



Local spatial log-Gaussian Cox processes for seismic data

Nicoletta D'Angelo¹ · Marianna Siino¹ · Antonino D'Alessandro² ·
Giada Adelfio¹

Received: 7 January 2021 / Accepted: 15 February 2022 / Published online: 25 April 2022
© The Author(s) 2022, corrected publication 2022

Abstract

In this paper, we propose the use of advanced and flexible statistical models to describe the spatial displacement of earthquake data. The paper aims to account for the external geological information in the description of complex seismic point processes, through the estimation of models with space varying parameters. A local version of the Log-Gaussian Cox processes (LGCP) is introduced and applied for the first time, exploiting the inferential tools in Baddeley (Spat Stat 22:261–295, 2017), estimating the model by the local Palm likelihood. We provide methods and approaches accounting for the interaction among points, typically described by LGCP models through the estimation of the covariance parameters of the Gaussian Random Field, that in this local version are allowed to vary in space, providing a more realistic description of the clustering feature of seismic events. Furthermore, we contribute to the framework of diagnostics, outlining suitable methods for the local context and proposing a new step-wise approach addressing the particular case of multiple covariates. Overall, we show that local models provide good inferential results and could serve as the basis for future spatio-temporal local model developments, peculiar for the description of the complex seismic phenomenon.

Keywords Log-Gaussian Cox process · Local composite likelihood · Spatial point processes · Palm likelihood · Seismology

1 Introduction

Interest in methods for analysing spatial point processes is increasing across many fields of science, notably in ecology, epidemiology, geoscience, astronomy, econometrics, and crime research (Baddeley et al. 2015; Diggle 2013). When the structure

✉ Nicoletta D'Angelo
nicoletta.dangelo@unipa.it

¹ Dipartimento di Scienze Economiche, Aziendali e Statistiche, Università degli Studi di Palermo, Palermo, Italy

² Istituto Nazionale di Geofisica e Vulcanologia (INGV), Palermo, Italy

of a given point pattern is observed, it is assumed as a realization of an underlying generating process, whose properties are estimated and then used to describe the structure of the observed pattern. Analyzing a spatial point process, the first step is to learn about its first-order characteristics, studying the relationship of the points with the underlying environmental variables that describe the observed heterogeneity. When the purpose of the analysis is to describe the possible interaction among points, that is, if the given data exhibit spatial inhibition or aggregation, the second-order properties of the process are analysed. However, in the analysis of spatial point process data it can be difficult to disentangle the two previous aspects, i.e., the heterogeneity corresponding to spatial variation of the intensity and the dependence structure amongst the points (Illian et al. 2008; Diggle 2013). For this reason, it is attractive and motivating to define and estimate models that account simultaneously for dependence structure among events, including also the effect of the observed covariates. The spatial seismic process belongs to the class of the point pattern data where the location of events in space are the observations of interest. The aim is typically to learn about the mechanism that generates these events (Møller et al. 1998; Diggle 2013; Illian et al. 2008).

The generator process of earthquakes is quite complex, since it is characterized by multidimensional and multiscale dynamics, dependence features of points both in space and time, spatial anisotropy and spatial variation of the intensity according to the seismogenic sources. Several spatial and spatio-temporal models have been proposed in the literature for describing the evolution of the earthquake process in space, in time and both the dimensions. The temporal clustered structure of earthquake data has been described by self-exciting point processes, i.e., the Hawkes model (Hawkes and Adamopoulos 1973) or the Epidemic Type Aftershock Sequence (ETAS) model (Ogata 1988; Adelfio and Chiodi 2015b). On the other hand, when an inhibitive pattern is observed for earthquake data, self-correcting processes based on the strain-release models (Vere-Jones 1978) are used, assuming the fault fracture generating earthquakes reduces the amount of strain present at the break point along the fault. In the context of spatial point patterns, an important example of self-correcting processes is the Strauss point process model (Strauss 1975) with a negative association, that is also the simplest inhibitive Gibbs point process (Baddeley and Møller 1989). Gibbs point processes are generally applied in the analysis of spatial interaction and inhomogeneity depending on covariate information. For instance, a multitype Strauss process and geological spatial covariates are combined to describe sequences of events in Anwar et al. (2012), Ye et al. (2015). In Siino et al. (2016), Hybrid of Gibbs models (Baddeley et al. 2013) are used to describe the Hellenic seismicity that shows interactions at different spatial scales (multiscale structure), characterising the spatial inhomogeneity of the processes as a function of the geological information available in the study area (the presence of volcanoes, plate boundaries and faults). As for the Cox processes, they are models used to describe environmentally driven processes (Cox 1955; Møller et al. 1998; Møller 2003). Møller and Toftaker (2014) propose a spatial model framework to estimate geometric anisotropy in spatial point process. In particular, they characterise the anisotropy of the main seismicity around Los Angeles (California) estimating an inhomogeneous

log-Gaussian Cox process (LGCP). Furthermore, in the spatio-temporal context, Siino et al. (2018) describe earthquake sequences comparing several Cox model specifications (with separable and non-separable spatio-temporal covariance functions) estimating parameters through the minimum contrast method.

All the aforementioned models can be referred to as ‘global’ models, as they are globally defined and the process properties and estimated parameters are assumed to be constant in all the study area. However, a model with constant parameters, may not represent detailed local variations in the data adequately, since the pattern may present spatial variation in the influence of covariates, in the scale or spacing between points and the abundance of points. Indeed, a different way of analysing a point pattern can be based on local techniques identifying specific and undiscovered local structure, for instance sub-regions characterised by different interactions among points, intensity and influence of covariates.

Throughout the paper we shall distinguish between ‘global’ models, in which the parameters are assumed constant as in regular regression models, and ‘local’ models, in which the parameters are allowed to vary with location.

For earthquake data, Ogata and Katsura (1988), Ogata (1989) used a space-adaptive space-time point process model. In Ogata and Katsura (1988) a method is developed for estimating both the spatial intensity of the point locations and the spatial variation of a characteristic parameter of the distributions for the attached marks, analysing seismological and ecological data. In Ogata (1989) the occurrence rate of earthquakes is extended to include multiple aftershock sequences.

For spatial point process, Baddeley (2017) presents a general framework based on the local composite likelihood to detect and model gradual spatial variation in any parameter of a spatial stochastic model (such as Poisson, Gibbs and Cox processes). In particular, the parameters in the model that govern the intensity, the dependence of the intensity on the covariates and the spatial interaction between points, are estimated locally. Moreover, this approach has the advantage to detect and model spatial variation in any property of a point process, within a formal likelihood framework providing space-varying parameter estimates, confidence intervals and hypothesis tests.

This paper aims at using these recent results for the local composite likelihood for spatial point processes to describe the spatial displacement of earthquake data, accounting for external geological information. We show that models with parameters that vary in space can be suitable for describing both seismic catalogues and sequences.

Seismicity *catalogs* are datasets with statistical units describing earthquakes, and variables consisting in their location, origin time, and magnitude, together with additional metadata such as associated uncertainties and focal mechanism information (Woessner et al. 2010). Earthquakes typically occur in *sequences* that may include foreshocks, the mainshock (the largest event or events), and aftershocks. Aftershocks are smaller earthquakes following the mainshock. They typically occur on or near the rupture plane of the mainshock, resulting from changes of stress and frictional properties of the fault zone caused by the mainshock. The duration of aftershock sequences is typically a few years for earthquakes at plate boundaries, but can last much longer within stable continental interiors (Liu and Stein 2019).

In this paper, we consider two case studies, analysing datasets concerning seismic activity in Greece and Italy.

The Central Mediterranean area is one of the most geodynamically complex areas and active regions of the world. In this framework are embedded the Italian peninsula and the Greek region and its Hellenic arc, which are considered the most tectonically complex and seismically active regions of Mediterranean, object of study in this work. The tectonic evolution of Greek region, is strongly related to the active subduction along the Hellenic arc which is the boundary between the African and the Anatolian-Aegean microplates (Le Pichon and Angelier 1979; Hall et al. 2005). The overall tectonism of the region can be attributed to several events, such as the collision of African and Eurasian plates, convergence between Arabian and Eurasian plates and displacement of the Anatolian-Aegean microplate (Ryan et al. 1982; Taymaz et al. 1991). The stress regime in Hellenic arc region is a multi-component one (Oya 2016). There are subduction zones governing the seismicity in the region, back arc events, and tension areas induced by the interference of continental and oceanic basins (Spakman et al. 1988; Taymaz et al. 1991; Papazachos et al. 1995). The shape and seismicity distribution of the Italian peninsula and Sicily island are consequence of the convergence between the African and Eurasian plates, active since at least 65 Ma and especially on the Quaternary evolution of Central Mediterranean (Amato et al. 1997; Di Stefano et al. 1999; Sgroi et al. 2012). In the Italian peninsula and Sicily, at least 5 tectonic districts can be identified: (1) the Alpine arc, with evidence of past subduction; (2) the North Apennines, where subduction of the oceanic lithosphere was followed by continental collision; (3) the Calabrian arc, with clear evidence of subducting slab of oceanic lithosphere; (4) the Southern Apennines, where a detached slab is revealed by seismicity and seismic tomography; (5) the Sicilian region that represents a portion of the Apennine-Maghrebide fold-and-thrust belt (Amato et al. 1997; Di Stefano et al. 1999; Sgroi et al. 2012).

Hence, when analysing seismic catalogues, the analysis of the dependence of the first-order intensity function on external covariates is a crucial issue. This is accounted properly by the local Poisson model, describing large scale variation by including external covariates and letting their effects vary along the studied area. This is shown with a case study for the description of the seismic events occurred in Greece between 2005 and 2014, an area of high seismic hazard that has been characterised by many destructive earthquakes in the last century. Furthermore, the local version of the Poisson model, fitted with a spatial varying effect of covariates, is compared to a global inhomogeneous Poisson model in terms of fitting and interpretability of results.

Describing seismic sequences, the analysis of the small scale variation is of interest, focusing on the second-order characteristics of the process. Therefore, we introduce and fit local log-Gaussian Cox Process to study a seismic sequence. At the best of our knowledge, our proposal is a novelty for this field of application, and represents a first step for suggesting the use of local complex models for describing the fractal seismic phenomena. The interaction among points is typically taken into account by the Log-Gaussian Cox Processes models through the estimation of the covariance parameters of the Gaussian Random Field. In their local version, these parameters vary in space, allowing the description and characterization of the study

area through a multiple underlying process, meaning that it would make sense to assume a generating process of the observed point pattern that is not unique. The peculiarity of the newly local spatial Log-Gaussian Cox Processes lies in the inhomogeneity of the point process being simultaneously addressed via spatial covariates, a latent error process, and their estimated local coefficients. Therefore, after providing an application of the local Poisson models to the Greek seismic catalogue, and identifying the regions that display the most inhomogeneous behaviour, a local Log-Gaussian Cox Process model is fitted to describe a selected Greek seismic sequence, with space varying effects of both intensity and interaction parameters. Indeed, one crucial issue in local analysis of spatial as well as spatio-temporal point patterns, is the identification of subregions where points behave differently in terms of clustering. In addition, based on the assumptions on the underlying processes, local estimates provide further insight into the relationship of the intensity of the process on external covariates, as well as on the correlation structure, that in this setting is allowed to vary with location. Local models are also applied to the Italian data, since as Greece, Italy is an area of high seismic risk and in particular, Central Italy, where our analysis focuses, has been characterised by a strong series of earthquakes between 2009 and 2017. The results are reported in “Appendix 1”, for the sake of brevity. The key assumption stating that for each spatial location there is an unobserved spatial subdomain inside which the template model is exactly true, allow us to exploit the statistical properties of the spatially varying parameters, as they were the ones of a model with constant parameters in that specific subregion. This feature is particularly appealing as it allows us to build diagnostic tools and interpret parameters as we are used in classical log-linear models. Besides this, we also test on the two case studies that local models provide good inferential results, if compared to the semi-parametric competitors.

Moreover, this work represents a contribution to the framework of diagnostics for local models.

Indeed, when a model is fitted to a set of random points, diagnostic measures are necessary to assess the goodness-of-fit and to evaluate the ability of that model to describe the random point pattern behaviour (Adelfio et al. 2019). However, tools for checking or criticizing the fitted model are quite limited. There is currently no analogue for spatial point patterns of the comprehensive strategy for model criticism in the linear model, which uses tools such as residual plots and influence diagnostics to identify unusual or influential observations, for assessing the model assumptions sequentially and to individuate the any departures from the model (Baddeley et al. 2005).

It is a widespread practice in the statistical analysis of spatial point pattern data to focus primarily on comparing the data with a homogeneous Poisson process (‘complete spatial randomness’), which is generally the null model in applications, rather than the fitted model. This approach considers a stationary Poisson residual process by randomly rescaling (Meyer 1971; Schoenberg 1999) or thinning (Schoenberg 2003), and investigates whether the second-order properties of the observed residuals are consistent with those of the stationary Poisson process.

An alternative approach is to define a weighted second-order statistic, where essentially to each observed point a weight inversely proportional to the

conditional intensity at that point is given. This method was adopted by Veen and Schoenberg (2006) in constructing a weighted version of the spatial K -function of Ripley (1977) and Veen (2006), although the spatial weighted analogue of Ripley's K -function which was first introduced by Baddeley and Turner (2000).

Therefore, establishing a coherent strategy for model criticism in spatial point process models, resembling the existing methods for the linear model, depends crucially on finding the right definition of residuals for a spatial point process model fitted to point pattern data.

In this paper, we propose the use of some diagnostics tools generally considered for global point processes, e.g. the smoothed raw residuals and the inhomogeneous K -function (Baddeley et al. 2005), to the local context, and properly review local tests and the test of departure from 'homogeneity' (Baddeley 2017), meaning that the observed pattern is generated by a template model with constant parameters. Our contribution concerns a bootstrap procedure, carried out to support the model selection process, retaining also local information. Furthermore, a stepwise procedure is also developed, assessing the single local effects of the considered covariates on the rejection of the given 'homogeneity' hypothesis. The method provides a ranking of the covariates that most contribute to the local features, together with p -values maps, useful for progressively identifying the variables' effects within the study area.

Therefore, this paper contributes to the framework of the local spatial processes, in two main directions: providing a formalization and characterization of local LGCP for earthquake data and contributing to diagnostic methods for this class of models.

This paper is structured as follows. A review of spatial point processes and log-linear Poisson models, in terms of both global and local methodology, is reported in Sect. 2. In Sect. 3 spatial local log-Gaussian Cox processes are introduced and in Sect. 4 their estimation procedure is discussed. Section 5 is dedicated to diagnostics methods for local models. In particular, Sect. 5.1 contains global diagnostics tools, well established in spatial point pattern analysis, and suitable for local models. In Sect. 5.2, we propose two local diagnostics procedures, based on the test of local departure from 'homogeneity' testing (Baddeley 2017). In Sect. 6, local models are applied to describe seismic activity in Greece, describing both the entire Greek seismic catalogue and a particular seismic sequence. Section 7 is devoted to conclusions and final remarks.

2 Local spatial Poisson Processes

2.1 Spatial point processes

Following Cressie (2015), we introduce point processes by a mathematical approach that uses the definition of a counting measure on a set $X \subseteq \mathbb{R}^d$, $d \geq 1$, with positive values in \mathbb{Z} : for each Borel set B this \mathbb{Z}_+ -valued random measure gives the number of events falling in B .

