



Including covariates in a space-time point process with application to seismicity

Giada Adelfio^{1,2}  · Marcello Chiodi^{1,2}

Accepted: 5 July 2020 / Published online: 18 July 2020
© The Author(s) 2020

Abstract

The paper proposes a spatio-temporal process that improves the assessment of events in space and time, considering a contagion model (branching process) within a regression-like framework to take covariates into account. The proposed approach develops the forward likelihood for prediction method for estimating the ETAS model, including covariates in the model specification of the epidemic component. A simulation study is carried out for analysing the misspecification model effect under several scenarios. Also an application to the Italian seismic catalogue is reported, together with the reference to the developed R package.

Keywords Space-time point processes · ETAS model · R package for seismic data · Covariates

1 Introduction

Contagious phenomena are well described in space and time by self-exciting point processes, such that the conditional intensity function is obtained as the sum of the long-term variation component (the so-called endemic) and the short-term variation one (the epidemic part). This kind of models have been widely used in the literature: infectious disease (Paul et al. 2008; Paul and Held 2011; Meyer et al. 2012, 2017), crime (Mohler et al. 2011), quakes (Ogata 1998; Adelfio and Chiodi 2009; Adelfio and Ogata 2010; Adelfio and Chiodi 2015a; Zhuang et al. 2002). To model earthquake activity in space and time accounting both for the endemic (background activity) and epidemic (aftershocks) effect, the Epidemic-Type Aftershock Sequences (ETAS) model (Ogata 1988, 1998) is widely used, describing events starting from their space-time coordinates (and magnitude as mark) and

✉ Giada Adelfio
giada.adelfio@unipa.it

Extended author information available on the last page of the article

incorporating seismological laws in a mechanistic approach, as a natural approach in the context of earthquake data.

In this paper, we aim at providing an improved framework for further computational and theoretical development of the study and the description of epidemic phenomena. In particular, extending the model formulation proposed by Meyer et al. (2012) in the context of infectious disease transmission, we suggest the use of a specific branching-type model for earthquake description (the ETAS model) in a regression-oriented version modelling, accounting also for external covariates, expected to explain some of the overall variability of the studied phenomenon and lead to a decrease in the unpredictable variability. We apply the Forward Likelihood for prediction (FLP) method (Chiodi and Adelfio 2011) for estimating the ETAS model components, also introducing a covariate vector for the epidemic part, crucial for a more realistic description of the observed activity. Indeed from previous studies (e.g., Adelfio and Chiodi 2015b) on the basis of the diagnostic results, the need of a more flexible model for the triggered component of the ETAS model revealed, noticing that even if the background seismicity is well described by the FLP estimated intensity, at least if compared with existing methods, something is still missing in the description of the space-time triggered part. In our opinion, considering external information (such as geological information related to faults distribution) for the description of spatio-temporal earthquakes could be a promising direction of study for this field of research.

Previous studies tried to incorporate the external information in the self-exciting model: Meyer et al. (2012) proposes a model for the epidemic forecast with a linear predictor, where background function is composed of a function of population density and of a vector of covariates; Schoenberg (2016) proposed an application to southern California earthquake forecasting using weather data as an additive component; Reinhart and Greenhouse (2018) incorporate piecewise constant covariates only for the background component of the model.

In the survival analysis, an example of such two-component temporal point process regression modelling for independent individuals is the additive-multiplicative model (Lin and Ying 1995; Sasieni 1996), such that the conditional intensity consist of additive endemic (the risk of infection from external sources, independent of the past) and epidemic components (individual-to-individual transmission of the disease and thus depends on the internal history of the process). In these models, the endemic risk can be expressed as a function of external covariates, introduced as a multiplicative function (Höhle 2009; Martinussen and Scheike 2002).

In this paper, using the terminology of the survival analysis, we propose a more general additive-multiplicative model for the conditional intensity function of a space-time self-exciting point process with covariates which may vary also continuously in space, incorporating their effect in the triggering effect, using the FLP approach as in Chiodi and Adelfio (2017b) for the estimation of the background intensity. The methodology here proposed can be extended in any context, different from the original seismic one, e.g., the credit risk one, where there is a contagious effect of the previous history in space and time, together with specific covariates,

such as macroeconomic characteristics of the considered countries, as in Adelfio et al. (2018).

This paper is organized as follows. In Sect. 2 the theoretical background is defined, recalling some basic theory of space-time point processes and the ETAS model. In Sect. 3, the proposed extension of the ETAS model with the inclusion of covariates is provided. A simulation study with several scenarios is carried out to assess the performance of the proposed method (see Sect. 4). Finally, we present a seismic application, where external variables with respect to the usual event's coordinates may provide information for the description of the seismic activity of the defined area. Section 6 is devoted to conclusive remarks.

2 Point processes and ETAS model

A spatial-temporal point process is a random collection of points, whose realisations consists of a finite or a countably infinite set of points in a space-time domain. We consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^{d-1} \times \mathbb{R}$, where a point $(\mathbf{s}, t) \in X$ corresponds to an event at $\mathbf{s} \in \mathbb{R}^{d-1}$ occurring at time $t \in \mathbb{R}$. We observe n events $\{(\mathbf{s}_i, t_i)\}_{i=1}^n$ of distinct points of X occurring in n distinct times $\{t_i\}_{i=1}^n$ within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^{d-1} \times \mathbb{R}$, with volume $|W| > 0$, and with length $|T| > 0$ where $n \geq 0$ is not fixed in advance (for more details see Diggle 2013).

A spatial-temporal point process is uniquely characterized by its associated *conditional intensity function* (CIF), $\lambda_\theta(t, \mathbf{s} | \mathcal{H}_t)$, (Daley and Vere-Jones 2003) i.e., the instantaneous rate or hazard for events at time t and location \mathbf{s} given all the observations up to time t , conditioning on the random past history of the process. Let $N(B)$ denote the number of events of the process falling in a bounded region $B \subset W \times T$. A completely stationary spatio-temporal point process has a constant intensity λ , defined as $\lambda = E[N(B)]$, i.e., λ is the mean number of points per unit volume and unit time (Illian et al. 2008).

To model events that are clustered together, self-exciting point processes are often used. These models are used mainly to describe earthquakes characteristics, assuming that the occurrence of an event increases the probability of occurrence of other events in time and space. The Hawkes model (Hawkes and Adamopoulos 1973) and the ETAS model (Ogata 1988) are examples of self-exciting point processes.

