Advanced spatio-temporal point processes for the Sicily seismicity analysis

Processi puntuali spatio-temporali avanzati per l'analisi della sismicità in Sicilia

Marianna Siino and Giada Adelfio

Abstract Due to the complexity of the generator process of seismic events, we study under several aspects the interaction structure between earthquake events using recently developed spatio-temporal statistical techniques and models. Using these advanced statistical tools, we aim to characterise the global and local scale cluster behaviour of the Easter Sicily seismicity considering the catalogue data since 2006, when the Italian National Seismic Network was upgraded and earthquake location was sensibly improved. Firstly, we characterise the global complex spatio-temporal interaction structure with the space-time ETAS model where background seismicity is estimated non-parametrically, while triggered seismicity is estimated by MLE. After identifying seismic sequences by a clustering technique, we characterise their spatial and spatio-temporal interaction structures using other advanced point process models. For the characterisation of the spatial interactions, a version of hybrid of Gibbs point process models is proposed as method to describe the multiscale interaction structure of several seismic sequences accounting for both the attractive and repulsive nature of data. Furthermore, we consider log-Gaussian Cox processes (LGCP), that are relatively tractable class of empirical models for describing spatio-temporal correlated phenomena. Several parametric formulation of spatio-temporal LGCP are estimated, by the minimum contrast procedure, assuming both separable and non-separable parametric specification of the correlation function of the underlying Gaussian Random Field.

Abstract *Il processo generatore dei fenomeni sismici è caratterizzato da una certa complessità. In questo lavoro studieremo sotto diversi punti di vista la struttura di interazione tra i terremonti utilizzando tecniche e modelli statistici di tipo spazio temporale. In particolare, le analisi si pongono come obiettivo quello di caratteriz-*

Marianna Siino
Università degli studi di Palermo, Dipartimento di Scienza Economiche, aziendali e statistiche e-mail: marianna.siino01@unipa.it

Giada Adelfio
Università degli studi di Palermo, Dipartimento di scienza economiche, aziendali e statistiche e-mail: giada.adelfio@unipa.it

zare il comportamento a piccola e a grande scale degli eventi sismici avvenuti a partire dal 2006 nella Sicilia Orientale. In prima istanza, si caratterizza l'andamento globale della sismicità nell'area di studio utilizzando il modello spazio-temporale ETAS, in cui la sismicità di fondo viene stimata in modo non parametrico mentre la sismicità indotta viene stimata mediante massima verosimiglianza. Dopo aver identificato sequenze sismiche, con un opportuno metodo di clustering, per ciascuna sequenza si studia la struttura di relazione che intercorre tra gli eventi dal punto di vista sia spaziale che spazio-temporale. Per quanto riguarda lo studio delle interazioni spaziali, una versione dei modelli ibridi Gibbs viene proposta per descrivere la struttura di interazione multiscala delle sequenze tenendo conto allo stesso tempo del comportamento attrattivo e repulsivo che intercorre tra i punti. Inoltre, abbiamo considerato la classe dei modelli di tipo log-Gaussian Cox, utili per la descrizione di fenomeni correlati nello spazio e nel tempo. Utilizzando il metodo dei momenti, sono stati stimati e in seguito confrontati diversi modelli parametrici assumendo una struttura di correlazione del processo Gaussiano sottostante sia separabile che non-separabile.

Key words: earthquakes; hybrid of Gibbs process; log-Gaussian Cox processes; minimum contrast method; non-separable covariance function; point process; spatio-temporal pair correlation function

1 Goals of the analysis and description of the study area

To describe and predict seismic events in space and time, a proper characterisation of both the intensity function and the second-order properties of the generator process is a crucial issue. In this work, we give a general overview of advanced statistical models that can be used in the seismic context, showing the structure of the models, the main diagnostic tools and the available code (Siino et al, 2018a,b). In an application example, we aim to characterise under several aspects the interaction structure observed in the Sicily catalogue events using proper point process models.

First, we describe the global complex interaction structure using the space-time Epidemic Type Aftershock Sequence model (Ogata, 1988). Focusing on a local scale, we identify some sequences of events characterising their spatial and spatio-temporal structure. In particular, hybrid of Gibbs point processes (Baddeley et al, 2015) are used to describe the distribution of epicentres that generally exhibit interaction at different spatial scales. On the other hand, the spatio-temporal cluster structure is characterised by spatio-temporal log-Gaussian Cox models (Diggle et al, 2013), where the main issue concerns the specification of the moments of the underlying Gaussian Random Field (GRF). Moreover, in the remain of the paragraph we give a general overview of the study area. It is focused on eastern Sicily: it extends from 36.5° to 39° Lat. N and from 14° to 16° Long. E. East Sicily and South Calabria is the area with greater deformation rate and seismic strain release in Italy. We consider the seismic catalogue since 2006, when the Italian National Seismic

Network was upgraded and earthquake location was sensibly improved. The instrumental seismicity recorded in the period from 2006 to 2016, consists of 12356 events; 4170 events (33.7 %) have M between 2.1 and 3.0, 333 events (2.7 %) have M between 3.1 and 4.0, and just 28 events (0.2 %) have $M > 4$. The most of the events (75.4%) are crustal, with hypocentral depth lower than 30 km (Figure 1).