Definition 1 *Point process* Let (Ω, \mathcal{A}, P) be a probability space and Φ a collection of locally finite counting measures on $X \subset \mathbb{R}^d$. Define \mathcal{X} as the Borel σ -algebra of X and let \mathcal{N} be the smallest σ -algebra on Φ , generated by sets of the form $\{\phi \in \Phi : \phi(B) = n\}$ for all $B \in \mathcal{X}$. A point process N on X is a measurable mapping of (Ω, \mathcal{X}) into (Φ, \mathcal{N}) . A point process defined on (Ω, \mathcal{A}, P) induces a probability measure $\Pi_N(Y) = P(N \in Y), \forall Y \in \mathcal{N}$.

Then, for any set $B \in \mathcal{X}$, $N(B)$ represents the number of points falling in B , such that if B is the union of disjoint sets $\tilde{B}_1, \tilde{B}_2, \dots$, then $N(B) = \sum N(\tilde{B}_i)$. A spatial point pattern N is an unordered set $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ of points \mathbf{x}_i where $n(\mathbf{x}) = n$ denotes the number of points, not fixed in advance. If \mathbf{x} is a point pattern and $D \subset \mathbb{R}^2$ is a bounded region, we write $\mathbf{x} \cap D$ for the subset of \mathbf{x} consisting of points that fall in D and $n(\mathbf{x} \cap D)$ for denoting the number of points of \mathbf{x} falling in D . A point process model assumes that \mathbf{x} is a realization of a finite point process N in D without multiple points. A point location in the plane is denoted by a lower case letter like \mathbf{u} . Any location \mathbf{u} can be specified by its Cartesian coordinates $\mathbf{u} = (u_1, u_2)$ in such a way that we do not need to mention the coordinates explicitly. The first-order property of N is described by the intensity function, defined as

$$\lambda(\mathbf{u}) = \lim_{|\mathbf{du}| \rightarrow 0} \frac{\mathbb{E}[N(\mathbf{du})]}{|\mathbf{du}|}$$

where \mathbf{du} is an infinitesimal region that contains the point $\mathbf{u} \in D$, $|\mathbf{du}|$ is its area and $\mathbb{E}[N(\mathbf{du})]$ denotes the expected number of events in \mathbf{du} . When the intensity is constant the process is called homogeneous. In the inhomogeneous case, the intensity is not constant in the study area but may depend, for instance, on the coordinates of points. A point process model, assuming independence, is completely described by its intensity function $\lambda(\mathbf{u})$ (Daley and Vere-Jones 2007). The inter-point interaction between events is measured by the second-order moment characteristics, studied through the second-order intensity function, defined as

$$\lambda_2(\mathbf{u}, \mathbf{v}) = \lim_{|\mathbf{du}||\mathbf{dv}| \rightarrow 0} \frac{\mathbb{E}[N(\mathbf{du})N(\mathbf{dv})]}{|\mathbf{du}||\mathbf{dv}|}$$

where $\mathbf{u} \neq \mathbf{v}$. Refer to Daley and Vere-Jones (2007), Preposition 3.3.I, for existence and convergence property of the point processes intensity.

Several functional summary statistics are used to study the second-order characteristics of a point pattern and to measure dependence.

A general spatial log-linear Poisson model (Cox 1972), generalizes both homogeneous and inhomogeneous models, such that:

$$\lambda(\mathbf{u}) = \lambda(\mathbf{u}, \theta) = \exp(B(\mathbf{u}) + \theta^T \mathbf{Z}(\mathbf{u})), \tag{1}$$

where $\mathbf{u} \in D$, $B(\mathbf{u})$ and $\mathbf{Z}(\mathbf{u}) = (Z_1(\mathbf{u}), \dots, Z_p(\mathbf{u}))$ are known functions (covariates), $\theta^T = (\theta_1, \dots, \theta_p)$ are unknown parameters, and therefore $\theta^T \mathbf{Z}(\mathbf{u}) = (\theta_1 Z_1(\mathbf{u}) + \dots + \theta_p Z_p(\mathbf{u}))$. These models have an especially convenient structure, since the log intensity is a linear function of the parameters and covariates

can be quite general functions, making them a very wide class of models. The estimation of the point process parameters is carried out through the maximization of the log-likelihood, defined by:

$$\log L(\theta) = \sum_i \log \lambda(\mathbf{x}_i; \theta) - \int_D \lambda(\mathbf{u}; \theta) d\mathbf{u} \tag{2}$$

where the sum is over all point \mathbf{x}_i in the point process \mathbf{x} (Daley and Vere-Jones 2007).

2.2 Local composite likelihood for Poisson processes

In this section, we report some basic definitions and results about the local composite likelihood applied to Poisson point processes, following the notation used in Baddeley (2017). In the local context, given a spatial point pattern \mathbf{x} , realization of a finite point process N_1 in D , the *template model* is defined as a point process N_2 , governed by a parameter $\theta \in \Theta \subseteq \mathbb{R}^p$. N_2 is defined as the “template” model rather than the “homogeneous” one, because the potential dependence of the distribution of N_2 on some spatial covariates would make N_2 a spatially inhomogeneous point process. The template model is assumed to provide good local approximations to N_1 , in the sense that, in any small region $B \subset D$, the distribution of $N_1 \cap B$ is well-approximated by the distribution of $N_2 \cap B$ of parameter θ which may depend on B .

When spatial log-linear models are used, the model parameters are usually assumed to be constant across the entire study region. This assumption may be too simplistic in contexts that are characterised by multi-scale and fractal features, like the seismic one, where the relationship between the intensity of earthquakes and other possible characteristics of the area where events occur. In this context, the features can be described by a Poisson process with intensity

$$\lambda(\mathbf{u}) = \lambda(\mathbf{u}; \theta(\mathbf{u})) = \exp(B(\mathbf{u}) + \theta^T(\mathbf{u})\mathbf{Z}(\mathbf{u})) \tag{3}$$

with local coefficients, where $\theta(\cdot)$ is a function of the spatial location. This is the “inhomogeneous” alternative to the template log-linear intensity (1). For estimating local models as in Eq. (3), the likelihood at \mathbf{u} is used, being a weighted version of the original likelihood in Eq. (2), with the greatest weight on locations close to \mathbf{u} .

Therefore, in the local context, the local log-likelihood associated with location \mathbf{s} (Loader et al. 1999) is:

$$\log L(\mathbf{s}; \theta) = \sum_{i=1}^n w_h(\mathbf{x}_i - \mathbf{s}) \log \lambda(\mathbf{x}_i; \theta) - \int_D \lambda(\mathbf{u}; \theta) w_h(\mathbf{u} - \mathbf{s}) d\mathbf{u} \tag{4}$$

where $w_h(\mathbf{u}) = h^{-d}w(\mathbf{u}/h)$ is a weight nonparametric function, and $h > 0$ is a smoothing bandwidth. It is not necessary to assume that w_h is a probability density. Throughout the paper, we will consider a kernel of fixed bandwidth h . In the local likelihood in Eq. (4) this is usually chosen by the cross-validation criterion (Loader et al. 1999). The optimal bandwidth h_{opt} maximises

$$\text{LCV}(h) = \sum_i \log \lambda(\mathbf{x}_i; \hat{\theta}_i(\mathbf{x}_i)) - \int_D \lambda(\mathbf{u}; \hat{\theta}(\mathbf{u})) d\mathbf{u} \quad (5)$$

where $\hat{\theta}(\mathbf{u}) = \hat{\theta}(\mathbf{u}, h)$ is the local estimate of θ at location \mathbf{u} using bandwidth $h > 0$, and $\hat{\theta}_i(\mathbf{x}_i) = \hat{\theta}_i(\mathbf{x}_i; h)$ is the corresponding ‘leave-one-out’ estimate at the location \mathbf{x}_i computed from the data omitting \mathbf{x}_i .

Maximizing the local likelihood for each fixed \mathbf{u} gives local parameter estimates $\hat{\theta}(\mathbf{u})$ of the original model, confidence intervals, hypothesis tests and other standard tools (Baddeley 2017). These fitted coefficients can be plotted as a function of spatial location \mathbf{u} . Also the fitted intensity can be obtained as a function of \mathbf{u} , as $\hat{\lambda}(\mathbf{u}) = \lambda(\mathbf{u}; \hat{\theta}(\mathbf{u}))$. It is worth to recall that the local likelihood approach assumes that the template model is a good local approximation to the data. On the one hand, this allows to use existing statistical techniques and theory for the template model to be easily modified in local estimation and inference. On the other hand, there is no explicit description of the “true” model, but only of the template model, which is the local approximation of the true one.

The key assumption is that, for each spatial location $\mathbf{u} \in D$, there is an unobserved spatial subdomain $D(\mathbf{u})$ containing \mathbf{u} where the template model is *exactly* true and parameter $\theta(\mathbf{u})$. If the support of the kernel centred at \mathbf{u} falls entirely inside $D(\mathbf{u})$, then the statistical properties of $\hat{\theta}(\mathbf{u})$ are the same as they would be if the entire process followed the template model with constant parameter value $\theta = \theta(\mathbf{u})$ (Baddeley 2017).

3 Local spatial log-Gaussian Cox processes

The main feature of the Poisson point process is independence and stationarity, that is a theoretical assumption often inappropriate for describing real data. However, considering more general point process models may be a complex issue in terms of inference and diagnostics (Adelfio and Schoenberg 2009; Adelfio and Chiodi 2015a), but necessary if dealing with real contexts where dependence structure among data is of main interest (Siino et al. 2016, 2018). Two large classes of alternative models for spatial point processes are the Gibbs and Cox processes (Van Lieshout 2000; Møller 2003; Illian et al. 2008). For clustered point patterns, the most commonly-used models are Cox and cluster processes (Baddeley et al. 2015, Chapter 12).

The Cox processes are a generalisation of the inhomogeneous Poisson processes, where the intensity is a realisation of a random field. For this reason, they are also called doubly stochastic Poisson processes. The point process N is said to be a Cox process driven by Λ , if the conditional distribution of the point process N given a realisation $\Lambda(\mathbf{u}) = \lambda(\mathbf{u})$ is a Poisson process on D with intensity function $\lambda(\mathbf{u})$. The most used Cox point processes are the shot-noise Cox processes (Møller 2003) and the log-Gaussian Cox processes (Møller et al. 1998). In a shot-noise Cox process, offsprings are observed around unseen parents representing the random unobservable variability.

Following the inhomogeneous specification in Diggle et al. (2013), the log-Gaussian Cox process for a generic point in space has the following intensity

$$\Lambda(\mathbf{u}) = \lambda(\mathbf{u}) \exp(S(\mathbf{u})) \quad (6)$$

where S is a Gaussian Random Field (GRF) with mean function $\mu = \mathbb{E}(S(\mathbf{u})) = -\frac{\sigma^2}{2}$ so that $\mathbb{E}[\exp(S(\mathbf{u}))] = 1$. The covariance function is $\mathbb{C}(S(\mathbf{u}), S(\mathbf{v})) = \mathbb{C}(\|\mathbf{u} - \mathbf{v}\|) = \sigma^2 \rho(r)$ under the stationary assumption, i.e. only depends on the variance parameter σ^2 , and distance $r = \|\mathbf{u} - \mathbf{v}\|$ between locations \mathbf{u} and \mathbf{v} .

$\rho(\cdot)$ is the correlation function of the GRF, completely specified by its first and second moments. Following the notation used in Baddeley et al. (2015), the random intensity $\Lambda(\mathbf{u}) = \exp(S(\mathbf{u}))$ is a log-Gaussian or log-normal distribution. If $Z \sim N(\mu, \sigma^2)$ and $\Lambda = \exp(Z)$ then Λ has mean $\exp(\mu + \sigma^2/2)$. In this paper, we assume the exponential structure for the covariance function as in Brix and Diggle (2001),

$$\mathbb{C}(r) = \sigma^2 \exp\left(-\frac{r}{\alpha}\right) \quad (7)$$

depending only on α and σ^2 that represent the scale parameter for the spatial distance and the variance, respectively. The effect of increasing σ^2 is to generate higher peaks in the surface intensity which leads to clusters of points. Increasing the spatial scale parameter α , the underlying GRF presents a strong spatial correlation and it corresponds to a diffuse aggregation of points of the LGCP. Other basic stationary and isotropic family models commonly used for the spatial component are the Matérn, Cauchy, Gaussian and Spherical covariance functions (Gelfand et al. 2010). In a spatial log-Gaussian Cox process, the intensity model can be expressed as a function of further covariates, according to a loglinear model as in Eq. (1). We know that $\lambda(\mathbf{u}) = \mathbb{E}(\Lambda(\mathbf{u})) = \mathbb{E}(\exp(S(\mathbf{u}))) = \exp((\mu(\mathbf{u}) + \sigma^2/2))$, where

$$\mu(\mathbf{u}) = -\frac{\sigma^2}{2} + B(\mathbf{u}) + \theta^T \mathbf{Z}(\mathbf{u}) \quad (8)$$

is the mean function of the Gaussian random field. This peculiarity, together with the possibility of detecting the clustered structure of the point process, makes the LGCP model attractive for modelling seismic data.

In this paper, we formalize the characteristics of the local version of the LGCP model, allowing also the additional parameters σ and α to vary in space (denoted by $\sigma(\mathbf{u})$ and $\alpha(\mathbf{u})$ throughout the rest of the paper), as well as the effects of the considered covariates, according to a loglinear model as in Eq. (3). Recalling the specification of the global LGCP, we know that now, as the interaction parameters can vary in space, the GRF and its moments will also depend on the locations \mathbf{u} . In other words, for each location \mathbf{u} we obtain a local GRF. Basically, while the global LGCP implies a unique generating process, the local LGCP implies that events are generated from multiple processes, i.e. with different correlation patterns, that can yet be estimated through a unique model. Therefore, the local LGCP represents an interesting model to analyse and describe those observed regions for which it is reasonable to assume

different generating GRFs, characterized by different covariances. Furthermore, this implies that a local version of a LGCP model can be useful to spot regions with different correlation structures, and the mean of the local process in Eq. (8), also depends on the space varying parameters $\theta^T(\mathbf{u})$ and $\sigma^2(\mathbf{u})$. Therefore, if in the global model the mean of the process is influenced just by the linear predictor and by the constant σ^2 , in the local LGCP, the mean varies also because of the local variation of $\theta^T(\mathbf{u})$ and $\sigma^2(\mathbf{u})$ parameters. Then, the covariance function, assumed in Eq. (7), now depends on the local variance parameter $\sigma(\mathbf{u})$ and spatial scale parameter $\alpha(\mathbf{u})$. A visual displaying of the space varying estimates of the interaction parameters may suggest the presence of a space varying clustered structure.

4 Model estimation

In this Section, the model estimation procedure based on the local version of the Palm likelihood of Ogata and Katsura (1991), and developed by Baddeley (2017), is reported. The Cox models are usually estimated by a two-step procedure, involving both the intensity and the cluster or correlation parameters. In the first step, a Poisson model is fitted to the point pattern data, providing the estimates of the coefficients of all the terms in the model formula characterizing the intensity, as the one in the most general log-linear model in Eq. (1). This intensity is then considered as the true one in the second step, and the cluster or correlation parameters are estimated by either the method of minimum contrast (Pfanzagl 1969; Eguchi 1983; Diggle 1979; Diggle and Gratton 1984; Siino et al. 2018), Palm likelihood (Ogata and Katsura 1991; Tanaka et al. 2008) or composite likelihood (Guan 2006). Hereafter we will denote by θ the vector of (first-order) intensity parameters, and by ψ the cluster parameters, also denoted as correlation or interaction parameters by some authors. In the case of a spatial Log-Gaussian Cox process with exponential covariance as the one in (7), the cluster parameters correspond to $\psi = (\alpha, \sigma)$.

In order to estimate the local version of the LGCP models, we refer to the adaptation of the Palm likelihood to a general nonstationary point process, considered by Baddeley (2017) because of its formal similarity to the Poisson likelihood. Ogata and Katsura (1991) developed a surrogate likelihood function called Palm likelihood for the analysis of stationary point processes.