The self-exciting process can be interpreted as a generalized Poisson cluster process associating to centres, of rate λ , a branching process of descendants. This kind of models are used to model reproduction phenomena and have been recently considered for the description of different applicative fields: biology (Caron-Lormier et al. 2006), demography (Johnson and Taylor 2008), epidemiology (Becker 1977; Balderama et al. 2012). According to a branching structure, the conditional intensity function of the self-exciting model is defined as the sum of a term describing the large-time scale variation (spontaneous activity or background, generally assumed homogeneous in time but not in space) and one relative to the

small-time scale variation due to the interaction with the events in the past (induced or triggered activity):

$$\lambda_{\theta}(t, \mathbf{s} | \mathcal{H}_t) = \mu f(\mathbf{s}) + \sum_{t_j < t} v_{\phi}(t - t_j, \mathbf{s} - \mathbf{s}_j) \quad (1)$$

with \mathcal{H}_t the past history of the process, $\theta = (\phi, \mu)'$, the vector of parameters of the induced intensity (ϕ) together with the parameter of the background general intensity (μ), $f(\mathbf{s})$ the spatial density and $v_{\phi}(\cdot)$ the spatial-temporal triggered density.

The triggering component of the model essentially provides a description of the intensity at a space-time location (t, \mathbf{s}) induced by each previous event, as a function of the spatial distances $\mathbf{s} - \mathbf{s}_j$ and temporal lags $t - t_j, \forall j$. In a clustered process, v_{ϕ} is a decreasing function of $\mathbf{s} - \mathbf{s}_j$ and $t - t_j$.

Simultaneously estimating the different components of the intensity function (large-time scale and small-time scale) in (1) is a main issue. If the large-time scale component $\mu f(\mathbf{s})$ is known, the parameters ϕ can be usually estimated by the Maximum Likelihood method. In applications, the large-time scale component $\mu f(\mathbf{s})$ is usually estimated through nonparametric techniques, like kernel estimators.

As introduced above, a branching process for earthquake description, widely used in seismological context, is the Epidemic Type Aftershocks-Sequences (ETAS) model (Ogata 1988, 1998). The ETAS conditional intensity function can be written, starting from model (1), as follows:

$$\lambda_{\theta}(t, \mathbf{s} | \mathcal{H}_t) = \mu f(\mathbf{s}) + \sum_{t_j < t} \frac{\kappa_0 \exp(\alpha(m_j - m_0))}{(t - t_j + c)^p} \left\{ (\mathbf{s} - \mathbf{s}_j)^2 + d \right\}^{-q} \quad (2)$$

The aftershock/induced component is the product of the density of aftershocks in time, i.e., the Omori law, representing the occurrence rate of aftershocks at time t , following the earthquake of time t_j and magnitude m_j , and the density of aftershocks in space. In particular, m_j is the magnitude of the j -th event and m_0 the threshold magnitude, that is, the lower bound for which earthquakes with higher values of magnitude are surely recorded in the catalogue, κ_0 is a normalizing constant, c and p are characteristic parameters of the seismic activity of the given region; p is useful for characterizing the pattern of seismicity, indicating the decay rate of aftershocks in time; d and q are two parameters related to the spatial influence of the mainshock. α is related to the expected number of offsprings generated by a single event, which is proportional to $\kappa_0 \exp\{\alpha(m_j - m_0)\}$.

For providing a simultaneous estimation of the background intensity and the triggered intensity components of an Epidemic-type model, in a previous paper, we developed the FLP approach (Adelfio and Chiodi 2015a). It is a nonparametric estimation procedure based on the subsequent increments of the log-likelihood obtained adding an observation one at a time, to account for the information of the observations until t_k on the next one. This approach allows the estimation of smoothing constants to get a reliable kernel estimate of the background intensity. The simultaneous estimation of the two parametric components of a branching-type

model alternating the standard parametric likelihood method with the FLP approach is here denoted by ETAS-FLP. Given the lack of specific open-source tools, the package *etasFLP* (Chiodi and Adelfio 2017a, b) provides tools to implement this mixed approach for a wide class of ETAS models for the description of seismic events, developed in a R environment.

In the provided methodology, the kernel background estimation is improved by weighting observations according to weights ρ_i , given by the ratio between the background and total intensity, for each observed event occurred at time t_i in location \mathbf{s}_i , according to the following relationship:

$$\hat{\rho}_i = \frac{\hat{\mu}f(\mathbf{s}_i)}{\lambda_{\hat{\theta}}(t_i, \mathbf{s}_i | \mathcal{H}_t)} \tag{3}$$

The quantity in Eq. (3) also gives an estimate of the probability that the i -th event E_i has been generated by the background component of the process.

3 The ETAS model with covariates

In this paper, we propose an additive-multiplicative model for the conditional intensity function of a space-time point process defined in $[0, T] \times W \subset \mathbb{R}^3, T > 0$, in the classical ETAS model framework.

In the ETAS model defined in (2), the expected number of offsprings generated by a single event is related to the magnitude of generating events, i.e., $\kappa_0 \exp\{\alpha(m_j - m_0)\}$. In this paper, we make it possible to include other covariates, besides the magnitude, related to every single event. Starting from the definition provided in Eq. (1), we propose to modify its additive formulation (sum of the “endemic” and “epidemic” parts), considering the offspring component in a novel regression view, that is, accounting for a vector of covariates. For interpretation convenience, in our proposal, we model covariates of the ETAS model as in a GLM framework, such that η_j is a classical linear predictor given by $\eta_j = \boldsymbol{\beta}'\mathbf{z}_j$, where \mathbf{z}_j is the vector of covariates observed for the j -th event and $\boldsymbol{\beta}$ is a vector of unknown parameters. This choice makes the methodology simple and easily extensible for any general linear predictor η .

As proposed by Meyer et al. (2012) in a context of infection occurrences, we incorporate the space-time phenomenological laws of the triggering part of ETAS model with the effects of covariates.

