (2.26)

$$\Gamma((u, v) | \mathbf{x}) = \gamma^{n[C_r^q(u, v); \mathbf{x}]},$$

where $n[C_r^q(u, v); \mathbf{x}] = \sum_i \mathbb{1}(|u - \xi_i| \leq r, |v - t_i| \leq q)$, for a LGCSHP we have

$$\Gamma((u, v) | \mathbf{x}) = \gamma^{n[C_r^q(u, v); \mathbf{x}]} \prod_i \mathbb{1}\{(\xi_i, t_i) \notin C_{hc_s}^{hc_t}(u, v)\},$$

and for a LGCGP we have

$$\Gamma((u, v) | \mathbf{x}) = \gamma^{\min\{s, n(C_r^q(u, v); \mathbf{x})\}} \prod_i \gamma^{\min\{s, n(C_r^q(\xi_i, t_i); \mathbf{x} \cup (u, v))\} - \min\{s, n(C_r^q(\xi_i, t_i); \mathbf{x})\}}.$$

Chapter 3

Inference

There are different ways in fitting point process models, basically: moment-, likelihood- and Bayesian-based methods. In general, the likelihood has no closed form expression (and thus is intractable) for most of density's models. To address this issue, a simple and quick inference procedure is using the composite likelihood-based inference which is mainly defined based on the Papangelou conditional intensity function. In this chapter, we focus on both global and local parameter estimation. For global estimation, we extend to the spatio-temporal context two composite likelihood-based inference methods for our new Gibbs models and design a Bayesian hierarchical approach for the Cox-Gibbs model. We then implement a local parameter estimation approach to take into account different local interaction structures for spatio-temporal Gibbs models.

3.1 Global parameter estimation

Gibbs point process models involve 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. Typically, regular parameters determine the 'strength' of the interaction, while irregular parameters determine the 'range' of the interaction.

Irregular parameters, like saturation threshold s and distances r and q in Geyer model (2.11) and also hc_s and hc_t in Strauss hardcore model (2.14), 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 the spatio-temporal framework, 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.

However, the hardcore interaction term $m(\cdot|\mathbf{x})$ in the conditional intensity (2.18) does not depend on the other parameters of the densities of Gibbs point processes. This implies that it can first be estimated and kept fixed for the sequel (Baddeley et al. 2019). In the spatial framework, the maximum likelihood estimate of the hardcore distance in $m(\cdot|\mathbf{x})$ corresponds to the minimum interpoint distance (Baddeley et al. 2013). The generalization to the spatio-temporal context with a cylindrical hardcore structure implies to consider a multi-objective minimization problem over the spatial and temporal hardcore distances hc_s and hc_t . The choice of our hardcore parameters needs to analyze the Pareto front of feasible solutions on the graph of spatial and temporal interpoint distances. We refer to Ehrgott (2005) for a description of multi-criteria optimization and the definition of Pareto optimality. To estimate the hardcore distance hc_s and hc_t , we consider a feasible solution on the Pareto front as large as possible and with a ratio consistent with our knowledge of interaction mechanisms in practice.

Regular parameters like trend λ and interaction γ in (2.11) and (2.14) 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 models and we compare their performance in Chapter 4.

3.1.1 Composite likelihoods

The likelihoods of Gibbs models are intractable; when a surrogate likelihood is required, the choice is usually a composite likelihood (Lindsay 1988, Varin et al. 2011) of which there are several kinds adapted to different classes of models. See Baddeley et al. (2019) for composite likelihood-based statistical inference in Gibbs point processes. We here implement two composite likelihoods; pseudo-likelihood and logistic likelihood for Geyer and Strauss hardcore models.

3.1.1.1 Pseudo-likelihood approach

Let θ 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 $\lambda_{\theta}((u, v)|\mathbf{x})$ over \mathcal{W} as follows

$$PL(\mathbf{x}; \boldsymbol{\theta}) = \exp\left(-\int_S \int_T \lambda_{\theta}((u, v)|\mathbf{x}) dv du\right) \prod_{(\xi, t) \in \mathbf{x}} \lambda_{\theta}((\xi, t)|\mathbf{x}). \quad (3.1)$$

The pseudo score is defined by

$$U(\mathbf{x}; \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log PL(\mathbf{x}; \boldsymbol{\theta}), \quad (3.2)$$

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

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log PL(\mathbf{x}; \boldsymbol{\theta}) = 0, \quad (3.3)$$

where

$$\log PL(\mathbf{x}; \boldsymbol{\theta}) = \sum_{(\xi, t) \in \mathbf{x}} \log \lambda_{\theta}((\xi, t)|\mathbf{x}) - \int_S \int_T \lambda_{\theta}((u, v)|\mathbf{x}) dv du, \quad (3.4)$$

and $\lambda_{\theta}(\cdot|\mathbf{x})$ is defined by (2.13) for hybrid Geyer model (2.12).

For sake of clarity, we now assume that $\boldsymbol{\theta} = [\log \gamma_1, \dots, \log \gamma_m]^\top$ the logarithm of interaction parameters in model (2.12). To estimate $\boldsymbol{\theta}$, we use the pseudo-likelihood approach. Equation (2.13) can be rewritten as $\lambda_{\theta}((u, v)|\mathbf{x}) = \lambda(u, v) \prod_{j=1}^m \exp(\theta_j S_j((u, v), \mathbf{x}))$ where

$$\begin{aligned} S_j((u, v), \mathbf{x}) &= \min\{s_j, n(C_{r_j}^{q_j}(u, v); \mathbf{x})\} \\ &+ \sum_{(\xi, t) \in \mathbf{x} \setminus (u, v)} [\min\{s_j, n(C_{r_j}^{q_j}(\xi, t); \mathbf{x} \cup (u, v))\} \\ &- \min\{s_j, n(C_{r_j}^{q_j}(\xi, t); \mathbf{x} \setminus (u, v))\}], \end{aligned} \quad (3.5)$$

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

$$\log \lambda_{\theta}((u, v)|\mathbf{x}) = \log \lambda(u, v) + \boldsymbol{\theta}^\top \mathbf{S}((u, v), \mathbf{x}) \quad (3.6)$$

is a linear model in $\boldsymbol{\theta}$ with offset $\log \lambda(u, v)$. Thus, equation (3.3) gives us the pseudo-likelihood equations

$$\frac{\partial}{\partial \boldsymbol{\theta}} \left[\sum_{(\xi, t) \in \mathbf{x}} [\log \lambda(\xi, t) + \sum_{j=1}^m \theta_j S_j((\xi, t), \mathbf{x})] - \int_S \int_T \lambda(u, v) \prod_{j=1}^m e^{\theta_j S_j((u, v), \mathbf{x})} dv du \right] = 0, \quad (3.7)$$

For each parameter $\theta_i, i = 1, \dots, m$, the equations (3.7) can be rewritten

$$\sum_{(\xi, t) \in \mathbf{x}} S_i((\xi, t), \mathbf{x}) = \int_S \int_T \lambda(u, v) S_i((u, v), \mathbf{x}) \prod_{j=1}^m e^{\theta_j S_j((u, v), \mathbf{x})} dv du, \quad (3.8)$$

The major difficulty is to estimate the integrals on the right hand side of equations (3.8). 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 (3.8). 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 (3.4) by a Riemann sum

$$\int_S \int_T \lambda_{\theta}((u, v) | \mathbf{x}) dv du \approx \sum_{k=1}^{n+p} w_k \lambda_{\theta}((\xi_k, t_k) | \mathbf{x}), \quad (3.9)$$

where (ξ_k, t_k) are points in $\{(\xi_1, t_1), \dots, (\xi_n, t_n), (\xi_{n+1}, t_{n+1}), \dots, (\xi_{n+p}, t_{n+p})\} \in \mathcal{W}$ 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

$$\log PL(\mathbf{x}; \boldsymbol{\theta}) \approx \sum_{(\xi, t) \in \mathbf{x}} \log \lambda_{\theta}((\xi, t) | \mathbf{x}) - \sum_{k=1}^{n+p} w_k \lambda_{\theta}((\xi_k, t_k) | \mathbf{x}). \quad (3.10)$$

Note that if the set of points $\{(\xi_k, t_k), k = 1, \dots, n + p\}$ includes all the points of $\mathbf{x} = \{(\xi_1, t_1), \dots, (\xi_n, t_n)\}$, we can rewrite (3.10) as

$$\log PL(\mathbf{x}; \boldsymbol{\theta}) \approx \sum_{k=1}^{n+p} w_k (y_k \log \lambda_{\theta}((\xi_k, t_k) | \mathbf{x}) - \lambda_{\theta}((\xi_k, t_k) | \mathbf{x})), \quad (3.11)$$

where

$$y_k = \begin{cases} 1/w_k, & \text{if } (\xi_k, t_k) \in \mathbf{x} \text{ is an event,} \\ 0, & \text{if } (\xi_k, t_k) \notin \mathbf{x} \text{ is a dummy point.} \end{cases} \quad (3.12)$$

The right hand side of (3.11), for fixed \mathbf{x} , is formally equivalent to the log-likelihood of independent Poisson variables $Y_k \sim \text{Poisson}(\lambda_{\theta}((\xi_k, t_k) | \mathbf{x}))$ 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 (3.11).

Note that in hybrid Geyer model (2.12), 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 \mathcal{W} and merge them with all the data points in \mathbf{x} to construct the set of quadrature points $(\xi_k, t_k) \in \mathcal{W}$ with $k = 1, \dots, n + p$.
- Compute the quadrature weights w_k and the indicators y_k defined in (3.12),
- Compute the sufficient statistics $\mathcal{S}((\xi_k, t_k), \mathbf{x})$ at each quadrature point,
- Fit a log-linear Poisson regression with explanatory variables $\mathcal{S}((\xi_k, t_k), \mathbf{x})$, 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 \mathcal{S} -vector and intercept $\hat{\theta}_0$,
- Return the maximum pseudo-likelihood-based parameter estimates $\hat{\gamma}_j = \exp(\hat{\theta}_j)$ for $j = 1, \dots, m$ and $\hat{\beta} = \exp(\hat{\theta}_0)$.

We define the quadrature scheme by defining a spatio-temporal partition of \mathcal{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 $\sum_k w_k = \ell(\mathcal{W})$, 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. 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).

3.1.1.2 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 \mathcal{W} having a density f_{θ} with respect to the unit rate Poisson process and with conditional intensity function $\lambda_{\theta}(\cdot|\mathbf{x})$. We consider an independent Poisson process for dummy points, with intensity function ρ , and we denote by \mathbf{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 $(\xi, t) \in \mathbf{x} \cup \mathbf{d}$, we obtain independent Bernoulli variables taking one for data points and zero for dummy points. We have

$$Pr(Y(\xi, t) = 1) = \frac{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t))}{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t)) + \rho(\xi, t)}, \quad (3.13)$$

By considering the log linearity assumption for the conditional intensity $\lambda_{\theta}(\cdot|\mathbf{x})$ in (3.6), the logit of $Pr(Y(\xi, t) = 1)$ is

$$\log \frac{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t))}{\rho(\xi, t)} = \log \frac{\lambda(\xi, t)}{\rho(\xi, t)} + \sum_{j=1}^m \theta_j S_j((\xi, t), \mathbf{x} \setminus (\xi, t)), \quad (3.14)$$

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

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

$$\begin{aligned} \log LL(\mathbf{x}, \mathbf{d}; \theta) &= \sum_{(\xi, t) \in \mathbf{x} \cup \mathbf{d}} Y(\xi, t) \log \frac{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t))}{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t)) + \rho(\xi, t)} \\ &\quad + \sum_{(\xi, t) \in \mathbf{x} \cup \mathbf{d}} [1 - Y(\xi, t)] \log \frac{\rho(\xi, t)}{\lambda_{\theta}((\xi, t)|\mathbf{x}) + \rho(\xi, t)} \\ &= \sum_{(\xi, t) \in \mathbf{x}} \log \frac{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t))}{\lambda_{\theta}((\xi, t)|\mathbf{x} \setminus (\xi, t)) + \rho(\xi, t)} \\ &\quad + \sum_{(\xi, t) \in \mathbf{d}} \log \frac{\rho(\xi, t)}{\lambda_{\theta}((\xi, t)|\mathbf{x}) + \rho(\xi, t)}. \end{aligned} \quad (3.15)$$

The maximum of the log-logistic likelihood exists and under regularity condition (Bad-

deley et al. 2019) is unique. Hence, estimation can be implemented in \mathbb{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 from the following algorithm.

Algorithm 2

- Generate dummy points \mathbf{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 $(\xi_k, t_k) \in \mathcal{W}$,
- Obtain the response variables y_k (1 for data points, 0 for dummy points),
- Compute the sufficient statistics $\mathbf{S}((\xi_k, t_k), \mathbf{x} \setminus (\xi_k, t_k))$ at each quadrature point,
- Fit a logistic regression model with explanatory variables $\mathbf{S}((\xi_k, t_k), \mathbf{x} \setminus (\xi_k, t_k))$, 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 \mathbf{S} -vector and intercept $\hat{\theta}_0$,
- Return the parameter estimator $\hat{\gamma} = \exp(\hat{\boldsymbol{\theta}})$ 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 $\hat{\beta} = c^{-1} \exp(\hat{\theta}_0)$.

In the same way, we assume that $\boldsymbol{\theta} = (\log \gamma_1, \log \gamma_2, \dots, \log \gamma_m)$ is the logarithm of interaction parameters in spatio-temporal hybrid Strauss hardcore point process (2.16). The Papangelou conditional intensity of the spatio-temporal hybrid Strauss hardcore process for $(u, v) \in \mathcal{W}$ is

$$\lambda((u, v) | \mathbf{x}) = \lambda(u, v) \prod_{j=1}^m \gamma_j^{n[C_{r_j}^{q_j}(u, v); \mathbf{x} \setminus (u, v)]} \mathbb{1}\{ \|\xi - u\| > hc_s \text{ or } |t - v| > hc_t; \forall (\xi, t) \in \mathbf{x} \setminus (u, v) \}. \quad (3.16)$$

To estimate $\boldsymbol{\theta}$, due to (2.18), we just consider the points (u, v) where $m((u, v) | \mathbf{x})$ is equal to 1 in (3.16). By defining $S_j((u, v), \mathbf{x}) := n[C_{r_j}^{q_j}(u, v); \mathbf{x} \setminus (u, v)]$ in (3.16), we can thus write $\lambda_{\boldsymbol{\theta}}((u, v) | \mathbf{x}) = \lambda(u, v) \prod_{j=1}^m \exp(\theta_j S_j((u, v), \mathbf{x}))$. Hence, the logarithm of the Papangelou conditional intensity of the spatio-temporal hybrid Strauss hardcore point process for $(u, v) \in \mathcal{W}$ which satisfies in hardcore condition, i.e. $m((u, v) | \mathbf{x}) = 1$ in (3.16), is

$$\begin{aligned} \log \lambda((u, v) | \mathbf{x}) &= \log \lambda(u, v) + \sum_{j=1}^m (\log \gamma_j) S_j((u, v), \mathbf{x}) \\ &= \log \lambda(u, v) + \boldsymbol{\theta}^\top \mathbf{S}((u, v), \mathbf{x}) \end{aligned} \quad (3.17)$$

corresponding to a linear model in θ with offset $\log \lambda(u, v)$ where

$$\mathbf{S}((u, v), \mathbf{x}) = [S_1((u, v), \mathbf{x}), S_2((u, v), \mathbf{x}), \dots, S_m((u, v), \mathbf{x})]^\top$$

is a sufficient statistics.

The logit for the models is

$$\log \frac{\lambda_\theta((\xi, t)|\mathbf{x})}{\rho(\xi, t)} = \log \frac{\lambda(\xi, t)}{\rho(\xi, t)} + \sum_{j=1}^m \theta_j S_j((\xi, t), \mathbf{x}), \quad (3.18)$$

which is a linear model in θ with offset $\log \frac{\lambda(\xi, t)}{\rho(\xi, t)}$. We finally implement *Algorithm 2* for quadrature points (data and dummy points) such that $m(\cdot|\mathbf{x}) = 1$.

3.1.2 Bayesian approach

The calculation of composite likelihoods: pseudo-likelihood and logistic likelihood involves complex, high-dimensional integrals for Cox-Gibbs models, which further need estimation methods to handle the latent (i.e., unobserved) Gaussian variables. The hierarchical structure of the Cox-Gibbs model ensures a Bayesian formulation, and inference can be achieved by using the Integrated Nested Laplace Approximation (Rue et al. 2009). In Gabriel et al. (2017), Serra et al. (2014b), Opitz et al. (2020) and Pimont et al. (2021), INLA-based estimation has been implemented for spatio-temporal log-Gaussian Cox process models (i.e., without Gibbs interactions) for wildfire ignitions. We here extend INLA-based inference to spatio-temporal Cox-Gibbs processes. It is coupled with the stochastic partial differential equations approach of Lindgren et al. (2011) for numerically convenient Gauss-Markov representations of spatial Matérn covariances, Alternatively, one could consider Markov chain Monte Carlo (MCMC) estimation for LGCPs (Taylor et al. 2015), whose use could be extended to the Cox-Gibbs models, but we do not pursue this approach here. The INLA method calculates the integrals by a set of carefully chosen deterministic approximations related to the classical Laplace approximation (Tierney and Kadane 1986). INLA is generally faster compared to MCMC methods when considering comparable approximation accuracy, and INLA works well even with very sophisticated hierarchical structures combining several Gaussian random effects. A comparison in the spatial LGCP setting (but without estimating hyperparameters such as the range and variance of the Gaussian) was conducted by Taylor and Diggle (2014).

3.1.2.1 Bayesian estimation with INLA

Rue et al. (2009) introduced INLA as an estimation method for generalized mixed additive regression models with multivariate Gaussian random effects, also called latent

Gaussian models. Opitz (2017) reviews INLA method for spatio-temporal applications and presents the approach as follows. The distribution of observation y_i of the response depends on a predictor η_i (with Gaussian prior distribution) and potentially on other parameters, so-called hyperparameters collected into a vector $\boldsymbol{\theta}$. We adopt standard Bayesian notation by writing $\pi(\cdot)$ for densities and conditional densities. Specifically, conditional to η_i and $\boldsymbol{\theta}$, we write the probability density of y_i as $\pi(y_i|\eta_i, \boldsymbol{\theta})$. Most often, the mean of y_i is related to η_i through a link function h^{-1} such that $\mathbb{E}(y_i|\eta_i, \boldsymbol{\theta}) = h(\eta_i)$, $i = 1, \dots, n$, with $h : \mathbb{R} \rightarrow \mathbb{R}$. The hierarchical model is as structured follows:

$$\begin{aligned} \boldsymbol{\theta} &\sim \pi(\boldsymbol{\theta}), && \text{hyperparameters,} \\ \mathbf{z}|\boldsymbol{\theta} &\sim \mathcal{N}_m(\mathbf{0}, \mathbf{Q}_\theta^{-1}), && \text{latent Gaussian process} \\ \mathbf{y} | \mathbf{z}, \boldsymbol{\theta} &\overset{\text{ind.}}{\sim} \prod_i \pi(y_i|\eta_i(\mathbf{z}), \boldsymbol{\theta}), && \text{observations} \end{aligned}$$

where \mathbf{Q}_θ is the precision matrix of the latent Gaussian vector \mathbf{z} . The predictor vector $\boldsymbol{\eta}$ is additively composed of components \mathbf{z} , that is, $\boldsymbol{\eta}(\mathbf{z}) = \mathbf{A}\mathbf{z}$, with the fixed observation matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ that maps latent variables \mathbf{z} to predictors $\eta_i = \eta_i(\mathbf{z})$ associated to observations y_i . The resulting joint posterior density of \mathbf{z} and $\boldsymbol{\theta}$ given \mathbf{y} is

$$\pi(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y}) \propto \exp\left(-\frac{1}{2}\mathbf{z}'\mathbf{Q}_\theta\mathbf{z} + \sum_i \log(\pi(y_i|\eta_i, \boldsymbol{\theta})) + \log \pi(\boldsymbol{\theta})\right),$$

in general, and in our case, this density does not correspond to some known and easily tractable multivariate distribution family.

3.1.2.2 Latent Gauss–Markov fields and Laplace approximation

We say that a random vector $\mathbf{z}|\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_\theta^{-1})$ is a Gauss–Markov random field if \mathbf{Q} is a sparse matrix, i.e. the number of non-null entries of its $n \times n$ covariance matrix $\mathbf{Q} = (q_{ij})_{1 \leq i, j \leq n}$ is $\mathcal{O}(n)$.

The core of the INLA approach is Laplace approximation which starts with the computation the integral $\int f(\mathbf{z})d\mathbf{z}$. This integral can be approximated by using the fact that a multivariate Gaussian density integrates to 1 as follows

$$\int f(\mathbf{z})d\mathbf{z} = \int \exp(kg(\mathbf{z}))d\mathbf{z} \approx \left(\frac{2\pi}{k}\right)^{d/2} |\mathbf{H}(g)(\mathbf{z}^*)|^{-1/2} \exp(kg(\mathbf{z}^*)),$$

where \mathbf{z}^* is the unique global maximum of g and $\mathbf{H}(g)(\mathbf{z}^*)$ is the Hessian matrix.

In the context of INLA, Let $k = 1$ and $f(\mathbf{z}) = \exp(g(\mathbf{z})) = \pi(\mathbf{z}, \boldsymbol{\theta})$. We have

$$\int \pi(\mathbf{z}, \boldsymbol{\theta}) d\mathbf{z} = \frac{1}{\pi_G(\mathbf{z}^*|\boldsymbol{\theta})} \exp(\pi(\mathbf{z}^*, \boldsymbol{\theta})),$$

where π_G is a Gaussian approximation with mean vector \mathbf{z}^* to the conditional density of $\mathbf{z}|\boldsymbol{\theta}$. Hence, to calculate the posterior marginal densities of hyperparameters

$$\pi(\theta_j|\mathbf{y}) = \int \int \pi(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y}) d\mathbf{z} d\boldsymbol{\theta}_{-j} = \int \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta}_{-j},$$

we use the Laplace approximation for $\int \pi(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y}) d\mathbf{z} = \pi(\boldsymbol{\theta}|\mathbf{y})$ such that the approximated density $\tilde{\pi}$ satisfies

$$\tilde{\pi}(\boldsymbol{\theta}|\mathbf{y}) \propto \frac{\pi(\mathbf{z}, \boldsymbol{\theta}, \mathbf{y})}{\pi_G(\mathbf{z}|\boldsymbol{\theta}, \mathbf{y})} \Big|_{\mathbf{z}=\mathbf{z}^*(\boldsymbol{\theta})}$$

where $\mathbf{z}^*(\boldsymbol{\theta})$ is the mode of the joint density $\pi(\mathbf{z}, \boldsymbol{\theta}, \mathbf{y})$ for fixed $(\boldsymbol{\theta}, \mathbf{y})$. Thus an approximation of the posterior marginal of θ_j is

$$\tilde{\pi}(\theta_j|\mathbf{y}) = \sum_{l=1}^L w_l \tilde{\pi}(\boldsymbol{\theta}_l|\mathbf{y}),$$

which is a numerical integration with a set of integration nodes $\boldsymbol{\theta}_l$ chosen from a numerical exploration of the surface of the density $\tilde{\pi}(\boldsymbol{\theta}_{-j}, \theta_j|\mathbf{y})$ with θ_j held fixed and weights w_l . To calculate the posterior marginal densities of the latent Gaussian field

$$\pi(z_i|\mathbf{y}) = \int \int \pi(\mathbf{z}, \boldsymbol{\theta}|\mathbf{y}) d\mathbf{z}_{-i} d\boldsymbol{\theta} = \int \pi(z_i|\boldsymbol{\theta}, \mathbf{y}) \pi(\boldsymbol{\theta}|\mathbf{y}) d\boldsymbol{\theta},$$

that just is enough to approximately evaluate $\pi(z_i|\boldsymbol{\theta}, \mathbf{y})$. Hence, an approximation of the posterior marginal of z_i is

$$\tilde{\pi}(z_i|\mathbf{y}) = \sum_{k=1}^K w_k \tilde{\pi}(z_i|\boldsymbol{\theta}_k, \mathbf{y}) \tilde{\pi}(\boldsymbol{\theta}_k|\mathbf{y}),$$

where

$$\tilde{\pi}(z_i|\boldsymbol{\theta}_k, \mathbf{y}) \propto \frac{\pi(\mathbf{z}|\boldsymbol{\theta}_k, \mathbf{y})}{\pi(\mathbf{z}_{-i}|z_i, \boldsymbol{\theta}_k, \mathbf{y})}$$

is a numerical integration with a set of integration nodes $\boldsymbol{\theta}_k$ and weights w_k .

3.1.2.3 Penalized complexity priors

Simpson et al. (2017) have proposed a principled, intuitive approach of choosing prior distributions for important hyperparameters of Bayesian models, such as the range and

variance parameters of the SPDE-based GMRF-representation of Gaussian fields with Matérn covariance.

3.2 Local parameter estimation

An alternative approach for modeling multi-scale point patterns relies on local likelihood approach (Loader 1999, Baddeley 2017) to obtain spatially-varying estimates of the parameters of a point process model. We here provide an extension to the spatio-temporal framework to explore local changes in the interaction structures, that can not be retrieved from models with global parameters.

The local likelihood at each spatio-temporal point $(\xi, t) \in \mathcal{W}$ is the likelihood of the restriction of point process to a cylinder centred at (ξ, t) . It may be investigated in pseudo-likelihood's or logistic likelihood's model fitting of spatio-temporal Gibbs point processes (Besag 1977, Lindsay 1988, Baddeley et al. 2014). Local pseudo-likelihood for Gibbs models was defined independently by Zhuang (2015) and Baddeley et al. (2015). Baddeley (2017) developed a general approach, the local composite likelihood. In what follows, we develop local versions of two composite likelihoods: pseudo-likelihood and logistic likelihood for our spatio-temporal Gibbs models (Section 3.1.1). In each case, the logarithm of the composite likelihood is a stochastic integral over the spatio-temporal domain, similar to the Poisson log-likelihood. We simply introduce a local weighting kernel into this stochastic integral, giving a local composite log-likelihood.

3.2.1 Local pseudo-likelihood

Zhuang (2015) extended the local likelihood to spatio-temporal point processes so that to each point (u^*, v^*) is assigned a spatial weight $W_{\sigma_s}(u^*)$ that depends on its relative spatial location u^* as follows

$$\log LL(u^*; \boldsymbol{\theta}) = \sum_{i=1}^n W_{\sigma_s}(u^* - \xi_i) \log \lambda_{\boldsymbol{\theta}}(\xi_i, t_i) - \int_S \int_T W_{\sigma_s}(u^* - u) \lambda_{\boldsymbol{\theta}}(u, v) dv du, \quad (3.19)$$

where W_{σ_s} is a weight function (usually a kernel function) and $\lambda_{\boldsymbol{\theta}}$ is intensity function. The local likelihood can be also used to estimate the spatio-temporal variation of the parameters of the model.

We extend the local likelihood approach to spatio-temporal Gibbs point process models by using pseudo-likelihood (Besag 1977) at each desired spatio-temporal point

$(u^*, v^*) \in \mathcal{W}$ by

$$\begin{aligned} \log LPL((u^*, v^*); \boldsymbol{\theta}) &= \sum_{i=1}^n W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \lambda_{\boldsymbol{\theta}}((\xi_i, t_i) | \mathbf{x}) \\ &\quad - \int_S \int_T W_{\sigma_s}(u^* - u) W_{\sigma_t}(v^* - v) \lambda_{\boldsymbol{\theta}}((u, v) | \mathbf{x}) dv du, \end{aligned} \quad (3.20)$$

where $W_{\sigma_s}(u) = \sigma_s^{-2} W(u/\sigma_s)$ and $W_{\sigma_t}(v) = \sigma_t^{-1} W(v/\sigma_t)$ are weight functions (usually a kernel function), $\sigma_s, \sigma_t > 0$ are the smoothing bandwidths, and $\lambda_{\boldsymbol{\theta}}(\cdot | \mathbf{x})$ is conditional intensity function of Gibbs model. By maximizing $\log LPL((u^*, v^*); \boldsymbol{\theta})$ and varying (u^*, v^*) , we can determine how parameter $\boldsymbol{\theta}$ changes with spatio-temporal points. Local pseudo-likelihood (3.20) can be maximised in the same way as pseudo-likelihood approach proposed in Section 3.1.1.1.

The local pseudo-likelihood (3.20) can be approximated by

$$\begin{aligned} \log LPL((u^*, v^*); \boldsymbol{\theta}) &\approx \sum_{i=1}^n W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \lambda_{\boldsymbol{\theta}}((\xi_i, t_i) | \mathbf{x}) \\ &\quad - \sum_{k=1}^{n+p} W_{\sigma_s}(u^* - \xi_k) W_{\sigma_t}(v^* - t_k) \lambda_{\boldsymbol{\theta}}((\xi_k, t_k) | \mathbf{x}) w_k, \end{aligned} \quad (3.21)$$

where (ξ_k, t_k) are points in $\{(\xi_1, t_1), \dots, (\xi_n, t_n), (\xi_{n+1}, t_{n+1}), \dots, (\xi_{n+p}, t_{n+p})\} \in \mathcal{W} = S \times T$ and w_k are quadrature weights that $\sum_{k=1}^{n+p} w_k = \ell(S \times T)$. Hence, we can rewrite (3.21) as

$$\log LPL((u^*, v^*); \boldsymbol{\theta}) \approx \sum_{k=1}^{n+p} (y_k \log \lambda_k - \lambda_k) w_k W_{\sigma_s}(u^* - \xi_k) W_{\sigma_t}(v^* - t_k), \quad (3.22)$$

where $\lambda_k = \lambda_{\boldsymbol{\theta}}((\xi_k, t_k) | \mathbf{x})$ and

$$y_k = \begin{cases} 1/w_k & \text{if } (\xi_k, t_k) \in \mathbf{x} \\ 0 & \text{if } (\xi_k, t_k) \notin \mathbf{x} \end{cases} \quad (3.23)$$

The right hand side of (3.22), for fixed \mathbf{x} , is formally equivalent to the log-likelihood of independent Poisson variables $Y_k \sim \text{Poisson}(\lambda_k)$ taken with weights $w_k \times W_{\sigma_s}(u^* - \xi_k) \times W_{\sigma_t}(v^* - t_k)$. Therefore (3.22) can be maximized using standard software for fitting GLMs, such as that in R (R Core Team, 2016).

Due to the advantage of logistic likelihood over pseudo-likelihood for spatio-temporal Gibbs point processes (see Iftimi et al. (2018) and Section 4.3.1), we aim to develop the local logistic likelihood for spatio-temporal Gibbs point processes as follows.

3.2.2 Local logistic likelihood

We assume that X is a spatio-temporal point process on $\mathcal{W} = S \times T$ whose distribution is given by a density $f_{\boldsymbol{\theta}}$ with respect to a unit rate Poisson process on \mathcal{W} and let $\mathbf{x} = \{(\xi_1, t_1), \dots, (\xi_n, t_n)\}$ be a realisation of X with conditional intensity function of a loglinear form, i.e. $\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) = \lambda(u, v) \exp(\boldsymbol{\theta}^\top \mathbf{S}((u, v), \mathbf{x}))$ for $(u, v) \in \mathcal{W}$. The local logistic log likelihood at point (u^*, v^*) can be defined by

$$\begin{aligned} \log LLL((u^*, v^*); \boldsymbol{\theta}) &= \sum_{(\xi_i, t_i) \in \mathbf{x}} W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \frac{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x})}{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \\ &\quad - \int_S \int_T W_{\sigma_s}(u^* - u) W_{\sigma_t}(v^* - v) \rho(u, v) \log \frac{\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) + \rho(u, v)}{\rho(u, v)} dv du, \end{aligned} \quad (3.24)$$

where $W_{\sigma_s}, W_{\sigma_t}$ are weight functions and ρ is a nonnegative real-valued function. By Campbell theorem, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{(\xi_i, t_i) \in \mathbf{d}} W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \frac{\rho(\xi_i, t_i)}{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \right] &= \\ \int_S \int_T W_{\sigma_s}(u^* - u) W_{\sigma_t}(v^* - v) \rho(u, v) \log \frac{\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) + \rho(u, v)}{\rho(u, v)} dv du, \end{aligned} \quad (3.25)$$

Hence, we can approximate (3.24) by

$$\begin{aligned} \log LLL((u^*, v^*); \boldsymbol{\theta}) &\approx \sum_{(\xi_i, t_i) \in \mathbf{x}} W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \frac{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x})}{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \\ &\quad + \sum_{(\xi_i, t_i) \in \mathbf{d}} W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \log \frac{\rho(\xi_i, t_i)}{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \end{aligned} \quad (3.26)$$

where $\mathbf{d} = \{(\xi_{n+1}, t_{n+1}), \dots, (\xi_{n+p}, t_{n+p})\}$ is a random pattern of p dummy points that is a realisation of a Poisson point process D independent of X with intensity function ρ . The local logistic score at point (u^*, v^*) is

$$\begin{aligned} U((u^*, v^*); \boldsymbol{\theta}) &= \frac{\partial}{\partial \boldsymbol{\theta}} \log LLL((u^*, v^*); \boldsymbol{\theta}) \\ &= \sum_{(\xi_i, t_i) \in \mathbf{x}} W_{\sigma_s}(u^* - \xi_i) W_{\sigma_t}(v^* - t_i) \frac{\rho(\xi_i, t_i) \zeta_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x})}{\lambda_{\boldsymbol{\theta}}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \\ &\quad - \int_S \int_T W_{\sigma_s}(u^* - u) W_{\sigma_t}(v^* - v) \frac{\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) \rho(u, v) \zeta_{\boldsymbol{\theta}}((u, v)|\mathbf{x})}{\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) + \rho(u, v)} dv du \end{aligned} \quad (3.27)$$

where ζ_{θ} is the first derivative of the log conditional intensity, $\log \lambda_{\theta}(\cdot|\mathbf{x})$, with respect to θ and local parameter estimates $\hat{\theta}(u^*, v^*)$ is a zero of the local score (3.27).

