



UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato in Scienze Economiche Aziendali e Statistiche

Dipartimento di Scienze Economiche Aziendali e Statistiche

Settore Scientifico Disciplinare: SECS/S01

**Models and methods for space and space-time interactions in complex
point processes with applications on earthquakes**

IL DOTTORE
Marianna Siino

IL DECANO DEL COLLEGIO
Prof. Fabio Mazzola

IL TUTOR
Prof.ssa Giada Adelfio

CICLO XXIX
ANNO CONSEGUIMENTO TITOLO 2017

To my family

Acknowledgements

I would like to thank all the people close to me, who helped me in the period of the PhD. First of all, I would like to thank my family for their love and for their support.

A special thank goes to my tutor Professor Giada Adelfio for her guidance, warm encouragement and patience. I want to express my thank to Professor Marcello Chiodi for believing me and for having been my supervisor during the first period of the PhD. My thank also goes to Professor Vito Muggeo for helping me in growing in the research and for offering valuable advice and critical comments. In general, all my acknowledgments go to all the people of my department, especially the professors in Statistics for having contributed to my professional and personal growth.

Thanks to all the colleagues of the PhD, especially Andrea, Federico, Veronica, Gianluca, Ola, Pietro, Roberto, Marco, Fabio for having had a nice time during our working hours and for have shared this experience.

My most sincere thank goes to Professor Jorge Mateu, who allowed me to spend a beautiful experience during my visiting period at Universidad Jaume I, for his kindness, professionalism, for his encouragement and his help in my research. Moreover, I would like to thank all the friends that I met in Spain (Mehdi, Ahmed, Leyla, Francisco, Pau, Rachel, Carolina, Diego, Fernando, Shema group and so on...) who helped me to grow, and

that made me feel like family.

Above all, my thank goes to God for giving me the power, spirit and endurance to complete this research, since during this period I have discovered his help to overcome difficulties.

Contents

Introduction	1
1 Analysis and modelling of spatio and spatio-temporal point processes	9
1.1 Spatial point processes	10
1.1.1 First-order and second-order properties	12
1.1.2 Model formulation	16
1.1.3 Gibbs and Hybrids of Gibbs point process models .	19
1.1.3.1 Model estimation	23
1.1.3.2 Diagnostics	25
1.2 Spatio-temporal point processes	28
1.2.1 First-order and second-order properties	30
1.2.2 Estimators of second-order characteristics	33
1.2.3 Model formulation	36
1.2.4 Spatio-temporal Log-Gaussian Cox process	37
1.2.4.1 Covariance Functions for the Gaussian random field	41
1.2.4.2 Geometric anisotropy in space	43
1.2.4.3 Model estimation	46

1.2.5	Epidemic type aftershocks stochastic model	52
1.2.5.1	Model estimation	54
2	Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity	57
2.1	Literature review	58
2.2	The study area and the geological information	59
2.3	Descriptive analysis	65
2.4	Statistical inference and modelling proposal	68
2.4.1	Inhomogeneous Poisson models	68
2.4.2	Inhomogeneous hybrid models	73
2.5	Final remarks	79
3	Spatio-temporal log-Gaussian Cox processes on earthquake events	83
3.1	Literature review	84
3.2	The study area and the dataset	89
3.3	Results	90
3.4	Some conclusions and future analysis	99
4	Testing procedure for local spatio-temporal structure	101
4.1	Literature review	102
4.2	Product density LISTA functions	105
4.2.1	Behaviour of LISTA functions	107
4.3	Statistical testing procedure for local interaction	113
4.4	Simulation study	118
4.4.1	Empirical level	118
4.4.2	Power	121
4.5	Application to earthquake data	124

Contents	iii
4.5.1 Estimation of the ETAS model: background and induced events	125
4.5.2 Application of the test based on the LISTA functions	127
4.6 Final remarks	128
5 Conclusions and future works	131
Bibliography	151

Introduction

Point process theory, a branch of the modern field of spatial and spatio-temporal statistics, is widely used to model data in form of events in a region of the plane showing a particular spatial structure, and also to represent the evolution of some random system through time. There are applications of point process models in a wide range of scientific fields such as forestry and plant ecology, epidemiology, zoology, human and social interaction, seismology and astronomy (Baddeley *et al.*, 2006; Diggle, 2013; Baddeley *et al.*, 2015).

The description of a real point process phenomena, like seismic events, often requires the definition of more complex models than the stationary Poisson process without assuming statistical independence of events. Therefore, the study of the second-order properties of a point process has a relevant role in the comprehension of the process and its realization.

Moreover, the generator process of an observed point pattern can be quite complex in different aspects: it can be a multidimensional and multiscale process, with dependence features of points both in space and time, and varying according to some underlying spatial environmental information available in the study area and temporal variables.

Given a point pattern, that is an observed realisation of a point process that

is the stochastic model governing the locations, the research attention can focus on two complementary aspects that are difficult to untangle in real data. It can be of interest to describe the first-order characteristics of the point pattern studying the relationship of the events with underlying environmental variables and/or temporal covariates that can determine the phenomena, capturing a trend or more global irregularities. The other aspect involves the description of the interaction structure between events and so of the second-order properties of the process. For instance, a specific point pattern may have various types of interaction between its points, the points may exhibit a cluster or a regularity structure or a combination of the above features.

This dissertation addresses the development of statistical methodologies to describe short-range interaction among the points in the presence of larger-scale inhomogeneity in both space and space-time contexts.

Indeed, seismic data represent an interesting situation where the study and the interpretation of features like self-similarity, long-range dependence and fractal dimension are crucial. In an observed area, earthquakes can be considered as a realization of a marked space-time point process, assuming that each event, with a given magnitude (the mark), is identified by a point in space (by its geographical coordinates) and time (fixing its time of occurrence). Moreover, the distribution of epicentres generally shows diverse interaction structures at different spatial and spatio-temporal scales varying according to the location of earthquake sources in the study area. The most prominent form of interaction structure in earthquake patterns is clustering. For example, the concentration of earthquakes in space is observed along the boundaries of major tectonic plates and along faults. On the other hand, clustering in time can be seen as a significant increase of seismic activity

immediately after large earthquakes leading to aftershock sequences.

Currently in the seismological context when events are clustered together, the modelling formulation is usually based on self-exciting point processes assuming that the occurrence of an event increases the probability of occurrence of other events in time and space (such as the Hawkes model, Hawkes and Adamopoulos (1973) or the ETAS model, Ogata (1988)).

Whereas, when an inhibitive pattern is observed, an important class of processes is represented by self-correcting processes. In these models, it is assumed that each event, that occurs in a given point in time, has an inhibiting effect to the following ones. In the seismological context, self-correcting processes are represented by the strain-release models (Vere-Jones, 1978), for which the fault fracture that generates earthquakes reduces the amount of strain present at the break point along the fault. The Strauss point process model (Strauss, 1975) with a negative association is an example of self-correcting processes.

The main characteristic of the aforementioned models is that they describe the process focusing on the interaction structure between events without considering other additional information such as geological characteristics of the study area, geographical distribution of eventual sources, etc. Only in the spatial context, Gibbs point processes have been used in the analysis of spatial interaction and inhomogeneity depending on covariate information (Anwar *et al.*, 2012; Ye *et al.*, 2015). Nevertheless, these models present some limitations such as they are not able to describe the multi-scale spatial structure, as explained in Illian *et al.* (2008) and Baddeley *et al.* (2013).

As a matter of fact, although the models cited above represent a wide variety of solutions to describe different types of spatio and spatio-temporal patterns, real phenomena often need more complex models able, for exam-

ple, to account for environmental heterogeneity while detecting interactions at several spatial scales.

For this reason, it is attractive and motivating the definition and estimation of models that account for both types of dependence (both attractive and repulsive, such as the Hybrid models or the Log-Gaussian Cox processes) while also including the effect of observed covariates in order to better describe the probability of occurrence of events in fixed regions.

The first contribution of this dissertation is the analysis of earthquake events in the Hellenic area in the spatial context (Siino *et al.*, 2016a). Using inhomogeneous hybrid models (Baddeley *et al.*, 2013), we characterise both the interaction between events happening at several spatial scales, and the spatial inhomogeneity of the process considering the geological information. This work is a very nice contribution to the literature on point processes and seismological applications, since the idea of using a hybrid approach and the distances to seismic sources as covariates is novel and potentially useful.

Nevertheless, the more natural context to study natural processes (such as earthquakes) is the spatio-temporal framework, where the events have an ordering according to the instants of occurrence. Moving from space to space-time point process, there are many open questions for the definition of new tools for testing specific assumptions in the data, and for the construction of new spatio-temporal models.

The well known statistical model applied to the analysis of spatio-temporal earthquakes is the space-time Epidemic Type Aftershock Sequence (ETAS) model (Ogata, 1988), that is a spatio-temporal marked point process model and a special case of the Hawkes process. This model can be classified as a mechanistic model (Diggle, 2013) since it embodies the underlying

scientific processes that generated the data using geological laws such as the Omori's law to explain the rate of aftershocks with time.

However, also in the spatio-temporal context, it could be crucial to include covariate information in the intensity function specification, that could be some of the determinants of the complex observed point pattern, different from the usual considered space-time-magnitude hypocentral information. Cox and log-Gaussian Cox processes (LGCPs) are models for point process phenomena that are environmentally driven (Møller *et al.*, 1998). The LGCP is the natural analogue for point process data of the linear Gaussian model for real-valued geostatistical data (Diggle *et al.*, 2013). Its principal virtue is that it provides a flexible and relatively tractable class of empirical models for describing spatially and spatio-temporal correlated phenomena specifying the moments of the underlying Gaussian Random Field. Covariate information varying in space and/or in time can be incorporated into the model in a quite straightforward fashion providing complex models able to mimic real behaviour of seismic events. These models can be useful in estimating the intensity surface of a spatio-temporal point process, in constructing spatially continuous maps of earthquake risk from spatially discrete data, and in real-time seismic activity surveillance.

Under the light of these considerations, the second contribution of this dissertation concerns the definition and estimation of several spatio-temporal log-Gaussian Cox models for the description of earthquake events in the Alaska Peninsula (Siino *et al.*, 2016b). We consider spatial geological information in the study area and different specification of the covariance structure of the Gaussian Random Field. More exactly, we detect the presence of spatial geometric anisotropy and assume a space-time non-separable covariance structure. We properly adapt minimum contrast method to es-

timate the parameters of interest using the spatio-temporal pair correlation function in the estimation procedure.

Moreover in this dissertation, we focus on spatio-temporal second-order properties of a process that can be described using a model formulation and/or second-order summary statistics (such as the K- and pair-correlation functions) (Gabriel and Diggle, 2009). These later tools are used to describe the interaction structure of a spatio-temporal point pattern and to detect deviations from the Poisson assumption. As a matter of fact, they play a fundamental role in the phase of descriptive analysis, in model validation and for testing procedures giving global information of the point pattern. An interesting question may concern if the same conclusions are valid locally, and thus, for example, in testing procedures if in subregions of the spatio-temporal window the pattern behaves differently identifying specific regions where the null hypothesis is not accepted. Indeed, a different way of looking interaction between events can be based on local versions of the summary statistics. In the analysis of spatial data, local second-order statistics play an important role giving information on how an event relates to nearby events. For instance, Cressie and Collins (2001a) define a local version of the product density function for each event, derived under a definition of a local indicator of spatial association (LISA) (Anselin, 1995). Instead, for the spatio-temporal case, Rodríguez-Cortés (2014) defines local indicator of spatio-temporal association functions based on the spatio-temporal product density function.

Driven by a research question in the analysis of earthquake events, we use the aforementioned local quantity to define a non-parametric testing approach identifying significant local structure extending the test proposed by Moraga and Montes (2011) into the spatio-temporal framework. Going

into details, to describe the seismicity of an area, it could be useful to study and compare local spatio-temporal second order-properties of independent events (that define the background seismicity) and those of the strongly correlated ones (also called aftershocks).

Therefore, the third contribution of the thesis (Siino *et al.*, 2016c) concerns the definition of a test statistic to analyse differences in local spatio-temporal structures. A simulation study is performed to illustrate that the defined tests have the prescribed size and their behaviour is reasonable under several alternative assumptions. Using the Greek earthquakes, we check which triggered events have a significant different local structure with respect to the underlying process. This analysis may reveal specific and undiscovered local structure, for instance sub-regions characterised by clusters. The developed procedure can be inserted in the framework of the techniques for spatio-temporal local detection comparing case and control data. The reference point pattern is represented by the background seismicity, instead the induced events are the case events.

The thesis is organised as follows. In Chapter 1, the theoretical framework of the spatial and spatio-temporal point process is introduced. In particular, we explain the statistical tools that are useful for the goals of the dissertation. We will explain the global and local second order summary statistics. As for the model formulation, we describe the spatial Gibbs and Hybrids of Gibbs models and the general settings of the log-Gaussian Cox process. The results for the application of the Hybrid processes on the Hellenic data are shown in Chapter 2. Chapter 3 contains the results for the application of spatio-temporal Log-Gaussian Cox process. In Chapter 4, the definition of the non-parametric test statistic for local structure, the assessment of the

performances in a simulation study and the application on earthquake data are shown. The last Chapter is devoted to a discussion of the main results and considerations for future works.

Chapter 1

Analysis and modelling of spatio and spatio-temporal point processes

In this Chapter, we introduce the fundamentals in point process theory in terms of definition of a point process, functional summary statistics and several model formulations for both spatial and spatio-temporal point processes. The aim of the explanation is to outline the tools that are useful for the goals of the thesis. Definitions and notations used throughout this Chapter were introduced in Møller and Waagepetersen (2003); Illian *et al.* (2008); Baddeley *et al.* (2015) for the spatial setting and in Gabriel and Diggle (2009); Møller and Ghorbani (2012); Diggle (2013); Gabriel *et al.* (2013); González *et al.* (2016) for the spatio-temporal one.

1.1 Spatial point processes

In this section, we report and revise concepts in spatial point pattern analysis, such as functional spatial summary statistics, some model formulations, estimation procedures and diagnostic techniques. More exactly, we specify Gibbs and Hybrids of Gibbs point process models that will be used and discussed in the first contribution of the dissertation.

The goal of spatial point process statistics is to analyse the geometrical structure of patterns formed by objects that are distributed randomly in one- two- and three-dimensional space using appropriate statistical methods (Illian *et al.*, 2008). The events in a real dataset may be biological cells, animals, locations of trees in a forest, blood particles, pores, or galaxies in the universe.

More formally, a spatial point pattern $\boldsymbol{v} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an unordered set of points in the region $W \subset \mathbb{R}^d$ where $n(\boldsymbol{v}) = n$ is the number of points, $|W| < \infty$ and usually $d = 2$. A point process model assumes that $\boldsymbol{v}$ is a realisation of a finite point process $\mathcal{X}$ in W without multiple points. We can equivalently view $\mathcal{X}$ as a random finite subset of W . Usually the following assumptions are considered for the analysis: the locations of points are measured exactly, there are not replicated points, points are mapped without omission and an event can occur at any location in the study region.

For an observed point pattern several research questions can arise and be on interest, such as: are the points uniformly distributed over the study window? Do some environmental variables explain the spatial density? Is the complete spatial randomness assumption valid for the observed pattern? Is there evidence of clustering? Is the spatial pattern consistent with some scientific hypotheses?

As it is also outlined in the introduction, four aspects may be relevant in the

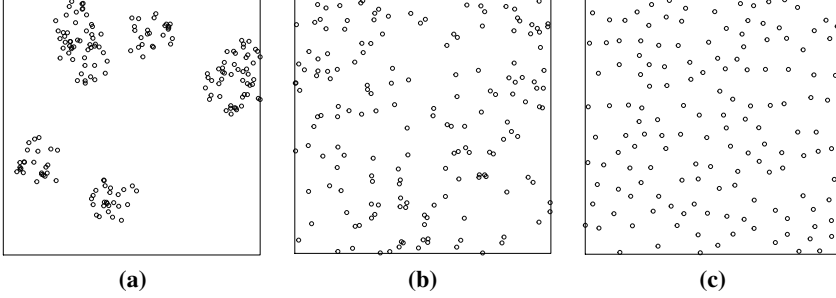


Figure 1.1: Simulated point patterns in $W = [0, 1] \times [0, 1]$ with an expected number of points equal to 150. (1.1a) Realisations from a Matérn's cluster process with 5 expected number of parents and radius parameter of the cluster equal to 0.1. (1.1b) Realisations from an homogeneous Poisson model. (1.1c) Realisations from an inhibitive process with inhibition distance equal to 0.05.

analysis of a point pattern and may explain the phenomena under several perspectives. The first one involves the identification of a **spatial trend** and so the description of the first-order characteristics of the pattern. In some contexts, it is also important the description of the spatial arrangement as a function of environmental information (such as a spatial variable $Z(\mathbf{u})$ defined at all locations $\mathbf{u} \in W$), in this second case, we aim to assess the **influence of spatial covariates**.

The third aspect concerns the study of **interaction between points** (and so the description of the second order characteristics) assessing the assumption of complete spatial randomness or characterise properly the interaction structure, considering a cluster or inhibitive interaction between points or a combination of both. As an example, Figure 1.1 shows some simulated point patterns in $W = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ with cluster (on the left), complete spatial randomness (in the middle) and inhibitive structures (on the right). Finally for the fourth aspect, we can have additional information (qualita-

12 Analysis and modelling of spatio and spatio-temporal point processes

tive or continuous attributes) for each point and we can do a **marked point analysis**. As a matter of fact, the research question concerns the identification of a correlation between the point distribution and the marks.

For instance, we describe the four aspects considering data from a plant community where the points are plant locations (Illian *et al.*, 2008). The description of the spatial displacement can involve spatial covariates, giving information about terrain, hydrology (depth to water table, annual recharge flow), soil (soil type, acidity) characteristics of the area. Furthermore, it is known that young trees in natural forests tend to show clustering as the seedlings sprout in proximity to the parent tree, while old trees in cultivated forests tend to be regular. Moreover, if we have additional information for each tree such as the marks plant size characteristics (diameter and height of each tree), we may expect that in areas of high point density the marks may tend to be smaller than in areas with low point density resulting from stronger competition for limited resources.

Analogous considerations can be done for the spatial analysis of earthquake data.

1.1.1 First-order and second-order properties

The first-order behavior is described by the intensity, and it is defined for a process $\mathcal{X}$ as

$$\rho(\mathbf{u}) = \lim_{|d\mathbf{u}| \rightarrow 0} \frac{\mathbb{E}(Y(d\mathbf{u}))}{|d\mathbf{u}|}$$

where $d\mathbf{u}$ is an infinitesimal region that contains the point $\mathbf{u} \in W$, $|d\mathbf{u}|$ is its area and $\mathbb{E}(Y(d\mathbf{u}))$ denotes the expected number of events in $d\mathbf{u}$.

A point process is called stationary, if the statistical properties of $\mathcal{X}$ and

$\mathcal{X} + \mathbf{v}$ are identical, for any choice of the translation vector $\mathbf{v}$. Stationarity implies, for example, that $\mathcal{X}$ must have homogeneous intensity $\rho(\mathbf{u}) = \rho$, that is the mean number of events per unit area. A point process is isotropic if its statistical properties are unchanged when it is rotated, through any given angle, around the origin. For example, the homogeneous Poisson process is a stationary and isotropic point process, because the observable statistical properties do not depend on the location or orientation of the study region in two-dimensional space.

Whereas intensity is a first moment property, interpoint interaction is measured by second-order moment quantities. Several functional summary statistics are used to study the second-order characteristics of a point pattern and so measuring dependence. Basically, they can be classified either in quantities that measure the spatial correlation (such as the product density, the Ripley's K-function and the pair correlation function) or functions that measure dependence in terms of spacings or shortest distances in the pattern (such as the nearest-neighbour function G and empty-space function F). We introduce the K and G-function because they are used in the descriptive analysis of real data in Chapter 2 and also their residual versions are fundamental for the model diagnostic of the Hybrid of Gibbs point process model as it is discussed in Section 1.1.3.2

Adelfio and Schoenberg (2009) introduce a diagnostic approach based on a weighted version of several second order statistics for general point processes.

When interest is in describing the correlations between points of a pattern, we have to consider pairs of events, and the corresponding characteristics are called second-order measures, such as the second-order intensity, that is defined as

14 Analysis and modelling of spatio and spatio-temporal point processes

$$\rho^{(2)}(\mathbf{u}_i, \mathbf{u}_j) = \lim_{|\mathbf{u}_i|, |\mathbf{u}_j| \rightarrow 0} \frac{\mathbb{E}[Y(d\mathbf{u}_i)Y(d\mathbf{u}_j)]}{|d\mathbf{u}_i||d\mathbf{u}_j|}$$

For a stationary and isotropic process, $\rho^{(2)}(\mathbf{u}_i, \mathbf{u}_j) = \rho^{(2)}(\|\mathbf{u}_i - \mathbf{u}_j\|)$.

Equivalently, the second-order properties of a point process can be described by the pair correlation function (also called g-function). It is defined as a normalisation of the second order product density by dividing with the corresponding probability for a Poisson process,

$$g(r) = \frac{\rho^{(2)}(\|\mathbf{u}_i - \mathbf{u}_j\|)}{\rho(\mathbf{u}_i)\rho(\mathbf{u}_j)} \quad \text{with } r > 0 \quad (1.1)$$

The K-function is connected to the pair correlation throughout the following relationship

$$K(r) = 2\pi \int_0^r u' g(u') du'. \quad (1.2)$$

For a stationary and isotropic process, one definition of the K-function is

$$K(r) = \rho^{-1} \mathbb{E}[Y_0(r)]$$

where $Y_0(r)$ is the number of further events within distance r of an arbitrary event. The value of $K(r)$ contains contributions for all interpoint distances less than or equal to r . When the process is stationary, the most commonly used estimator is due to Ripley (1976) :

$$\hat{K}(r) = \frac{|W|}{n(n-1)} \sum_{i=1}^n \sum_{j=1, j \neq i}^n I(d_{ij} \leq r) w(\mathbf{u}_i, \mathbf{u}_j) \quad (1.3)$$

where n is the number of points in the pattern, $|W|$ is the area of the window, $w(\mathbf{u}_i, \mathbf{u}_j)$ is the Ripley's edge correction weight and $d_{ij} = \|\mathbf{u}_i - \mathbf{u}_j\|$ is the pairwise distances between all distinct pairs of points $\mathbf{u}_i$ and $\mathbf{u}_j$ in

the pattern. The estimate of K is compared to the value of the function for the assumption of completely random (Poisson) point process, that is $K(r) = \pi r^2$. Deviations between the empirical and theoretical K curves may suggest spatial clustering (Figure(1.2a)) or spatial inhibition (Figure(1.2c)). Baddeley *et al.* (2000) compute a generalisation of the K -function for an inhomogeneous point process. Moreover for further details of the theoretical K - and g -functions and their estimators see Baddeley *et al.* (2015). Additional information on a point pattern can also be revealed by measuring spacings or shortest distances in the pattern, such as the nearest-neighbour distance from each data point to the nearest other data point. A good statistical analysis of spatial point patterns should therefore study both the correlation (throughout for example the Ripley K -function) and the shortest distances.

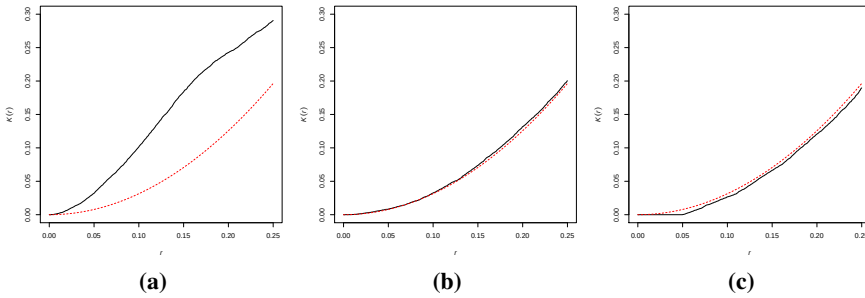


Figure 1.2: The black lines are the estimated homogeneous K -functions for the simulated point patterns in Figure 1.1. The red curves indicates the value of the K -function for a completely random point process, that is $K(r) = \pi r^2$.

The G -function is a summary statistic that is based on the spacing between points in a point pattern. It is a function of $d_i = d(\mathbf{u}_i, \mathbf{v} \setminus \mathbf{u}_i)$ that indicates the shortest distance from $\mathbf{u}_i$ to the pattern $\mathbf{v} \setminus \mathbf{u}_i$ consisting of all points of

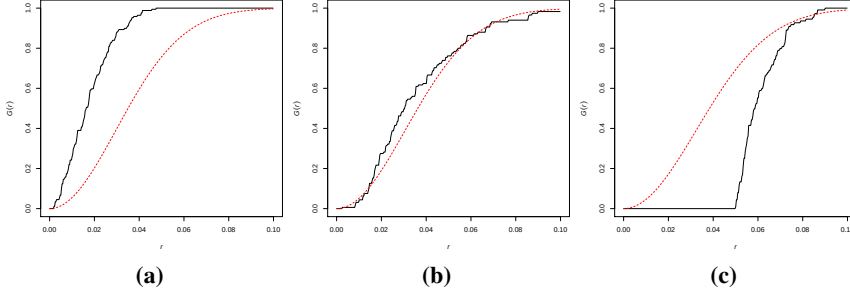


Figure 1.3: The black lines are the estimated G-functions for the simulated point patterns in Figure 1.1. The red curves indicate the value of the G-function for a completely random point process, that is $G(r) = 1 - \exp(-\rho\pi r^2)$.

$\mathbf{u}$ except $\mathbf{u}_i$. The G-function is the cumulative distribution function of the nearest-neighbour distance at a typical point of $\mathcal{X}$, and it is defined as

$$G(r) = \mathbb{P}\{d(\mathbf{u}, \mathcal{X} \setminus \mathbf{u}) \leq r\} / \mathcal{X} \text{ has a point in } \mathbf{u} \quad (1.4)$$

for any distance $r \geq 0$ and any location $\mathbf{u}$. The estimator is based on empirical cumulative distribution functions of the nearest-neighbour distances at all data points. $G(r)$ is a non-decreasing step function and $G(0) = 0$. For a homogeneous Poisson process, the nearest-neighbour distance distribution function is $G(r) = 1 - \exp(-\rho\pi r^2)$ where ρ is the intensity (expected number of points per unit area). Deviations between the empirical and theoretical G curves may suggest spatial clustering (Figure 1.3a) or spatial regularity (Figure 1.3c).

1.1.2 Model formulation

The statistical description of point patterns are often in combination with suitable point process models because analysing data using only summary

statistic could be unsatisfactory. The main purpose of statistical modelling is to make explicit assumptions and to assess them, as matter of fact it is the best way to investigate relationships while taking account of other relationships, and while accounting for variability. In this section, we briefly summarise the main classes of spatial point process models and which aspects of the first and second order characteristics of a process can describe. With the **Poisson model**, we can describe an observed inhomogeneity assuming a parametric, non-, or semi-parametric structure for the intensity (ρ) that completely describes the process. The key assumption is that the random points are independent of each other.

Other class of models usually allow the description of dependence between points in terms of clustering (positive dependence) or regularity (negative dependence).

Two flexible and convenient models for clustering structure are the **Cox point processes** and **Neyman-Scott cluster processes** that are modifications of the Poisson process incorporating additional random influences and source of variability.

Cox process is a Poisson process with a random intensity function that depends on unobservable external factors, such as the log-Gaussian Cox process and the shot-noise Cox process. In a log-Gaussian Cox process, it is assumed that a Gaussian Random Field (GRF) governs the random intensity and the estimation of the mean, the variance and the covariance of the GRF are of interest. The characteristics of log-Gaussian Cox process for the spatio-temporal context are detailed in Section 1.2.4.

Instead, in a cluster process, we observe offspring points that occur around unseen parents that represent the random unobservable variability. For example, in the Matérn's cluster process, the parents are generated consid-

ering a homogeneous Poisson point process, then each parent point is replaced by a random cluster of offspring points and their positions being placed and uniformly inside a disc of radius r centered on the parent point. In Figure 1.1a, we have realisations from a Matér's cluster process with 5 expected number of parents, radius parameter of the cluster equal to $r = 0.1$ and total expected number of point equal to 150.

For inhibitive structure, we can consider the **Simple Sequential Inhibition point process**. Starting with an empty window, a new point is generated uniformly in the window and independently of preceding points. If the new point lies closer than r units from an existing point, then it is rejected. In Figure 1.1c, there is an observed point pattern from the Simple Sequential inhibition point process.

Finally, **Gibbs point process** is a class of flexible and natural models for point patterns that postulates interactions between the points of the process defining a density for a point process with respect to a Poisson process of unit intensity. These models allow to characterise a form of local or Markovian dependence amongst events, they are flexible since they can produce a wide range of spatial patterns, ranging from strongly regular to strongly clustered patterns. Moreover using hybrid of Gibbs point process, it is possible to combine repulsion and attraction at different scales.

Another important characteristics of Gibbs point process is that it is possible to distinguish between inhomogeneity and interaction parameters considering several formulations.

This last class of point process model are described with more details in the following section.

1.1.3 Gibbs and Hybrids of Gibbs point process models

The class of Gibbs processes $\mathcal{X}$ also called Markov point processes is determined through a probability density function $f : \mathcal{X} \rightarrow [0, \infty)$ (with respect to a Poisson process of unit intensity), where $\mathcal{X} = \{\nu \subset W : n(\nu) < \infty\}$ is a set of point configurations contained in W . Generally any unnormalised density $f(\nu)$ that is integrable and not identical to zero can be used to define a finite point process but $f(\cdot)$ should also satisfy other properties such as the local stability. For further details about desirable properties for an unnormalized density see Møller and Waagepetersen (2003) and Baddeley *et al.* (2015). A finite Gibbs point process is a finite simple point process defined with a density that satisfies the hereditary condition $f(\nu) > 0 \Rightarrow f(\nu') > 0$ for all $\nu' \subset \nu$. Its density function can be written in the form

$$f(\nu) = \exp\left\{V_0 + \sum_{\mathbf{u} \in \nu} V_1(\mathbf{u}) + \sum_{\mathbf{u}, \mathbf{u}' \in \nu} V_2(\mathbf{u}, \mathbf{u}') + \dots\right\} \quad (1.5)$$

where V_0 is a constant and $V_k : \mathcal{X}^k \rightarrow \mathbb{R} \cup \{-\infty\}$ is a symmetric function called the potential of order k , with $k = \{1, 2, \dots\}$ (Ripley and Kelly, 1977; Møller and Waagepetersen, 2003). An unnormalized density f has interaction order m if $V_k = 0$ for all $k > m$. It has an interaction range equal to r , if all $V_k(\nu) = 0$ whenever ν contains two points $\mathbf{u}_1, \mathbf{u}_2 \in \nu$ such that $\|\mathbf{u}_1 - \mathbf{u}_2\| > r$. In this paper, we refer to specific Gibbs models and combine them under a hybrid approach. In particular, we focus on Strauss and Geyer models that account for different types of point interactions.