The Palm distribution formalises the concept of conditioning on a point of the process. Assume that the process has the intensity function $\lambda(\mathbf{u}; \phi)$ for $\mathbf{u} \in D$ and second moment intensity $\lambda_2(\mathbf{u}, \mathbf{v}; \phi)$ for $\mathbf{u}, \mathbf{v} \in D$, depending on parameters $\phi \in \Phi$. Define the intensity of the Palm distribution at \mathbf{u} as

$$\lambda_p(\mathbf{u}|\mathbf{v}; \phi) = \lambda_2(\mathbf{u}, \mathbf{v}; \phi) / \lambda(\mathbf{u}; \phi). \tag{9}$$

The Palm log-likelihood can be written as

$$\log \text{PAL}(\phi) = \sum_i \left[\sum_{j \neq i} Q(\mathbf{x}_i, \mathbf{x}_j) \log \lambda_p(\mathbf{x}_j | \mathbf{x}_i; \phi) - \int_D Q(\mathbf{x}_i, \mathbf{u}) \lambda_p(\mathbf{u} | \mathbf{x}_i; \phi) d\mathbf{u} \right] \tag{10}$$

where $Q(\mathbf{u}, \mathbf{v})$ is a 0/1 valued constraint function, designed to simplify computation and optimise statistical properties, typically taken to be $Q(\mathbf{u}, \mathbf{v}) = I\{\|\mathbf{u} - \mathbf{v}\| \leq R\}$ where $I\{\cdot\}$ is the indicator function, such that $I\{x\} = 1$ if x is true, and $R > 0$ is an upper bound to the correlation distance of the model. The quantity in square braces is recognisable as the log-likelihood of a Poisson point process with intensity $\lambda_p(\cdot|\mathbf{x}_i; \boldsymbol{\phi})$ restricted to the set where $Q(\mathbf{x}_i, \cdot) = 1$.

Dealing with the Cox processes, it is usual to work with models in which the intensity and interaction parameters are ‘separable’

$$\lambda_2(\mathbf{u}, \mathbf{v}; \boldsymbol{\phi}) = \lambda(\mathbf{u}; \boldsymbol{\theta})\lambda(\mathbf{v}; \boldsymbol{\theta})\rho(\mathbf{u}, \mathbf{v}; \boldsymbol{\psi}) \tag{11}$$

where $\boldsymbol{\phi} = (\boldsymbol{\theta}, \boldsymbol{\psi})$, with $\boldsymbol{\theta}$ representing the intensity parameters and $\boldsymbol{\psi}$ the interaction parameters, and ρ is the correlation function. As stated by Baddeley (2017), examples include log-Gaussian Cox processes and certain inhomogeneous Neyman–Scott processes (Waagepetersen 2007; Waagepetersen and Guan 2009). Recalling what previously introduced, specifying a local Log-Gaussian Cox process, $\boldsymbol{\theta}$ is the vector of parameters representing the effects of the covariates included in the intensity, while $\boldsymbol{\psi}$ is the vector of the interaction parameters σ and α , representing the spatial scale parameter and the variance of the underlying GRF.

The Palm log-likelihood (10) involves summation and integration over pairs of spatial points. To define a local version of the Palm likelihood it is possible to apply local weights to the second element of each pair, that is, the weights are applied just to the points for which ‘predictions’ are made. Summing over the first element \mathbf{x}_i , we get:

$$\begin{aligned} \log \text{PAL}(\mathbf{s}; \boldsymbol{\phi}) &= \sum_i \sum_{j \neq i} w_h(\mathbf{x}_j - \mathbf{s})Q(\mathbf{x}_i, \mathbf{x}_j) \log \lambda_p(\mathbf{x}_j|\mathbf{x}_i; \boldsymbol{\phi}) \\ &\quad - \sum_i \int_D w_h(\mathbf{u} - \mathbf{s})Q(\mathbf{x}_i, \mathbf{u})\lambda_p(\mathbf{u}|\mathbf{x}_i; \boldsymbol{\phi})d\mathbf{u} \end{aligned}$$

that is basically the same Eq. (10) obtained replacing $Q(\mathbf{u}, \mathbf{v})$ by $w_h(\mathbf{v} - \mathbf{s})Q(\mathbf{u}, \mathbf{v})$. This must be maximised numerically. We assume a separable isotropic model as in Eq. (11). Recalling Eq. (9), then we have

$$\lambda_p(\mathbf{v}|\mathbf{u}; \boldsymbol{\phi}) = \lambda(\mathbf{v}; \boldsymbol{\theta})\rho(\|\mathbf{u} - \mathbf{v}\|; \boldsymbol{\psi})$$

so that the local Palm log-likelihood to be maximised is

$$\begin{aligned} \log \text{PAL}(\mathbf{s}; \boldsymbol{\phi}) &= \sum_i \sum_{j \neq i} w_h(\mathbf{s} - \mathbf{x}_i)Q(\mathbf{x}_i, \mathbf{x}_j) \{ \log \lambda(\mathbf{x}_j; \boldsymbol{\psi}) + \log \rho(\|\mathbf{x}_i - \mathbf{x}_j\|; \boldsymbol{\theta}) \} \\ &\quad - \sum_i \int_D w_h(\mathbf{s} - \mathbf{u})Q(\mathbf{x}_i, \mathbf{u})\lambda(\mathbf{u}; \boldsymbol{\psi})\rho(\|\mathbf{x}_i - \mathbf{x}_j\|; \boldsymbol{\theta})d\mathbf{u}. \end{aligned} \tag{12}$$

A two step estimation procedure is used, as follows:

- The intensity parameters θ are estimated by maximising the Palm likelihood of the Poisson process with intensity $\lambda(\cdot; \theta)$, that is equivalent to a weighted Poisson likelihood.
- The interaction parameters ψ are estimated separately at each location \mathbf{u} by maximising the local Palm likelihood (12) using the simplex algorithm.

If the Palm likelihood is taken in the form (10), the natural analogue of the Poisson likelihood cross-validation criterion (5) for Palm likelihood is

$$\text{LCV}(h) = \sum_i \sum_{j \neq i} Q(\mathbf{x}_i, \mathbf{x}_j) \log \lambda_p(\mathbf{x}_j | \mathbf{x}_i; \hat{\phi}_{-j}(\mathbf{x}_j)) - \sum_i \int_D Q(\mathbf{x}_i, \mathbf{u}) \lambda_p(\mathbf{u} | \mathbf{x}_i; \hat{\phi}(\mathbf{u})) d\mathbf{u} \quad (13)$$

for bandwidth h , where $\hat{\phi}_{-j}(\mathbf{x}_j)$ is the ‘leave-one-out’ estimate, that is, the local parameter estimate at $\mathbf{s} = \mathbf{x}_j$ based on the data $\mathbf{x} \setminus \mathbf{x}_j$.

Alternative approaches for fitting local Cox and cluster processes include the methods of minimum contrast (Pfanzagl 1969; Diggle and Gratton 1984) and model-based clustering (Banfield and Raftery 1993; Dasgupta and Raftery 1998; Walsh and Raftery 2002, 2005). In particular, a local version of the minimum contrast approach is developed using the local K-functions or the local pair correlation functions by Baddeley (2017), bearing a very close resemblance to the local Palm likelihood approach.

5 Diagnostics

This section outlines relevant methods for carrying out diagnostics for local models. A particular focus is given to the model selection, dealing with multiple covariates. First, some well-known methods for diagnostics are reviewed and proposed as model selection approaches in the context of local models. These basically concern intensity-based methods. After that, we focus on local tests, that are specific diagnostic tools for models with space varying parameters, testing the significance of parameters and the properness of a local model. Finally, we propose a stepwise procedure to identify the spatial covariates that most contribute to the choice of the local model, that is the covariates whose effects can be correctly considered as varying in space.

5.1 Global diagnostics methods for local model selection

This section reviews some methods widely used for diagnostics of global spatial point processes models, here proposed as diagnostic methods for spatial models in the local framework. As stated in some previous papers (Adelfio et al. 2019; Adelfio and Schoenberg 2009), the main problem when dealing with residual analysis for point processes is to find a correct definition of residuals, since the one used in dependence models cannot be used for point processes. Two of the mostly used methods for diagnostics of spatial point processes are the inhomogeneous K-function

and the smoothed raw residuals, here reported. A widely used summary statistics, for descriptive analysis and diagnostics, is the Ripley's K -function (Ripley 1976, 1988), which is defined as

$$K(r) = 2\pi\lambda^{-2} \int_0^r \lambda_2(x)xdx,$$

for a stationary and isotropic process for which $\lambda_2(r) = \lambda^2$ (Schabenberger and Gotway 2017). It is also known as the reduced second moment measure (Cressie 2015), as the second reduced moment function (Chiu et al. 2013), and as the second order reduced moment measure (Møller 2003).

Given that in a stationary process the distribution of N is the same as the distribution of the shifted process $N + \mathbf{v}$, for any vector \mathbf{v} , the K -function is defined as

$$K(r) = \frac{1}{\lambda} \mathbb{E} [\text{number of } r\text{-neighbours of } \mathbf{u} | N \text{ has a point at location } \mathbf{u}]$$

for any distance $r \geq 0$ and any location \mathbf{u} . Being a measure of the distribution of the inter-point distances, $K(r)$ captures the spatial dependence between different regions of a point process. Under the homogeneous Poisson assumption the following theoretical relation holds $K_{pois}(r) = \pi r^2$. When the process is stationary, the most commonly used estimator is introduced in Ripley (1976):

$$\hat{K}(r) = \frac{|D|}{n(n-1)} \sum_{i=1}^n \sum_{i=1, i \neq j}^n I\{d_{ij} \leq r\} e(\mathbf{x}_i, \mathbf{x}_j, r) \tag{14}$$

where $I\{\cdot\}$ is the indicator function, n is the number of points in the pattern, $|D|$ is the area of the window, $e(\mathbf{x}_i, \mathbf{x}_j, r)$ is the Ripley's edge correction and $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$ is the pairwise distances between all distinct pairs of points \mathbf{x}_i and \mathbf{x}_j in the pattern. When a process has $\hat{K}(r) < \pi r^2$, it means that the points tend to cluster, and when $\hat{K}(r) > \pi r^2$, it means that the points tend to inhibit each other.

Let $\lambda(\mathbf{x}_i)$ be the intensity function characterizing the generating process N , each point \mathbf{x}_i can be weighted by the inverse of the intensity function, i.e. $1/\lambda(\mathbf{x}_i)$, the reciprocal of the intensity at \mathbf{x}_i , and each pair of points $\mathbf{x}_i, \mathbf{x}_j$ will be weighted by $1/(\lambda(\mathbf{x}_i)\lambda(\mathbf{x}_j))$. The inhomogeneous K -function is defined as:

$$K_{inhom}(r) = \mathbb{E} \left[\sum_{\mathbf{x}_i \in \mathbf{X}} \frac{1}{\lambda(\mathbf{x}_j)} I\{0 < \|\mathbf{x}_i - \mathbf{x}_j\| \leq r\} \mid \mathbf{x}_i \in \mathbf{X} \right]$$

assuming that this does not depend on the location \mathbf{x}_i (Baddeley et al. 2000). The standard estimator of the K -function in Eq. (14) can be extended to the inhomogeneous K -function as follows:

$$\hat{K}_{inhom}(r) = \frac{1}{D} \sum_i \sum_{j \neq i} \frac{I\{\|\mathbf{x}_i - \mathbf{x}_j\| \leq r\}}{\hat{\lambda}(\mathbf{x}_i)\hat{\lambda}(\mathbf{x}_j)} e(\mathbf{x}_i, \mathbf{x}_j; r) \tag{15}$$

where $e(\mathbf{x}_i, \mathbf{x}_j, r)$ is an edge correction weight as before, and $\hat{\lambda}(\mathbf{u})$ is an estimate of the intensity function $\lambda(\mathbf{u})$. If $\hat{\lambda}(\mathbf{u})$ is the intensity estimated through a fitted model, then the inhomogeneous version of the K -function can be used as a diagnostic tool.

The inhomogeneous K -function is useful for interpreting the local features of data, since if the estimated intensity $\hat{\lambda}(\mathbf{u})$ is close to the generating one $\lambda(\mathbf{u})$, the estimated $\hat{K}_{\text{inhom}}(r)$ should behave like the corresponding function under a Poisson model. Moreover, $\hat{K}_{\text{inhom}}(r)$ values greater than the expected value under the Poisson model $K_{\text{inhom,pois}}(r) = \pi r^2$ (exactly as for the homogeneous case) indicate that the fitted model is not appropriate, since the distances computed among points exceed the Poisson theoretical ones, suggesting those locations where departures are more evident. In practice, given the fitted competitive models, the model with the estimated inhomogeneous K -function which most resembles the theoretical Poisson one, has to be preferred.

For an inhomogeneous Poisson process model, with fitted intensity $\hat{\lambda}(\mathbf{u})$, the predicted number of points falling in any region D is $\int_D \hat{\lambda}(\mathbf{u}) d\mathbf{u}$. Hence, the residual in each region $D \subset \mathbb{R}^2$ is the ‘observed minus predicted’ number of points falling in D (Alm 1998), that is $R(D) = n(\mathbf{x} \cap D) - \int_D \hat{\lambda}(\mathbf{u}) d\mathbf{u}$, where \mathbf{x} is the observed point pattern, $n(\mathbf{x} \cap D)$ the number of points of \mathbf{x} in the region D , and $\hat{\lambda}(\mathbf{u})$ is the intensity of the fitted model. A simple residuals visualization can be obtained by smoothing them. The ‘smoothed residual fields’ are defined as

$$s(\mathbf{u}) = \tilde{\lambda}(\mathbf{u}) - \lambda^\dagger(\mathbf{u}) \quad (16)$$

where $\tilde{\lambda}(\mathbf{u}) = e(\mathbf{u}) \sum_{i=1}^{n(\mathbf{x})} \kappa(\mathbf{u} - \mathbf{x}_i)$ is the nonparametric, kernel estimate of the fitted intensity $\hat{\lambda}(\mathbf{u})$, while $\lambda^\dagger(\mathbf{u})$ is a correspondingly-smoothed version of the (typically parametric) estimate of the intensity of the fitted model, $\lambda^\dagger(\mathbf{u}) = e(\mathbf{u}) \int_W \kappa(\mathbf{u} - \mathbf{v}) \hat{\lambda}(\mathbf{v}) d\mathbf{v}$. Here, κ is the smoothing kernel and $e(\mathbf{u})$ is the edge correction. The smoothing bandwidth for the kernel estimation of the raw residuals is selected by cross-validation, as the value that minimises the Mean Squared Error criterion defined by Diggle (1985), by the method of Berman and Diggle (1989). See Diggle (2013) for further details. The difference in Eq. (16) should be approximately zero when the fitted model is close to the real one. Therefore, the best model is the one with the lowest values of the smoothed raw residuals.

In the next sections, these residuals are used for diagnostics, both for their flexibility and their simplicity of implementation, and the complexity of model estimation when the parameters vary on the space domain.

5.2 Local diagnostics methods: testing the local departure from ‘homogeneity’

In this section, we introduce local methods for local models diagnostics, namely the tests of local departure from the ‘homogeneity’ assumption, as proposed in Baddeley (2017). For further confuting this hypothesis, still retaining local features, and for assessing the contribution of each covariate to the rejection of the null hypothesis, dealing with multiple linear predictor, we propose a bootstrap and step-wise procedure, in Sects. 5.2.1 and 5.2.2, respectively.