By considering the log linearity assumption for conditional intensity $\lambda_{\theta}(\cdot|\mathbf{x})$, the negative Hessian matrix is thus obtained

$$\begin{aligned}
H((u^*, v^*); \theta) &= -\frac{\partial}{\partial \theta^{\top}} U((u^*, v^*); \theta) = -\frac{\partial^2}{\partial \theta \partial \theta^{\top}} \log LLL((u^*, v^*); \theta) \\
&= -\sum_{(\xi, t) \in \mathbf{x}} W_{\sigma_s}(u^* - \xi) W_{\sigma_t}(v^* - t) \zeta_{\theta}((\xi, t)|\mathbf{x}) \zeta_{\theta}((\xi, t)|\mathbf{x})^{\top} \\
&\quad \times \frac{\rho(\xi, t) \mathbf{S}((\xi, t), \mathbf{x})}{(\mathbf{S}((\xi, t), \mathbf{x}) + \rho(\xi, t))^2} \\
&+ \int_S \int_T W_{\sigma_s}(u - u^*) W_{\sigma_t}(v - v^*) \zeta_{\theta}((u, v)|\mathbf{x}) \zeta_{\theta}((u, v)|\mathbf{x})^{\top} \\
&\quad \times \frac{\rho(u, v)^2 \mathbf{S}((u, v), \mathbf{x})}{(\mathbf{S}((u, v), \mathbf{x}) + \rho(u, v))^2} dv du.
\end{aligned} \tag{3.28}$$

We can also rewrite (3.26) by

$$\begin{aligned}
\log LLL((u^*, v^*); \theta) &= \sum_{(\xi, t) \in \mathbf{x} \cup \mathbf{d}} W_{\sigma_s}(\xi - u^*) W_{\sigma_t}(t - v^*) Y(\xi, t) \log p_{\theta}((\xi, t)|\mathbf{x}) \\
&+ \sum_{(\xi, t) \in \mathbf{x} \cup \mathbf{d}} W_{\sigma_s}(\xi - u^*) W_{\sigma_t}(t - v^*) [1 - Y(\xi, t)] \log(1 - p_{\theta}((\xi, t)|\mathbf{x})),
\end{aligned} \tag{3.29}$$

where $p_{\theta}((\xi, t)|\mathbf{x}) = \frac{\lambda_{\theta}((\xi, t)|\mathbf{x})}{\lambda_{\theta}((\xi, t)|\mathbf{x}) + \rho(\xi, t)}$ and $Y(\xi, t) = \mathbb{1}_{\{(\xi, t) \in \mathbf{x}\}}$ for $(\xi, t) \in \mathbf{x} \cup \mathbf{d}$, with

$$Pr(Y(\xi, t) = 1) = \frac{\lambda_{\theta}((\xi, t)|\mathbf{x})}{\lambda_{\theta}((\xi, t)|\mathbf{x}) + \rho(\xi, t)}. \tag{3.30}$$

Hence, the local logistic likelihood (3.26) is a weighted logistic regression with offset term $\log \frac{\lambda(\xi, t)}{\rho(\xi, t)}$ and weights $W_{\sigma_s}(u^* - \cdot) W_{\sigma_t}(v^* - \cdot)$ that can be maximised in the same way as logistic likelihood approach in Section 3.1.1.2 at point $(u^*, v^*) \in \mathcal{W}$ as follows.

Algorithm 3

- Generate a set of dummy points according to 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 $(u_k, v_k) \in \mathcal{W}$,
- Compute the weights $W_{\sigma_s}(u^* - u_k) W_{\sigma_t}(v^* - v_k)$,
- Obtain the response variables y_k (1 for data points, 0 for dummy points),

- Compute the values $\mathcal{S}((u_k, v_k), \mathbf{x})$ of the vector of sufficient statistics at each quadrature point,
- Fit a logistic regression model with explanatory variables $\mathcal{S}((u_k, v_k), \mathbf{x})$ and offset $\log[\lambda(u_k, v_k)/\rho(u_k, v_k)]$ and weights $W_{\sigma_s}(u^* - u_k)W_{\sigma_t}(v^* - v_k)$ on the responses y_k to obtain local estimates $\hat{\boldsymbol{\theta}}(u^*, v^*)$ for the \mathcal{S} -vector and intercept $\hat{\theta}_0(u^*, v^*)$,
- Return the local parameter estimator $\hat{\gamma}(u^*, v^*) = \exp(\hat{\boldsymbol{\theta}}(u^*, v^*))$ and $\hat{\beta}(u^*, v^*) = \exp(\hat{\theta}_0(u^*, v^*))$ and in case that $\lambda(u_k, v_k)/\rho(u_k, v_k)$ is constant c , the offset parameter may be omitted and return $\hat{\gamma}(u^*, v^*) = \exp(\hat{\boldsymbol{\theta}}(u^*, v^*))$ and $\hat{\beta} = c^{-1} \exp(\hat{\theta}_0(u^*, v^*))$.

Algorithm 3 can be implemented in the same way as *Algorithm 2* in R. However, it requires to compute the weights $W_{\sigma_s}(u^* - u_k)W_{\sigma_t}(v^* - v_k)$. We assume the weight functions are kernel densities for simplicity. For a density estimation and bandwidth selection, we develop proposed approach in Baddeley (2017) to the spatio-temporal framework. Hence, first, we have to select bandwidths in (3.24).

3.2.3 Bandwidth selection

In the context of point processes, a range of different methods for bandwidth selection have been proposed in the literature (see e.g. Baddeley et al. (2015) and Davies et al. (2018) for overviews), and most noteworthy are perhaps the recent method of Cronie and van Lieshout (2018) and the Poisson processes likelihood cross-validation method in Loader (1999) which we here develop it to the logistic likelihood.

The logistic likelihood cross-validation criterion is defined by

$$LLCV(\sigma_s, \sigma_t) = \sum_{i=1}^n \log \frac{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x})}{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} - \int_S \int_T \rho(u, v) \log \frac{\hat{\lambda}((u, v)|\mathbf{x}) + \rho(u, v)}{\rho(u, v)} dv du, \quad (3.31)$$

where $\hat{\lambda}((u, v)|\mathbf{x}) = \lambda_{\hat{\boldsymbol{\theta}}(u, v)}((u, v)|\mathbf{x})$, and $\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) = \lambda_{\hat{\boldsymbol{\theta}}_{-i}(\xi_i, t_i)}((\xi_i, t_i)|\mathbf{x})$ is the plug-in estimation of conditional intensity at a point (ξ_i, t_i) , using the leave-one-out parameter estimation $\hat{\boldsymbol{\theta}}_{-i}(\xi_i, t_i)$ based on $\mathbf{x} \setminus (\xi_i, t_i)$, and $\hat{\boldsymbol{\theta}}(\xi, t) = \hat{\boldsymbol{\theta}}((\xi, t); \sigma_s, \sigma_t)$ is the local estimate of $\boldsymbol{\theta}$ at spatio-temporal point (ξ, t) using bandwidths $\sigma_s, \sigma_t > 0$. The optimal bandwidths $\sigma_s^{opt}, \sigma_t^{opt}$ are the maximiser of (3.31).

It is difficult to approximate the integrals on the right hand side of (3.31). To address

this issue, we have by (3.26)

$$\begin{aligned}
LLCV(\sigma_s, \sigma_t) &\approx \sum_{i=1}^n \log \frac{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x})}{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \\
&\quad + \sum_{i=n+1}^{n+p} \log \frac{\rho(\xi_i, t_i)}{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)} \\
&= \sum_{i=1}^n \log \hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) \\
&\quad + \sum_{i=1}^{n+p} \log \frac{\rho(\xi_i, t_i)}{\hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) + \rho(\xi_i, t_i)}.
\end{aligned} \tag{3.32}$$

Note that, in principle, it is required to fit $n + p$ local logistic likelihoods for sample size $n + p - 1$ for obtaining a single evaluation of (3.32). Hence, it has a high computational cost and for overcoming it, we approximate (3.32) as follows. As in the case of density estimation in page 90 of Loader (1999), we may approximate $\hat{\theta}_{-i}(\xi_i, t_i)$ to first order using leverage and influence functions (Baddeley 2017, Baddeley et al. 2019). Hence, we have

$$\begin{aligned}
\log \hat{\lambda}_{-i}((\xi_i, t_i)|\mathbf{x}) - \log \hat{\lambda}((\xi_i, t_i)|\mathbf{x}) &\approx -\zeta_{\hat{\theta}(\xi_i, t_i)}((\xi_i, t_i)|\mathbf{x})^\top H((\xi_i, t_i); \hat{\theta}(\xi_i, t_i))^{-1} \\
&\quad \times \zeta_{\hat{\theta}(\xi_i, t_i)}^*((\xi_i, t_i)|\mathbf{x}),
\end{aligned} \tag{3.33}$$

where

$$\begin{aligned}
\zeta_{\theta}^*((\xi_i, t_i)|\mathbf{x}) &= W_{\sigma_s}(0)W_{\sigma_t}(0)\zeta_{\theta}((\xi_i, t_i)|\mathbf{x}) \\
&\quad + \sum_{j \neq i, j \in \mathbf{x}} W_{\sigma_s}(\xi_j - \xi_i)W_{\sigma_t}(t_j - t_i) \Delta_{(\xi_i, t_i)} \zeta_{\theta}((\xi_j, t_j)|\mathbf{x}) \\
&\quad - \sum_{j \neq i, j \in \mathbf{x} \cup \mathbf{d}} W_{\sigma_s}(\xi_j - \xi_i)W_{\sigma_t}(t_j - t_i) \Delta_{(\xi_i, t_i)} \pi_{\theta}((\xi_j, t_j)|\mathbf{x}),
\end{aligned} \tag{3.34}$$

with $\Delta_{(u,v)}g((u', v'), \mathbf{x}) := g((u', v'), \mathbf{x} \cup (u, v)) - g((u', v'), \mathbf{x} \setminus (u, v))$, for $(u, v), (u', v') \in \mathcal{W}$ and

$$\pi_{\theta}((\xi_j, t_j)|\mathbf{x}) = \zeta_{\theta}((\xi_j, t_j)|\mathbf{x}) \frac{\lambda_{\theta}((\xi_j, t_j)|\mathbf{x})}{\lambda_{\theta}((\xi_j, t_j)|\mathbf{x}) + \rho(\xi_j, t_j)}.$$

Hence, we approximate (3.32) by

$$\begin{aligned}
LLCV(\sigma_s, \sigma_t) &\approx \sum_{i=1}^n \log \lambda_{\hat{\boldsymbol{\theta}}(\xi_i, t_i)}((\xi_i, t_i) | \mathbf{x}) \\
&\quad - \sum_{i=1}^n \zeta_{\hat{\boldsymbol{\theta}}(\xi_i, t_i)}((\xi_i, t_i) | \mathbf{x})^\top H((\xi_i, t_i); \hat{\boldsymbol{\theta}}(\xi_i, t_i))^{-1} \zeta_{\hat{\boldsymbol{\theta}}(\xi_i, t_i)}^*((\xi_i, t_i) | \mathbf{x}) \\
&\quad - \sum_{i=1}^{n+p} \frac{\hat{\lambda}_{-i}((\xi_i, t_i) | \mathbf{x})}{\hat{\lambda}_{-i}((\xi_i, t_i) | \mathbf{x}) + \rho(\xi_i, t_i)},
\end{aligned} \tag{3.35}$$

The third term of right side in (3.35) is the first order of Taylor series expansion of $\log \frac{\rho}{\hat{\lambda}_{-i} + \rho}$ ($\log(1 - x) = -x - \frac{x^2}{2} - \dots$, $-1 < x < 1$) by

$$\log \frac{\rho}{\hat{\lambda}_{-i} + \rho} = \log\left(1 - \frac{\hat{\lambda}_{-i}}{\hat{\lambda}_{-i} + \rho}\right) \approx -\frac{\hat{\lambda}_{-i}}{\hat{\lambda}_{-i} + \rho}, \tag{3.36}$$

where $\hat{\lambda}_{-i}((\xi_i, t_i) | \mathbf{x})$ is obtained from (3.33).

Chapter 4

Simulations

Simulation-based techniques play an important role in the analysis of point patterns which are used in calculation of summary statistics and goodness-of-fit tests for point process models and also visualisation of point process models (Illian et al. 2008). This chapter treats both an algorithm based on MCMC methods and two-step simulation-based procedure for Gibbs and Cox-Gibbs models proposed in Chapter 2. We implemented simulation algorithms and inference methods in R code. Most of them can be found here <http://edith.gabriel.pagesperso-orange.fr/software.html>.

4.1 Gibbs models' 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.

4.1.1 Birth-death Metropolis-Hastings algorithm

For \mathbf{x} a spatio-temporal point pattern in \mathcal{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 $(u, v) \in \mathcal{W}$ 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 $d(\mathbf{x}, \cdot)$ 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 $q(\mathbf{x}) = \frac{1}{2}$, $b(\mathbf{x}, \cdot) = 1/\ell(\mathcal{W})$ 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 $r((u, v); \mathbf{x}) = \frac{\ell(\mathcal{W})}{n(\mathbf{x})+1} \times \lambda((u, v)|\mathbf{x})$ is the Hastings ratio (Iftimi et al. 2018), we obtain the following birth-death Metropolis-Hastings algorithm.

Algorithm 4

For $n = 0, 1, \dots$, 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 $r((u, v); \mathbf{x}) = \frac{\ell(\mathcal{W})}{n(\mathbf{x})+1} \lambda((u, v)|\mathbf{x})$, $(u, v) \notin \mathbf{x}$.
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 $r((\xi, t); \mathbf{x} \setminus (\xi, t)) = \frac{\ell(\mathcal{W})}{n(\mathbf{x})} \lambda((\xi, t)|\mathbf{x} \setminus (\xi, t))$, $(\xi, t) \in \mathbf{x}$.
If $y_2 < 1/r((\xi, t); \mathbf{x} \setminus (\xi, t))$ then $X_{n+1} = \mathbf{x} \setminus (\xi, t)$ 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. 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 4*. 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 Cox-Gibbs models' simulation

For generating the simulations from the hybrid Cox-Gibbs models, we follow two different procedures in simulation study and model validation.

For simulation study, we simulate point patterns under this model in two steps: first, a realization of Gaussian random field is simulated by the function `rLGCP` from the R-package `spatstat`. Second, a realization of hybrid Cox-Gibbs model given that realization is simulated using the birth–death Metropolis–Hastings algorithm.

For model validation, we simulate point patterns under fitted hybrid Cox-Gibbs model in two steps too: first, we sample from the posterior distribution of the fitted

INLA model and extract the values of its latent field as a first-order interaction function (or trend function) of the Gibbs model. Second, a realization of hybrid Cox-Gibbs model given computed trend is simulated using the birth–death Metropolis–Hastings algorithm by the function `rmh.default` from the R-package `spatstat`. Indeed, we need the predictions from the model anywhere in the domain as we do not have any other covariate. To do this, an option is computation the linear predictor at the mesh nodes and project it onto the grid taking into account the uncertainty by sampling from its posterior distribution of the fitted INLA model. Hence, we could sample directly from the posterior marginal of the linear predictor at the observation scale. Indeed, when the likelihood depends not only in the latent field but also on some hyperparameters (as in the Gaussian case) we need to sample both $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ jointly. In R-INLA this can be done using the `inla.posterior.sample` function which needs the output from the `inla` function by setting `config=TRUE` in the `control.compute` option.

Note that these simulations are motivated by our application on temperature anomalies (see Section 5.2) in which we consider spatial interactions during time intervals. Hence simulations above are performed in space but could be easily extended to the spatio-temporal framework in the case of continuous times, by using `stpp` package for simulating log-Gaussian Cox processes and estimating K -function.

4.3 Simulation study

The aim of our simulation study is threefold: first we want to compare the performance of the composite likelihood approaches defined in the previous chapter for the Gibbs models, second we want to test the simulation procedure of the Cox-Gibbs model and third we want to test the local logistic likelihood method.

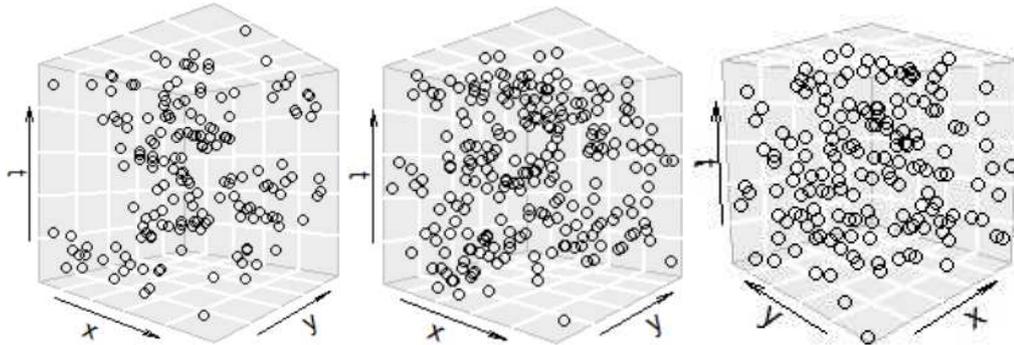
4.3.1 Simulation study for Gibbs models

We compare the performance of the pseudo-likelihood and logistic likelihood approaches on the spatio-temporal multi-scale Geyer point process. Due to obtained results, we then perform a simulation study for the spatio-temporal hybrid Strauss hardcore model based on logistic likelihood approach.

We implement the estimation and simulation algorithms in R and generate 100 simulated realizations in the unit cube from three models of hybrid Geyer point process. 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 4.1. We consider a burn-in period of 20,000 steps in *Algorithm 4*. Fig. 4.1 shows one realization of each model.

Table 4.1 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
<i>Model 1</i>	70	(1.5,1.5)	(0.05,0.1)	(0.05,0.1)	(2,2)
<i>Model 2</i>	100	(0.5,1.5)	(0.05,0.1)	(0.05,0.1)	(1,3)
<i>Model 3</i>	200	(0.8,0.8)	(0.05,0.1)	(0.05,0.1)	(1,1)

**Figure 4.1** Realizations of *Model 1* (left); *Model 2* (middle); *Model 3* (right) from hybrid Geyer model.

According to Baddeley et al. (2014), we generate a spatio-temporal Poisson process with intensity $\rho = 4n(\mathbf{x})$ (resp. $4n(\mathbf{x})/\ell(\mathcal{W})$) 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 4.3) and plot the related boxplots (Fig. 4.2). In Table 4.3 the lowest RMSE value is in bold and in Fig. 4.2 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.

Table 4.2 RMSE of parameter estimates from 100 simulated realizations of the multi-scale Geyer point process model.

Method	<i>Model 1</i>			<i>Model 2</i>			<i>Model 3</i>		
	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\lambda}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
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

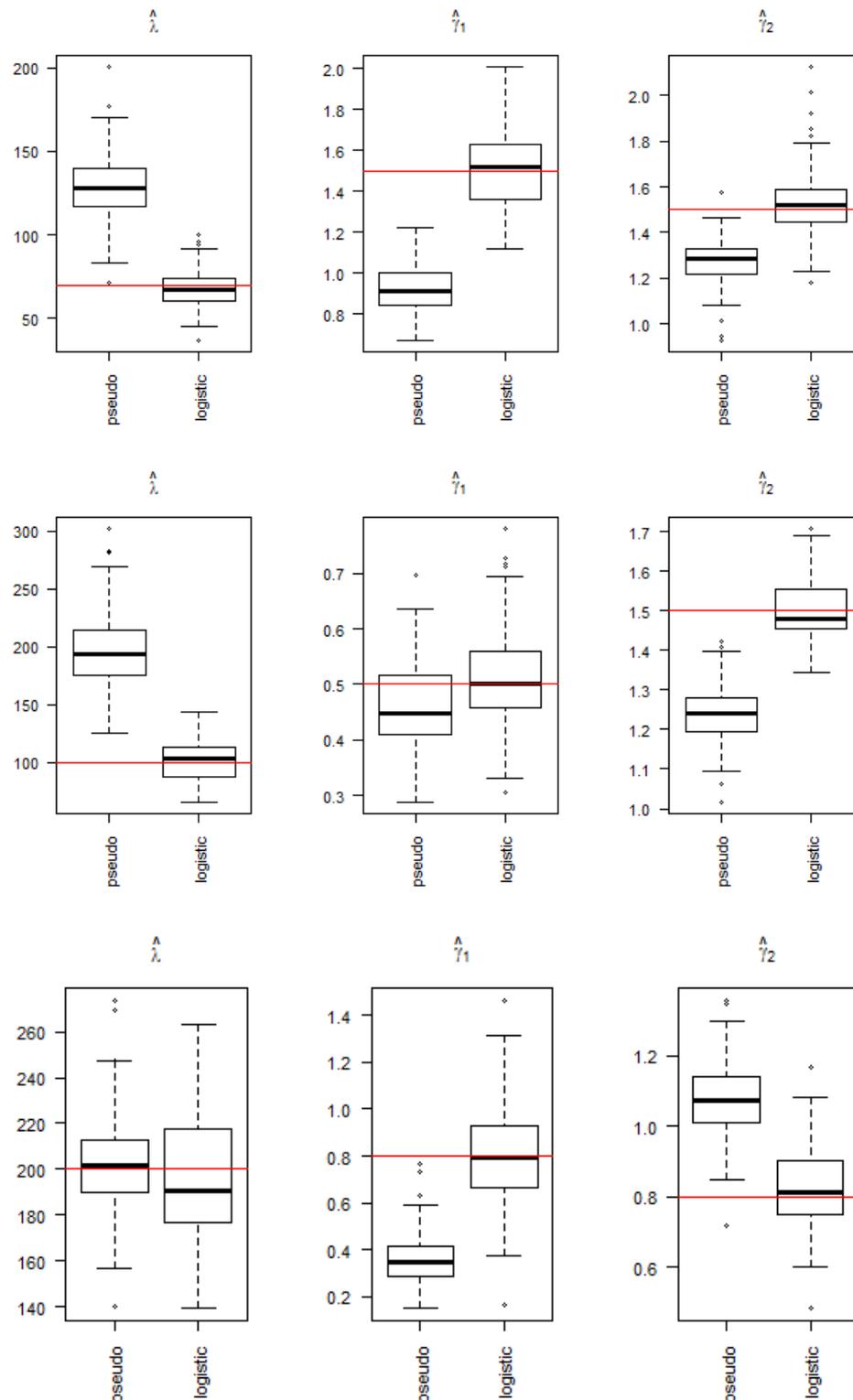


Figure 4.2 Boxplots of regular parameters estimated from the pseudo-likelihood and logistic likelihood approaches for *Model 1* (first row), *Model 2* (second row) and *Model 3* (third row) from hybrid Geyer model. True values are represented by horizontal red lines.

Table 4.3 Parameter combinations of three hybrid Strauss hardcore point process models used in simulation study.

Model	Values of parameter			
	Regular parameters		Irregular parameters	
	λ	γ	r, q	hc_s, hc_t
<i>Model 4</i>	70	(0.8,.08)	(0.05,0.1)	(0.01,0.01)
<i>Model 5</i>	50	(1.5,1.5)	(0.05,0.1)	(0.01,0.01)
<i>Model 6</i>	70	(0.5,1.5)	(0.05,0.1)	(0.01,0.01)

Table 4.4 Point and interval parameter estimates of three hybrid Strauss hardcore point process models used in simulation study.

True values	Mean	95% CI
<i>Model 4</i>		
$\lambda = 70$	71.43	(69.16,73.70)
$\gamma_1 = 0.8$	0.89	(0.78,1.00)
$\gamma_2 = 0.8$	0.78	(0.74,0.82)
<i>Model 5</i>		
$\lambda = 50$	50.84	(48.99,52.68)
$\gamma_1 = 1.5$	1.41	(1.23,1.60)
$\gamma_2 = 1.5$	1.46	(1.38,1.54)
<i>Model 6</i>		
$\lambda = 70$	71.67	(69.18,74.15)
$\gamma_1 = 0.5$	0.50	(0.43,0.57)
$\gamma_2 = 1.5$	1.49	(1.42,1.55)

(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.

Hence, we consider the logistic likelihood approach for a simulation study of hybrid Strauss hardcore model. We generate simulations of three stationary spatio-temporal hybrid Strauss hardcore point processes specified by a conditional intensity of the form (3.16) in $\mathcal{W} = [0, 1]^3$. The parameter values used for the simulations are reported in Table 4.3. The spatial and temporal radii r and q , spatial and temporal hardcores hc_s and hc_t are treated as known parameters.

We generate 100 simulations of each specified model. Boxplots of parameter estimates λ , γ_1 , and γ_2 obtained from the logistic likelihood estimation method for each model are shown in Fig. 4.3. The red horizontal lines represent the true parameter values. Point and interval parameter estimates λ , γ_1 , and γ_2 are reported in Table 4.4. Most of the estimated parameter values are close to the true values for three models. Due to visual and computational comparisons, we conclude that the logistic likelihood approach performs well for spatio-temporal hybrid Strauss hardcore point processes.

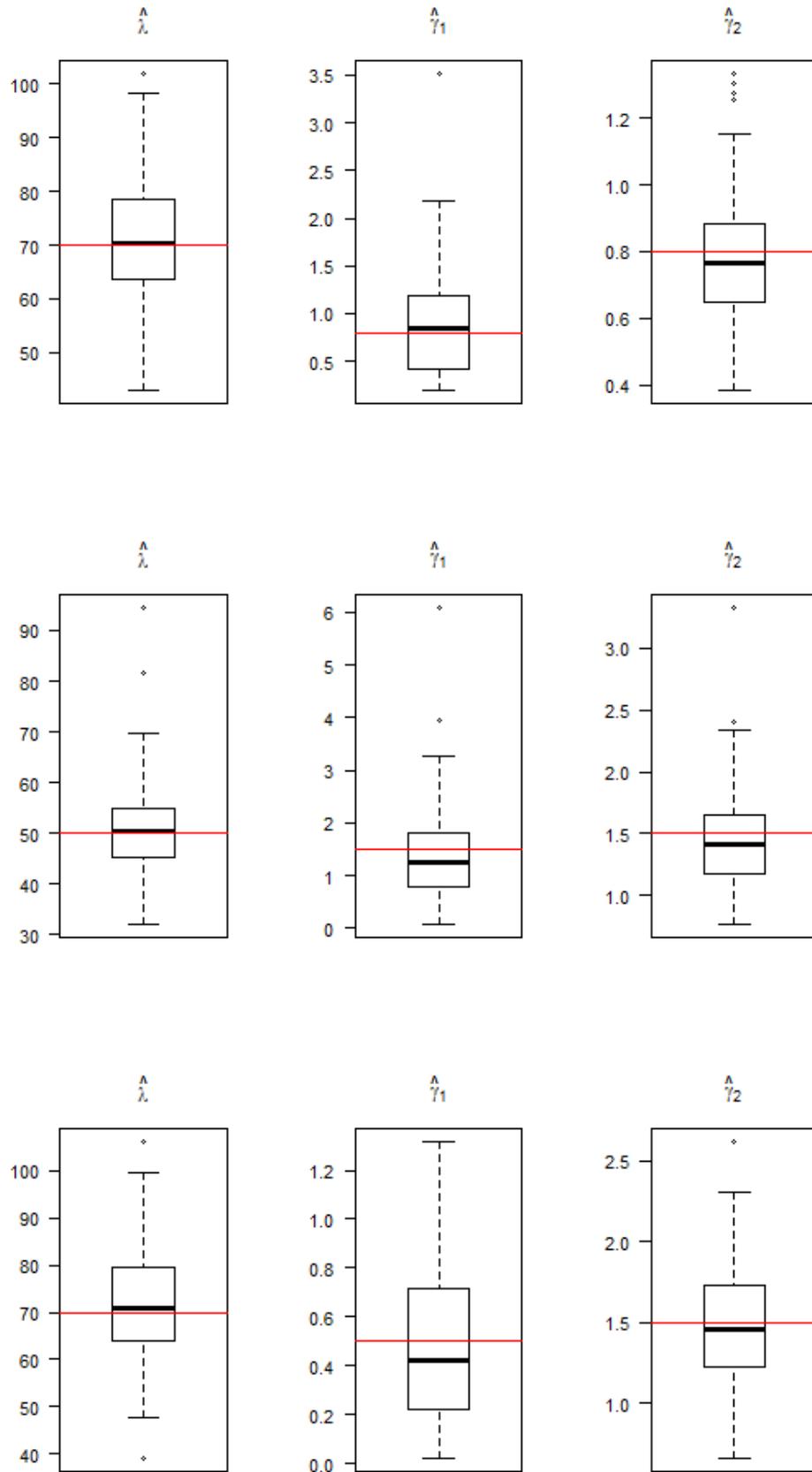


Figure 4.3 Boxplots of parameter estimates of the hybrid Strauss hardcore point process obtained from the logistic likelihood estimation methods. Up to down: *Model 4*, *Model 5*, and *Model 6*.

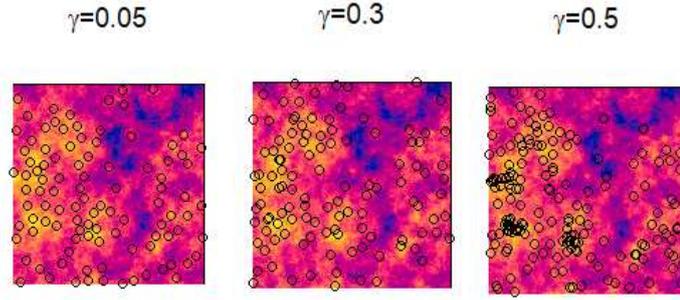


Figure 4.4 The simulated LGCGPs on the unit square (black points) and the corresponding realisation of \mathbf{Z} (blue to yellow scale image).

4.3.2 Simulation study for Cox-Gibbs models

Consider a spatial LGCGP on the observation window $\mathcal{W} = [0, 1]^2$ with density

$$f(\mathbf{x}) = \mathbb{E} \left[\frac{1}{c(\mathbf{Z})} \prod_{i=1}^n \exp(Z(\xi_i)) \gamma^{\min\{s, t_r(\xi_i, \mathbf{x} \setminus \xi_i)\}} \right], \quad (4.1)$$

where $t_r(\xi_i, \mathbf{x} \setminus \xi_i) = \sum_{j \neq i} \mathbb{1}\{\|\xi_j - \xi_i\| \leq r\}$. For simulation of point pattern under this model, first, a realisation \mathbf{z} of \mathbf{Z} is simulated with the function `rLGCP` from the R-package `spatstat` where \mathbf{Z} is a Gaussian random field with constant mean 5 and exponential covariance function

$$C(u) = Q^2 \exp(-u/\alpha)$$

with $Q^2 = 2$ and $\alpha = 0.1$. Second, three realisations of LGCGP given $\mathbf{Z} = \mathbf{z}$ is simulated using a birth-death Metropolis-Hastings algorithm is implemented in function `rmh.default` in `spatstat` package from a Geyer point process model with $r = 0.05$, $s = 2$ and three different value of $\gamma = 0.05, 0.3, 0.5$.

To assess the degree of clustering and regularity we consider $L(u) - u$ which its negative (resp. positive) values show point process is repulsive (resp. aggregated) at inter-point distances u .

Fig. 4.4 is three examples of simulated realisations of the LGCGP on $\mathcal{W} = [0, 1]^2$ with mentioned parameters. Fig. 4.4 shows the effect of more intensity of Gaussian random field, as first-order interaction of LGCGP, in some part of the window leads to more density in same parts when the value of γ is increased. Fig. 4.5 shows plots of empirical estimates of the L -functions minus the identity of the point patterns with above parameters which each of these plots comes with the empirical L -functions of 19 different realisations of same process with their mean to assess the general behaviour of the LGCGP related to value of γ . The point patterns exhibit both regularity and aggregation in different scale with a decreasing degree of regularity at small to moderate distances as increases, but a similar degree of aggregation at large distances. However,

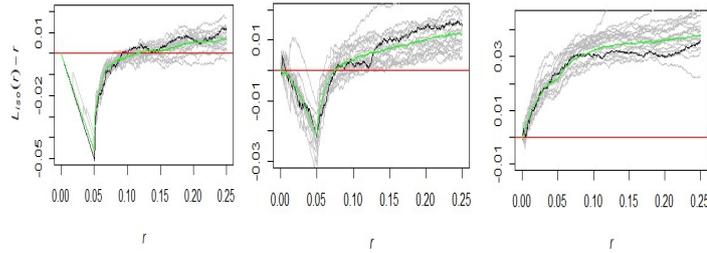


Figure 4.5 The empirical L -function minus the identity for the point patterns of Fig. 4.4 (black curve) and for 19 simulations of the same process (grey curves) plus their mean (green curve).

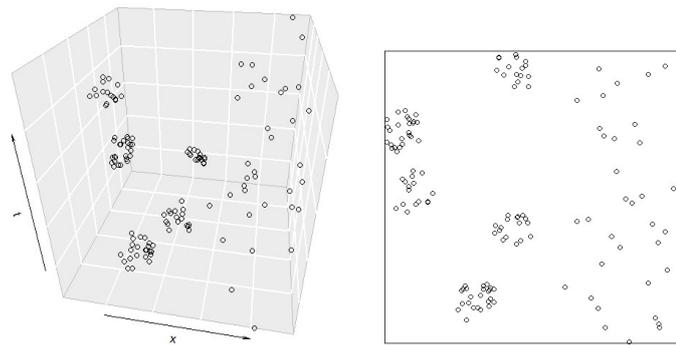


Figure 4.6 (Left) A simulated multi-structure point pattern in unit cube with 132 points. (Right) A spatial projection of point pattern.

the general behaviour of the empirical L -function suggests a tendency to a higher degree of clustering at large distances when the value of γ is increased. Fig. 4.5 also suggests that for realisations of an LGCGP, the empirical estimate of $L(u) - u$ often has a global minimum when u is close to the interaction radius r , at least when there is strong to moderate repulsion in the model.

4.3.3 First evaluation of the local logistic likelihood approach

To assess the performance of the local logistic likelihood procedure, we generate a multi-structure point pattern in an unit cube. We consider a pattern which is a combination of strong clustering and randomness. Fig. 4.6 (left) is the plot of spatio-temporal multi-structure point pattern in unit cube divided to two regions by plane $x = 0.5$ that the left region (region I) appears to be strongly clustered while the right region (region II) appears to be randomised. Fig. 4.6 (right) is a spatial projection of point pattern.

We consider a homogeneous spatio-temporal Strauss hardcore point process as a template model with density (2.14). Due to the value of γ , the model is clustered if $\gamma > 1$ and inhibited if $\gamma < 1$.