Apart from the Poisson point process, that has interaction order equal to 1 and null interaction range, the most common Gibbs processes are the pairwise interaction processes that have $m = 2$ and a density function of the

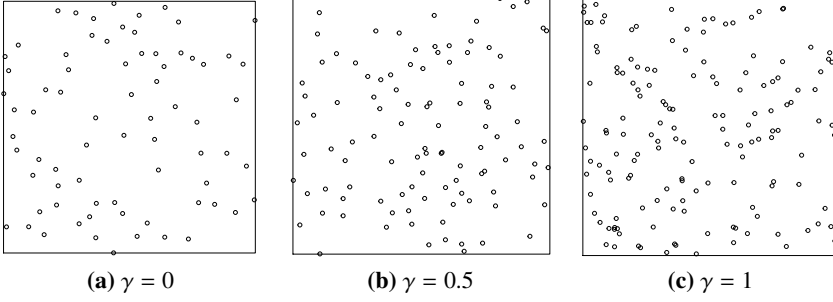


Figure 1.4: Simulated realisations of Strauss processes in the unit square with different values of γ ranging from 0 to 1, $\beta = 150$ and interaction radius $r = 0.05$. In Figure (1.4a) the generating process is equivalent to hard core process; instead, in Figure (1.4c), the generating process is equivalent to a Poisson one.

form

$$f(\mathbf{v}) = \alpha \prod_{i=1}^n b(\mathbf{u}_i) \prod_{i < j} c(\mathbf{u}_i, \mathbf{u}_j) \quad (1.6)$$

where $b : W \rightarrow [0, \infty)$, $c : W \times W \rightarrow [0, \infty)$ and $\alpha > 0$ is the normalizing constant.

A simple case of (1.6) is the homogeneous **Strauss Process** (Strauss, 1975), with $b(\mathbf{u}) = \beta > 0$ and

$$c(\mathbf{u}_i, \mathbf{u}_j) = \begin{cases} \gamma & \text{if } \|\mathbf{u}_i - \mathbf{u}_j\| < r \\ 1 & \text{otherwise} \end{cases}$$

where r is a fixed interaction distance and $0 \leq \gamma \leq 1$ is the interaction parameter. It follows that the probability density is

$$f(\mathbf{v}) = \beta^{n(\mathbf{v})} \gamma^{t(\mathbf{v}, r)}, \quad (1.7)$$

where $t(\mathbf{v}, r)$ is the number of unordered pairs of distinct points in the pat-

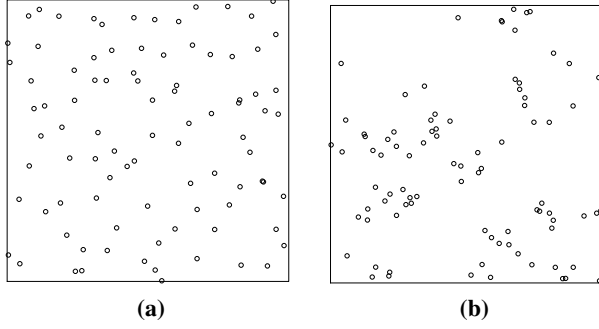


Figure 1.5: Simulated realisations of Geyer saturation process in the unit square. In Figure (1.5a), there is an inhibitive structure and the parameters are equal to $\{\beta, \gamma, r, s\} = \{300, 0.5, 0.07, 3\}$. In Figure (1.5b), there is a cluster structure and the parameters are equal to $\{\beta, \gamma, r, s\} = \{15, 2, 0.07, 3\}$.

tern $\mathbf{v}$ that lie closer than r units apart. The Strauss process is intermediate between the hard-core process and the complete spatial randomness and its behaviour depends on the value of γ (Baddeley *et al.*, 2015). Figure 1.4 shows simulated realizations of the Strauss process with $\beta = 150$ and different values of $\gamma = \{0, 0.5, 1\}$. As a special case, the Strauss process with $\gamma = 1$ is a Poisson process with completely random points, see Figure 1.4c. Instead, if γ is chosen to be equal to 0, the Strauss process is the hard core process, see Figure 1.4a.

Pairwise interaction point processes are commonly used to describe spatial events with a regular pattern (Illian *et al.*, 2008).

On the other hand, a proper Gibbs process to describe attractive interaction is the homogeneous **Geyer saturation process** (Geyer, 1999). It has an unnormalized density equal to

$$f(\mathbf{v}) = \beta^{n(\mathbf{v})} \prod_{i=1}^{n(\mathbf{v})} \gamma^{\min(s, t(\mathbf{u}_i, \mathbf{v} \setminus \mathbf{u}_i))} \quad (1.8)$$

22 Analysis and modelling of spatio and spatio-temporal point processes

where $s > 0$ is the saturation parameter and $t(\cdot)$ is the number of unordered pairs of distinct points in the pattern $\mathbf{v}$ that lie closer than r units apart.

The process is inhibitive when $\gamma \leq 1$, see Figure 1.5a, and clustered when $\gamma > 1$, see Figure 1.5b.

However Gibbs processes have some drawbacks when points have a strong clustering and show spatial dependence at multiple scales (Illian *et al.*, 2008; Møller and Waagepetersen, 2003; Baddeley *et al.*, 2013).

In Baddeley *et al.* (2013), the authors propose hybrid models as a general way to generate multi-scale processes combining Gibbs processes.

Given m unnormalized densities $f_1, f_2, \dots, f_m$, the hybrid density is defined as $f(\mathbf{v}) = f_1(\mathbf{v}) \times \dots \times f_m(\mathbf{v})$. The new unnormalized density and the component densities f_i have to respect some assumptions, for more details see Baddeley *et al.* (2013). For example the density of the stationary hybrid process obtained considering m Geyer components (1.8) (with interaction ranges $r_1, \dots, r_m$ and saturation parameters $s_1, \dots, s_m$) is

$$f(\mathbf{v}) = \beta^{n(\mathbf{v})} \prod_{i=1}^{n(\mathbf{v})} \prod_{j=1}^m \gamma_j^{\min(s_j, t(\mathbf{u}_i, \mathbf{v} \setminus \mathbf{u}_i; r_j))} \quad (1.9)$$

where $t(\mathbf{u}_i, \mathbf{v} \setminus \mathbf{u}_i; r_j) = \sum_i \{\mathbf{1} \|\mathbf{u} - \mathbf{u}_i\| \leq r_j\}$. This density indicates that the spatial interaction between points changes with the distances r_j and the parameters that capture this information are the interaction parameters γ_j .

If an inhomogeneous version of (1.9) is considered, the parameter β is replaced by a function $\beta(\mathbf{u}_i)$ that expresses a spatial trend and it can be a function of the coordinates of the points and covariate information defined in all the study area. Generally, we can specify that the density f is a function of a vector of regular parameters θ and a vector of irregular parameters η . This distinction is important since these parameters are estimated into two differ-

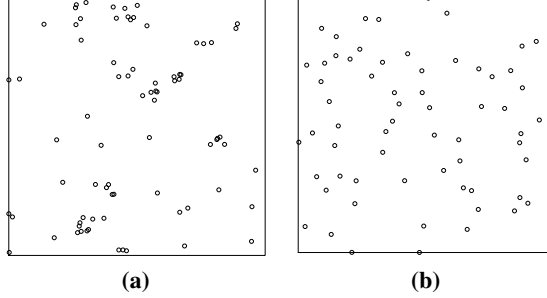


Figure 1.6: Simulated realisations from a hybrid Geyer process in the unit square defined by two components, and irregular parameters equal to $\eta = \{(r_1, s_1), (r_2, s_2)\} = \{(0.05, 3), (0.1, 1)\}$. In Figure (1.6a), the regular parameters are $\exp(\theta) = \{\beta, \gamma_1, \gamma_2\} = \{40, 2, 0.5\}$, and in Figure (1.6b), they are set to $\exp(\theta) = \{\beta, \gamma_1, \gamma_2\} = \{40, 0.5, 2\}$.

ent steps, as it is shown in the next section. In the case of (1.9), the previous two parametric vectors are $\theta = \{\log(\beta), \log(\gamma_1), \dots, \log(\gamma_m)\} = \{\theta_1, \theta_2\}$ for regular parameters and $\eta = \{(r_1, s_1), \dots, (r_m, s_m)\}$ for the irregular ones. Moreover, the regular vector can be subdivided into parameters for the description of the spatial trend (θ_1) and parameters for the interaction effects (θ_2). Figure 1.6 shows two simulated point patterns that combine repulsion and attraction at different spatial scales considering hybrid models of Geyer processes.

1.1.3.1 Model estimation

Generally the density function f of the point process is governed by a parameter vector ϕ , but, for a wide class of spatial process models, the likelihood is difficult to evaluate and to maximize. The pseudolikelihood of a point process is then used as an alternative (Besag, 1975, 1977; Jensen

24 Analysis and modelling of spatio and spatio-temporal point processes

et al., 1991), and it is given by

$$PL(\phi; \mathbf{v}) = \left[\prod_{i=1}^{n(\mathbf{v})} \rho_{\phi}(\mathbf{u}_i; \mathbf{v}) \right] \exp \left[- \int_W \rho_{\phi}(\mathbf{u}; \mathbf{v}) d\mathbf{u} \right]$$

where

$$\rho_{\phi}(\mathbf{u}; \mathbf{v}) = \frac{f_{\phi}(\mathbf{v} \cup \{\mathbf{u}\})}{f_{\phi}(\mathbf{v} \setminus \{\mathbf{u}\})}$$

is the so called Papangelou conditional intensity at location $\mathbf{u} \in W$ given $\mathbf{v}$. In hybrid models $\phi = \{\theta, \eta\}$ and the conditional intensity is the product of the conditional intensities of the components (Baddeley *et al.*, 2013)

$$\rho_{\phi}(\mathbf{u}; \mathbf{v}) = \exp\{B(\mathbf{u}) + \theta_1^T V_1(\mathbf{u}, \eta) + \theta_2^T G(\mathbf{u}, \mathbf{v}, \eta)\} \quad (1.10)$$

where $B(\mathbf{u})$ is an offset term, $\theta_1^T V_1(\mathbf{u}, \eta)$ is the first-order potential and $\theta_2^T G(\mathbf{u}, \mathbf{v}, \eta)$ is the sum of terms with $k \geq 2$ in (1.5). The term $G(\cdot)$ accounts for the interaction effects, and in Chapter 2 following a hybrid approach, it will be a combination of Strauss and Geyer processes. The technique that is considered in this work for computing approximate maximum pseudolikelihood estimates is described in Baddeley and Turner (2000). The irregular parameters are estimated through the profile pseudolikelihood, maximizing $p(\eta, \mathbf{v}) = \max_{\theta} \log PL(\theta, \eta)$ over η . Their computation is a crucial issue both for a numerical aspect and because they may affect the interpoint interaction or the spatial trend parameters estimation. In fact, after computing $\hat{\eta}$, the vector θ is obtained by maximizing $\log PL(\theta, \hat{\eta})$.

Different specifications can be considered for the trend part ($B(\mathbf{u})$ and $\theta_1^T V_1(\mathbf{u})$) of (1.10). In Chapter 2, we apply these models on earthquake data, and we consider a parametric and a non-parametric formulation for the spatial term depending on the coordinates. Moreover, we explain the trend part as a

parametric function of a set of spatial covariates that are defined in all the study region.

For the non-parametric specification of the spatial trend, we use a kernel estimator. Generally, given n observed events $\mathbf{u}_1, \mathbf{u}_2 \dots, \mathbf{u}_n$ in $\mathbb{R}^d$, the kernel estimator of the unknown density $f(\cdot)$ (Silverman, 1986) is

$$\hat{f}_{\Sigma}(\mathbf{u}) = \frac{\sum_{i=1}^n w_i K(\mathbf{u} - \mathbf{u}_i, \Sigma)}{\sum_{i=1}^n w_i}$$

where $K(\cdot, \cdot)$ is a multivariate kernel function centred at the observed points, Σ is a matrix of smoothing constants with $\text{diag}(\Sigma) = \{h_x, h_y\}$ when $d = 2$ and w_i are weights. Usually the normal multivariate density is chosen for $K(\cdot, \cdot)$. In our study, we consider some of the options proposed in the `spatstat` package Baddeley and Turner (2005) of R (R Development Core Team, 2005) to determine the bandwidth for a Gaussian kernel: cross-validated bandwidth selection, likelihood cross-validation and Scott's rule. Moreover, we use the FLP (Forward Likelihood for Prediction) method proposed in Chiodi and Adelfio (2011), and developed in Adelfio and Chiodi (2015a), for the estimation of the bandwidth vector $\mathbf{h} = (h_x, h_y)$, assuming a bivariate normal kernel for the spatial distribution of the background seismicity.

1.1.3.2 Diagnostics

After fitting a model, the next step is to assess the goodness-of-fit. In this work, the diagnostic tools on spatial residuals described in Baddeley *et al.* (2005) have been used to check the general fitting and if the specification of the spatial trend, covariate effects and interpoint interactions are appropriate.

Additional diagnostic tools useful for hybrid models are proposed in Baddeley *et al.* (2011) and they are based on the second-order summary statistics, K (1.2) and G (1.4) functions. Usually the Ripley's K-function (Ripley, 1976, 1988), that is determined by the second-order moment properties of the process, is used to detect the presence of clustering or spatial regularity comparing its value to the theoretical one under the Poisson assumption. Also the G-function, that is the nearest neighbor distance of a point pattern, is used to study clustering and inhibition between points.

In Baddeley *et al.* (2011) , they propose new tools for model validation in the analysis of spatial point pattern data based on residual summary statistics namely, the residual K- and G-functions. The residual version of one of the classical functional summary statistics is obtained by subtracting a compensator from the functional summary statistic. The compensator depends on the fitted model, and may also depend on the observed data. For example, suppose that the fitted model is the homogeneous Poisson process, then (ignoring some details) the compensator of the empirical K-function $\hat{K}(r)$ is its expectation $K_0(r) = \pi r^2$ under the model. So, they show that the score test for models comparison is equivalent to the residual K- and G-functions, that are the difference of a summary statistic $S(\nu)$ and its predicted value $\mathbb{E}(S(X))$ under a null model.

For example, the residual K-function is computed as the difference between the non-parametric estimated K-function ($\hat{K}_\nu(r)$) such as the estimator in Equation 1.3 and its compensator, $C\hat{K}_\nu(r)$). Following Baddeley *et al.* (2011), under CSR (complete spatial randomness) condition and the intensity being estimated by $n/|W|$, the pseudo-residual is approximately given by

$$\begin{aligned}
2(\hat{K}_v(r) - C\hat{K}_v(r)) = & \quad (1.11) \\
2\left(\frac{|W|}{n(v)(n(v)-1)} \sum_{i=1}^n \sum_{i=1, i \neq j}^n I(d_{ij} \leq r) w(\mathbf{u}_i, \mathbf{u}_j) \right. \\
& \left. - \frac{1}{\hat{\rho}^2(n(v)+1)|W|} \int_W t^w(\mathbf{u}, \mathbf{v}, r) \lambda_{\hat{\phi}}(\mathbf{u}, \mathbf{v}) d\mathbf{u} \right)
\end{aligned}$$

where $\hat{\rho}^2(n(v)) = n(v)(n(v)-1)/|W|^2$, $n(v)$ number of points in W , $w(\mathbf{u}_i, \mathbf{u}_j)$ is the edge correction, $t^w(\mathbf{u}, \mathbf{v}, r) = \sum_j w(\mathbf{u}, \mathbf{u}_j) I(\|\mathbf{u} - \mathbf{u}_j\| \leq r)$ is a weighted count of the points of $\mathbf{v}$ that are r -close to the location $\mathbf{u}$. In Baddeley *et al.* (2011), (1.11) is shown to become under the CSR condition

$$\frac{n(v)^2}{2|W|} (\hat{K}_v(r) - \pi r^2)$$

The compensator K-function is based on the fitted model and, if it is correct, the residual summary statistics should be on average zero. In terms of hypothesis testing using the residual K-function, the comparison is between the fitted model under the null hypothesis, and a hybrid of the current fitted model with a Strauss process under the alternative hypothesis. This connections derive from the fact that the estimate of the K-function is essentially the ‘sufficient statistics’ for the Strauss process. In this case, a peak of the residual K-function at a particular distance r suggests that the current model may be improved including a Strauss process component, with interaction range r .

The residual G-function is the difference between the non-parametric estimate of the G-function and its compensator version based on the fitted model. It is equivalent to a score test where the fitted model is assumed

under the null hypothesis, and a hybrid of the current model with a Geyer process with saturation parameter $s = 1$ is assumed under the alternative hypothesis. The nearest-neighbour distance function G is essentially the "sufficient statistics" for the Geyer saturation process.

Usually the residual K- and G-functions are represented graphically with upper and lower limits. The null distribution of the standardized residual is not known and the critical bands are approximate two-standard deviation limits under the fitted model, however the bands are unreliable for small values of r . It is possible to use simulations to obtain exact significance bands, but it is computationally demanding (Baddeley *et al.*, 2013).

Summing up, in the analysis in Chapter 2, we compare and assess the goodness-of fit of the estimated models in terms of AIC, spatial raw residuals and number of simulated points under the estimated model. Furthermore, the diagnostic plots based on the residual K- and G-functions are used to decide which component has to be added at each step to the hybrid model. Indeed, these graphs show for which spatial distances the current model has a lack of fit in describing the interaction between points.

In the `spatstat` package (Baddeley and Turner, 2005) of R (R Development Core Team, 2005), there are most of the functions that have been used for fitting, prediction, simulation and validation of hybrid models.

1.2 Spatio-temporal point processes

In this section, we report and revise concepts in spatio-temporal point pattern analysis. More exactly, we specify the (global) second order summary statistics to describe interaction between points because they are the bases for the study and the application of Local indicator of Spatio-Temporal

association (Chapter 4). Moreover, we describe the spatio-temporal Log-Gaussian Cox process because we use this model formulation in Chapter 3. Finally, we also outline the structure of the Epidemic Type Aftershock Sequence model since some results in applying this model are used for bandwidth selection to describe the spatial density of the events in the analysis in Chapter 2 and to classify earthquakes in aftershock and mainshock events (Chapter 4). For the two cited spatio-temporal models, we briefly describe the methods of estimation and model validation.

Modelling real problems through spatio-temporal point processes becomes essential in many scientific and engineering fields such as environmental sciences, climate prediction and meteorology, image analysis, seismology, astronomy, epidemiology and criminology.

Moving from the spatial to the spatio-temporal (or spatial-temporal) context, we deal with a process that evolves in a two-plus-one-dimensional space-time where the time dimension is different from either of the two spatial dimensions (Diggle, 2013). As explained in Illian *et al.* (2008), a temporally evolving process can be classified according to the type of time dependence and they identify three main generating processes that can be overlapped in an observed phenomena.

In a given pattern, the events can occur at a random instant of time and spatial location and their duration times are negligible or not relevant for the analysis. In this classification, we deal (strictly speaking) with spatio-temporal point process data and for instance earthquake events can be considered in this first category. When for example, the goal of the study concerns the analysis of paths of movements of physical particles or storm centres, we deal with events that move in the spatial domain or appear and remain for a random interval of time. Instead, the third classification in-

volves birth-death processes since the points appear randomly and disappear after a random length of time, and for example the modelling of plant communities or natural fires can be studied within this framework.

We analyse data that are in the first classification, giving the definition and the main properties of spatio-temporal point processes. We consider a spatio-temporal point process with no multiple points as a random countable subset $\mathcal{X}$ of $\mathbb{R}^2 \times \mathbb{R}$, where for a point $(\mathbf{u}, t) \in \mathcal{X}$, $\mathbf{u} \in \mathbb{R}^2$ is the location and $t \in \mathbb{R}$ is the time of occurrence. In practice, an observed spatio-temporal pattern is a finite set $\{(\mathbf{u}_i, t_i)\}_{i=1}^n$ of distinct points within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, where usually W is a polygon with area $|W| > 0$ and T a single closed interval with length $|T| > 0$. Considering a bounded spatio-temporal region $A \subset W \times T$, $Y(A)$ denotes the number of the events of the process falling in A .

A spatio-temporal point process $\mathcal{X}$ is *stationary* when the distribution of $\mathcal{X}$ is invariant under translations in $\mathbb{R}^2 \times \mathbb{R}$. Instead, we say that $\mathcal{X}$ is *isotropic* when for any rotation around the origin, the rotated point process has the same distribution. These two properties determine specific characteristics of the first- and second-order properties of the point pattern and they will be recalled with further details later.

1.2.1 First-order and second-order properties

Informally, the first-order properties of a point process are described by the intensity function and it is the analogue of the mean for a real-valued stochastic process. Instead, the analogue of the covariance function of a real-valued process is described by the second-order properties in a point process. In the analysis of spatio-temporal point process data is common to investigate the interactions among the points of the process (analysing

the second-order properties) after adjusting for spatial and/or temporal inhomogeneity, however it is difficult to untangle these two aspect because only a realisation of the process is observed.

The intensity of a process is defined as

$$\rho(\mathbf{u}, t) = \lim_{|d\mathbf{u}| \rightarrow 0; |dt| \rightarrow 0} \frac{\mathbb{E}[Y(d\mathbf{u}, dt)]}{|d\mathbf{u}||dt|}$$

where $d\mathbf{u} \times dt$ is a spatio-temporal region around the point $(\mathbf{u}, t)$, $|d\mathbf{u}|$ is the area of the spatial region, $|dt|$ is the length of the time interval and, $\mathbb{E}(Y(d\mathbf{u}, dt))$ denotes the expected number of events in the infinitesimal spatio-temporal region. The process is called homogeneous or stationary when the intensity is constant, $\rho(\mathbf{u}, t) = \rho$ for all $(\mathbf{u}, t) \in W \times T$.

When the interest is in describing the spatio-temporal variability and correlations between points of a pattern, we have to consider second-order measures, such as the product density $\rho^{(2)}(\cdot, \cdot)$. This quantity provides an interpretable measure of the spatio-temporal dependence structure and it is defined as

$$\rho^{(2)}((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) = \lim_{|D_i||D_j| \rightarrow 0} \frac{\mathbb{E}[Y(D_i)Y(D_j)]}{|D_i||D_j|} \quad (1.12)$$

where $D_i = d\mathbf{u}_i \times dt_i$ and $D_j = d\mathbf{u}_j \times dt_j$ are small cylinders around two distinct points $(\mathbf{u}_i, t_i)$ and $(\mathbf{u}_j, t_j)$. Equivalently, the second-order properties of a point process can be described by the covariance density and the pair correlation function that are respectively equal to

$$\varphi((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) = \rho^{(2)}((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) - \rho(\mathbf{u}_i, t_i)\rho(\mathbf{u}_j, t_j) \quad (1.13)$$

$$g((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) = \frac{\rho^{(2)}(\mathbf{u}_i, t_i; \mathbf{u}_j, t_j)}{\rho(\mathbf{u}_i, t_i)\rho(\mathbf{u}_j, t_j)} \quad (1.14)$$

32 Analysis and modelling of spatio and spatio-temporal point processes

Informally, the pair correlation function can be interpreted as the standardised probability density that there is a point of $\mathcal{X}$ in each of two small disjoint spatio-temporal cylinders with centres in the points $(\mathbf{u}_i, t_i)$ and $(\mathbf{u}_j, t_j)$.

Stationarity

A process for which $\rho(\mathbf{u}, t) = \rho$ for all $(\mathbf{u}, t) \in \mathcal{X}$ is called homogeneous or first-order stationary in both space and time. A point process $\mathcal{X}$ is second-order intensity-reweighted stationary (SOIRS), if $\rho^{(2)}(\mathbf{u}_i, t_i; \mathbf{u}_j, t_j) = \rho^{(2)}(\mathbf{u}_i - \mathbf{u}_j; t_i - t_j)$. Further, if the process is SOIRS and isotropic, then the pair correlation function depends only on the spatial and temporal distances between points $g(\mathbf{u}_i - \mathbf{u}_j, t_i - t_j) = g_0(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|)$ for some non-negative function $g_0(\cdot, \cdot)$.

Separability

A spatio-temporal point process is first-order separable if its intensity can be factorised as $\rho(\mathbf{u}, t) = \rho_1(\mathbf{u})\rho_2(t)$ for all $(\mathbf{u}, t) \in W \times T$, where ρ_1 and ρ_2 are non-negative functions. Often, in the analysis of real data, the pragmatic working assumption that first-order effects are separable is done, because in the absence of independent replication, we cannot in general make an empirical distinction between first-order and second-order effects without additional assumptions. Under this assumption, any non-separable effects are interpreted as second-order, rather than first-order properties (Gabriel and Diggle, 2009).

Analogous to the first-order case, a spatio-temporal point process is second-order separable if the covariance density can be factorised into

$$\varphi((\mathbf{u}_i, t_i), (\mathbf{u}_j, t_j)) = \bar{\varphi}_1(\mathbf{u}_i, \mathbf{u}_j)\bar{\varphi}_2(t_i, t_j), \quad (\mathbf{u}_i, t_i), (\mathbf{u}_j, t_j) \in \mathbb{R}^2 \times \mathbb{R}$$

where $\bar{\varphi}_1^{(2)}(\cdot, \cdot), \bar{\varphi}_2^{(2)}(\cdot, \cdot)$ are non-negative functions.

Spatio-temporal K-function

In practice, the space-time inhomogeneous pair correlation function and K-function are used to analyse the spatio-temporal interaction of a point process as measure of spatio-temporal clustering/regularity. In Gabriel and Diggle (2009), the authors extended Ripley's K -function to the spatio-temporal inhomogeneous context for a SOIRS and isotropic point process

$$K(r, h) = 2\pi \int_0^r \int_0^h g(r', h') r' \, d(r', h'), \quad r > 0, \quad h > 0, \quad (1.15)$$

where $g(r, h)$ is the spatio-temporal pair correlation function in Equation (1.14) with $r = \|\mathbf{u} - \mathbf{u}'\|$ and $h = |t - t'|$.

Considerations under the homogeneous Poisson assumption

The homogeneous Poisson process that is constituted by events that are distributed randomly and independently, is the reference process in almost all the type of analysis of spatio-temporal point patterns. Under this assumption, we have that $g(r, h) = 1$ and $K(r, h) = 2\pi r^2 h$.

1.2.2 Estimators of second-order characteristics

We show some estimators proposed in the literature to describe (global) second-order characteristics of a point process. These quantities are useful because the local test introduced in Chapter 4 is based on individual contributions of each point to the (global) product density. Furthermore, the spatio-temporal K - and g -functions are used to study the second-order

34 Analysis and modelling of spatio and spatio-temporal point processes

characteristics of the observed dataset and in the estimation procedure of the parameters of the log-Gaussian Cox process model in Chapter 3.

Under the stationary case, and ignoring edge-effects, a global naive non-parametric kernel estimator for $\rho^{(2)}(r, h)$ in (1.12) (Rodríguez-Cortés, 2014; Rodríguez-Cortés *et al.*, 2014) is given by

$$\widehat{\rho^{(2)}}_{\epsilon, \delta}(r, h) = \frac{1}{4\pi r |B|} \sum_{i=1}^n \sum_{j \neq i}^n \kappa_{\epsilon, \delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h), \quad (1.16)$$

where the sum is over all pairs $(\mathbf{u}_i, t_i) \neq (\mathbf{u}_j, t_j)$ of the data points, $B = W \times T$, $r > \epsilon > 0$ and $h > \delta > 0$. The kernel function κ has a multiplicative form $\kappa_{\epsilon, \delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h) = \kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r) \kappa_{2\delta}(|t_i - t_j| - h)$, where $\kappa_{1\epsilon}$ and $\kappa_{2\delta}$ are kernel functions with bandwidths ϵ and δ , respectively. For an approximately unbiased edge-corrected estimator for the spatio-temporal product density see Rodríguez-Cortés *et al.* (2014). In Section 4.2, we discuss about local indicator of spatio temporal association that strictly speaking are based on the global spatio-temporal product density, and using the individual contributions, we define a statistical test for local structure detection.

An unbiased estimator of the inhomogeneous K-function (1.15), proposed in Gabriel and Diggle (2009) under the assumption of a SOIRS and isotropy process, takes the following form

$$\widehat{K}(r, h) = \frac{1}{|B|} \sum_{i=1}^n \sum_{j \neq i}^n \frac{\mathbf{1}[\|\mathbf{u}_i - \mathbf{u}_j\| \leq r, |t_i - t_j| \leq h]}{\hat{\rho}(\mathbf{u}_i, t_i) \hat{\rho}(\mathbf{u}_j, t_j) w(\mathbf{u}_i, \mathbf{u}_j) w(t_i, t_j)} \quad (1.17)$$

where $B = |W \times T|$, $w(\mathbf{u}_i, \mathbf{u}_j)$ and $w(t_i, t_j)$ denote the Ripley's spatial and temporal edge correction factors, respectively and $\hat{\rho}(\cdot, \cdot)$ is an estimate of

the first-order intensity. The spatial edge-correction, $w(\mathbf{u}_i, \mathbf{u}_j)$ is the proportion of the circumference of a circle centred at $\mathbf{u}_i$ with radius $\|\mathbf{u}_i - \mathbf{u}_j\|$ lying in W . On the other hand, the temporal edge correction factor is equal to 1, if both ends of the interval of length $2|t_i - t_j|$ centred at t_i are within T , and 0.5 otherwise. Finally, an estimator of the space-time pair correlation function defined in Equation (1.14) has been proposed in Gabriel (2014)

$$\widehat{g}_{\epsilon, \delta}(r, h) = \frac{1}{4\pi r|B|} \sum_{i=1}^n \sum_{j \neq i} \frac{\kappa_{\epsilon, \delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h)}{\hat{\rho}(\mathbf{u}_i, t_i) \hat{\rho}(\mathbf{u}_j, t_j) w(\mathbf{u}_i, \mathbf{u}_j) w(t_i, t_j)} \quad (1.18)$$

where $w(\mathbf{u}_i, \mathbf{u}_j)$ and $w(t_i, t_j)$ are the spatial and temporal edge correction factors defined in Equation (1.17), and $\kappa_{\epsilon, \delta}$ is a multiplicative kernel function as defined in Equation (1.16).

The R package `stpp` (Gabriel *et al.*, 2013) implements the main code for the computation of the three estimators in (1.16), (1.17) and (1.18). Usually, the K-function and pair correlation function are used for model checking and for the estimation of the parameters to describe the second-order characteristics of a process (such as in the Log-Gaussian Cox process model), while the product density is used for explanatory analysis. The form of these functions helps to understand the type of interaction in the point pattern and to find suitable point process models. However, the K- and g-functions require a careful estimation of the first-order characteristics which are unknown in practice and an estimate may be biased. Instead, the product density does not show this problem, even though it is based on the same information, it does not need an estimate of the intensity function.