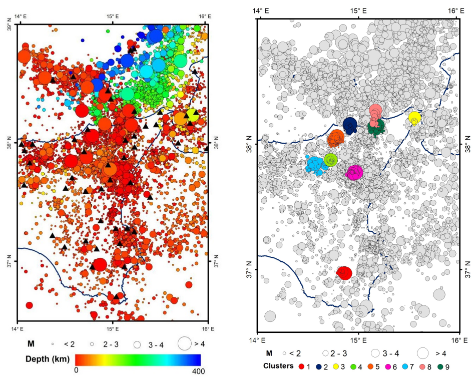


Fig. 1: (left) Seismicity map from 2006 to 2016 period, black triangles indicate the seismic stations. (right) Distribution of the detected sequences.

2 Global scale analysis: Epidemic Type Aftershocks-Sequences model

Let $\mathscr{X}$ be a random countable subset of $\mathbb{R}^2 \times \mathbb{R}^+$ where for a point $(\mathbf{u},t) \in \mathscr{X}$, $\mathbf{u} \in \mathbb{R}^2$ is the location and $t \in \mathbb{R}^+$ is the time of occurrence. In practice, an observed spatio-temporal pattern is a finite set $\{(\mathbf{u}_i,t_i)\}_{i=1}^n$ of distinct points within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$, where usually W is a polygon with area $|W| > 0$ and T a single closed interval with length $|T| > 0$. The space-time point process is defined as a random point pattern, completely characterized by its conditional intensity function (CIF). The CIF of the ETAS model (Ogata, 1988), in time t and location $\mathbf{u}$ conditional to $\mathscr{H}_t$ (the history until t) is defined as the sum of a term describing the long-term variation and one relative to the short-term variation,

$$\lambda_{\boldsymbol{\theta}}(t,\mathbf{u}|\mathscr{H}_t) = \mu f(\mathbf{u}) + \sum_{t_j<t} \frac{\kappa_0 \, e^{(\alpha)\;(m_j-m_0)}}{(t-t_j+c)^p} \left\{||\mathbf{u}-\mathbf{u}_j||^2 + d\right\}^{-q} \tag{1}$$

The background activity is characterized by μ, that indicates the general background intensity, and $f(\mathbf{u})$ that is the space density. Instead, for the induced intensity, κ_0 is a constant which measures the aftershocks productivity, c and p are parameters of the modified Omori's law (p characterises the pattern seismicity in the given region indicating the decay rate of aftershocks in time), α measures the influence on the relative weight of each sequence, m_j is the magnitude of the inducing event and m_0 is the completeness threshold of magnitude. The parameters are estimated with the Forward Likelihood-based predictive approach implemented in the `R` package `etasFLP` (Chiodi and Adelfio, 2014). Figure 2 shows the estimated spatial background intensity and the induced one with $m_0 = 2.7$. According to the estimated model (Figure 2), there is quicker decay in space than time since $q > p$, and there is a positive effect of the magnitude of the previous events since $\alpha > 0$.

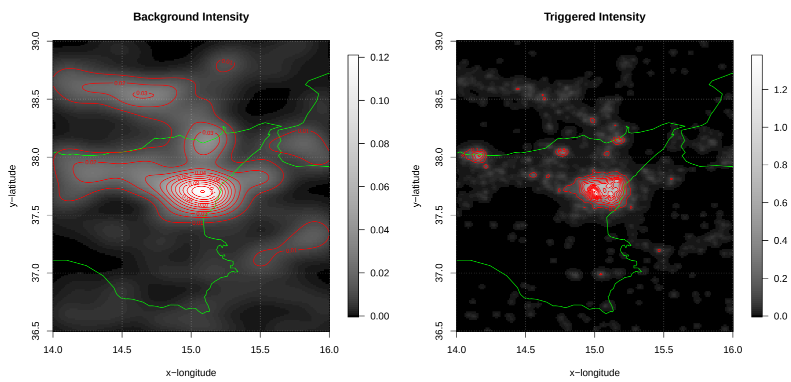


Fig. 2: Estimated background and triggered space intensities, with estimates $\boldsymbol{\theta} = \{0.051, 0.049, 0.005, 1.075, 0.931, 7.510, 1.796\}$ and corresponding standard errors $s.e.(\boldsymbol{\theta}) = \{0.003, 0.017, 0.001, 0.013, 0.113, 1.174, 0.073\}$.