We fitted the Strauss hardcore point process model with $r = q = 0.05$ and spatial and temporal hardcore distances equal to 0.005 to the pattern, yielding $\hat{\lambda} = 45.68$ and

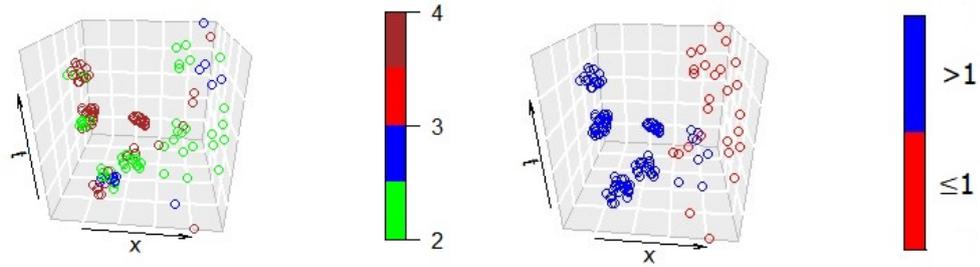


Figure 4.7 Local estimates of parameters λ (left) and γ (right) of Strauss hardcore point process.

$\hat{\gamma} = 3.45$. The optimal bandwidth are $\sigma_s = 0.2$ and $\sigma_t = 0.03$ by approximating the logistic likelihood cross-validation criterion (3.35). Fig. 4.7 is the local estimates of the first-order interaction λ (left) and second-order interaction γ (right) for the Strauss hardcore point process using these optimal bandwidths. The points in region I have the large estimated values for γ (larger than one with blue color) that suggests a strong clustered pattern. For most of points in region II, γ is locally estimated smaller or close to one with red color. There are some points in region II that γ is estimated larger than one which is due to existence other points (data or dummy) near them.

Chapter 5

Applications

In this chapter, we consider the Gibbs models developed in Section 2.1 to describe two different patterns of forest fire occurrences, the first in South of France, the second in centre Spain. The Cox-Gibbs model proposed in Section 2.2 is considered in an innovative application of spatio-temporal modeling of hotspots of temperature, and in particular of temperature anomalies, in the United States.

5.1 Gibbs models for 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 modelling wildfire occurrences to better understand their behaviors. In this section, first, we focus on the modelling of forest fire occurrences in the Bouches-du-Rhône county (Southern France) between 2001 and 2015 available from the Prométhée database¹ and the Castilla-La Mancha in Spain between 2004 to 2007 available from the `clmfires` dataset in R-package `spatstat` by hybrid Geyer model and hybrid Strauss hardcore model, respectively.

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.

Hence, we propose to fit the spatio-temporal hybrid Geyer point process model

¹<https://www.promethee.com/en>

(2.12) on wildfire occurrences to take into account both the inhomogeneities induced by covariates and the multi-scale structure of interactions.

However, the complexity of forest fire occurrences is due in particular to the existence of multi-scale structures and also hardcore distances in space and time. For instance, changes in vegetation due to forest fires burnt areas affect the probability of fire occurrences during the regeneration period leading to the existence of hardcore distances in space-time.

The main second focus of our forest fire pattern analysis is to quantify the interactions across a range of spatio-temporal scales with the presence of hardcore distances which can be done by using the spatio-temporal hybrid Strauss hardcore process model (2.16). We apply the hybrid Strauss hardcore model on a forest fire pattern of Castilla-La Mancha in Spain.

5.1.1 Hybrid Geyer model for forest fires in France

Our data set is of the form (ξ_i, t_i) , $i = 1, \dots, 434$, where (ξ_i, t_i) corresponds to a wildfire with more than 1 *hectare* of burnt surface spatially indexed by a DFCI² cell center ξ_i in the Lambert 93 projection system and year $t_i \in \{2001, \dots, 2015\}$. 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. 5.1 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. 5.2 in 2001.

5.1.1.1 Model fitting

Here we first estimate the spatio-temporal trend and then fit the multi-scale spatio-temporal Geyer model to forest fire occurrences. This two-step model fitting procedure

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

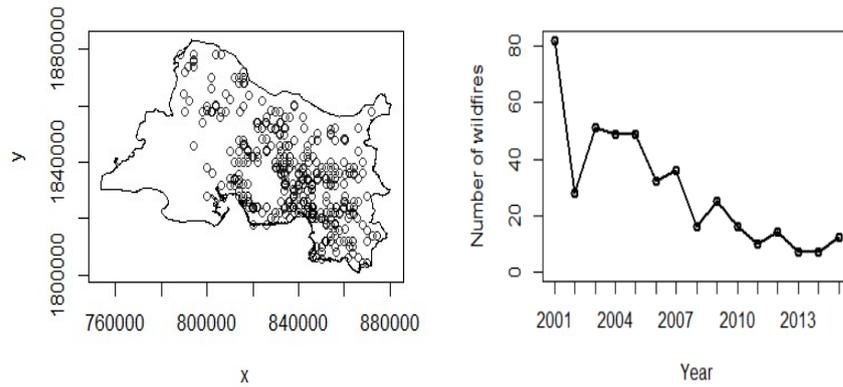


Figure 5.1 (Left) Forest fire locations in UTM coordinate system (distance in meters), with more than 1 *hectare* of burnt area, recorded during the years 2001 to 2015 in the Bouches-du-Rhône county in France. (Right) Number of recorded forest fires per year.

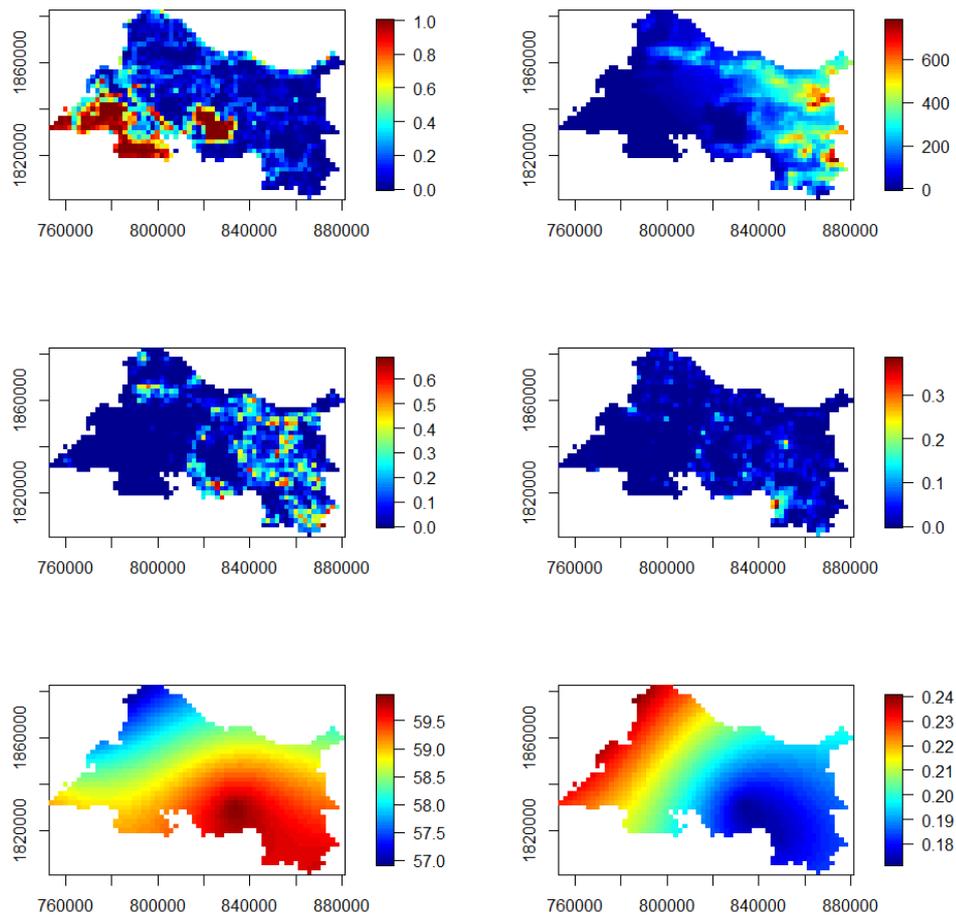


Figure 5.2 Maps of covariates: water coverage (top left), elevation (top right), coniferous cover (middle left), building cover (middle right), temperature average (bottom left) and square root of precipitation (bottom right) in 2001.

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.

Spatio-temporal trend estimation

We express the spatio-temporal trend (2.12) as $\lambda(\xi, t) = \beta\mu(\xi, t)$ where $\log \mu(\xi, t)$ is assumed to linearly depend on covariates:

$$\log \mu(\xi, t) = \beta_0 + \sum_{k=1}^4 \beta_k^S Z_k^S(\xi) + \sum_{l=1}^2 \beta_l^{ST} Z_l^{ST}(\xi, t) + \alpha t \quad (5.1)$$

with $Z_k^S(\xi)$, $k = 1, \dots, 4$, 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 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 ξ_i , $i = 1, \dots, 1320$ and years $t_j = 2001, \dots, 2015$ for $j = 1, \dots, 15$, and we consider a Poisson response for our model $N_{ij} | \mu(\xi_i, t_j) \sim \text{Poisson}(\mu(\xi_i, t_j))$, where N_{ij} is the number of forest fires in i^{th} DFCI cell at year t_j . The coefficient β will be estimated simultaneously with the others regular parameters by the logistic likelihood approach. Table 5.1 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 favour (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 obtained more significance of the covariates than under our hybrid Geyer saturation model.

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 dis-

Table 5.1 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	$< 2 \times 10^{-16}$ ***

tance, 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 advantage of logistic likelihood over pseudo-likelihood (see Section 4.3.1), 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 , $j = 1, \dots, m$, with varying m in $\{1, 2, 3, 4, 5\}$. The minimum AIC is obtained for the combination given in Table 5.2. 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.

5.1.1.2 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 (2.12) by *Algorithm 4* 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 $\hat{K}_{obs}^{inh}(h_s, h_t)$ and

Table 5.2 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]

$\hat{K}_i^{inh}(h_s, h_t), i \in \{1, \dots, n_{sim}\}$, 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

$$L(h_s, h_t) = \min_{1 \leq i \leq n_{sim}} \hat{K}_i^{inh}(h_s, h_t), \quad U(h_s, h_t) = \max_{1 \leq i \leq n_{sim}} \hat{K}_i^{inh}(h_s, h_t). \quad (5.2)$$

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 $\{\hat{K}_1^{inh}(h_s, h_t), \dots, \hat{K}_{n_{sim}}^{inh}(h_s, h_t), \hat{K}_{obs}^{inh}(h_s, h_t)\}$. We define the local p -value for each pair (h_s, h_t) by

$$p(h_s, h_t) = \frac{1 + \sum_{i=1}^{n_{sim}} \mathbb{1}\{T_i(h_s, h_t) > T_{obs}(h_s, h_t)\}}{n_{sim} + 1}, \quad (5.3)$$

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

$$T(h_s, h_t) = \sqrt{\frac{(\hat{K}^{inh}(h_s, h_t) - E(h_s, h_t))^2}{V(h_s, h_t)}}. \quad (5.4)$$

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

$$\tilde{T} = \int_0^{h_{t,max}} \int_0^{h_{s,max}} T(h_s, h_t) dh_s dh_t, \quad (5.5)$$

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) recommends to compare the results for several

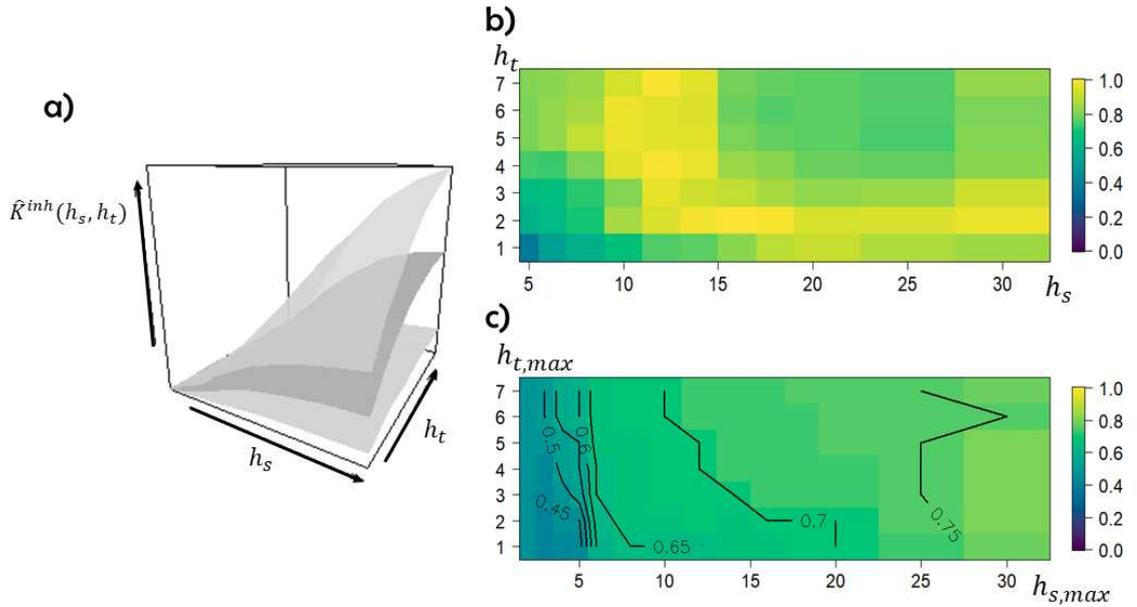


Figure 5.3 Temporal separations h_t are in *year* and spatial distances h_s are in *kilometer*. *a)* Envelopes of the spatio-temporal inhomogeneous K -function for the simulated spatio-temporal multi-scale Geyer point process according to the estimated parameters. *b)* Image plot of the local p -value. *c)* Image plot of the global p -value for any pairs of $(h_{s,max}, h_{t,max})$.

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. 5.3.a) shows the spatio-temporal inhomogeneous K function computed on our dataset (dark grey) and the envelopes obtained from our hybrid Geyer model (light grey); $\hat{K}_{obs}^{inh}(h_s, h_t)$ 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. 5.3.b). Global p -values are given in Fig. 5.3.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 , $i = 1, \dots, n_{sim}$ (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

$$E_i = \frac{1}{n_{ns}} \sum_{j=1}^{n_{st}} \mathbb{1}\{R_j \prec R_i\}, \quad (5.6)$$

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

$$p_{erl} = \frac{1 + \sum_{i=1}^{n_{sim}} \mathbb{1}\{E_i \geq E_{obs}\}}{n_{sim} + 1}. \quad (5.7)$$

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 $\mu(\xi, t)/(2000 \times 2000 \times 1)$ (5.8). 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.

5.1.2 Hybrid Strauss hardcore model for forest fires in Spain

Castilla La Mancha is located in the middle of the Iberian peninsula and the third largest of Spain's autonomous regions representing 15.7% of the Spanish national territory.

5.1.2.1 Data and covariates description

The `clmfires` dataset available in `spatstat` package records the occurrence of forest fires in the region of Castilla-La Mancha, Spain (Figure 5.4, left) from 1998 to 2007. The study area is approximately 400 by 400 *km*. The `clmfires` dataset has already been used in some academic works devoted to the point process theory (see e.g. Juan et al. (2010), Gomez-Rubio (2020), Myllymäki et al. (2021)). The dataset has two levels of precision: from 1998 to 2003 locations were recorded as the centroid of the corresponding “district unit”, while since 2004 locations coorespond to the exact UTM coordinates of the centroids of the fires.

Due to the low precision of fire locations for the years 1998 to 2003 (Gomez-Rubio 2020), we focus on fires in the period 2004 to 2007. In this period, we consider large forest fires with burnt areas larger than 5 *ha*. Fig. 5.4 (middle) shows the point pattern of 432 wildfire locations onto the spatial region.

Due to memory constraints and availability of climate covariates in months, we consider monthly fire occurrences. The temporal component of the process takes integer values from 1 to 48. We thus consider $\mathcal{W} = S \times T$ where S is the region of Castilla-la-Mancha and $T = \{1, 2, \dots, 48\}$ corresponds to the months since January 2004. Fig. 5.4 (right) shows the monthly number of fires occurring during our time period. We

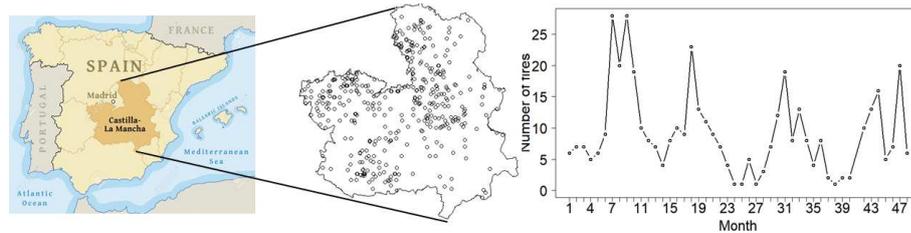


Figure 5.4 Left: Map of region of Castilla-La Mancha (Spain). Middle: Forest fire locations. Right: monthly numbers of fires recorded between January 2004 and December 2007 with burnt areas, spatial distances and time distances respectively bigger than 5 ha, 0.2 km and 100 days.

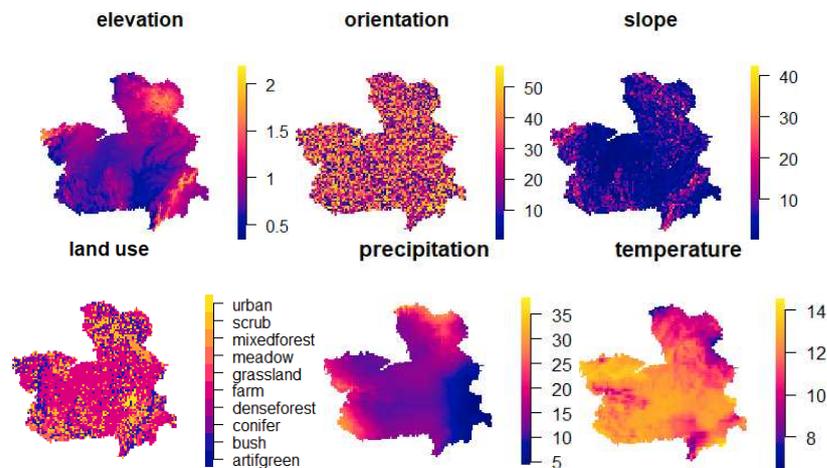


Figure 5.5 Image plot of environmental covariates (*elevation*, *orientation*, *slope* and *land use*) and climate covariates (*precipitation* and *temperature*) in January 2007.

observe seasonal effects with notably large numbers of fires in summer that could be caused by high temperatures and low precipitations in this period and also by human activities.

In point pattern analysis, the spatial (spatio-temporal) inhomogeneity of patterns is notably driven by covariates. The `clmfires` dataset contains four environmental covariates that we include in our analysis: *elevation*, *orientation*, *slope* and *land use*. The covariates are known on a spatial grid with pixels of 4×4 km, resulting in a total of 10,000 pixels. The *land use* is a factor-valued covariate whereas the others are real-valued covariates. We also consider weather data freely provided by the *WorldClim* database³ and containing monthly maximum temperatures ($^{\circ}C$) and total precipitations (*mm*). Fig. 5.5 illustrates the environmental covariates, which are considered fixed during our temporal period, and the climate covariates in January 2007.

³<https://www.worldclim.org>

5.1.2.2 Estimation

First, we estimate the trend function by considering a generalized linear model (GLM) on covariates. Then, by an exploratory analysis using spatio-temporal summary statistics we approximate the hardcore parameters and the interaction ranges. Finally, we use the logistic likelihood approach described in Section 3.1.1.2 for the estimation of regular parameters of our model with the trend function estimated in a preliminary step.

Trend estimate

Since covariates are available on a spatial grid, we restrict our attention on the related grid centers $\xi_i, i = 1, \dots, 10000$ and months $\{t_j\}_{j=1, \dots, 48} \in T$ and consider $N_{ij} | \lambda(\xi_i, t_j) \sim \text{Poisson}(\lambda(\xi_i, t_j))$ where N_{ij} is the number of forest fires in the i^{th} grid center at month t_j .

Following last section, by considering a GLM with Poisson response, we obtain:

$$\log \lambda(\xi_i, t_i) = \beta_0 + \sum_{k=1}^6 \beta_k Z_k(\xi_i, t_i), \quad (5.8)$$

where $Z_k(\xi_i, t_i), k = 1, \dots, 6$, are the environmental and climatic covariates at point (ξ_i, t_i) and $\beta_0, \beta_k, k = 1, \dots, 6$ are the coefficients to estimate. As said before, we consider the same values for environmental covariates over time. A straightforward way to fit a GLM in \mathbb{R} is to use the function `glm`. Table 5.3 reports the estimated coefficients in (5.8) and their significance level by a two-tailed Student's t-test. Coefficients higher (respectively lower) than zero imply an increase (resp. decrease) of the expected mean number of forest fires when the covariate value increase (resp. decrease). Those related to *elevation* and *temperature* are positively significant, showing that these two covariates favors the ignition of wildfires. At the opposite, the covariate *precipitation* as a negative significant coefficient indicating that an increase of the amount of precipitations induces a decrease in the mean number of forest fires. The *land use* appears not significantly different from zero, it can be explained by the low spatial resolution of the covariates.

Irregular parameter estimates

We have two types of irregular parameters in our spatio-temporal Gibbs point process. On the one hand, the hardcore distances that we can choose among the feasible solutions on the Pareto front of spatial and temporal interpoint distances. According to Fig. 5.6, we choose on the Pareto front the unique feasible solution in our case that gives non-zero values for the two hardcore distances, i.e. $hc_s = 0.35 \text{ km}$ and $hc_t = 1 \text{ month}$. On the other hand, for the nuisance parameters m, r_j and $q_j, j = 1, \dots, m$,

Table 5.3 Estimated coefficients, standard errors and p -values based on two-tailed Student's t -tests of significant differences from zero.

Coefficients	Estimate	Standard error	p -value
β_0 (intercept)	-8.468	0.298	$< 2 \times 10^{-16}$ ***
β_1 (elevation)	0.546	0.164	0.001 ***
β_2 (orientation)	0.005	0.003	0.114
β_3 (slope)	-0.019	0.01	0.054
β_4 (land use)	-0.009	0.024	0.689
β_5 (precipitation)	-0.007	0.002	0.003 **
β_6 (temperature)	0.054	0.006	$< 2 \times 10^{-16}$ ***

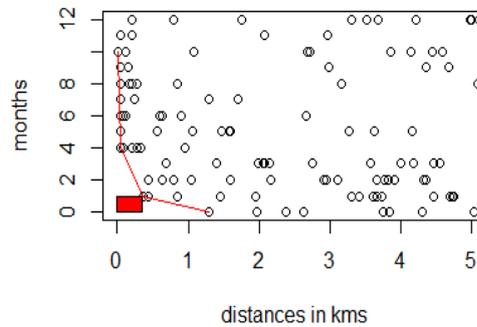


Figure 5.6 Spatial and temporal interpoint distances respectively lower than 5 kms and 12 months (black circles). The red line corresponds to the Pareto front and the red rectangle to the hardcore domain.

there is no common method for estimating them. 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 after the regular parameter estimation step. We chose 25 configurations of reasonable range for the nuisance parameters using a preliminary spatio-temporal exploratory analysis of the interaction ranges done with the inhomogeneous pair correlation function, the maximum nearest neighbor distance and the temporal auto-correlation function. We fitted the spatio-temporal hybrid Strauss point process model for a range of ad-hoc values $r_j \in (0.35, 20]$, $q_j \in \{2, \dots, 15\}$, $j = 1, \dots, m$ and $m \in \{1, \dots, 6\}$. The minimum AIC is obtained for the combination given in Table 5.4.

Regular parameter estimates

We consider the logistic likelihood method investigated in Section 3.1.1.2 to estimate the regular parameters. We simulate dummy points from an inhomogeneous Poisson

Table 5.4 Parameter estimates for $m = 6$.

		Irregular parameters				
r	0.5	1	1.5	6	15	20
q	2	4	6	8	12	15
Estimated regular parameters						
	$\hat{\gamma}_1 = 2.56$	$\hat{\gamma}_2 = 2.24$	$\hat{\gamma}_3 = 4.65$	$\hat{\gamma}_4 = 0.88$	$\hat{\gamma}_5 = 1.17$	$\hat{\gamma}_6 = 0.81$

point process with intensity $\rho(\xi, t) = C\hat{\lambda}(\xi, t)/\nu$ where $C = 4$ by a classical rule of thumb in the logistic likelihood approach, $\hat{\lambda}$ is the estimated trend and $\nu = 4 \times 4 \times 1$ is the volume of a grid cell on one month. In order to satisfy the hardcore condition in (2.18), we remove dummy points at spatial and temporal distances respectively less than hc_s and hc_t . Estimated regular parameters are provided in Table 5.4.

5.1.2.3 Goodness-of-fit

The goodness-of-fit is accomplished by simulating point patterns from the fitted model. The first diagnostic can be formulated by summary statistics of point processes. As the second-order characteristics carry most of the information on the spatio-temporal structure (Stoyan 1992, Gonzalez et al. 2016), we only consider the pair correlation function (g -function).

We generate $n_{sim} = 99$ simulations from the fitted hybrid Strauss hardcore model and compute the corresponding second-order summary statistics $g_i(h_s, h_t)$, $i = 1, \dots, n_{sim}$, for fixed spatio-temporal distances (h_s, h_t) . We then build upper and lower envelopes:

$$U(h_s, h_t) = \max_{1 \leq i \leq n_{sim}} g_i(h_s, h_t), \quad L(h_s, h_t) = \min_{1 \leq i \leq n_{sim}} g_i(h_s, h_t), \quad (5.9)$$

and compare the summary statistics obtained from the data, $g_{obs}(h_s, h_t)$, to the pointwise envelopes. If it lies outside the envelopes at some spatio-temporal distances (h_s, h_t) , then we reject at these distances the hypothesis that our data come from our fitted model. Fig. 5.7 shows the spatio-temporal inhomogeneous g -function computed on our dataset (blue) and the envelopes obtained from the fitted model (light grey); $g_{obs}(h_s, h_t)$ lies inside the envelopes for all (h_s, h_t) , meaning that the hybrid Strauss hardcore model is suitable for the data.

In addition, we compute global envelopes and p -value of the spatio-temporal g -functions based on the Extreme Rank Length (ERL) measure defined in Section 5.1.1.2. For each point pattern, we consider the long vector T_i , $i = 1, \dots, n_{sim}$ (resp. T_{obs}) merging the $g_i(\cdot, h_t)$ (resp. $g_{obs}(\cdot, h_t)$) estimates for all considered values h_t . The ERL measure of vector T_i (resp. T_{obs}) of length n_{st} is defined in (5.6). Due to the global p -value $p_{erl} = 0.59$ and the absence of significant regions, that corresponds here to pairs of spatial and temporal distances where the statistics is significantly above or below the

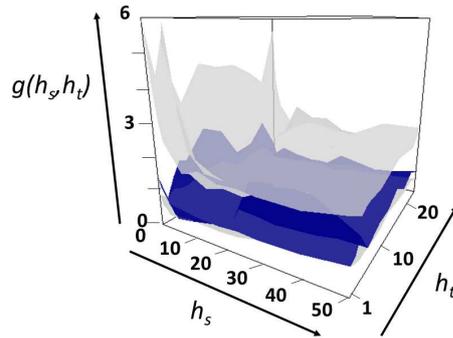


Figure 5.7 Envelopes of the spatio-temporal inhomogeneous g -function obtained from simulations of the fitted spatio-temporal hybrid Strauss hardcore point process (light grey). The blue surface corresponds to g_{obs} . Temporal separations are in *month* and spatial distances are in *kilometer*.

envelopes (see Fig. 5.8 and GET package), we conclude that our hybrid Strauss hardcore model can not be rejected a significance level of 1%.

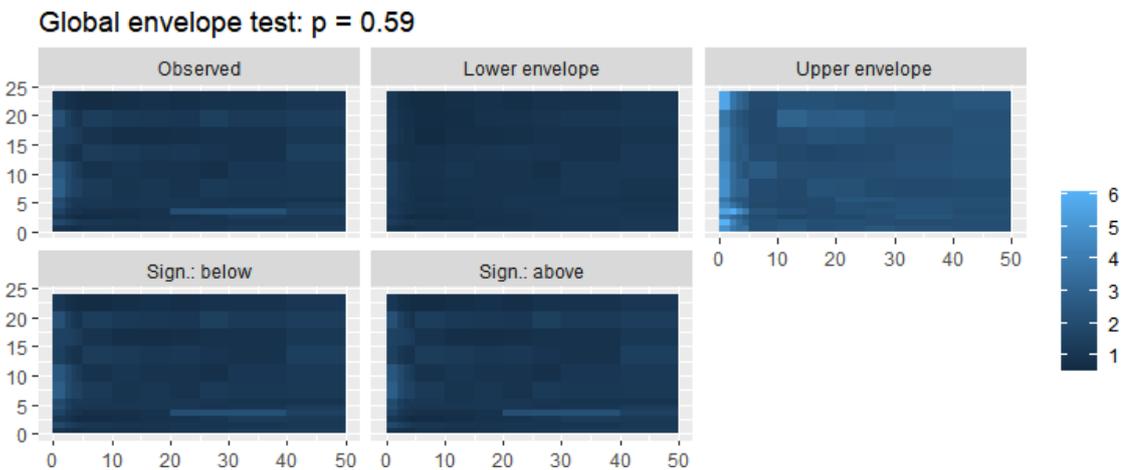


Figure 5.8 Top: estimated pair correlation function \hat{g}_{obs} , lower E_L and upper E_U bounds of the 99% global rank envelope (ERL). Bottom: differences $E_{obs} - E_L$ and $E_U - E_{obs}$. Negative values (if any) are represented in red and lead to reject the fitted model. Values on the horizontal axis are in kilometers and those on the vertical axis are in months.

5.2 Application to local temperature hotspots

We develop an innovative application of spatio-temporal modeling of hotspots of temperature, and in particular of temperature anomalies, in the United States, where we remove Alaska and islands such as Hawaii from the observation window. Based on a space-time temperature dataset available on a regular grid of latitude, longitude and months, we define a hotspot as a spatio-temporal point, i.e., as the combination of a spatial grid cell and a month and year. A hotspot arises if the temperature value observed

at the grid cell is higher than the values in all of the neighboring grid cells during the same month and year. We can either limit the spatial neighborhood of a grid cells to all its adjacent grid cells, or we can use larger neighborhood composed of all the grid cells whose center point lies within a certain distance radius R_0 of the center point of the reference grid cell. Moreover, to focus on hotspots corresponding to relatively large temperatures, we can further add the condition that the temperature value at the grid cell must be higher than a certain quantile of the temperature distribution. Specifically, we aim to extract hotspots that arise in temperature anomalies, such that the occurrence of a hotspot indicates a temperature value that is relatively high with respect to local climatic conditions. Therefore, we apply the hotspot extraction algorithm after a pretransformation of the original temperature values.

5.2.1 Hotspot extraction

This work is based on a gridded reanalysis dataset (at spatial 0.1° resolution, i.e., approximately 10 km) of monthly means of temperatures at 2 m height for the period 1981–2019 spanning 39 years, provided by the ERA5 climate model and postprocessed by the Copernicus Climate Data service of the European Union (Copernicus Climate Data Store 2021). We further use auxiliary variables related to altitude from the Shuttle Radar Topography Mission dataset. We use the temperature grid as basis for point process modeling of hotspots, and we refer to it as the PP grid. To upscale altitude data (at 1 km resolution) to the resolution of the temperature grid, we generate two synthetic variables: empirical mean and standard deviation of the altitude values in each of the PP grid cells.

We denote by $s(i_1, i_2, a, m)$ the temperature value at grid cell (i_1, i_2) , $i_1 = 1, \dots, 250$, $i_2 = 1, \dots, 700$ during year $a = 1, \dots, 39$ and month $m = 1, \dots, 12$. The coordinates of the center of grid cell (i_1, i_2) are denoted by u_{i_1, i_2} . There is a total of $39 \times 12 \times 250 \times 700 = 60900000$ observed values.

Temperature anomalies $\tilde{s}(i_1, i_2, a, m)$ are computed as follows.

$$\tilde{s}(i_1, i_2, a, m) = \frac{s(i_1, i_2, a, m) - \hat{\mu}(i_1, i_2, m)}{\hat{\sigma}(i_1, i_2, m)}, \quad (5.10)$$

where $\hat{\mu}(i_1, i_2, m)$ and $\hat{\sigma}(i_1, i_2, m)$ are empirical means and the standard deviations, respectively, of temperatures, and they are calculated by pooling together the temperature data of the 39 observations years for each configuration (i_1, i_2, m) . By construction, temperature anomalies have approximately mean zero and variance one.

Our algorithm for extracting temperature hotspots proceeds as follows:

INPUTS: Monthly gridded reanalysis data; neighborhood radius $R_0 > 0.1^\circ$; quantile level $p \in [0, 1]$.

1. Estimate

$$\hat{\mu}(i_1, i_2, m) = \frac{1}{39} \sum_{a=1}^{39} s(i_1, i_2, a, m)$$

and

$$\hat{\sigma}(i_1, i_2, m) = \sqrt{\frac{1}{39} \sum_{a=1}^{39} (s(i_1, i_2, a, m) - \hat{\mu}(i_1, i_2, m))^2}.$$

2. Compute anomalies using formula (5.10).
3. Extract all hotspot locations $x_i = (i_1, i_2, a, m)$ satisfying

$$\min_{\tilde{x} \in \mathcal{B}(x_i)} (\tilde{s}(x_i) - \tilde{s}(\tilde{x})) > 0, \quad \text{where } \mathcal{B}(x_i) = \{(\tilde{i}_1, \tilde{i}_2, a, m) \mid \|(\tilde{i}_1, \tilde{i}_2) - (i_1, i_2)\| \leq R_0\}.$$

4. RETURN the space-time point pattern $\{x_i, i = 1, \dots, N\}$.

Steps 1 and 3 in the algorithm can be implemented efficiently using raster representations (Hijmans 2020).

5.2.2 Model fitting and validation

We fit an extension of model (4.1) with $s = \infty$ by considering a spatial hybrid Strauss model (Baddeley et al., 2013) rather than Strauss model. The irregular parameters of the model (e.g. r) are estimated by profile pseudo-likelihood (Baddeley and Turner, 2000). We consider $r = \{0.5, 0.6, 0.7, 0.8, 0.9, 1\}$. The estimated interaction parameters are then $\hat{\gamma} = \{0.12, 0.77, 0.65, 0.93, 0.85, 0.40\}$. We extract the results of the spatial random field. The resulting maps are reported in Fig. 5.9.