1.2.3 Model formulation

In González *et al.* (2016), the authors do a self-contained review describing statistical models in the spatio-temporal context such as the Spatio-temporal Poisson, Neyman-Scott, inhibition, Strauss, Cox, Self-exciting processes.

In the next section, we describe the main characteristics of the log-Gaussian Cox process and the Epidemic Type Aftershock-Sequences (ETAS) model. As in all statistical model setting, it is possible to classify a model into a mechanistic or a deterministic model, even though practically it is not a sharp classification and it depends on the research goal (Diggle, 2013).

In the first class, there are models that encapsulate the underlying scientific processes that generated the data, instead the other class of models aim only to describe the pattern of a data-set. According to this classification, the ETAS model should be considered as a mechanistic model since incorporates seismologic information in the conditional intensity more exactly the Omori's law. On the other hand, the Log-Gaussian Cox process should be considered much more like a deterministic model since the interactions structure between points is expressed throughout a Gaussian process. It describes the pattern in a data-set without pointing to any particular underlying mechanism and it is useful in applications where the scientific focus is on prediction rather than on testing specific scientific (mechanistic) hypotheses.

However, it is important to point out that this distinction is a false dichotomy, the distinction between empirical and mechanistic models is not sharp and according on the research goal a model can be more suitable than an other.

1.2.4 Spatio-temporal Log-Gaussian Cox process

After outlining in a general way a Cox process, we report the main properties of log-Gaussian Cox processes (LGCP). Cox processes are an important class of flexible models used for point patterns that show clustered structure caused by an environmental stochastic heterogeneity. They are a generalisation of inhomogeneous Poisson processes where the intensity is a realisation of a random field. For this reason Cox processes are also called doubly stochastic Poisson processes (Cox, 1955). Suppose that $\Lambda = \{\Lambda(\mathbf{u}, t) : (\mathbf{u}, t) \in W \times T \subseteq \mathbb{R}^2 \times \mathbb{R}\}$ is a non negative and locally integrable stochastic process. The point process $\mathcal{X}$ is said to be a Cox process driven by Λ , if the conditional distribution of the point process $\mathcal{X}$ given a realization $\Lambda(\mathbf{u}, t) = \lambda(\mathbf{u}, t)$ is a Poisson process on W with intensity function $\lambda(\mathbf{u}, t)$. The most common Cox processes are the log Gaussian Cox processes (Møller *et al.*, 1998) and shot-noise Cox processes (Møller, 2003). Some examples of shot-noise Cox processes are the Neyman-Scott and Poisson-cluster processes.

Considering a generic point in space and time, the log-Gaussian Cox process has the following intensity $\Lambda(\mathbf{u}_i, t_i) = \exp\{S(\mathbf{u}_i, t_i)\}$ where S is a Gaussian Random Field (GRF) with mean $\mathbb{E}(S(\mathbf{u}_i, t_i)) = \mu$ and variance covariance matrix $\mathbb{C}(S(\mathbf{u}_i, t_i), S(\mathbf{u}_j, t_j)) = \mathbb{C}(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|) = \sigma^2 \gamma(r, h)$ under the stationary case where γ is the correlation function of the Random Field. In Diggle *et al.* (2013), they consider a re-parametrisation of the previous model such that

$$\Lambda(\mathbf{u}_i, t_i) = \lambda(\mathbf{u}_i, t_i) \exp\{S(\mathbf{u}_i, t_i)\} \quad (1.19)$$

In an epidemiology context, the components $\lambda(\mathbf{u}_i, t_i)$ determines the endemic spatial and temporal component of the population at risk; while the

38 Analysis and modelling of spatio and spatio-temporal point processes

random fields captures the residual variation, that can be seen as the epidemic component.

Given $(\mathbf{u}_i, t_i)$, thus the n -th order product density of a LGCP is given by

$$\begin{aligned} \rho^{(n)}((\mathbf{u}_1, t_1), \dots, (\mathbf{u}_n, t_n)) &= \mathbb{E} \prod_i^n \Lambda(\mathbf{u}_i, t_i) = \mathbb{E} \prod_i^n \lambda(\mathbf{u}_i, t_i) \exp(S(\mathbf{u}_i, t_i)) \\ &= \prod_i^n \lambda(\mathbf{u}_i, t_i) \exp\left(n\mu + \sigma^2 \left[\frac{n}{2} + \sum_{1 \leq i < j \leq n} \gamma((\mathbf{u}_i, t_i), (\mathbf{u}_j, t_j))\right]\right) \\ &= \prod_i^n \lambda(\mathbf{u}_i, t_i) \exp\left(n\mu + \sigma^2 \left[\frac{n}{2} + \sum_{1 \leq i < j \leq n} \gamma(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|)\right]\right) \end{aligned}$$

and this result is obtained considering the connection between the n -th moments of a Normal and a log-Normal distribution. Indeed, if $X \sim N(\mu, \sigma^2)$ then $\exp(X)$ has a log-Normal distribution with mean and variance equal to $(\exp(\mu + 0.5\sigma^2))$ and $\exp(2\mu + \sigma^2)(\exp \sigma^2 - 1)$, respectively. So the first- and the second-order product density of a log-Gaussian Cox process are:

$$\begin{aligned} \mathbb{E}(\Lambda(\mathbf{u}_i, t_i)) &= \rho(\mathbf{u}_i, t_i) = \lambda(\mathbf{u}_i, t_i) \exp\left(\mu + \sigma^2 \frac{1}{2}\right) \\ \mathbb{E}(\Lambda(\mathbf{u}_i, t_i); \Lambda(\mathbf{u}_j, t_j)) &= \rho^{(2)}((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) \\ &= \lambda(\mathbf{u}_i, t_i) \lambda(\mathbf{u}_j, t_j) \exp\left(2\mu + \sigma^2(1 + \gamma((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)))\right) \end{aligned}$$

Moreover, the covariance function (1.13), the pair correlation function (1.14), and the K-function (1.15) of a SOIRS and isotropic log-Gaussian Cox process are, respectively

$$\begin{aligned}
\varphi((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) &= \rho^{(2)}((\mathbf{u}_i, t_i), (\mathbf{u}_j, t_j)) - \rho(\mathbf{u}_i, t_i)\rho(\mathbf{u}_j, t_j) \\
&= \lambda(\mathbf{u}_i, t_i)\lambda(\mathbf{u}_j, t_j) \exp(2\mu + \sigma^2) \\
&\quad [\exp(\sigma^2\gamma(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|)) - 1]
\end{aligned}$$

$$\begin{aligned}
g((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) &= \frac{\rho^{(2)}((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j))}{\rho(\mathbf{u}_i, t_i)\rho(\mathbf{u}_j, t_j)} \\
&= \exp(\sigma^2\gamma((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j))) \\
&= \exp(\sigma^2\gamma(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|)) \\
&= \exp(\sigma^2\gamma(r, h)) = \exp(C(r, h)) \quad (1.20)
\end{aligned}$$

$$K(r, h) = 2\pi \int_0^r \int_0^h \exp(\sigma^2\gamma(u', t')) u' \, d(u', t') \quad r > 0, \quad h > 0,$$

Under the following assumption for the GRF

$$\mathbb{E}(S(\mathbf{u}_i, t_i)) = \mu = (-0.5\sigma^2), \quad (1.21)$$

we have

$$\mathbb{E}(\exp(S(\mathbf{u}_i, t_i))) = 1 \quad \mathbb{E}(\Lambda(\mathbf{u}_i, t_i)) = \lambda(\mathbf{u}_i, t_i) = \rho(\mathbf{u}_i, t_i).$$

In this way there is a clear separation between first-order properties that are described by the deterministic part $\lambda(\cdot)$, and second-order properties characterised by the parameters of the covariance matrix of the stochastic

component $S(\cdot)$.

This parametrisation allows to add covariate information in the model modifying properly λ . It also gives to the model a natural interpretation, by considering a multiplicative decomposition of the overall intensity into deterministic, $\lambda(\mathbf{u}_i, t_i)$, and stochastic components, $\exp(S(\mathbf{u}_i, t_i))$. If we wished to model a spatially and temporal varying intensity by including spatial and temporal covariates $Z_u(\mathbf{u})$ and $Z_t(t)$, the deterministic part in the spatio-temporal intensity (Equation (1.19)) is replaced by a regression model

$$\lambda(\mathbf{u}, t) = \lambda(Z_s(\mathbf{u}); Z_t(t); \boldsymbol{\beta}_s \boldsymbol{\beta}_t) = \lambda(Z_s(\mathbf{u}); \boldsymbol{\beta}_s) \lambda(Z_t(t); \boldsymbol{\beta}_t)$$

where $\boldsymbol{\beta} = \{\boldsymbol{\beta}_s, \boldsymbol{\beta}_t\}$ are the parameters related to the covariates. If it is assumed a separable structure of intensity, i.e. $\lambda(\mathbf{u}_i, s_i) = \lambda(\mathbf{u}_i) \lambda(s_i)$, any non-separable effects are interpreted as second-order, rather than first-order properties (Gabriel and Diggle, 2009). In the application described in Chapter 3, we use a combination of a non-parametric kernel estimate for the temporal part and a semi-parametric log-linear model for the spatial component.

Spatial and temporal components are often assumed discrete for convenient analysis purposes. In spatio-temporal processes and patterns, it is often convenient, and useful, for analysis purposes to allow either the spatial or temporal components to be discrete (Diggle, 2013; Diggle *et al.*, 2013). Indeed, in the case of earthquake data, both the spatial and temporal regions are essentially continuous; however in a long-term analysis, it may be meaningful changing the time resolution and considering a generic instant of time as k-days and we can refer to temporally discrete spatio-temporal point processes.

1.2.4.1 Covariance Functions for the Gaussian random field

A key aspect in LGCP models is the definition, and then the estimation, of the moments of the Gaussian Random Fields, S in Equation (1.19). Any valid family of spatio-temporal correlation functions for a Gaussian random field used in Geostatistics can define a class of spatio-temporal log-Gaussian Cox processes (Diggle *et al.*, 2013). The characteristics of the spatio-temporal covariance of the Random Field (separability, non-separability, isotropy or anisotropy structure) are inherited by the corresponding LGCP model.

If no further assumptions are considered, the covariance function for the Random Field depends on the space-time coordinates. The theoretical requirement for a function to be a valid covariance function is that the property of positive-definiteness holds, that is

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j C((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) \geq 0$$

for all finite sets of real coefficients a_i and points $(\mathbf{u}_i, t_i) \in W \times T$, $i = 1, \dots, n$. Other two important assumptions are the separability and the full symmetry. A stationary space-time covariance function is called separable if the spatio-temporal covariance structure can be factorised into purely spatial and purely temporal components. Instead, a stationary covariance function is called fully symmetric if it is spatially isotropic and temporally symmetric, i.e.

$$C(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|) = C(r, h)$$

If the covariance function has a parametric structure, it depends on a vector of parameter ϕ . For example, a separable structure for the covariance func-

42 Analysis and modelling of spatio and spatio-temporal point processes

tion can assume an exponential form for both the spatial and the temporal component,

$$C(r, h) = \sigma^2 \exp\left(\frac{r}{\phi_s}\right) \exp\left(\frac{h}{\phi_t}\right) \quad (1.22)$$

where $\phi = \{\sigma^2, \phi_s, \phi_t\}$ and more exactly σ^2 is the variance, ϕ_s is the scale parameter for the spatial distance and ϕ_t is the scale parameter for the temporal one. Another separable covariance function can have a Matérn form in space and a exponential form in time,

$$C(r, v) = \sigma^2 \left[\frac{2^{1-\tau}}{\Gamma(\tau)} \left(\sqrt{2\tau} \frac{r}{\phi_s} \right)^\tau K_\tau \left(\sqrt{2\tau} \frac{r}{\phi_s} \right) \right] \left[\exp\left(-\frac{h}{\phi_t}\right) \right] \quad (1.23)$$

where $\sigma^2, \phi_s, \phi_t, \tau > 0$, $\Gamma(\cdot)$ is the Gamma function, $K_\tau(\cdot)$ is the modified Bessel function of second kind and $\phi = \{\sigma^2, \phi_s, \phi_t, \tau\}$. Other basic stationary and isotropic family models are the Cauchy, Gaussian, Spherical covariance functions (Gelfand *et al.*, 2010).

In this thesis, we want to account also for non-separable properties, in order to describe more general situations. A non-separable structure of the covariance function is proposed by Gneiting (2002). Generally, the Gneiting covariance function can be written as

$$C(r, h) = \frac{1}{\psi(h^2)^{d/2}} \varphi\left(\frac{r^2}{\psi(h^2)}\right) \quad r \geq 0, h \geq 0$$

where φ is a complete monotone function associated to the temporal structure, and ψ a positive function with a completely monotone derivative associated to the spatial structure of data. For example, the choices $\varphi(r) = \sigma^2 \exp(-cr^\gamma)$ and $\psi(r) = (ar^\alpha + 1)^\delta$ yield to the parametric family

$$C(r, h) = \frac{\sigma^2}{(ah^{2\alpha} + 1)^{(\delta d)/2}} \exp\left(\frac{-cr^{2\gamma}}{(ah^{2\alpha} + 1)^{\delta\gamma}}\right) \quad (1.24)$$

of space-time covariance functions. Here, $a > 0$, $c > 0$ are scale parameters of time and space, α and γ are the smoothness parameters, and δ is the space-time interaction parameter that takes values in $(0, 1]$, and σ^2 is the process variance. So we have that $\phi = \{\sigma^2, a, c, \delta, \alpha, \gamma\}$. A separable covariance function is obtained when $\delta = 0$. The spatial margin is of the powered exponential form, and the temporal margin belongs to the Cauchy class. The aforementioned covariance functions have a parametric form; instead, Rodrigues and Diggle (2010) define a class of non-separable and non-parametric covariance function based on low-rank, convolution-based models.

In the analysis of Chapter 3, we use the covariance function in equation (1.24) setting some parameters and obtaining

$$C(r, h) = \sigma^2(h^{2z} + 1)^{-1} \exp\left(-\frac{r}{\alpha(h^{2z} + 1)^{\delta/2}}\right) \quad (1.25)$$

with $\sigma^2, \alpha > 0, z \in (0, 1], \delta \in [0, 1]$ and $\phi = \{\sigma^2, \alpha, z, \delta\}$.

Summing up, in the phase of model estimation the aim is the estimation of the two parameters vectors, β that is related to the first-order intensity and ϕ in order to explain the second-order properties of the Gaussian Random Field. Before explaining the main methods to estimate the parameters for the LGCP, we describe the case of geometric anisotropy for the spatial dimension.

1.2.4.2 Geometric anisotropy in space

The assumption of isotropy is often made in practice for simpler interpretation of the observed phenomena. Even these models are easier to build and estimate, isotropy is many times hard to find in real applications because

many point processes are indeed anisotropic, i.e. the spatial behaviour of the pattern varies with direction.

In this section, firstly we describe the geometric anisotropy in the spatial context following the formulation in Møller and Toftaker (2014). The geometric anisotropy is the simplest way to describe directionality in the process, because it is typically dealt by a transformation (in terms of rotation and scaling) of the spatial coordinates. Basically, the anisotropy for a LGCP is defined through an anisotropic structure of the covariance function $\mathbb{C}(S(\mathbf{u}_i, t_i), S(\mathbf{u}_j, t_j))$ of the Gaussian Random Field, S in Equation (1.19). In the context of spatial log-Gaussian Cox process, a geometric anisotropic process is obtained considering ellipsoids rather than circles or spheres isotropic covariance curves (centered at the point with which we are computing the covariance) (Gelfand *et al.*, 2010).

As a matter of fact, any valid spatial isotropic covariance function model can be generalized to make it geometrically anisotropic, by replacing the argument related to distance in space $r = (\mathbf{r}'\mathbf{r})^{1/2} = \|\mathbf{u}_i - \mathbf{u}_j\|$ (that is the Euclidean distance between two point in space and $\mathbf{r} = \mathbf{u}_i - \mathbf{u}_j$) with $r = (\mathbf{r}'\Sigma^{-1}\mathbf{r})^{1/2}$, where Σ is a 2×2 positive definite matrix.

More exactly, following the formulation in Møller and Toftaker (2014), the positive defined matrix Σ is equal to

$$\Sigma = U_\theta \text{diag}(\omega^2, \omega^2\zeta^2) U_\theta^\top = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \omega^2 & 0 \\ 0 & \omega^2\zeta^2 \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$

where $0 \leq \theta \leq \pi$, $\omega > 0$, $0 < \zeta \leq 1$ and U_θ is an orthonormal matrix ($UU^\top = I$). Considering the equation of the ellipse $E = \{\mathbf{u} : \mathbf{u}\Sigma^{-1}\mathbf{u}^\top = 1\}$, the parameter ω is the semi-major axis corresponding to the angle θ and $\omega\zeta$ is the semi-minor axis at the angle $\theta + \pi/2$. The parameter ζ is called the

anisotropic factor because it is the ratio between the minor axis and the major axis. Isotropy can be regarded as a special case of geometric anisotropy in which $\zeta = 1$ and $\Sigma = \omega^2 I$. In the next steps, we want to show why $r = (\mathbf{r}'\Sigma^{-1}\mathbf{r})^{1/2}$ can be rewritten in terms of distance in a transformation of the coordinates system. For the positive defined property of Σ , its square root matrix is $\Sigma^{1/2} = U_\theta \text{diag}(\omega, \omega\zeta)$ and the following relationships hold

$$\Sigma^{-1/2} = U_\theta \text{diag}(1/\omega, 1/(\omega\zeta))$$

$$\Sigma^{-1} = (\Sigma^{1/2}\Sigma^{1/2\top})^{-1} = \Sigma^{-1/2}\Sigma^{-1/2\top}$$

. We can write the argument of the covariance function as

$$\begin{aligned} r &= (\mathbf{r}'\Sigma^{-1}\mathbf{r})^{1/2} \\ &= (\mathbf{r}'\Sigma^{-1/2}\Sigma^{-1/2\top}\mathbf{r})^{1/2} \\ &= (\mathbf{r}'\Sigma^{-1/2}(\mathbf{r}'\Sigma^{-1/2})^\top)^{1/2} \\ &= \|\mathbf{r}'\Sigma^{-1/2}\| \end{aligned}$$

This means that the covariance can be seen as an isotropic covariance in a new system obtained rotating by an angle equal to θ and scaling the original y-coordinates by ζ . In Møller and Toftaker (2014), they study and develop estimation procedure using summary statistics to describe geometric anisotropy in particular in Cox processes. However, there are several possible settings to define the anisotropy, such as geometric, zonal or separable models of anisotropy. For a general overview of the recent advances see Allard *et al.* (2016), where they propose semi-parametric class of the directional mixture, defined through kernel functions.

In a spatio-temporal context, the covariance function of the Gaussian Random Field, $C(r, h) = \sigma^2 \gamma(r, h)$, and the corresponding pair correlation function for the LGCP, $g(r, h) = \exp(C(r, h))$, can be expressed as follow

$$C(\mathbf{r}, h) = C(\sqrt{\mathbf{r}^\top \Sigma^{-1} \mathbf{r}}, h) \quad g(\mathbf{r}, h) = g(\sqrt{\mathbf{r}^\top \Sigma^{-1} \mathbf{r}}, h)$$

where $\mathbf{r} \in R^2$ is a row vector, $\mathbf{r}^\top$ is its transpose, $g : R \times R \rightarrow [0; +\infty]$ satisfies the integrability condition $\int_0^r \int_0^t r g_0(r, h) dr dh < \infty$ for r and $h \in (0, \infty)$.

When we assume anisotropy in space-time, the elliptical anisotropic structure is constant over time, so the process has a main direction that does not change along time. The covariance function of the Gaussian Random Fields, such as those described in Equations (1.22)(1.23)(1.25), depends on additional parameters that describes the geometric anisotropy, and so the vector of parameters ϕ contains also θ , ζ and ω , this last parameter is equivalent to the spatial scale parameter ϕ_s .

1.2.4.3 Model estimation

In Diggle *et al.* (2013), they distinguish two main inferential aspects, estimation of model parameters and prediction of the realisations of the unobserved stochastic process. They describe three different parameter estimation methods: moment-based estimation, maximum likelihood estimation, and Bayesian estimation. We describe in details the first approach because, even it has *ad hoc* quality and the results depend on the initial settings, it is easy to apply and it has operative and computational advantages. The main goal in using this method is to give preliminary estimates of the parameters of the model.

We first describe the estimation procedure used under the isotropic assumption, that has been proposed in Brix and Diggle (2001); Diggle *et al.* (2005) and assessed with simulations in Davies and Hazelton (2013).

The method has two steps, the first one concerns the estimation of the first-order intensity function $\lambda(\mathbf{u}_i, t_i) = \lambda(Z_s(\mathbf{u}); Z_t(s); \boldsymbol{\beta}_s \boldsymbol{\beta}_t)$. Assuming an inhomogeneous Poisson point Process, the following log-likelihood is maximised

$$\sum_{i=1}^n \log \lambda(\mathbf{u}_i, t_i) - \int_0^T \int_W \lambda(\mathbf{u}, t) d\mathbf{u} dt$$

with respect to $\boldsymbol{\beta} = \{\boldsymbol{\beta}_s \boldsymbol{\beta}_t\}$ and so obtaining $\hat{\lambda}(Z_s(\mathbf{u}); Z_t(t); \hat{\boldsymbol{\beta}}_s \hat{\boldsymbol{\beta}}_t)$.

In the second step, the parameters of the covariance function of the Gaussian Random field ($\boldsymbol{\phi}$) are estimated with the moment-based approach. It consists of minimising a measure of the discrepancy between the empirical and theoretical second-moment properties of the process. One class of such measures can be a weighted least squares criterion, and the following quantity is minimize with respect to $\boldsymbol{\phi}$

$$D(\boldsymbol{\phi}) = \int_0^{r_0} \int_0^{h_0} w(r, h) \{ [\hat{f}(r, h)]^c - [f(r, h; \boldsymbol{\phi})]^c \}^2 dr dh \quad (1.26)$$

where $\hat{f}$ is an estimate of second order-summary statistics (it can be the spatio-temporal K-function or pair correlation function), $f(\cdot)$ is the theoretical corresponding second-order summary statistics that depends on $\boldsymbol{\phi}$, (r_0, h_0) are the fixed space-time maximum distances, $w(r, h)$ are weights that depend on the space-time distance and c is the power to apply on the function under consideration. Basically, substituting properly quantity for $\hat{f}$ and f , the parameter vector $\boldsymbol{\phi}$ are the parameters that define the covariance structure of the Gaussian Random Field of the LGCP. The quantities

(r_0, h_0) , $w(r, h)$ and c are user-specific, and the last two are used to stabilise the sampling variation.

The method of moments is easy to apply, but the results are sensible to the initial setting values and it can be used when the number of parameters in ϕ is low.

The R (R Development Core Team, 2005) package `lgcp` (Taylor *et al.*, 2013, 2015) includes routine model estimation of the spatial and spatio-temporal modelling of point patterns using the log-Gaussian Cox process based on the moment-based approach and on the Bayesian inference. It also implements code for point process simulation and predictive inference using MCMC methods. However, for the spatio-temporal setting the option for a non-separable covariance function for the Gaussian process has not been implemented yet. Furthermore, the package `stpp` Gabriel *et al.* (2013) implements code for point pattern simulation for both the separable and non-separable assumption of the covariance function of the Gaussian Random Field. For the analysis of Section 3.3, we use and modify some routines and functions of the `lgcp` and `stpp` packages.

In the case of geometric anisotropy, the moment-based method previously described can be used after rotating and scaling the x- and y-coordinates system according to the parameters that characterised the geometric anisotropy, that is assumed constant in time. In the remain part, we explain the estimation procedure proposed in Møller and Toftaker (2014) and after we describe our proposed extension in space and time.

The estimation procedure proposed in Møller and Toftaker (2014) in case of spatial inhomogeneous point process can be summarised in the following steps:

1. In the first step, the spatial intensity is estimated. When ρ is assumed

to be constant on W , the usual non-parametric estimate is $\hat{\rho} = n/|W|$, where n denotes the observed number of points. In the case of inhomogeneity, a kernel non-parametric estimation can be used.

2. The second step concerns the estimation of the angle θ and the anisotropic parameter ζ through the anisotropic spatial pair correlation function.

Following Stoyan and Stoyan (1994), we briefly describe the anisotropic version of the pair correlation function, its estimator and its usage to compute $\hat{\theta}$ and $\hat{\zeta}$.

The Equation (1.1) can be extended considering

$$g_1(r, \psi) = \frac{\rho^{(2)}(r, \psi)}{\rho^2} \quad r \geq 0; 0 \leq \psi < \pi$$

where in the second-order product density $r = \|\mathbf{u}_i - \mathbf{u}_j\|$ is as usual the euclidean distance between two pairs of points and ψ is the angle that defines the direction of the line through $\mathbf{u}_i$ and $\mathbf{u}_j$ and the abscissa axis. As explained in Stoyan and Stoyan (1994), the form of $g(r, \psi)$ for different ψ closely reflects the form of the anisotropy. As a matter of fact, the values of ψ at which the differences are greater demonstrates the main directions of the points. It is sufficient to consider only values of ψ between 0 and π because same values are obtained for ψ and $\psi + \pi$. The estimator of the anisotropic pair correlation function proposed by Stoyan and Stoyan (1994) is

$$\hat{g}_1(r, \phi) = \hat{g}_1(r, \phi + \pi) = \frac{1}{2} \sum_{i=1}^n \sum_{j \neq i} \frac{H(\mathbf{u}_i - \mathbf{u}_j, (r, \phi)) + H(\mathbf{u}_i - \mathbf{u}_j, (r, \phi + \pi))}{\hat{\rho}(\mathbf{u}_i) \hat{\rho}(\mathbf{u}_j) w(\mathbf{u}_i, \mathbf{u}_j)}$$

where $\hat{\rho}$ is an estimate of the first-order intensity, $w(\mathbf{u}_i, \mathbf{u}_j)$ is an edge correction, and

$$H(\mathbf{u}_i - \mathbf{u}_j, (r, \phi)) = \frac{1}{r} \kappa_{\epsilon_r}(\|\mathbf{u}_i - \mathbf{u}_j\| - r) \kappa_{\epsilon_\phi}(\alpha(\mathbf{u}_i, \mathbf{u}_j) - \phi)$$

where $\alpha(\mathbf{u}_i, \mathbf{u}_j)$ is the angle between the directed line from $\mathbf{u}_i$ to $\mathbf{u}_j$ and the abscissa-axis, κ_{ϵ_r} and κ_{ϵ_ϕ} are kernel functions with bandwidth ϵ_r and ϵ_ϕ , respectively.

First the parameter θ is estimated and then an estimate of ζ is obtained. Given a spatial point pattern x with a generic point $\mathbf{u}_i$, the following discrepancy measure is considered

$$D(\phi) = \int_{a_1}^{b_1} [g_1(r, \phi) - g_1(r, \phi + \pi/2)] dr, \quad 0 \leq \phi \leq \pi, \quad (1.27)$$

We consider a set of distances $\{r_i : i = 1, \dots, n_r\}$, a set of angle values $\{\phi_j = (j - 0.5)\pi/n_\phi : j = 1, \dots, n_\phi\}$ and the following quantity is maximised with respect to ϕ_j

$$\hat{\theta} = \arg_{\phi_j} \max \hat{D}(\phi_j) \approx \sum_{i_1}^{r_i} [\hat{g}_1(r_i, \phi_j) - \hat{g}_1(r_i, \phi_j + \pi/2)] / n_r \quad j = 1, \dots, n_\phi$$

After computing $\hat{\theta}$, we transform the points in the pattern $\mathbf{x}$ (where $\mathbf{u}_i$ is a generic point) and the window W in the following way

$$\mathbf{v}_{i\zeta}^\top = \mathbf{u}_i^\top \begin{bmatrix} \cos \hat{\theta} & -\sin \hat{\theta} \\ \sin \hat{\theta} & \cos \hat{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta \end{bmatrix}$$

$$V_\zeta = W \begin{bmatrix} \cos \hat{\theta} & -\sin \hat{\theta} \\ \sin \hat{\theta} & \cos \hat{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\zeta \end{bmatrix}$$

We consider a new point pattern $\mathbf{y}_\zeta$ formed by the n points $\mathbf{v}_{i\zeta}$ and a new window V_ζ that depend on the parameter ζ . We fix the set of values $\{\zeta_h = h/n_\zeta : h = 1, \dots, n_\zeta\}$. The parameter $\hat{\zeta}$ is the ζ_h value that minimises the following quantity

$$\hat{\zeta} = \arg_{\zeta_h} \min \left[\sum_{j=1}^{n_\phi} \left(\hat{D}_{y_{\zeta_h}}(\phi_j) - \frac{1}{n_\phi} \sum_{i=1}^{n_\phi} \hat{D}_{y_{\zeta_h}}(\phi_i) \right)^2 \right] \quad (1.28)$$

where $\hat{D}_{y_{\zeta_h}}$ is defined (1.27) computed with respect to the transformed point pattern y_{ζ_h} that depends on $\hat{\theta}$ and the anisotropic factor ζ_h . The reason behind the minimisation of formula in (1.28) is that, in the isotropic case of y_{ζ_h} , the anisotropic pair correlation function $\hat{g}(r, \phi)$ is constant with respect to ϕ , and so $\hat{D}(\phi)$ is expected to deviate less from its mean than it would be in the anisotropic case (Møller and Toftaker, 2014).

3. Having the following estimates $\{\hat{\theta}, \hat{\zeta}\}$, the anisotropic point pattern $\mathbf{x}$ and the window W is transformed as follow

$$\mathbf{v}_i^\top = \mathbf{u}_i^\top \begin{bmatrix} \cos \hat{\theta} & -\sin \hat{\theta} \\ \sin \hat{\theta} & \cos \hat{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\hat{\zeta} \end{bmatrix}$$

$$V = W \begin{bmatrix} \cos \hat{\theta} & -\sin \hat{\theta} \\ \sin \hat{\theta} & \cos \hat{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\hat{\zeta} \end{bmatrix}$$

where now the new point pattern $\mathbf{y}$ is formed by the n $\mathbf{v}_i$ points and it

has an isotropic structure.

4. The remain parameters in ϕ are ω and the parameters for the covariance function. Their estimation is based on the transformed point pattern y . Basically, in this step the method of the moments described before is applied on the isotropic point pattern y and we want to point out that the parameter ω is the scale parameter in the covariance function for the spatial distance.

In the spatio-temporal framework, if the time is discrete and we have several observations for each time t with $j = 1, \dots, T$, we can consider our observed point pattern as composed of T snapshots with n_t points such that $\sum_{j=1}^T n_j = n$. For each time of occurrence t , the parameters θ_j, ζ_j with $j = 1, \dots, T$ are computed considering the previously explained method. Finally, the space coordinates of the points in the spatio-temporal point pattern are rotated and scaling according to an overall measure of the previous quantities. The set of the estimated directions θ_j are axial data. The standard approach to analysing axial data is to double their angles, and to apply circular statistics. We summarise the directions computing the circular mean and by dividing it by 2, obtaining $\bar{\theta}$. Instead the values of ζ_j are summarised with their mean, $\bar{\zeta}$.