Let us introduce the following system:

$$\begin{cases} \mathcal{H}_0 : \theta \in \Theta_0 \\ \mathcal{H}_1 : \theta \in \Theta \end{cases}$$

where $\Theta_0 \subset \Theta$, a 'local test' can be introduced assessing for parameters dependence on the spatial location $\mathbf{u} \in D$. In the simplest case, suppose $\theta = (\theta_1, \dots, \theta_p)$ and assume that we want to verify

$$\begin{cases} \mathcal{H}_0 : \theta_j = 0 \\ \mathcal{H}_1 : \theta_j \neq 0 \end{cases}$$

The local version of the Wald test can be computed to assess the significance of the spatial covariates effects as a function of \mathbf{u} . This test is based on the standardised local coefficient estimate or 't-statistic'

$$t_j(\mathbf{u}) = \hat{\theta}_j(\mathbf{u})/se(\hat{\theta}_j(\mathbf{u})), \quad (17)$$

whose asymptotic null distribution is the standard Normal. For asymptotic properties of local tests refer to Baddeley (2017).

Local tests can be also used for clusters detection, defining a formal hypothesis test of local departure from 'homogeneity'. In this context, the pattern is considered 'homogeneous' if it is generated by the template model, called 'global model', with a constant parameter vector θ . Therefore, let us consider the null hypothesis \mathcal{H}_0 of a Poisson process with the template log-linear global model as in Eq. (1). The 'inhomogeneous' alternative \mathcal{H}_1 is a Poisson process with the intensity function as in Eq. (3), called 'local model', with $\theta(\cdot)$ function of the spatial location. So the test of local departure from 'homogeneity' refers to the following hypotheses system:

$$\begin{cases} \mathcal{H}_0 : \lambda(\mathbf{u}) = \lambda(\mathbf{u}, \theta) = \exp(B(\mathbf{u}) + \theta^T \mathbf{Z}(\mathbf{u})) \\ \mathcal{H}_1 : \lambda(\mathbf{u}) = \lambda(\mathbf{u}; \theta(\mathbf{u})) = \exp(B(\mathbf{u}) + \theta^T(\mathbf{u}) \mathbf{Z}(\mathbf{u})) \end{cases}$$

To assess this hypothesis, a Monte Carlo test can be carried out for a locally-fitted Poisson point process model, using either the local likelihood ratio test statistic or, as in this paper, the local score test statistic:

$$T^2(\mathbf{u}) = U(\mathbf{u}; \hat{\theta}(\mathbf{u}))I(\mathbf{u}; \hat{\theta}(\mathbf{u}))^{-1}U(\mathbf{u}; \hat{\theta}(\mathbf{u}))$$

where $\hat{\theta}(\mathbf{u})$ is the maximizer of $\log L(\mathbf{u}; \theta)$ subject to $\theta \in \Theta_0$. These statistics require a separate computation at each spatial location \mathbf{u} , and their asymptotic null distribution is χ_d^2 . By visual inspection of the maps of the resulting local p-values (one value for each spatial location \mathbf{u}) we individuate those regions of the analysed area where the hypothesis is confuted. Therefore, it can be possible to identify the regions where the local model is more appropriate than the global one, gathering information on the areas with the most inhomogeneous behaviour, that is, where the effect of the considered covariates varies. Of course, if the hypothesis is rejected in the majority of the area under study, the local model should be preferred to the global one.

5.2.1 Bootstrap procedure for 'homogeneity' testing

However, when the rejection of the null hypothesis, based on the above tests, is particularly due to some specific areas, it becomes interesting to individuate those regions with inhomogeneous behaviour, that is splitting the study area into sub-regions requiring further study or different models. Therefore, for confuting the hypothesis of 'homogeneity', however retaining information about the local features, a bootstrap procedure is proposed and carried out in this paper. Indeed, the bootstrap procedure (Efron 1982) is often used as an alternative to the statistical inference based on the assumption of a parametric model when that assumption is in doubt. The procedure used in this paper can be outlined as follows:

1. Set M as the number of bootstrap samples to be considered.
2. Bootstrap M times from the analysed point process, that is, random sample with replacement the points of the process, obtaining M new point processes.
3. The candidate local model is fitted to each of the M bootstrapped samples.
4. For each of the M resulting local models, the test of 'homogeneity' is carried out, obtaining M sets of local p-values for the given confidence level a .
5. The local p-values are averaged over the M samples.

The resulting averaged local p-values can be plotted as a function of the spatial location \mathbf{u} , together with their standard deviations. This map visually identifies the regions with inhomogeneous behaviour, i.e. where there is evidence against the hypothesis of 'homogeneity'. Contrary to the global p-value, computed as the average of the local p-values of one 'homogeneity' test, this bootstrap procedure contributes to the model selection process, retaining also local information. Fixing a confidence level a , e.g. 0.05, the areas where the local p-values are less than a confute the hypothesis of 'homogeneity', while values greater than a provide evidence in favour of the null hypothesis, that is, where the global model could be more preferable than the local one. An average of the resulting local p-values can also be computed, contributing to the model selection procedure, empirically comparing this value with the fixed level a (Diggle and Gratton 1984).

5.2.2 Assessing the effect of covariates on the local departure from 'homogeneity'

When the previously discussed local tests provide evidence against the null hypothesis of 'homogeneity', we can describe the available point pattern by a local model. At this point, we wonder which are the covariates that most contribute to the rejection of this hypothesis, that is to say, what covariates' effects really vary in space. Indeed, the 'homogeneity' test compares the linear predictor of the local model to the one of the corresponding global model, and therefore it does not provide information on each covariate individually.

In this section, we contribute to the framework of diagnostic tools for local models by proposing a stepwise procedure that highlights the effect of each

covariate on the rejection of the hypothesis of ‘homogeneity’. The procedure that we describe in this section can be used for local Poisson models with a multiple linear predictor, that is, dealing with multiple covariates.

Suppose that the chosen model is a local Poisson model with spatial varying intercept $\theta_0(\mathbf{u})$ and parameters $\theta(\mathbf{u}) = (\theta_1(\mathbf{u}), \dots, \theta_p(\mathbf{u}))$ corresponding to the spatial varying effects of the p covariates $\mathbf{Z}(\mathbf{u}) = (Z_1(\mathbf{u}), \dots, Z_p(\mathbf{u}))$. Several local Poisson models are fitted, starting from the one in which only the space varying intercept is added into the linear predictor, and adding one spatial covariate at a time. The covariates are added starting from the most significant in explaining the intensity marginally, based, for instance, on the information obtained from results of Berman’s tests (Berman 1986). This test is used to assess the dependence of a point process on a spatial continuous covariate $Z(\mathbf{u})$. It can be carried out before formulating and fitting a model, to get a first indication of the variables influencing the intensity of the process, marginally. Then, the test of ‘homogeneity’ is carried out in sequence to each of the local Poisson models, basically comparing each time the local Poisson model to its global counterpart. This procedure can be outlined as follows:

1. The Berman tests are computed on each available covariate, marginally. The covariates $\mathbf{Z}(\mathbf{u}) = (Z_1(\mathbf{u}), \dots, Z_p(\mathbf{u}))$ are then ranked from the most significant in explaining the intensity of the analysed process (therefore the one with the lowest p-value), to the least significant, obtaining the ordered vector $\mathbf{Z}^*(\mathbf{u}) = (Z_1^*(\mathbf{u}), \dots, Z_p^*(\mathbf{u}))$. The corresponding parameters are also ranked accordingly.
2. The model $\lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}))$ is fitted and the first test of local departure from ‘homogeneity’ is carried out for comparing

$$\begin{cases} \mathcal{H}_0 : \lambda(\mathbf{u}) = \exp(\theta_0) \\ \mathcal{H}_1 : \lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u})) \end{cases}$$

obtaining the first map of p-values, where values smaller than 0.05 indicate the regions where the local model should be preferred to the global one.

3. The second local model is fitted, adding the covariate that resulted the most significant in explaining the intensity of the process marginally, to the linear predictor $\lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1^*(\mathbf{u})Z_1^*(\mathbf{u}))$. Therefore, the second test of local departure from ‘homogeneity’ is carried out for comparing

$$\begin{cases} \mathcal{H}_0 : \lambda(\mathbf{u}) = \exp(\theta_0 + \theta_1^*Z_1^*(\mathbf{u})) \\ \mathcal{H}_1 : \lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1^*(\mathbf{u})Z_1^*(\mathbf{u})). \end{cases}$$

4. Repeat the last step for each $Z(\mathbf{u})$ following the ranking given in step 2, until the test of local departure from ‘homogeneity’ is carried out on the chosen model

$$\begin{cases} \mathcal{H}_0 : \lambda(\mathbf{u}) = \exp(\theta_0 + \theta_1^*Z_1^*(\mathbf{u}) + \dots + \theta_k^*Z_k^*(\mathbf{u})) \\ \mathcal{H}_1 : \lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1^*(\mathbf{u})Z_1^*(\mathbf{u}) + \dots + \theta_k^*(\mathbf{u})Z_k^*(\mathbf{u})). \end{cases}$$

The result is a set of maps, as many as the performed tests, that is, the number of covariates in the chosen model, plus the intercept. Each map will display the

p-values associated to the test of ‘homogeneity’ that compares the local model considered at that step with its global counterpart. Looking at each map individually, we obtain information about the regions in which the ‘homogeneity’ hypothesis is rejected. Of course, the wider the area leading to rejection, the stronger the evidence against the global model. By comparing the rejection areas obtained through the subsequent tests, carried out adding one variable at time, we are able to identify the most contributing variables to the overall rejection of the ‘homogeneity’ hypothesis in the final chosen model. In particular, if adding one variable to the model, say $Z_j(\mathbf{u})$, the rejection area of the corresponding map gets wider than the one relative to the model without $Z_j(\mathbf{u})$, then $Z_j(\mathbf{u})$ strongly contributes to the overall rejection, that is, the effect of $Z_j(\mathbf{u})$ varies in space. Otherwise, if no significant differences are evident between to subsequent maps, then, the variable $Z_j(\mathbf{u})$ does not contribute significantly to the overall rejection, suggesting that the effect of $Z_j(\mathbf{u})$, though considered as local in the model, could actually be considered as constant. This procedure is validated on the basis of the chosen local Poisson model fitted to the Greek seismic catalogue data in Sect. 6.1.

6 Analysis of the Greek seismicity

In this section, local models are applied to study the recent seismicity in Greece, assessing their goodness-of-fit by the diagnostics tools previously introduced.

All the analysis in this paper are carried out using the software R (R Core Team 2019). Global models, the diagnostic tools reviewed in Section 4.1 and Berman’s tests are computed with functions implemented in the `spatstat` package (Baddeley and Turner 2005). As for the local tests, and the fitting of the local models we refer to the `spatstat.local` package (Baddeley 2019). All the codes of the analyses carried out throughout the paper are available on request.

The analysed data (1105 events) concern earthquakes occurred in Greece between 2005 and 2014. Only seismic events with a magnitude larger than 4 are considered in this study, and the analyses in this paper are marginal with respect to time, focusing on the spatial dependence of events. In Fig. 1a earthquakes are reported together with the location of the five active volcanoes of the area (in green), the plate boundary (in red), and the faults (in blue). Data about the seismic events come from the Hellenic Unified Seismic Network (H.U.S.N.), while covariate information come from the Greek Database of Seismogenic Sources (GreDaSS).

In Fig. 1b the observed K -function, is represented. The dotted red line represents the values of the K -function for a homogeneous Poisson process, while the black one represents the K -function of the observed point process. The K statistic of the observed point pattern, with larger values than the theoretical one of a Poisson process, suggests that distances in the observed pattern are shorter than the Poisson ones. In other words, events are more clustered than an homogeneous Poisson pattern, as it is also evident in Fig. 1a.

Starting from the information of the area under study, we are interested in assessing if the presence of seismic sources affects the intensity of the process, that is if the effect of the available covariates can improve the fitting of the model, by fitting

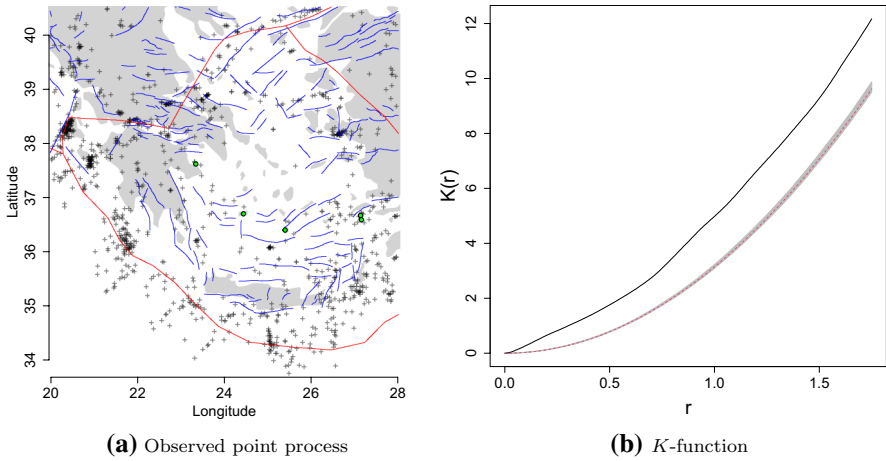


Fig. 1 Earthquakes occurred in Greece between 2005 and 2014. Volcanoes, in green, faults, in blue, and plate boundary, in red **(a)**; K -function of the observed point process of the Greek data **(b)**. The dotted red line represents the values of the K -function for a homogeneous Poisson process, while the black one represents the K -function of the observed point process (colour figure online)

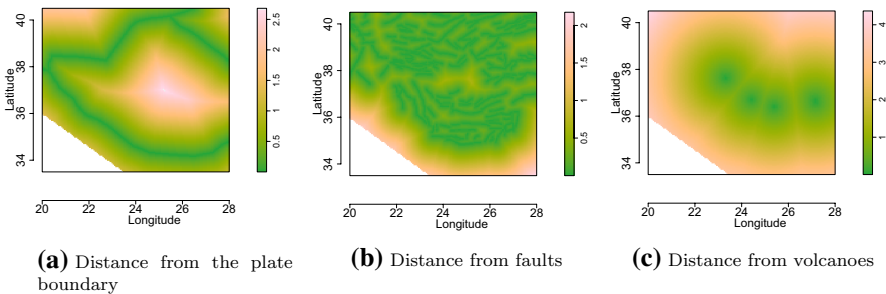


Fig. 2 The available covariates for the Greek data: *Distance from the faults* ($D_f(\mathbf{u})$), *Distance from the plate boundary* ($D_{pb}(\mathbf{u})$) and *Distance from volcanoes* ($D_v(\mathbf{u})$), computed as the Euclidean distances from the spatial location \mathbf{u} of events and the map of geological information

a global inhomogeneous Poisson process. Therefore, the considered explanatory variables are *Distance from the faults* ($D_f(\mathbf{u})$), *Distance from the plate boundary* ($D_{pb}(\mathbf{u})$) and *Distance from volcanoes* ($D_v(\mathbf{u})$), computed as the Euclidean distances from the spatial location \mathbf{u} of events and the map of geological information (Baddeley et al. 2015). In particular, only the information about the top segment of the fault is used to compute the distances. The covariate surfaces are displayed in Fig. 2.

The Berman’s tests, are computed for assessing the dependence of the intensity of the process from the continuous spatial covariates. The corresponding p-values are displayed in Table 1.