This triggering function is factorized into separate effects of marks, time, and relative location:

$$\lambda_{\theta}(t, \mathbf{s} | \mathcal{H}_t) = \mu f(\mathbf{s}) + \sum_{t_j < t} \frac{\kappa_0 \exp(\eta_j)}{(t - t_j + c)^p} \left\{ (\mathbf{s} - \mathbf{s}_j)^2 + d \right\}^{-q} \tag{4}$$

where (t_j, \mathbf{s}_j) is the time and location of individual occurrence j , $\eta_j = \boldsymbol{\beta}'\mathbf{Z}_j$ is a linear predictor, with \mathbf{Z}_j the external known covariate vector, including the magnitude (usually coinciding with the first covariate), acting in a multiplicative fashion on the

base risk and $\tilde{\theta} = (\mu, \kappa_0, c, p, d, q, \boldsymbol{\beta})'$, with $\boldsymbol{\beta}$ a k -components vector, to be estimated.

More in details, in the usual ETAS model, $k = 1$, $\mathbf{Z}_{j1} = m_j - m_0$, and $\beta_1 = \alpha$. In this model formulation, for an easier correspondence with the ETAS parametrization, in the $\boldsymbol{\beta}$ vector an intercept term is not included, because of the presence of the parameter κ_0 in the model.

In the seismic context, this extension would provide a more general formalism for the earthquake occurrence in space and time. Indeed, the main idea is that the effect on the future activity depends not only on the closeness of the previous events, but also on other characteristics of the main event, like magnitude, as usual, and also quadratic components like $(m_j - m_0)^2$, or the geometric distance from the generating faults, or other geological sources.

The extended version of the package *etasFLP* v. 2.0 for generalized offspring component is going to appear into the CRAN (Chiodi and Adelfio 2020). Indeed, the introduction of a general linear predictor did not introduce serious computational or theoretical difficulties, since $\boldsymbol{\beta}$ has been estimated with the parametric component in the *etasFLP* algorithm using the semi-parametric FLP approach.

4 Simulation study

The accuracy of the ETAS-FLP model with covariates can be evaluated under various conditions using simulations. In particular, we aim to assess the misspecification model effect, that is the properties of estimators of an ETAS model when the estimated linear predictor η_j is different from the real one.

Specifically, we consider the case where an ETAS model is simulated in the presence of some covariate (as in Eq. 4) and then for each simulation the model is estimated by the proposed approach as though the covariates were completely ignored and compared with the estimation under the correct model. The issue is whether the parameters of the ETAS model can be accurately estimated though the model is misspecified.

Indeed, inferential theoretical results for the proposed model are quite difficult to perform, since properties of the sampling distribution of the estimated quantities in the ETAS model are not known, but asymptotic general results are known (Ogata 1978; Rathbun 1996). Furthermore, in observed seismic catalogues, the presence of the nonparametric estimation of the background seismicity represents a further complication for the study of the parametric component properties.

In this section, we provide results of a simulation study for describing the properties of the estimators of the ETAS model parameters. However, this study holds under some assumptions, and a reasonable set of *true* values of the parameters. For example, the choice of the values of the parameters $\mu, \kappa_0, c, p, d, q$ for the model with intensity in (4) could be an issue, since their influence is not separable from the one of η ; specifically the choice of μ, κ_0 influences the ratio between the background events (i.e., Poisson-like) and the triggered ones (i.e the clustered events).

Concerning the background intensity, in our simulations we assume a constant intensity, proportional to μ , such that in our parametrization, the expected number of events in a given region is $E(N) = \mu\Delta t$, avoiding the influence of the estimation of the spatial background intensity. For the choice of $\mu, \kappa_0, c, p, d, q$, we use values close to those estimated for the Italian catalogue of the seismic events of magnitude greater than 2.5 (used in Sect. 5), assuming a constant background intensity, although this hypothesis may be unrealistic for the given area. However, different values of κ_0 are used to assess the effect of different weights of the aftershock component.

In the linear predictor η we consider two covariates: the first covariate is always the magnitude, while the second covariate is an artificial variable generated in two different ways:

- (a) a geometrical choice, in which the covariate associated to each event is the distance of the event from the main diagonal of the space region, as it were a seismic fault;
- (b) a random choice, in which the covariate is the square of a standard normal random number.

In all considered scenarios, magnitudes are obtained generating random numbers from a Gutenberg-Richter distribution with a parameter $b = 1.0789$, which is the estimated value for the used Italian catalogue. Moreover, a rectangular space region approximately equivalent to the rectangle embedding Italy is considered.

Therefore, we developed the following algorithm to simulate one catalogue from an ETAS process with conditional intensity function as in (4), which strictly depends on the branching nature of the generating process:

1. Input the *true* parameters values: $\mu, \kappa_0, c, p, d, q$, the parameters β_1 and β_2 related to covariates, and some other control parameters, that are the boundaries of the space-time region and the parameter b of the Gutenberg-Richter distribution;
2. generate a random number n_0 from a Poisson distribution with parameter $E(N) = \mu\Delta t$;
3. generate n_0 triples of space-time uniform coordinates in the spatial-temporal region together with n_0 random magnitudes and n_0 covariates, according to method (a) or (b);
4. for each point P_j generate a random number n_j from a Poisson distribution with parameter proportional to $\exp \eta_j$;
5. generate n_j triples of space-time coordinates of aftershocks in the spatial-temporal region together with n_j random magnitudes and n_j covariates, according to methods (a) or (b), reported above;
6. add the n_j new points to the set of events. Proceed with step (2) until all the events inside the region are involved in the simulation process as possible generators of further events.

Details of the method to generate random sequences from a branching process are in Zhuang and Touat (2015).

For simulating the triggered component, we consider 36 different scenarios:

- two different values of κ_0 : 0.003, 0.006;
- three different values of p : 1.05, 1.10, 1.15;
- three different values of q : 1.2, 1.3, 1.4;
- two different methods to generate the second covariate, that is (a) and (b) mentioned above.

Moreover, for the r -th simulated sample ($r = 1, \dots, N$, with N the number of the simulated samples for each scenario, fixed to 100 in this paper), two different ETAS models are fitted: the first one (Model1) is a misspecified model considering the magnitude as the only covariate, and the second one (Model2) is the right model, including both the magnitude and the further covariate, really used.

Let I_{ir} be the indicator variable assigned to each simulated event, that is 1 if the i -th event of the r -th simulated sample belongs to the Poisson generated set, 0 if it belongs to the aftershocks set. For the r -th sample, the following quantities are computed: the estimates of the parameters θ under Model1, $\hat{\theta}_{r(1)}$, and under Model2, $\hat{\theta}_{r(2)}$; the length of the simulated sample n_r , the number of events generated according to the Poisson background process n_{r0} and the number of aftershock events n_{r1} , such that $n_{r0} + n_{r1} = n_r$.