1.2.5 Epidemic type aftershocks stochastic model

In a model specification, it can be of interest to specify the properties of a process conditionally on its realisations up to the current time, i.e. by conditioning on the past. A widely used branching-type model for earthquake description is the Epidemic Type Aftershock-Sequences (ETAS) model based on the Gutenberg-Richter law and Omori law (Ogata, 1988). It is defined

throughout a conditioned intensity function, that is the expected frequency of events in a space-time unit region, conditional on the history $\mathcal{H}_t$ of the point process up to time t , i.e.,

$$\rho_c(\mathbf{u}, t | \mathcal{H}_t) = \lim_{|d\mathbf{u}| \rightarrow 0; |dt| \rightarrow 0} \frac{\mathbb{E}[Y(d\mathbf{u}, dt) | \mathcal{H}_t]}{|d\mathbf{u}||dt|}$$

where $\mathbf{u} = \{x, y\}$. Generally, the conditional intensity $\rho_c(\mathbf{u}, t)$ depends on some unknown parameters. The conditional intensity function of the branching model in a space-time point $(\mathbf{u}, t)$ is defined as the sum of a term describing the long-term variation (spontaneous activity or background) and one relative to the short-term variation (induced activity or offsprings):

$$\rho_{c\theta}(\mathbf{u}, t | \mathcal{H}_t) = \mu f(\mathbf{u}) + \tau_\phi(\mathbf{u}, t) \quad (1.29)$$

with $\theta = (\mu, \phi)'$, the vector of parameters of the induced intensity (ϕ) together with the parameter of the general background intensity (μ), $f(\mathbf{u})$ is the space density, and $\tau_\phi(\mathbf{u}, t)$ is the induced intensity, given by $\tau_\phi(\mathbf{u}, t) = \sum_{t_j \leq t} \nu_\phi(\mathbf{u} - \mathbf{u}_j, t - t_j)$, with $\nu(\cdot)$ a space-time density function.

If we consider a seismic catalog of n events, the space-time-magnitude coordinates of the i -th event are $(\mathbf{u}_i, t_i, m_i)$. The ETAS conditional intensity function can be written as follows:

$$\rho_{c\theta}(\mathbf{u}, t | \mathcal{H}_t) = \mu f(\mathbf{u}) + \sum_{t_j \leq t} \nu_\phi(\mathbf{u} - \mathbf{u}_j, t - t_j | m_j) \quad (1.30)$$

where m_j is the magnitude of the j -th event, $\mu f(\mathbf{u})$ is the time stationary background seismicity, and

$$\nu_\phi(\mathbf{u} - \mathbf{u}_j, t - t_j | m_j) = g(t - t_j | m_j) \ell(\mathbf{u} - \mathbf{u}_j | m_j).$$

54 Analysis and modelling of spatio and spatio-temporal point processes

The $g(\cdot)$ function is the probability function of the occurrence times of triggered events following the earthquake at time t_j with magnitude m_j . It takes the following form

$$g(t - t_j | m_j) = \frac{\kappa_0 e^{(\alpha - \gamma)(m_j - m_0)}}{(t - t_j + c)^p}, \text{ with } t > t_j,$$

where κ_0 is a constant which measures the aftershocks productivity, c and p are parameters of the modified Omori's law, p characterises the pattern seismicity in the given region, indicating the decay rate of aftershocks in time, α measures the influence on the relative weight of each sequence and m_0 is the completeness threshold of magnitude i.e., the lower bound for which earthquakes with higher magnitude are surely recorded in the catalogue.

The $\ell(\cdot)$ function describes the spatial rate of the locations of the triggered events conditioned to the magnitude of the generating events

$$\ell(x - x_j, y - y_j | m_j) = \left\{ \frac{(x - x_j)^2 + (y - y_j)^2}{e^{\gamma(m_j - m_0)}} + d \right\}^{-q}.$$

where γ measures the influence on the relative weight of each sequence, d and q are two parameters related to the spatial influence of mainshocks. With this parametrisation, we have $\theta = (\mu, \kappa_0, c, p, \alpha, \gamma, d, q)$.

1.2.5.1 Model estimation

The log-likelihood for this model is defined as

$$\sum_{i=1}^n \log \rho_{c\theta}(\mathbf{u}_i, t_i | \mathcal{H}_{t_i}) - \int_0^T \int_W \rho_{c\theta}(\mathbf{u}, t | \mathcal{H}_t) d\mathbf{u} dt$$

In such models, the main issue is the simultaneous estimate of the two different components of the intensity function (large-time scale and small-time scale). If the large-time scale component $\mu f(\mathbf{u})$ in (1.29) is known, the remain parameters θ is usually estimated by Maximum Likelihood method (Ogata, 1988). In applications, the background component $\mu f(\mathbf{u})$ is usually estimated by non-parametric techniques, like kernel estimators. A procedure to maximize the expected complete data log-likelihood function, based on the expectation-maximization algorithm, was introduced in Veen and Schoenberg (2008).

In Adelfio and Chiodi (2015a,b) the FLP (Forward Likelihood for Prediction) approach is used to evaluate this bandwidth vector of the background intensity of a space-time ETAS model. Moreover for each event, the probability of being a background event is estimated in order to provide an estimate of weights for a non-parametric density estimation of the spatial background component.

We have describe the ETAS model since it is used for some analysis. In particular, in Chapter 2, the FLP procedure is considered to get valid estimates of the spatial bandwidths for the description of the inhomogeneous spatial intensity. Furthermore in Chapter 4, the ETAS model is estimated for the spatio-temporal Greek data and the final probabilities provided in the last iteration step are used to classify the earthquakes in background and triggered events.

Chapter 2

Spatial pattern analysis using hybrid models: an application to the Hellenic seismicity

In this chapter, we present the main results obtained applying Hybrid of Gibbs point process model (explained in Section 1.1.3) for the description of the spatial displacement of earthquakes in the Greek area over the last 10 years. The novel aspect of this contribution is to account, for the first time in the geological context of complex point process models, for the geological information in the description of earthquake spatial displacement. As a matter of fact, this is a new field of research for the description of the characteristics of spatial and spatio-temporal distribution of earthquakes, including additional external information.

After a brief literature review in Section 2.1, the study area and the available data are described in Section 2.2. The results of the descriptive analysis and the estimated models are presented in Section 2.3 and Section 2.4, respec-

tively. After comparing several models using appropriate diagnostic tools, we find two competing suitable hybrid models, one with a full parametric structure and the other one based on nonparametric kernel estimators for the spatial inhomogeneity.

2.1 Literature review

Since the earthquake data show cluster structure, the modelling formulation is usually based on self-exciting point processes. These models assume that the occurrence of an event increases the probability of occurrence of other events in time and space. In particular, Hawkes model (Hawkes and Adamopoulos, 1973) or ETAS model (Ogata, 1988) are examples of self-exciting point processes. On the other hand, when an inhibitive pattern is observed, an important class of processes is represented by self-correcting processes. In these models, it is assumed that each event, that occurs in a given point in time, has an inhibiting effect to the following ones. In the seismological context, self-correcting processes are represented by the strain-release models (Vere-Jones, 1978), for which the fault fracture that generates earthquakes reduces the amount of strain present at the break point along the fault. An important example of self-correcting processes is the Strauss point process model (Strauss, 1975) with a negative association (Equation (1.7)). It is also the simplest inhibitive Markov (or Gibbs) point process, which describes patterns characterised by a local interaction between neighbouring points.

In the context of spatial point patterns, Gibbs point processes are generally applied in the analysis of spatial interaction and inhomogeneity depending on covariate information. For example, Anwar *et al.* (2012) and Ye *et al.*

(2015) use a multitype Strauss process considering geological spatial covariates to describe the spatial pattern of the aftershocks epicentres following two strong earthquakes occurred in Kashmir (Pakistan-2005) and in the Yellowstone National Park in 1975. However, Gibbs models could not be appropriate to describe strong clustering and multi-scale variations that are characteristics of most natural processes, such as earthquakes (Illian *et al.*, 2008; Møller and Waagepetersen, 2003).

Although the models cited above represent a wide variety of solutions to describe different types of spatio and spatio-temporal patterns, seismological events need complex models that are able to account for environmental heterogeneity while detecting interactions at several spatial scales, usually in terms of clustering.

Motivated by these needs, we propose using hybrid models (Baddeley *et al.*, 2013) to analyse earthquake data. Hybrid models are used for patterns that show interactions at different spatial scales, such as human social interaction (Baddeley *et al.*, 2013) and spatial distribution of plants (Badreldin *et al.*, 2015). More exactly, we describe the spatial intensity of the events with a magnitude greater or equal than 4 occurred in the Hellenic area between 2005 and 2014. We characterise in a proper way the spatial interaction between events happening at several spatial scales, and the spatial inhomogeneity of the process considering the geological information available in the study area (the presence of volcanoes, plate boundaries and faults).

2.2 The study area and the geological information

Earthquakes are the most unpredictable damaging natural disaster. An earthquake is a sudden break and movement in the earth lithosphere due

to the release of stress accumulated along tectonic plates, faults or volcanic areas. Earthquake damage is caused by the shaking of the ground generated by seismic waves that reach the earth surface. The level of damages mainly depends on the earthquake magnitude, the hypocentral depth, the epicentral distance and on the local shallow geology, which can generate local amplification.

In this study, the class of hybrid models proposed in Baddeley *et al.* (2013) are applied to the study of the Greek-Hellenic seismicity, an area of high seismic hazard characterised by both tectonic and volcanic seismogenic sources. The Hellenic arc is the most seismically active area of the European-Mediterranean region having experienced many destructive earthquakes. In the last century at least a dozen of earthquakes have affected the Greek area with devastating consequences (Papazachos *et al.*, 1997). The most destructive earthquake occurred in 1926 (M_S 7.4) at about 115 km of depth just beneath the Hellenic volcanic Arc (Ambraseys and Adams, 1998), in 1933 (M_W 6.6) with the epicentre near the Kos island (M_W 6.6, Papazachos *et al.* (1997)) and in 1956 (M_S 7.4) located in the central Santorini-Amorgos graben. Recently, Greece was affected by an important seismic sequence characterised by two main earthquakes (both M_W 6.0) occurred within eight days, on the 26th January and the 3rd of February, 2014 (Sokos *et al.*, 2015).

Hellenic seismicity is generated by relative compressional motion between the African and Eurasian plates started at Oligo-Miocene time, about 26 Ma (million years) ago Meulenkamp *et al.* (1988). The African plate is subducting beneath the Aegean lithosphere with a relative velocity of about 4 cm per year Le Pichon and Angelier (1979) (see Fig. 2.1) forming the Hellenic Subduction Zone (HSZ) and the chain of volcanic island in the

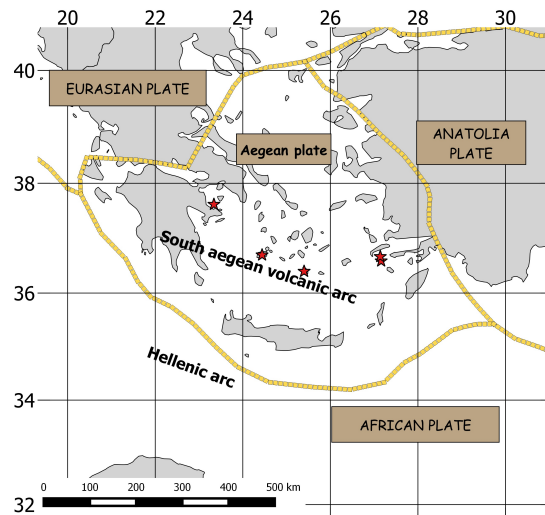


Figure 2.1: Tectonic plates in the Hellenic region and red stars indicate the main volcanic centers.

South Aegean Sea (Hellenic Volcanic Arc, HVA). The HSZ is the main source of earthquakes dominated by shallow intraplate seismicity within the Aegean plate, interplate seismicity along the plate contact and by intermediate deep seismicity along the subducting African slab. Moving from West to East, the main volcanoes of the HVA are Methana, Milos, Santorini, Yali and Nisyros (Siebert and Simkin, 2014) (see Table 2.1). In general, the seismicity activity along the HVA is smaller in terms of magnitude than the seismic activity along the HSZ. Most of the volcanoes of the HVA are not active in terms of eruptions but are sources of microseismic activity (Bohnhoff *et al.*, 2006; Dimitriadis *et al.*, 2009).

Hellenic seismicity is monitored by four seismological centres: the Institute of Geodynamics and the Seismological Laboratories of the Universities of Athens, Thessaloniki and Patras. In the beginning of 2005, a national

Name	Lat	Long	Elevation (meters)	Type
Methana	37.62	23.34	760	Lava domes
Milos	36.70	24.44	751	Stratovolcano
Santorini	36.40	25.40	367	Shield volcano
Yali	36.67	27.14	180	Lava domes
Nisyros	36.59	27.16	698	Stratovolcano

Table 2.1: The main volcanoes in the south Aegean volcanic arc moving from West to East.

project was launched to unify the seismological networks in Greece (the Hellenic Unified Seismological Network, HUSN). The HUSN is nowadays constituted by about 150 seismic stations adequately distributed over the area to be monitored and by portable stations moved according to the investigation requirements (D'Alessandro *et al.*, 2011).

Focal parameters of the earthquake occurring in the Hellenic region are routinely determined from the analysis of data recorded by the HUSN, and the earthquakes catalogue for the area contains events from 1964 up to nowadays. The accuracy and precision with which such parameters (hypocenter coordinates and event magnitude) are determined is mainly a function of the earthquake magnitude and of the number and distribution of the seismic stations. Small magnitude earthquakes are generally recorded from few stations (only the nearest ones) and consequently, focal parameters are not estimated reliably. Given the purpose of the work, in order to ensure a dataset as homogeneous as possible, in terms of data quality and reliability of focal parameters, in our analysis we have considered only earthquakes occurred in the Hellenic region between 2005 and 2014, with magnitude equal or greater than 4. So fixing this value of the magnitude, we study the spatial intensity of the main events occurred in the area excluding the

micro and the minor events according to a qualitative classification of the magnitude.

Considering a period of 10 years, we also analyze and describe in a proper way the spatial interactions between points, neglecting time. We selected this time period not only for a technical reason due to the unification of the seismological networks, but also for having a good trade-off between a sufficient number of points and the model complexity that is related to the computational efficiency. Moreover aggregating over a longer time interval, the scale of aggregation observed in Figure 2.2 changes and the clusters or gaps observed could be cancelled out giving a different interpretation of the phenomenon (Illian *et al.*, 2008).

We additionally considered spatial geological information available in the area to understand dependence between events and the different sources of earthquakes.

In Fig.2.2, the dotted line is the Aegean plate boundary coming from a digitalized global set of present plate boundaries on the earth Bird (2003). Furthermore, we used the Greek Database of Seismogenic Sources (GreDaSS) that concerns tectonic and active-fault data in Greece and its surroundings Caputo *et al.* (2013). Generally, a ‘fault’ is a fracture in the earth along which the two blocks of crust have been moved relative to each other. The term ‘active-faults’ refers to faults in which a displacement has occurred in the recent geologic time (‘seismogenic sources’) or to areas assumed to be capable of producing moderate to strong earthquakes (‘capable faults’). A capable fault will eventually become seismogenic sources in the future, and a seismogenic source was a capable fault, and usually the study of these is of interest in the seismic hazard analysis. An active fault has been identified integrating different available piece of information coming from historical

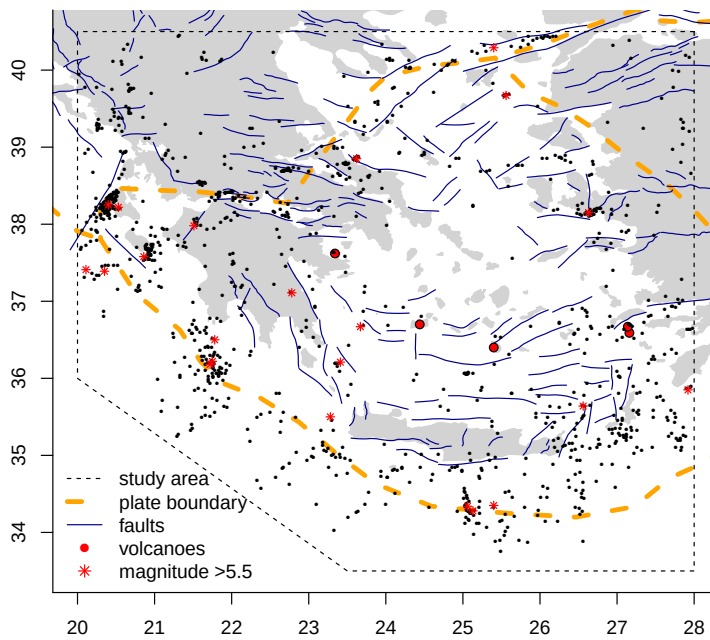


Figure 2.2: Faults, plate boundary and main volcanoes in the Hellenic area. The dotted black line indicates the study window. The points are the observed events with a magnitude greater or equal than 4, red stars indicate points with a magnitude greater than 5.5.

and instrumental seismicity data, geological, structural, morphotectonic, paleoseismological and geophysical investigations.

The database provides metadata associated to each fault such as geometric (strike, dip, width, depth) and kinematic (rake) parameters and descriptive information (e.g. comments, latest earthquakes). In our analysis we considered the spatial shape file about the faults associated to the Composite Seismogenic Sources (CSS), see Fig.2.2. We note that the seismogenic sources are not homogeneous in all the Hellenic area because the historical information is richer in the most populated areas and it is difficult to detect

in offshore regions, e.g. the Aegean Sea.

Since in our analysis we want to detect whether and how the intensity depends on the geological information, we converted the spatial information on the faults, the plate boundary and the volcanoes into distance functions. For example using the information about coordinates of the volcanoes, we computed a spatial covariate that indicates for a generic location $\mathbf{u}$ in W the distance to the nearest volcano. In Table 2.2, the three spatial covariates are reported.

Var	Description	P-value
D_{pb}	Distance to the plate boundary	0
D_f	Distance to the nearest fault	0
D_v	Distance to the nearest volcano	0.880

Table 2.2: List of the spatial covariates and p-values for the Berman test.

2.3 Descriptive analysis

The events in the study window with a magnitude greater or equal than 4 between 2005 and 2014 define a spatial point pattern with 1105 events. Distances were computed considering geographical coordinates (longitude and latitude) and expressed in kilometres to interpret and comment the main results. The pattern has a clear spatial inhomogeneity and multi-scale interactions between points, for instance most of the events are around the plate boundary and we can visually identify clusters of events with a magnitude greater than 5.5 (Fig.2.2).

Figure 2.3a shows the non-parametric empirical K-function (Equation (1.3)) and the corresponding envelope that was obtained considering the K-function for a large number of simulated patterns assuming complete spatial random-

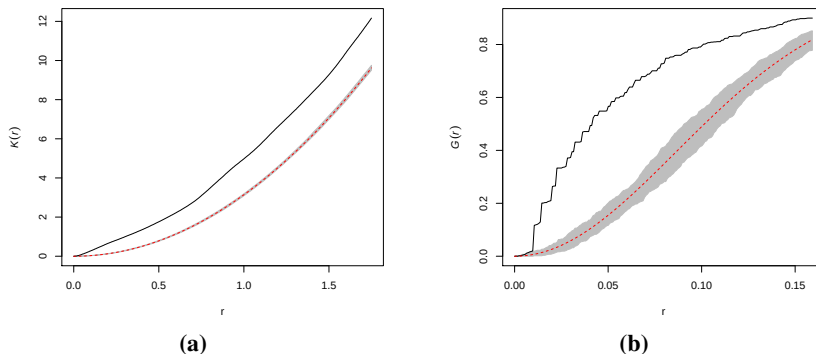


Figure 2.3: (2.3a) The envelope of the K-function (shaded region) to test Complete Spatial Randomness (CSR) condition of the earthquake events. The black curve is the estimated K-function instead the red one correspond to the function under the CSR assumption. (2.3b) The envelope of the G-function (shaded region) to test Complete Spatial Randomness (CSR) condition of the earthquake events. The black curve is the estimated G-function instead the red one correspond to the function under the CSR assumption.

ness and the same intensity of the observed pattern. The upper and lower envelopes are the pointwise maximum and minimum of the simulated values of the K-function. Moreover, Figure 2.3b shows the non-parametric empirical G-function (theoretical form in Equation (1.4)) and its corresponding envelope. It is clear that the observed pattern is not a Poisson process since the empirical K- and G-functions are outside the shadow region.

The Berman test (Berman, 1986; Baddeley and Turner, 2005) is used to assess the dependence of a point process on a spatial continuous covariate. Table 2.2 shows the corresponding p-values indicating that the variable D_v (distance to the nearest volcano) indicates that this variable is not significant to add valued information to the spatial intensity of the process. In Fig. 2.4b, the previous result is confirmed as the smoothed estimate of

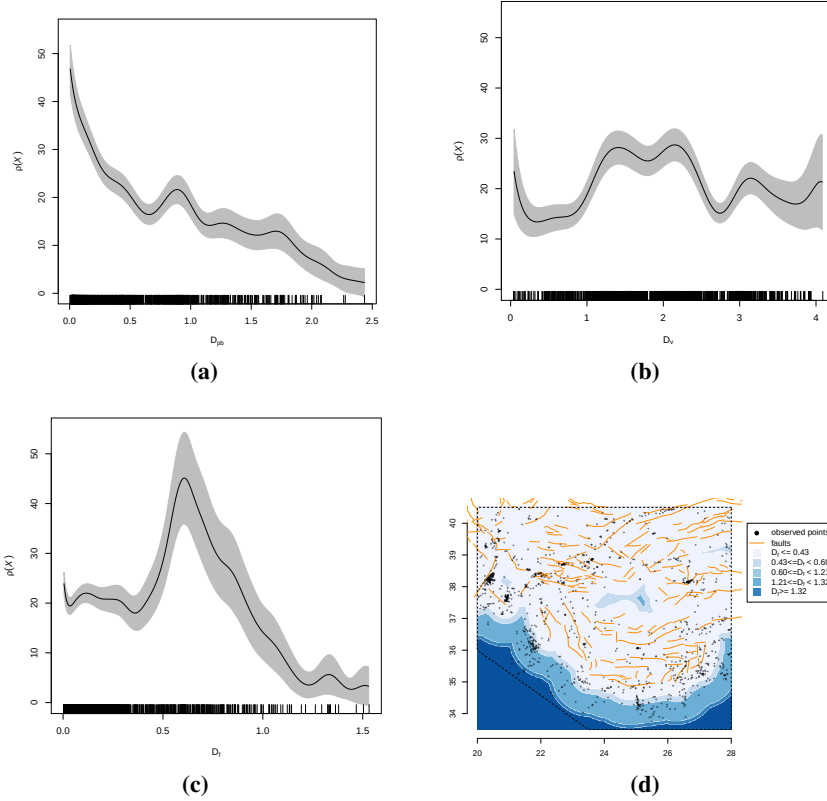


Figure 2.4: Smoothed estimates of the intensity of the point process, as a function of the spatial covariates, distance to the plate boundary (2.4a), the nearest volcano (2.4b) and the nearest fault (2.4c). Discretisation of the spatial variable D_f (distance to the nearest fault) (2.4d).

the intensity as a function of the distance to the nearest volcano is basically constant. Instead, the intensity of the point process decreases exponentially moving away from the plate boundary (variable D_{pb}). The intensity has a piecewise trend with respect to D_f (distance to the nearest fault) with a peak at degree distance of 0.6 (around 65 km) that corre-

sponds to the events that lie in the south part of the study region around the plate boundary. Considering a segmented linear regression between D_f and the intensity, we found the following breakpoints in terms of degree distances $\phi = \{0.43, 0.60, 1.21, 1.32\}$ for which there is a change in the sign of the slope (Muggeo, 2003, 2008). Figure 2.4d shows the discretization of the spatial covariate D_f , where the points in the study region for which $(\phi_i \leq D_f < \phi_{i+1})$ have a distance to the nearest fault between ϕ_i and ϕ_{i+1} .

2.4 Statistical inference and modelling proposal

2.4.1 Inhomogeneous Poisson models

The first attempt in model estimation is to fit an inhomogeneous Poisson model with intensity depending on the spatial covariates, that is equal to equation (1.10) where the term $G(\cdot)$ is null.

According to the descriptive results, we considered the spatial covariates assuming that the log of the intensity depends linearly on D_{pb} and piecewise linearly on D_f .

As for the general trend with respect to the spatial coordinates ($\mathbf{u} = \{x, y\}$), we considered both a parametric polynomial trend and a non-parametric trend, this last computed using a Gaussian kernel estimator selecting the bandwidth values with the methods indicated in Section 1.1.3.1 (cross-validated bandwidth selection, likelihood cross-validation, Scott's rule and the FLP method). Comparing different semi-parametric models in terms of AIC, the model with kernel estimated using the FLP method performed better than the others and the obtained bandwidth values were $(h_x, h_y) = (0.22, 0.19)$. On the other hand, for the full parametric model, a second-order polynomial in x and y for the spatial trend fits reasonably good (see

Table 2.3: *ModPoisson_a* - Estimated parameters for the inhomogeneous Poisson point process model with Gaussian kernel estimator for x and y and bandwidth values $(h_x, h_y) = (0.22, 0.19)$.

	Estimate		Zval
(Intercept)	0.2433	***	3.57
D_{pb}	-0.1153		-1.78
$I(D_f < \phi_1)D_f$	-1.2557	***	-3.72
$I(\phi_1 \leq D_f < \phi_2)D_f$	-0.0633		-0.28
$I(\phi_2 \leq D_f < \phi_3)D_f$	-0.4000	**	-3.21
$I(\phi_3 \leq D_f < \phi_4)D_f$	-0.8222		-1.78
$I(D_f \geq \phi_4)D_f$	-0.3021		-1.28
AIC	-6382.877		
N^o of simulated points	1101		
Range of raw residuals	[-1.91;2.33]		

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2.4 and Figures 2.5c and 2.5d). We note that using a higher-order polynomial drives into computational errors in the estimation of the Fisher information matrix.

We selected the significant covariates comparing nested models in terms of the residual deviance. The estimated parameters θ_1 in (1.10) for the inhomogeneous Poisson process both with a kernel estimation (*ModPoisson_a*) and a second-order polynomial (*ModPoisson_b*) with respect to x and y are reported in Tables 2.3 and 2.4, respectively. Using the spatial kernel, the significant variables are D_{pb} and some of the component of the piecewise linear regression in D_f , whilst using a polynomial trend all the spatial covariates are significant. The estimated parameters of the models are not directly comparable since they have two different spatial trends, however, in both cases, we can assert that increasing the distance to the nearest fault and to the plate boundary the intensity decreases. For both models, the number of simulated points is close to the observed one, but their spatial

Table 2.4: *ModPoisson_b* - Estimated parameters for the inhomogeneous Poisson point process model with a second-order polynomial with respect to the coordinates x and y .

	Estimate		Zval
(Intercept)	-146.4664	**	-3.01
D_{pb}	-1.0124	***	-13.69
$I(D_f < \phi_1)D_f$	-2.8391	***	-8.15
$I(\phi_1 \leq D_f < \phi_2)D_f$	-1.3013	***	-5.10
$I(\phi_2 \leq D_f < \phi_3)D_f$	-1.6692	***	-10.52
$I(\phi_3 \leq D_f < \phi_4)D_f$	-2.4947	***	-5.28
$I(D_f \geq \phi_4)D_f$	-2.0786	***	-8.25
D_v	0.1887	*	2.01
x	0.3864		0.42
y	8.0901	***	3.90
x^2	0.0384	***	3.37
xy	-0.0611	***	-5.08
y^2	-0.0920	***	-3.63
AIC	-5247.943		
N^o of simulated points	1168		
Range of raw residuals	[-5.77;6.14]		

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

displacement does not capture the observed clusters (Figure 2.5a and 2.5c) even with *ModPoisson_a* that identifies the hotspots (Figure 2.5d). For both models, the residual K and G functions (explained in Section 1.1.3.2) show a specific trend that is far from zero for short distances. Indeed, these diagnostic tools show that the use of the inhomogeneous Poisson point process model is inappropriate since there is an unexplained interaction between points (Figure 2.6), that reflects a positive association between the events. Reminding that the residual K and G functions are related to score test for model comparisons, the current Poisson models may be improved adding some components accounting for interaction, such as a Strauss component or a Geyer one. Therefore it could be interesting to consider more complex

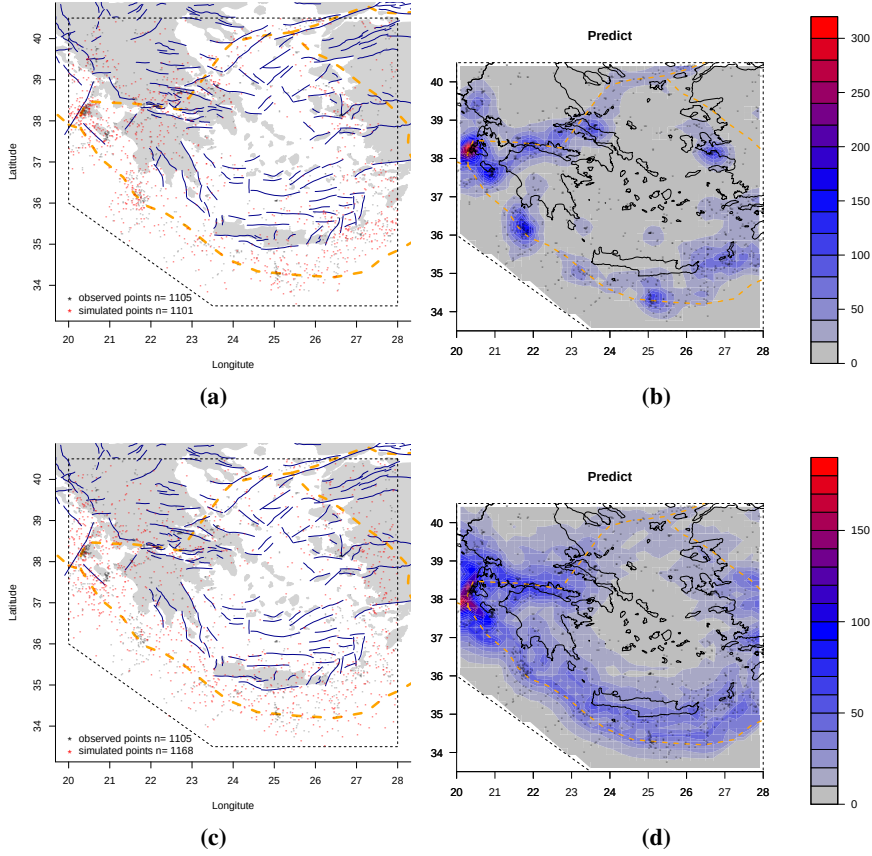


Figure 2.5: Plot of observed and simulated events (2.5a) and estimated spatial trend (2.5b) for the model $ModPoisson_a$. Plot of observed and simulated events (2.5c) and estimated spatial trend (2.5d) for the model $ModPoisson_b$

models, like hybrids, that account for the interaction between points, also at multiple scales, more properly.