Therefore, the marginal effects of $D_{pb}(\mathbf{u})$ and $D_f(\mathbf{u})$ seem significant. Otherwise, $D_v(\mathbf{u})$ does not affect significantly the earthquakes intensity (p-value 0.68). To assess the variables influence on the intensity of the process, we plot the marginal

Table 1 Berman’s Tests carried out for the available covariates for the Greek data

Variable	<i>p</i> -value
Distance from the faults	0.0000
Distance from the plate boundary	0.0000
Distance from the volcanoes	0.6862

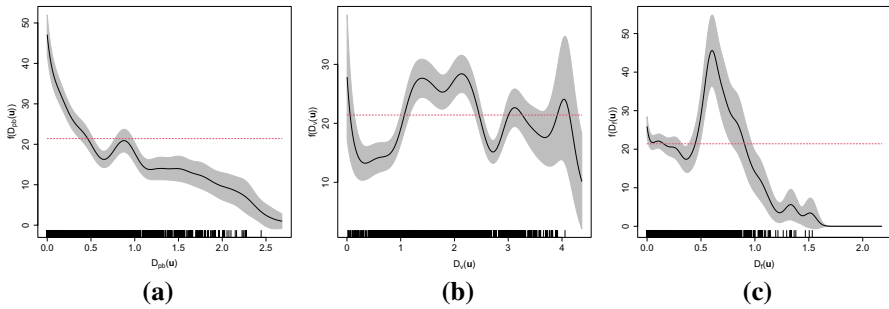


Fig. 3 Nonparametric estimate of the intensity of the analysed point process, as a function of the available variables for the Greek data: Distance from the plate boundary (a), Distance from the volcanoes (b) and Distance from the faults (c). The considered smoothing procedure is based on fixed-bandwidth kernel density estimation

smoothed intensity function for each covariate. Let $\lambda(\mathbf{u})$ be the intensity of the process under study, and $Z(\mathbf{u})$ the considered covariate. Then, assuming $\lambda(\mathbf{u}) = f(Z(\mathbf{u}))$ where f is a nonparametric estimate of the intensity of the analysed point process, we wonder if the observed intensity depends on each spatial covariate, at least marginally. Smooth estimators of $f(Z(\mathbf{u}))$ were proposed by Baddeley et al. (2012), and the smoothing procedure we consider is based on fixed-bandwidth kernel density estimation. The Poisson confidence bands are also computed.

The analysis of the smoothed functions confirms the results of the Berman’s tests. Indeed, $D_v(\mathbf{u})$ does not seem to affect the intensity, and Fig. 3b shows a quite constant relationship with the intensity. Looking at the smoothed functions of $D_{pb}(\mathbf{u})$ and $D_f(\mathbf{u})$, it is evident that the effect of both covariates varies as a function of the scale. In detail, the intensity exponentially decreases moving away from the plate boundary, while it has a piece-wise trend with respect to $D_f(\mathbf{u})$.

In Sect. 6.1, a local Poisson model is fitted to the Greek catalogue data and results are compared to those referred to a global model fitting, assuming an inhomogeneous Poisson models. In Sect. 6.2, a local LGCP model is fitted to a seismic sequence selected from this catalogue of events.

6.1 Local Poisson model for the Greek catalogue

On the basis of the previous results and for the given data, we want to assess if the local features cannot be ignored, fitting an inhomogeneous Poisson model as in Eq.

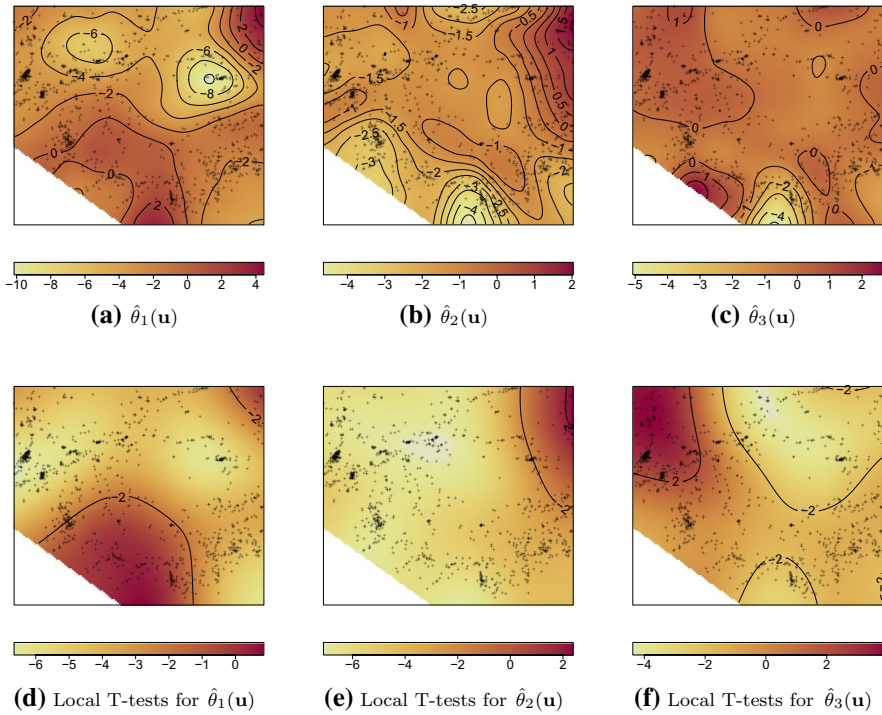


Fig. 4 On panels (a, b, c): Spatial varying coefficients $\hat{\theta}_1(\mathbf{u})$, $\hat{\theta}_2(\mathbf{u})$ and $\hat{\theta}_3(\mathbf{u})$ of the local Poisson Model in Eq. (18), the coefficients of the variables $D_f(\mathbf{u})$, $D_{pb}(\mathbf{u})$, and $D_v(\mathbf{u})$, respectively. On panels (d, e, f): the corresponding T-tests. T level curves correspond to the ± 1.96 threshold, associated to a 0.05 confidence level. The darkest (and lighter) regions identify where the null hypothesis is rejected, that is where the local coefficients are significantly different from zero

(18). Starting from the available variables, the choice of the final model is based on diagnostic approaches, as reported in Sects. 5.1 and 5.2. The chosen model is the one including the *Distance from the faults* (D_f), the *Distance from the plate boundary* (D_{pb}) and the *Distance from the volcanoes* (D_v), that is

$$\lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1(\mathbf{u})D_f(\mathbf{u}) + \theta_2(\mathbf{u})D_{pb}(\mathbf{u}) + \theta_3(\mathbf{u})D_v(\mathbf{u})) \tag{18}$$

where the smoothing parameter h , obtained maximizing the LCV in Eq. (5) used in the local likelihood in Eq. (4), is 0.73 degrees, corresponding to approximately 81.03 kilometers. The maps of the varying coefficients $\hat{\theta}_1(\mathbf{u})$, $\hat{\theta}_2(\mathbf{u})$ and $\hat{\theta}_3(\mathbf{u})$ are shown on the top three panels in Fig. 4.

The summary statistics of the local coefficients are reported in Table 2. As expected, the estimated coefficients are negative quite in all the area, since usually, as the distance from the seismic source increases the intensity decreases, but it is important to notice that the coefficients take quite different values along the whole areas, reaching the lowest values of -10.39 for $\hat{\theta}_1(\mathbf{u})$. The lowest values correspond to the earthquakes occurred along the seismic sources.

Table 2 Summary of the space varying coefficients of the local Poisson model (18)

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Intercept	- 0.8465	3.7821	4.6536	4.7818	5.4126	13.7942
D_f	- 10.3862	- 4.0474	- 2.3888	- 2.4529	- 0.6133	4.6906
D_{pb}	- 4.735	- 1.925	- 1.413	- 1.448	- 1.04	2.17
D_v	- 5.3552	- 0.6595	- 0.2799	- 0.2987	0.2351	2.9546

To study the effect of each covariate separately and to asses if the estimated values of the spatial varying coefficients are significant, T-tests of the parameters of the model are computed, as in Eq. (17). The T-test output is shown in the bottom panels of Fig. 4, where the level curves correspond to the ± 1.96 threshold, associated to a 0.05 confidence level. Therefore, in the darkest (and lighter) regions the null hypothesis is rejected, that is, the coefficients corresponding to these locations are significantly different from zero. It is interesting to notice that the coefficients of $D_f(\mathbf{u})$ take negative values along all the study area but for a wide area on the bottom-left on the window in which coefficients are positive and not significant. The coefficients of $D_{pb}(\mathbf{u})$ are significant and negative basically in all the study area, but for the two areas at the boundary of the window. Finally, the coefficients of $D_v(\mathbf{u})$ are significant only in three bounded areas of the window, taking both positive and negative values.

These results suggest that a local model might be suitable for describing the features of the study area. To get more evidence supporting this hypothesis, the test of the ‘homogeneity’ is carried out, bootstrapping 100 datasets from the original one, and computing the average of the resulting p-values obtained for each location, as discussed in Sect. 5.2.1. These values are shown in Fig. 5, together with their standard deviations, suggesting the inadequacy of the null hypothesis for all the study area: indeed, few local p-values slightly greater than the fixed level 0.05 are observed, but less than 0.1, in the southern region, corresponding also to the sub-area with the largest standard errors. Besides, the mean of all the previous p-values is 0.01, with standard deviation equal to 0.05, further corroborating the hypothesis of the absence of ‘homogeneity’ in these data.

This result suggests that the proposed local model in Eq. (18) should be preferred to its global counterpart

$$\lambda(\mathbf{u}) = \exp(\theta_0 + \theta_1 D_f(\mathbf{u}) + \theta_2 D_{pb}(\mathbf{u}) + \theta_3 D_v(\mathbf{u})).$$

To understand the effect of each variable in rejecting this hypothesis, we propose the stepwise comparison between the different ‘homogeneity’ settings, introduced in Sect. 5.2.2. Therefore, four local Poisson models are fitted, starting from the one in which only the local intercept is considered, and adding one spatial covariate at a time. The covariates are ranked from the most significant in explaining the intensity marginally, based on the results of Berman’s tests in Table 1, such that $\mathbf{Z}^*(\mathbf{u}) = (D_f(\mathbf{u}), D_{pb}(\mathbf{u}), D_v(\mathbf{u}))$. These are subsequently added to the linear predictor and the test of ‘homogeneity’ is carried out each time comparing the template global Poisson model to its local counterpart. Values smaller than 0.05 in Fig. 6 are

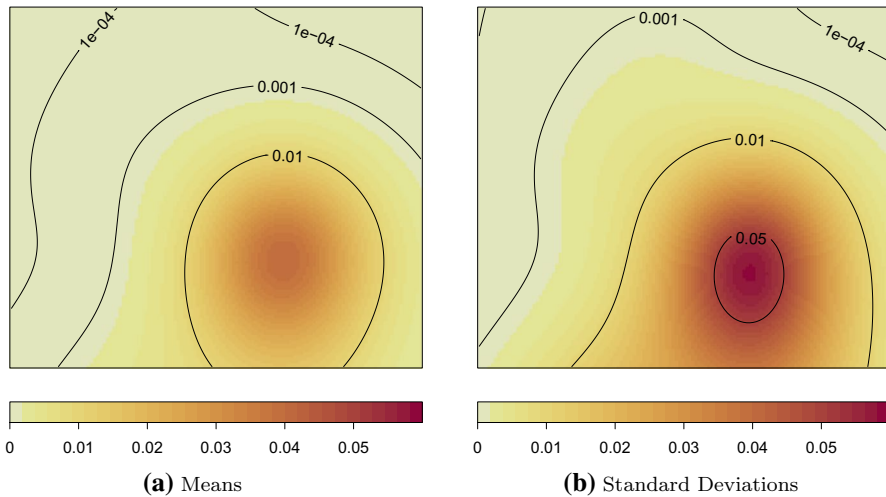


Fig. 5 Averaged local p-values and their standard deviations for the test of ‘homogeneity’, computed over 100 bootstrapped datasets

plotted in lighter colors, and they are found in those regions where the hypothesis of ‘homogeneity’ is rejected. Of course, we may notice that the more covariates with local effects, the more areas ‘needing’ a local model. In particular, as the greater differences between subsequent maps are in correspondence of the addition of $D_f(\mathbf{u})$ and $D_{pl}(\mathbf{u})$, this means that these are the covariates with spatially varying effects. On the contrary, $D_v(\mathbf{u})$ slightly contributes to the rejection of the hypothesis and this may suggest that its effect could be considered as constant. Overall, we may notice that the best fitting for all the area would require an even more complex model, but introducing local coefficients makes the model more flexible and more useful for a deeper interpretation.

A feasible trade-off between the flexibility of a local model and the parsimony of a global one could be the model with a linear predictor that includes a non-parametric term for spatial coordinates and parametric expression for the spatial covariates, as

$$\lambda(\mathbf{u}) = \exp(f(\mathbf{u}) + \theta_1 D_f(\mathbf{u}) + \theta_2 D_{pb}(\mathbf{u}) + \theta_3 D_v(\mathbf{u})) \quad (19)$$

where $f(\cdot)$ is a nonparametric function for $\mathbf{u} \in D$, estimated here through thin plate regression splines with 30 knots. Comparisons between the models (18) and (19) are first carried out by the inspection of the smoothed raw residuals, as defined in Eq. (16). The smoothing bandwidth used for the kernel estimation of the raw residuals is 0.025 degrees, corresponding to approximately 27.75 kilometers. The smoothed raw fields of the two models are shown in Fig. 7 and their ranges are reported in Table 3.

From these results, we see that both models achieve a good fitting to the data, even if the global model slightly presents smaller residuals than the local one.

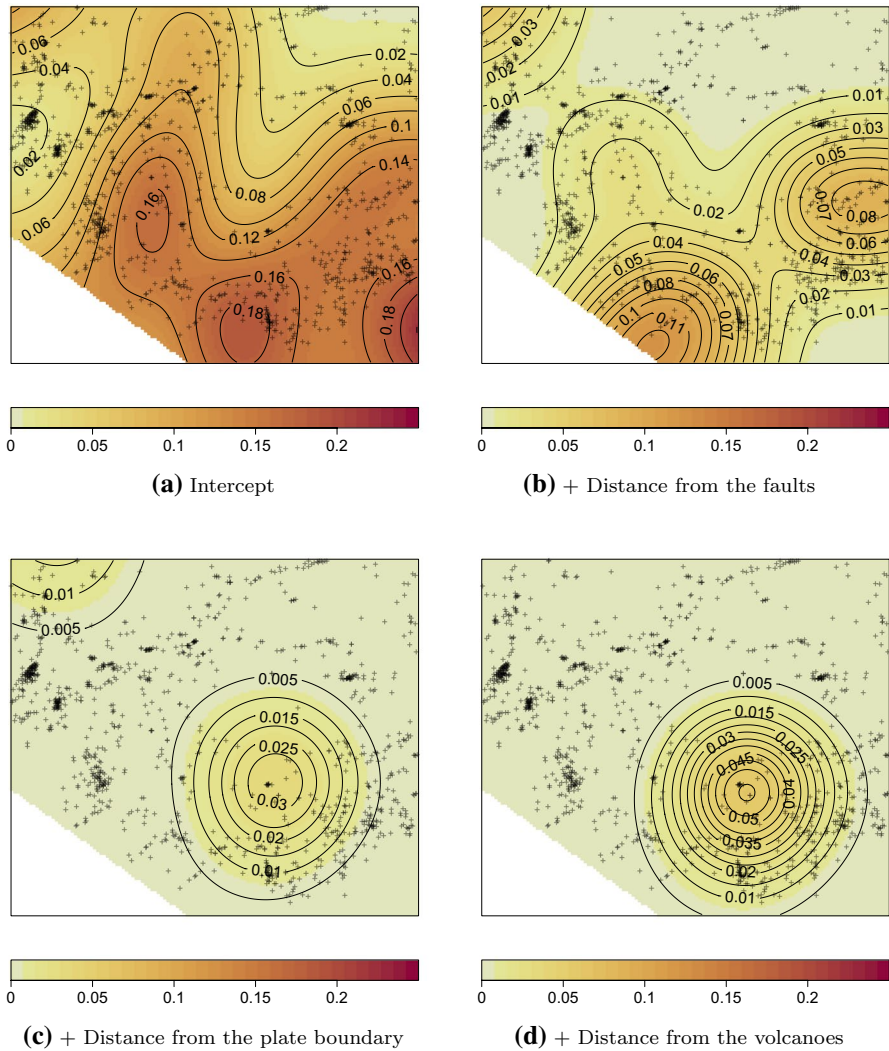


Fig. 6 Test of ‘homogeneity’ change: p-values of the test of ‘homogeneity’ are plotted, in reference to the four local Poisson models fitted for describing Greek data. Starting from the one which includes only $\hat{\theta}_0(\mathbf{u})$ in panel (a), the covariates are added in the models gradually. The p-values are computed for the tests of ‘homogeneity’ carried out on the models including, step-by-step, the Distance from the faults (b), the Distance from the plate boundary (c) and the Distance from the volcanoes (d)