Therefore, for the r -th sample and for the model M ($M = 1, 2$), we compute the area under the ROC curve, denoted by $AUC_{r(M)}$, as a measure of the properness of $\hat{\rho}_{ir(M)}$ to classify induced and background events ($I_{ir} = 0, 1$).

Furthermore, for each event of the r -th sample and for each model $M = 1, 2$, the intensities λ_{ir} according to the true values of the parameters are computed, and compared with the intensities $\hat{\lambda}_{ir(M)}$ estimated under the model M for $M = 1, 2$, computing the following mean absolute difference:

$$\Delta_{r(M)}(\lambda) = \frac{\sum_{i=1}^{n_r} |\hat{\lambda}_{ir(M)} - \lambda_{ir}|}{n_r}$$

Eventually, for each scenario we get $N = 100$ simulations, summarized in Tables 1, 2, 3 and 4. In particular, in these tables we report the following quantities:

- average of the simulated background events, n_{r0} ,
- average of the simulated induced events, n_{r1} ,
- relative ratios between the averages values of the two models of $\Delta_{r(M)}(\lambda)$, that is:

$$RR_{\Delta(\lambda)} = \frac{\sum_{r=1}^N \Delta_{r(1)}(\lambda) - \sum_{r=1}^N \Delta_{r(2)}(\lambda)}{\sum_{r=1}^N \Delta_{r(2)}(\lambda)}$$

- relative ratios between the averages values of the $AUC_{r(M)}$:

Table 1 Average of n_0 , n_1 , RR_{AUC} , $RR_{\Delta(\lambda)}$ and RR_ϕ between the mean square errors for the parameters of models Model1.1 and Model1.2, with the other parameters true values $\mu = 0.079$, $\kappa_0 = 0.003$, $c = 0.013$, $d = 0.5$, $\beta_1 = 0.39$, $\beta_2 = -0.03$, using a geometrical covariate

p	q	n_0	n_1	RR_{AUC}	$RR_{\Delta(\lambda)}$	RR_μ	RR_{κ_0}	RR_c	RR_β	RR_d	RR_q	RR_{β_1}
1.05	1.2	236.4	71.5	0.0074	0.011	-0.032	-0.922	-0.212	-0.099	-0.121	-0.015	0.079
1.10	1.2	234.9	71.5	0.0067	0.018	0.044	-0.207	-0.097	-0.072	0.144	0.100	0.014
1.15	1.2	236.3	67.0	0.0057	0.025	-0.011	-0.937	-0.063	-0.004	-0.205	-0.130	0.084
1.05	1.3	236.6	51.9	0.0059	0.013	0.021	-0.999	-0.034	0.016	-0.540	-0.388	-0.122
1.10	1.3	235.6	50.0	0.0046	0.018	0.086	-0.754	-0.019	0.101	0.325	0.201	0.151
1.15	1.3	236.7	50.2	0.0038	0.022	0.017	-0.860	0.087	0.186	0.276	0.245	-0.044
1.05	1.4	235.4	42.2	0.0047	0.010	-0.087	16.135	-0.108	-0.059	1.020	0.851	0.171
1.10	1.4	234.6	39.4	0.0031	0.020	0.014	-0.978	-0.264	0.003	-0.377	-0.035	0.072
1.15	1.4	237.4	40.1	0.0026	0.022	0.023	-0.827	0.172	0.287	0.068	0.089	0.051

Table 2 Average of n_0 , n_1 , RR_{AUC} , $RR_{\Delta(z)}$ and RR_ϕ between the mean square errors for the parameters of models Model1.1 and Model1.2, with the other parameters true values $\mu = 0.079$, $\kappa_0 = 0.003$, $c = 0.013$, $d = 0.5$, $\beta_1 = 0.39$, using a simulated covariate

p	q	n_0	n_1	RR_{AUC}	$RR_{\Delta(z)}$	RR_μ	RR_{κ_0}	RR_ϵ	RR_ρ	RR_d	RR_q	RR_{β_1}
1.05	1.2	236.6	173.8	0.0009	0.007	0.037	0.094	- 0.007	0.148	0.116	0.117	0.022
1.10	1.2	240.8	167.9	0.0011	0.011	- 0.003	0.181	- 0.035	- 0.050	0.151	0.065	0.114
1.15	1.2	238.4	165.5	0.0008	0.013	0.024	0.088	- 0.012	0.018	0.006	- 0.005	0.007
1.05	1.3	236.3	115.6	0.0011	0.007	- 0.004	- 0.394	0.034	0.045	0.079	0.116	0.051
1.10	1.3	236.2	114.3	0.0005	0.010	- 0.023	0.003	- 0.039	0.014	0.056	0.068	0.030
1.15	1.3	236.9	115.5	0.0006	0.015	- 0.011	0.274	- 0.012	0.039	0.036	0.035	- 0.039
1.05	1.4	236.9	92.9	0.0005	0.007	- 0.000	0.091	- 0.020	- 0.032	- 0.015	- 0.025	- 0.024
1.10	1.4	234.7	88.9	0.0003	0.011	0.011	- 0.306	0.025	0.024	0.018	0.064	- 0.029
1.15	1.4	237.2	89.7	0.0003	0.015	- 0.013	- 0.527	- 0.037	0.019	0.031	0.008	0.045

Table 3 Average of n_0 , n_1 , RR_{AUC} , $RR_{\Delta(\lambda)}$ and RR_ϕ between the mean square errors for the parameters of models Model1.1 and Model1.2, with the other parameters true values $\mu = 0.079$, $\kappa_0 = 0.006$, $c = 0.013$, $d = 0.5$, $\beta_1 = 0.39$, $\beta_2 = -0.03$, using a geometrical covariate.