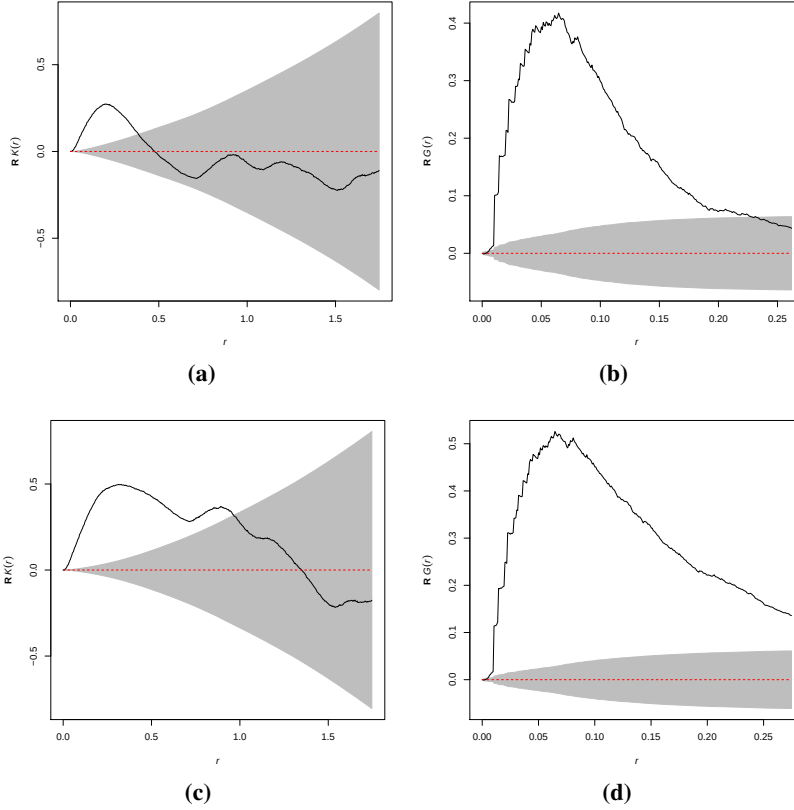


Figure 2.6: The residual K and G functions for the inhomogeneous Poisson process models, with kernel-based spatial trend $ModPoisson_a$ (a-b), and second-order polynomial $ModPoisson_b$ (c-d) in the coordinates.

2.4.2 Inhomogeneous hybrid models

We then considered hybrid models introduced in Section 2.1 in order to better explain the observed interpoint interaction. The challenge here is to choose a combination of Gibbs process components that explains well the spatial observed pattern. Hybrid models are fitted incrementally adding one component density at a time. The form of the residual K or G functions helps in identifying the appropriate Gibbs component. For both models (*ModPoisson_a* and *ModPoisson_b*), the residual G function has a specific trend that is far from zero for short distances with higher values for r between 0.05 and 0.1 (around 5.6 and 10.1 km). We thus tried several combinations of Strauss (1.7) and Geyer (1.8) processes selecting the irregular parameters (η) by maximum profile pseudolikelihood. We searched the interaction radius in the range $0.01 \leq r \leq 2$, and for the Geyer process the saturation parameter varied from 1 to 20, $s = 1, 1.5, \dots, 19.5, 20$. To consider the inhomogeneity degree, we estimated two models taking into account the spatial covariates, one with a non-parametric kernel-based trend in the coordinates (*ModHybrid_a*) and the other one with a quadratic polynomial trend with respect to the coordinates of the points (*ModHybrid_b*). A number of nested models were compared in terms of the residual deviance. Table 2.5 shows the irregular parameters (η) for the best selected models, and Tables 2.6 and 2.7 show the corresponding estimated regular parameters θ .

ModHybrid_a represents a hybrid process based on multitype Geyer processes (four components) and a Strauss process which describes adequately the multi-scale interactions between the events. For this model the estimated irregular parameters (η) are nine as shown in Table 2.5. Moreover the parametric vector θ in (1.10) has six elements, the intercept term (θ_1) and

Table 2.5: Estimated irregular parameters for the selected hybrid models: *ModHybrid_a* and *ModHybrid_b*

Gibbs component		Irregular paramters					
		<i>ModHybrid_a</i>			<i>ModHybrid_b</i>		
		r_j		s_j	r_j		s_j
G_1	Geyer	0.04	(4.45 km)	11	0.05	(5.56km)	15.5
G_2	Geyer	0.05	(5.56 km)	4.5	0.06	(6.67km)	1.5
G_3	Geyer	0.06	(6.67 km)	2.5	0.09	(10.00km)	11.5
G_4	Geyer	0.08	(8.90 km)	1.5	0.15	(16.68km)	3.5
S	Strauss	0.66	(73.39 km)				

Table 2.6: *ModHybrid_a* - Estimated parameters for the Hybrid model with a kernel term for the spatial trend in x and y .

	Estimate	γ_j		Zval
(Intercept)	-0.2858		***	-3.62
G_1	0.0835	1.09	***	8.24
G_2	0.1523	1.16	***	4.27
G_3	0.1835	1.20	**	2.92
G_4	0.3004	1.35	***	4.13
S	-0.0112	0.97	***	-9.25
AIC	-7371.777			
N^o of simulated points	1191			
Range of raw residuals	[-1.63;2.03]			

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

five interaction parameters (θ_2), one for each process in the hybrid model. For the Geyer processes, the interaction parameters are greater than 1, confirming a multi-scale clustering dependence between points for interaction distances less or equal than 8.9 km. For a larger interaction radius, the events show a CSR behavior since the γ parameter associated to the Strauss process is not significantly different from 1. Using a hybrid model with a non-parametric kernel-based term in x and y and smoothing parameters estimated by FLP, the covariates are not significant.

Table 2.7: *ModHybrid_b* - Estimated parameters for the Hybrid model with a second-order polynomial trend in x and y

	Estimate	γ_j		Zval
(Intercept)	8.7863			0.45
G_1	0.0841	1.09	***	12.06
G_2	0.4627	1.59	***	8.06
G_3	0.1193	1.13	***	12.22
G_4	0.1968	1.22	***	6.51
D_{pb}	-0.2656		***	-6.25
$I(D_f < \phi_1)D_f$	-1.0196		***	-4.15
$I(\phi_1 \leq D_f < \phi_2)D_f$	-0.4963		*	-2.38
$I(\phi_2 \leq D_f < \phi_3)D_f$	-0.5362		***	-4.39
$I(\phi_3 \leq D_f < \phi_4)D_f$	-1.1964		*	-2.54
$I(D_f \geq \phi_4)D_f$	-1.0059		***	-4.97
x	-1.1664		*	-2.42
y	0.5004			0.63
x^2	0.0273		***	5.48
xy	-0.0045			-0.55
y^2	-0.0065			-0.75
AIC	-7121.841			
N^o of simulated points	1475			
Range of raw residuals	[-1.62;2.29]			

Note: *p<0.1; **p<0.05; ***p<0.01

On the other hand, *ModHybrid_b* represents a model based on a hybrid of four Geyer components that have an interaction parameter greater than 1. For this model, the total number of estimated irregular parameters (η) are eight and the parametric vector (θ) has eleven parameters to explain the trend part and four interaction effects one for each Geyer process. In this case, the variables D_f and D_{pb} are significant and both of them are negatively related to the log spatial intensity.

More in the detail, increasing the distance from the plate boundary to 50 km ($r=0.4496$), the spatial conditional intensity decreases by 1.13. In an area around the faults less than 50 km, the log intensity decreases by 0.046

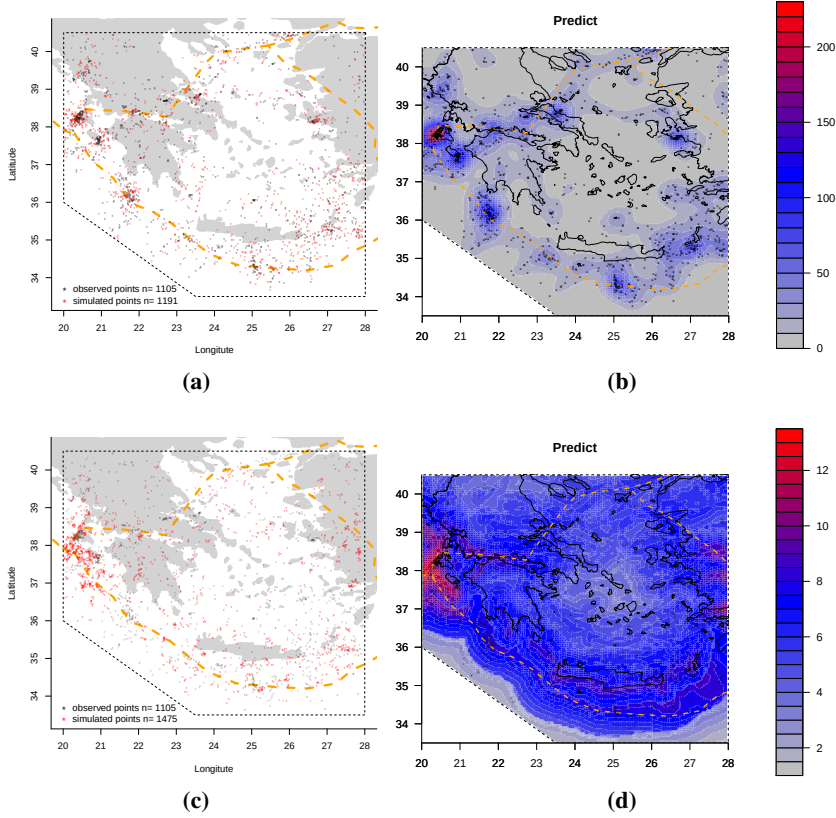


Figure 2.7: Plot of observed and simulated points (2.7a) and estimated spatial trend (2.7b) for the model *ModHybrid_a*. Plot of observed and simulated points (2.7c) and estimated spatial trend (2.7d) for the model *ModHybrid_b*

increasing the distance to the nearest fault by 5 km. The trend for the intensity is decreasing for an area that has points with a distance greater than 50 km to the faults; indeed, all the parameters for the piecewise linear trend with respect to D_f are negative and significant. Close to the faults, we have a more rapid reduction of the intensity than in the neighborhood of the plate boundary.

In both the hybrid models, we can observe that the number of simulated points is greater than the number of observed ones, especially for *ModHybrid_b*. However the higher number of simulated points affects marginally the residual K and G functions and their upper and lower limits. We notice that under the parametric formulation assuming both a Poisson and a hybrid process, the number of simulated points are greater than 1105. It is due to the fact that considering covariates information in the model all the different seismogenetic source of earthquakes can generate points, even if in reality only some of them had generated earthquakes.

Considering both hybrid models, we have an improvement in terms of AIC with respect to both models based on the inhomogeneous Poisson process. Indeed the AIC decreases by 988 for model *ModHybrid_a* and by 1874 for model *ModHybrid_b*. Moreover we observe a sensible reduction of the range of the spatial raw residuals for the model *ModHybrid_b* even though the number of simulated points is slightly larger than the number of events in the observed pattern.

Now the residual K and G functions, shown in Figure 2.8, oscillate around zero and they are inside the envelopes indicating that the interaction structure between the earthquakes is well described by both models.

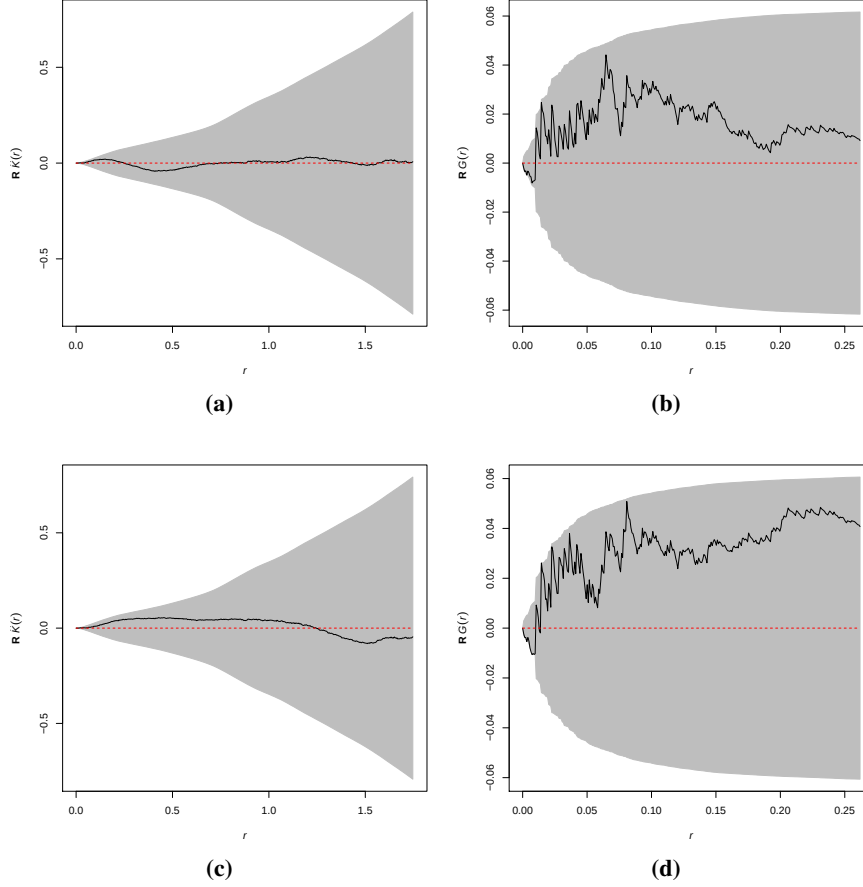


Figure 2.8: The residual K and G-functions for the hybrid models with kernel-based spatial trend $ModHybrid_a$ (a-b), and second-order polynomial (c-d) in the coordinates, $ModHybrid_b$.

2.5 Final remarks

After considering simple inhomogeneous Poisson models (too simple to provide a good fit), we found two suitable inhomogeneous hybrid models that are good in terms of AIC, of range of raw residuals, and of residual K and G-functions. The first hybrid model (*ModHybrid_a*) has a kernel smoothed trend in x and y with bandwidth values estimated with the *FLP* (Forward likelihood prediction) method. On the other hand, we also fitted a hybrid parametric model with a second-order spatial polynomial trend and spatial covariates (*ModHybrid_b*). We found that the variable that indicates the distance to the nearest volcano is not significant in explaining the spatial intensity, in fact the volcanic Hellenic arc area is mostly characterised by microseismic activity. Increasing the distance to the plate boundary and to the fault, the conditional intensity decreases, and this reduction is greater if we are close to faults. As for the interaction aspect between points, both models agree in identifying a multi-scale clustering between points for interpoint distances approximately less than 10 km.

The estimated spatial trends for the two models are significantly different because the first model reproduces exactly the observed trend and hotspots, while the second model gives a smoother information for all the study area. To further compare the models, we computed image plots and boxplots of the raw residuals, that are shown in Figure 2.9 and Figure 2.10, respectively. According to these results, moving from Poisson to hybrid models we have improvements in terms of range, variability and reduction of specific spatial trend in the raw residuals. Comparing the two final models, *ModHybrid_a* seems slightly better than *ModHybrid_b* since it has a higher AIC. Moreover the spatial raw residuals of *ModHybrid_a* have a smoother spatial trend (Figure 2.9 (c)) compared to what happens in Figure 2.9(d), in fact *ModHybrid_b*

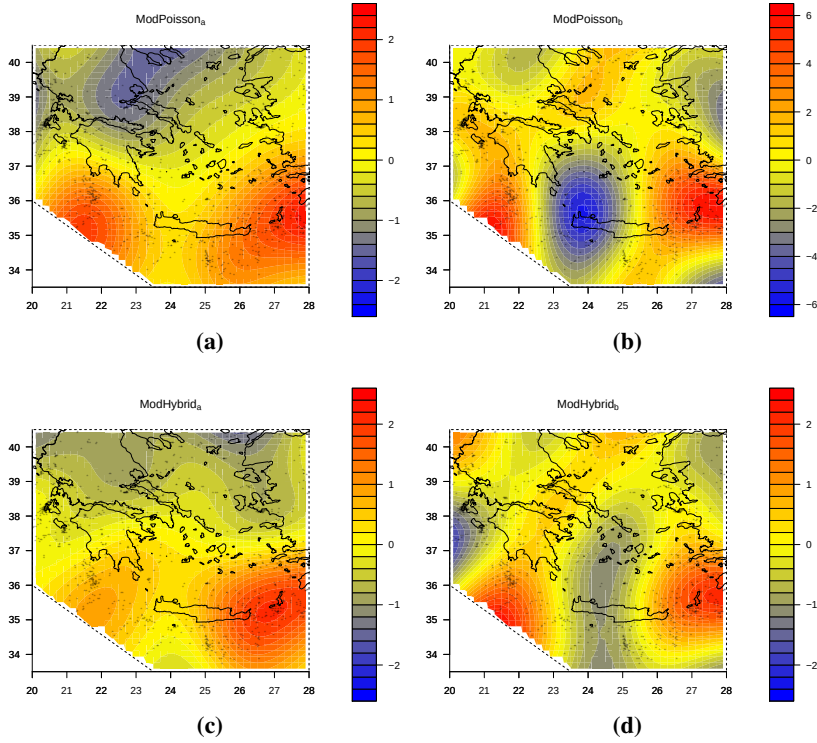


Figure 2.9: Plot of smoothed raw residuals for $ModPoisson_a$ (a), $ModPoisson_b$ (b), $ModHybrid_a$ (c) and $ModHybrid_b$ (d).

has a few number of residuals with high values.

Summing up, if the research interest focuses on the description of the spatial intensity of the phenomena and therefore on the study of the interaction between earthquake events, the model based on a non-parametric spatial trend estimated by FLP seems more suitable. On the other hand, if the main interest is to study and to understand how the several seismogenic sources of earthquakes influence the spatial intensity of the phenomena, a full parametric model would be more appropriate.

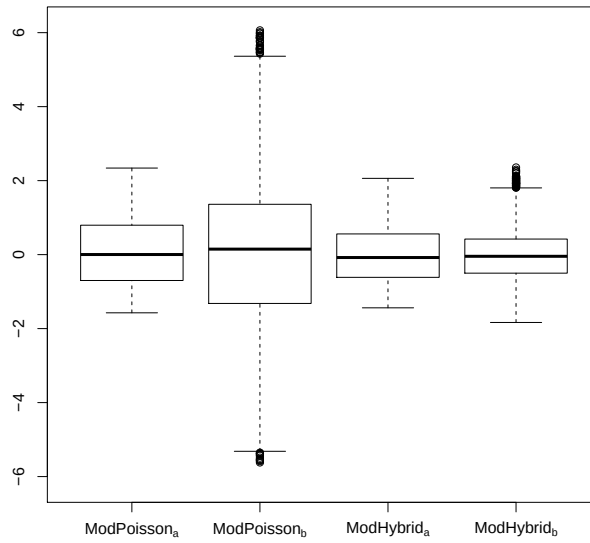


Figure 2.10: Boxplots of smoothed raw residuals for $ModPoisson_a$ (a), $ModPoisson_b$ (b), $ModHybrid_a$ (c) and $ModHybrid_b$ (d).

Chapter 3

Spatio-temporal log-Gaussian Cox processes on earthquake events

In this work, we formulate some spatio-temporal log-Gaussian Cox models for the description of earthquake events in the Alaska Peninsula. We consider spatial geological information in the study area and different specification of the covariance structure of the Gaussian Random Field (GRF). More exactly, we detect the presence of spatial geometric anisotropy and describe the interaction features between events assuming a separable and non-separable spatio-temporal covariance structure. In the Section 3.1, there is a literature review of the main works on spatio-temporal log-Gaussian Cox processes (LGCPs). In Section 3.2, the study area and the data are described. The main results are in Section 3.3, where we report the results obtained fitting inhomogeneous spatio-temporal Poisson models. Since the inhomogeneous spatio-temporal pair correlation function presents cluster

structure for short distances, LGCP models are considered to describe the interaction structure between events. The parameters that characterise the covariance function of the GRF are estimated using the method of moments taking into account the spatial anisotropy. This application represents a first attempt in the application of LGCP models on earthquake data using covariate information. The presented results are interim and partial and they will need further developments, insights and considerations. For this reason in Section 3.4, the possible analysis, that should be done to improve the obtained results and to have more reliable conclusions, are outlined.

3.1 Literature review

Cox processes are natural models for point process phenomena that are environmentally driven and Møller *et al.* (1998) introduce the class of log-Gaussian processes (LGCPs). They are a flexible and relatively tractable class of models for describing spatially and spatio-temporal correlated phenomena specifying the moments of an underlying Gaussian Random Field (GRF). The intensity has the following form $\Lambda(\mathbf{u}_i, t_i) = \exp\{S(\mathbf{u}_i, t_i)\}$ where S is a Gaussian process with mean $\mathbb{E}(S(\mathbf{u}_i, t_i)) = \mu$ and variance covariance matrix $\mathbb{C}(S(\mathbf{u}_i, t_i), S(\mathbf{u}_j, t_j)) = \mathbb{C}(\|\mathbf{u}_i - \mathbf{u}_j\|, |t_i - t_j|) = \sigma^2 \gamma(r, h)$. All the theoretical framework has been introduced in Section 1.2.4.

Usually Cox processes do not focus on the direct modelling of interactions between events since they have an historical independent formulation. In presence of spatial and temporal trends, Cox processes model can describe clustered patterns by ascribing local spatio-temporal concentrations of events to peaks in the stochastic process. So, the main open research aspects are in the direction of the specification of either the non-stationary

aspect and the covariance function of the GRF. Indeed in this dissertation, we investigate both aspects, introducing spatial covariates accounting for non-stationary models, and analysing second-order non-separability and anisotropy of the covariance function.

There are other several open research aspects in LGCP model for the definition of: spatio-temporal LGCP model for surveillance setting, complex covariance structure, anisotropy, conditional specification of the intensity to the past history and proper covariance functions for circular time.

Spatio-temporal LGCP models to analyse real data

The first main spatio-temporal model specification is in Brix and Diggle (2001) where they describe a flexible class of space-time LGCP model whose stochastic intensity is a space-time Ornstein-Uhlenbeck process that is estimated with a moment-based method and the underlying intensity is predicted by using a Markov chain Monte Carlo approach. In Brix and Møller (2001), they extend the previous model proposing a space-time point process based on a bivariate log-Gaussian Cox birth process for multi-type point patterns. On the other hand, in Diggle *et al.* (2005), they consider non-stationary version for an on-line spatio-temporal surveillance for a gastroenteric disease for the Hampshire county in UK.

Most of the spatio-temporal applications are for patterns that show clustering behaviour, such as infectious diseases data. Usually, it is analysed the role of spatially and temporally covariates and often the models are applied to a surveillance system for detecting emergent spatio-temporal clusters of events. The general aims are to estimate the risk map of the studied phenomena and to study the dependence on covariates. For example in Beneš *et al.* (2011), they investigate the association between tick-borne encephalitis

and spatially varying covariates of population census data, vegetation and altitude in Bohemia. In Ahn *et al.* (2014), they consider a Bayesian two-stage space-time point process model based on LGCP model for diarrheal disease in northwestern Ecuador. Considered in the model spatially and temporally referenced covariates (such as precipitation and temperature), they estimate the risk-related parameters and conduct predictive inference considering a disease mapping model. As for wildfire data, Serra *et al.* (2014) estimate a spatio-temporal LGCP obtaining risk map that can contribute to the prevention and management of those wildfires which are not random in space or time. Moreover, Giorgi *et al.* (2016) study the relationship of plague with temperature and precipitation anomalies, and with elevation through a spatio-temporal analysis of plague incidence in Madagascar from 1980 to 2007.

Also in the spatial context, an interesting contribution is in Liang *et al.* (2008) where they propose a marked LGCP for differentiating cancer types incorporating both spatial covariates (local poverty rate and location within the metro area) and non-spatial individual level covariates (patient age and cancer stage).

Complex covariance structure for the undelying GRF

Nevertheless, most of the previous models account only for a separable covariance function for the Gaussian Random Field. Although separability is a convenient working assumption, it may be too restrictive for some applications. So other parametric and non-parametric families for the covariance function such as a non-separable structure may be considered (Gneiting, 2002; Rodrigues and Diggle, 2010). In particular, Rodrigues and Diggle (2012) use a class of low-rank convolution-based models (Rodrigues and

Diggle, 2010) that have computationally efficient construction, in spatio-temporal LGCP models for detecting emergent spatio-temporal clusters of homicides in Belo Horizonte, Brazil.

LGCP models that account for anisotropy

Another possible characterisation can be in describing the directional behaviour of the points. For example in Møller and Toftaker (2014), they consider spatial point processes with an elliptical (and not spherical) pair correlation, so they model a geometric anisotropy in the framework of Cox processes.

Specify an intensity that depends on the past history

The formulation of the LGCP intensity may have a conditional specification on the past history. As explained in Diggle (2013), the LGCP is classified as a deterministic model since usually it does not incorporate any mechanistic law. However, it may be of interest to apply both contagion models and history independent correlation models to disentangle contagion from other types of correlation to describe the observed clustering structure. In Mohler (2013), there is an interesting model formulation on the line, they specify a mechanistic contagious (Hawkes) model with a background rate driven by a log-Gaussian Cox process describing the presence of clustering in time series of crime and security data.

Spatio-temporal LGCP for circular time data

Furthermore, another relevant aspect in model formulation concerns in how the time dimension is treated. In all of the aforementioned spatio-temporal model specifications, the time is defined for positive real values, however in

some applications it is treated in a circular way (i.e. it may be of interest the time of occurrence in terms of day of the week or hour in crime data analysis). Shiota and Gelfand (2016) develop models to accommodate the time as a circular variable in the analysis of crime events. The key aspect for defining a LGCP consists in the definition of a spatio-temporal covariance function for the underlying Gaussian Random field, so they define valid separable and non-separable covariance functions over a bounded spatial region crossed with circular time.

Available R packages

Some R packages implement routines to simulate and fit LGCPs. In particular, the package `stpp` (Gabriel *et al.*, 2013) implements code to simulate spatio-temporal LGCP with a separable and non-separable covariance structure for the Gaussian Random Field. Instead the package `lgcp` (Taylor *et al.*, 2013, 2015) implements code to fit LGCP models using methods of the moments and also a Bayesian inference for spatial, spatio-temporal, multivariate and aggregated point processes. Although we mainly used the available packages, they do not handle for non-separable and anisotropic correlation structures of the covariance structure of the Gaussian Random Field. In `lgcp` package, the minimum contrast method is used for the estimation of parameters assuming a separable structure of the covariance of the GRF. Then, new code has been implemented and used to account for these aspects. Davies and Hazelton (2013) provide a set of simulation studies to assess the performance of minimum contrast methods for parameter estimation for spatial and spatio-temporal LGCP models.

3.2 The study area and the dataset

Alaska is the most seismically active state in the USA country. We focus our analysis to the Alaska Peninsula, that is the north-western protrusion of the North American continent extending about 900 km from the mainland of the country and ending in the Aleutian Islands. It separates the Pacific Ocean (in the South) and the Bristol Bay in the Bering Sea that is in the North. In the area, a highly active volcanic mountain range runs along the entire length of the Peninsula, that is called Aleutian Range and it is a part of the Aleutian Arc. Most of the earthquakes and all major earthquakes in the area are due to the movement of tectonic plates of the Pacific Ocean towards the North American Plate with a subduction process. Quaternary deformation is the result of the oblique subduction of the Pacific plate on the North American plate. The largest earthquakes, occur primarily in the shallow part of the Alaska-Aleutian subduction zone, such as the 1938 M8.3 Alaska Peninsula earthquake, 1946 M7.8 Unimak earthquake, 1957 M8.6 Fox Islands earthquake, 1964 M9.2 Good Friday earthquake, and 1965 M8.7 R at Islands earthquake.

We focus on the Alaska Peninsula selecting an area with latitude ranging between 48 and 60 and longitude between -180 and -150. In the analysis, we project all the WGS84 georeferenced information using NAD83/Alaska Albers (EPSG:3338) projection coordinate system. The study area, properly rescaled, is in Figure 3.1a, that reports the lineament of the Quaternary faults (Koehler, 2013) and the main volcanoes with red points.

The Alaska Earthquake Center monitors earthquakes in the area and provides the catalogue of events. We consider earthquakes occurred between 2004 and 2006, with a magnitude greater or equal to 2.5. For a long-term analysis, the time of occurrence is transformed in a discrete time resolution,

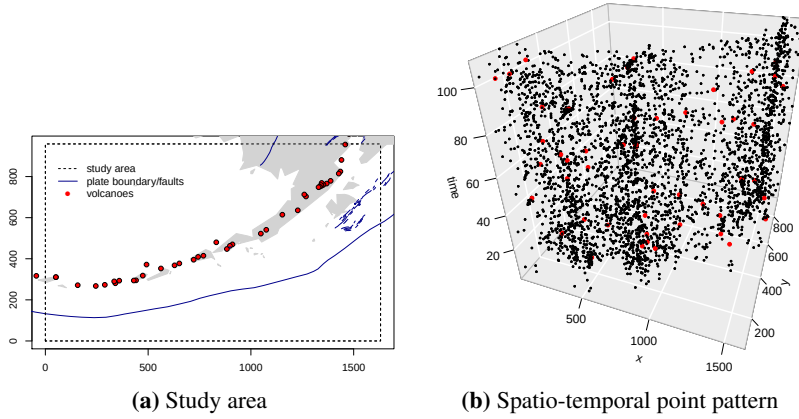


Figure 3.1: (3.1a) The study area and the geological information such as the volcanoes and the Quaternary faults. (3.1b) Plot of the spatio-temporal pattern, red points are events with a magnitude greater than 4.5.

a generic t refers to 10-days. The mean number of observed points for each t is 34.35, for a total number of 110 time periods. Figure 3.3b shows the events and the red points are those with a magnitude greater than 4.5.

With the following analysis, we aim to address some research questions such as: which are the main determinants of the observed spatial inhomogeneity? Is there an orientation in the point pattern? Which are the parameters that characterised the spatial geometric anisotropy? How do the results change taking into account the anisotropy? Which covariance functions (of the Gaussian Random Field) do better describe the second-order interaction structure between events?

3.3 Results

The main steps in the estimation procedure are outlined.