Secondly, a comparison is carried out in terms of the inhomogeneous K -function introduced in Eq. (15). In Fig. 8a, we report the inhomogeneous K -function for the local model (18), while in Fig. 8b, the one of the global inhomogeneous semi-parametric model (19) is displayed. In both Figures, the solid black line represents the inhomogeneous K -function $\hat{K}_{\text{inhom}}(r)$ estimated with the intensity of the fitted model, while the dotted red line represents the theoretical one $K_{\text{inhom,pois}}(r)$. The

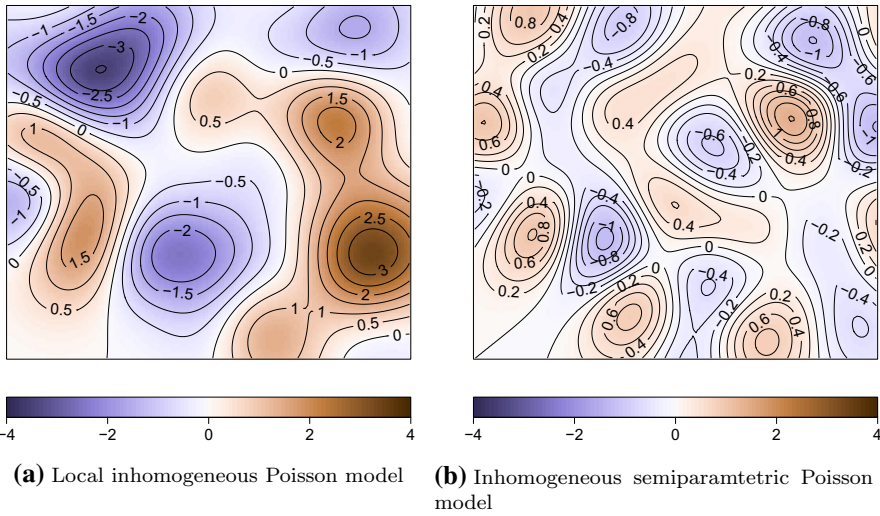


Fig. 7 Smoothed raw residuals of the local inhomogeneous Poisson Model (18) and the global inhomogeneous semiparametric Poisson Model (19)

Table 3 Range of the smoothed raw fields of the local Poisson Model (18) and the global Poisson Model (19)

	Min. range	Max. range
GlobalPoisson	- 1.27	1.41
LocalPoisson	- 3.51	3.42

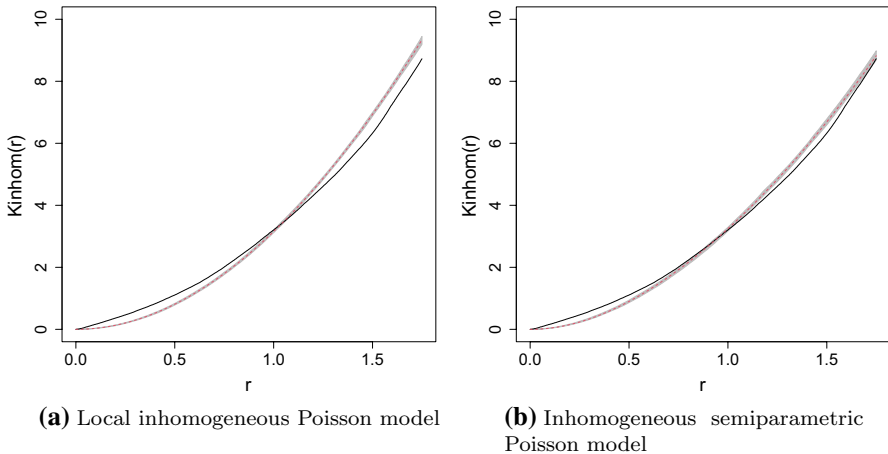


Fig. 8 Inhomogeneous K -functions of the local inhomogeneous Poisson Model (18) (a) and the global Inhomogeneous semiparametric Poisson Model (19) (b). In both panels, the inhomogeneous K -function estimated on the observed pattern $\hat{K}_{inhom}(r)$ is the solid black line and the theoretical Poisson one $\hat{K}_{inhom,pois}(r)$ is the dotted red line, together with the corresponding envelopes, in grey (colour figure online)

corresponding envelopes, that are obtained considering the inhomogeneous K -function for a given number of simulated patterns, assuming the same intensity of the observed patterns, are displayed in grey. For their computation, we simulated 100 realizations from the fitted model in Eq. (18). The inhomogeneous K -function of both models lay very close to the theoretical Poisson curves, even if the one of the global model is a bit closer to its theoretical one, even for larger distances. Based on these results, the global model in Eq. (19) should be preferred to the local one in Eq. (18).

Overall, we might now conclude that both the global model (19), in which the spatial component is modelled non-parametrically, and the local model (18) achieve good fitting to the analysed data. Nevertheless, with both models, it is not possible to get a concise comment about the effect of the spatial covariates, earning in terms of fitting but losing interpretation. On the one hand, the great advantage of local models is that one can get a whole map of parameters, one for each covariate added in the linear predictor, and this allows us to detect those regions in which covariates have a different effect of the phenomenon under study. Therefore, if we wish to get a general overview of the effect of covariates a global model is to be preferred. On the other hand, if we are interested in spotting those areas in which covariates may have a different effect, and therefore, highlighting the multi-scale nature of the observed generating process, local models could be more appropriate. Moreover, as we add more covariates into the linear predictor, complexity increases, as each one of them carries a whole map of parameters. Finally, local models are useful also when covariates are not available. Indeed, since the test of ‘homogeneity’ can be carried out to any kind of linear predictor, this may spot those regions that really need a local varying intercept, that is those areas that need more attention and further analysis.

6.2 Local log-Gaussian Cox process model for a Greek seismic sequence

When analysing seismic sequences, the analysis of the small scale variation is of crucial interest (Giunta et al. 2009). For this purpose, a Log-Gaussian Cox Process is a candidate model, accounting for the interaction among points through a Gaussian Random Field. In their local version, parameters of the Gaussian Random Field vary in space, allowing to describe the complex dependence structure of the generating process. Starting from the Greek catalogue, we selected a smaller area in order to emphasize the dependence features at short distances. The chosen area is situated in the Ionian Sea, and it consists of three of ‘the Seven Islands’, namely Ithaki, Kefalonia and Zakynthos, from North to South. In this area, earthquakes appear to be clustered in the Western area of Kefalonia island and South to the Zakynthos island. For this analysis, only $D_f(\mathbf{u})$ and the $D_{pb}(\mathbf{u})$ are considered, since the volcanoes are very far from this analysed area. In Fig. 9 earthquakes, together with plate boundary (in red) and faults (in blue), are displayed. Given that previous analyses have highlighted the highly inhomogeneity of the chosen subregion, and therefore the need of more complex models than the inhomogeneous Poisson, we fit the log-Gaussian Cox process, as our objective here is to find the model that better describes the clustered structure of the point pattern.

Fig. 9 Earthquakes occurred in Ithaki, Kefalonia and Zakynthos between 2005 and 2014. The plate boundary is displayed in red while the faults are displayed in blue

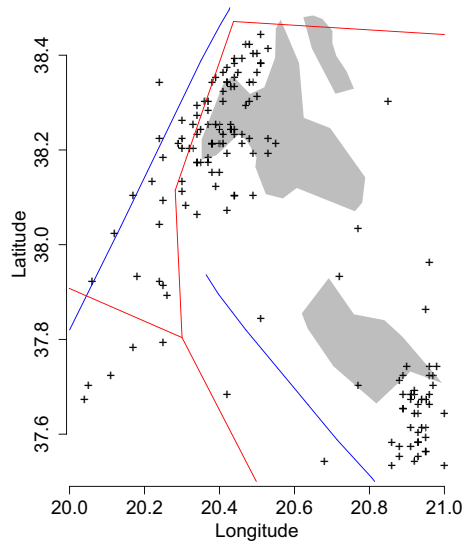


Table 4 Summary statistics of the local coefficients of the deterministic component of the local LGCP model for the Greek data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Intercept	- 11.264	2.372	5.797	2.943	6.396	10.052
D_f	- 28.259	1.318	8.095	8.303	14.777	51.888
D_{pb}	- 23.820	- 8.782	- 0.575	6.843	25.126	55.091
$D_f D_{pb}$	- 306.57	- 84.03	- 45.48	- 47.90	- 12.10	63.96

After choosing the following multiplicative local Poisson model for the deterministic component, then, it is used to fit the local LGCP model:

$$\lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1(\mathbf{u})D_f(\mathbf{u}) + \theta_2(\mathbf{u})D_{pb}(\mathbf{u}) + \theta_3(\mathbf{u})D_f(\mathbf{u})D_{pb}(\mathbf{u})) \quad (20)$$

where the local weight function w in the local Palm likelihood is an isotropic Gaussian density, with smoothing parameter equal to 0.135, chosen by the cross-validation criterion (13). The summary statistics of the estimated parameters of the deterministic component of the local LGCP model are reported in Table 4.

The estimates of the local interaction parameters $\hat{\sigma}^2(\mathbf{u})$ and $\hat{\alpha}(\mathbf{u})$ are displayed in Fig. 10 and their summary statistics are reported in Table 5.

The results in Table 5 are quite reasonable, since the estimates of the coefficients $\hat{\sigma}^2$ and $\hat{\alpha}$ of the global counterpart of the LGCP model, estimated maximizing the (global) Palm Likelihood, are respectively equal to 1.92 and 0.2.

We know that the effect of increasing σ^2 is to generate higher peaks in the surface intensity, inducing a clustering effect. Increasing the spatial scale parameter α , the underlying GRF presents a strong spatial correlation and it corresponds to a diffuse aggregation of points of the LGCP. Therefore, from the visual inspection of the maps

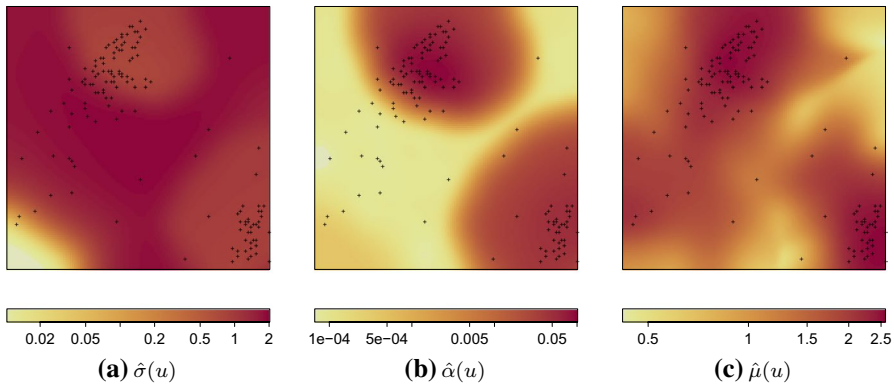


Fig. 10 Local coefficients $\hat{\sigma}(\mathbf{u})$ and $\hat{\alpha}(\mathbf{u})$ of the global LGCP model, respectively **(a, b)** and the mean of the process $\hat{\mu}(\mathbf{u})$ **(c)**

Table 5 Summary statistics of the local coefficients of the covariance function of the local LGCP model for the Greek data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
σ^2	0.000105	0.802612	2.451644	2.192861	3.627219	4.157830
α	0.00001	0.00009	0.00012	0.01988	0.03764	0.11281

of the interaction parameters in Fig. 10, we are able to identify regions with different underlying Gaussian processes, driven by different covariance structures. Figure 10a shows the map of the estimates $\hat{\sigma}(\mathbf{u})$; we may notice that, unexpectedly, the highest values are found where events are not clustered. Looking at Fig. 10b, the highest values of $\hat{\alpha}(\mathbf{u})$ correspond to those regions exhibiting clustered points. From the exponential expression of the covariance chosen for the GRF in Eq. (7) we know that the values of σ^2 contribute to the covariance in a multiplicative way, while the values of α affect the covariance exponentially. Then, a unit increase of α has a greater effect on the computation of the covariance, if compared to a unit increase of σ^2 . Therefore, inspecting Fig. 10, we know that the difference in the correlation among points is evident, and this is higher in those areas in which clusters are observed, i.e. on the top and on the bottom-right of the region.

A further application of the proposed local models is provided in “Appendix 1”, focusing on the study of an important seismic sequence of Abruzzo, a region of central Italy, characterized by different geological features with respect to the Greek ones.

7 Discussion and conclusion

In this paper, we have introduced models with space varying parameters, estimated by the composite likelihood for spatial point processes (Baddeley 2017), for the description of complex earthquake data characterized by multiple and multifractal

dependence structure. Analyzing the recent seismicity of a highly seismic active area, local models seem to be suitable for accounting the complexity of the seismic process, since they describe its local features fitting a unique model to the whole area under study.

First, a log-linear relationship between the earthquakes' intensity and the distances from the nearest seismic sources has been modelled, fitting a local inhomogeneous Poisson process model. From the related obtained results, it seems reasonable assuming local coefficients, that is, location depending effects of the considered covariates, reflecting the complex seismic activity of the given areas. Furthermore, we have found that although the local models are useful in finding spatially varying covariate effects, they may not be optimal for summarizing these different effects. However, this information could be used in order to split the study region into different areas and separately interpreting the estimated local coefficients, following a multiple global approach. Secondly, interaction among points, crucial in the context of the analysis of seismic events, has been taken into account by fitting a local LGCP model.

This represents the first main contribution of the paper, as the local Log-Gaussian Cox Process is comprehensively addressed and applied here, for the first time. We use the LGCP for describing a seismic sequence rather than an entire catalogue, since there is the possibility to account for the interaction among points by the estimation of the parameters of the underlying Gaussian Random Field. This makes the LGCP particularly suitable for the description of highly inhomogeneous areas, as the one chosen in the application provided in this paper. In addition, fitting a local version of the LGCP, it becomes possible to account both for the effect of covariates and for the interaction among points, as a function of the spatial location. Results of the reported application confirm the crucial role of the proposed approach for describing and characterising the study area through a multiple underlying process, with both the first- and second-order characteristics vary with location.

Furthermore, the second main innovation of this paper refers to our contribution to the framework of diagnostics for local models, outlining all the relevant methods for diagnostics and model selection, also proposing a bootstrap procedure for driving the model selection process, and a stepwise procedure, that is able to assess single local effects of covariates on the rejection of the hypothesis of 'homogeneity'. These procedures have been validated by the application of the proposed methods to real seismic data, showing that, as the number of covariates in the linear predictor increases, the number of areas actually needing a local model increases too. The use of existing diagnostic tools for global models in the context of local fitting and the proposed stepwise approach provide valid results and useful hints for assessing the estimated models and selecting the best one, testing the need of a local model in the analysed area.