We implement a method for model validation based on posterior predictions and comparison with their second-order summary statistics according to the global envelopes. We draw 19 samples from the posterior joint distribution which we interested in the values of the latent field of each sample. We need to query the index for each latent field component. Hence, we can obtain the predicted values for each response. We compute L -function for each sample and also for the point patterns during 1981 to 2019. Fig. 5.10 compares L -function for point patterns in during 1981 to 2019 in August and mean of L -function for 19 samples from the posterior joint distribution of fitted LGCSP and confirms that the fitted LGCSP captures the behavior of the point pattern very well.

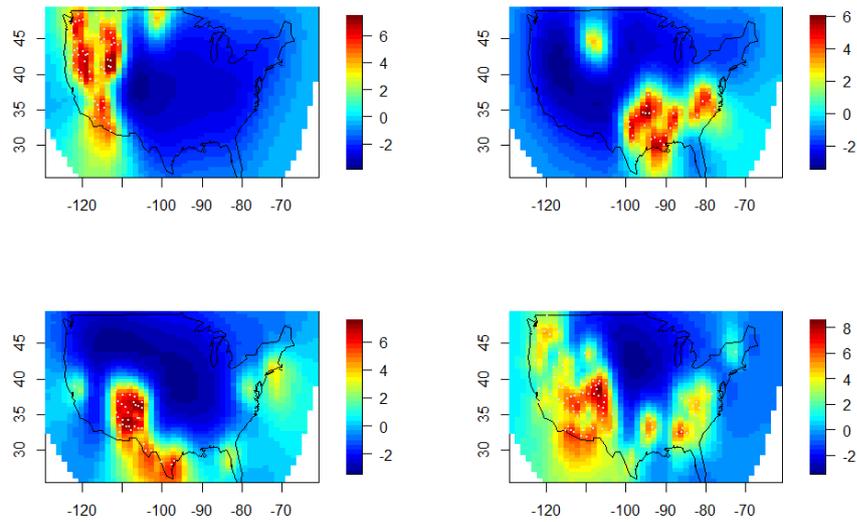


Figure 5.9 Posterior mean of the spatial random fields for August of 1981 (top-left), 1999 (top-right), 2009 (bottom-left), and 2019 (bottom-right) .

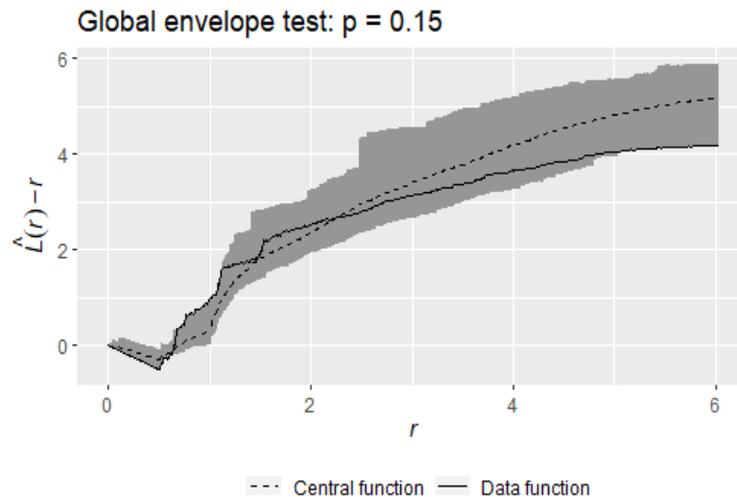


Figure 5.10 Combined global envelopes based on the empirical L -function for LGCSP fitted by our INLA-based approach. The solid curve is the empirical functional summary statistics for the observed point pattern and the dashed curve is the means obtained from 19 posterior predictions. Each shaded area indicates a 95% global envelope based on the extreme rank length. At the top of each plot, the p -value of the corresponding global envelope test is stated.

Chapter 6

Conclusion and future work

6.1 Conclusion

A spatial (and spatio-temporal) point pattern, as a realization of a point process, is a collection of events for which locations (and times) of occurrence have been observed in a specified spatial region (and temporal period). Point patterns are often classified into three classes of single interaction structure: randomness, clustering, and inhibition that can be modeled for instance by Poisson process, Cox processes, and Gibbs processes, respectively. These single-structure point process models can be too simplistic to describe some complex phenomena in seismology, epidemiology, and forestry as they involve several structures at different spatial (and spatio-temporal) scales, thus requiring multi-structure point processes to describe them. The main concern of this Ph.D. thesis is the spatio-temporal modeling of such complex point patterns taking into account the spatio-temporal inhomogeneity driven by covariates and the complexity of their interaction structures at different scales.

In the spatial point processes literature, three general approaches are considered for constructing multi-structure point process models: thinning, superposition, and hybridization. The key contribution of the Ph.D. thesis is to introduce spatio-temporal hybrid point processes based on Gibbs and Cox models using hybridization approach and to develop their statistical inference.

Extending hybridization approach to the spatio-temporal framework, we develop new hybrid Gibbs models, namely the spatio-temporal hybrid Geyer saturation point process and spatio-temporal hybrid Strauss hardcore point process, that combine both multi-scaling and hardcore distances. A different approach, leading to more flexibility in the model and challenging inference, consists of Gibbs models that contain both random and fixed effects to take into account complex patterns of heterogeneity. We propose to embed spatio-temporally structured Gaussian random effects in the Gibbs trend function. Therefore, this approach focuses on models derived from the multi-scale classes of combinations of Gibbs and log-Gaussian Cox point processes which we refer to as Cox-Gibbs models.

We also extend to the spatio-temporal framework, and implement, inference meth-

ods for these new models. In this Ph.D. thesis we classify the inference procedure into two approaches: global and local estimation methods. For the global statistical inference of Gibbs models we tailor the pseudo-likelihood and logistic likelihood approaches. Because, the models based on global parameter estimates can not take into account different local interaction structure, we extend the local likelihood approach to the spatio-temporal context. It can be viewed as an alternative method for modeling multi-structure point patterns with spatially and/or temporally varying parameters in Gibbs point process models. Finally, for Cox-Gibbs models, the calculation of the likelihood variants (composite likelihoods: pseudo-likelihood and logistic likelihood) would involve complex, high-dimensional integrals, and we would need estimation methods that allow handling the latent Gaussian variables. Due to the hierarchical structure of Cox-Gibbs models, we can formulate and estimate them within a Bayesian hierarchical approach, using techniques as the Integrated Nested Laplace Approximation.

We propose to simulate the hybrid Gibbs point process models with a birth-death Metropolis-Hasting algorithm. We develop a two-step procedure for simulating the hybrid Cox-Gibbs model by simulating, firstly, a realisation of a Gaussian random field and then simulating a realisation of hybrid Cox-Gibbs model given that Gaussian realisation using the birth-death Metropolis-Hasting algorithm. Estimation, simulation and validation methods proposed in the Ph.D. thesis have been carried out using R together with the `spatstat`, `stpp`, `splanes`, `fields`, `sparr`, `raster`, `INLA` and `GET` packages.

Finally we use these models to describe the complex interaction structure observed in different data sets. We particularly focus on wildfires in France and in Spain using hybrid Gibbs models. The spatio-temporal distribution of forest fires is very complex in nature with non-separability in space and time and multiple structures (repulsion and aggregation) at different spatial and/or temporal scales. Spatio-temporal variations of fire occurrences further depend on the spatial distribution of current land use and meteorological conditions, but also depends on past events (changes in vegetation due to fires affect the probability of fire occurrence during a regeneration period). We also develop an innovative application of spatio-temporal modeling of hotspots of temperature, and in particular of temperature anomalies, in the United States using Cox-Gibbs models.

6.2 Future work

We defined two new spatio-temporal multi-scale Gibbs point process models and detailed extensions of classical statistical inference methods and MCMC simulation techniques to the spatio-temporal framework. Simulation algorithms and inference methods are implemented in R code¹ and will be added to the `stpp` package (Gabriel et al.

¹<http://edith.gabriel.pagesperso-orange.fr/software.html>

2013).

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 forest fire occurrences application study in France (resp. Spain), 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 (resp. grid) scale and at the yearly (resp. monthly) 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. Indeed, when fitting inhomogeneous Gibbs models to patterns by this two-step procedure there is the problem of deciding which covariates should be included in the final model. Recently, Choiruddin et al. (2018) and Daniel et al. (2018) presented a general framework for regularizing inhomogeneous spatial Gibbs point process models via penalized composite likelihoods, incorporating both the pseudo-likelihood and logistic likelihood approaches to model fitting. Regularization is an attractive procedure that performs variable selection and parameter estimation simultaneously by maximizing a penalized likelihood function. This methodology could be extended in several ways, first, extension to the spatio-temporal Gibbs models and proposing the penalized logistic likelihood in a spatio-temporal framework. Second, regularization could be applied on the spatio-temporal Gibbs models with more complex interaction terms, such as multi-scale and hybrid interactions as we developed in this thesis.

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.

Our method for choosing irregular parameters in hybrid Geyer model (resp. hybrid Strauss hardcore model) relies on 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, 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 , $j = 1, \dots, m$ (resp. m , r_j , q_j , $j = 1, \dots, m$). 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 models could further be applied to describe complex point processes in other fields, like seismology and epidemiology for example, because several mechanisms exhibit interaction between points at multiple scales in space and time.

Bibliography

Albert-Green, A. (2016), Joint Models for Spatial and Spatio-Temporal Point Processes, PhD thesis, University of Western Ontario, Canada.

Albert-Green, A., Braun, W. J., Dean, C. B. and Miller, C. (2019), A hierarchical point process with application to storm cell modelling, *Canadian Journal of Statistics* **47**(1), 46–64.

Ambler, G. K. (2002), Dominated Coupling from the Past and Some Extensions of the Area-Interaction Process, PhD thesis, University of Bristol, UK.

Ambler, G. K. and Silverman, B. (2004), Perfect simulation of spatial point processes using dominated coupling from the past with application to a multiscale area-interaction point process. University of Bristol, Department of Mathematics, Bristol. Available at <http://www.stats.ox.ac.uk/~silverma/pdf/amblersilverman1.pdf>.

Ambler, G. K. and Silverman, B. (2010), Perfect simulation using dominated coupling from the past with application to area-interaction point processes and wavelet thresholding, in N. H. Bingham and C. M. Goldie, eds, *Probability and Mathematical Genetics*, Cambridge University Press, Cambridge, pp. 64–90.

Andersen, I. T. and Hahn, U. (2016), Matérn thinned Cox processes, *Spatial Statistics* **15**, 1–21.

Anwar, S. and Stein, A. (2015), Spatial pattern development of selective logging over several years, *Spatial Statistics* **13**, 90–105.

Arago, P., Juan, P., Diaz-Avalos, C. and Salvador, P. (2016), Spatial point process modeling applied to the assessment of risk factors associated with forest wildfires incidence in Castellon, Spain, *European Journal of Forest Research* **135**(3), 451–464.

Baddeley, A. (2017), Local composite likelihood for spatial point processes, *Spatial Statistics* **22**, 261–295.

Baddeley, A., Coeurjolly, J. F., Rubak, E. and Waagepetersen, R. (2014), Logistic regression for spatial Gibbs point processes, *Biometrika* **101**(2), 377–392.

- Baddeley, A., Rubak, E. and Turner, R. (2015), *Spatial Point Patterns: Methodology and Applications with R*, Chapman and Hall/CRC Press, London.
- Baddeley, A., Rubak, E. and Turner, R. (2019), Leverage and influence diagnostics for Gibbs spatial point processes, *Spatial Statistics* **29**, 15–48.
- Baddeley, A. and Turner, R. (2000), Practical maximum pseudolikelihood for spatial point patterns (with discussion), *Australian and New Zealand Journal of Statistics* **42**, 283–322.
- Baddeley, A. and Turner, R. (2005), spatstat: a R package for analyzing spatial point patterns, *Journal of Statistical Software* **12**(6), 1–42.
- Baddeley, A. and Turner, R. (2006), Modelling spatial point patterns in R, in A. Baddeley, P. Gregori, J. Mateu, R. Stoica and D. Stoyan, eds, *Case Studies in Spatial Point Process Modeling*. Lecture Notes in Statistics, Vol. 185, Springer, New York, pp. 23–74.
- Baddeley, A., Turner, R., Mateu, J. and Bevan, A. (2013), Hybrids of Gibbs point process models and their implementation, *Journal of Statistical Software* **55**(11), 1–43.
- Baddeley, A., Turner, R. and Rubak, E. (2016), Adjusted composite likelihood ratio test for spatial Gibbs point processes, *Journal of Statistical Computation and Simulation* **86**(5), 922–941.
- Badreldin, N., Uria-Diez, J., Mateu, J., Youssef, A., Stal, C., El-Bana, M., Magdy, A. and Goossens, R. (2015), A spatial pattern analysis of the halophytic species distribution in an arid coastal environment, *Environmental Monitoring and Assessment* **187**, 1–15.
- Baile, R., Muzy, J. F. and Silvani, X. (2021), Multifractal point processes and the spatial distribution of wildfires in French Mediterranean regions, *Physica A: Statistical Mechanics and its Applications* **568**, 125697.
- Bartlett, M. S. (1954), Processus stochastiques ponctuels, **16**, 35–60.
- Bartlett, M. S. (1955), *An Introduction to Stochastic Processes*, Cambridge University Press, Cambridge.
- Bates, B. C., McCaw, L. and Dowdy, A. J. (2018), Exploratory analysis of lightning-ignited wildfires in the Warren Region, Western Australia, *Journal of Environmental Management* **225**, 336–345.
- Berman, M. and Turner, R. (1992), Approximating point process likelihoods with GLIM, *Applied Statistics* **41**(1), 31–38.

- Besag, J. (1977), Some methods of statistical analysis for spatial data, *Bulletin of the International Statistical Institute* **47**, 77–92.
- Borrajo, M. I., Fuentes-Santos, I. and González-Manteiga, W. (2018), Nonparametric first-order analysis of spatial and spatio-temporal point processes, in M. L. Rocca, B. Liseo and L. Salmaso, eds, *Nonparametric Statistics*, Vol. 4, Springer, Italy, pp. 101–112.
- Borrajo, M. I., González-Manteiga, W. and Martínez-Miranda, M. (2017), Testing first-order intensity model in non-homogeneous Poisson point processes with covariates. arXiv:1709.07716.
- Borrajo, M. I., González-Manteigaand, W. and Martínez-Miranda, M. D. (2020a), Bootstrapping kernel intensity estimation for inhomogeneous point processes with spatial covariates, *Computational Statistics and Data Analysis* **144**, 106875.
- Borrajo, M. I., González-Manteigaand, W. and Martínez-Miranda, M. D. (2020b), Testing for significant differences between two spatial patterns using covariates, *Spatial Statistics* **40**, 100379.
- Bosq, D. (1998), *Nonparametric Statistics for Stochastic Processes*, Springer-Verlag New York, New York.
- Brillinger, D. R., Haiganoush, K. P. and Benoit, J. W. (2003), Risk assessment: a forest fire example, in D. R. Goldstein, ed., *Statistics and Science: A Festschrift for Terry Speed*, Institute of Mathematical Statistics, Beachwood, pp. 177–196.
- Brix, A. and Chadœuf, J. (2000), Spatio-temporal modeling of weeds and shot-noise G Cox processes, *Biometrical Journal* **44**, 83–99.
- Brix, A. and Diggle, P. J. (2001), Spatiotemporal prediction for log-Gaussian Cox processes, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **63**(4), 823–841.
- Brix, A. and Kendall, W. S. (2002), Simulation of cluster point processes without edge effects, *Advances in Applied Probability* **34**, 267–280.
- Brix, A. and Møller, J. (2001), Space-time multitype log Gaussian Cox processes with a view to modelling weed data, *Journal Scandinavian Journal of Statistics* **28**, 471–488.
- Chiu, S. N., Stoyan, D., Kendall, W. S. and Mecke, J. (2013), *Stochastic Geometry and its Applications*, 3 edn, Wiley, Chichester.

- Choiruddin, A., Coeurjolly, J. F. and Letué, F. (2018), Convex and non-convex regularization methods for spatial point processes intensity estimation, *Electronic Journal of Statistics* **12**(1), 1210–1255.
- Clyde, M. and Strauss, D. (1991), Logistic regression for spatial pair-potential models, *Lecture Notes-Monograph Series* **20**, 14–30.
- Coeurjolly, J. F. and Lavancier, F. (2019), Understanding spatial point patterns through intensity and conditional intensities, in D. Coupier, ed., *Stochastic Geometry*, Springer, pp. 45–85.
- Coeurjolly, J. F., Møller, J. and Waagepetersen, R. P. (2017), A tutorial on palm distributions for spatial point processes, *International Statistical Review* **85**(3), 404–420.
- Comas, C., Costafreda-Aumedes, S., López, N. and Vega-Garcia, C. (2019), On the correlation structure between point patterns and linear networks, *Spatial Statistics* **29**, 192–203.
- Comas, C., Costafreda-Aumedes, S. and Vega-Garcia, C. (2014), Characterizing configurations of fire ignition points through spatiotemporal point processes, *Natural Hazards and Earth System Sciences* **2**, 2891–2911.
- Copernicus Climate Data Store (2021), Agrometeorological indicators from 1979 to present derived from reanalysis. doi:10.24381/cds.6c68c9bb.
- Costafreda-Aumedes, S., Comas, C. and Vega-Garcia, C. (2016), Spatio-temporal configurations of human-caused fires in Spain through point patterns, *Forests* **7**(9), 185.
- Cox, D. R. (1955), Some statistical methods connected with series of events (with discussion), *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **17**, 129–164.
- Cox, D. R. (1962), *Renewal Theory*, Methuen, London.
- Cox, D. R. (1972), The statistical analysis of dependencies in point processes, in P. Lewis, ed., *Stochastic Point Processes*, Wiley, New York, pp. 55–66.
- Cox, D. R. and Isham, V. (1980), *Point Processes*, Chapman and Hall, London.
- Cronie, O. and van Lieshout, M. (2018), A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions, *Biometrika* **105**, 455–462.
- Cronie, O. and van Lieshout, M. N. M. (2015), A J-function for inhomogeneous spatio-temporal point processes, *Scandinavian Journal of Statistics* **42**(2), 562–579.

- Daley, D. J. and Vere-Jones, D. (2003), *An Introduction to the Theory of Point Processes*, Vol. I, Springer-Verlag, New York.
- Daniel, J., Horrocks, J. and Umphrey, G. J. (2018), Penalized composite likelihoods for inhomogeneous Gibbs point process models, *Computational Statistics and Data Analysis* **124**, 104–116.
- Davies, T., Hazelton, M. and Marshall, J. (2011), `sparr`: analyzing spatial relative risk using fixed and adaptive kernel density estimation in R, *Journal of Statistical Software* **39**(1), 1–14.
- Davies, T., Marshall, J. and Hazelton, M. (2018), Tutorial on kernel estimation of continuous spatial and spatiotemporal relative risk with accompanying instruction in R, *Statistics in Medicine* **37**(7), 1191–1221.
- Dayananda, P. W. A. (1977), Stochastic models for forest fires, *Ecological Modelling* **3**, 309–313.
- del Hoyo, L. V., Isabel, M. P. M. and Vega, F. J. M. (2011), Logistic regression models for human-caused wildfire risk estimation: analysing the effect of the spatial accuracy in fire occurrence data, *European Journal of Forest Research* **130**, 983–996.
- Dereudre, D. (2019), Introduction to the theory of Gibbs point processes, in D. Coupier, ed., *Stochastic Geometry*, Springer, pp. 181–229.
- Diaz-Avalos, C., Juan, P. and Mateu, J. (2013), Similarity measures of conditional intensity functions to test separability in multidimensional point processes, *Stochastic Environmental Research and Risk Assessment* **27**(5), 1193–1205.
- Diaz-Avalos, C., Juan, P. and Mateu, J. (2014), Significance tests for covariate-dependent trends in inhomogeneous spatio-temporal point processes, *Stochastic Environmental Research and Risk Assessment* **28**, 593–609.
- Diaz-Avalos, C., Juan, P. and Serra, L. (2016), Modeling fire size of wildfires in Castellon (Spain), using spatiotemporal marked point processes, *Forest Ecology and Management* **381**, 360–369.
- Diggle, P. J. (1983), *Statistical Analysis of Spatial Point Patterns*, Academic Press, London.
- Diggle, P. J. (2003), *Statistical Analysis of Spatial Point Patterns*, Vol. 2, Hodder Arnold, London.
- Diggle, P. J. (2013), *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, Vol. 3, Chapman and Hall/CRC, New York.

- Diggle, P. J. and Gabriel, E. (2010), Spatio-temporal point processes, *in* Handbook of Spatial Statistics, Chapman and Hall/CRC Handbooks of Modern Statistical Methods, pp. 449–461.
- Diggle, P. J., Moraga, P., Rowlingson, B. and Taylor, B. M. (2013), Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm, *Statistical Science* **28**(4), 542–563.
- Ehrgott, M. (2005), *Multicriteria Optimization*, Springer Verlag, Berlin.
- Feller, W. (1950), *An Introduction to Probability Theory and Its Applications*, Wiley, New York.
- Fuentes-Santos, I., González-Manteiga, W. and Mateu, J. (2016), Consistent smooth bootstrap kernel intensity estimation for inhomogeneous spatial Poisson point processes, *Scandinavian Journal of Statistics* **43**(2), 416–435.
- Fuentes-Santos, I., González-Manteiga, W. and Mateu, J. (2017), A nonparametric test for the comparison of first-order structures of spatial point processes, *Spatial Statistics* **22**, 240–260.
- Fuentes-Santos, I., González-Manteiga, W. and Mateu, J. (2018), A first-order ratio-based nonparametric separability test for spatio-temporal point processes, *Environmetrics* **29**(1), 1–18.
- Fuentes-Santos, I., González-Manteiga, W. and Mateu, J. (2021), Testing similarity between first-order intensities of spatial point processes. a comparative study, *Communications in Statistics - Simulation and Computation* . doi: 10.1080/03610918.2021.1901118.
- Fuentes-Santos, I., Marey-Pérez, M. and González-Manteiga, W. (2013), Forest fire spatial pattern analysis in Galicia (NW Spain), *Journal of Environmental Management* **128**, 30–42.
- Gabriel, E. (2014), Estimating second-order characteristics of inhomogeneous spatio-temporal point processes, *Methodology and Computing in Applied Probability* **16**(2), 411–431.
- Gabriel, E. and Diggle, P. J. (2009), Second-order analysis of inhomogeneous spatio-temporal point process data, *Statistica Neerlandica* **63**(1), 43–51.
- Gabriel, E., Opitz, T. and Bonneau, F. (2017), Detecting and modeling multi-scale space-time structures: the case of wildfire occurrences, *Journal of the French Statistical Society* **158**(3), 86–105.

- Gabriel, E., Rowlingson, B. and Diggle, P. J. (2013), `stpp`: a R package for plotting, simulating and analyzing spatio-temporal point patterns, *Journal of Statistical Software* **53**(2), 1–29.
- Ganteaume, A. and Jappiot, M. (2013), What causes large fires in Southern France, *Forest Ecology and Management* **294**, 76–85.
- Genton, M., Butry, D., Gumpertz, M. and Prestemon, J. (2006), Spatio-temporal analysis of wildfire ignitions in the St Johns River water management district, Florida, *International Journal of Wildland Fire* **15**, 87–97.
- Geyer, C. J. (1999), Likelihood inference for spatial point processes, in O. Barndorff-Nielsen, W. Kendall and M. van Lieshout, eds, *Stochastic Geometry: Likelihood and Computation*, Vol. 80, Chapman and Hall, London, pp. 79–140.
- Geyer, C. J. and Møller, J. (1994), Simulation procedures and likelihood inference for spatial point processes, *Scandinavian Journal of Statistics* **18**, 505–544.
- Goldstein, J., Haran, M., Simeonov, I., Fricks, J. and Chiaromonte, F. (2015), An attraction-repulsion point process model for respiratory syncytial virus infections, *Biometrics* **71**(2), 376–385.
- Gomez-Rubio, V. (2020), *Bayesian Inference with INLA*, Chapman and Hall-CRC, Boca Raton.
- Gonzalez, J. A., Rodriguez-Cortes, F. J., Cronie, O. and Mateu, J. (2016), Spatio-temporal point process statistics: A review, *Spatial Statistics* **18**, 505–544.
- Gonzalez-Olabarria, J. R., Mola, B. and Coll, L. (2015), Different factors for different causes: Analysis of the spatial aggregations of fire ignitions in Catalonia (Spain), *Risk Analysis* **35**, 1197–1209.
- Gonzalez-Olabarria, J. R., Mola-Yudego, B., Pukkala, T. and Palahi, M. (2011), Using multi-scale spatial analysis to assess fire ignition density in Catalonia, Spain, *Annals of Forest Science* **68**, 861–871.
- Gralewicz, N. J., Nelson, T. A. and Wulder, M. A. (2011), Spatial and temporal patterns of wildfire ignitions in Canada from 1980 to 2006, *International Journal of Wildland Fire* **21**(3), 230–242.
- Gua, F., Hu, H., Sun, L. and Ma, Z. (2011), Spatial patterns of lightning-ignited forest fires in daxing’an mountains, Heilongjiang Province, China, 1973–1997, *Advanced Materials Research* **183-185**, 2268–2274.

- Gua, F., Innes, J., Wang, G., Ma, X., Sun, L., Hu, H. and Su, Z. (2015), Historic distribution and driving factors of human-caused fires in the Chinese boreal forest between 1972 and 2005, *Journal of Plant Ecology* **8**(5), 480–490.
- Gua, F., Su, Z., Wang, G., Sun, L., Lin, F. and Liu, A. (2016), Wildfire ignition in the forests of southeast China: identifying drivers and spatial distribution to predict wildfire likelihood, *Applied Geography* **66**, 12–21.
- Gua, F., Su, Z., Wang, G., Sun, L., Tigabu, M., Yang, X. and Hu, H. (2017), Understanding fire drivers and relative impacts in different Chinese forest ecosystems, *The Science of the Total Environment* **606**, 411–425.
- Habel, H., Sarkka, A., Rudemo, M., Blomqvist, C. H., Olsson, E. and Nordin, M. (2019), Colloidal particle aggregation in three dimensions, *Journal of Microscopy* **275**(3), 149–158.
- Hahn, U., Jense, A. B. V., Rudemo, M., Blomqvist, C. H., Olsson, E. and Nordin, M. (2015), Hidden second-order stationary spatial point processes, *Scandinavian Journal of Statistics* **43**(2), 455–475.
- Hawkes, A. G. (1971), Spectra of some self-exciting and mutually exciting point processes, *Biometrika* **58**, 83–90.
- Hering, A. S., Hering, C. L. and Genton, M. G. (2009), Modeling spatio-temporal wildfire ignition point patterns, *Environmental and Ecological Statistics* **16**, 225–250.
- Hijmans, R. J. (2020), *raster*: Geographic data analysis and modeling. R package version 3.1-5. <https://CRAN.R-project.org/package=raster>.
- Iftimi, A., Montes, F., Mateu, J. and Ayyad, C. (2017), Measuring spatial inhomogeneity at different spatial scales using hybrids of Gibbs point process models, *Stochastic Environmental Research and Risk Assessment* **31**(6), 1455–1469.
- Iftimi, A., van Lieshout, M. C. and Montes, F. (2018), A multi-scale area-interaction model for spatio-temporal point patterns, *Spatial Statistics* **26**, 38–55.
- Illian, J. and Hendrichsen, D. (2010), Gibbs point process models with mixed effects, *Environmetrics* **21**, 341–353.
- Illian, J., Martino, S., Sørbye, S., Gallego-Fernandez, J. B., Zunzunegui, M., Esquivias, M. P. and Travis, J. M. (2013), Fitting complex ecological point process models with integrated nested Laplace approximation, *Methods in Ecology and Evolution* **4**, 305–315.

- Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008), *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley and Sons, Chichester.
- Illian, J., Sørbye, S. and Rue, H. (2012a), A toolbox for fitting complex spatial point process models using integrated nested laplace approximation (INLA), *Annals of Applied Statistics* **6**(4), 1499–1530.
- Illian, J., Sørbye, S., Rue, H. and Hendrichsen, D. (2012b), Using INLA to fit a complex point process model with temporally varying effects - a case study, *Journal of Environmental Statistics* **3**, 1–25.
- Juan, P. (2019), Enhancing the SPDE modeling of spatial point processes with INLA, applied to wildfires. Choosing the best mesh for each database, *Communications in Statistics - Simulation and Computation* . doi: 10.1080/03610918.2019.1618473.
- Juan, P. (2020), Spatio-temporal hierarchical Bayesian analysis of wildfires with Stochastic Partial Differential Equations. A case study from Valencian Community (Spain), *Journal of Applied Statistics* **47**(5), 927–946.
- Juan, P., Mateu, J. and Díaz-Avalos, C. (2010), Characterizing spatial-temporal forest fire patterns, METMA V: International Workshop on Spatio-Temporal Modelling, Santiago de Compostela, Spain.
- Juan, P., Mateu, J. and Saez, M. (2012), Pinpointing spatio-temporal interactions in wildfire patterns, *Stochastic Environmental Research and Risk Assessment* **26**(8), 1131–1150.
- Kingman, J. F. C. (1993), *Poisson Processes*, Vol. 3, Oxford University Press, Oxford.
- Kingman, J. F. C. (2006), Poisson processes revisited, *Probability and Mathematical Statistics* **26**(1), 77–95.
- Koh, J., Pimont, F., Dupuy, J.-L. and Opitz, T. (2021), Spatiotemporal wildfire modeling through point processes with moderate and extreme marks. arXiv:2105.08004.
- Koutsias, N., Allgöwer, B., Kalabokidis, K., Mallinis, G., Balatsos, P. and Goldammer, J. G. (2015), Fire occurrence zoning from local to global scale in the European Mediterranean basin: implications for multi-scale fire management and policy, *iForest - Biogeosciences and Forestry* **9**, 195–204.
- Koutsias, N., Balatsos, P. and Kalabokidis, K. (2014), Fire occurrence zones: kernel density estimation of historical wildfire ignitions at the national level, Greece, *Journal of Maps* **10**(4), 630–639.

- Koutsias, N., Kalabokidis, K. and Allgöwer, B. (2004), Fire occurrence patterns at landscape level: beyond positional accuracy of ignition points with kernel density estimation methods, *Natural Resource Modeling* **17**(4), 359–375.
- Kwak, H., Lee, W. K., Lee, S. Y., Won, M. S., Koo, K. S., Lee, B. and Lee, M. B. (2010), Cause-specific spatial point pattern analysis of forest fire in Korea, *Journal of Korean Forest Society* **99**(3), 259–266.
- Kwak, H., Lee, W. K., Won, M. S., Koo, K. S., Lee, M. B. and Lee, S. H. (2009), Spatial and temporal pattern of the human-caused forest fire occurrences in Korea, in *Proceeding of the ESRI user conference, San Diego*, pp. 662–673.
- Lavancier, F. and Møller, J. (2016), Modelling aggregation on the large scale and regularity on the small scale in spatial point pattern datasets, *Scandinavian Journal of Statistics* **43**, 587–609.
- Lavancier, F., Møller, J. and Rubak, E. (2015), Determinantal point process models and statistical inference, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **77**, 853–877.
- Levin, C. A. (1992), The problem of pattern and scale in ecology: The Robert H. Macarthur award lecture, *Ecology* **73**, 1943–1967.
- Li, S. and Banerjee, T. (2020), Spatial and temporal patterns of wildfires in California, *Earth and Space Science Open Archive ESSOAr*. doi: 10.1002/essoar.10504419.1.
- Lindgren, F. and Rue, H. (2015), Bayesian spatial modelling with R-INLA, *Journal of Statistical Software* pp. 1–25.
- Lindgren, F., Rue, H. and Lindstrom, J. (2011), An explicit link between Gaussian fields and Gaussian Markov random fields the SPDE approach, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **73**(4), 423–498.
- Lindsay, B. (1988), Composite likelihood methods, in N. Prabhu, ed., *Statistical Inference from Stochastic Processes*, number 80, American Mathematical Society, pp. 221–239.
- Liu, Z., Yang, J., Chang, Y., Weisberg, P. J. and He, H. S. (2012), Spatial patterns and drivers of fire occurrence and its future trend under climate change in a boreal forest of Northeast China, *Global Change Biology* **18**, 2041–2056.
- Loader, C. (1999), *Local Regression and Likelihood*, Springer Verlag, New York.
- Macchi, O. (1975), The coincidence approach to stochastic point processes, *Advances in Applied Probability* **7**, 83–122.

- Matérn, B. (1960), *Spatial Variation. Lectures Notes in Statistics*, Springer-Verlag, Chichester.
- Miranda, B. R., Sturtevant, B. R., Stewart, S. I. and Hammer, R. B. (2012), Spatial and temporal drivers of wildfire occurrence in the context of rural development in northern Wisconsin, USA, *International Journal of Wildland Fire* **21**, 141–154.
- Møller, J. (2003), Shot noise Cox processes, *Advances in Applied Probability* **35**(3), 614–640.
- Møller, J. and Diaz-Avalos, C. (2010), Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires, *Scandinavian Journal of Statistics* **37**(1), 2–25.
- Møller, J., Syversveen, A. R. and Waagepetersen, R. P. (1998), Log Gaussian Cox processes, *Scandinavian Journal of Statistics* **25**(3), 451–482.
- Møller, J. and Torrisi, G. L. (2005), Generalised shot noise Cox processes, *Advances in Applied Probability* **37**, 48–74.
- Møller, J. and Waagepetersen, R. P. (2004), *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.
- Mundo, I. A., Wiegand, T., Kanagaraj, R. and Kitzberger, T. (2013), Environmental drivers and spatial dependency in wildfire ignition patterns of northwestern Patagonia, *Journal of Environmental Management* **123**, 77–87.
- Myllymäki, M., Kuronen, M. and Mrkvička, T. (2021), Testing global and local dependence of point patterns on covariates in parametric models, *Spatial Statistics* **42**, 100436.
- Myllymäki, M. and Mrkvička, T. (2019), GET: Global envelopes in R. arXiv:1911.06583.
- Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H. and Hahn, U. (2017), Global envelope tests for spatial processes, *Journal of the Royal Statistical Society: Series B* **79**, 381–404.
- Najafabadi, A. T. P., Gorgani, F. and Najafabadi, M. O. (2015), Modeling forest fires in Mazandaran Province, Iran, *Journal of Forestry Research* **26**, 851–858.
- Neyman, J. and Scott, E. L. (1958), Statistical approach to problems of cosmology, *Journal of the Royal Statistical Society. Series B (Methodological)* **20**, 1–29.