- The first step concerns the estimation of a separable spatio-temporal intensity of an inhomogeneous Poisson process that has a log-linear form, $\lambda(\mathbf{u}, t) = \lambda(\mathbf{u})\lambda(t) = \exp(\beta_u Z_s(\mathbf{u}))\exp(\beta_t Z_t(t))$. The geological information (Quaternary faults and volcanoes) are introduced in the model as distance to the nearest fault (D_f) and distance to the nearest volcano (D_v). We introduce these variables linearly since it is expected that the intensity decreases as the distance to the seismological source increases.

Following Diggle (2013), we estimate separately the spatial and temporal trend, $\lambda(\mathbf{u}) = \exp(\beta_0 + \beta_1 D_f(\mathbf{u}) + \beta_2 D_v(\mathbf{u}) + \log(\text{offset}(k_{\delta_s}(\mathbf{u})))$ and $\lambda(t)$ is estimated considering a negative binomial model for count data based on kernel. A two dimensional Gaussian kernel estimator with bandwidth is used for the spatial trend and a one-dimensional Gaussian kernel intensity estimator is used for the temporal term. Both the spatial variables turn out to be significant and $\hat{\beta}_s = \{1.53, -0.0063, -0.0084\}$. Figure 3.2a and 3.2a illustrate the estimates of the spatial (left) and temporal (right) intensity, considering the product between the spatial and temporal part. The spatial quantity is rescaled such that $\int_W \lambda(\mathbf{u}) du = 1$.

- In the second step, the presence of anisotropy is detected assuming that it is constant over time. Considering the 3775 events in a collection of 110 screenshots, for each t , the orientation of the data is estimated according to the procedure proposed in Møller and Tof-taker (2014) and described in Section 1.2.4.3. Figure 3.3a shows the main orientation of the data at each time t (θ_i) and the blue arrow is the estimated mean ($\bar{\theta} = 0.34$) (treating properly the data as ax-

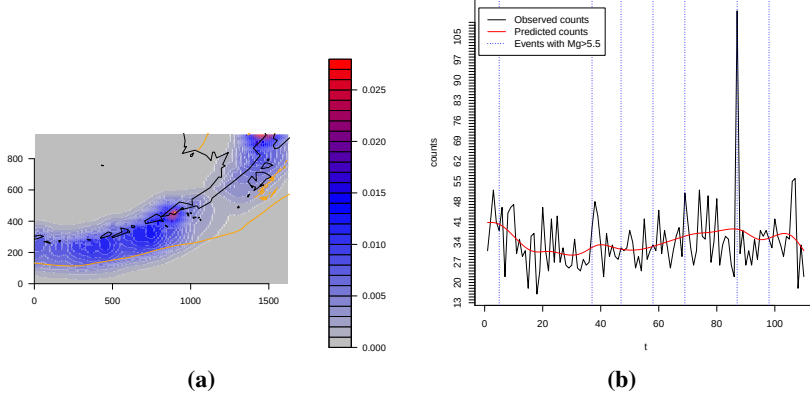


Figure 3.2: (3.2a) Estimated spatial intensity $\hat{\lambda}(\mathbf{u})$. (3.2b) The red line is $\hat{\lambda}(t)$ and the black line represents the observed number of events for each time t .

ial information). Figure 3.3b shows the boxplot of the parameters ζ_i . Their mean ($\bar{\zeta} = 0.51$) indicates the ratio between the major and the minor axes in an elliptical anisotropic structure (Figure 3.3c). In the next step, the spatial window (W) and the generic observation ($\mathbf{u}_i$) are transformed in the following way

$$\mathbf{v}_i = \mathbf{u}_i^T \begin{bmatrix} \cos \bar{\theta} & -\sin \bar{\theta} \\ \sin \bar{\theta} & \cos \bar{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\bar{\zeta} \end{bmatrix}$$

$$V = W \begin{bmatrix} \cos \bar{\theta} & -\sin \bar{\theta} \\ \sin \bar{\theta} & \cos \bar{\theta} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/\bar{\zeta} \end{bmatrix}$$

and the estimation of the spatial intensity $\lambda_A(\mathbf{v})$ is in Figure 3.3d, with estimated parameters $\hat{\beta}_s = \{1.30, -0.003, -0.004\}$.

We continuous the analysis using two set of events, the original data ($\mathbf{u}_i \in W \subset R^2$) and the transformed data according to $\bar{\theta}$ and $\bar{\zeta}$,

$\mathbf{v}_i \in V \subset \mathbb{R}^2$. Basically, we want to assess if the two inhomogeneous Poisson models estimated on $\{\mathbf{u}_i\}_{i=1}^n$ and $\{\mathbf{v}_i\}_{i=1}^n$, respectively $\hat{\lambda}(\mathbf{u})\hat{\lambda}_t(t)$ and $\hat{\lambda}_A(\mathbf{v})\hat{\lambda}_t(t)$ (with the spatial term properly rescaled such that it integrates to 1), explain the observed data.

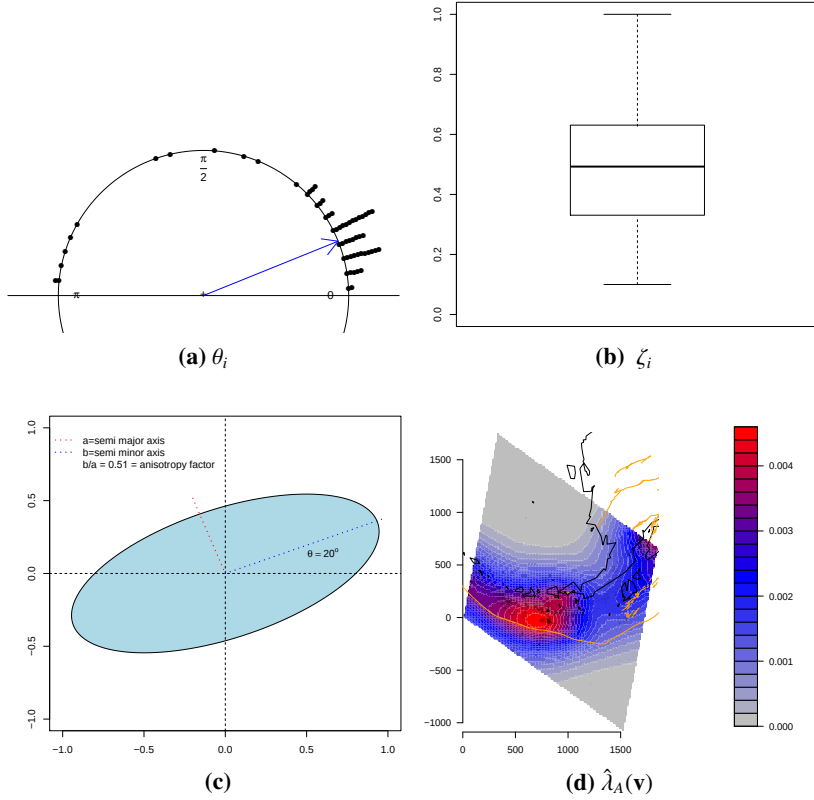


Figure 3.3: (3.3a) Semi-circular data plot of the orientation parameters θ_i for each time t . (3.3b) Boxplot of the ζ_i that represent the ratio between the semi-minor axis and the semi-major axis. (3.3c) Oriented ellipse according to $\bar{\theta}$ and $\bar{\zeta}$ (3.3d) Estimated spatial intensity, $\hat{\lambda}_A(\mathbf{v})$ rotating the coordinates and the window according to $\bar{\theta}$ and $\bar{\zeta}$

- For both the models, the inhomogeneous pair correlation function ($\hat{g}(r, h)$) is estimated considering the estimator in Equation (1.18). These descriptive summary statistics give a general indication if the fitted models describe well the observed data. Both the estimated inhomogeneous pair correlation functions show clustering structure for some spatio-temporal distances, Figure 3.4a and Figure 3.4b. Indeed, the surfaces are not close to 1, that is the theoretical surface under the Poisson assumption, and they give a indication of spatio-temporal clustering for $r \leq 50$ Km and $h \leq 10$ (100 days). However, comparing $\hat{g}(r, h)$ for the original point pattern and for the rotated one, the inhomogeneous pair correlation function has smaller values for the latter case.

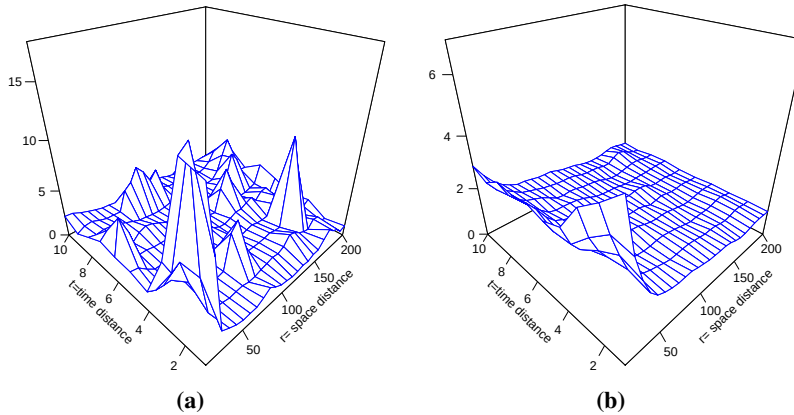


Figure 3.4: (3.4a) Estimated inhomogeneous $\hat{g}(r, h)$ based on the estimated intensity $\hat{\lambda}(\mathbf{u})\hat{\lambda}_t(t)$ (3.4b) Estimated inhomogeneous $\hat{g}(r, h)$ based on the estimated intensity $\hat{\lambda}_A(\mathbf{v})\hat{\lambda}_t(t)$

- Since there is an unexplained clustering structure, more complex models than the inhomogenous Poisson one are considered. We as-

sume LGCP models with the parametrisation indicated in (1.19), $\Lambda(\mathbf{u}_i, t_i) = \lambda(\mathbf{u}_i, t_i) \exp\{S(\mathbf{u}_i, t_i)\}$ and the assumption in (1.21). The parameters of the covariance function of the Gaussian Random Field (ϕ) are estimated using the method of moments, (1.26), that depends on f and $\hat{f}$.

We remember the theoretical relationship in Equation (1.20), that the theoretical pair correlation function of the LGCP is equal to the exponential of the covariance function of the GRF. In our case, the theoretical quantity f is equal to the theoretical pair correlation function of a LGCP, $g((\mathbf{u}_i, t_i); (\mathbf{u}_j, t_j)) = \exp(\sigma^2 \gamma(r, h)) = \exp(C(r, h))$. The quantity $\hat{f}$ is the non-parametric estimation of the inhomogenous spatio-temporal pair correlation function using the estimator in Equation (1.18). Three parametric formulations are considered for $f(r, h)$. In particular, it is considered a separable spatio-temporal exponential covariance function (1.22), a separable covariance function with a Matérn component for the spatial term and a exponential component for the temporal one (1.23), and a non-separable Gneiting covariance function (1.25). Minimising the discrepancy measure in (1.26), $\hat{\phi}$ is estimated settings $w(r, h) = 1$, $c=1$ and $(r_0, h_0) = (200, 10)$.

Table 3.1 summarises the estimated parameters under the isotropic and the anisotropic context. We remind that in the latter case, the parameters that explain the anisotropy have been estimated in the previous step $(\bar{\theta}, \bar{\zeta})$ and the usual procedure for an isotropic point patter is applied on the rotated point pattern.

In Figure 3.5a and Figure 3.5b, the estimated spatio-temporal pair correlation function with respect to $\hat{\lambda}(\mathbf{u})\hat{\lambda}_t(t)$ is plotted with a blue surface. Figure 3.5a shows also the estimated parametric surfaces

Covariance function	Parameters ϕ in $\exp(\text{Cov}(u, v))$	Estimates under the ISOTROPIC assumption	Estimates under ANISOTROPY
(1)	$(\sigma^2, \phi_s, \phi_t)$	(7.72, 8.43, 2.70)	(5.76, 12.15, 1.88)
(2)	$(\sigma^2, \phi_s, \phi_t, \tau)$	(10.81, 34.78, 3.84, 0.11)	(9.41, 30.70, 3.28, 0.10)
(3)	$(\sigma^2, \alpha, \delta, z)$	(9.39, 11.01, 1, 0.49)	(5.43, 17.77, 0.75, 0.27)

Table 3.1: Estimated ϕ with the method of the moments assuming that the covariance function of the Gaussian Random field has a (1) Exponential-Exponential, (2) Exponential-Matérn and (3) Gneiting structure. In the anisotropic case the usual procedure is applied to the rotated point pattern.

under the separable assumption; the red surface is the estimated parametric $g(h, r)$ based on Exponential-Exponential covariance of the GRF, instead yellow surface is the estimated parametric $g(h, r)$ based on Exponential-Matérn covariance of the GRF.

Figure 3.5b shows the estimated parametric surface of $g(h, r)$ under the non-separable assumption; the violet surface is the estimated parametric $g(h, r)$ based on the Gneiting covariance of the GRF.

On the other hand, the Figure 3.6a and 3.6b are analogous to the previous two plots but they are computed under the anisotropic assumption.

For both the isotropic and anisotropic assumptions, we have a better description of $\exp(C(r, h))$ with Gneiting covariance function (1.25), and it seems that the assumption of second order separability is too restrictive. In fact for both the cases, the space-time interaction parameter δ is close to 1.

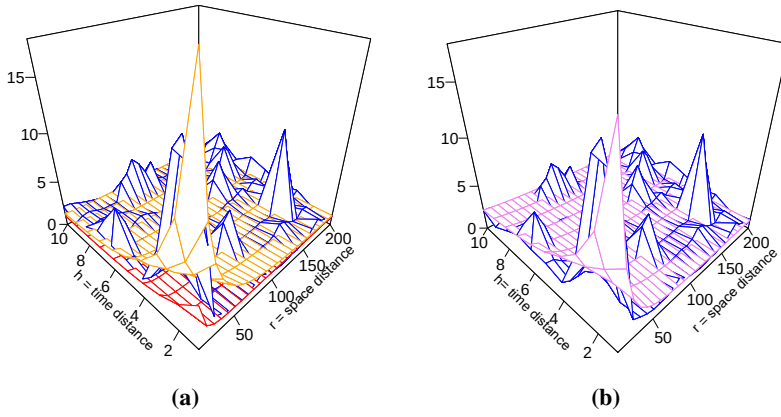


Figure 3.5: Analysis under the isotropic assumption. (3.5a) The blue surface is the inhomogenous pair correlation function with respect to $\hat{\lambda}(\mathbf{u})\hat{\lambda}_t(t)$. The red surface is the estimated parametric $g(h, r)$ based on Exponential-Exponential covariance of the GRF. The yellow surface is the estimated parametric $g(h, r)$ based on Exponential-Matérn covariance of the GRF. (3.5b) The blue surface is the inhomogenous pair correlation function with respect to $\hat{\lambda}(\mathbf{u})\hat{\lambda}_t(t)$. The violet surface is the estimated parametric $g(h, r)$ based on Gneiting covariance of the GRF.

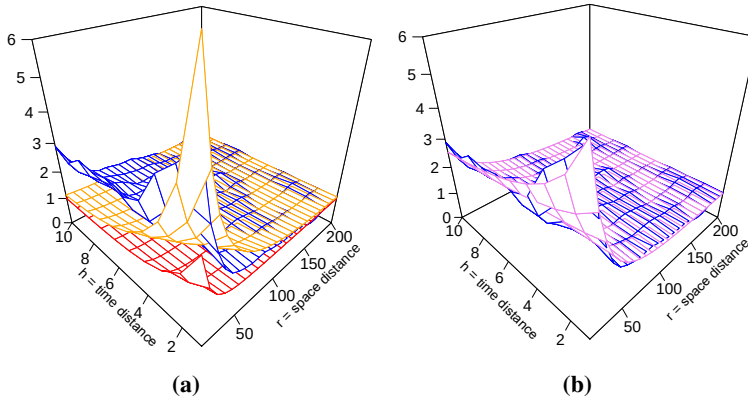


Figure 3.6: Analysis under the anisotropic assumption. (3.6a) The blue surface is the inhomogeneous pair correlation function with respect to $\hat{\lambda}_A(\mathbf{v})\hat{\lambda}_t(t)$. The red surface is the estimated parametric $g(h, r)$ based on Exponential-Exponential covariance of the GRF. The yellow surface is the estimated parametric $g(h, r)$ based on Exponential-Matérn covariance of the GRF. (3.6b) The blue surface is the inhomogeneous pair correlation function with respect to $\hat{\lambda}_A(\mathbf{v})\hat{\lambda}_t(t)$. The violet surface is the estimated parametric $g(h, r)$ based on Gneiting covariance of the GRF.

3.4 Some conclusions and future analysis

Summing up, both distance to the nearest faults and distance to the nearest volcano are relevant covariates to explain the spatial intensity. There is a prevalent orientation in the data; considering properly the anisotropy in the model formulation we improve the results: the residual pair correlation function has smaller values. In both cases, the parameter δ that indicates the strength of separability, in the Gneiting covariance function is close to 1. Furthermore, it seems that there is a better fitting to the inhomogeneous pair correlation function when we assume a non-separable covariance rather than a separable structure.

Starting from a new estimation approach for more general LGCP models, the previous results can be seen as a descriptive attempt to capture clustering behaviour and anisotropy in spatio-temporal data. However, with the previous analysis, we are not able to identify and select the best model. Further analysis are indeed necessary to obtain more reliable conclusions. For instance, after estimating the inhomogeneous Poisson model, it would be interesting to assess the presence of clustering and of interaction structure between space and time using the testing procedures proposed in Tamayo-Uria *et al.* (2014).

To select the final LGCP model, more formal model checking can be proposed by using other second-order summary statistics than that one used for estimating the parameters. Moreover, on the transformed data it should be checked by proper test whether the anisotropy has been adequately explained by the model.

Chapter 4

Testing procedure for local spatio-temporal structure

In the following chapter, we present a brief literature review on the local indicator of spatial association and testing procedures to detect clusters in both the spatial and spatio-temporal contexts (section 4.1). In section 4.2, we present an unbiased estimator of Local Indicator of Spatio Temporal Association (here after LISTA) based on the global spatio-temporal product density (1.16), (Rodríguez-Cortés, 2014; Siino *et al.*, 2016c). Also, the behaviour of LISTA surfaces is studied under several simulated point patterns making graphical comparisons between global and local second-order estimates. Furthermore, motivated by an application on the analysis of earthquake data, we develop a testing procedure for local structure comparing spatio-temporal point patterns based on LISTA functions (section 4.3) performing simulations to assess its properties (section 4.4). The results of the application of the test on earthquake data are shown in section 4.5.

4.1 Literature review

Spatio-temporal cluster analysis is a key aspect of the practical analysis of spatio-temporal point patterns. The use of global spatio-temporal second-order summary statistics is now a common practice. Global spatial and spatio-temporal second-order measures, such as the quantities introduced in Equation 1.14 are used to detect deviations from the Poisson assumption. The idea of considering individual contributions of a global estimator as a measure of clustering emerged in the mid 90th under the name of Local Indicators of Spatial Association (LISA) coined by Anselin (1995), and they have been used as an exploratory data tool to examine individual points in a point pattern in terms of how they relate to their neighbouring points. Generally, as explained in Anselin (1995), the individual contributions to a global statistic can be used to identify outlying components measuring the influence of each contribution to the global statistic, this is the case of the scatter plot based on the local Moran index (Anselin, 1993). On the other hand, the individual contributions, as given by the G statistics (Getis and Ord, 1992), can be used to test for specific local structure, such as testing for spatial association and hot spots detection in areal data.

Basically, the aforementioned local statistics are often used to analyse areal data. In the context of spatial point processes, Getis and Franklin (1987) introduce a local version of the K -function to show that trees exhibit different heterogeneity when examined at different scales of analysis. In Stoyan and Stoyan (1994), the notion of individual functions for the L and G statistics has been introduced. Cressie and Collins (2001a) propose a local product density function to identify points with similar neighbourhood structure at various scales developing theoretical properties, namely first- and second-order moments, of these functions. This local density is considered more

sensitive to identify different local structures and unusual points than the local K -function (Mateu *et al.*, 2010).

Applications of LISA functions range from detecting features in images with noise (Mateu *et al.*, 2007) to detection of disease clusters (Moraga and Montes, 2011). In Rodríguez-Cortés (2014), the author extends the concept of LISA function to the spatio-temporal point pattern context defining the LISTA functions presenting edge-corrected estimators, and developing their first theoretical moments. Moreover, motivated by an application on the analysis of earthquake data, we develop a testing procedure for local structure comparing spatio-temporal point patterns based on LISTA functions (that are described in Section 4.2)

Indeed, several tests have been proposed in the literature to analyse second-order properties of a set of points or for comparing point patterns. A classification of them can be done in accordance to what is postulated under the null hypothesis, for example tests for no disease clustering (Diggle *et al.*, 2007), absence of space-time interaction (Diggle *et al.*, 1995) and approximately independence of two point patterns (Allard *et al.*, 2001). Depending on the goal of the study, the testing procedure may be applied to evaluate the global evidence against the null hypothesis, or to conduct local tests identifying specific regions where the null hypothesis is not accepted. In the later case, the tests should depend on local measures of association.

The methodology for local cluster detection has been developed to find out whether there is a local tendency of events to group together throughout the region under study, and usually most of the applications are in the public health and epidemiological fields (Tango, 2010). For example, a disease cluster detection for point patterns allows to evaluate if cases are geographically randomly distributed (taking properly the population at risk

into account), or if significant clusters are due to the relationship with an underlying environmental risk factor or to an existing spatial correlation between points. The scan statistics are well established methodological tools to identify clusters and to assess their significance in case events or in comparisons of case-control data. They have been developed for both spatial (Kulldorff, 1997; Kulldorff and Nagarwalla, 1995) and spatio-temporal data (Kulldorff *et al.*, 1998, 2005; Kulldorff, 2001). Working with space-time data, the analysis may be retrospective (Kulldorff *et al.*, 1998; Assunção *et al.*, 2007) or also for prospective surveillance analysis of space-time clusters (Kulldorff, 2001; Assunção and Correa, 2009).

However, there are some differences between the results of significant local structure using scan methods and the results that may be obtained with tests based on the local test statistics in the context of point pattern analysis. With the scan method the whole study window is analysed considering several sizes and shapes of the scanning window and a p -value is associated to an area. Instead, with the latter approach the significant level is associated to each observed point (Tango, 2010; Moraga and Montes, 2011). Moreover the scan method identifies significant cluster structure, but using local second-order summary statistics, it is possible to summarise information on the local association in a wide sense. In fact, the standard tests for cluster detection are in some sense restrictive when the goal of the study is to detect differences in the local structure of a point pattern with respect to an underlying process, and when in the observed pattern the elements of aggregation and regularity are combined.

Driven by the analysis of earthquake events, in section 4.3, we define a non-parametric testing approach identifying significant local structure extending the test proposed by Moraga and Montes (2011) into a spatio-temporal

framework. A simulation study is performed to illustrate that the defined tests have the prescribed size and their behaviour is reasonable under several alternative assumptions (section 4.4). In the analysis of earthquakes in section 4.5, we aim to detect which triggered events have a significant different local cluster structure with respect to the underlying process, represented by the background events, linking the results with the geological information available in the study area.

4.2 Product density LISTA functions

Local indicators of spatio-temporal association (LISTA) are a set of functions which are individually associated with each one of the events of the point pattern, and they can provide information about the local behaviour of the pattern. For the spatial case, the operational definition of local indicator was introduced by Anselin (1995) and it has been extended as follows. Considering the spatio-temporal product density in (1.12), its local version is denoted by $\rho^{(2)i}(\cdot, \cdot)$. Then for fixed r and h , it holds that

$$\widehat{\rho^{(2)}_{\epsilon, \delta}(r, h)} = \frac{1}{n-1} \sum_{i=1}^n \widehat{\rho^{(2)i}_{\epsilon, \delta}(r, h)}, \quad (4.1)$$

where $\widehat{\rho^{(2)i}_{\epsilon, \delta}(r, h)}$ is given by

$$\widehat{\rho^{(2)i}_{\epsilon, \delta}(r, h)} = \frac{n-1}{4\pi r |B|} \sum_{j \neq i} \kappa_{\epsilon, \delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h), \quad (4.2)$$

with $r > \epsilon > 0$, $h > \delta > 0$. Any second-order spatio-temporal summary statistic, that satisfies the operational definition in equation (4.1), which means that the sum of spatio-temporal local indicator functions is propor-

tional to the global statistic, can be called a LISTA statistic.

An unbiased edge-corrected kernel-based estimator for the product density LISTA function (Rodríguez-Cortés, 2014) is given by

$$\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r, h) = \frac{n-1}{4\pi r|B|} \sum_{j \neq i} \frac{\kappa_{\epsilon,\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h)}{w(\mathbf{u}_i, \mathbf{u}_j)w(t_i, t_j)}, \quad (4.3)$$

with $r > \epsilon > 0$, $h > \delta > 0$, for $(\mathbf{u}_i, t_i) \in W \times T$, $i = 1, \dots, n$, $w(\mathbf{u}_i, \mathbf{u}_j)$ and $w(t_i, t_j)$ are the edge-effect factors. In space, $w(\mathbf{u}_i, \mathbf{u}_j)$ is the proportion of the circumference of a circle centred at $\mathbf{u}_i$ with radius $\|\mathbf{u}_i - \mathbf{u}_j\|$ lying in W . On the other hand, the temporal edge correction factor $w(t_i, t_j)$ is equal to 1, if both ends of the interval of length $2|t_i - t_j|$ centred at t_i are within T , and two otherwise. The kernel function κ has a multiplicative form $\kappa_{\epsilon\delta}(\|\mathbf{u}_i - \mathbf{u}_j\| - r, |t_i - t_j| - h) = \kappa_{1\epsilon}(\|\mathbf{u}_i - \mathbf{u}_j\| - r) \kappa_{2\delta}(|t_i - t_j| - h)$, where $\kappa_{1\epsilon}$ and $\kappa_{2\delta}$ are kernel functions with bandwidths ϵ and δ , respectively.

For a homogeneous spatio-temporal Poisson process with intensity ρ , the expectation of the product density LISTA function has a closed form, and it is given by

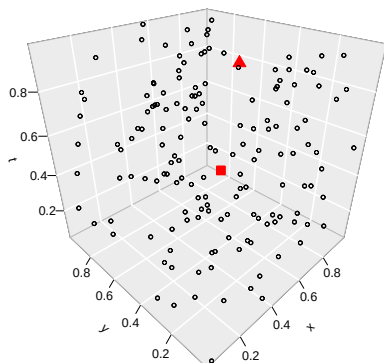
$$\mathbb{E}_! \left[\widehat{\rho^{(2)i}}_{\epsilon,\delta}(r, h) \right] = \rho^2 + \frac{\rho}{|B|}. \quad (4.4)$$

where $\mathbb{E}_![\cdot]$ denotes the Palm expectation. It means expectation under the condition that there is a point of the spatio-temporal point process occurring at a fixed location. Generally, the Palm distribution is the conditional distribution of a point process under the condition that there are points of the point process at a given set of locations. For more precise details about the Palm distribution in a general setting, see Chiu *et al.* (2013), and for its extension to the spatio-temporal context, see Cronie and Van Lieshout (2015). Finally, for formal theoretical details on the LISTA functions see Rodríguez-Cortés *et al.* (2014); Rodríguez-Cortés (2014).

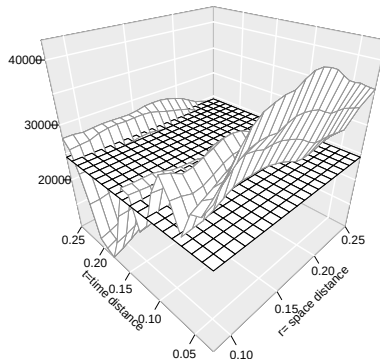
4.2.1 Behaviour of LISTA functions

The knowledge of the local spatio-temporal correlation between points in a neighbourhood of each event in a point pattern provides valuable information for cluster detection. Local indicators of association have a long trajectory as a powerful tool in spatial and spatio-temporal data mining. LISTA functions are theoretically known under Poisson processes. We briefly want to highlight their behaviour when the point pattern comes from a Poisson and clustered processes, and analyse the potential of their use in discovering events with local cluster structure. Thus we considered two types of spatio-temporal clustered processes, and estimated the global product density, and the LISTA functions for two selected events (using the function `LISTAfunc` in the GitHub repository `pdLISTA` (Rodríguez-Cortés, 2016)) and compared them with those under the Poisson case. We note that the global function inherits the shape of the local ones. Similar LISTA functions can be then grouped into bundles indicating points with similar local structure.

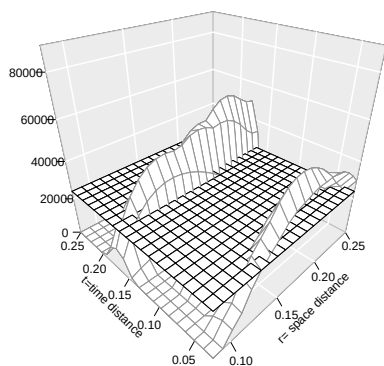
We conducted a simulation experiment to analyse the shapes of the estimated surfaces related to each model. We set the spatio-temporal window as $W \times T = [0, 1]^2 \times [0, 1]$ and simulated a number of spatio-temporal point patterns, using the R package `stpp` (Gabriel *et al.*, 2013) and the GitHub repository `mvstpp` (Rodríguez-Cortés, 2015). A grid was used for each spatial and temporal distances r and h spanning the sequence starting from $r > \epsilon > 0$ to 0.25 and $h > \delta > 0$ to 0.25 with small increments of distances. Both spatial and temporal components are computed using the Epanechnikov kernel (Illian *et al.*, 2008) and the bandwidths were estimated with a direct plug-in method (Sheather and Jones, 1991) using the `dpik` function of the R package `Kernsmooth` (Wand, 2015). Note that the selected band-



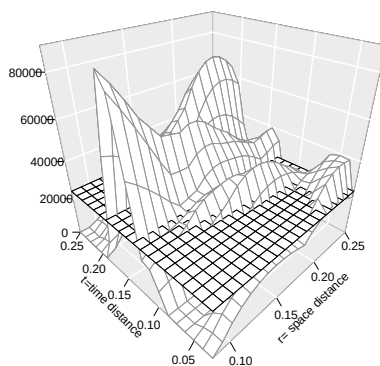
(a) Spatio-temporal point pattern



(b) Estimated global product density function

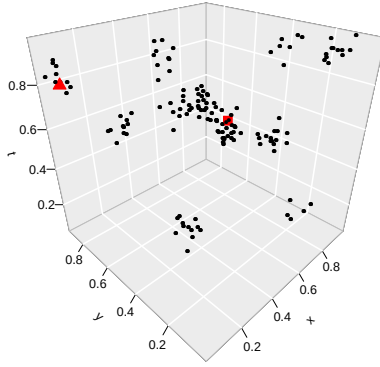


(c) Estimated LISTA function for the selected triangle point (▲)

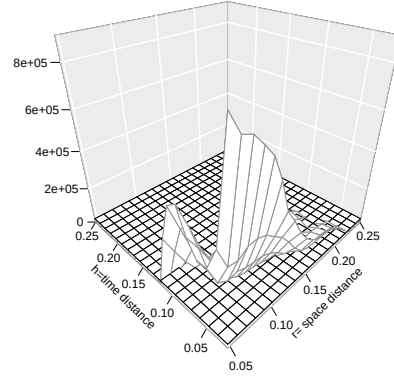


(d) Estimated LISTA function for the selected square point (■)

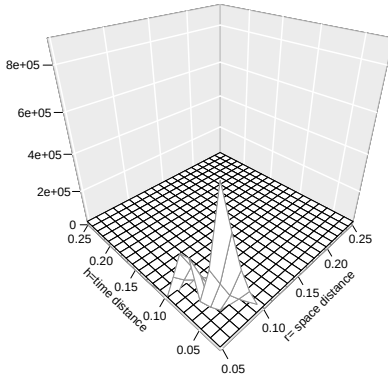
Figure 4.1: (4.1a) A realisation of a spatio-temporal Poisson model with expected number of points equal to 150 in the unit cube. (4.1b) Estimated global spatio-temporal product density (grey surface) together with the theoretical black surface under a Poisson process, and bandwidths $\epsilon = 0.080$ and $\delta = 0.026$. (4.1c)-(4.1d) Estimated LISTA functions (grey surface) for two selected points in (4.1a), together with the theoretical LISTA under a Poisson process (black surface) with same bandwidths as before.