p	q	n_0	n_1	RR_{AUC}	$RR_{\Delta(\lambda)}$	RR_μ	RR_{κ_0}	RR_ϵ	RR_p	RR_d	RR_q	RR_{β_1}
1.05	1.2	237.3	446.1	0.0041	0.037	0.178	0.651	- 0.101	- 0.048	0.021	- 0.022	0.296
1.10	1.2	234.9	376.6	0.0036	0.051	0.576	0.628	- 0.045	0.254	0.010	- 0.042	0.340
1.15	1.2	235.4	344.4	0.0032	0.062	0.082	0.528	- 0.075	0.230	- 0.022	- 0.044	0.124
1.05	1.3	237.2	189.1	0.0050	0.042	- 0.089	0.677	0.022	0.050	0.047	0.022	0.341
1.10	1.3	240.1	178.1	0.0044	0.051	- 0.116	0.692	- 0.188	- 0.083	0.037	- 0.017	0.325
1.15	1.3	234.7	170.7	0.0030	0.065	0.151	0.527	- 0.014	0.091	- 0.007	0.029	0.129
1.05	1.4	235.6	122.7	0.0035	0.043	0.019	0.367	- 0.101	0.099	- 0.007	0.031	0.468
1.10	1.4	237.7	118.8	0.0034	0.060	0.011	0.149	- 0.130	- 0.017	- 0.169	- 0.064	0.051
1.15	1.4	235.8	116.7	0.0025	0.070	0.003	- 0.020	- 0.022	0.011	0.080	0.089	0.242

Table 4 Average of $n_0, n_1, RR_{AUC}, RR_{A(\lambda)}, RR_{\psi}$ between the mean square errors for the parameters of models Mode1.1 and Mode1.2, with the other parameters true values $\mu = 0.079, \kappa_0 = 0.006, c = 0.013, d = 0.5, \beta_1 = 0.39, \beta_2 = -0.4$, using a simulated covariate.

p	q	n_0	n_1	RR_{AUC}	$RR_{A(\lambda)}$	RR_{μ}	RR_{κ_0}	RR_c	RR_{β}	RR_d	RR_q	RR_{β_1}
1.05	1.2	238.6	1251.8	0.0010	0.020	- 0.049	0.315	0.090	0.223	0.069	0.073	0.107
1.10	1.2	235.6	1129.3	0.0009	0.024	- 0.036	0.340	- 0.005	0.022	0.052	0.019	0.104
1.15	1.2	241.4	1135.1	0.0006	0.032	- 0.036	0.306	- 0.052	- 0.031	0.102	0.083	0.084
1.05	1.3	234.4	482.4	0.0008	0.020	- 0.023	0.338	- 0.009	0.015	0.005	0.002	0.039
1.10	1.3	234.6	461.4	0.0006	0.027	0.001	0.292	- 0.029	0.009	0.031	- 0.000	- 0.018
1.15	1.3	236.9	448.5	0.0006	0.033	- 0.015	0.320	0.013	0.002	- 0.022	- 0.022	0.135
1.05	1.4	238.1	307.4	0.0005	0.021	- 0.002	0.300	0.006	0.021	0.003	- 0.036	0.137
1.10	1.4	238.0	286.6	0.0005	0.023	- 0.016	0.325	- 0.018	- 0.015	- 0.001	- 0.003	0.028
1.15	1.4	236.1	287.0	0.0004	0.033	- 0.004	0.355	- 0.022	- 0.009	0.016	- 0.004	- 0.022

$$RR_{AUC} = \frac{\sum_{r=1}^N AUC_{r(2)} - \sum_{r=1}^N AUC_{r(1)}}{\sum_{r=1}^N AUC_{r(1)}}$$

- relative ratios between the simulated mean square errors of the two models for each parameter estimate $\hat{\phi}$, where ϕ is any of the parameter $\mu, \kappa_0, c, p, d, q, \beta_1$

$$RR_{\hat{\phi}} = \frac{\sum_{r=1}^N (\hat{\phi}_{r(1)} - \phi)^2 - \sum_{r=1}^N (\hat{\phi}_{r(2)} - \phi)^2}{\sum_{r=1}^N (\hat{\phi}_{r(2)} - \phi)^2}$$

For more details, the tables reporting the simulated mean square errors (MSE) of the parameter estimates, estimated under the wrong model (Model1) and the right model (Model2) are reported in section “[Appendix](#)” (see Tables 8, 9, 10, 11, 12, 13, 14, 15).

The simulations results using a geometrical covariate, as described in the point (a) above, for $\kappa_0 = 0.003$, are reported in Table 1. The corresponding results using a simulated covariate, as described in the point (b) above, are reported in Table 2.

The tables reporting the same comparisons for a greater value of κ_0 are reported in Table 3, using the geometric covariate, and in Table 4, for the case including a simulated covariate.

It could be noticed that all the AUC values used for the RR_{AUC} values (see Tables 1, 2, 3 and 4) are however in the range 0.8881–0.9999.

The results reported in the Tables 1 and 2 show that, in general, the estimation of the parameters can still be performed in the absence of data on covariates, since it does not greatly affect the conditional intensity of the process, confirming results reported in Schoenberg (2016) for different formulations of self-exciting models including covariates. Conversely, the misspecification has a peculiar effect on the parameter κ_0 behaviour, comparing to the other parameters of the model. In particular, for a lower value of $\kappa_0 = 0.003$, the averages of the MSE estimating the Model1 are generally lower than the corresponding ones estimating the Model2 (see the negative sign of the RR reported in the Table 1). This effect is less evident if we use a random covariate in the simulation procedure (see the Table 2). Moreover, as described above, similar simulations are carried out accounting for a higher value of $\kappa_0 = 0.006$, that is with a greater influence of the aftershock component (Tables 3 and 4).

Here we observe that the κ_0 estimation provides higher MSE when the wrong model Model1 is estimated than estimating the true model Model2. We also note that this effect, though of minor size, seems to be opposite for the parameter μ . In particular, the smaller the true value κ_0 is (that is, the *more* Poissonian the generating process is), the more variable its estimator is (see the Tables 8, 9, 10, 11, 12, 13, 14, 15 in the [Appendix](#)). In general, the misspecification model effect on the estimation of κ_0 can be explained as the consequence of the different number of simulated events (n_0) and induced events (n_1), depending on the true value of κ_0 . Indeed, for $\kappa_0 = 0.003$ the simulated catalogues are mostly composed by

background events, and therefore the estimation of the κ_0 parameter appears to be more unstable. Otherwise, increasing κ_0 the simulated catalogues have more aftershock events, providing more information for the triggered component estimation and more evidence in favour of `Model2`.

Actually, this last result can be also affected by our assumption related to the use of a constant background density in space and therefore, could require deeper analyses.