- Nightingale, G. F., Illian, J., King, R. and Nightingale, P. (2019), Area interaction point processes for bivariate point patterns in a Bayesian context, *Journal of Environmental Statistics* **9**(2), 1–22.
- Nychka, D., Furrer, R., Paige, J. and Sain, S. (2017), `fields`: Tools for spatial data. R package version 11.6, <https://github.com/NCAR/Fields>.
- Ogata, Y. and Tanemura, M. (1981), Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure, *Annals of the Institute of Statistical Mathematics* **33**, 315–338.
- Opitz, T. (2009), Simulating and fitting of Gibbs processes in 3D-models, algorithms and their implementation, Masters thesis, Ulm University, Germany.
- Opitz, T. (2017), Latent Gaussian modeling and INLA: A review with focus on space-time applications, *Journal of the French Statistical Society (Special Issue on Space-Time Statistics)* **158**(3).
- Opitz, T., Bonneu, F. and Gabriel, E. (2020), Point-process based Bayesian modeling of space-time structures of forest fire occurrences in Mediterranean France, *Spatial Statistics* **40**, 100429.
- Palm, C. (1943), Intensitätsschwankungen im Fernsprechverkehr, *Ericsson Technics* **44**, 1–189.
- Papangelou, F. (1974), The conditional intensity of general point processes and an application to line processes, *Probability Theory and Related Fields* **28**(3), 207–226.
- Pei, T., Gao, J., Ma, T. and Zhou, C. (2012), Multi-scale decomposition of point process data, *GeoInformatica* **16**(4), 625–652.
- Peng, R. D., Schoenberg, F. P. and Woods, J. (2005), A space-time conditional intensity model for evaluating a wildfire hazard index, *Journal of the American Statistical Association* **100**(469), 26–35.
- Penttinen, A. (1984), *Modelling Interaction in Spatial Point Patterns: Parameter Estimation by the Maximum Likelihood Method*, University of Jyväskylä, Finland.
- Pereira, J. and Turkman, K. F. (2018), Statistical models of vegetation fires-spatial and temporal patterns, in A. E. Gelfand, M. Fuentes, J. A. Hoeting and R. L. Smith, eds, *Handbook of Environmental and Ecological Statistics*, Chapman and Hall/CRC, Boca Raton, chapter 17, pp. 401–420.

- Pereira, P., Turkman, K., Amaral-Turkman, M., Sa, A. and Pereira, J. (2013), Quantification of annual wildfire risk; a spatio-temporal point process approach, *Statistica* **73**(1), 55–68.
- Pereira, P. and Turkman, K. F. (2012), Preliminary analysis of the forest fires in Portugal using point processes, Technical report, CEAUL, Lisboa.
- Picard, N., Bar-Hen, A., Mortier, F. and Chadoeuf, J. (2009), The multi-scale marked area-interaction point process: a model for the spatial pattern of trees, *Scandinavian Journal of Statistics* **36**, 23–41.
- Pimont, F., Fargeon, H., Opitz, T., Ruffault, J., Barbero, R., Martin-StPaul, N., Rigolot, E., Riviere, M. and Dupuy, J.-L. (2021), Prediction of regional wildfire activity in the probabilistic bayesian framework of firelihood, *Ecological Applications* . doi: 10.1002/eap.2316.
- Podur, J., Martell, D. L. and Csillag, F. (2003), Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998, *Ecological Modelling* **164**(1), 1–20.
- Preston, C. J. (1976), *Random Fields*, Springer-Verlag, Berlin.
- Quinlan, J. J., Díaz-Avalos, C. and Mena, R. H. (2021), Modeling wildfires via marked spatio-temporal Poisson processes, *Environmental and Ecological Statistics* . doi: 10.1007/s10651-021-00497-1.
- R Core Team (2016), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <https://www.R-project.org>.
- Raeisi, M. (2018), On models for complex spatio-temporal point process data, Masters thesis, Sorbonne University, France.
- Raeisi, M., Bonneu, F. and Gabriel, E. (2019), On spatial and spatio-temporal multi-structure point process models, *Les Annales de IISUP* **63**(2–3), 97–114.
- Raeisi, M., Bonneu, F. and Gabriel, E. (2021a), A spatio-temporal hybrid strauss hard-core point process for forest fire occurrences, *hal-03193464* .
- Raeisi, M., Bonneu, F. and Gabriel, E. (2021b), A spatio-temporal multi-scale model for Geyer saturation point process: Application to forest fire occurrences, *Spatial Statistics* **41**, 100492.
- Rajala, T., Murrell, D. J. and Olhede, S. C. (2018), Detecting multivariate interactions in spatial point patterns with gibbs models and variate selection, *Journal of the Royal Statistical Society, Applied Statistics, Series C* **67**(5), 1237–1273.

- Rihan, W., Zhao, J., Zhang, H., Guo, X., Ying, H., Deng, G. and Li, H. (2019), Wildfires on the Mongolian Plateau: identifying drivers and spatial distributions to predict wildfire probability, *Remote Sensing* **11**(20), 2361.
- Ripley, B. D. and Kelly, F. P. (1977), Markov point processes, *Journal of the London Mathematical Society* **15**(1), 188–192.
- Rowlingson, B. and Diggle, P. (1993), Splancs: spatial point pattern analysis code in S-PLUS, *Computers and Geosciences* **19**, 627–655.
- Rue, H., Martino, S. and Chopin, N. (2009), Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion), *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **71**(2), 319–392.
- Ruelle, D. (1969), *Statistical Mechanics: Rigorous Results*, W.A. Benjamin, Amsterdam.
- Schoenberg, F. P. (2004), Testing separability in spatial-temporal marked point processes, *Biometrics* **60**(2), 471–481.
- Schoenberg, F. P., Chang, C. H., Keeley, J. E., Pompa, J., Woods, J. and Xu, H. (2007), A critical assessment of the burning index in Los Angeles County, California, *International Journal of Wildland Fire* **16**(4), 473–483.
- Schoenberg, F. P., Pompa, J. and Chang, C. H. (2009), A note on non-parametric and semi-parametric modeling of wildfire hazard in Los Angeles County, California, *Environmental and Ecological Statistics* **16**, 251–269.
- Serra, L., Juan, P., Varga, D., Mateu, J. and Saez, M. (2013), Spatial pattern modelling of wildfires in Catalonia, Spain 2004–2008, *Environmental Modelling and Software* **40**, 235–244.
- Serra, L., Saez, M., Juan, P., Varga, D. and Mateu, J. (2014a), A spatio-temporal Poisson hurdle point process to model forest fires, *Stochastic Environmental Research and Risk Assessment* **28**(7), 1671–1684.
- Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P. and Diaz-Avalos, C. (2014b), Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of Catalonia, 1994–2008, *Environmental and Ecological Statistics* **21**(3), 531–563.
- Serra, L., Saez, M., Varga, D., Tobias, A. and Mateu, J. (2012), Spatio-temporal modelling of wildfires in Catalonia, Spain, 1994–2008, through log Gaussian Cox processes, *WIT Transactions on Ecology and the Environment* **158**, 34–49.

- Siino, M., Adelfio, G. and Mateu, J. (2018a), Joint second-order parameter estimation for spatio-temporal log-Gaussian Cox processes, *Stochastic Environmental Research and Risk Assessment* **32**, 3525–3539.
- Siino, M., Adelfio, G., Mateu, J., Chiodi, M. and DAlessandro, A. (2017), Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity, *Stochastic Environmental Research and Risk Assessment* **31**(7), 1633–1648.
- Siino, M., DAlessandro, A., Adelfio, G., Scudero, S. and Chiodi, M. (2018b), Multi-scale processes to describe the eastern sicily seismic sequences, *Annals of Geophysics* **61**(2), 1–17.
- Simpson, D., Rue, H., Riebler, A., Martins, T. G. and Sørbye, S. H. (2017), Penalising model component complexity: A principled, practical approach to constructing priors, *Statistical Science* **32**(1), 1–28.
- Sørbye, S. H., Illian, J., Simpson, D. P., Burslem, D. and Rue, H. (2019), Careful prior specification avoids incautious inference for log-Gaussian Cox point processes, *Journal of the Royal Statistical Society, Applied Statistics, Series C* **68**(3), 543–564.
- Stoyan, D. (1979), Interrupted point processes, *Biometrical Journal* **21**, 607–610.
- Stoyan, D. (1992), Statistical estimation of model parameters of planar Neyman-Scott cluster processes, *Metrika* **39**(1), 67–74.
- Strauss, D. J. (1975), A model for clustering, *Biometrika* **62**(2), 467–475.
- Tamayo-Uria, I., Mateu, J. and Diggle, P. J. (2014), Modelling of the spatiotemporal distribution of rat sightings in an urban environment, *Spatial Statistics* **9**, 192–206.
- Taylor, B. M., Davies, T. M., Rowlingson, B. S. and Diggle, P. J. (2015), Bayesian inference and data augmentation schemes for spatial, spatiotemporal and multivariate log-Gaussian Cox processes in R, *Journal of Statistical Software* **63**, 1–48.
- Taylor, B. M. and Diggle, P. J. (2014), INLA or MCMC? a tutorial and comparative evaluation for spatial prediction in log-Gaussian Cox processes, *Journal of Statistical Computation and Simulation* **84**, 2266–2284.
- Taylor, S. W., Woolford, D. G., Dean, C. B. and Martell, D. L. (2013), Wildfire prediction to inform fire management: Statistical science challenges, *Statistical Science* **28**(4), 586–615.
- Thomas, M. (1949), A generalization of Poissons binomial limit for use in ecology, *Biometrika* **36**, 18–25.

- Tierney, L. and Kadane, J. B. (1986), Accurate approximations for posterior moments and marginal densities, *Journal of the American Statistical Association* **81**(393), 82–86.
- Tonini, M., Pereira, M. G., Parente, J. and Orozco, C. V. (2017), Evolution of forest fires in Portugal: from spatio-temporal point events to smoothed density maps, *Natural Hazards* **85**, 1189–1510.
- Trilles, S., Juan, P., Diaz, L., Arago, P. and Huerta, J. (2013), Integration of environmental models in spatial data infrastructures: a use case in wildfire risk prediction, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **6**(1), 128–138.
- Turner, R. (2009), Point patterns of forest fire locations, *Environmental and Ecological Statistics* **16**, 197–223.
- Uppala, M. and Handcock, M. S. (2014), Modeling wildfire ignition origins in southern California using linear network point processes, *Annals of Applied Statistics* **14**(1), 339–356.
- van Lieshout, M. (2000), *Markov Point Processes and Their Applications*, Imperial College Press, London.
- van Lieshout, M. (2019), *Theory of Spatial Statistics*, Chapman and Hall/CRC, New York.
- Varin, C., Reid, N. and Firth, D. (2011), An overview of composite likelihood methods, *Statistica Sinica* **21**, 5–42.
- Vihrs, N., Møller, J. and Gelfand, A. E. (2021), Approximate Bayesian inference for a spatial point process model exhibiting regularity and random aggregation, *Scandinavian Journal of Statistics: Theory and Applications* . doi: 10.1111/sjos.12509.
- Wang, X., Zheng, G., Yun, Z. and Moskal, L. M. (2020), Characterizing tree spatial distribution patterns using discrete Aerial Lidar Data, *Remote Sensing* **12**(4), 712.
- Wang, Y. and Anderson, K. R. (2010), An evaluation of spatial and temporal patterns of lightning- and human-caused forest fires in Alberta, Canada, 1980–2007, *International Journal of Wildland Fire* **19**, 1059–1072.
- Wiegand, T., Gunatillekema, N. and Okudam, T. (2007), Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering, *Ecology* **88**, 3088–3102.
- Wiegand, T., Huth, A. and Martinez, I. (2009), Recruitment in tropical tree species: Revealing complex spatial pattern, *The American Naturalist* **174**, E106–E140.

- Wilkins, C. W. (1977), A Stochastic Analysis of the Effect of Fire on Remote Vegetation, PhD thesis, University of Adelaide, USA.
- Woo, H., Chung, W., Graham, J. M. and Lee, B. (2017), Forest fire risk assessment using point process modelling of fire occurrence and Monte Carlo fire simulation, *International Journal Wildland Fire* **26**, 789–805.
- Xi, D. D. Z., Taylor, S. W., Woolford, D. G. and Dean, C. B. (2019), Statistical models of key components of wildfire risk, *Annual Review of Statistics and Its Application* **6**, 197–222.
- Xu, H. and Schoenberg, F. P. (2011), Point process modeling of wildfire hazard in Los Angeles County, California, *Annals of Applied Statistics* **5**, 684–704.
- Yang, J., He, H. S. and Shifley, S. R. (2008), Spatial control of occurrence and spread of wildfires in the Missouri Ozark Highlands, *Ecological Applications* **18**(5), 1212–1225.
- Yang, J., He, H. S., Shifley, S. R. and Gustafson, E. J. (2007), Spatial patterns of modern period human-caused fire occurrence in the Missouri Ozark Highlands, *Forest Science* **53**(1), 1–15.
- Yang, J., Weisberg, P. J., Dilts, T. E., Loudermilk, E. L., Scheller, R. M., Stanton, A. and Skinner, C. (2015), Predicting wildfire occurrence distribution with spatial point process models and its uncertainty assessment: a case study in the Lake Tahoe Basin, USA, *International Journal of Wildland Fire* **24**, 380–390.
- Yau, C. Y. and Loh, J. M. (2012), A generalization of the Neyman-Scott process, *Statistica Sinica* **22**, 1717–1736.
- Yin, L., Shen, Z., Yang, M. and Piao, S. (2019), Wildfire detection probability of MODIS fire products under the constraint of environmental factors: a study based on confirmed ground wildfire records, *Remote Sensing* **11**, 3031.
- Yin, Z. Y., Estberg, J., Hallisley, E. J. and Cayan, D. R. (2007), Spatial patterns of lightning at different spatial scales in the western United States during August of 1990 - a case study using the geographic information systems technology, *Journal of Environmental Informatics* **9**(1), 4–17.
- Zhang, T. and Zhuang, Q. (2014), On the local odds ratio between points and marks in marked point processes, *Spatial Statistics* **9**, 20–37.
- Zhang, Z., Zhang, H., Li, D., Xu, J. and Zhou, D. (2013), Spatial distribution pattern of human-caused fires in Hulunbeir grassland, *Acta Ecologica Sinica* **33**(7), 2023–2031.

Zhuang, J. (2015), Weighted likelihood estimators for point processes, *Spatial Statistics* **14**, 166–178.

Appendices

Appendix A

Review article published in *Annales of ISUP*

On spatial and spatio-temporal multi-structure point process models

Morteza Raeisi^{*,†}, Florent Bonneu^{†,‡} and Edith Gabriel[‡]

LMA EA2151, Avignon University, F-84000 Avignon, France[†]
and INRAE, BioSP, F-84914, Avignon, France[‡]

Abstract: Spatial and spatio-temporal single-structure point process models are widely used in epidemiology, biology, ecology, seismology... However, most natural phenomena present multiple interaction structure or exhibit dependence at multiple scales in space and/or time, leading to define new spatial and spatio-temporal multi-structure point process models. In this paper, we investigate and review such multi-structure point process models mainly based on Gibbs and Cox processes.

1. Introduction

Fundamental concepts of the theory of point processes emerged from life tables, renewal theory and counting problems [28]. The modern theory has mainly been developed between 1940's and 1970's (see e.g. the monographs by Palm [69], Feller [36], Bartlett [12], Matérn [59] and Cox [23, 24]) and is linked to nonlinear techniques in stochastic process theory [13, 14]. From 1980's spatial and spatio-temporal point processes have then become a subject on their own right. Today, they cover a plethora of applications in ecology, forestry, astronomy, epidemiology, seismology, fishery...

Spatial (and spatio-temporal) point process data are a collection of points for which locations (and times) of occurrence have been observed in a specified spatial region (and temporal period). Usually, the terms *points* and *events* are respectively used for arbitrary locations and for observations. The main goals in the analysis of point patterns concern the specification of intensity variations (first-order moment), interaction between events (second-order moment) and model identification for the underlying process. Processes are often classified into three classes of interaction structure [30]:

- *randomness*: In the absence of any interaction between events, a point pattern is said Completely Spatially (or Spatio-Temporally) Random in the sense that the probability that an event occur at any point is equally likely to occur anywhere within a bounded region and that its location (and time) is independent of each any other event. This property provides the standard baseline against

*Corresponding author

AMS 2000 subject classifications: Primary 60G55, 62M30; secondary 62H11

Keywords and phrases: Spatio-temporal point process models, Cox process, Gibbs process, Multi-scale process, Multi-structure process

which point patterns are often compared. The simplest and most fundamental point process for modelling a complete random distribution of points is the Poisson point process [53, 54]. It is used as null hypothesis for statistical test of interaction [31, 50].

- *clustering or aggregation*: In a clustered distribution, events tend to be closer than would be expected under complete randomness. Clustered patterns are mainly modelled by Cox processes [25], in particular log-Gaussian Cox processes [16, 17, 34, 60], Poisson Cluster processes [18, 38, 65] and Shot-Noise Cox processes [15, 63, 64].
- *inhibition or regularity*: In a regular distribution, events are more evenly spaced than would be expected under complete randomness. This structure can be modelled by Strauss processes [27, 82], Matérn hard core processes [37, 59] or determinantal point processes [55, 58].

Gibbs processes [29, 74, 77] offer a large class of models which allow any of the above interaction structure.

These single-structure point process models are too simplistic to describe phenomena with interactions at different spatial or spatio-temporal scales. That is for instance the case of seismic data as the different sources of earthquakes (faults, active tectonic plate and volcanoes) produce events with different displacements [78] and can be seen as the superposition of background earthquakes (which are distributed over a large spatio-temporal scale with low density) and clustered earthquakes (which are distributed over a small spatio-temporal scale with high density) [71]. Such multi-structure phenomena motivate statisticians to construct new spatial point process models, e.g. in ecology [57, 73, 87], in epidemiology [47] and in seismology [78, 79], mainly based on Gibbs processes, but not only [56]. There are very few spatio-temporal models: [40] and [76] modeled the multi-scale spatio-temporal structure of forest fires occurrences by log-Gaussian Cox processes (LGCP) and multi-scale Geyer saturation process respectively, [48] developed a multi-scale area-interaction model for varicella cases and [52] modelled the locations of muskoxen herds by LGCP with a constructed covariate measuring local interactions.

In the spatial point processes literature, three general approaches are considered for constructing multi-structure point process models: hybridization [10], thinning and superposition [19]. Hybridization consists in combining two or more point process models [9]. Spatial hybrids of Gibbs models are defined in [10] and hybrids of area-interaction potentials in [73]. Extension of the hybridization approach to the spatio-temporal framework has recently been considered in [48, 76]. Thinning consists in deleting points of a point process according to some probabilistic rule which is either independent or dependent of thinning other points [19]. This operation allows to get point processes with inhibition at small scales and attraction at large scales [6, 56]. Superposition of several processes is the union of the points of each process. It can be useful to model multi-scale clustered processes [87].

In this paper, we give a thorough overview of available methods and models

for spatial and spatio-temporal multi-structure point process data. In Section 2, we review the required preliminaries which include definitions and properties of point processes and single-structure models. In Section 3, we investigate the spatial and spatio-temporal multi-structure point process models based on Gibbs and Cox processes and other methods for introducing new multi-structure models. Finally, Section 4 provides concluding remarks and discusses directions for future research.

2. Inhomogeneity and structures in point patterns

2.1. Definitions

We consider a finite spatial or spatio-temporal point process X observed in \mathcal{W} , where \mathcal{W} denotes either a spatial region $W \subset \mathbb{R}^d$ or a spatio-temporal region $W \times T \subset \mathbb{R}^d \times \mathbb{R}$. We denote \mathbf{x} a realization of the point process, i.e. a collection of events $\{x_i\}_{i=1,\dots,n}$ (or $\{(x_i, t_i)\}_{i=1,\dots,n}$) $\subset \mathcal{W}$. Let ξ be any point in \mathcal{W} . We refer to [19, 28] (resp. [33, 35, 43]) for more formal definitions of spatial (resp. spatio-temporal) point processes. Without loss of generality, we set $d = 2$ throughout this paper. The main characteristics driving the spatial (resp. spatio-temporal) distribution of points are the *intensity function*, which governs the univariate distribution of the points of X , and the *pair correlation function*, which governs the bivariate distribution of the points of X , i.e. the interaction between events. In the following we remind some definitions and properties when X is a spatial or a spatio-temporal point process.

Campbell's theorem [19] relates the expectation of a function, h assumed to be non-negative and measurable, summed over a point process X to an integral involving the mean measure of the point process :

$$\mathbb{E} \left[\sum_{\xi_1, \dots, \xi_k \in X}^{\neq} h(\xi_1, \dots, \xi_k) \right] = \int \dots \int h(\xi_1, \dots, \xi_k) \lambda^{(k)}(\xi_1, \dots, \xi_k) \prod_{i=1}^k d\xi_i,$$

where $\xi_i \in \mathcal{W}$ and $\lambda^{(k)}$, $k \geq 1$, are the product densities. For a simple point process, i.e. $\xi_i \neq \xi_j$ for $i \neq j$, if they exist, the product densities are related to the counting measure N in infinitesimal spatial or spatio-temporal regions $d\xi_1, \dots, d\xi_k \subset \mathcal{W}$, around ξ_1, \dots, ξ_k , with volumes $|d\xi_1|, \dots, |d\xi_k|$: $\mathbb{P} [N(d\xi_1) = 1, \dots, N(d\xi_k) = 1] = \lambda^{(k)}(\xi_1, \dots, \xi_k) \prod_{i=1}^k d\xi_i$. Thus, the intensity function is related to the expected number of points in infinitesimal regions

$$\lambda(\xi) = \lambda^{(1)}(\xi) = \lim_{|d\xi| \rightarrow 0} \frac{\mathbb{E}[N(d\xi)]}{|d\xi|}$$

and the pair correlation function is defined by

$$(2.1) \quad g(\xi_i, \xi_j) = \frac{\lambda^{(2)}(\xi_i, \xi_j)}{\lambda(\xi_i)\lambda(\xi_j)}.$$

A point process is *homogeneous* when its intensity is constant, $\lambda(\xi) = \lambda, \forall \xi$, *inhomogeneous* otherwise. In practice, the inhomogeneity is often driven by environmental covariates and we account for them by using parametric models for the intensity function [9]. Under the assumption of *stationarity*, the properties of the point process are invariant under translation and the process is homogeneous. The *second-order stationarity* states that the second-order intensity only depends on the difference between points $\lambda^{(2)}(\xi_i, \xi_j) = \lambda^{(2)}(\xi_i - \xi_j)$. Because in practice most of processes are inhomogeneous, [8, 39] weakened it and defined the *second-order intensity-reweighted stationary* assumption for which the pair correlation function (2.1) is well-defined and a function of $\xi_i - \xi_j$. [85] provides general concepts of factorial moment properties. The previous definition of inhomogeneous processes is not unique, [45] defined inhomogeneous model classes (including the class of reweighted second-order stationary processes) into the common general framework of hidden second-order stationary processes. The pair correlation function describes the structure of dependence/interaction between points : $g(\xi_i, \xi_j) = 1, > 1$ and < 1 indicates that the pattern is, respectively, completely random, clustered and regular.

Assume that the distribution of the point process is defined by a *probability density* $f(\mathbf{x})$ with respect to the distribution of a unit rate Poisson process. The probability density can be used to study point processes. It can be viewed as the probability of getting the point pattern \mathbf{x} , divided by the same probability under Complete Randomness [9]. The mathematical form of the probability density determines the structure of the point process, see [21, 22] about formulation of the density of point processes. A closely related concept is the Papangelou conditional intensity function [70], which has been extended to the spatio-temporal framework by [27]. It is defined by

$$(2.2) \quad \lambda(\xi|\mathbf{x}) = \frac{f(\mathbf{x} \cup \xi)}{f(\mathbf{x})},$$

for $\xi \notin \mathbf{x}$ provided $f(\mathbf{x}) \neq 0$.

2.2. Classical point process models

We refer to [9, 19, 31, 50, 64] and [27, 33, 35, 37, 43] for a presentation of most of spatial and spatio-temporal point process models. Hereafter we only focus on the ones mentioned/used in Section 3 to construct multi-structure point process models, namely the Poisson, Cox and Gibbs processes.

Poisson point processes

The Poisson point process is the reference model for independence of the locations of events, i.e. for complete spatial (or spatio-temporal) randomness. It is also the simplest and most widely used inhomogeneous point process model. Poisson point processes with intensity function $\lambda(\xi)$ are defined by two postulates :

- The number of points in any region $B \subseteq \mathcal{W}$, $N(B)$, follows a Poisson distribution with parameter $\int_B \lambda(\xi) d\xi$,
- For all $B \subseteq \mathcal{W}$, given $N(B) = n$, the n events in B form an independent random sample from the distribution on B with probability density function $\lambda(\xi) / \int_B \lambda(\xi) d\xi$.

The probability density of a Poisson point process with respect to the unit rate Poisson process is

$$f(\mathbf{x}) = \exp\left(|\mathcal{W}| - \int_{\mathcal{W}} \lambda(\xi) d\xi\right) \prod_{\xi \in \mathbf{x}} \lambda(\xi).$$

Then, from Equation (2.2), the Papangelou conditional intensity is $\lambda(\xi|\mathbf{x}) = \lambda(\xi)$ and $\lambda^{(2)}(\xi_i, \xi_j) = \lambda(\xi_i)\lambda(\xi_j)$, so that $g(\xi_i, \xi_j) = 1$.

Cox processes

Cox processes, so-called doubly stochastic point processes [23], are considered as a generalization of inhomogeneous Poisson processes where the intensity is a realization of a random field $\Lambda = \{\Lambda(\xi)\}_{\xi \in \mathcal{W}}$. These models are particularly useful as soon as spatial variation in events density reflects both the environment and dependence between events. Moreover, their first- and second-order moments being tractable, they are very attractive. We have

$$(2.3) \quad \lambda(\xi) = \mathbb{E}[\Lambda(\xi)] \quad \text{and} \quad g(\xi_i, \xi_j) = \frac{\mathbb{E}[\Lambda(\xi_i)\Lambda(\xi_j)]}{\lambda(\xi_i)\lambda(\xi_j)} = 1 + \frac{\text{cov}(\Lambda(\xi_i), \Lambda(\xi_j))}{\lambda(\xi_i)\lambda(\xi_j)}.$$

The probability density $f(\mathbf{x}) = \mathbb{E}[\exp(|\mathcal{W}| - \int_{\mathcal{W}} \Lambda(\xi) d\xi) \prod_{\xi \in \mathbf{x}} \Lambda(\xi)]$ is intractable for these processes. Consequently, the Papangelou conditional intensity is not known. The second-order intensity function $\lambda^{(2)}(\xi_i, \xi_j) = \mathbb{E}[\Lambda(\xi_i)\Lambda(\xi_j)]$ is only tractable for two special cases of Cox processes, that we present below, the *Shot Noise Cox process* and the *log-Gaussian Cox process*.

Shot noise Cox processes [61] (SNCP) are a wide class of Cox processes associated to

$$\Lambda(\xi) = \sum_{(c, \gamma) \in \Phi} \gamma k(c, \xi),$$

where Φ is a Poisson point process on $\mathcal{W} \times [0, \infty)$ with intensity measure ζ and $k(c, \cdot)$ is a density function on \mathcal{W} , $\forall c \in \mathcal{W}$. The intensity and pair correlation function are

$$\lambda(\xi) = \int \gamma k(c, \xi) d\zeta(c, \gamma) \quad \text{and} \quad g(\xi_i, \xi_j) = 1 + \frac{\int \gamma^2 k(c, \xi_i) k(c, \xi_j) d\zeta(c, \gamma)}{\lambda(\xi_i)\lambda(\xi_j)}.$$

SNCP include Poisson cluster processes, i.e. a Poisson process in which each point is replaced by a cluster of points, the original point is considered as the cluster center [26]. When the points in the cluster are independently and identically distributed

about the cluster centre, the process is referred to as a Neyman-Scott process [65]. Two mathematically tractable models of Neyman-Scott processes are the *Thomas process* [83], where k is a zero-mean normal density, and the *Matérn cluster process*, where k is a uniform density on a ball centered at the origin.

Log-Gaussian Cox processes (LGCP) have been introduced in [60], considering that the intensity is a log-Gaussian process : $\Lambda(\xi) = \exp(Y(\xi))$, where Y is a real-valued Gaussian random field, with mean function $\mu(\xi)$ and covariance function $C(\xi_i, \xi_j)$. In that case, from Equation (2.3) we have

$$\lambda(\xi) = \exp(\mu(\xi) + C(\xi, \xi)/2), \quad \forall \xi \in \mathcal{W} \quad \text{and} \quad g(\xi_i, \xi_j) = \exp(C(\xi_i, \xi_j)), \quad \forall \xi_i, \xi_j \in \mathcal{W}.$$

The expression of the pair correlation function shows that the interaction is controlled by the second-order moment of Y . If $C(\xi_i, \xi_j) \geq 0$, we get $g(\xi_i, \xi_j) > 1$ and clustering. As they are based on a latent random field describing the intensity, LGCPs have a hierarchical structure making them particularly flexible [50]. Note that the interaction is controlled through the second-order moment of the Gaussian random field, so that LGCPs do not describe the mechanistic process generating the points what is the case of most of Gibbs processes (see below) for which the dependence between points is controlled through local interaction between pairs of points.

Gibbs point processes

A finite Gibbs point process on \mathcal{W} admits a density

$$(2.4) \quad f(\mathbf{x}) = \exp(-\Psi(\mathbf{x}))$$

w.r.t. the Poisson process of unit intensity on \mathcal{W} . The potential function Ψ is often specified as the sum of pair potentials :

$$(2.5) \quad \Psi(\xi_1, \dots, \xi_n) = \alpha_0 + \sum_i \alpha_1(\xi_i) + \sum_{i < j} \alpha_2(\xi_i, \xi_j) + \dots + \alpha_n(\xi_1, \dots, \xi_n),$$

with α_0 a normalizing constant for the density and the pair potentials $\alpha_1, \alpha_2, \dots$ which determine the contribution to the potential from each δ -uple of points. Note that, if the α_δ , $\delta \geq 2$ are identically zero, the process is Poisson with intensity $\lambda(\xi) = \exp(-\alpha_1(\xi))$. Hence, α_1 can be viewed as controlling a spatial (or spatio-temporal) trend, while the α_δ , $\delta \geq 2$ control the interactions between events. The normalizing constant is generally intractable, so it is often impossible to compute the intensity and pair correlation function of Gibbs processes. However, the Papangelou conditional intensity can be computed [22].

When the interaction between points is restricted to pairs, i.e. for

$$f(\mathbf{x}) = \alpha \prod_i \beta(\xi_i) \prod_{i < j} \gamma(\xi_i, \xi_j),$$

with $\alpha > 0$, β an intensity function and γ a symmetric interaction function, the process is called *pairwise interaction process* [30, 84]. A well-known example of such processes is the *Strauss process* [82] for which

$$f(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\gamma^{s(\mathbf{x})},$$

where $\beta, \gamma > 0$, $n(\mathbf{x})$ is the number of points in \mathbf{x} and $s(\mathbf{x})$ the number of neighbour pairs of \mathbf{x} at distances less than a given distance R . When $\gamma = 0$, we get the *Hard Core process*. Note that in the Strauss process, γ should be smaller than 1 otherwise the density is no integrable. [41] modified the Strauss process and proposed the *Geyer saturation process* in which the overall contribution from each point is trimmed to never exceed a maximum value. We thus have

$$f(\mathbf{x}) = \alpha\beta^{n(\mathbf{x})}\prod_{\xi \in \mathbf{x}} \gamma^{\min(s, t(\xi, r, \mathbf{x}))},$$

where $\alpha, \beta, \gamma, r, s$ are parameters and $t(\xi, r, \mathbf{x})$ is the number of other events lying with a distance r of the point ξ .

3. Multi-structure point process models

Spatial and spatio-temporal single-structure point process models presented in the previous section are generally used when only one type of interaction governs the structure of the point pattern. When there are indications that the spatial or spatio-temporal structure combines several structures or varies with ranges of distances, we need to consider multi-structure point process models. We present in this section some of these models derived from the classes of Gibbs and Cox processes. By nature, few spatial point processes can exhibit directly several structures and/or scales of interaction and we recall some useful construction techniques to incorporate the multi-structure: hybridization, thinning, superposition or clustering.

3.1. Models based on Gibbs processes

Gibbs point processes are mainly used to model repulsion structure in point patterns, even if some examples exist for modelling low clustering [19]. Their definition through the potential function Ψ fit well in the statistical mechanics framework where the spatial modelling of particles needs often to consider their interaction. It is common in various domains (mechanics, biology. . .) to observe repulsion at short range and aggregation at medium-long range of entities, leading to define multi-structure point processes models.

For pairwise interaction processes, some parametric potential functions can be defined to take into account multiple scales of interaction, see e.g. [20, 44, 67, 72, 77]. We consider in the sequel the homogeneous case, i.e. when α_1 is constant and the pair potential function $\alpha_2(\xi_i, \xi_j) = \alpha_2(\|\xi_i - \xi_j\|)$ in (2.5).

The Lennard-Jones pair potential function, well-known in statistical mechanics, is given by

$$\alpha_2(r) = \epsilon_1 \left(\frac{\sigma}{r}\right)^{m_1} - \epsilon_2 \left(\frac{\sigma}{r}\right)^{m_2}, \quad \forall r > 0$$

where $m_1 > m_2$, $\epsilon_1, \sigma > 0$ and in the multi-structure case $\epsilon_2 > 0$. Another one is the step potential function given by

$$\alpha_2(r) = c_l \quad \text{if } R_{l-1} < r \leq R_l \quad \forall l = 1, \dots, m$$

where $R_0 = 0$, $R_m = \infty$, $c_1 = \infty$, $c_m = 0$ and $c_l \in \mathbb{R}$ for $l = 2, \dots, m-1$. The resulting model is an extension of the Strauss process to the multi-scale framework [72]. The square-well potential is obtained with $l = 2$. More recently, [42] introduced a pair potential function varying smoothly over distance with scale interactions defined through a differential system of equations. Other pair potential functions can be found in the literature for modeling multi-structure phenomena, e.g. in [19, 67].