3 Local scale analysis

For the local scale analysis, we report the main results obtained for the cluster 5 of Figure 1 occurred close to the village San Salvatore di Fitalia in 2011. The main-shock event was the 23^{rd} of June 2011 with M=4.5 and 7 km depth. For this sequence, the aims are to characterise the spatial multiscale interaction structure using hybrid models and the spatio-temporal evolution using log-Gaussian Cox models.

3.1 Hybrid of spatial Gibbs process

The class of Gibbs processes $\mathscr{X}$ is determined through a probability density function $f : \mathscr{X} \to [0, \infty)$, where $\mathscr{X} = \{\boldsymbol{\upsilon} \subset W : n(\boldsymbol{\upsilon}) < \infty\}$ is a set of point configurations contained in W. In the literature several Gibbs models have been proposed such as the area-interaction, Strauss, Geyer, hard core processes. However Gibbs processes have some drawbacks when points have a strong clustering and show spatial dependence at multiple scales (Baddeley et al, 2013). Baddeley et al (2013) propose hybrid models as a general way to generate multi-scale processes combining Gibbs processes. Given m unnormalized densities $f_1, f_2, \ldots, f_m$, the hybrid density is defined as $f(\boldsymbol{\upsilon}) = f_1(\boldsymbol{\upsilon}) \times \ldots \times f_m(\boldsymbol{\upsilon})$, respecting some properties. For example, the density of the inhomogeneous hybrid process obtained considering m Geyer components (with interaction ranges $r_1, \ldots r_m$ and saturation parameters $s_1, \ldots, s_m$) is

$$f(\boldsymbol{\upsilon}) = \prod_{i=1}^{n(\boldsymbol{\upsilon})} \beta(\boldsymbol{u}_i) \prod_{j=1}^{m} \gamma_j^{min(s_j, t(\mathbf{u}_i, \boldsymbol{\upsilon} \backslash \mathbf{u}_i; r_j))} \tag{2}$$

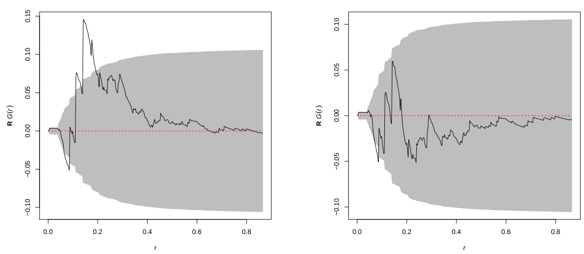
where $t(\mathbf{u}_i, \boldsymbol{\upsilon} \backslash \mathbf{u}_i; r_j) = \sum_i \{\mathbf{1} \|\mathbf{u} - \mathbf{u}_i\| \leq r_j\}$. This density indicates that the spatial interaction between points changes with the distances r_j and the parameters that capture this information are the interaction parameters γ_j. Gibbs models can be fitted to data by pseudolikelihood that is function of the Papangelou conditional intensity $\rho_\phi(\mathbf{u}|\boldsymbol{\upsilon})$ at location $\mathbf{u} \in W$ given $\boldsymbol{\upsilon}$, where ϕ are the parameters to estimate (Baddeley et al, 2015). In hybrid models, the conditional intensity is

$$\rho_\phi(\mathbf{u},; \upsilon) = \exp\{B(\mathbf{u}) + \theta_1^T V_1(\mathbf{u}, \eta) + \theta_2^T G(\mathbf{u}, \boldsymbol{\upsilon}, \eta)\}, \tag{3}$$