5 Application to the Italian earthquakes and comments

In this section, we report some of the results of the proposed ETAS-FLP approach with covariates, starting from the Italian catalogue of the space-time Italian seismicity, from May 5th, 2012 to May 7th, 2016, with 2.5 as the threshold magnitude (i.e. the lower bound for which earthquakes with higher values of magnitude are surely recorded in the catalogue). The catalogue reports 2226 events with the usual hypocentral coordinates (longitude, latitude, depth, time) together with the events magnitude, and also some additional information, such as: the hypocentral uncertainty, the distance from the nearest station (for shallow earthquakes, this distance should be sufficiently small), a measure of the quality of the location (named *rms*), the number of stations that recorded the event (this number is heavily influenced by the magnitude of the event and strongly influences the accuracy of the location) and the distance from the nearest fault (i.e., the identified earthquakes sources in that area).

The ETAS-FLP estimating results of `Model1`, as in Eq. (2), accounting both for the epicentral coordinates (longitude, latitude and time) and the magnitude of the inducing event, are not completely satisfying, suggesting, as usual, some lack of fitting mostly due to the triggered component (see Table 5). However, adding also the available covariates (`Model2`, reported in Fig. 1), that is considering the model in Eq. (4) with covariates, starting from the complete model, including all the available covariates, the best selected one includes, together with the magnitude m , also the depth z , the number of stations n^{st} , the distance from the nearest station δ^{st} and the distance from the nearest fault δ , such that we have a model with $k = 5$ covariates and a total of 11 parameters. In particular, the last four covariates have both a negative effect on the space-time reproducing activity (see Table 6).

Table 5 Estimates, and standard errors within brackets, of the parameters of models `Model1`. In Eq. (2) α corresponds to β_1 of the linear predictor in Eq. (4)

	μ	κ_0	α	c	p	d	q
θ	0.67	0.02	0.74	0.01	1.11	1.91	1.95
std.err	(0.0226)	(0.0058)	(0.0926)	(0.0027)	(0.0157)	(0.2604)	(0.0776)

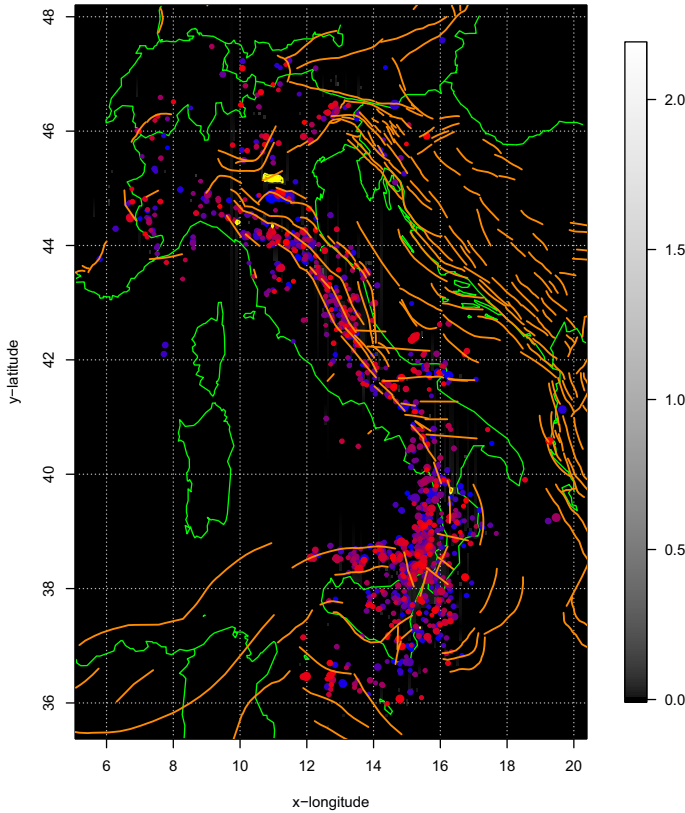


Fig. 1 Output of the estimated ETAS-FLP model with covariates: estimated total intensity together with the observed points (old in blue and recent in red) and available line-faults

Table 6 Estimates, and standard errors within brackets, of the parameters of models *Model2*

	μ	κ_0	c	p	d	q
$\hat{\theta}$	0.71	0.07	0.02	1.15	1.94	2.00
std.err.	(0.0227)	(0.0212)	(0.0034)	(0.0169)	(0.2724)	(0.0848)
$exp(\eta_j)$	$1.16m_j - 0.04z_j - 0.01n_j^{st} - 0.01\delta_j^{st} - 1.83\delta_j$					
std.err.	(0.0854)	(0.0058)	(0.0028)	(0.0018)	(0.2986)	

Looking at Tables 5 and 6, as already observed from our simulation study, the estimation of the parameters μ , c , p , d , and q is not strongly affected by the introduction of the covariates; conversely, the estimation of κ_0 is more sensitive to the choice of the model, since it is strictly related to the linear predictor value, that affects in turn the triggered component. Indeed, coherently with results observed in

Sect. 4, the relative standard error of $\hat{\kappa}_0$ (w.r.t. its estimated value) is much higher than the relative standard errors of the other estimators.

For the temporal domain diagnostics of the fitted model, the marginal time process is obtained by integrating the estimated intensity function with respect to the observed space region (see Adelfio and Chiodi 2015b). As in Adelfio and Chiodi (2015b), let $\{t_i\}$ be data of a temporal point process identified by the conditional intensity function $\lambda(t)$. Therefore, the residual temporal process can be obtained moving each point $(t_i) \forall i$ on the real half-line to the point:

$$\tau_i = \Lambda(t_i) = \int_0^{t_i} \lambda(t) dt \quad (5)$$

(Meyer 1971; Cox and Isham 1980; Ogata 1988). The process $\{\tau_i\}$ in (5) is a homogeneous Poisson process with unit intensity. Then, a plot of τ_i versus i can inform about bad fitting in time. In particular, this plot, together with a plot of the estimated time intensities, informs on the time at which departures from model assumptions are more evident.