Some of these pair potential functions define multi-scale generalizations of single scale Gibbs processes. Indeed, the step potential functions of homogeneous pairwise interaction processes in [30] and [72] represent multi-scale extensions of the Strauss process where the density is given by

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{l=1}^m \gamma_l^{s_l(\mathbf{x})},$$

where $s_l(\mathbf{x}) = \sum_{i < j} \mathbb{1}(R_{l-1} < \|\xi_i - \xi_j\| \leq R_l)$.

In the same way, the multi-scale generalization of the area-interaction model has been introduced in [3-5] with a two-scale structure and in [73] for multi-scale marked area-interaction processes. Its density function in a homogeneous multi-scale case is given by

$$f(\mathbf{x}) = \alpha \beta^{n(\mathbf{x})} \prod_{l=1}^m \exp(-\kappa_l U(\mathbf{x}, r_l))$$

where $U(\mathbf{x}, r_l)$ is the d -dimensional volume of the set $\mathcal{W} \cap \bigcup_{\xi \in \mathbf{x}} b(\xi, r_l)$, with $b(\xi, r_l)$ the ball centered at ξ_i of radius $r_l > 0$. The sign of κ_l defines the l th structure : inhibition if negative, clustering otherwise. [66] used area-interaction point processes for bivariate point patterns for modelling both attractive and inhibitive intra- and inter-specific interactions of two plant species.

[10] defined a new class of multi-scale Gibbs point processes named hybrid models and including the two previous generalization examples. This unified framework allows to define properly generalizations of single-scale Gibbs point processes by preserving Ruelle and local stability [84]. This 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, so that

$$f(\mathbf{x}) = cf_1(\mathbf{x})\dots f_m(\mathbf{x})$$

where c is a normalization constant and f_l is a Gibbs density function for $l = 1, \dots, m$. The choice of the normalization constant allows to well define a probability density in the case where the product $f_1\dots f_m$ is integrable. The integrability condition is of course essential and induced by others conditions on the f_l (Ruelle stability, local stability or hereditary, see [10]) which play an important role in simulation algorithms and are established in general to demonstrate the model validity of the hybrid process.

[10] introduced the spatial multi-scale Geyer saturation point process that was applied in epidemiology by [47] and in seismology by [78] and [79]. [76] extend the definition and the estimation procedure in the general case of an inhomogeneous spatio-temporal multi-scale Geyer saturation process which density is given by

$$(3.1) \quad f(\mathbf{x}) = c \prod_{\xi \in \mathbf{x}} \lambda(\xi) \prod_{l=1}^m \gamma_l^{\min\{s_l, n(C_r^q(\xi); \mathbf{x})\}}$$

where $\lambda \geq 0$ is a measurable and bounded function, γ_l, r_l, q_l and $s_l > 0$ are the model parameters and $n(C_r^q(\xi); \mathbf{x}) = \sum_{\xi_i \in \mathbf{x} \setminus \xi} \mathbb{1}\{\|x_i - x\| \leq r_l, |t_i - t| \leq q_l\}$ is the number of other points in \mathbf{x} which are in a cylinder centred on $\xi \in \mathbf{x}$ with spatial and temporal radii r_l and q_l . For fixed $l \in \{1, \dots, m\}$, when $0 < \gamma_l < 1$ we would expect to see inhibition between events at spatio-temporal scales. On the other hand, when $\gamma_l > 1$ we expect clustering between events. We observe that Equation (3.1) reduces to an inhomogeneous Poisson process when $s_l = 0 \forall l \in \{1, \dots, m\}$. [75] used a multitype generalization of Gibbs point processes with point-to-point interactions at different spatial scales in order to model a complex rainforest data of 83 species.

The definition of hybrid Gibbs models does not impose to consider the same m Gibbs models which is emphasized in [9]. In this way, [11] applied a hybrid model with three model structures at different ranges of distance to the spatial pattern of halophytic species distribution in an arid coastal environment. They considered a hardcore process at very short distances, a Geyer process at short to medium distances and a Strauss process for the structure at large distances.

3.2. Models based on Cox processes

Cox processes are mainly defined from additive or log-linear random intensity functions. Their hierarchical structure allows to quantify the various sources of variation governing the spatial or spatio-temporal distribution of the pattern of interest. They are widely used for modelling environmental and ecological patterns.

Cluster Cox processes and superposition

Some Cox processes are obtained by clustering of *offspring* points around *parent* points and correspond to specific cases of cluster processes. This two-step construction allows to consider easily different structures for the patterns of parents and offspring.

[62] introduced the class of Generalized Shot Noise Cox processes (GSNCP), extending the definition of SNCP, and allowing relevant multi-structure point processes for modelling regularity and clustering in many applications. This class has two advantages. Firstly, the parent process is not restricted to be Poisson, as in Neyman-Scott processes, and can be a repulsive Gibbs point process in order to add inhibition between the clusters. Secondly, in each cluster, the intensity and the bandwidth of the dispersion kernel can be random. By consequence, a GSNCP is a Cox process driven by a random field of the form

$$\Lambda(\xi) = \sum_{(c,\gamma,h) \in \Phi} \gamma k_h(c, \xi),$$

where Φ is a point process on $\mathcal{W} \times [0, \infty) \times [0, \infty)$ and h is a bandwidth for the kernel density $k_h(c, \cdot)$. So, given Φ , a GSNCP is distributed as the superposition $\cup_l X_l$ of independent Poisson processes with intensity functions $\gamma_l k_{h_l}(c_l, \cdot)$ where $\{\gamma_l\}_l$, $\{h_l\}_l$ are random and $\Phi_{cent} = \{c_l\}_l$ is the parent process. In population dynamics, with G_0 a Poisson process for the initial population and G_{n+1} a GSNCP where the cluster centers are given by G_n , the superposition of GSNCPs G_0, G_1, \dots is a spatial Hawkes process [46]. The GSNCP class contains the special cluster Cox process defined in [88], where the parents process is a Strauss process. This model coupling inhibition at medium/long range and aggregation in cluster is applied to tree locations in a rain-forest, in order to consider the competition and reproduction mechanisms. [1] and [2] generalized the Neymann-Scott process by considering a log-Gaussian Cox process model for the parents, instead of a homogeneous Poisson process, leading to two scales of clustering, inter- and intra-clusters. This hierarchical model is applied to storm cell modelling in North Dakota.

Wiegand and co-authors' papers [86, 87] consider several construction of Cox processes incorporating clustering at multiple scales. The nested double-cluster process is an extension of the Thomas process in an multi-generation evolution of the population where the offspring become parents and generate offspring. They consider also the superposition of cluster processes, like the Thomas process.

Cox processes with constructed covariate

Another way to incorporate both small and large spatial scale structure in Cox processes is to define a constructed covariate measuring the local structure of a point pattern associated to an additional spatial effect at medium-long range. This methodology developed in [51] and applied to koala data is used again in [49, 52] for

other spatial ecological data. They consider a log-Gaussian Cox process in a Bayesian framework in order to apply the INLA approach for speeding up the estimation of parameters in comparison to MCMC approaches that are very time-consuming. [40] used also this approach in the context of wildfire modelling in Mediterranean France. In the case of a spatial LGCP model, the method consists in estimating the random field Λ on grid cells s_i as follow

$$\Lambda(s_i) = \exp \left(\beta_0 + f(z_c(s_i)) + \sum_{k=1}^p f_k(z_k(s_i)) + Y(s_i) \right)$$

where β_0 is the intercept, $f(z_c(\cdot))$ is a function of the constructed covariate z_c , $f_k, k = 1, \dots, p$ are functions of the observed covariates z_k and Y is a Gaussian random field taking into account the spatial autocorrelation not explained by the covariates. This intensity is estimated for each cell s_i of a grid partitioning the observation window.

In [51], the constructed covariate at each center point c of the grid cell s is the distance from c to the nearest point in the pattern outside the grid cell, i.e $z_c(s) = \min_{\xi \in \mathbf{x} \setminus s} (\|c - \xi\|)$. This constructed covariate describes small scale inter-individual behavior whereas the random field Y captures the spatial autocorrelation at a large spatial scale. The space-time and space-mark extensions of the constructed covariate definition are respectively introduced in [52] and [49]. In [40] the constructed covariate corresponds to a temporal intensity index given by the ratio between the number of wildfires observed spatially close to an other in a specified period and the total number of closed wildfires observed outside this given period. This covariate measures the temporal wildfire inhibition at close spatial distances induced by the local burn of vegetation after a wildfire occurrence. [80] fitted a LGCP to rainforest tree species by adding to the combination of covariates in the log-intensity a spatial random field and error field. The first random field captures the spatial autocorrelation in point counts among neighboring grid cells and the second one the clustering within grid cells, as a nugget effect in geostatistics. The intensity in $s \in \mathcal{W}$ is thus given by

$$\Lambda(s) = \exp \left(\beta_0 + \sum_{k=1}^p \beta_k z_k(s) + \frac{1}{\sqrt{\tau}} \left\{ \sqrt{\rho} \times Y(s) + \sqrt{1 - \rho} \times \epsilon(s) \right\} \right)$$

where β_k are linear effects of observed covariates z_k , Y is a spatial random field with autocorrelation between grid cells and ϵ the error field driving the aggregation structure within grid cells.

Thinned point processes

Thinning is a an operation allowing to delete points in a point process in order to obtain a new one with different characteristics. Each point of a point process has a

probability $1 - \pi$ of deletion, where the retention probability π can be constant or not, independent of the location point or depending on one to several points. For Cox processes, this technique is generally applied to create random local regularity. For example, [6] applied a Matérn hard core dependent thinning to a Shot Noise Cox process to obtain short range repulsion with medium range clustering. For a given point pattern and a specified distance h , Matérn hard core thinning acts by first attaching random positive marks (arrival times) to each point. Subsequently a point is removed if it has a neighbour within distance h and with a smaller mark (i.e. the neighbour arrived earlier). In that way, for a given location ξ , the retention probability $\pi(\xi)$ is the ratio between the intensities of the thinned process and the original process at ξ . [56] extended the definition of interrupted point processes in [81] and [19] and considered a spatial point process X obtained by an independent thinning driven by a random process Z on a regular point process Y . An example is given with Y a Matérn hard core process and Z the transformation by a characteristic function of a Boolean disc model [19].

4. Discussion and conclusion

This paper presents a review of methods for constructing multi-structure point processes for modelling aggregation and/or inhibition at different spatial or spatio-temporal scales. We focus our attention on the main two classes of point processes, namely the Gibbs and Cox processes. Some multi-structure techniques are specific to a family of point processes, as the hybridization approach for Gibbs processes or the double-cluster process for Cox processes; others are more global, as the superposition or the thinning method, even if they are respectively more adapted to Gibbs or Cox processes. We could also consider determinantal point processes to model regularity as in [56] who considered it instead of the Matérn hard core process. Spatio-temporal point processes can also be defined by conditioning on the past, often used in epidemiology or seismology. For instance, the definition of the conditional intensity in [32] allows an aggregation of cases in the spatio-temporal spread of the foot and mouth disease and also a random occurrence of cases in the entire observation domain.

We selected the most relevant references for us in the state-of-the-art of these types of Gibbs and Cox models to describe these approaches for introducing regularity in cluster processes and aggregation in repulsive processes. Because these models are suitable in an environmental and ecological framework, due to the complexity of mechanisms governing attraction and repulsion of entities (particles, cells, plants...), we can expect a wide use of these models in many studies.

References

- [1] ALBERT-GREEN, A. (2016). *Joint Models for Spatial and Spatio-Temporal Point Processes*. PhD thesis, University of Western Ontario.
- [2] ALBERT-GREEN, A., BRAUN, W.J., DEAN, C. B. and MILLER, C. (2019). A hierarchical point process with application to storm cell modelling. *Canad. J. Stat.* **47(1)** 46–64.
- [3] AMBLER, G.K. (2002). *Dominated Coupling from the Past and Some Extensions of the Area-Interaction Process*. PhD thesis, University of Bristol.
- [4] AMBLER, G.K. and SILVERMAN, B. (2004). *Perfect simulation of spatial point processes using dominated coupling from the past with application to a multiscale area-interaction point process*. Preprint, University of Bristol, Department of Mathematics, Bristol. Available at <http://www.stats.ox.ac.uk/~silverma/pdf/amblersilverman1.pdf>
- [5] AMBLER, G.K. and SILVERMAN, B. (2010). *Perfect simulation using dominated coupling from the past with application to area-interaction point processes and wavelet thresholding*. In: Bingham, N.H., Goldie, C.M. (Eds.), *Probability and Mathematical Genetics*. Cambridge University Press, Cambridge.
- [6] ANDERSEN, I.T. and HAHN, U. (2016). Matern thinned Cox processes. *Spat. Stat.* **15** 1–21.
- [7] BADDELEY, A.J. and VAN LIESHOUT, M-C. (1995). Area-interaction point processes. *Ann. I. Stat. Math.* **47(4)** 601–619.
- [8] BADDELEY, A., MOLLER, J., and WAAGEPETERSEN, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat. Neer.* **54** 329–350.
- [9] BADDELEY, A., RUBAK, E. and TURNER, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*, Chapman and Hall/CRC Press, London.
- [10] BADDELEY, A., TURNER, R., MATEU, J. and BEVAN, A. (2013). Hybrids of gibbs point process models and their implementation. *J. Stat. Soft.* **55(11)** 1–43.
- [11] BADRELDIN, N., URÍA-DIEZ, J., MATEU, J., YOUSSEF, A., STAL, C., EL-BANA, M., MAGDY, A. and GOOSSENS, R. (2015). A spatial pattern analysis of the halophytic species distribution in an arid coastal environment. *Env. Mon. Ass.* **187** 1–15.
- [12] BARTLETT, M.S. (1954). Processus stochastiques ponctuels. *Ann. Inst. Henri Poincaré*, **14** 35–60.
- [13] BARTLETT, M.S. (1955). *An Introduction to Stochastic Processes*. Cambridge University Press.
- [14] BOSQ, D. (1998). *Nonparametric Statistics for Stochastic Processes*. Lecture Notes in Statistics. Springer.
- [15] BRIX, A. and CHADÉUF, J. (2000). Spatio-temporal modeling of weeds and shot-noise G Cox processes. *Biom. J.* **44** 83–99.
- [16] BRIX, A. and MØLLER, J. (2001). Space-time multitype log Gaussian Cox

- processes with a view to modelling weed data. *Scand. J. Stat.* **28** 471–488.
- [17] BRIX, A., DIGGLE, P.J. (2001). Spatiotemporal prediction for log-Gaussian Cox processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63**(4) 823–841.
- [18] BRIX, A. and KENDALL, W.S. (2002). Simulation of cluster point processes without edge effects. *Adv. Appl. Prob.* **34** 267–280.
- [19] CHIU, S.N., STOYAN, D., KENDALL, W.S. and MECKE, J. (2013). *Stochastic Geometry and its Applications*, 3rd ed. Wiley, Chichester.
- [20] CLYDE, M. and STRAUSS, D. (1991). Logistic regression for spatial pair-potential models. *Spat. Stat. Imaging* **20**, 14–30.
- [21] COEURJOLLY, J-F., MØLLER, J. and WAAGEPETERSEN, R.P. (2017). A tutorial on Palm distributions for spatial point processes. *Int. Stat. Review.* **85**(3), 404–420.
- [22] COEURJOLLY, J-F. and LAVANCIER, F. (2019). *Understanding Spatial Point Patterns Through Intensity and Conditional Intensities*. In: Coupier D. (eds) *Stochastic Geometry. Lecture Notes in Mathematics* **2237** Springer. pp 45–85.
- [23] COX, D.R. (1955). Some statistical methods connected with series of events (with discussion). *J. Roy. Statist. Soc. Ser. B*, **17** 129–164.
- [24] COX, D.R. (1962). *Renewal Theory*. Methuen, London.
- [25] COX, D. R. (1972). The statistical analysis of dependencies in point processes. In *Stochastic Point Processes* (ed Lewis, P. A. W.), Wiley, New York, 55–66.
- [26] COX, D.R. and ISHAM, V. (1980). *Point Processes*. Chapman & Hall, London.
- [27] CRONIE, O. and VAN LIESHOUT, M. (2015). A J-function for inhomogeneous spatio-temporal point processes. *Scand. J. Stat.* **42**(2) 562–579.
- [28] DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*. Vol. I. New York: Springer-Verlag.
- [29] DEREUDRE, D. (2019). *Introduction to the theory of Gibbs point processes.* In *Stochastic Geometry*. In: Coupier D. (eds) *Stochastic Geometry. Lecture Notes in Mathematics* **2237** Springer. pp. 181–229.
- [30] DIGGLE, P.J. (1983). *Statistical Analysis of Spatial Point Patterns*, Academic Press, London.
- [31] DIGGLE, P.J. (2003). *Statistical Analysis of Spatial Point Patterns*, 2nd ed. Hodder Arnold, London.
- [32] DIGGLE, P.J. (2006). Spatio-temporal point processes, partial likelihood, foot-and-mouth. *Stat. Meth. Med. Res.* **15** 325–336.
- [33] DIGGLE, P.J. and GABRIEL, E. (2010). *Spatio-temporal point processes*. Handbook of Spatial Statistics. Chapman and Hall/CRC Handbooks of Modern Statistical Methods, 449–461.
- [34] DIGGLE, P.J., MORAGA, P., ROWLINGSON, B. and TAYLOR, B.M. (2013). Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. *Statist. Sci.* **28**(4) 542–563.
- [35] DIGGLE, P.J. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, 3rd ed. Chapman and Hall/CRC, New York.
- [36] FELLER, W. (1950). *An Introduction to Probability Theory and Its Applications*.

- Wiley, New-York.
- [37] GABRIEL, E., ROWLINGSON, B., and DIGGLE, P.J. (2013). stpp: An R package for plotting, simulating and analyzing spatio-temporal point patterns. *J. Stat. Softw.* **53(2)** 1–29.
 - [38] GABRIEL, E. (2014). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes. *Meth. and Comp. App. Prob.* **16(2)** 411–431.
 - [39] GABRIEL, E. and DIGGLE, P.J. (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Stat. Neerl.* **18** 505–544.
 - [40] GABRIEL, E., OPITZ, T. and BONNEU, F. (2017). Detecting and modeling multi-scale space-time structures: the case of wildfire occurrences. *J. Fran. Stat. Soc.* **158(3)** 86–105.
 - [41] GEYER, C.J. (1999). *Likelihood Inference for Spatial Point Processes: Likelihood and Computation*. In W. Kendall, O. Barndroff-Nielsen, & M. N. van Lieshout (Eds.), *Stochastic Geometry: Likelihood and Computation* 141–172. London, UK: Chapman and Hall/CRC.
 - [42] GOLDSTEIN, J., HARAN, M., SIMEONOV, I., FRICKS, J. and CHIAROMONTE, F. (2015). An attraction-repulsion point process model for respiratory syncytial virus infections. *Biom.* **71(2)** 376–385.
 - [43] GONZALEZ, J.A., RODRIGUEZ-CORTES, F.J., CRONIE, O. and MATEU, J. (2016). Spatio-temporal point process statistics: A review. *Spat. Stat.* **18** 505–544.
 - [44] HABEL H., SARKKA A., RUDEMO M., BLOMQUIST C.H., OLSSON E. and NORDIN M. (2019). Colloidal particle aggregation in three dimensions. *J. Micro.* **275(3)** 149–158.
 - [45] HAHN, U. and VEDEL JENSEN, E. B. (2015). Hidden second-order stationary spatial point processes. *Scand. J. Stat.* **43(2)** 455–475.
 - [46] HAWKES, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90.
 - [47] IFTIMI, A., MONTES, P., MATEU, J. and AYYAD, C. (2017). Measuring spatial inhomogeneity at different spatial scales using Hybrids of Gibbs point process models. *Stoch. Env. Res. Ris. Ass.* **31(6)** 1455–1469.
 - [48] IFTIMI, A., VAN LIESHOUT, M.C. and MONTES, F. (2018). A multi-scale area-interaction model for spatio-temporal point patterns. *Spat. Stat.* **26** 38–55.
 - [49] ILLIAN, J., MARTINO, S., SORBYE, S., GALLEGU-FERNANDEZ, J.B., ZUNZUNEGUI, M., ESQUIVIAS, M.P. and TRAVIS, J.M. (2013). Fitting complex ecological point process models with integrated nested Laplace approximation. *Meth. Eco. Evo.* **4** 305–315.
 - [50] ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley & Sons, Chichester.
 - [51] ILLIAN, J., SORBYE, S. and RUE, H. (2012a). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA) *Ann. Appl. Stat.* **6(4)** 1499–1530.

- [52] ILLIAN, J., SORBYE, S., RUE, H. and HENDRICHSEN, D. (2012b). Using INLA to fit a complex point process model with temporally varying effects - a case study. *J. Env. Stat.* **3** 1–25.
- [53] KINGMAN, J. F. C. (1993). *Poisson Processes*. Oxford Studies in Probability **3**. Oxford University Press, Oxford.
- [54] KINGMAN, J. F. C. (2006). *Poisson processes revisited*. *Probab. Math. Statist.* **26** 77–95.
- [55] LAVANCIER, F., MØLLER, J. and RUBAK, E. (2015). Determinantal point process models and statistical inference. *J. Roy. Stat. Soc. B* **77** 853–877.
- [56] LAVANCIER, F. and MØLLER, J. (2016). Modelling aggregation on the large scale and regularity on the small scale in spatial point pattern datasets. *Scan. J. Stat.* **43** 587–609.
- [57] LEVIN, S.A. (1992). The problem of pattern and scale in ecology. *Ecology* **73** 1943–1967.
- [58] MACCHI, O. (1975). The coincidence approach to stochastic point processes. *Adv. Appl. Probab.* **7** 83–122.
- [59] MATÉRN, B. (1960). *Spatial Variation*. Lectures Notes in Statistics. Springer-Verlag.
- [60] MØLLER, J., SYVERSVEEN, A.R. and WAAGEPETERSEN, R.P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482.
- [61] MØLLER, J. (2003). Shot noise Cox processes. *Adv. Appl. Prob.* **35(3)** 614–640.
- [62] MØLLER, J. and TORRISI, G.L. (2005). Generalised shot noise Cox processes. *Adv. Appl. Prob.*, **37** 48–74.
- [63] MØLLER, J. and DIAZ-AVALOS, C. (2010). Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. *Scand. J. of Stat.* **37(1)** 2–25.
- [64] MØLLER, J. and WAAGEPETERSEN, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.
- [65] NEYMAN, J. and SCOTT, E.L. (1958). Statistical approach to problems of cosmology. *J. Roy. Statist. Soc. Ser. B* **20** 1–29.
- [66] NIGHTINGALE, G.F., ILLIAN J.B., KING, R. and NIGHTINGALE, P. (2019). Area interaction point processes for bivariate point patterns in a Bayesian context. *J. of Envir. Stat.*, **9(2)**
- [67] OGATA, Y. and TANEMURA, M. (1981). Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Ann. Inst. Stat. Math* **33(1)** 315–338.
- [68] OPITZ, T., BONNEU, F. and GABRIEL, E. (2020). Point-process based Bayesian modeling of space-time structures of forest fire occurrences in Mediterranean France. *Spat. Stat.*, 100429.
- [69] PALM, C. (1943). *Intensitätsschwankungen in Fernsprechverkehr*. Ericsson, Technics **44**.
- [70] PAPANGELOU, F. (1974). The conditional intensity of general point processes and an application to line processes. *Prob. Theo. Rel. Fiel.* **28(3)** 207–226.

- [71] PEI, T., GAO, J., MA, T. AND ZHOU, C. (2012). Multi-scale decomposition of point process data. *GeoInformatica*, **16**(4) 625–652.
- [72] PENTTINEN, A. (1984). *Modelling Interaction in Spatial Point Patterns: Parameter Estimation by the Maximum Likelihood Method*. Number 7 in Jyväskylä Studies in Computer Science, Economics, and Statistics, University of Jyväskylä.
- [73] PICARD, N., BAR-HEN, A., MORTIER, F. and 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.
- [74] PRESTON, C.J. (1976). *Random Fields*. Lecture Notes in Mathematics 534. Springer-Verlag, Berlin.
- [75] RAJALA, T., MURRELL, D.J. and OLHEDE, S.C. (2018) Detecting multivariate interactions in spatial point patterns with Gibbs models and variate selection. *Appl. Statist.* **67**(5), 1237–1273.
- [76] RAEISI, M., BONNEU, F. and GABRIEL, E. (2019). A spatio-temporal multi-scale model for Geyer saturation point process: application to forest fire occurrences. Unpublished results. <http://arxiv.org/abs/1911.06999>.
- [77] RUELLE, D. (1969). *Statistical Mechanics: Rigorous Results*, W.A. Benjamin, Reading, Massachusetts.
- [78] SIINO, M., ADELFO, G., MATEU, J. and D’ALESSANDRO, A. (2017). Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity. *Stoch. Env. Res. Ris. Ass.* **31**(7) 1633–1648.
- [79] SIINO, M., D’ALESSANDRO, A., ADELFO, G., SCUDERO, S. and CHIODI, M. (2018). Multiscale processes to describe the eastern sicily seismic sequences. *Ann. Geo.* **61**(2)
- [80] SORBYE, S.H., ILLIAN, J.B., SIMPSON, D.P., BURSLEM, D. and RUE, H. (2019). Careful prior specification avoids incautious inference for log-Gaussian Cox point processes. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **68**(3) 543–564.
- [81] STOYAN, D. (1979). Interrupted point processes. *Biometrical. J.* **21** 607–610.
- [82] STRAUSS, D.J. (1975). A model for clustering. *Biometrika* **62** 467–475.
- [83] THOMAS, M. (1949). A generalization of Poisson’s binomial limit for use in ecology. *Biometrika* **36** 18–25.
- [84] VAN LIESHOUT, M-C. (2000). *Markov Point Processes and Their Applications*. Imperial College Press, London.
- [85] VAN LIESHOUT, M-C. (2019). *Theory of Spatial Statistics*. New York: Chapman and Hall/CRC.
- [86] WIEGAND, T., HUTH, A. and MARTINEZ, I. (2009). Recruitment in tropical tree species: Revealing complex spatial patterns. *Ame. Nat.* **174** E106–E140.
- [87] WIEGAND, T., GUNATILLEKEM, N. and OKUDAM, T. (2007). Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. *Eco.* **88** 3088–3102.
- [88] YAU, C.Y. and LOH, J.M. (2012). A generalization of the Neyman-Scott process. *Stat. Sinica.* **22** 1717–1736.

*Laboratoire de Mathématiques d'Avignon
Avignon University
Campus Jean-Henri Fabre
301 Rue Baruch de Spinoza,
BP 21239
F-84916 AVIGNON Cedex 9*

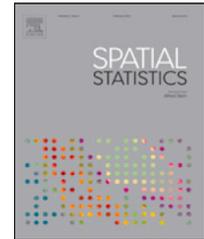
e-mail: morteza.raeisi@univ-avignon.fr
florent.bonneu@univ-avignon.fr
edith.gabriel@inrae.fr
url: math.univ-avignon.fr/annuaire

Appendix B

Original research published in *Spatial Statistics*

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

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.

© 2021 Elsevier B.V. All rights reserved.

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 Kendall, 2002](#); [Gabriel, 2014](#)) and

* Corresponding author.

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

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 Ræisi 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 $f = cf_1 \times \dots \times f_m$ 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 database¹, 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. $\mathbf{x} = \{(\xi_1, t_1), \dots, (\xi_n, t_n)\}$ where (ξ_i, t_i) describes the location and time, respectively, associated with the i th event.

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

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

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

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

with $a/0 := 0$ for $a \geq 0$ and $(u, v) \in W$ (Cronie and van Lieshout, 2015). Hence, we have $\lambda((u, v)|\mathbf{x}) = \frac{f(\mathbf{x} \cup (u, v))}{f(\mathbf{x})}$ if $(u, v) \notin \mathbf{x}$ and $\lambda((u, v)|\mathbf{x}) = \frac{f(\mathbf{x})}{f(\mathbf{x} \setminus (u, v))}$ 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 $\tilde{n}(C_r^q(u, v); \mathbf{x}) = \sum_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{\|u - \xi\| \leq r, |v - t| \leq q\}$ 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 $n(C_r^q(\xi, t); \mathbf{x}) = \sum_{(u, v) \in \mathbf{x} \setminus (\xi, t)} \mathbb{1}(\|u - \xi\| \leq r, |v - t| \leq q)$ 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 $\gamma_j > 0, j = 1, \dots, m$, are the interaction parameters, and $r_1 < \dots < r_m, q_1 < \dots < q_m$ are spatial and temporal interaction ranges.

For any $j \in \{1, \dots, m\}$, 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 $j \in \{1, \dots, m\}$, 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 $\gamma_j, j = 1, \dots, m$ 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 $\max_{1 \leq j \leq m} \{2 \max\{r_j, q_j\}\} = 2 \max\{r_m, q_m\}$ (Baddeley et al., 2013). \square

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

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

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 θ 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 $\lambda_\theta((u, v)|\mathbf{x})$ over W as follows

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

The pseudo score is defined by

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

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

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

where

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

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

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

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

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

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

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

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

For each parameter $\theta_i, i = 1, \dots, m$, 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

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

where (ξ_k, t_k) are points in $\{(\xi_1, t_1), \dots, (\xi_n, t_n), (\xi_{n+1}, t_{n+1}), \dots, (\xi_{n+p}, t_{n+p})\} \in W$ 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

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

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

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

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 $Y_k \sim \text{Poisson}(\lambda_\theta((\xi_k, t_k)|\mathbf{x}))$ 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 $k = 1, \dots, n + p$.
- Compute the quadrature weights w_k and the indicators y_k defined in (18),
- Compute the sufficient statistics $\mathbf{S}((\xi_k, t_k), \mathbf{x})$ at each quadrature point,
- Fit a log-linear Poisson regression with explanatory variables $\mathbf{S}((\xi_k, t_k), \mathbf{x})$, and offset $\log \lambda(\xi_k, t_k)$ on the responses y_k with weights w_k to obtain estimates $\hat{\theta}$ for the \mathbf{S} -vector and intercept $\hat{\theta}_0$,
- Return the maximum pseudo-likelihood-based parameter estimates $\hat{\gamma}_j = \exp(\hat{\theta}_j)$ for $j = 1, \dots, m$ 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 v 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 $w_k = v/n_k$ 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_θ with respect to the unit rate Poisson process and with conditional intensity function $\lambda_\theta(\cdot|\mathbf{x})$. We consider an independent Poisson process for dummy points, with intensity function ρ , and we denote by \mathbf{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 $(\xi, t) \in \mathbf{x} \cup \mathbf{d}$, we obtain independent Bernoulli variables taking one for data points and zero for dummy points. We have

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

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

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

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

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

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

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 from the following algorithm.

Algorithm 2

- Generate dummy points \mathbf{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 $\mathbf{S}((\xi_k, t_k), \mathbf{x} \setminus (\xi_k, t_k))$ at each quadrature point,
- Fit a logistic regression model with explanatory variables $\mathbf{S}((\xi_k, t_k), \mathbf{x} \setminus (\xi_k, t_k))$, and offset $\log(\mu(\xi_k, t_k)/\rho(\xi_k, t_k))$ on the responses y_k to obtain estimates $\hat{\theta}$ for the \mathbf{S} -vector and intercept $\hat{\theta}_0$,

- Return the parameter estimator $\hat{\gamma} = \exp(\hat{\theta})$ 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 $d(\mathbf{x}, \cdot)$ 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 $n = 0, 1, \dots$, 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 $r((\xi, t); \mathbf{x} \setminus (\xi, t)) = \frac{\ell(W)}{n(\mathbf{x})} \lambda((\xi, t)|\mathbf{x} \setminus (\xi, t))$, $(\xi, t) \in \mathbf{x}$.
If $y_2 < 1/r((\xi, t); \mathbf{x} \setminus (\xi, t))$ then $X_{n+1} = \mathbf{x} \setminus (\xi, t)$ 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

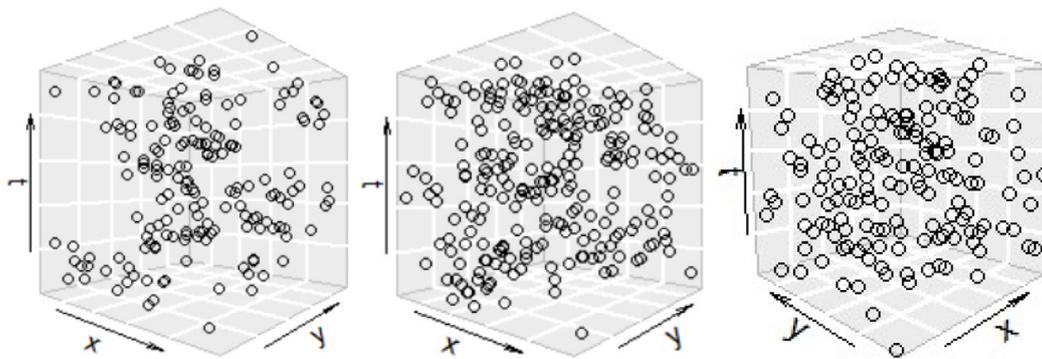


Fig. 1. Realizations of Model 1 (left); Model 2 (middle); Model 3 (right).