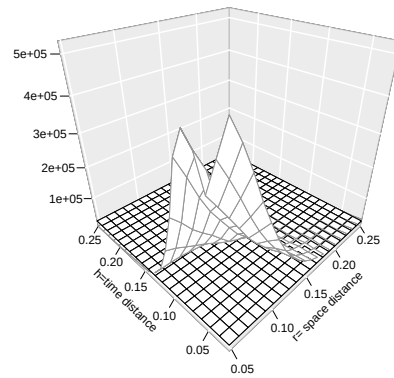
(a) Spatio-temporal point pattern



(b) Estimated global product density function



(c) Estimated LISTA function for the selected triangle point(▲)



(d) Estimated LISTA function for the selected square point (■)

Figure 4.2: (4.2a) A realisation of a spatio-temporal Poisson cluster model for a uniform distribution in space and time with parameters $n_p = 15$, $m_c = 10$, dispersion = $c(0.075, 0.075)$ in the unit cube. (4.2b) Estimated global spatio-temporal product density (grey surface) together with the theoretical black surface under a Poisson process, and bandwidths $\epsilon = 0.045$ and $\delta = 0.022$. (4.2c)-(4.2d) Estimated LISTA functions (grey surface) for two selected points in (4.2a), together with the theoretical LISTA under a Poisson process (black surface) with same bandwidths as before.

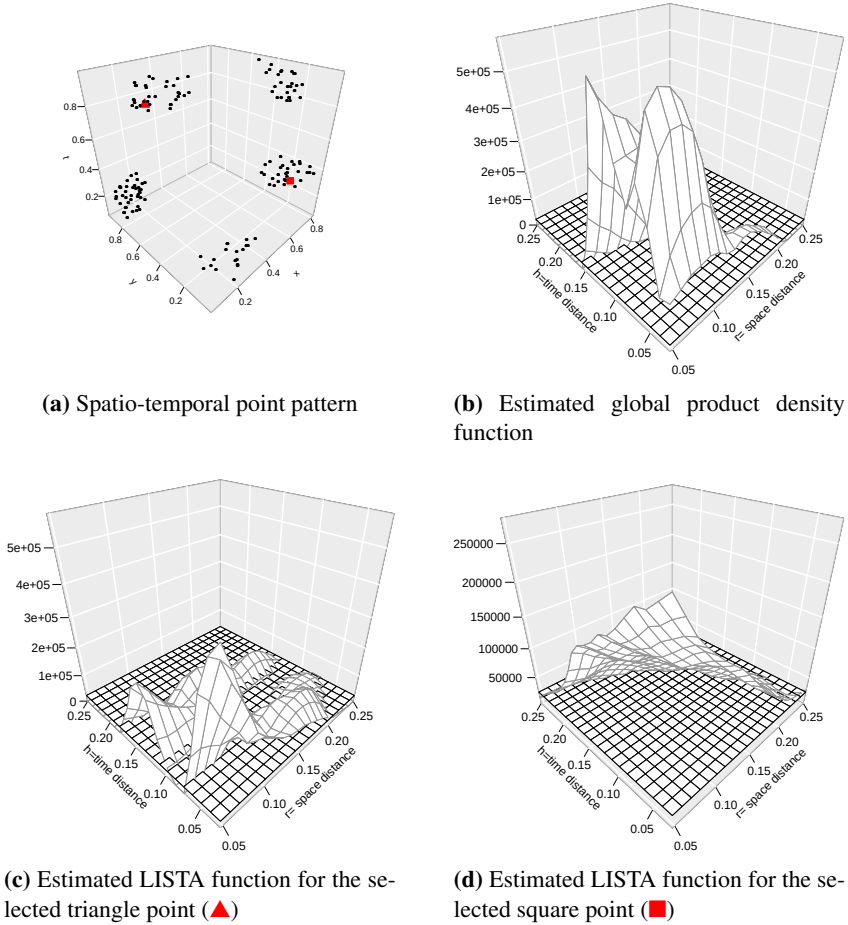


Figure 4.3: (4.3a) A realisation of a spatio-temporal double-cluster model with parameters $\lambda_p = 5$, $\mu = 30$, $a = 0.2$, $b = 0.1$, $c = 0.15$ in the unit cube. (4.3b) Estimated global spatio-temporal product density (grey surface) together with the theoretical black surface under a Poisson process, and bandwidths $\epsilon = 0.045$ and $\delta = 0.024$. (4.3c)-(4.3d) Estimated LISTA functions (grey surface) for two selected points in (4.3a), together with the theoretical LISTA under a Poisson process (black surface) with same bandwidths as before.

widths (ϵ and δ) were the same for all the estimated LISTA functions for each chosen point of the patterns. Figure 4.1a shows a realisation of a homogeneous spatio-temporal Poisson process (generated using the function `rpp` in R package `stpp`) with an intensity given by $\rho = n$ (n is the expected number of points in the unit cube). Figure 4.2a shows a realisation of a spatio-temporal cluster Poisson process (generated using the function `rpcp` in the R package `stpp` with uniform spatial and temporal distributions). This spatio-temporal model has the distinctiveness that the shape of its cluster are cylinders with central axis parallel to the time axis. Therefore the spatial dispersion parameter is equivalent to a radius r of a circle with centre in the parent points, and the time dispersion parameter is half of the height h of a cylinder. This model is governed by a set of parameters which determine its intensity $\lambda = n_p m_c$ (n_p is the expected number of parents, and m_c is the mean of offspring per parent) and the cluster degree (scale parameter).

Figure 4.3a shows a realisation of a spatio-temporal double-cluster model introduced by Stoyan *et al.* (2016) and generated using the function `stdcpp` in the GitHub repository `mvstpp` (Rodríguez-Cortés, 2015). This model is a straightforward extension of the planar Matérn cluster process (Matérn, 1986) that considers spheres or ellipsoids as the shape of its clusters, and takes the z -coordinate as times. The intensity is given by $\rho = \rho_p \mu$ where ρ_p is the number of clusters and μ is the mean of points per cluster, and the cluster degree is controlled by the length of the semi-axes given by a , b and c .

For the realisation from a homogeneous spatio-temporal Poisson process, we point out that the estimated surfaces of the global product density in Figure 4.1b is around the theoretical value under the Poisson assumption

$\rho^{(2)}(r, t) = \rho^2$ and the local indicator surfaces are around the plane $\rho^2 + \rho/|B|$, that is the theoretical surface under the Poisson assumption how is expressed in Equation (4.4).

On the other hand, we note that the estimated surfaces of the product density in Figure 4.2b and Figure 4.3b are higher than the theoretical $\rho^{(2)}(r, t) = \rho^2$ near to the origin in both space and time, confirming that the inter-event distances around (r, t) are relatively more frequent than under the Poisson case, which is typical of a cluster point process. Figures 4.2c, 4.2d and Figures 4.3c, 4.3d show that the shape of the local indicator surfaces can provide information about the clustering degree in a neighbourhood of each point. Both global and local surfaces show cluster behaviour, the difference being that the global one smooths out the local dependencies.

In general with the analysis of these surfaces for a given point pattern, we can explore how individual points are related to their neighbouring events and classify points with similar spatio-temporal local structure, identifying so areas that differ in terms of point densities. For example, in a spatial context, Mateu *et al.* (2010) use the spatial local product density function to classify the events in a point pattern in terms of features and clutter points using a multidimensional scaling approach. Another potential way to put in value these surfaces defined for each observed point is in conducting tests for local structure. For example to identify spatial disease clusters, Moraga and Montes (2011) propose a test based on the spatial local product density to detect kidney disease clusters in a case-control study, comparing the local structure of the two point patterns of cases and controls. In the following section, we propose a test in the spatio-temporal framework extending the idea in Moraga and Montes (2011).

4.3 Statistical testing procedure for local interaction

We propose a method for detecting significant differences in the local structure of two given point spatio-temporal patterns X and Z . We test the null hypothesis of no difference in the spatio-temporal local structure of X and Z with respect to the i -th point $(\mathbf{u}_i, t_i) \in X$ where $N(X) = n$ and $N(Z) = m$. A permutation test approach is considered, where the null distribution of a given statistical test is estimated by randomly permuting the labels associated to the two point patterns. With the previous hypothesis in mind, we assume that the point pattern X is of interest and then a p -value for each point in X is computed. The steps to define our testing procedure are as follows.

1. The first step is the estimation of the corresponding LISTA function $\widehat{\rho^{(2)i}_{\epsilon, \delta}}(r, h)$ for each of the $(\mathbf{u}_i, t_i) \in X$, for $i = 1, \dots, n$. We consider the settings explained in Section 4.2.
2. The next step concerns, for each fixed point $(\mathbf{u}_i, t_i) \in X$, the generation of k point patterns under the null hypothesis that the events in $X \setminus \{(\mathbf{u}_i, t_i)\}$ and Z are independently and randomly sampled from the same population. The set $Q = \{X \setminus \{(\mathbf{u}_i, t_i)\}\} \cup Z$ is a random labelled spatio-temporal point pattern with $n - 1$ labels equal to lab_X and m labels equal to lab_Z . We randomly permute the labels of the points in Q identifying a new set of events with labels lab_X , and adding $(\mathbf{u}_i, t_i)$ we define a new point pattern under the null hypothesis. For each fixed point $(\mathbf{u}_i, t_i)$, k local spatio-temporal product density surfaces are estimated based on the patterns coming from the permutation procedure. Moreover the estimated local product density surfaces under the null hypothesis, $\widehat{\rho^{(2)i_q}_{\epsilon, \delta}}(r, h)$ for $q = 1, \dots, k$ are summarised in terms

of the average surface, denoted by $\bar{\rho}_{H_0}^i(r, h)$, which for the Poisson case is given by (4.4), and the sample variance surface, denoted by $\sigma_{H_0}^{2i}(r, h)$.

3. Since the goal is to identify for which points in X the local structure with respect to Z is different, we consider the squared differences between the estimated local product density and its mean surface under the null hypothesis, and we propose the statistic

$$T_1^i = \int_0^{h_0} \int_0^{r_0} \left(\widehat{\rho^{(2)i}}_{\epsilon, \delta}(r, h) - \bar{\rho}_{H_0}^i(r, h) \right)^2 dr dh, \quad (4.5)$$

where r_0 and h_0 are chosen using the Diggle's rule (Diggle, 2013).

For illustrative purposes, we report graphically the estimated spatio-temporal local product density and its sample mean value under the null hypothesis for two particular situations, when the local structure of X and Z around a selected point $(\mathbf{u}_i, t_i) \in X$ is similar and when it is different. We generate X and Z from a spatio-temporal homogeneous Poisson process in the same unit cube with the same intensity ($\rho = 150$, see Figure 4.4a). For a selected point $(\mathbf{u}_i, t_i)$ of X , Figure 4.4b shows the estimated product density LISTA function and its sample mean surface $\bar{\rho}_{H_0}^i(r, h)$ under the null hypothesis. On the other hand, Figure 4.4d shows the LISTA function for a point belonging to a set of events generated from a spatio-temporal Poisson cluster point process in the unit cube with an expected number of parents $n_p = 5$, $m_c = 30$ the expected offspring per parent, with `cluster=c("uniform", "uniform")` and `dispersion=c(0.075, 0.075)` and having generated Z from a homogeneous spatio-temporal Pois-

son process with $\rho = 150$.

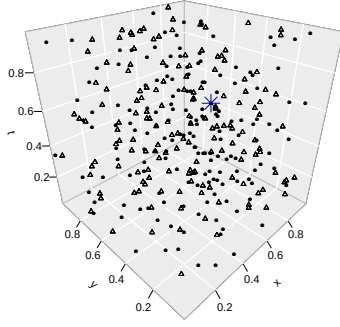
Since we expect that the variability of the surfaces is not the same and it is directly related with the range of spatio-temporal distances, we considered an additional standardised statistic that takes into account the variability of the estimated surfaces under the null hypothesis via the empirical variance, given by

$$T_2^i = \int_0^{h_0} \int_0^{r_0} \frac{\left(\widehat{\rho^{(2)i}}_{\epsilon, \delta}(r, h) - \bar{\rho}_{H_0}^i(r, h) \right)^2}{\sigma_{H_0}^{2i}(r, h)} dr dh. \quad (4.6)$$

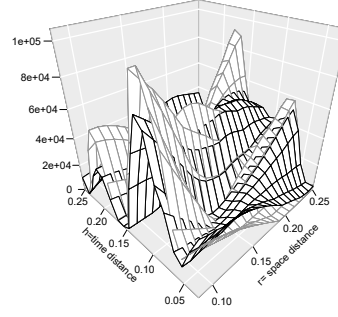
Summing up for a generic point $(\mathbf{u}_i, t_i) \in X$, the two statistics in (4.5) and (4.6) are proposed. We now explain how to obtain the corresponding empirical distributions for T_1^i and T_2^i under the null hypothesis.

4. The developed procedure is based on the permutation approach, that makes no theoretical assumptions on the distribution of the statistics. The theoretical distribution of our statistics under the null hypothesis is not known, so we rely on simulation-based empirical distributions. Fixing a point $(\mathbf{u}_i, t_i) \in X$, the estimated value of the j -th statistic for $j = 1, 2$, is compared with the empirical distribution of the k values of $T_{jH_0}^{iq}$ with $q = 1, \dots, k$ that are obtained computing the test between the q -th generated LISTA surfaces under the null hypothesis and their sample mean function. The p -value of T_j^i is calculated considering the following ratio $p_j^i = \sum_{q=1}^k \mathbf{I}(T_{jH_0}^{iq} \geq T_j^i) / k$. Given a fixed nominal value of α for the type I error, the null hypothesis is rejected if $p_j^i \leq \alpha$. In the procedure of local hypothesis testing, it is important to point

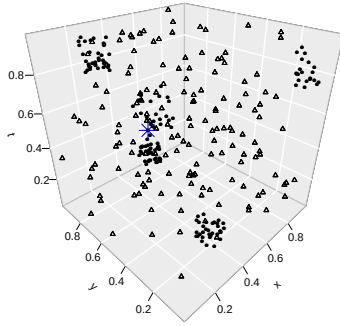
out that the statistics for individual locations will tend to be correlated (Anselin, 1995). Since multiple hypotheses are tested on the same data, a correction for the theoretical alpha (e.g. Bonferroni correction) should be adopted. However, we cannot determine exactly how many tests are really dependent, since a test T_i will be correlated with some of the other statistics. Indeed, the correlation between individual statistics is due to the common elements in the neighbourhood sets and the spatio-temporal product density is really sensitive for short distances (Mateu *et al.*, 2010). Performing simulations to calibrate the tests, we obtain an empirical type I error reasonable close to the theoretical one without considering an adjustment for multiple comparisons. On the other hand using a Bonferroni correction, we have obtained too conservative results under the null hypothesis.



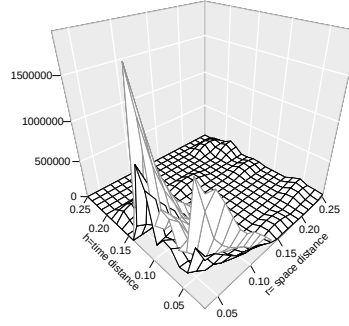
(a) Spatio-temporal point patterns



(b) Estimated LISTA function for the selected point together with the sample mean surface



(c) Spatio-temporal point patterns



(d) Estimated LISTA function for the selected point together with the sample mean surface

Figure 4.4: (4.4a) Realisations of two spatio-temporal homogeneous Poisson point patterns, X formed by the dotted points and Z formed by the triangle points. The “*” indicates a selected point $(\mathbf{u}_i, t_i) \in X$. (4.4b) Estimated LISTA function for the point $(\mathbf{u}_i, t_i)$ (grey surface) and sample mean of the LISTAs under the null hypothesis (black surface) for the pattern in 4.4a with bandwidths $\epsilon = 0.081$ and $\delta = 0.028$. (4.4c) Realisations of two spatio-temporal point patterns, the triangles are generated from a homogeneous Poisson process (Z) and the dots are coming from a Poisson cluster process (X). The “*” indicates a selected point $(\mathbf{u}_i, t_i) \in X$. (4.4d) Estimated LISTA function for the point $(\mathbf{u}_i, t_i)$ (grey surface) and sample mean of the LISTAs under the null hypothesis (black surface) for the pattern in 4.4c with bandwidths $\epsilon = 0.031$ and $\delta = 0.017$.

4.4 Simulation study

A simulation study with several scenarios is carried out to assess the performances in terms of type I error and power of the tests introduced in the previous section.

4.4.1 Empirical level

The two proposed statistical tests in (4.5) and (4.6) are compared in terms of type I error simulating point patterns under the null hypothesis of no difference in the spatio-temporal local structure of two given point patterns X and Z with respect to the i -th point $(\mathbf{u}_i, t_i) \in X$. We set the spatio-temporal window study as $W \times T = [0, 1]^2 \times [0, 1]$ and considered several scenarios varying the type of process (Poisson, Poisson cluster, double-cluster) and the corresponding parameters providing different intensities. We also considered $\mathbb{E}[N(W \times T)] = n + m = \{150, 300\}$ for $X \cup Z$. Under the null hypothesis, a pattern is generated with expected number of points equal to $n + m$, and the points are randomly associated to the pattern X or Z such that the number of points for the two sets is equal, and the tests are computed for all the points belonging in X . For each point, the number of permutations is equal to $k = 99$.

The spatio-temporal Poisson point patterns were generated using the function `rpp` in the package `stpp` of R. The Poisson cluster processes were simulated using `rpcp` in `stpp`. Finally, we used the function `stdcpp` in the GitHub repository `mvstpp` to generate double-cluster point processes. Given the values of $n + m$ and the dispersion parameters, we control the degree of clustering by changing the expected number of parents ($n_p = \{5, 10\}$) and the number of offspring points with respect to each parent.

For instance, with the Poisson cluster process, for each simulated parent the offspring points are generated in a cylinder with parameters $\{h, r\} = \{0.26, 0.13\}$; instead for the double-cluster model, they are generated in an ellipsoid with semi-axes lengths given by $\{a, b, c\} = \{0.2, 0.1, 0.15\}$.

For each of the resulting 10 scenarios under the null hypothesis, 100 pairs of patterns of (X, Z) are generated, and the type I error probability is defined as the proportion of points belonging to X for which the null hypothesis is rejected considering a fixed nominal value of α . Table 4.1 presents the average and the variance of the p -values under H_0 with the rejection rate for $\alpha = 0.05$. The two considered statistical tests exhibit acceptable empirical rejection rates for the several scenarios. There are no remarkable differences in the results when changing the intensity, the association structure between points and the degree of clustering. Moreover the mean and the variance of the empirical distribution of the p -values are reasonable comparable with the values of a uniform variable. It seems that T_2^i tends to overestimate the true significance level, since the null hypothesis is rejected more often. We remark that considering a correction for multiple comparisons (e.g. Bonferroni), the results for the two tests turns out to be too conservative with estimated rejection rate in the order of 0.01, even though the theoretical α is fixed at 0.05.

Scenarios		Dispersion	$n + m$		n_p	ϵ	δ	Rej.	T_1	Var	Rej.	T_2	Mean	Var
X	Z	parameters							Mean			Mean		
P	P	-	150	-		0.134	0.080	0.057	0.490	0.087	0.059	0.487	0.087	
	-	-	300	-		0.111	0.069	0.055	0.487	0.086	0.058	0.484	0.086	
PC	PC	$\{h, r\} = (0.26; 0.13)$	150	5		0.094	0.069	0.053	0.500	0.086	0.057	0.493	0.088	
			15			0.114	0.074	0.048	0.493	0.084	0.050	0.489	0.084	
			300	5		0.082	0.061	0.047	0.500	0.085	0.051	0.497	0.085	
			15			0.095	0.065	0.047	0.511	0.085	0.051	0.502	0.086	
DC	DC	$\{a, b, c\} = (0.2; 0.1; 0.15)$	150	5		0.086	0.065	0.054	0.503	0.086	0.054	0.499	0.087	
			15			0.107	0.074	0.049	0.501	0.084	0.051	0.495	0.085	
			300	5		0.076	0.060	0.052	0.507	0.087	0.055	0.500	0.088	
			15			0.090	0.065	0.054	0.488	0.085	0.060	0.480	0.086	

Table 4.1: Rejection rates (Rej.) at $\alpha = 0.05$, the mean and the variance (Var.) of the p -values for the statistics T_1 and T_2 . The spatio-temporal processes considered for X and Z are homogeneous Poisson point processes (P), Poisson cluster point processes (PC) and double-cluster point processes (DC). Dispersion parameters for the cluster models are given in the table, $n + m$ is the total expected number of points for $X \cup Z$, and n_p is the expected number of parents for the cluster models. For each scenario, 100 simulations are considered, ϵ and δ are the bandwidths in space and time, respectively.

4.4.2 Power

The power of a test is the probability of deciding on the alternative when it is the correct choice. We evaluated the power of the tests considering the proportion of the points $(\mathbf{u}_i, t_i)$ in the pattern X for which the null hypothesis is rejected, simulating them under the alternative assumption of difference in the local structure between X and Z . We defined scenarios that gradually moved far from the null hypothesis of local similarity between the two point patterns X and Z .

In the unit cube, the point pattern Z is generated from a homogeneous spatio-temporal Poisson process, and the pattern X differs in terms of the cluster structure from the previous one. The pattern X is generated from a spatio-temporal Poisson cluster model (PC) and from a double-cluster model (DC) such that the degree of clustering is comparable. For each type of process, the scenarios differ in terms of cluster degree, changing the size of the cylinder (ellipsoid) for the PC (DC) process, and the number of parents.

Both X and Z have the same expected number of points, such that the total expected number of points is $\mathbb{E}[N(W \times T)] = n + m = \{150, 300\}$. Table 4.2 shows the rejection rates of the tests T_1 and T_2 for the several considered scenarios, where $n + m$ is the total expected number of points in the spatio-temporal window, n_p indicates the number of parents, and we also varied the dispersion parameters controlling the dimension of the clusters.

For both tests, the power increases when the strength of the cluster structure increases, i. e., when moving from $n_p = 15$ down to $n_p = 5$ and when decreasing the values of the dispersion parameters that define the size of the clusters. Furthermore, the power improves when increasing the intensity in the unit cube (going from $n + m = 150$ to $n + m = 300$). Note that the

setting of the cluster degree of both PC and DC is comparable and thus with both statistics the performances are better in the case of cylindrical clusters. Comparing the two tests, the test T_2 tends to reject more than T_1 .

Scenarios X	$n + m$ Z	Dispersion parameters	$n_p = 15$				$n_p = 5$			
			ϵ	δ	Rej. T_1	Rej. T_2	ϵ	δ	Rej. T_1	Rej. T_2
PC	P	$\{h, r\} = (0.5; 0.25)$	0.081	0.028	0.246	0.271	0.072	0.027	0.585	0.623
		$\{h, r\} = (0.26; 0.13)$	0.077	0.028	0.390	0.391	0.061	0.025	0.843	0.842
		$\{h, r\} = (0.13; 0.065)$	0.071	0.026	0.575	0.551	0.048	0.023	0.943	0.931
PC	P	$\{h, r\} = (0.5; 0.25)$	0.058	0.018	0.320	0.363	0.052	0.017	0.659	0.709
		$\{h, r\} = (0.26; 0.13)$	0.055	0.018	0.585	0.613	0.043	0.016	0.927	0.943
		$\{h, r\} = (0.13; 0.065)$	0.048	0.017	0.860	0.865	0.032	0.014	0.969	0.977
DC	P	$\{a, b, c\} = (0.25; 0.15; 0.22)$	0.080	0.028	0.205	0.221	0.074	0.027	0.483	0.502
		$\{a, b, c\} = (0.2; 0.1; 0.15)$	0.078	0.028	0.264	0.267	0.069	0.026	0.661	0.657
		$\{a, b, c\} = (0.1; 0.05; 0.075)$	0.073	0.026	0.465	0.461	0.054	0.023	0.924	0.913
DC	P	$\{a, b, c\} = (0.25; 0.15; 0.22)$	0.058	0.018	0.285	0.328	0.054	0.017	0.623	0.670
		$\{a, b, c\} = (0.2; 0.1; 0.15)$	0.056	0.018	0.431	0.463	0.049	0.016	0.847	0.878
		$\{a, b, c\} = (0.1; 0.05; 0.075)$	0.050	0.017	0.792	0.808	0.035	0.014	0.990	0.994

Table 4.2: Rejection rates (Rej.) at $\alpha = 0.05$ for the statistics T_1 and T_2 . The spatio-temporal processes considered for X and Z are homogenous Poisson point processes (P), Poisson cluster point processes (PC) and double-cluster point processes (DC). Dispersion parameters for the cluster models are given in the table, $n + m$ is the total expected number of points in X and Z , and n_p is the expected number of parents for the cluster models. For each scenario, 100 simulations are considered, ϵ and δ are the bandwidths in space and time, respectively.

4.5 Application to earthquake data

To describe the seismicity of an area in space, time and magnitude domains, it could be useful to study separately the features of independent events of those that are strongly correlated. In the seismological context, these two set of points are connected to background seismicity and to aftershock sequences. Indeed, a background event refers to an earthquake that has not been triggered by another and that might be related to changes in the tectonic field. On the other hand, triggered events are thought to have been caused by a previous earthquake.

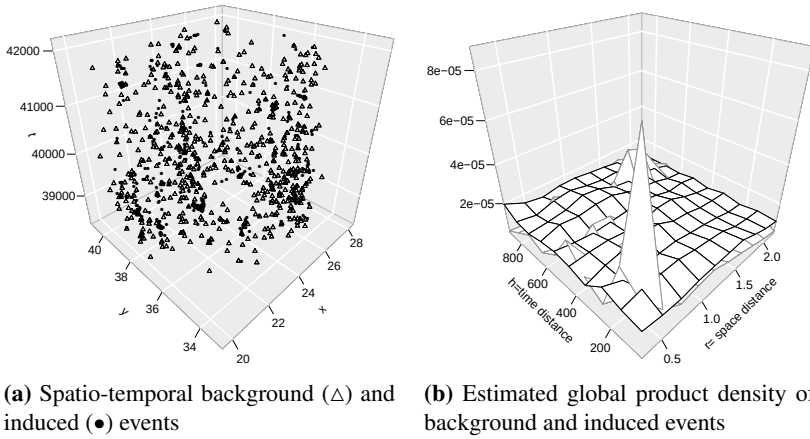


Figure 4.5: (4.5a) Scatterplot of the spatio-temporal earthquake data classified in background events (triangles) and induced events (points) according to the procedure of declustering using the ETAS model. (4.5b) Estimated global product density for the background events (black surface), and induced ones (grey surface) with bandwidths $\epsilon = 30.44$ km and $\delta = 44.14$ days.

Globally, these two set of events present different spatio-temporal global interaction structure, however it can be of interest to compare them focusing on a local scale. Using the defined tests, we want to identify the

induced events that are significantly different in terms of spatio-temporal local structure to the background events relating the results to the geological information available in the study area (such as faults, volcanoes, plate boundaries). Indeed, the earthquake generator process is very complex in different aspects: it is a multidimensional and multiscale process, with dependence features of points both in space and time, that vary according to their seismogenetic sources. In other words, the distribution of earthquake epicentres generally shows diverse interaction structures at different spatial and spatio-temporal scales and it can be related to the underlying seismological characteristics of the area.

4.5.1 Estimation of the ETAS model: background and induced events

In this application, we consider earthquake events occurred in the Greek area between 2011 and 2014 with a magnitude greater than 4 (Figure 4.5a), for a total number of 1105 events in the studied spatio-temporal window. The area is characterised by high seismic-hazard due to both tectonic and volcanic seismogenic sources, and Figure 2.2 shows information on the faults, the plate boundary and main volcanoes in the Greek area. The complex spatial multiscale structure of the data has been analysed in Chapter 2. The earthquakes are classified into background and induced events using a declustering procedure, a probability of being independent events is assigned to each one and it comes from an algorithm for the estimation of Epidemic Type Aftershocks-Sequences (ETAS) model (Ogata, 1988). See Section 1.2.5 that describes the ETAS model formulation.

We fitted the model using the R package `etasFLP` (Chiodi and Adelfio, 2014) based on the method developed in Chiodi and Adelfio (2011) and

	Estimates	std.err.
μ	0.186	0.008
k_0	0.037	0.021
c	0.029	0.008
p	1.095	0.026
α	1.478	0.088
γ	0.604	0.174
d	12.390	3.209
q	1.780	0.108

Table 4.3: Main output of the estimated parameters in the ETAS model and their standard errors.

Adelfio and Chiodi (2015a). The main output in terms of parametric estimation is shown in Table 4.3.

Interpreting, the main results fitting the ETAS model on the spatio-temporal Greek data, we have that the estimated expected intensity (μ) for the background is 0.186 and there are relevant induced events since the estimate of k_0 is significant and positive. α indicates the impact of the magnitude of the main shock to the following events in time. Instead, γ measures the effect in space of the magnitude of the previous event. Finally, since $p > 1$, this indicates slow decay of the aftershock sequences.

For each event, the probability of being a background event is estimated in order to provide an estimate of weights for a non-parametric density estimation of the spatial background component. We use the final probabilities provided in the last step of the iteration procedure to classify the events with a magnitude greater than 4 into the two groups, obtaining 580 background events and 525 triggered events (Figure 4.5a). Considering a classification between independent and induced earthquakes, we would ask the following research question: Is there a different global structure between the two point

pattern? Which triggered events have a significant different local structure with respect to the underlying process (background events)? Is there any geological justification of the identified clusters?