Now, let N be a spatial process \mathbf{S} , with $\mathbf{s} = \{s_1, \dots, s_n\}$ the vector of its observed locations in a bounded region W of the plane \mathbb{R}^2 . For a spatial point process, a useful definition of residuals has been introduced by Baddeley et al. (2005), based on the innovation process, extended by the Georgii–Nguyen–Zessin identity:

$$E \left[\sum_{i=1}^n b(u_i, \mathbf{s}) \right] = E \int_W b(u, \mathbf{s}) \lambda(u, \mathbf{s}) du \quad (6)$$

where $b(u, \mathbf{s})$ is a non-negative integrable function. For an exploratory data analysis, Pearson residuals are obtained considering in (6) the integrand function defined by:

$$\hat{b}(u, \mathbf{x}) = 1 / \sqrt{\hat{\lambda}(u, \mathbf{x})} \quad (\text{Baddeley et al. 2005}).$$

In this paper, as described in Adelfio and Chiodi (2015b), due to their interpretative simplicity, the above defined Pearson spatial residuals are computed in a space-time region Q , which extends over the whole time interval and on a rectangular space area. These tools have been also implemented in the version 2.0 of the package *etasFLP*. This area is determined by an equally spaced grid of q intervals for each dimension, such that the observed space area is divided in $q \times q$ rectangular cells. According to this approach, the classical comparison between observed and theoretical frequencies is considered and a χ^2 -measure, based on the computed Pearson residuals, can be easily computed.

As further investigation, we also perform a spatio-temporal diagnostics based on the second-order statistics coming from weighted measures (Adelfio and Schoenberg 2009; Adelfio et al. 2020). The used method assesses goodness-of-fit of spatio-temporal models by using weighted second-order statistics, computed after weighting the contribution of each observed point by the inverse of the conditional intensity function that identifies the process. Weighted second-order statistics directly apply to data without assuming homogeneity nor transforming the data into residuals, eliminating thus the variability due to transforming procedures, such as a random thinning procedure (Baddeley et al. 2006). In particular, in this paper we

use the weighted K -function, obtained by weighting the process by the inverse of the conditional intensity function of the estimated model. Indeed, if departures from a such a behaviour were observed, then the data would be supposed to come from a model identified by a conditional intensity function different from the one used in the weighting procedure.

For other relevant diagnostics methods for self-exciting point process models see also Schoenberg (2003), Ogata and Tanemura (2003), Bray et al. (2014).

In our application, diagnostic results suggest a more satisfying fitting for Model2 than Model1, as shown in the Figs. 2 (see the better time residuals using Model2), 3 and 4, and summarized in Table 7. This result is more evident from the total intensity diagnostic plots than the background ones, suggesting that the real fit improvement is due to the introduction of covariates in the induced component of the model, since the observed data are characterised by the presence of big sequences, affecting the estimation procedure.

Therefore, the omission of influential external variables may be relevant on the characterisation of the conditional rate of seismicity, mostly when the analysed area is characterised by severe seismicity, with characteristics of evident dependence both in space and time, like Italy.

Moreover, Fig. 5 reports the difference between the observed values of the global weighted spatio-temporal K -function and the expected ones, weighting by the MLE intensity under Model1 and Model2, respectively. As a χ^2 -type statistic, we evaluate this distance computed over a fixed grid that covers the region of interest, and obtain $\chi^2 = 120.58$ (under the Model1), $\chi^2 = 11.35$ (under the Model2). It is evident that this difference is significantly larger in the first case corresponding to

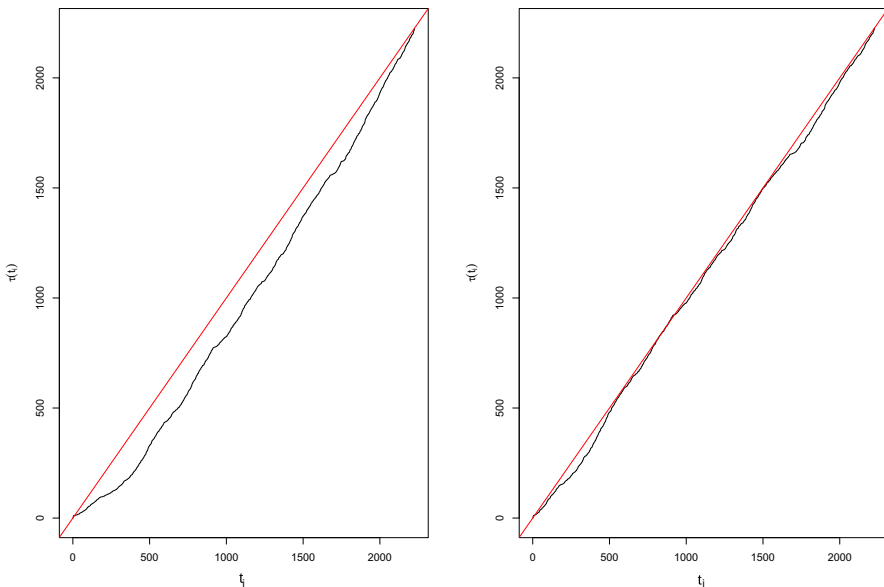


Fig. 2 Temporal residuals for Model1 (on the left); Temporal residuals for Model2 (on the right)

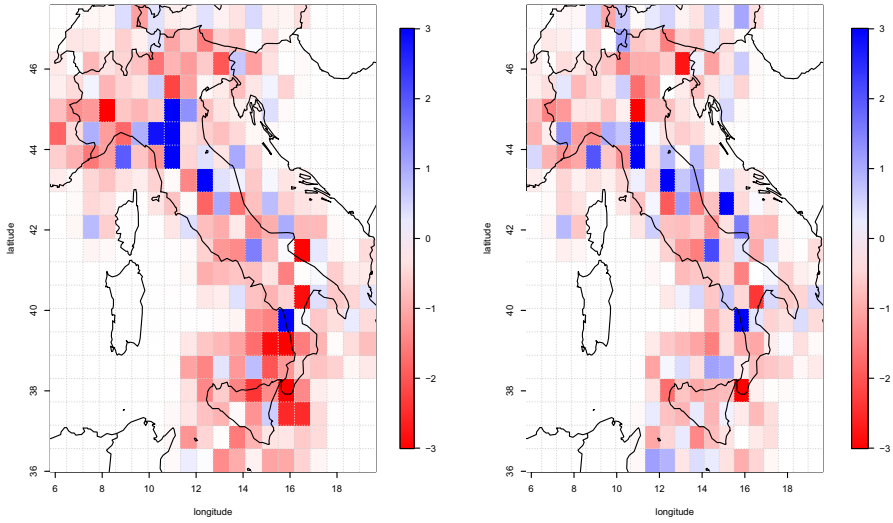


Fig. 3 Spatial residuals for the total intensity Model1 (on the left); Spatial residuals for the total intensity Model2 (on the right)