Overall, starting from the results of the proposed application, local models provide good inferential results in the seismic point processes context. The complex earthquake phenomenon is characterized by multiple and multifractal dependence structure, and its description demands local tools and models to address its peculiar features. It seems that fitting a model with locally varying parameters is reasonable when analysing both seismic catalogues and sequences. In particular, the local

LGCP models more succeed in identifying regions where the events are clustered, than the global LGCP models that, through the estimation of constant interaction parameters, can just identify the presence of a global clustered structure. Moreover, local models provide a good alternative in terms of fitting to global models, even to nonparametric ones. A further advantage of the application of local models to the seismic context proposed in our paper, is the possibility of fitting a unique fully parametric model, considering in a proper way the interaction between covariates, peculiar in the describing the complex seismic phenomenon, compared to the most classical semi-parametric model proposed in the literature, and reviewed in the Introduction. Furthermore, as already pointed out by Baddeley (2017), some methodological problems still remain, such as the need for a theory of misspecified models and for variance estimators for Cox and cluster processes. Earthquake data, typically show extremely high concentrated points in small zones of high seismicity. Alternative approaches include the division of the study area into Dirichlet cells, fitting a proper model to the union of such cells (e.g the spatially-adaptive space-time point process models for earthquakes (Ogata and Katsura 1988; Ogata 1989)), with the drawback of providing high dimensional models. Spatial clustering techniques may work differently in different tectonic environments. In this work we have applied the proposed technique to a seismicity typical of a compressive tectonics and in particular in an area such as the Greek one where, as is well known, a lithospheric subduction process is underway. It is certainly desirable in the future to test the spatial clustering procedure here proposed also to seismicity recorded in areas characterized by extensional or strike-slip tectonics, as well as in areas characterized by seismicity linked to volcanic processes. This would make possible to verify the performance of the proposed technique in different tectonic regimes and its possible fine-tuning aimed at optimizing the spatial clustering.

Finally, a relevant drawback of this work is that we have not accounted for the temporal dimension of the seismic events, whose realization depends on their past history, as proved by the existence of aftershocks. Given the results of this paper on the application of spatial local models to the seismic context, we believe that these models could be the basis for future development for introducing local models and methods both for spatial and temporal dimension, referring to the theory of the space-time point processes. Indeed, few classical models widely used in the spatio-temporal seismic context account for external information such as spatial or temporal covariates (a recent proposal for the Epidemic Type Aftershock sequence model is provided by Adelfio and Chiodi (2021)). Therefore, it would be interesting to develop spatio-temporal extensions of the spatial local models used in this paper, for instance exploiting the alternative estimation procedure for the local LGCP models with minimum contrast based on local second-order summary statistics.

Appendix 1: Analysis of the Italian seismicity

In this section, the local LGCP model is used to describe an important Italian seismic sequence. The aim of specifying a LGCP model is to take into account the existing interactions among events, that are not accounted by a Poisson model, even

local. The analysis of the Greek seismic events has been carried out assuming that the intensity of earthquakes could depend only on spatial covariates, such as the distance to the seismic sources. This has been useful in order to test local models and to make different comparisons, but these variables only, all representing distances from some source of seismic activity, may not be enough in order to explain the phenomenon. Indeed, one could assume that other relevant variables could be useful to explain the intensity of the observed earthquakes, such as variables related to the network dislocation and measurement errors.

The Italian catalogue considered in this section is downloaded from the Istituto Nazionale di Geofisica e Vulcanologia (INGV) archive,¹ and it describes the space-time Italian seismicity, from May 5th 2012 to May 7th 2016 with 2.5 as the threshold magnitude, and reports 2226 events and several variables, including hypocentral coordinates (longitude, latitude, depth, time), magnitude, hypocentral uncertainty, distance from the nearest station, rms, that is a measure of the quality of the location and the number of stations that recorded the event. As the location of faults along the study area is also available, the distance from the nearest fault is computed, in the same way as for the Greek data, and it is denoted by $D_f(\mathbf{u})$.

The Italian data is quite different with respect to the Greek one. While the Greek covariates have been obtained from information external to the analysed point pattern, the Italian catalogue carries more complete information about seismic events. Indeed, Greek covariates are defined for all the region under study, while the Italian covariates refer only to those seismic events already occurred. While the structure of the Italian catalogue may lead to think to the available variables as marks of the point process, some of this variables can be actually thought as spatial covariates, if interpolated along the whole study area. Indeed, it makes sense to think that variables such as the Hypocentral uncertainty, the RMS and the Distance to the nearest station (Err , Rms and D_{ns}) could depend on the location, and therefore it makes sense to get an interpolation of these variables along all the study area. On the other hand, variables such as the Magnitude, Depth and Time, are characteristics of the points, known as 'marks' in spatial point process theory. The spatial covariates obtained from the interpolation, through the application of Cubic spline interpolation, are treated as potential covariates in the fitted models and denoted by $Err(\mathbf{u})$, $Rms(\mathbf{u})$ and $D_{ns}(\mathbf{u})$.

It is worth to notice that, for computational reasons, it was not possible to describe the other Italian areas, such as the one corresponding to the Emilia sequence. Indeed, analysing the given point process would lead to both unsatisfactory interpolation and computational problems when fitting local models, due to the strong irregularity of data. Moreover, as already stated by Baddeley (2017), the local composite likelihood works well if kernel smoothing performs well. Dealing with earthquake data, this point can not be guaranteed, since their well known features of inhomogeneity and clustering in space.

Consequently, we focus on a seismic sequence. The selected region is Abruzzo, sadly well known for its seismic history. In April 2009, in the mountainous Abruzzo

¹ <https://istituto.ingv.it/it/risorse-e-servizi/archivi-e-banche-dati.html>.

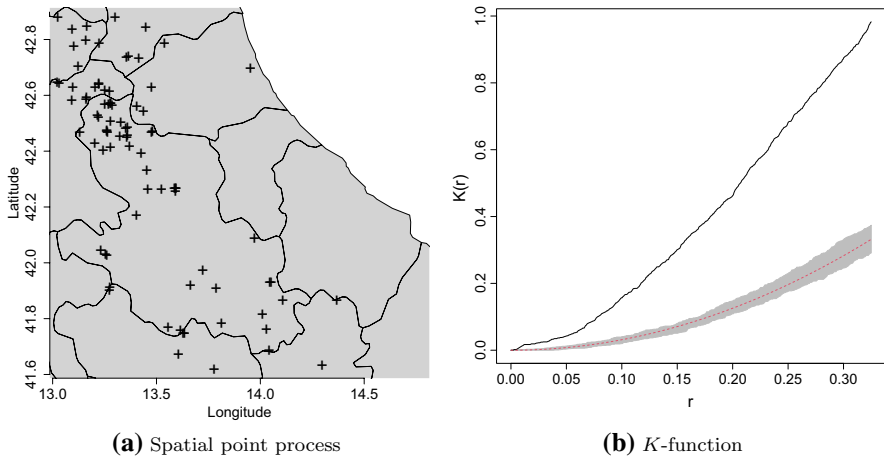


Fig. 11 Earthquakes occurred in Abruzzo between May 2012 and May 2016 **(a)**, consisting of 85 events with 2.5 as the threshold magnitude; K -function of the observed **(b)** Italian seismic sequence. The dotted red line represents the values of the K -function for a homogeneous Poisson process, while the black one represents the K -function of the observed point process

region, a 6.3-magnitude quake killed nearly 300 people and destroyed or seriously damaged thousands of buildings. During a period of three months in 2016, Central Italy was hit by three large earthquakes and numerous aftershocks. The first one, with the magnitude of 6.0, occurred in Amatrice on August 24 and killed 298 people. On October 26 evening, a series of shocks with the strongest one having a magnitude of about 6, took place in approximately the same region (Lutikov et al. 2018). Those sequences are actually missing from the data analysed in this paper.

The earthquakes occurred between May 2012 and May 2016 in Abruzzo are displayed in Fig. 11a, consisting of 85 events with 2.5 as the threshold magnitude. As we may notice, a concentration of events is observed on the top-left of the region, in correspondence of the city of L'Aquila. So, even if the events occurred in 2009 are not analysed in this paper, their consequences on the following sequences can be observed.

The K -function of the analysed point process is shown in Fig. 11b, highlighting a clear clustered pattern of the events. The Berman's tests are carried out in order to check if these variables influence the intensity of earthquakes. As we may notice from Table 6, the variables that seem to influence the spatial intensity are $D_f(\mathbf{u})$, $D_{ns}(\mathbf{u})$, $Err(\mathbf{u})$, while $Rms(\mathbf{u})$ doesn't influence the intensity values significantly.

Model proposal

Starting from the Italian data, a local Poisson model is estimated to describe the local features of data, including the available covariates, assuming that they may influence the intensity as a function of the spatial location. After having compared different models through the inhomogeneous K -function and the smoothed raw residuals, the chosen local Poisson model is the following

Table 6 Berman's Tests for the available spatial variables in the Italian catalogue

Variable	<i>p</i> value
Distance from the faults	0.002958
Distance from the nearest station	0.0002208
Hypocentral error	0.001353
RMS	0.1441

Table 7 Summary statistics of the coefficients of the local Poisson model (21) for the Italian data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Intercept	3.17	4.04	4.51	4.52	4.99	5.82
D_{ns}	-0.10	-0.08	-0.06	-0.06	-0.05	-0.03
D_f	-7.78	-5.91	-2.67	-2.90	0.10	3.03

$$\lambda(\mathbf{u}) = \exp(\theta_0(\mathbf{u}) + \theta_1(\mathbf{u})D_f(\mathbf{u}) + \theta_2(\mathbf{u})D_{ns}(\mathbf{u})) \quad (21)$$

where only the Distance from the faults and the Distance from the nearest station are involved. The local coefficients of the model are shown in Fig. 12 and their summary statistics are reported in Table 7. Negative values of the coefficients of both variables are reasonable. Indeed, usually, as the distance from the faults increases, the intensity decreases. Moreover, as the distance from the nearest station recording the event increases, the intensity decreases, as detection of earthquakes is usually more accurate if they occur not far from the network station. Looking at the left panels of Fig. 12 both variables take different values of their coefficients along all the study area, though looking at their T-test, they are not significant along the whole area. Indeed, it seems that the area could be divided into two sub-regions (right panels of Fig. 12): one, on the top-right, in which the variable affecting the intensity is $D_{ns}(\mathbf{u})$, and one, that is all the remaining area, in which only $D_f(\mathbf{u})$ affects the intensity.

For testing local departure from 'homogeneity', the stepwise approach already used for the Greek data is here considered. In order to assess the influence of each variable to the lack of the 'homogeneity', the space varying p-values of the 'homogeneity' tests are visually inspected. In Fig. 13a the p-values for the 'homogeneity' test in the Null Model are less than 0.05 in two separate areas, namely the one on the top-left and the one on the right of the window, corresponding to those areas where events appear to be clustered or are missing. Looking at Fig. 13b we see that, including the variable $D_{ns}(\mathbf{u})$ in the model, the 'homogeneity' hypothesis is rejected in the middle of the window. Finally, in Fig. 13c, including also $D_f(\mathbf{u})$, the necessity of a local model is more evident. Therefore, as for the Greek data, adding covariates into the model makes the local model more adequate for describing observed data, with respect to a global model with constant regression coefficients.

As already stated in the analysis of the Greek data, we fit a local LGCP model, with the same linear predictor selected from the above local Poisson model.

In Fig. 14 the estimates of the two additional parameters $\hat{\sigma}(\mathbf{u})$ and $\hat{\alpha}(\mathbf{u})$ are shown, and their summary statistics are reported in Table 8. The results are

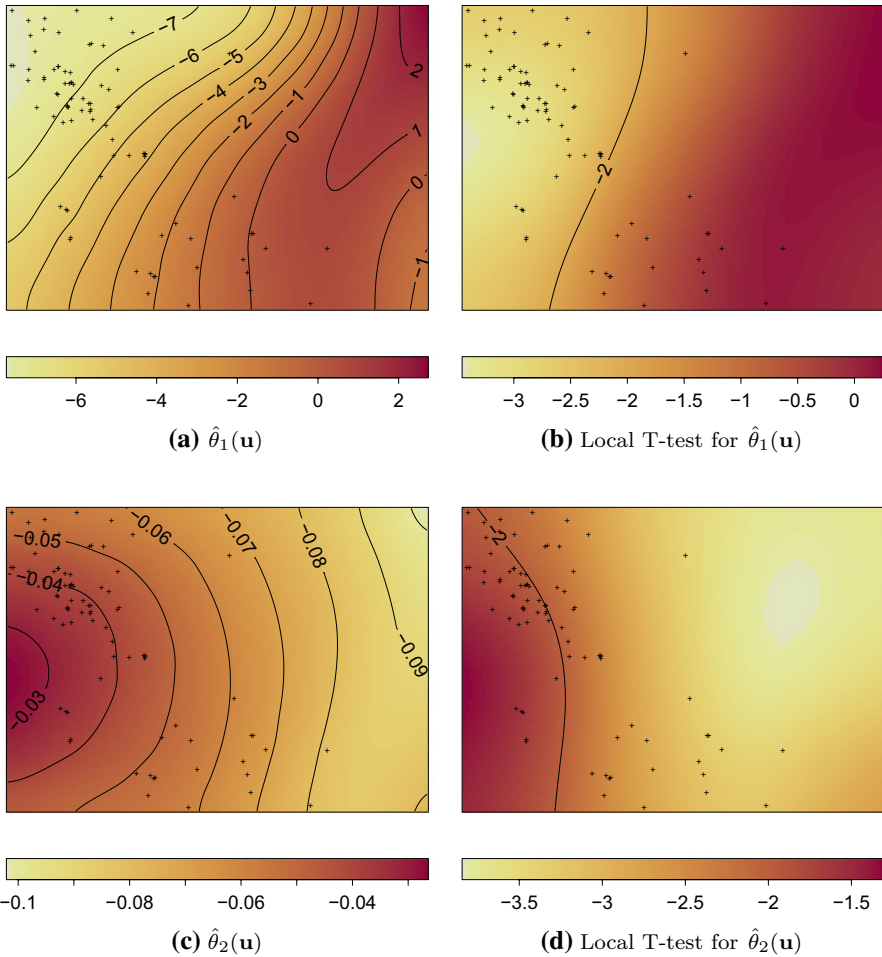


Fig. 12 Coefficients of the local Poisson model (21) and their T-tests for the Italian data. The coefficients of the Distance from the nearest station are displayed on the top-left panel while the coefficients of the Distance from the faults are displayed on the bottom-left panel

coherent with the estimates of the coefficients $\hat{\sigma}^2$ and $\hat{\alpha}$ of the global LGCP model that are respectively equal to 1.47 and 0.21.

The results indicate that these two parameters have different behaviour on the top-left of the selected window, with respect to the remaining region. This area corresponds to the region where events are more clustered, suggesting, as obvious, that in this area the correlation among points is higher and therefore the corresponding regions should be treated differently from the rest.