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

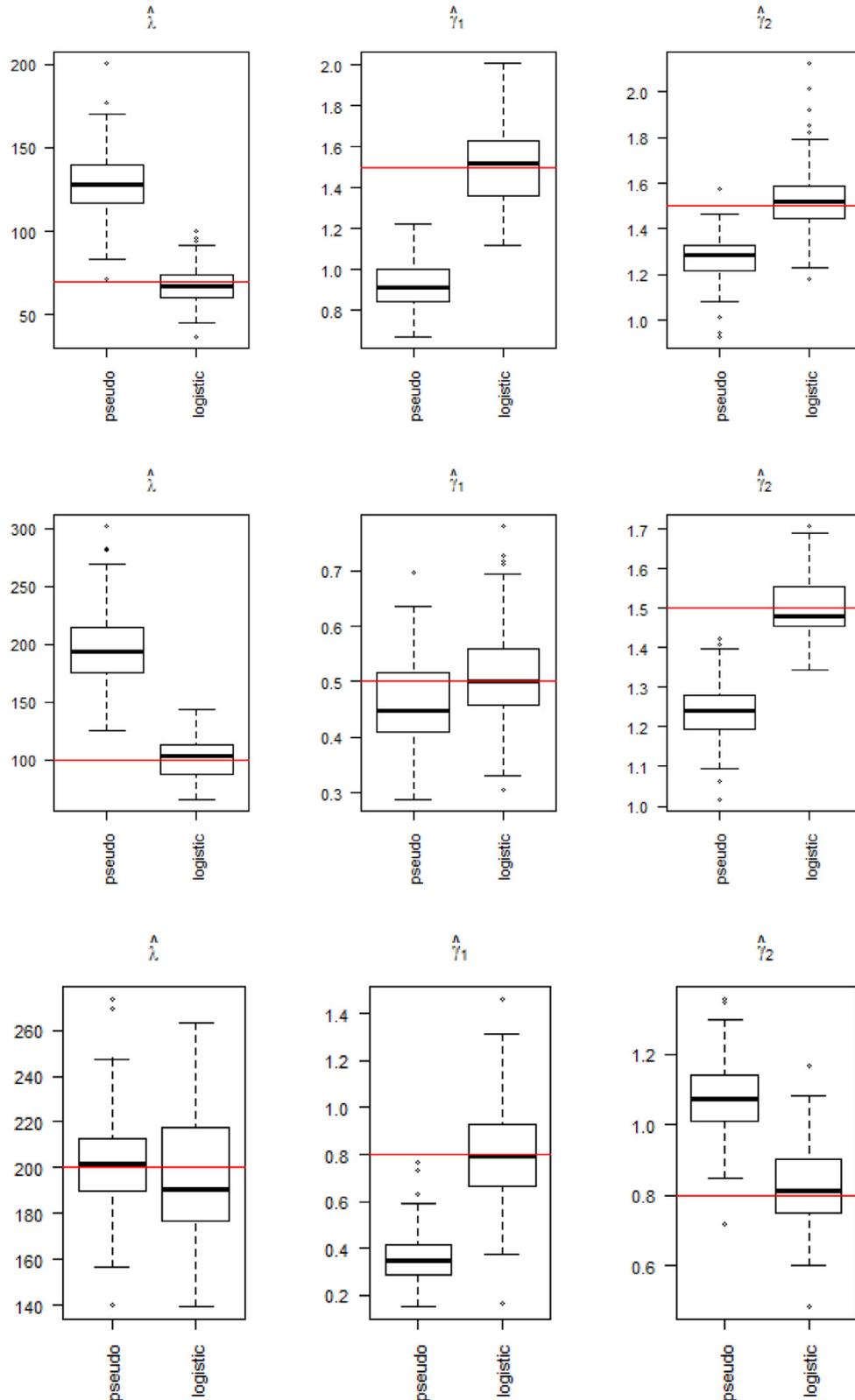


Fig. 2. Boxplots of regular parameters estimated from the pseudo-likelihood and logistic likelihood approaches for *Model 1* (first row), *Model 2* (second row) and *Model 3* (third row). True values are represented by horizontal red lines. (For interpretation of the references to color in this figure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

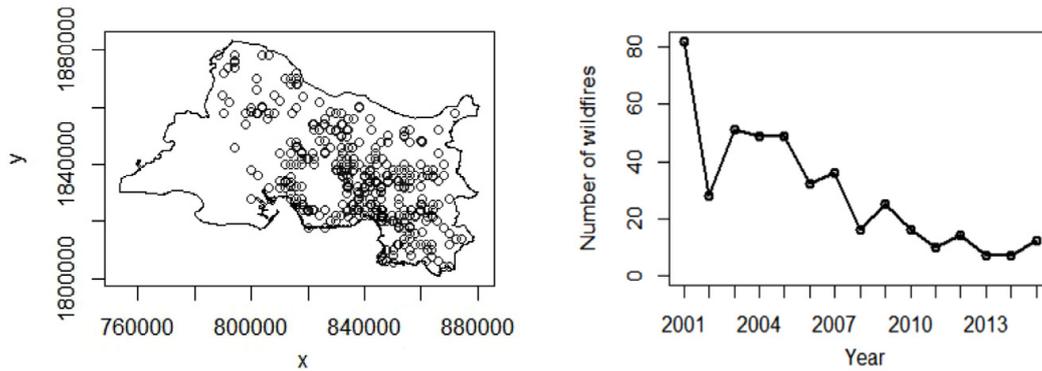


Fig. 3. (Left) Forest fire locations in UTM coordinate system (distance in meters), with more than 1 *hectare* of burnt area, recorded during the years 2001 to 2015 in the Bouches-du-Rhône county in France. (Right) Number of recorded forest fires per year.

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) , $i = 1, \dots, 434$, where (ξ_i, t_i) corresponds to a wildfire with more than 1 *hectare* of burnt surface spatially indexed by a DFCI² cell center ξ_i in the Lambert 93 projection system and year $t_i \in \{2001, \dots, 2015\}$. 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. \(2020\)](#).

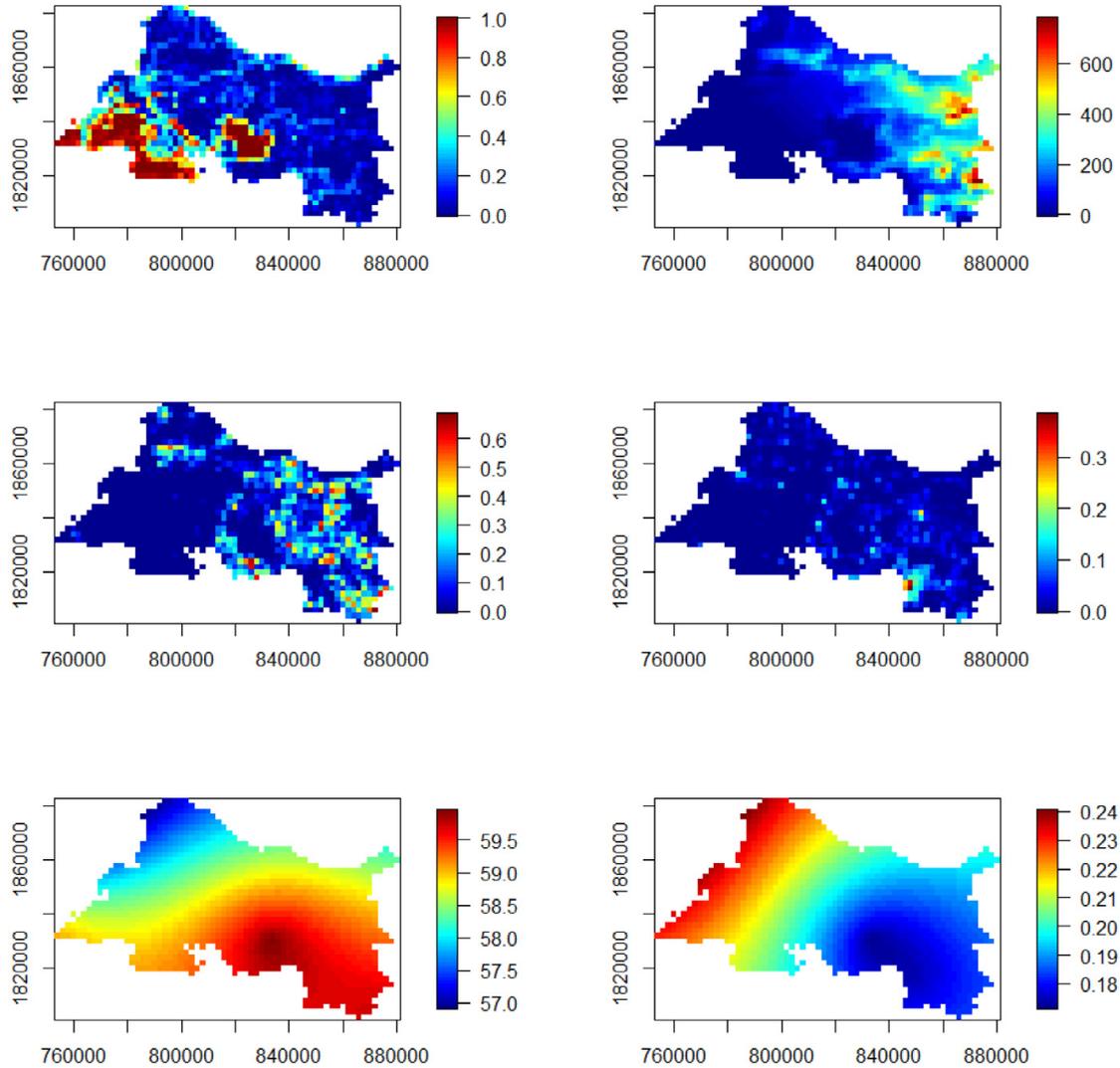


Fig. 4. Maps of covariates: water coverage (top left), elevation (top right), coniferous cover (middle left), building cover (middle right), temperature average (bottom left) and square root of precipitation (bottom right) in 2001.

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 $Z_k^S(\xi)$, $k = 1, \dots, 4$, 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 $\xi_i, i = 1, \dots, 1320$ and years $t_j = 2001, \dots, 2015$ for $j = 1, \dots, 15$, and we consider a Poisson response for our model $N_{ij} | \mu(\xi_i, t_j) \sim \text{Poisson}(\mu(\xi_i, t_j))$, where N_{ij} is the number of forest fires in i th 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 $\beta_0, \beta_k^S, \beta_l^{ST}$ 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 $s_j, j = 1, \dots, m$, 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				
$\hat{\beta} = 0.66$ [0.442, 0.968]	$\hat{\gamma}_1 = 2.73$ [1.818, 3.405]	$\hat{\gamma}_2 = 0.93$ [0.820, 0.994]	$\hat{\gamma}_3 = 1.07$ [1.020, 1.120]	$\hat{\gamma}_4 = 0.98$ [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 $\hat{K}_{obs}^{inh}(h_s, h_t)$ and $\hat{K}_i^{inh}(h_s, h_t), i \in \{1, \dots, n_{sim}\}$, 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

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

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 $\{\hat{K}_1^{inh}(h_s, h_t), \dots, \hat{K}_{n_{sim}}^{inh}(h_s, h_t), \hat{K}_{obs}^{inh}(h_s, h_t)\}$. 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 i th 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); $\hat{K}_{obs}^{inh}(h_s, h_t)$ 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

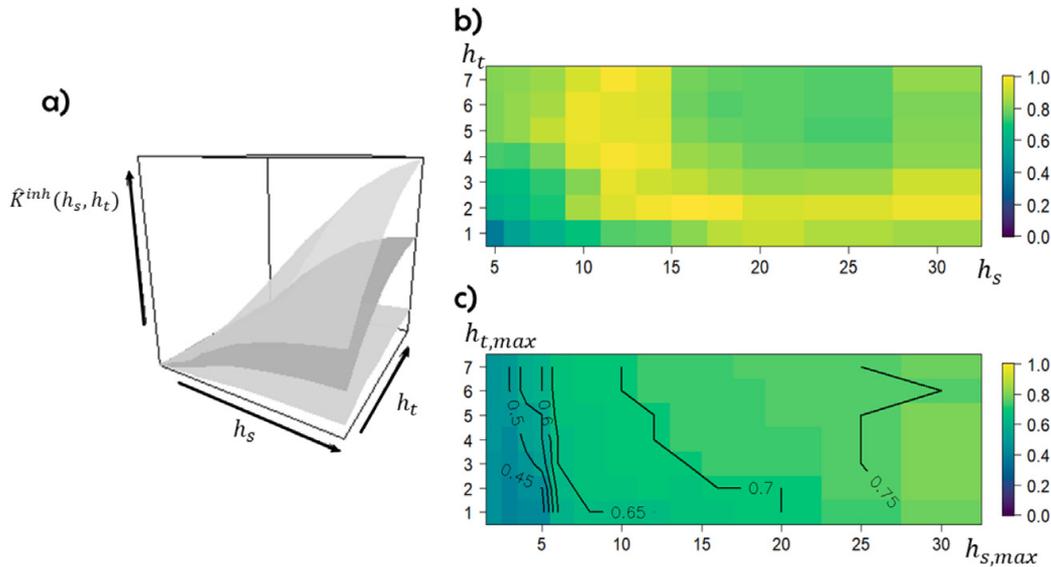


Fig. 5. Temporal separations h_t are in year and spatial distances h_s are in kilometer. (a) Envelopes of the spatio-temporal inhomogeneous K -function for the simulated spatio-temporal multi-scale Geyer point process according to the estimated parameters. (b) Image plot of the local p -value. (c) Image 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 $T_i, i = 1, \dots, n_{sim}$ (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

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

where R_i is the vector of pointwise ordered ranks and $<$ 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 code³ 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 , $j = 1, \dots, m$. 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., Stoyan, 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](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](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., Kendall, 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., Letu e, 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=scholar>.
- 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., Gunatillekum, 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>.

Appendix C

**Original research submitted in *Stochastic Environmental Research
and Risk Assessment***

A spatio-temporal hybrid Strauss hardcore point process for forest fire occurrences

Morteza Raeisi · Florent Bonneu ·
Edith Gabriel

Received: date / Accepted: date

Abstract We propose a new point process model that combines, in the spatio-temporal setting, both multi-scaling by hybridization and hardcore distances. Our so-called hybrid Strauss hardcore point process model allows different types of interaction, at different spatial and/or temporal scales, that might be of interest in environmental and biological applications. The inference and simulation of the model are implemented using the logistic likelihood approach and the birth-death Metropolis-Hastings algorithm. Our model is used to describe forest fire occurrences in Spain.

Keywords Spatio-temporal Gibbs point processes · Hybridization · Multi-scale point processes · Forest fires

1 Introduction

In point process modeling, most of existing models yield point patterns with mainly single-structure, but only a few with multi-structure. Interactions with single-structure are often classified into three classes: randomness, clustering and inhibition. Among inhibition processes is the hardcore process. It has some hardcore distance h in which distinct points are not allowed to come closer than a distance h apart. This type of interaction can be modelled by Gibbs point processes as the hardcore or Strauss hardcore point processes and also

M. Raeisi
LMA EA2151, Avignon University, F-84000 Avignon, France
Tel.: +334 90 84 35 41
E-mail: morteza.raeisi@univ-avignon.fr

F. Bonneu
LMA EA2151, Avignon University, F-84000 Avignon, France
INRAE, BioSP, F-84914 Avignon, France

E. Gabriel
INRAE, BioSP, F-84914 Avignon, France

by Cox point processes as Matérn's hardcore (Matérn, 1960; 1986) or Matérn thinned Cox point processes (Andersen and Hahn, 2016). Here, we focus on the former, i.e. Gibbs models implemented by a hardcore component as in the Strauss hardcore model. The form of Strauss hardcore density indicates that the hardcore parameter only rules at least distance between points, and has no effect on the interaction terms of the density (Dereudre and Lavancier, 2017, sect. 2.3).

In several domains, there exist point patterns with hardcore distances that have to be modelled. Spatial point patterns with hardcore property can be found in capillaries studies (Mattfeldt et al., 2006; 2007; 2009), in texture synthesis (Hurtut et al., 2009), in forest fires (Turner, 2009), in cellular networks (Taylor et al., 2012 and Ying et al., 2014), in landslides (Das and Stein, 2016), in modern and contemporary architecture and art (Stoyan, 2016) and in location clustering econometrics (Sweeney and Gomez-Antonio, 2016).

There also exist point patterns with either clustering and inhibition like hardcore interactions at different scales simultaneously (Badreldin et al., 2015; Andersen and Hahn, 2016 and Wang et al., 2020). Wang et al. (2020, sect. 2.4) investigated effect of the hardcore distance on spatial patterns of trees by comparing the pair correlation function curves for different values of hardcore distances in the fitted hybrid Geyer hardcore model. Raeisi et al. (2019) review spatial and spatio-temporal point processes that model both inhibition and clustering at different scales. Such multi-structure interactions can be modelled by the spatial hybrid Gibbs point process (Baddeley et al., 2013). In this paper, we aim to extend the spatial Strauss hardcore point process (Ripley, 1988) to the spatio-temporal framework and introduce a multi-scale version of it using hybridization approach. We use this model to describe one of the most complex phenomena from the spatio-temporal modeling point of view: forest fire occurrences.

The complexity of forest fire occurrences is due in particular to the existence of multi-scale structures and hardcore distances in space and time. For instance, spatio-temporal variations of fire occurrences depend on the spatial distribution of current land use and weather conditions. Changes in vegetation due to forest fires burnt areas further affect the probability of fire occurrences during the regeneration period leading to the existence of hardcore distances in space-time. The multi-scale structure of clustering and inhibition in the spatio-temporal pattern of the forest fire occurrences is discussed in Gabriel et al. (2017). Wildfires have mainly been modelled by Cox processes and inferred by Bayesian hierarchical approaches, as the integrated nested Laplace approximation (INLA) approach (Rue et al., 2009). See Møller and Diaz-Avalos (2010), Pereira et al. (2013), Serra et al. (2012, 2014a,b), Najafabadi et al. (2015), Juan (2020) and Pimont et al. (2021) for single-structure models and Gabriel et al. (2017), Opitz et al. (2020) for multi-structure models. Recently, Raeisi et al. (2021) modelled the multi-structure of forest fire occurrences by a spatio-temporal Gibbs process and use a composite likelihood approach for its inference.

This paper is organized as follows. In Section 2 we introduce in the spatio-temporal framework the notations and definitions of Gibbs point processes in order to introduce our multi-scale version of the Strauss hardcore model. Section 3 is devoted to the inference of our model. It describes techniques to determine the irregular parameters (hardcore and interaction distances) and the logistic-likelihood approach generalized to the spatio-temporal setting to estimate the regular parameters (strength of interactions). Section 4 illustrates the goodness-of-fit of the logistic likelihood approach on simulated patterns of our model obtained by an extended Metropolis-Hastings algorithm. Finally in Section 5, we apply our model to monthly records of forest fires in the center of Spain.

2 Towards multi-scale Strauss hardcore point processes

Gibbs models are flexible point processes that allow the specification of point interactions via a probability density defined with respect to the unit rate Poisson point process. These models allow to characterize a form of local or Markovian dependence amongst events. Gibbs point processes contain a large class of flexible and natural models that can be applied for:

- Postulating the interaction mechanisms between pairs of points,
- Taking into account clustering, randomness or inhibition structures,
- Combining several structures at different scales with the hybridization approach.

Let $\mathbf{x} = \{(\xi_1, t_1), \dots, (\xi_n, t_n)\}$ be a spatio-temporal point pattern where $(\xi_i, t_i) \in W = S \times T \subset \mathbb{R}^2 \times \mathbb{R}$. We consider $(W, d(\cdot, \cdot))$ where $d((u, v), (u', v')) := \max\{\|u - u'\|, |v - v'|\}$ for $(u, v), (u', v') \in W$ is a complete, separable metric space. The cylindrical neighbourhood $C_r^q(u, v)$ centred at $(u, v) \in W$ is defined by

$$C_r^q(u, v) = \{(a, b) \in W : \|u - a\| \leq r, |v - b| \leq q\}, \quad (1)$$

where $r, q > 0$ are spatial and temporal radius and $\|\cdot\|$ 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 centre (u, v) , radius r , and height $2q$.

A finite Gibbs point process is a finite simple point process defined with a density $f(\mathbf{x})$ that satisfies the hereditary condition, i.e. $f(\mathbf{x}) > 0 \Rightarrow f(\mathbf{y}) > 0$ for all $\mathbf{y} \subset \mathbf{x}$.

A closely related concept to density functions is Papangelou conditional intensity function (Papangelou, 1974) which is a tool for simulating Gibbs models and inferring its parameters. The Papangelou conditional intensity of a spatio-temporal point process on W with density f is defined, for $(u, v) \in W$, by

$$\lambda((u, v)|\mathbf{x}) = \frac{f(\mathbf{x} \cup (u, v))}{f(\mathbf{x} \setminus (u, v))}, \quad (2)$$

with $a/0 := 0$ for all $a \geq 0$ (Cronie and van Lieshout, 2015).

The Papangelou conditional intensity is also very useful to describe local interactions in a point pattern, and leads to the notion of a Markov point process which is the basis for the implementation of MCMC algorithms used for simulating of Gibbs models. We say that the point process has "interactions of range R at (ξ, t) " if points further than R away from (ξ, t) do not contribute to the conditional intensity at (ξ, t) . A spatio-temporal Gibbs point process X has a finite interaction range R if the Papangelou conditional intensity satisfies

$$\lambda((u, v)|\mathbf{x}) = \lambda((u, v)|\mathbf{x} \cap C_R^R(u, v)) \quad (3)$$

for all configurations \mathbf{x} of X and all $(u, v) \in W$, where $C_R^R(u, v)$ denotes the cylinder of radius $R > 0$ and height $2R > 0$ centred at (u, v) . Spatio-temporal Gibbs models usually have finite interaction range property (spatio-temporal Markov property) and are called in this case Markov point processes (van Lieshout 2000). The finite range property of a spatio-temporal Gibbs model implies that the probability to insert a point (u, v) into \mathbf{x} depends only on some cylindrical neighborhood of (u, v) .

Here, we first review spatio-temporal Gibbs models and then extend the spatial Strauss hardcore model to the spatio-temporal and multi-scale context. We further refer to Dereudre (2019) for more formal introduction of Gibbs point processes.

2.1 Single-scale Gibbs point process models

In the literature, several spatio-temporal Gibbs point process models have been proposed such as the hardcore (Cronie and van Lieshout, 2015), Strauss (Gonzalez et al., 2016), area-interaction (Iftimi et al., 2018), and Geyer (Raeisi et al., 2021) point processes.

A Gibbs point process model explicitly postulates that interactions traduce dependencies between the points of the pattern. The hardcore interaction is one of the simplest type of interaction, which forbids points being too close to each other. The homogeneous spatio-temporal hardcore point process is defined by the density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})}\mathbb{1}\{|\xi - \xi'| > h_s \text{ or } |t - t'| > h_t; \forall(\xi, t) \neq (\xi', t') \in \mathbf{x}\}, \quad (4)$$

with respect to a unit rate Poisson point process on W , where $c > 0$ is a normalizing constant, $\lambda > 0$ is an activity parameter, $h_s, h_t > 0$ are, respectively, the spatial and the temporal hardcore distances and $n(\mathbf{x})$ is the number of points in \mathbf{x} . The Papangelou conditional intensity of a homogeneous spatio-temporal hardcore point process for $(u, v) \notin \mathbf{x}$ is obtained

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda\mathbb{1}\{|\xi - u| > h_s \text{ or } |t - v| > h_t; \forall(\xi, t) \in \mathbf{x}\} \\ &= \lambda \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{|\xi - u| > h_s \text{ or } |t - v| > h_t\} \\ &= \lambda \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{(\xi, t) \notin C_{h_s}^{h_t}(u, v)\}. \end{aligned} \quad (5)$$

. The homogeneous spatio-temporal Strauss point process is defined by density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})}\gamma^{S_r^q(\mathbf{x})}, \quad (6)$$

with respect to a unit rate Poisson point process on W , where

$$S_r^q(\mathbf{x}) = \sum_{(\xi,t) \neq (\xi',t') \in \mathbf{x}} \mathbb{1}\{|\xi - \xi'| \leq r, |t - t'| \leq q\}$$

and the Papangelou conditional intensity of the model is

$$\text{for } (u, v) \notin \mathbf{x}, \quad \lambda((u, v)|\mathbf{x}) = \lambda\gamma^{n[C_r^q(u,v);\mathbf{x}]}, \quad (7)$$

and

$$\text{for } (\xi, t) \in \mathbf{x}, \quad \lambda((\xi, t)|\mathbf{x}) = \lambda\gamma^{n[C_r^q(\xi,t);\mathbf{x} \setminus (\xi,t)]}, \quad (8)$$

where $n[C_r^q(y, z); \mathbf{x}] = \sum_{(\xi,t) \in \mathbf{x}} \mathbb{1}\{||y - \xi|| \leq r, |z - t| \leq q\}$ is the number of points in \mathbf{x} which are in a cylinder $C_r^q(y, z)$. Although the Strauss point process was originally intended as a model of clustering, it can only be used to model inhibition, because the parameter γ cannot be greater than 1. If we take $\gamma > 1$, the density function of Strauss model is not integrable, so it does not define a valid probability density.

As mentioned, the Strauss point process model only achieves the inhibition structure. In the spatial framework, two ways are introduced to overcome this problem that we extend to the spatio-temporal framework hence defining two new spatio-temporal Gibbs point process models.

A first way is to consider an upper bound for the number of neighboring points that interact. In this case, Raeisi et al. (2021) defined a homogeneous spatio-temporal Geyer saturation point process by density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})} \prod_{(\xi,t) \in \mathbf{x}} \gamma^{\min\{s, n^*[C_r^q(\xi,t);\mathbf{x}]\}}, \quad (9)$$

with respect to a unit rate Poisson point process on W , where s is a saturation parameter and $n^*[C_r^q(\xi, t); \mathbf{x}] = n[C_r^q(\xi, t); \mathbf{x} \setminus (\xi, t)] = \sum_{(u,v) \in \mathbf{x} \setminus (\xi,t)} \mathbb{1}\{||u - \xi|| \leq r, |v - t| \leq q\}$.

A second way is to introduce a hardcore condition to the Strauss density (6). Hence, we can define a Strauss hardcore model in the spatio-temporal context.

Definition 1 We define the *spatio-temporal Strauss hardcore point process* as the point process with density

$$f(\mathbf{x}) = c\lambda^{n(\mathbf{x})}\gamma^{S_r^q(\mathbf{x})}\mathbb{1}\{||\xi - \xi'|| > h_s \text{ or } |t - t'| > h_t; \forall (\xi, t) \neq (\xi', t') \in \mathbf{x}\}, \quad (10)$$

where $0 < h_s < r$ and $0 < h_t < q$.

The model could be used to model clustering patterns with a softer attraction between the points like a pattern with a combination of interaction terms that show repulsion between the points at a small scale and attraction between the points at a larger scale. The Papangelou conditional intensity of a homogeneous spatio-temporal Strauss hardcore point process for $(u, v) \notin \mathbf{x}$ is obtained

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda \gamma^{n[C_r^q(u, v); \mathbf{x}]} \mathbb{1}\{|\xi - u| > h_s \text{ or } |t - v| > h_t; \forall (\xi, t) \in \mathbf{x}\} \\ &= \lambda \gamma^{n[C_r^q(u, v); \mathbf{x}]} \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{(\xi, t) \notin C_{h_s}^{h_t}(u, v)\}. \end{aligned} \quad (11)$$

We can define inhomogeneous versions of all above models by replacing the constant λ by a function $\lambda(\xi, t)$, inside the product operator over $(\xi, t) \in \mathbf{x}$, that expresses a spatio-temporal trend, which can be a function of the coordinates of the points and depends on covariate information.

2.2 Multi-scale Gibbs point process models

Since most natural phenomena exhibit dependence at multiple scales as earthquake (Siino et al., 2017;2018) and forest fire occurrences (Gabriel et al., 2017), single-scale Gibbs point process models are unrealistic in many applications. This motivates us and other statisticians to construct multi-scale generalizations of the classical Gibbs models. Baddeley et al. (2013) proposed hybrid models as a general way to generate multi-scale processes combining Gibbs processes. Given m densities f_1, f_2, \dots, f_m of Gibbs point processes, the hybrid density is defined as $f(\mathbf{x}) = c f_1(\mathbf{x}) \times f_2(\mathbf{x}) \times \dots \times f_m(\mathbf{x})$ where c is a normalization constant.

Iftimi et al. (2018) extended the hybrid approach for an area-interaction model to the spatio-temporal framework where the density is given by

$$f(\mathbf{x}) = c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{-\ell(\cup_{(\xi, t) \in \mathbf{x}} C_{r_j}^{q_j}(\xi, t))}, \quad (12)$$

with respect to a unit rate Poisson process on W , where (r_j, q_j) are pairs of irregular parameters of the model and γ_j are interaction parameters, $j = 1, \dots, m$.

In the same way, Raeisi et al. (2021) defined a spatio-temporal multi-scale Geyer saturation point process with density

$$f(\mathbf{x}) = c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{\min\{s_j, n(C_{r_j}^{q_j}(\xi, t); \mathbf{x})\}} \quad (13)$$

with respect to a unit rate Poisson process on W , where $c > 0$ is a normalizing constant, $\lambda \geq 0$ is a measurable and bounded function, $\gamma_j > 0$ are the interaction parameters.

Similarly, a hybrid version of spatio-temporal Strauss model can be defined by hybridization.

Definition 2 We define the *spatio-temporal hybrid Strauss point process* with density

$$f(\mathbf{x}) = c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{S_{r_j}^{q_j}(\mathbf{x})}, \quad (14)$$

with respect to a unit rate Poisson process on W .

Note that we called the model (14) hybrid rather than multi-scale. The model (14) can cover inhibition structure because $0 < \gamma_j < 1, \forall j \in \{1, \dots, m\}$. However, it can take into account clustering if one of densities in hybrid is the one of a hardcore process.

2.3 Hybrid Strauss hardcore point process

The hybrid Gibbs point process models do not necessarily include m same Gibbs point process models (see Baddeley et al., 2015 sect. 13.8). Badreldin et al. (2015) applied a spatial hybrid model including a hardcore density to model strong inhibition at very short distances, Geyer density for cluster structure in short to medium distances and a Strauss density for a randomness structure in larger distances to the spatial pattern of the halophytic species distribution in an arid coastal environment. Wang et al. (2020) fitted a spatial hybrid Geyer hardcore point process on the tree spatial distribution patterns. In this section, we extend this type of hybrids to the spatio-temporal context.

Definition 3 We define the *spatio-temporal hybrid Strauss hardcore point process* with density

$$f(\mathbf{x}) = c \prod_{(\xi, t) \in \mathbf{x}} \lambda(\xi, t) \prod_{j=1}^m \gamma_j^{S_{r_j}^{q_j}(\mathbf{x})} \times \mathbb{1}\{ \|\xi' - \xi''\| > h_s \text{ or } |t' - t''| > h_t; \forall (\xi', t') \neq (\xi'', t'') \in \mathbf{x} \}, \quad (15)$$

where $0 < h_s < r_1 < \dots < r_m$ and $0 < h_t < q_1 < \dots < q_m$.

The Papangelou conditional intensity of an inhomogeneous spatio-temporal hybrid Strauss hardcore process is then, for $(u, v) \notin \mathbf{x}$,

$$\begin{aligned} \lambda((u, v)|\mathbf{x}) &= \lambda(u, v) \prod_{j=1}^m \gamma_j^{n[C_{r_j}^{q_j}(u, v); \mathbf{x}]} \mathbb{1}\{ \|\xi - u\| > h_s \text{ or } |t - v| > h_t; \forall (\xi, t) \in \mathbf{x} \} \\ &= \lambda(u, v) \prod_{j=1}^m \gamma_j^{n[C_{r_j}^{q_j}(u, v); \mathbf{x}]} \prod_{(\xi, t) \in \mathbf{x}} \mathbb{1}\{ (\xi, t) \notin C_{h_s}^{h_t}(u, v) \}. \end{aligned} \quad (16)$$

Because, the conditional intensity of Gibbs models including a hardcore interaction term takes the value zero at some locations, we can rewrite it as

$$\lambda((u, v)|\mathbf{x}) = m((u, v)|\mathbf{x}) \lambda^+((u, v)|\mathbf{x}), \quad (17)$$

where $m((u, v)|\mathbf{x})$ takes only the values 0 and 1, and $\lambda^+((u, v)|\mathbf{x}) > 0$ everywhere.

The spatio-temporal hybrid Strauss hardcore point process (15) is a Markov point process in Ripley-Kelly's (1977) sense at interaction range $\max\{r_m, q_m\}$. This can be shown as in Iftimi et al. (2018) and Raeisi et al. (2021).

3 Inference

Gibbs point process models involve two types of parameters: regular and irregular parameters. A parameter is called regular if the log likelihood of density is a linear function of that parameter otherwise it is called irregular. Typically, regular parameters determine the 'strength' of the interaction, while irregular parameters determine the 'range' of the interaction. As an example, in the Strauss hardcore point process (10), the trend parameter λ and the interaction γ are regular parameters and the interaction distances r and q and the hardcore distances h_s and h_t are irregular parameters.

To determine the interaction distances r and q , there are several practical techniques, but no general statistical theory available. A useful technique is the maximum profile pseudo-likelihood approach (Baddeley and Turner, 2000). In the spatio-temporal framework, Iftimi et al. (2018) and Raeisi et al. (2021) selected feasible range of irregular parameters by analyzing the behavior of some summary statistics and the goodness-of-fit of several models with different combinations of parameters.

The hardcore interaction term $m(\cdot|\mathbf{x})$ in the conditional intensity (17) does not depend on the other parameters of the densities of Gibbs point processes. This implies that it can first be estimated and kept fixed for the sequel (Baddeley et al., 2019, p. 26). In the spatial framework, the maximum likelihood estimate of the hardcore distance in $m(\cdot|\mathbf{x})$ corresponds to the minimum interpoint distance (Baddeley et al., 2013, Lemma 7). The generalization to the spatio-temporal context with a cylindrical hardcore structure implies to consider a multi-objective minimization problem over the spatial and temporal hardcore distances h_s and h_t . The choice of our hardcore parameters needs to analyze the Pareto front of feasible solutions on the graph of spatial and temporal interpoint distances. We refer the reader to Ehrgott (2005) for a description of multi-criteria optimization and the definition of Pareto optimality. To estimate the hardcore distance h_s and h_t , we consider a feasible solution on the Pareto front as large as possible and with a ratio consistent with our knowledge of interaction mechanisms in practice.

Regular parameters can be estimated using the pseudo-likelihood method (Baddeley and Turner, 2000) or logistic likelihood method (Baddeley et al., 2014) rather than the maximum likelihood method (Ogata and Tanemura, 1981). Due to the advantage of the logistic likelihood over pseudo-likelihood for spatio-temporal Gibbs point processes (Iftimi et al., 2018; Raeisi et al., 2021), we implement the former approach in Raeisi et al. (2021, *Algorithm*

2) for regular parameter estimation of the spatio-temporal hybrid Strauss hardcore point process.