where $B(\mathbf{u})$ is an offset term, $\theta_1^T V_1(\mathbf{u}, \eta)$ is the first-order potential and $\theta_2^T G(u, \boldsymbol{\upsilon}, \eta)$ accounts for the interaction effects. In the following analysis, it will be a combination of Geyer processes, and in this case the irregular parameter η accounts both for an interaction distance r_j and a saturation parameter s_j for each j-th Geyer component. Usually, to assess the goodness-of fit, the estimated models are compared in terms of AIC, spatial raw residuals and number of simulated points under the estimated model (Baddeley et al, 2015). Furthermore, the diagnostic plots based on the residual K- and G-functions are used to decide which component has to be added at each step to the hybrid model (Baddeley et al, 2015). In the `spatstat` package (Baddeley et al, 2015) of R (R Development Core Team, 2005), there are the most of the functions that have been used for fitting, prediction, simulation and validation of Hybrid models. Given a set of points, the first attempt in model estimation is to fit an inhomogeneous Poisson model with intensity depending on the spatial coordinates in which the points do not interact with each other, that is equal to equation (3) where the term $G(\cdot)$ is null. For the selected sequence, the corresponding AIC and the range of the spatial raw residuals of the fitted inhomogeneous Poisson models with a non-parametric spatial trend are in Table 1. In Figure 3a, the residual G-function for the estimated inhomogeneous Poisson model is reported. For distances up to about 200 meters, it wanders substantially outside the limits showing peaks, so there is a positive association between points unexplained by the Poisson model. Therefore, we consider hybrid of Geyer processes, and the main estimation results are in Table 1. Considering the hybrid model, there is an improvement in terms of AIC, of range of raw residuals, and of residual G-function (Figure 3b), it oscillates around zero and it is inside the envelopes indicating that the multiscale interaction structure between the earthquakes is well described by the hybrid model of two Geyer processes.

Table 1: Inhomogeneous Poisson and the hybrid of Gibbs model: AIC, range of spatial raw residuals, vector of irregular parameter η and interaction parameter for each Geyer component γ_j.

Inhom. Poisson model		Inhom. hybrid of Gibbs processes				
AIC	Range res.	Comp.	η	γ	AIC	Range res.
-961.829	[-0.142 ; 0.531]	G_1	$(r_1;s_1) = (0.060;0.2)$	0.413	-990.565	[-0.258 ; 0.171]
		G_2	$(r_2;s_2) = (0.240;3.5)$	1.157		



(a) (b)

Fig. 3: For the sequence 5, (a) residual G-functions for the inhomogeneous Poisson process model and for the inhomogeneous hybrid model of Geyer processes (b).

3.2 *Spatio-temporal log-Gaussian Cox process*

The class of log-Gaussian Cox processes (LGCPs) (Møller et al, 1998) is a flexible and relatively tractable class of models for describing correlated phenomena specifying the moments of an underlying Gaussian Random Field (GRF). Considering a spatio-temporal point process as in Section 2, $\mathscr{X}$ is said to be a Cox process driven by Λ (a non-negative and locally integrable stochastic process), if the conditional distribution of the point process $\mathscr{X}$ given a realization $\Lambda(\mathbf{u},t) = \lambda(\mathbf{u},t)$ is a Poisson process on $W \times T$ with intensity function $\lambda(\mathbf{u},t)$. Following the homogeneous specification in Diggle et al (2013), the log-Gaussian Cox process for a generic point in space and time has the following intensity $\Lambda(\mathbf{u},t) = \exp\{\beta + S(\mathbf{u},t)\}$ where S is a Gaussian process with $\mathbb{E}(S(\mathbf{u},t)) = \mu = -0.5\sigma^2$ and so $\mathbb{E}(\exp\{S(\mathbf{u},t)\}) = 1$ and with variance-covariance matrix $\mathbb{C}(S(\mathbf{u},t), S(\mathbf{v},s)) = \mathbb{C}(||\mathbf{u}-\mathbf{v}||, |t-s|) = \sigma^2\gamma(r,h)$ under the stationary assumption, where $\gamma(\cdot)$ is the correlation function of the GRF. Following Møller et al (1998), the first-order product density and the pair correlation function of the log-Gaussian Cox process are $\mathbb{E}(\Lambda(\mathbf{u},t)) = \lambda = \exp(\beta)$ and $g(r,h) = \exp(\sigma^2\gamma(r,h))$, respectively.

To describe the spatio-temporal correlation structure of the selected sequence (Figure 1), we consider several formulation of LGCP models changing the specification of the covariance structure of the underline GRF assuming both separable (exponential structure in space and time) and non-separable parametric specifications (Gneiting and Iaco-Cesare families). Table 2 shows the estimated parameters for each model obtained with the minimum contrast method proposed in Siino et al (2018a). To assess the goodness-of-fit of the estimated models, we use the 95% envelopes of the non-parametric spatio-temporal K-function (Figure 4) and a global test based on Monte Carlo procedure to check if the spatio-temporal point pattern is a realisation of a LGCP with the specified parameters. The fitted models assuming a separable structure and the Gneting families seem to do not describe properly the data because the observed spatio-temporal K-function is outside the envelope, see Figures 4a and 4c. This is confirmed by the global p-values in Table 2. A LGCP model with Iaco-Cesare structure for the underlying GRF describes better the spatio-temporal cluster structure according to Figure 4b and the global p-value that is 0.10 (Table 2). According to the estimates for the selected homogeneous LGCP model, there is a decay of the interaction in space and time governed mainly by the scale parameters α and β.