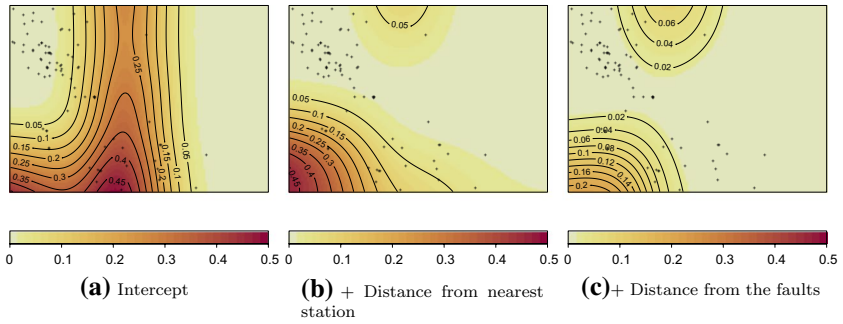


Fig. 13 Test of ‘homogeneity’ change for the model (21) fitted for describing Italian data: p-values of the test of ‘homogeneity’ are plotted, for the Null Model (a) that involves only $\theta_0(\mathbf{u})$ into the linear predictor, the model in which also the Distance from the nearest station is added (b) and the model in which both the Distance from the nearest station and the Distance from the faults are added into the linear predictor (c)

Table 8 Summary statistics of the spatial varying coefficients of the local LGCP model for the Italian data

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
σ^2	0.000026	2.666119	3.455831	3.055668	3.889292	4.343586
α	0.0005174	0.0094810	0.0113980	0.0194697	0.0157996	0.0770704

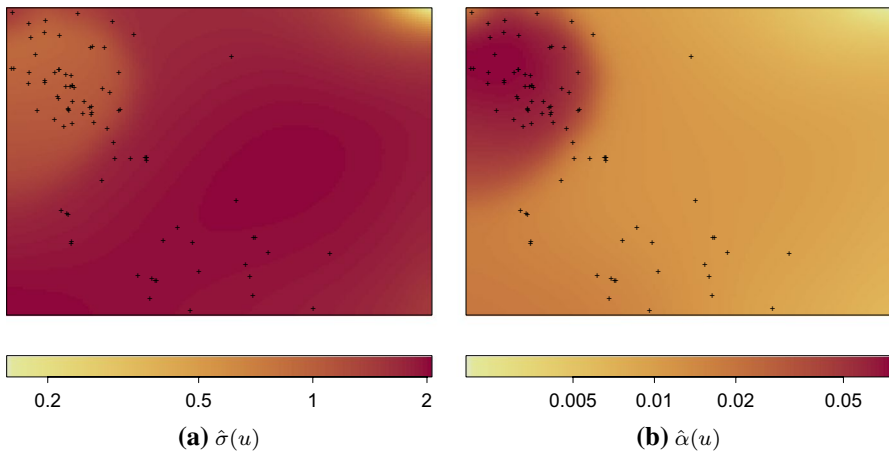


Fig. 14 Spatial varying additional parameters $\hat{\sigma}(\mathbf{u})$ and $\hat{\alpha}(\mathbf{u})$ of the local LGCP model fitted for describing the Italian seismic sequence

Funding Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement. This work was supported by 'FFR 2021 GIADA ADELFFIO'.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adelfio, G., Chiodi, M.: Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stoch. Environ. Res. Risk Assess.* **29**(2), 443–450 (2015)
- Adelfio, G., Chiodi, M.: Flp estimation of semi-parametric models for space-time point processes and diagnostic tools. *Spat. Stat.* **14**, 119–132 (2015)
- Adelfio, G., Chiodi, M.: Including covariates in a space-time point process with application to seismicity. *Stat. Methods Appl.* **30**, 947–971 (2021)
- Adelfio, G., Schoenberg, F.P.: Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Ann. Inst. Stat. Math.* **61**(4), 929–948 (2009)
- Adelfio, G., Siino, M., Mateu, J., Rodríguez-Cortés, F.J.: Some properties of local weighted second-order statistics for spatio-temporal point processes. *Stoch. Environ. Res. Risk Assess.* **34**, 1–20 (2019)
- Alm, S.E.: Approximation and simulation of the distributions of scan statistics for poisson processes in higher dimensions. *Extremes* **1**(1), 111–126 (1998)
- Amato, A., Chiarabba, C., Selvaggi, G., et al.: Crustal and deep seismicity in Italy (30 years after) (1997)
- Anwar, S., Stein, A., van Genderen, J.: Implementation of the marked Strauss point process model to the epicenters of earthquake aftershocks. In: Leung, Y. (ed.) *Advances in Geo-Spatial Information Science*, pp. 125–140. Taylor & Francis, London (2012)
- Baddeley, A.: Local composite likelihood for spatial point processes. *Spat. Stat.* **22**, 261–295 (2017)
- Baddeley, A.: spatstat.local: Extension to 'spatstat' for Local Composite Likelihood. R package version 3.6-0 (2019)
- Baddeley, A., Chang, Y.-M., Song, Y., Turner, R.: Nonparametric estimation of the dependence of a spatial point process on spatial covariates. *Stat. Interface* **5**(2), 221–236 (2012)
- Baddeley, A., Møller, J.: Nearest-neighbour markov point processes and random sets. *Int. Stat. Rev.* **57**(2), 89–121 (1989)
- Baddeley, A., Møller, J., Waagepetersen, R.: Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat. Neerlandica* **54**(3), 329–350 (2000)
- Baddeley, A., Rubak, E., Turner, R.: *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London (2015)
- Baddeley, A., Turner, R.: spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* **12**(6), 1–42 (2005)
- Baddeley, A., Turner, R., Mateu, J., Bevan, A.: Hybrids of gibbs point process models and their implementation. *J. Stat. Softw.* **55**(11), 1–43 (2013)
- Baddeley, A., Turner, R., Møller, J., Hazelton, M.: Residual analysis for spatial point processes (with discussion). *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **67**(5), 617–666 (2005)
- Baddeley, A., Turner, T.R.: Practical maximum pseudo likelihood for spatial point patterns (with discussion). *Aust. New Zealand J. Stat.* **42**(3), 283–322 (2000)
- Banfield, J.D., Raftery, A.E.: Model-based gaussian and non-gaussian clustering. *Biometrics* **49**, 803–821 (1993)
- Berman, M.: Testing for spatial association between a point process and another stochastic process. *Appl. Stat.* **35**(1), 54–62 (1986)

- Berman, M., Diggle, P.: Estimating weighted integrals of the second-order intensity of a spatial point process. *J. R. Stat. Soc. Ser. B (Methodol.)* **51**(1), 81–92 (1989)
- Brix, A., Diggle, P.J.: Spatiotemporal prediction for log-gaussian cox processes. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **63**(4), 823–841 (2001)
- Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J.: *Stochastic Geometry and Its Applications*. Wiley (2013)
- Cox, D.: *The Statistical Analysis of Dependencies in Point Processes, Stochastic Point Processes*, pp. 55–66. Wiley, New York (1972)
- Cox, D.R.: Some statistical methods connected with series of events. *J. R. Stat. Soc. Ser. B (Methodol.)* **17**, 129–164 (1955)
- Cressie, N.: *Statistics for Spatial Data*. Wiley (2015)
- Daley, D.J., Vere-Jones, D.: *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, 2nd edn. Springer, New York (2007)
- Dasgupta, A., Raftery, A.E.: Detecting features in spatial point processes with clutter via model-based clustering. *J. Am. Stat. Assoc.* **93**(441), 294–302 (1998)
- Di Stefano, R., Chiarabba, C., Lucente, F., Amato, A.: Crustal and uppermost mantle structure in italy from the inversion of p-wave arrival times: geodynamic implications. *Geophys. J. Int.* **139**(2), 483–498 (1999)
- Diggle, P.: A kernel method for smoothing point process data. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **34**(2), 138–147 (1985)
- Diggle, P.J.: On parameter estimation and goodness-of-fit testing for spatial point patterns. *Biometrics* **35**, 87–101 (1979)
- Diggle, P.J.: *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*. CRC Press (2013)
- Diggle, P.J., Gratton, R.J.: Monte carlo methods of inference for implicit statistical models. *J. R. Stat. Soc. Ser. B (Methodol.)* **46**(2), 193–212 (1984)
- Diggle, P.J., Moraga, P., Rowlingson, B., Taylor, B.M., et al.: Spatial and spatio-temporal log-gaussian cox processes: extending the geostatistical paradigm. *Stat. Sci.* **28**(4), 542–563 (2013)
- Efron, B.: *The Jackknife, the Bootstrap, and other Resampling Plans*, vol. 38. SIAM (1982)
- Eguchi, S.: Second order efficiency of minimum contrast estimators in a curved exponential family. *Ann. Stat.* **11**, 793–803 (1983)
- Gelfand, A.E., Diggle, P., Guttorp, P., Fuentes, M.: *Handbook of Spatial Statistics*. CRC Press (2010)
- Giunta, G., Luzio, D., Agosta, F., Caldò, M., Di Trapani, F., Giorgianni, A., Oliveri, E., Orioli, S., Pernicciaro, M., Vitale, M., Chiodi, M., Adelfio, G.: An integrated approach to investigate the seismotectonics of northern sicily and southern tyrrhenian. *Tectonophysics* **476**, 13–21 (2009)
- Guan, Y.: A composite likelihood approach in fitting spatial point process models. *J. Am. Stat. Assoc.* **101**(476), 1502–1512 (2006)
- Hall, J., Aksu, A., Yaltirak, C.: Miocene to recent tectonic evolution of the eastern mediterranean; new pieces of the old mediterranean puzzle. *Mar. Geol.* **221**(1–4), 1–13 (2005)
- Hawkes, A., Adamopoulos, L.: Cluster models for earthquakes-regional comparison. *Bull. Int. Stat. Inst.* **45**(3), 454–461 (1973)
- Illian, J., Penttinen, A., Stoyan, H., Stoyan, D.: *Statistical Analysis and Modelling of Spatial Point Patterns*, vol. 70. Wiley (2008)
- Le Pichon, X., Angelier, J.: The hellenic arc and trench system: a key to the neotectonic evolution of the eastern mediterranean area. *Tectonophysics* **60**(1), 1–42 (1979)
- Liu, M., Stein, S.: Earthquake, aftershocks. *Encyclopedia Solid Earth Geophys.*, 1–4 (2019)
- Loader, C.R., et al.: Bandwidth selection: Classical or plug-in? *Ann. Stat.* **27**(2), 415–438 (1999)
- Lutikov, A., Rogozhin, E., Dontsova, G.Y., Kuchai, M.: Strong earthquakes in 2009–2016 in central italy: tectonic position, seismic history, and aftershock processes. *Izvestiya Phys. Solid Earth* **54**(2), 233–251 (2018)
- Meyer, P.-A.: Démonstration simplifiée d'un théorème de knight. *Séminaire de probabilités de Strasbourg* **5**, 191–195 (1971)
- Møller, J.: Shot noise cox processes. *Adv. Appl. Probab.* **35**, 614–640 (2003)
- Møller, J., Syversveen, A.R., Waagepetersen, R.P.: Log gaussian cox processes. *Scand. J. Stat.* **25**(3), 451–482 (1998)
- Møller, J., Toftaker, H.: Geometric anisotropic spatial point pattern analysis and cox processes. *Scand. J. Stat.* **41**(2), 414–435 (2014)
- Ogata, Y.: Statistical models for earthquake occurrences and residual analysis for point processes. *J. Am. Stat. Assoc.* **83**(401), 9–27 (1988)
- Ogata, Y.: Statistical model for standard seismicity and detection of anomalies by residual analysis. *Tectonophysics* **169**(1–3), 159–174 (1989)

- Ogata, Y., Katsura, K.: Likelihood analysis of spatial inhomogeneity for marked point patterns. *Ann. Inst. Stat. Math.* **40**(1), 29–39 (1988)
- Ogata, Y., Katsura, K.: Maximum likelihood estimates of the fractal dimension for random spatial patterns. *Biometrika* **78**, 463–474 (1991)
- Oya, P.: Geodynamic assessment of eastern mediterranean region: a joint gravity and seismic b value approach. *Arab. J. Geosci.* **9**(5), 360 (2016)
- Papazachos, C., Hatzidimitriou, P., Panagiotopoulos, D., Tsokas, G.: Tomography of the crust and upper mantle in southeast europe. *J. Geophys. Res. Solid Earth* **100**(B7), 12405–12422 (1995)
- Pfanzagl, J.: On the measurability and consistency of minimum contrast estimates. *Metrika* **14**(1), 249–272 (1969)
- R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2019)
- Ripley, B.D.: The second-order analysis of stationary point processes. *J. Appl. Probab.* **13**, 255–266 (1976)
- Ripley, B.D.: Modelling spatial patterns (with discussion). *J. R. Stat. Soc. Ser. B* **39**(2), 172–212 (1977)
- Ripley, B.D.: *Statistical Inference for Spatial Processes*. Cambridge University Press (1988)
- Ryan, W.B., Kastens, K.A., Cita, M.B.: Geological evidence concerning compressional tectonics in the eastern mediterranean. *Tectonophysics* **86**(1–3), 213–242 (1982)
- Schabenberger, O., Gotway, C.A.: *Statistical Methods for Spatial Data Analysis*. Chapman and Hall/CRC (2017)
- Schoenberg, F.: Transforming spatial point processes into poisson processes. *Stoch. Process. Appl.* **81**(2), 155–164 (1999)
- Schoenberg, F.P.: Multidimensional residual analysis of point process models for earthquake occurrences. *J. Am. Stat. Assoc.* **98**(464), 789–795 (2003)
- Sgroi, T., de Nardis, R., Lavecchia, G.: Crustal structure and seismotectonics of central sicily (southern italy): new constraints from instrumental seismicity. *Geophys. J. Int.* **189**(3), 1237–1252 (2012)
- Siino, M., Adelfio, G., Mateu, J.: Joint second-order parameter estimation for spatio-temporal log-gaussian cox processes. *Stoch. Environ. Res. Risk Assess.* **32**(12), 3525–3539 (2018)
- Siino, M., Adelfio, G., Mateu, J., Chiodi, M., D'Alessandro, A.: Spatial pattern analysis using hybrid models: an application to the hellenic seismicity. *Stoch. Environ. Res. Risk Assess.* **31**, 1633–48 (2016)
- Spakman, W., Wortel, M., Vlaar, N.: The hellenic subduction zone: a tomographic image and its geodynamic implications. *Geophys. Res. Lett.* **15**(1), 60–63 (1988)
- Strauss, D.J.: A model for clustering. *Biometrika* **62**(2), 467–475 (1975)
- Tanaka, U., Ogata, Y., Stoyan, D.: Parameter estimation and model selection for neyman-scott point processes. *Biometrical J. J. Math. Methods Biosci.* **50**(1), 43–57 (2008)
- Taymaz, T., Jackson, J., McKenzie, D.: Active tectonics of the north and central aegean sea. *Geophys. J. Int.* **106**(2), 433–490 (1991)
- Van Lieshout, M.: *Markov Point Processes and Their Applications*. World Scientific (2000)
- Veen, A.: *Some Methods of Assessing and Estimating Point Processes Models for Earthquake Occurrences*. University of California, Los Angeles (2006)
- Veen, A., Schoenberg, F.P.: Assessing spatial point process models using weighted k-functions: analysis of california earthquakes. In: *Case Studies in Spatial Point Process Modeling*. Springer, pp. 293–306 (2006)
- Vere-Jones, D.: Earthquake prediction—a statistician's view. *J. Phys. Earth* **26**(2), 129–146 (1978)
- Waagepetersen, R., Guan, Y.: Two-step estimation for inhomogeneous spatial point processes. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **71**(3), 685–702 (2009)
- Waagepetersen, R.P.: An estimating function approach to inference for inhomogeneous neyman-scott processes. *Biometrics* **63**(1), 252–258 (2007)
- Walsh, D.C., Raftery, A.E.: Classification of mixtures of spatial point processes via partial bayes factors. *J. Comput. Graph. Stat.* **14**(1), 139–154 (2005)
- Walsh, D.C.I., Raftery, A.E.: Detecting mines in minefields with linear characteristics. *Technometrics* **44**(1), 34–44 (2002)
- Woessner, J., Hardebeck, J., Hauksson, E.: What is an instrumental seismicity catalog. *Commun. Online Resour. Stat. Seism. Anal.* (2010)
- Ye, X., Yu, J., Wu, L., Li, S., Li, J.: Open source point process modeling of earthquake. In: *Geo-Informatics in Resource Management and Sustainable Ecosystem*. Springer, pp. 548–557 (2015)