As expected, the spatio-temporal product density of the spontaneous events does not show any particular behaviour, and on the other hand for the induced seismicity there is a spatio-temporal clustering at around temporal distances $t < 300$ days and for spatial distances $r < 65$ kilometres (Figure 4.5b). However, we further aim to detect if we can obtain different conclusions focusing on a local scale detecting spatio-temporal clusters.

4.5.2 Application of the test based on the LISTA functions

We apply the testing procedure based on T_1 since according to the simulation results in Section 4.4, it has slightly better performance than T_2 . The point pattern Z is represented by the background events and the events in X are the triggered ones. To estimate the product density LISTA functions, we used the settings explained in Section 4.2.1 and the values of the bandwidths are $\epsilon = 12.99$ km and $\delta = 43.98$ days. Figure 4.6a and Figure 4.6c show the 3D density estimation of the significant and non-significant triggered points, respectively. Furthermore the representation of the results only in space allows to interpret them in relation to the geological information available in the study area (Figure 4.6b and Figure 4.6d).

We can identify some areas in which the induced events (with a magnitude greater than 4) are different in terms of spatio-temporal local structure than the background seismicity. Going into details, the earthquakes occurred in the western Greek islands of Kefalonia and Zakynthos and the Samos area (East Aegean Sea) have been characterised by clustered spatio-temporal sequences of events with a magnitude greater than 4. Whereas, along the

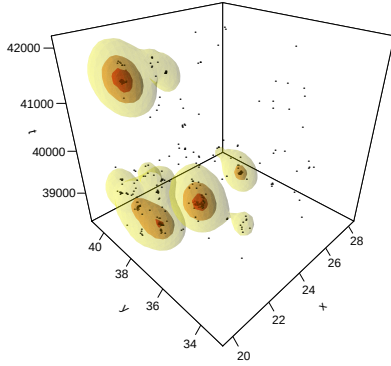
south part of the Hellenic Arc and in the area characterised by faults, the induced events with a magnitude greater than 4 do not differ from the background seismicity. The different behaviour of the significant sequences cannot be ascribed to the magnitude of the corresponding mainshocks because they are not higher than the mainshock magnitudes of the non-significant sequences. Probably it is due to their specific geological characteristics and, in particular, to a higher fracturation degree of their seismogenetic volumes. Otherwise, the central area, that is the Aegen Volcanic Arc, is characterised by a visco-plastic crust generating shallow microseismic activity, that may be the reason of the spread seismicity observed there. These results confirm our idea that the observed seismicity is generated by a complex model, characterised by spatial-temporal interaction, with events happening at several scales, and with spatial inhomogeneity related to the geological information available in the study area (the presence of volcanoes, plate boundaries and faults). Moreover, although with a given probability we could assign events to one of the two main components of seismicity (background or triggering), inside this classification we could also characterise events that have further different dependence features according to their generator source.

4.6 Final remarks

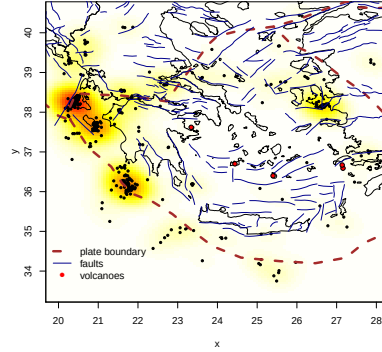
We deal with a non-parametric testing approach for spatio-temporal point processes, in order to compare the local structure of two spatio-temporal point patterns (say X and Z). With the defined tests, we obtain a p -value for each point in X that indicates if in the neighbourhood of the point the two point patterns under analysis have the same local structure. The defined

statistics T_1 and T_2 lead to approximately valid tests and the results in terms of type I error and power are reasonably good (also changing the sample size, the cluster size and the number of points inside the clusters). The test that does not take into account in its formulation the variability of the LISTA surfaces under the null hypothesis performs slightly better than the other one. We point out that the defined tests can be used when we have spatio-temporal case-control data that we want to compare, and we aim to assess where the cases are more clustered in a spatio-temporal window of interest. Furthermore, we suggest using the defined setting in contexts where the total number of points in X and Z is close.

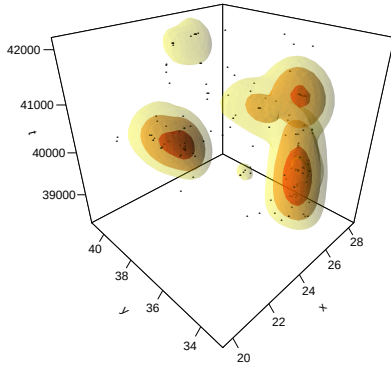
Using the aforementioned test, we compare background and induced seismicity with a magnitude greater than 4 in the Greek area identifying the points, and so the sequences of events, that have a local structure different to the background group. Besides, the results have been related to the sources of earthquakes in the study area. It seems that the sequences of events that are strongly different to the underlying process, are placed in specific regions and along the plate boundary, and these areas may be characterised by a higher fracturation degree. Instead, the sequence of events that have local structure similar to the background events are mostly in the area characterised by faults.



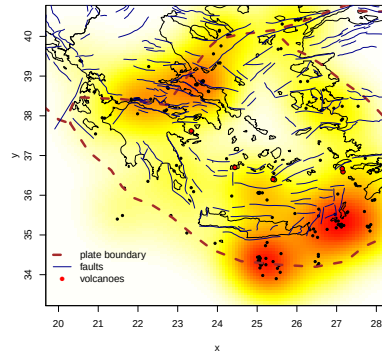
(a) Spatio-temporal density of the significant induced events



(b) Spatial density of the significant induced events and the seismic sources



(c) Spatio-temporal density of the non-significant induced events



(d) Spatial density of the significant induced events and the seismic sources

Figure 4.6: (4.6a) 3D scatterplot and estimated density of the significant induced events that have a different spatio-temporal local structure than the background seismiscity. (4.6b) Image plot in space of the significant induced events and seismogenetic sources. (4.6c) 3D scatterplot and estimated density of the non-significant induced events that have a different spatio-temporal local structure than the background seismiscity. (4.6d) Image plot in space of the non-significant induced events and seismogenetic sources.

Chapter 5

Conclusions and future works

In this dissertation, we propose some developments in the context of spatial and spatio-temporal point process analysis with emphasis on the description of complex interaction structure between events. Basically, using the contributions of this thesis, we describe in a novel way some aspects of the complex natural process that characterise the occurrence of tectonic events. After explaining the main statistical framework in the spatial and spatio-temporal point process model (Chapter 1), the thesis is articulated into three main works. There are several links among them, such as the description of interaction between points especially clustering, and the use of spatial covariates available in the study area to better explain the inhomogeneity and interpret the results.

In the spatial case, we identify properly a multi-scale cluster structure fitting hybrid models of Geyer processes (Chapter 2). In the spatio-temporal case, we propose some LGCP model specifications where the covariance function of the underlying Random field is ascribed to capture the spatio-temporal interaction structure between points, proposing a new estimation

approach that involves the covariance function characteristics (Chapter 3). Moreover, we detect local spatio-temporal cluster structure defining a non-parametric test (Chapter 4).

Going into details for each work, we summarise the obtained results underlining the advantages of the proposed analysis, indicating the weak aspects and the possible developments for future works.

The work in Chapter 2 deals with the interesting problem of modelling the spatial interaction for the epicentres of earthquakes with magnitude above 4 over a period of 10 years in Greece using hybrid of Gibbs point process model Baddeley *et al.* (2013). Since the spatial distribution of the events is relatively complex and it seems that the different sources of earthquakes (faults, active tectonic plate and volcanoes) produce events with different spatial displacements and orientations, we characterise the multi-scale interaction between the earthquakes and the relationship between the conditional spatial intensity and the geological information available in the study area. The spatial variables used in the analysis are: distance to the nearest fault, the distance to the nearest volcano and the distance to the plate boundary. We find two competing inhomogeneous hybrid of Geyer processes, with respectively a full-parametric and a semi-parametric formulation. These models describe properly the spatial multi-scale clustering aspect between points. As for future work, some possible extensions could regard an improvement of the structure of the full parametric model considering in a proper way the interaction between covariates. In fact the distance to the nearest fault and the distance to the plate boundary for the observed points are not independent and they present an exponential relationship. Moreover, it would be of interest trying to subdivide the study region into macro areas characterising the structure of interaction between

points in these different areas. Indeed, we expect that the multi-scale cluster between events is different for points that occur along the plate boundary than points that occur in the neighbourhood of faults. Another possible extension could be to consider a spatio-temporal formulation of hybrid models and also including depth information.

Although, the ability to predict earthquakes is of huge importance and a very difficult problem from the modelling aspect, this work does not contribute towards this goal, but rather on the description of the spatial intensity of earthquake events. However performing a spatial partition, and assuming a multi-type hybrid Gibbs process, as well as considering an extension in space and time, needs a more detailed methodological formulation (Baddeley *et al.*, 2013).

In Chapter 3, we describe the occurrence of events in Alaska Peninsula using a historical independent model describing the general observed inhomogeneity as a function of the geological spatial covariates and introducing a new estimation approach for LGCP. The interaction structure between events is described through a separable or a non-separable covariance function for the Gaussian Random Field taking into account properly the observed anisotropy. We want to point out that in the analysis of earthquake events, there are no application based on LGCP models and usually most of model framework in space-time is based on the mechanistic branching model, such as the ETAS model. For this reason, our work may be seen as a first attempt in the application of LGCP models for the description of the complex earthquake phenomena. Some of the analysis reported in this dissertation are provisional and partial, they need further insights in terms of model selection, model checking also for temporal prediction. For instance, it would be notable to set up a surveillance system for detecting emergent

spatio-temporal clusters of events.

In Chapter 4, we have introduced and studied the behaviour of the Local indicator of spatio-temporal association (LISTA) based on the spatio-temporal product density for some simulated dataset. Motivated by an application on earthquake events, we deal with a non-parametric testing approach for spatio-temporal point processes, in order to compare the local structure of two spatio-temporal point patterns (say X and Z), assuming that one of the two patterns (for instance X) is our focus of interest. With the defined tests, we assess if the events in X are an independent random sample from the superposition of X and Z (null hypothesis). Briefly, we obtain a p -value for each point in X that indicates if the null hypothesis for a point is not rejected, namely if in the neighbourhood of the point the two point patterns under analysis have the same local structure. The two proposed tests are based on the permutation approach and they differ in the way in which the LISTA surface for a point is summarised. We assess the defined tests setting up a simulation study, and the results in terms of type I error and power are reasonably good

However, the proposed tests have some limitations that are common to all the tests based on the permutation approach. A drawback is that not all the enumeration of the possible permutations are considered but randomly a sample of them is used to get a suitable approximation to the permutation distribution under the null hypothesis. In our case for each point, we consider $k=100$ permutations but increasing this value deals to a more time-consuming computation. Additionally, using Monte Carlo based approach, two individuals using the same statistical method on the same data may reach different conclusions.

As possible future development, we may consider the definition of other

local tests based on LISTA surfaces, for instance to assess locally a fitted model in terms of second-order characteristics. Furthermore, we may define other LISTA surfaces based on other spatio-temporal second-order summary statistics, such as the spatio-temporal pair correlation function. Finally, other possible analysis can be done to classify the points according to their local structure, and also to identify outlying components measuring the influence of each event to the global second-order statistics.

Finally, we want to give some general considerations that are common to all the works and they open the possibility for further future works. For all the analysis, the spatial geological information are transformed into spatial covariates considering the distance of a generic point to the sources of earthquakes. We want to point out that the available geological information are in shape-files that are a simple representation of more complex geological structures. For instance, the line is the commonly way to represent a fault on geologic maps that corresponds to the fault trace, the intersection of a fault with the ground surface. However, a fault has a more complex 3-D structure since it is a fracture along which the blocks of crust on either side have moved relative to one another parallel to the fracture. It would be of interest to analyse interaction between events and the geological information considering also the depth information. For all the analysis, we apply methodologies defined for processes in the plane. Nevertheless, it would be of interest to do some analysis, and compare them with those obtained in this thesis, considering the earthquakes as a realisation of a point process defined in a region of a sphere (Lawrence *et al.*, 2016). In fact, the definition and application of statistical methodologies on a process defined on the sphere (or on a subset) are one of the most prominent research field in point process theory.

Bibliography

- Adelfio, G. (2010). An analysis of earthquakes clustering based on a second-order diagnostic approach. In *Data Analysis and Classification*, pages 309–317. Springer.
- Adelfio, G. and Chiodi, M. (2009). Second-order diagnostics for space-time point processes with application to seismic events. *Environmetrics*, **20**(8), 895–911.
- Adelfio, G. and Chiodi, M. (2015a). Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stochastic Environmental Research and Risk Assessment*, **29**(2), 443–450.
- Adelfio, G. and Chiodi, M. (2015b). Flp estimation of semi-parametric models for space-time point processes and diagnostic tools. *Spatial Statistics*, **14**, 119–132.
- Adelfio, G. and Schoenberg, F. P. (2009). Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Annals of the Institute of Statistical Mathematics*, **61**(4), 929–948.
- Ahn, J., Johnson, T. D., Bhavnani, D., Eisenberg, J. N. and Mukherjee, B.

- (2014). A space-time point process model for analyzing and predicting case patterns of diarrheal disease in northwestern ecuador. *Spatial and Spatio-Temporal Epidemiology*, **9**, 23–35.
- Allard, D., Brix, A. and Chadoeuf, J. (2001). Testing local independence between two point processes. *Biometrics*, **57**(2), 508–517.
- Allard, D., Senoussi, R. and Porcu, E. (2016). Anisotropy models for spatial data. *Mathematical Geosciences*, **48**(3), 305–328.
- Ambraseys, N. and Adams, R. (1998). The rhodes earthquake of 26 june 1926. *Journal of Seismology*, **2**(3), 267–292.
- Anselin, L. (1993). *The Moran scatterplot as an ESDA tool to assess local instability in spatial association*. Regional Research Institute, West Virginia University Morgantown, WV.
- Anselin, L. (1995). Local indicators of spatial association-lisa. *Geographical analysis*, **27**(2), 93–115.
- Anwar, S., Stein, A. and van Genderen, J. (2012). Implementation of the marked strauss point process model to the epicenters of earthquake aftershocks. In Y. Leung, editor, *Advances in Geo-Spatial Information Science*., pages 125–140. Taylor & Francis, London.
- Assunção, R. and Correa, T. (2009). Surveillance to detect emerging space-time clusters. *Computational Statistics & Data Analysis*, **53**(8), 2817–2830.
- Assunção, R., Tavares, A., Correa, T. and Kulldorff, M. (2007). Space-time cluster identification in point processes. *Canadian Journal of Statistics*, **35**(1), 9–25.

- Baddeley, A. and Møller, J. (1989). Nearest-neighbour markov point processes and random sets. *International Statistical Review*, **57**(2), 89–121.
- Baddeley, A. and Turner, R. (2005). Spatstat: An r package for analyzing spatial point patterns. *Journal of Statistical Software*, **12**(i06).
- Baddeley, A. and Turner, T. R. (2000). Practical maximum pseudo likelihood for spatial point patterns (with discussion). *Australian & New Zealand Journal of Statistics*, **42**(3), 283–322.
- Baddeley, A., Møller, J. and Waagepetersen, R. (2000). Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**(3), 329–350.
- Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005). Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, **67**(5), 617–666.
- Baddeley, A., Gregori, P., Mateu, J., Stoica, R. and Stoyan, D. (2006). *Case Studies in Spatial Point Pattern Modelling*. Lecture Notes in Statistics, 185. Springer, New York.
- Baddeley, A., Rubak, E. and Møller, J. (2011). Score, pseudo-score and residual diagnostics for spatial point process models. *Statistical Science*, **26**(4), 613–646.
- 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., Rubak, E. and Turner, R. (2015). *Spatial Point Patterns:*

- Methodology and Applications with R*. London: Chapman and Hall/CRC Press.
- 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**(5), 1–15.
- Beneš, V., Bodlák, K., Møller, J. and Waagepetersen, R. (2011). A case study on point process modelling in disease mapping. *Image Analysis & Stereology*, **24**(3), 159–168.
- Berge, C. (1971). *Principles of combinatorics*. New York: Academic Press.
- Berman, M. (1986). Testing for spatial association between a point process and another stochastic process. *Applied Statistics*, **35**(1), 54–62.
- Besag, J. (1975). Statistical analysis of non-lattice data. *The Statistician*, **24**(3), 179–195.
- Besag, J. (1977). Some methods of statistical analysis for spatial data. *Bulletin of the International Statistical Institute*, **47**(2), 77–92.
- Bird, P. (2003). An updated digital model of plate boundaries. *Geochemistry, Geophysics, Geosystems*, **4**(3).
- Bohnhoff, M., Rische, M., Meier, T., Becker, D., Stavrakakis, G. and Harjes, H.-P. (2006). Microseismic activity in the hellenic volcanic arc, greece, with emphasis on the seismotectonic setting of the santorini–amorgos zone. *Tectonophysics*, **423**(1), 17–33.

- 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 Moller, J. (2001). Space-time multi type log gaussian cox processes with a view to modelling weeds. *Scandinavian Journal of Statistics*, **28**(3), 471–488.
- Caputo, R., Chatzipetros, A., Pavlides, S. and Sboras, S. (2013). The greek database of seismogenic sources (gredass): state-of-the-art for northern greece. *Annals of Geophysics*, **55**(5).
- Chiodi, M. and Adelfio, G. (2011). Forward likelihood-based predictive approach for space-time processes. *Environmetrics*, **22**(6), 749–757.
- Chiodi, M. and Adelfio, G. (2014). etasflp: Estimation of an etas model. mixed flp (forward likelihood predictive) and ml estimation of non-parametric and parametric components of the etas model for earthquake description. *R package version 1.0*.
- Chiu, S. N., van de Weygaert, R. and Stoyan, D. (1996). The sectional Poisson-Voronoi tessellation is not a Voronoi tessellation. *Advances in Applied Probability*, **28**(2), 356–376.
- Chiu, S. N., Stoyan, D., Kendall, W. S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*. John Wiley & Sons, third edition.
- Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 129–164.
- Cox, D. R. and Isham, V. (1980). *Point Processes*. Chapman and Hall.

- Cressie, N. and Collins, L. B. (2001a). Analysis of spatial point patterns using bundles of product density lisa functions. *Journal of Agricultural, Biological, and Environmental Statistics*, **6**(1), 118–135.
- Cressie, N. and Collins, L. B. (2001b). Patterns in spatial point locations: Local indicators of spatial association in a minefield with clutter. *Naval Research Logistics*, **48**(5), 333–347.
- Cronie, O. and Van Lieshout, M. (2015). A J-function for inhomogeneous spatio-temporal point processes. *Scandinavian Journal of Statistics*, **42**(2), 562–579.
- D'Alessandro, A., Papanastassiou, D. and Baskoutas, I. (2011). Hellenic unified seismological network: an evaluation of its performance through snes method. *Geophysical Journal International*, **185**(3), 1417–1430.
- Daley, D. J. and Vere-Jones, D. (2007). *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*. Springer-Verlag, New York, second edition.
- Davies, T. M. and Hazelton, M. L. (2013). Assessing minimum contrast parameter estimation for spatial and spatiotemporal log-gaussian cox processes. *Statistica Neerlandica*, **67**(4), 355–389.
- Diggle, P., Chetwynd, A., Häggkvist, R. and Morris, S. (1995). Second-order analysis of space-time clustering. *Statistical Methods in Medical Research*, **4**(2), 124–136.
- Diggle, P., Rowlingson, B. and Su, T.-I. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics*, **16**(5), 423–434.

- Diggle, P. J. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press.
- Diggle, P. J., Gómez-Rubio, V., Brown, P. E., Chetwynd, A. G. and Gooding, S. (2007). Second-order analysis of inhomogeneous spatial point processes using case-control data. *Biometrics*, **63**(2), 550–557.
- Diggle, P. J., Moraga, P., Rowlingson, B., Taylor, B. M. *et al.* (2013). Spatial and spatio-temporal log-gaussian cox processes: extending the geo-statistical paradigm. *Statistical Science*, **28**(4), 542–563.
- Dimitriadis, I., Karagianni, E., Panagiotopoulos, D., Papazachos, C., Hatzidimitriou, P., Bohnhoff, M., Rische, M. and Meier, T. (2009). Seismicity and active tectonics at coloumbo reef (aegean sea, greece): Monitoring an active volcano at santorini volcanic center using a temporary seismic network. *Tectonophysics*, **465**(1), 136–149.
- 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., Rowlingson, B. S. and Diggle, P. J. (2013). stpp: An R package for plotting, simulating and analyzing Spatio-Temporal Point Patterns. *Journal of Statistical Software*, **53**(2), 1–29.
- Gelfand, A. E., Diggle, P., Guttorp, P. and Fuentes, M. (2010). *Handbook of spatial statistics*. CRC press.

- Getis, A. and Franklin, J. (1987). Second-order neighborhood analysis of mapped point patterns. *Ecology*, **68**(3), 473–477.
- Getis, A. and Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, **24**(3), 189–206.
- Geyer, C. J. (1999). Likelihood inference for spatial point processes. *Stochastic Geometry: Likelihood and Computation*, **80**, 79–140.
- Giorgi, E., Kreppel, K., Diggle, P. J., Caminade, C., Ratsitorahina, M., Rajerison, M. and Baylis, M. (2016). Modeling of spatio-temporal variation in plague incidence in madagascar from 1980 to 2007. *Spatial and Spatio-temporal Epidemiology*, **19**, 125–135.
- Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, **97**(458), 590–600.
- González, J. A., Rodríguez-Cortés, F., Cronie, O. and Mateu, J. (2016). Spatio-temporal point process statistics: a review. *Spatial Statistics*, **18**, 505–544.
- Hawkes, A. and Adamopoulos, L. (1973). Cluster models for earthquakes-regional comparison. *Bulletin of the International Statistical Institute*, **45**(3), 454–461.
- Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, volume 70. John Wiley & Sons.

- Jensen, J. L., Moller, J. *et al.* (1991). Pseudolikelihood for exponential family models of spatial point processes. *The Annals of Applied Probability*, **1**(3), 445–461.
- Koehler, R. (2013). Quaternary faults and folds (qff): Alaska division of geological.
- Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics-Theory and methods*, **26**(6), 1481–1496.
- Kulldorff, M. (2001). Prospective time periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **164**(1), 61–72.
- Kulldorff, M. and Nagarwalla, N. (1995). Spatial disease clusters: detection and inference. *Statistics in Medicine*, **14**(8), 799–810.
- Kulldorff, M., Athas, W. F., Feurer, E. J., Miller, B. A. and Key, C. R. (1998). Evaluating cluster alarms: a space-time scan statistic and brain cancer in los alamos, new mexico. *American Journal of Public Health*, **88**(9), 1377–1380.
- Kulldorff, M., Heffernan, R., Hartman, J., Assunção, R. and Mostashari, F. (2005). A space-time permutation scan statistic for disease outbreak detection. *Plos med*, **2**(3), e59.
- Lawrence, T., Baddeley, A., Milne, R. K. and Nair, G. (2016). Point pattern analysis on a region of a sphere. *Stat*, **5**(1), 144–157.
- Le Pichon, X. and Angelier, J. (1979). The hellenic arc and trench system: a key to the neotectonic evolution of the eastern mediterranean area. *Tectonophysics*, **60**(1), 1–42.

- Liang, S., Carlin, B. P. and Gelfand, A. E. (2008). Analysis of minnesota colon and rectum cancer point patterns with spatial and nonspatial covariate information. *The Annals of Applied Statistics*, **3**(3), 943.
- Matérn, B. (1986). *Spatial Variation*, volume 36. Springer-Verlag, Berlin.
- Mateu, J., Lorenzo, G. and Porcu, E. (2007). Detecting features in spatial point processes with clutter via local indicators of spatial association. *Journal of Computational and Graphical Statistics*, **16**(4), 968–990.
- Mateu, J., Lorenzo, G. and Porcu, E. (2010). Features detection in spatial point processes via multivariate techniques. *Environmetrics*, **21**(3-4), 400–414.
- Meulenkamp, J., Wortel, M., Van Wamel, W., Spakman, W. and Strating, E. H. (1988). On the hellenic subduction zone and the geodynamic evolution of crete since the late middle miocene. *Tectonophysics*, **146**(1), 203–215.
- Mohler, G. (2013). Modeling and estimation of multi-source clustering in crime and security data. *The Annals of Applied Statistics*, **7**(3), 1525–1539.
- Møller, J. (2003). Shot noise cox processes. *Advances in Applied Probability*, pages 614–640.
- Møller, J. and Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica*, **66**(4), 472–491.
- Møller, J. and Toftaker, H. (2014). Geometric anisotropic spatial point

- pattern analysis and cox processes. *Scandinavian Journal of Statistics*, **41**(2), 414–435.
- Møller, J. and Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
- Møller, J., Syversveen, A. R. and Waagepetersen, R. P. (1998). Log gaussian cox processes. *Scandinavian Journal of Statistics*, **25**(3), 451–482.
- Moraga, P. and Montes, F. (2011). Detection of spatial disease clusters with lisa functions. *Statistics in Medicine*, **30**(10), 1057–1071.
- Muggeo, V. M. (2003). Estimating regression models with unknown breakpoints. *Statistics in Medicine*, **22**(19), 3055–3071.
- Muggeo, V. M. (2008). Segmented: an r package to fit regression models with broken-line relationships. *R news*, **8**(1), 20–25.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, **83**(401), 9–27.
- Papazachos, V., Papazachos, B., Papazachou, C. and Papazachou, K. (1997). *The Earthquakes of Greece*. Editions Ziti.
- R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Ripley, B. D. (1976). The second-order analysis of stationary point processes. *Journal of Applied Probability*, **13**, 255–266.

- Ripley, B. D. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society Series B*, **39**(2), 172–212.
- Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press.
- Ripley, B. D. and Kelly, F. P. (1977). Markov point processes. *Journal of the London Mathematical Society*, **2**(1), 188–192.
- Rodrigues, A. and Diggle, P. J. (2010). A class of convolution-based models for spatio-temporal processes with non-separable covariance structure. *Scandinavian Journal of Statistics*, **37**(4), 553–567.
- Rodrigues, A. and Diggle, P. J. (2012). Bayesian estimation and prediction for inhomogeneous spatiotemporal log-gaussian cox processes using low-rank models, with application to criminal surveillance. *Journal of the American Statistical Association*, **107**(497), 93–101.
- Rodríguez-Cortés, F. J. (2014). *Modelling, Estimation and Applications of Second-Order Spatio-Temporal Characteristics of Point Processes*. Ph.D. thesis, Departament de Matemàtiques; Universitat Jaume I.
- Rodríguez-Cortés, F. J. (2015). mvstpp: Mark Variogram for Spatio-Temporal Point Processes. GitHub repository.
- Rodríguez-Cortés, F. J. (2016). pdLISTA: Second-order product density local indicator of spatio-temporal association function. GitHub repository.
- Rodríguez-Cortés, F. J., Ghorbani, M., Mateu, J. and Stoyan, D. (2014). On the expected value and variance for an estimator of the spatio-temporal product density function. Research Report Series No. R-2014-03, Department of Mathematical Sciences, Aalborg University.

- Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C. and Rue, H. (2014). Spatio-temporal log-gaussian cox processes for modelling wildfire occurrence: the case of catalonia, 1994–2008. *Environmental and Ecological Statistics*, **21**(3), 531–563.
- Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, **53**(3), 683–690.
- Shirota, S. and Gelfand, A. E. (2016). Space and circular time log gaussian cox processes with application to crime event data. *arXiv preprint arXiv:1611.08719*.
- Siebert, L. and Simkin, T. (2014). Volcanoes of the world: an illustrated catalog of holocene volcanoes and their eruptions. *Smithsonian Institution, Global Volcanism Program Digital Information Series, GVP-3*.
- Siino, M., Adelfio, G., Mateu, J., Chiodi, M. and D’Alessandro, A. (2016a). Spatial pattern analysis using hybrid models: an application to the hellenic seismicity. *Stochastic Environmental Research and Risk Assessment*, (DOI: 10.1007/s00477-016-1294-7).
- Siino, M., Mateu, J. and Adelfio, G. (2016b). Spatio-temporal log-gaussian cox processes on earthquake events. In *Proceedings of the Eight International Workshop on Spatio-Temporal Modelling (METMA8)*.
- Siino, M., Rodríguez-Cortés, F. J., Mateu, J. and Adelfio, G. (2016c). Testing for local structure in spatio-temporal point pattern data. *Submitted*.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, volume 26. Chapman and Hall/CRC, London.

- Sokos, E., Kiratzi, A., Gallovič, F., Zahradník, J., Serpetsidaki, A., Plicka, V., Janský, J., Kostecký, J. and Tselentis, G.-A. (2015). Rupture process of the 2014 cephalonia, greece, earthquake doublet (mw6) as inferred from regional and local seismic data. *Tectonophysics*, **656**, 131–141.
- Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields: methods of geometrical statistics*. Wiley, Chichester.
- Stoyan, D. and Stoyan, H. (2000). Improving ratio estimators of second order point process characteristics. *Scandinavian Journal of Statistics*, **27**(4), 641–656.
- Stoyan, D., Rodríguez-Cortés, F. J., Mateu, J. and Gille, W. (2016). Mark variograms for spatio-temporal point processes. *submitted*, pages 1–34.
- 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 spatio-temporal distribution of rat sightings in an urban environment. *Spatial Statistics*, **9**, 192–206.
- Tango, T. (2010). *Statistical Methods for Disease Clustering*. Springer-Verlag, New York.
- Taylor, B. M., Davies, T. M., Rowlingson, B. S. and Diggle, P. J. (2013). lgcp: An R package for inference with spatial and spatio-temporal log-Gaussian Cox processes. *Journal of Statistical Software*, **52**(4), 1–40.
- Taylor, B. M., Davies, T. M., Rowlingson, B. S. and Diggle, P. J. (2015). Bayesian inference and data augmentation schemes for spatial, spa-

- tiotemporal and multivariate log-Gaussian Cox processes in R. *Journal of Statistical Software*, **63**(7), 1–48.
- Veen, A. and Schoenberg, F. P. (2008). Estimation of space–time branching process models in seismology using an em–type algorithm. *Journal of the American Statistical Association*, **103**(482), 614–624.
- Vere-Jones, D. (1978). Earthquake prediction-a statistician’s view. *Journal of Physics of the Earth*, **26**(2), 129–146.
- Wand, M. (2015). *KernSmooth: Functions for Kernel Smoothing Supporting Wand & Jones (1995)*. R package version 2.23-15.
- Ye, X., Yu, J., Wu, L., Li, S. and Li, J. (2015). Open source point process modeling of earthquake. In *Geo-Informatics in Resource Management and Sustainable Ecosystem*, pages 548–557. Springer.