We assume that $\boldsymbol{\theta} = (\log \gamma_1, \log \gamma_2, \dots, \log \gamma_m)$ is the logarithm of interaction parameters in spatio-temporal hybrid Strauss hardcore point process (15). To estimate $\boldsymbol{\theta}$, due to (17), we just consider the points (u, v) where $m((u, v)|\mathbf{x})$ is equal to 1 in (16). By defining $S_j((u, v), \mathbf{x}) := n[C_{r_j}^{q_j}(u, v); \mathbf{x} \setminus (u, v)]$ in (16), we can thus write $\lambda_{\boldsymbol{\theta}}((u, v)|\mathbf{x}) = \lambda(u, v) \prod_{j=1}^m \exp(\theta_j S_j((u, v), \mathbf{x}))$. Hence, the logarithm of the Papangelou conditional intensity of the spatio-temporal hybrid Strauss hardcore point process for $(u, v) \in W$ which satisfies in hardcore condition, i.e. $m((u, v)|\mathbf{x}) = 1$ in (16), is

$$\begin{aligned} \log \lambda((u, v)|\mathbf{x}) &= \log \lambda(u, v) + \sum_{j=1}^m (\log \gamma_j) S_j((u, v), \mathbf{x}) \\ &= \log \lambda(u, v) + \boldsymbol{\theta}^\top \mathbf{S}((u, v), \mathbf{x}) \end{aligned} \quad (18)$$

corresponding to a linear model in $\boldsymbol{\theta}$ with offset $\log \lambda(u, v)$ where $\mathbf{S}((u, v), \mathbf{x}) = [S_1((u, v), \mathbf{x}), S_2((u, v), \mathbf{x}), \dots, S_m((u, v), \mathbf{x})]^\top$ is a sufficient statistics.

By considering a set of dummy points \mathbf{d} from an independent Poisson process with intensity function ρ , we obtain by defining the Bernoulli variables $Y((\xi, t)) = \mathbb{1}_{\{(\xi, t) \in \mathbf{x}\}}$ for $(\xi, t) \in \mathbf{x} \cup \mathbf{d}$ that the logit of $P(Y((\xi, t))) = 1$ is equal to $\log \frac{\lambda_{\boldsymbol{\theta}}((\xi, t)|\mathbf{x})}{\rho(\xi, t)}$. Under regularity conditions, the log-logistic likelihood

$$\begin{aligned} \log LL(\mathbf{x}, \mathbf{d}; \boldsymbol{\theta}) &= \sum_{(\xi, t) \in \mathbf{x}} \log \frac{\lambda_{\boldsymbol{\theta}}((\xi, t)|\mathbf{x})}{\lambda_{\boldsymbol{\theta}}((\xi, t)|\mathbf{x}) + \rho(\xi, t)} \\ &\quad + \sum_{(\xi, t) \in \mathbf{d}} \log \frac{\rho(\xi, t)}{\lambda_{\boldsymbol{\theta}}((\xi, t)|\mathbf{x}) + \rho(\xi, t)}, \end{aligned} \quad (19)$$

admits a unique maximum. By consequence, the estimation of $\boldsymbol{\theta}$ in the Papangelou conditional intensity is equivalent to the estimation of logistic regression parameters, already implemented by using standard software for GLMs. The logistic regression

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

is a linear model in $\boldsymbol{\theta}$ with offset $\log \frac{\lambda(\xi, t)}{\rho(\xi, t)}$. We use the approach of Raiesi et al. (2021) for data and dummy points such that $m(\cdot|\mathbf{x}) = 1$. We also consider that $\lambda(\xi, t) = \beta \mu(\xi, t)$, where $\mu(\xi, t)$ is a trend preliminary estimated with spatio-temporal covariates.

4 Simulation study

Due to the markovian property of the spatio-temporal hybrid Strauss hardcore point process (15), its Papangelou conditional intensity at a point thus depends

Table 1 Parameter combinations of three hybrid Strauss hardcore point process models used in simulation study.

Model	Values of parameter			
	Regular parameters		Irregular parameters	
	λ	γ	r, q	h_s, h_t
<i>Model 1</i>	70	(0.8, 0.8)	(0.05, 0.1)	(0.01, 0.01)
<i>Model 2</i>	50	(1.5, 1.5)	(0.05, 0.1)	(0.01, 0.01)
<i>Model 3</i>	70	(0.5, 1.5)	(0.05, 0.1)	(0.01, 0.01)

Table 2 Mean and 95% interval regular parameter estimates of the three hybrid Strauss hardcore point process models used in simulation study.

True values	Mean	95% CI
<i>Model 1</i>		
$\lambda = 70$	71.43	(69.16, 73.70)
$\gamma_1 = 0.8$	0.89	(0.78, 1.00)
$\gamma_2 = 0.8$	0.78	(0.74, 0.82)
<i>Model 2</i>		
$\lambda = 50$	50.84	(48.99, 52.68)
$\gamma_1 = 1.5$	1.41	(1.23, 1.60)
$\gamma_2 = 1.5$	1.46	(1.38, 1.54)
<i>Model 3</i>		
$\lambda = 70$	71.67	(69.18, 74.15)
$\gamma_1 = 0.5$	0.50	(0.43, 0.57)
$\gamma_2 = 1.5$	1.49	(1.42, 1.55)

only on that point and its neighbors in \mathbf{x} . Hence, We can design simulation approach by Markov chain Monte Carlo algorithms.

Gibbs point process models can be simulated a birth-death Metropolis-Hastings algorithm that typically requires only computation of the Papanagelou conditional intensity (Møller and Waagepetersen, 2004). Raeisi et al. (2021) extended the birth-death Metropolis-Hastings algorithm to the spatio-temporal context that we adapt here for simulating the spatio-temporal hybrid Strauss hardcore point process.

We implement the estimation and simulation algorithms in R (R Core Team, 2016) and generate simulations of three stationary spatio-temporal hybrid Strauss hardcore point processes specified by a conditional intensity of the form (16) in $W = [0, 1]^3$. The parameter values used for the simulations are reported in Table 1. The spatial and temporal radii r and q , spatial and temporal hardcores h_s and h_t , are treated as known parameters.

We generate 100 simulations of each specified model. Boxplots of parameter estimates λ , γ_1 , and γ_2 obtained from the logistic likelihood estimation method for each model are shown in Figure 1. The red horizontal lines represent the true parameter values. Point and interval parameter estimates λ , γ_1 , and γ_2 are reported in Table 2. Most of the estimated parameter values are close to the true values for three models. Due to visual and computational comparisons, we conclude that the logistic likelihood approach performs well for spatio-temporal hybrid Strauss hardcore point processes.

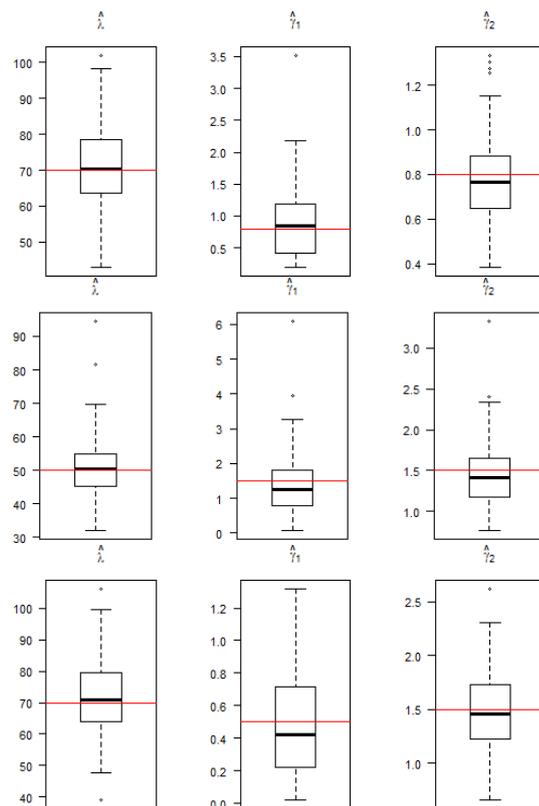


Fig. 1 Boxplots of regular parameter estimates of the hybrid Strauss hardcore point process obtained from the logistic likelihood estimation methods. Up to down: *Model 1*, *Model 2*, and *Model 3*

5 Application

In this section we aim to model the interactions of forest fire occurrences across a range of spatio-temporal scales.

5.1 Data description

The `clmfires` dataset available in `spatstat` package records the occurrences of forest fires in the region of Castilla-La Mancha, Spain (Figure 2, left) from 1998 to 2007. The study area is approximately $400 \text{ km} \times 400 \text{ km}$. The `clmfires` dataset has already been used in some academic works devoted to the point process theory (see e.g. Juan et al., 2010; Gomez-Rubio, 2020, sect. 7.4.2; Myllymäki et al., 2020). The dataset has two levels of precision: from 1998 to 2003 locations were recorded as the centroids of the corresponding “district

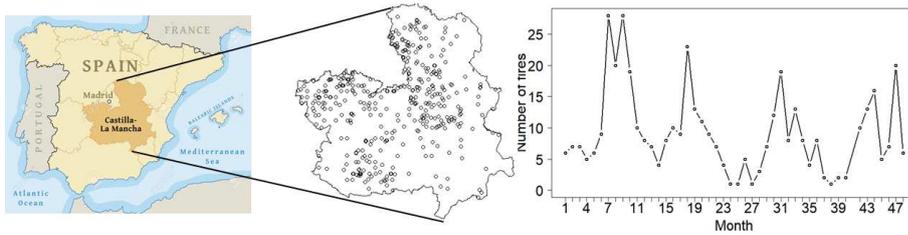


Fig. 2 Left: Map of region of Castilla-La Mancha (Spain). Middle: Forest fire locations. Right: monthly numbers of fires recorded between January 2004 and December 2007 with burnt areas, spatial distances and time distances respectively bigger than 5 *ha*, 0.2 *km* and 100 *days*.

units”, while since 2004 locations correspond to the exact UTM coordinates of the fire locations.

Due to the low precision of fire locations for the years 1998 to 2003 (Gomez-Rubio 2020, sect. 7.4.2), we focus on fires in the period 2004 to 2007. In this period, we consider large forest fires with burnt areas larger than 5 *ha*. Figure 2 (middle) shows the point pattern of 432 wildfire locations onto the spatial region.

Due to memory constraints and availability of climate covariates in months, we consider monthly fire occurrences. The temporal component of the process takes integer values from 1 to 48. We thus consider $W = S \times T$ where S is the region of Castilla-la-Mancha and $T = \{1, 2, \dots, 48\}$ corresponds to the months since January 2004. Figure 2 (right) shows the monthly number of fires occurring during our time period. We observe seasonal effects with notably large numbers of fires in summer that could be caused by high temperatures and low precipitations in this period and also by human activities.

In point pattern analysis, the spatial (spatio-temporal) inhomogeneity of patterns is notably driven by covariates. The `clmfires` dataset contains four environmental covariates that we include in our analysis: *elevation*, *orientation*, *slope* and *land use*. The covariates are known on a spatial grid with pixels of 4 *km* \times 4 *km*, resulting in a total of 10,000 pixels. The *land use* is a factor-valued covariate whereas the others are real-valued covariates. We also consider weather data freely provided by the *WorldClim* database¹ and containing monthly maximum temperatures ($^{\circ}C$) and total precipitations (*mm*). Figure 3 illustrates the environmental covariates, which are considered fixed during our temporal period, and the climate covariates in January 2007.

5.2 Estimation

First, we estimate the trend function by considering a generalized linear model (GLM) on covariates. Then, by an exploratory analysis using spatio-temporal

¹ <https://www.worldclim.org>

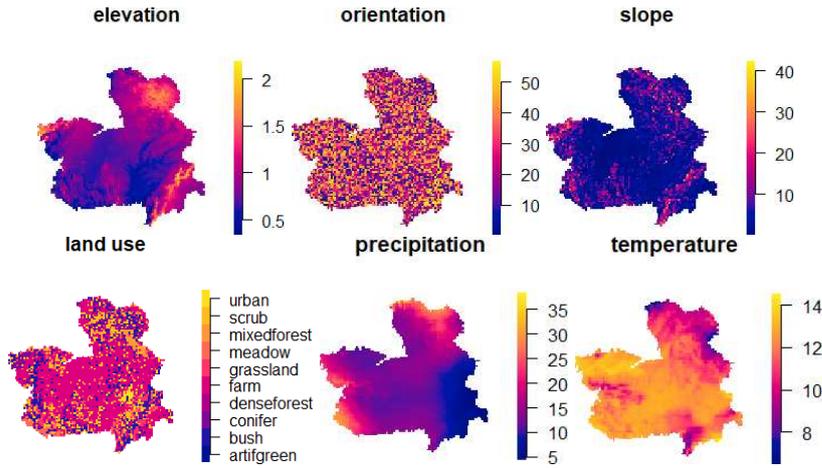


Fig. 3 Image plot of environmental covariates (*elevation*, *orientation*, *slope* and *land use*) and climate covariates (*precipitation* and *temperature*) in January 2007.

summary statistics we approximate the hardcore parameters and the interaction ranges. Finally, we use the logistic likelihood approach described in Section 3 for the estimation of regular parameters of our model with the trend function estimated in the preliminary step.

5.2.1 Trend estimate

Since covariates are available on a spatial grid, we restrict our attention on the related grid centers $\xi_i, i = 1, \dots, 10000$ and months $\{t_j\}_{j=1, \dots, 48} \in T$ and consider $N_{ij} | \lambda(\xi_i, t_j) \sim \text{Poisson}(\lambda(\xi_i, t_j))$ where N_{ij} is the number of forest fires in the i^{th} grid center at month t_j .

Following Raeisi et al. (2021), by considering a GLM with Poisson response, we obtain:

$$\log \lambda(\xi_i, t_i) = \beta_0 + \sum_{k=1}^6 \beta_k Z_k(\xi_i, t_i), \quad (21)$$

where $Z_k(\xi_i, t_i)$, $k = 1, \dots, 6$, are the environmental and climatic covariates at point (ξ_i, t_i) and β_0, β_k , $k = 1, \dots, 6$ are the coefficients to estimate. As said before, we consider the same values for environmental covariates over time. A straightforward way to fit a GLM in R is to use the function `glm`. Table 3 reports the estimated coefficients in (21) and their significance level by a two-tailed Student's t-test. Coefficients higher (respectively lower) than zero imply an increase (resp. decrease) of the expected mean number of forest fires when the covariate value increase (resp. decrease). Those related to *elevation* and *temperature* are positively significant, showing that these two covariates favors the ignition of wildfires. At the opposite, the covariate *precipitation* has a negative significant coefficient indicating that an increase of the amount of

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

Coefficients	Estimate	Standard error	p -value
β_0 (intercept)	-8.468	0.298	$< 2 \times 10^{-16}$ ***
β_1 (elevation)	0.546	0.164	0.001 ***
β_2 (orientation)	0.005	0.003	0.114
β_3 (slope)	-0.019	0.01	0.054
β_4 (land use)	-0.009	0.024	0.689
β_5 (precipitation)	-0.007	0.002	0.003 **
β_6 (temperature)	0.054	0.006	$< 2 \times 10^{-16}$ ***

precipitation induces a decrease in the mean number of forest fires. The *land use* appears not significantly different from zero, it can be explained by the low spatial resolution of the covariates.

5.2.2 Irregular parameter estimates

We have two types of irregular parameters in our spatio-temporal Gibbs point process. On the one hand, the hardcore distances that we can choose among the feasible solutions on the Pareto front of spatial and temporal interpoint distances. According to Figure 4, we choose on the Pareto front the unique feasible solution in our case that gives non-zero values for the two hardcore distances, i.e. $h_s = 0.35$ km and $h_t = 1$ month. On the other hand, for the nuisance parameters m , r_j and q_j , $j = 1, \dots, m$, there is no common method for estimating them. 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 after the regular parameter estimation step (Raeisi et al., 2021). We chose 25 configurations of reasonable range for the nuisance parameters using a preliminary spatio-temporal exploratory analysis of the interaction ranges done with the inhomogeneous pair correlation function, the maximum nearest neighbor distance and the temporal auto-correlation function. We fitted the spatio-temporal hybrid Strauss point process model for a range of ad-hoc values $r_j \in (0.35, 20]$, $q_j \in \{2, \dots, 15\}$, $j = 1, \dots, m$ and $m \in \{1, \dots, 6\}$. The minimum AIC is obtained for the combination given in Table 4.

5.2.3 Regular parameter estimates

We consider the logistic likelihood method investigated in Section 3 to estimate the regular parameters. We simulate dummy points from an inhomogeneous Poisson point process with intensity $\rho(\xi, t) = C\hat{\lambda}(\xi, t)/\nu$ where $C = 4$ by a classical rule of thumb in the logistic likelihood approach, $\hat{\lambda}$ is the estimated trend and $\nu = 4 \times 4 \times 1$ is the volume of a grid cell on one month. In order to satisfy the hardcore condition in (17), we remove dummy points at spatial and temporal distances respectively less than h_s and h_t . Estimated regular parameters are provided in Table 4.

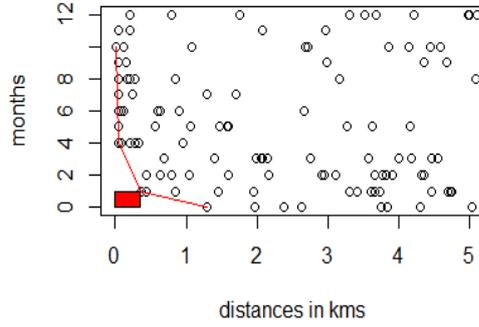


Fig. 4 Spatial and temporal interpoint distances respectively lower than 5 kms and 12 months (black circles). The red line corresponds to the Pareto front and the red rectangle to the hardcore domain.

Table 4 Parameter estimates for $m = 6$.

		Irregular parameters					
r		0.5	1	1.5	6	15	20
q		2	4	6	8	12	15
		Estimated regular parameters					
		$\hat{\gamma}_1 = 2.56$	$\hat{\gamma}_2 = 2.24$	$\hat{\gamma}_3 = 4.65$	$\hat{\gamma}_4 = 0.88$	$\hat{\gamma}_5 = 1.17$	$\hat{\gamma}_6 = 0.81$

5.3 Goodness-of-fit

The goodness-of-fit is accomplished by simulating point patterns from the fitted model. The first diagnostic can be formulated by summary statistics of point processes. As the second-order characteristics carry most of the information on the spatio-temporal structure (Stoyan, 1992 ; Gonzalez et al., 2016), we only consider the pair correlation function (g -function).

We generate $n_{sim} = 99$ simulations from the fitted hybrid Strauss hardcore model and compute the corresponding second-order summary statistics $g_i(u, v)$, $i = 1, \dots, n_{sim}$, for fixed spatio-temporal distances (u, v) . We then build upper and lower envelopes:

$$U(u, v) = \max_{1 \leq i \leq n_{sim}} g_i(u, v), \quad L(u, v) = \min_{1 \leq i \leq n_{sim}} g_i(u, v), \quad (22)$$

and compare the summary statistics obtained from the data, $g_{obs}(u, v)$, to the pointwise envelopes. If it lies outside the envelopes at some spatio-temporal distances (u, v) , then we reject at these distances the hypothesis that our data come from our fitted model. Figure 5 shows the spatio-temporal inhomogeneous g -function computed on our dataset (blue) and the envelopes obtained from the fitted model (light grey); $g_{obs}(u, v)$ lies inside the envelopes for all

(u, v) , meaning that the hybrid Strauss hardcore model is suitable for the data.

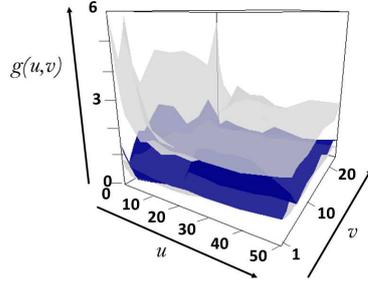


Fig. 5 Envelopes of the spatio-temporal inhomogeneous g -function obtained from simulations of the fitted spatio-temporal hybrid Strauss hardcore point process (light grey). The blue surface corresponds to g_{obs} . Temporal separations v are in *month* and spatial distances u are in *kilometer*.

In addition, we compute global envelopes and p -value of the spatio-temporal g -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, 2020). For each point pattern, we consider the long vector T_i , $i = 1, \dots, n_{sim}$ (resp. T_{obs}) merging the $g_i(\cdot, v)$ (resp. $g_{obs}(\cdot, v)$) 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, 2020). The final p -value is obtained by

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

Due to the global p -value $p_{erl} = 0.59$ and the absence of significant regions, that corresponds here to pairs of spatial and temporal distances where the statistics is significantly above or below the envelopes (see Figure 6 and `GET` package), we conclude that our hybrid Strauss hardcore model can not be rejected a significance level of 1%.

Conclusion

In this paper, we introduced the spatio-temporal Strauss hardcore point process. The Strauss hardcore model is a Gibbs model for which points are pushed to be at a hardcore distance apart and repel up to a interaction distance which is larger than the hardcore distance. As in Raeisi et al. (2021), inference and

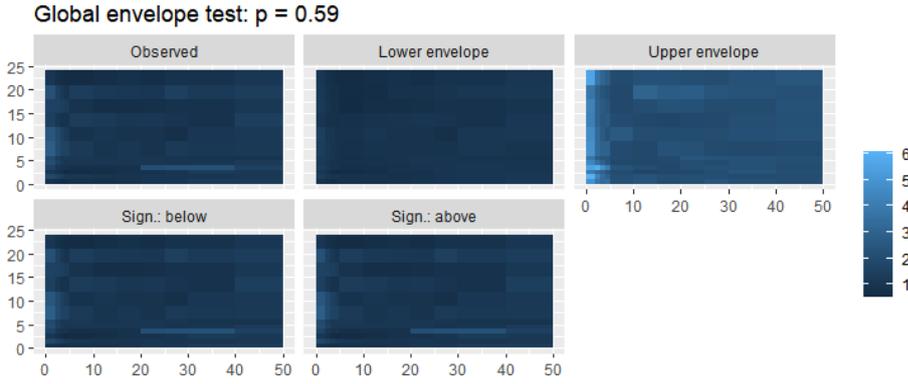


Fig. 6 Top: estimated pair correlation function \hat{g}_{obs} , lower E_L and upper E_U bounds of the 99% global rank envelope (ERL). Bottom: differences $E_{obs} - E_L$ and $E_U - E_{obs}$. Negative values (if any) are represented in red and lead to reject the fitted model. Values on the horizontal axis are in kilometers and those on the vertical axis are in months.

simulation of the model were performed with logistic likelihood and birth-death Metropolis-Hasting algorithm, respectively. A multi-scale version of the model was introduced and applied to wildfires to take into account structural complexity of forest fire occurrences in space and time. We based our model validation on both pointwise and global envelopes and p -value based on the Extreme Rank Length (ERL) measure of the spatio-temporal inhomogeneous pair correlation function. Our model could be suitable in other environmental and ecological frameworks, when we want to deal with the complexity of mechanisms governing attraction and repulsion of entities (particles, cells, plants...).

In spatio-temporal Gibbs point process models, the heterogeneity can be captured by estimating a non-constant trend. This spatio-temporal trend is typically considered as a function of covariates by estimating fixed effects in a generalized linear model as we carried out it in this paper and also in Ifitimi et al. (2018) and Raeisi et al. (2021). A different approach consists in considering Gibbs models with both random and fixed effects (e.g. see Illian and Hendrichsen, 2010) to take into account complex patterns of spatio-temporal heterogeneity. Vihrs et al. (2020) proposed a new modeling approach for this case and embedded spatially structured Gaussian random effects in trend function of a pairwise interaction process. They introduced the spatial log-Gaussian Cox Strauss point process to capture both structures; aggregation in small-scale and repulsion in large-scale. Rather than spatial pairwise interaction processes in single-scale, we now focus on models derived from the multi-scale classes of combinations of Gibbs and log-Gaussian Cox point processes in space and time, to which we refer to as Cox-Gibbs models. We propose to embed spatio-temporally structured Gaussian random effects in the Gibbs trend function. Due to the hierarchical structure of such models, we can

formulate and estimate them within a Bayesian hierarchical framework, using the INLA approach.

References

1. Andersen IT, Hahn U (2016) Matérn thinned Cox processes. *Spatial Statistics* 15:1–21.
2. Baddeley A, Coeurjolly JF, Rubak E, Waagepetersen R (2014) Logistic regression for spatial Gibbs point processes. *Biometrika* 101(2):377–392.
3. Baddeley A, Rubak E, Turner R (2015) *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London.
4. Baddeley A, Turner R (2000) Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics* 42:283–322.
5. Baddeley A, Turner R, Mateu J, Bevan A (2013) Hybrids of Gibbs point process models and their implementation. *Journal of Statistical Software* 55(11):1–43.
6. Baddeley A, Rubak E, Turner R (2019) Leverage and influence diagnostics for Gibbs spatial point processes. *Spatial Statistics* 29:15–48.
7. Badreldin N, Uria-Diez J, Mateu J, Youssef A, Stal C, El-Bana M, Magdy A, Goossens R (2015) A spatial pattern analysis of the halophytic species distribution in an arid coastal environment. *Environmental Monitoring and Assessment* 187: 1-15.
8. Cronie O, van Lieshout M (2015) A J-function for inhomogeneous spatio-temporal point processes. *Scandinavian Journal Statistics* 42(2):562–579.
9. Das I, Stein A (2016) Application of the Multitype Strauss Point Model for Characterizing the Spatial Distribution of Landslides. *Mathematical Problems in Engineering*. <http://dx.doi.org/10.1155/2016/1612901>.
10. Dereudre D (2019) Introduction to the theory of Gibbs point processes. In: Coupier D (ed) *Stochastic Geometry*. Lecture Notes in Mathematics vol 2237, Springer, Cham, pp 181–229.
11. Dereudre D, Lavancier F (2017) Consistency of likelihood estimation for Gibbs point processes. *Annals of Statistics* 45(2):744–770.
12. Ehrgott M (2005) *Multicriteria Optimization*. Springer Verlag, Berlin.
13. Juan P (2020) Spatio-temporal hierarchical Bayesian analysis of wildfires with Stochastic Partial Differential Equations. A case study from Valencian Community (Spain). *Journal of Applied Statistics* 47(5): 927-946. <https://doi.org/10.1080/02664763.2019.1661360>.
14. Juan P, Mateu J, Díaz-Avalos C (2010) Characterizing spatial-temporal forest fire Patterns. In METMA V: International Workshop on Spatio-Temporal Modelling. Santiago de Compostela, Spain.
15. Hurtut T, Landes PE, Thollot J, Gousseau Y, Drouilhet R, Coeurjolly JF (2009) Appearance-guided synthesis of element arrangements by example, *Proceedings SIG-GRAPH Symposium on Non-Photorealistic Animation and Rendering (NPAR)*, 51–60.
16. Illian JB, Hendrichsen DK (2010) Gibbs Point Process Models with Mixed Effects. *Environmetrics* 21:341-353.
17. Illian J, Penttinen A, Stoyan H, Stoyan D (2008) *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley & Sons, Chichester.
18. Iftimi A, van Lieshout MC, Montes F (2018) A multi-scale area-interaction model for spatio-temporal point patterns. *Spatial Statistics* 26:38-55.
19. Gabriel E, Opitz T, Bonneau F (2017) Detecting and modeling multi-scale space-time structures: the case of wildfire occurrences. *Journal of the French Statistical Society* 158(3):86-105.
20. Glass L, Tobler WR (1971) Uniform distribution of objects in a homogeneous field: Cities on a plain. *Nature* 233:67-68.
21. Gomez-Rubio V (2020) *Bayesian Inference with INLA*. Chapman & Hall-CRC, Boca Raton.
22. Gonzalez JA, Rodriguez-Cortes FJ, Cronie O, Mateu J (2016) Spatio-temporal point process statistics: A review. *Spatial Statistics* 18:505–544.

23. Matérn B (1960) Spatial variation. Stochastic models and their application to some problems in forest surveys and other sampling investigations. *Medd. Statens Skogsforskningsinst* 49(5):1–144.
24. Matérn B (1986) Spatial Variation. In: *Lecture Notes in Statistics*, vol 36. Springer, New York.
25. Mattfeldt T, Eckel S, Fleischer F, Schmidt V (2006) Statistical analysis of reduced pair correlation functions of capillaries in the prostate gland. *Journal of Microscopy* 223(2):107–119.
26. Mattfeldt T, Eckel S, Fleischer F, Schmidt V (2007) Statistical modelling of the geometry of planar sections of prostatic capillaries on the basis of stationary Strauss hard-core processes. *Journal of Microscopy* 228:272–281.
27. Mattfeldt T, Eckel S, Fleischer F, Schmidt V (2009) Statistical analysis of labelling patterns of mammary carcinoma cell nuclei on histological sections. *Journal of Microscopy* 235:106–118.
28. Møller J, Diaz-Avalos C (2010). Structured spatio-temporal shot-noise cox point process models, with a view to modelling forest fires. *Scandinavian Journal of Statistics* 37(1):2–25. <https://www.jstor.org/stable/41000913>
29. Møller J, Waagepetersen R P (2004) *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
30. Myllymäki M, Kuronen M, Mrkvička T (2021) Testing global and local dependence of point patterns on covariates in parametric models. *Spatial Statistics* 42:100436.
31. Myllymäki M, Mrkvička T (2020) *GET: Global envelopes in R*. submitted to *Journal of Statistical Software*. <http://arxiv.org/abs/1911.06583>
32. Myllymäki M, Mrkvička T, Grabarnik P, Seijo H, Hahn U (2017) Global Envelope Tests for Spatial Processes. *Journal of the Royal Statistical Society: Series B* 79:381–404. doi:10.1111/rssb.12172.
33. Najafabadi ATP, Gorgani F, Najafabadi MO (2015) Modeling forest fires in Mazandaran Province, Iran. *Journal of Forestry Research* 26:851–858. <https://doi.org/10.1007/s11676-015-0107-z>
34. Ogata Y, Tanemura M (1981) Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics* 33:315–338.
35. Opitz T, Bonneu F, Gabriel E (2020) Point-process based Bayesian modeling of space-time structures of forest fire occurrences in Mediterranean France. *Spatial Statistics* 40:100429. <https://doi.org/10.1016/j.spasta.2020.100429>
36. Papangelou F (1974) The conditional intensity of general point processes and an application to line processes. *Probability Theory and Related Fields* 28(3):207–226.
37. Pereira P, Turkman K, Amaral-Turkman M, Sa A, Pereira J (2013) Quantification of annual wildfire risk; a spatio-temporal point process approach. *Statistica* 73(1):55–68. <https://doi.org/10.6092/issn.1973-2201/3985>
38. Pimont F, Fargeon H, Opitz T, Ruffault J, Barbero R, Martin-StPaul N, Rigolot E, Riviere M, Dupuy J-L (2021) Prediction of regional wildfire activity in the probabilistic Bayesian framework of firelihood, *Ecological Applications*. doi:10.1002/eap.2316.
39. R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.
40. Raeisi M, Bonneu F, Gabriel E (2019) On spatial and spatio-temporal multi-structure point process models. *Annales de l'Institut de Statistique de l'Université de Paris* 63(2-3):97–114.
41. Raeisi M, Bonneu F, Gabriel E (2021) A spatio-temporal multi-scale model for Geyer saturation point process: application to forest fire occurrences. *Spatial Statistics* 41:100492.
42. Ripley BD (1988) *Statistical Inference for Spatial Processes*. Cambridge University Press.
43. Ripley BD, Kelly FP (1977). Markov point processes. *Journal of the London Mathematical Society* 15(1):188–192. <https://doi.org/10.1112/jlms/s2-15.1.188>
44. Rue H, Martino S, Chopin N (2009) Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71(2):319–92. doi:10.1111/j.1467-9868.2008.00700.x.

45. Serra L, Saez M, Mateu J, Varga D, Juan P, Diaz-Avalos C, Rue H (2014a) Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: the case of Catalonia, 1994–2008. *Environmental and Ecological Statistics* 21(3):531–563. <https://doi.org/10.1007/s10651-013-0267-y>
46. Serra L, Saez M, Varga D, Tobías A, Juan P, Mateu J (2012) Spatio-temporal Modelling Of Wildfires In Catalonia, Spain, 1994–2008, Through Log Gaussian Cox Processes. *WIT Transactions on Ecology and the Environment* 158:39–49. <https://doi.org/10.2495/FIVA120041>
47. Serra L, Saez M, Juan P, Varga D, Mateu J (2014b) A spatio-temporal Poisson hurdle point process to model forest fires. *Stochastic Environmental Research and Risk Assessment* 28(7):1671–1684. <https://doi.org/10.1007/s00477-013-0823-x>
48. Siino M, Adelfio G, Mateu J, Chiodi M, D’Alessandro A (2017) Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity. *Stochastic Environmental Research and Risk Assessment* 31(7):1633–1648.
49. Siino M, D’Alessandro A, Adelfio G, Scudero S, Chiodi M (2018) Multiscale processes to describe the eastern sicily seismic sequences. *Annals of Geophysics* 61(2).
50. Stoyan D (1992) Statistical estimation of model parameters of planar Neyman-Scott cluster processes. *Metrika* 39(1):67–74. <https://doi.org/10.1007/BF02613983>
51. Stoyan D (2016) Point process statistics: application to modern and contemporary art and design. *Journal of Mathematics and the Arts* 10:20–34.
52. Sweeney S, Gómez-Antonio M (2016) Localization and industry clustering econometrics: An assessment of Gibbs models for spatial point processes. *Journal of Regional Science* 56:257–287.
53. Taylor DB, Dhillon HS, Novlan TD, Andrews J G (2012) Pairwise interaction processes for modeling cellular network topology. *Proceedings IEEE GLOBECOM in Anaheim*.
54. Teichmann J, Ballani F, Boogaart KG van den (2013) Generalizations of Matérn’s hard-core point processes. *Spatial Statistics* 9:33–53.
55. Turner R (2009) Point patterns of forest fire locations. *Environmental and Ecological Statistics* 16:197–223. <https://doi.org/10.1007/s10651-007-0085-1>
56. van Lieshout MNM (2000) *Markov Point Processes and Their Applications*. Imperial College Press, London.
57. Vihrs N, Møller J, Gelfand AE (2021) Approximate Bayesian inference for a spatial point process model exhibiting regularity and random aggregation, *Scandinavian Journal of Statistics: Theory and Applications*. <https://doi.org/10.1111/sjos.12509>
58. Wang X, Zheng G, Yun Z, Moskal LM (2020). Characterizing tree spatial distribution patterns using discrete Aerial Lidar Data. *Remote Sensing* 12(4): 712. <https://doi.org/10.3390/rs12040712>
59. Ying Q, Zhao Z, Zhou Y, Li R, Zhou X, Zhang H (2014). Characterizing spatial patterns of base stations in cellular networks. *Proceedings IEEE/CIC International Conference on communications in China (ICCC)*, 490–495.