4 Comments

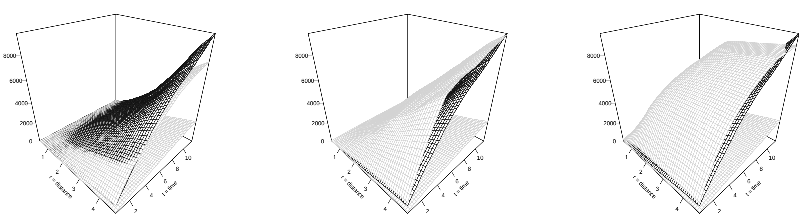
We propose the use of different models for the analysis of seismic data, focussing both on the global and local scales. In the first case the ETAS model is considered, providing a space-time characterization of the seismic activity in the considered area. For the local scale analysis of the sequence, we consider both Hybrid and LGCP models. Their results are not comparable, since the two model formulation focus on different aspects. In particular, with the hybrid model, we focus on a spatial scale and the multiscale interaction structure is properly described by a hybrid of Geyer processes after taking into account the spatial inhomogeneity with a non-parametric kernel. Moreover, there is evidence of a strong clustering structure for short distances up to 200 meters. On the other hand, by using the LGCP model, spatial and temporal evolution are simultaneously described specifying the moments of a GRF, without assuming a deterministic part in the intensity and considering a homogeneous formulation of the model.

References

Baddeley A, Turner R, Mateu J, Bevan A (2013) Hybrids of gibbs point process models and their implementation. Journal of Statistical Software 55(11):1–43

Table 2: Estimates of the LGCP models assuming several covariance parametric families for the GRF. The p-values are shown to assess if the point pattern comes from the assumed spatio-temporal LGCP model.

Cluster	Family	$\hat{\sigma}^2$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}_s$	$\hat{\gamma}_t$	$\hat{\delta}$	p-value
Exp-Exp	$\mathbb{C}(r,h) = \sigma^2 \exp\left(-\frac{r}{\alpha}\right) \exp\left(-\frac{h}{\beta}\right)$	3.075	12.025	14.934				0.01
Iaco-Cesare	$\mathbb{C}(r,h) = \sigma^2 \left(1 + \left(\frac{r}{\alpha}\right)^{\gamma_s} + \left(\frac{h}{\beta}\right)^{\gamma_t}\right)^{-\delta}$	3.237	13.709	15.910	0.983	1.022	1.500	0.10
Gneiting	$\mathbb{C}(r,h) = \frac{\sigma^2}{\left(\left(\frac{h}{\beta}\right)^{\gamma_t}+1\right)^{\delta/\gamma_t}} \exp\left(-\frac{\left(\frac{r}{\alpha}\right)^{\gamma_s}}{\left(\left(\frac{h}{\beta}\right)^{\gamma_t}+1\right)^{\delta/(2\gamma_t)}}\right)$	3.093	2.252	2.955	1.504	1.701	0.354	0.08



(a) Exp-Exp (b) Iaco-Cesare (c) Gneiting

Fig. 4: 95% envelope of the spatio-temporal K-function based on simulated spatio-temporal LGCP patterns according to the estimated parameters in Table 2.

Baddeley A, Rubak E, Turner R (2015) Spatial Point Patterns: Methodology and Applications with R. London: Chapman and Hall/CRC Press

Chiodi M, Adelfio G (2014) etasflp: Estimation of an etas model. mixed flp (forward likelihood predictive) and ml estimation of non-parametric and parametric components of the etas model for earthquake description. R package version 10

Diggle PJ, Moraga P, Rowlingson B, Taylor BM (2013) Spatial and spatio-temporal log-Gaussian Cox processes: extending the geostatistical paradigm. Statistical Science 28(4):542–563

Møller J, Syversveen AR, Waagepetersen RP (1998) Log-Gaussian Cox processes. Scandinavian Journal of Statistics 25(3):451–482

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association 83(401):9–27

R Development Core Team (2005) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org, ISBN 3-900051-07-0

Siino M, Adelfio G, Mateu J (2018a) Joint second-order parameter estiamation for spatio-temporal log-gaussian cox processes. (submitted)

Siino M, D'Alessandro A, Adelfio G, Scudero S, Chiodi M (2018b) Multiscale processes to describe the eastern sicily seismic sequences. submitted