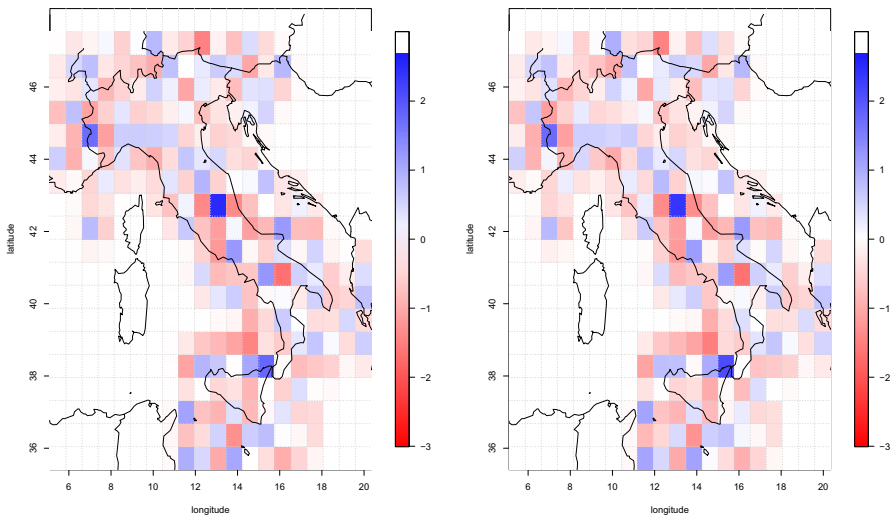


Fig. 4 Spatial residuals for the background intensity Model1 (on the left); Spatial residuals for the background intensity Model2 (on the right)

Table 7 AIC and χ^2 -statistic (20 × 20 grid), for the two estimated models

	Model1	Model2
AIC	43249.19	42706.16
χ^2	498.25	353.44

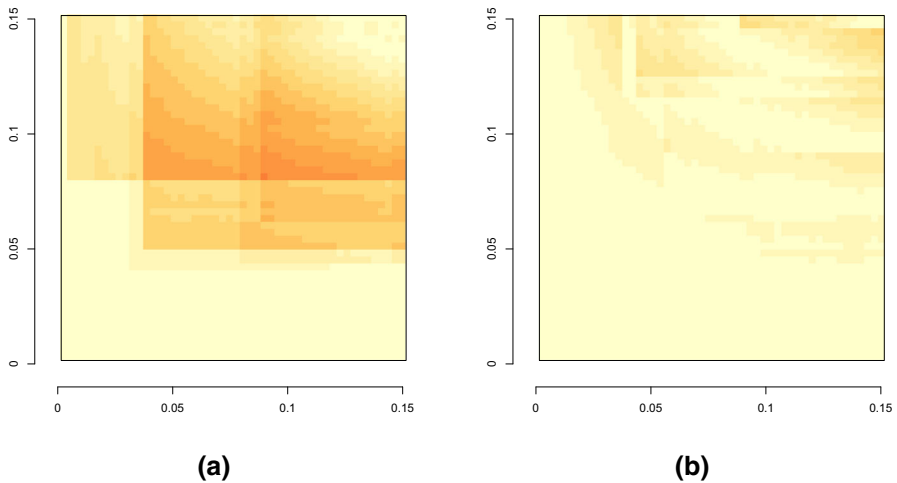


Fig. 5 Difference between the global weighted spatio-temporal K -function and the expected ones, weighting by the MLE intensity of the `Model1` (left) and of the `Model2` (right), respectively.

the model without covariates, providing further evidence about the necessity of a more complex model for the description of the observed data.

6 Conclusive remarks

In this paper, we have proposed a novel contagion risk measurement model, aimed at improving earthquake risk measurement taking both space and time dependence into account, by means of the ETAS space-time point process, estimated by the ETAS-FLP approach.

The simulation study confirms the importance of well specifying the model, properly accounting for further variables, mostly in case of catalogues with several sequences and many events and, in particular, for parameters describing the size of the clustered component. Indeed, the misspecification effect is more evident for the estimation of the parameters κ_0 , and slightly for μ , that characterize, respectively, the induced and the background components of the conditional intensity function of the ETAS model. When the catalogue is compounded by several sequences, then the estimation of κ_0 is more sensitive to the misspecification of the model. Comments invert in case of a lower number of induced events. However, we have to stress that in our simulation study we assume just a constant background, and therefore the real model misspecification effect on μ should require deeper reflections.

The importance of the proposed method is also confirmed by the reported application, related to the Italian area. Indeed, Italy is characterized by a strong clustered seismicity (Adelfio and Chiodi 2009; Giunta et al. 2009), for which further information related to the spatial-temporal characterization can be relevant for a better description of the seismic activity. Indeed, in the seismic context, the proposed approach would provide a more general formalism for the earthquake

occurrence in space and time, since the main idea is that the effect on the future activity does not depend only on the closeness of the previous events, but also on specific characteristics of the main event, like magnitude, as usual, and also further information, such as geological features.

The reported results, though partial and provisional, confirm our intuition reported in previous studies (e.g., Adelfio and Chiodi 2015b). Indeed, the need of a more flexible model for the space-time triggered component of the ETAS model is often revealed, although the background seismicity is well described by the FLP estimated intensity. In our opinion, considering external information (such as geological information related to faults distribution) for the description of spatio-temporal earthquakes is an innovative and promising perspective of study, even relevant in different fields of research.

Future research developments include evaluating the robustness of the results with respect to different spatial regions (that is to extend the proposed study also to areas with different characteristics) and to the inclusion of other specific covariates, to study the dependence of observed intensity in space and time from other external factors (e.g., GPS information, external marks, etc).

Acknowledgements Open access funding provided by Università degli Studi di Palermo within the CRUI-CARE Agreement. This paper has been partially supported by the national grant of the Italian Ministry of Education University and Research (MIUR) for the PRIN-2015 program, ‘Complex space-time modelling and functional analysis for probabilistic forecast of seismic events’.

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/>.

Appendix

See the Tables 8, 9, 10, 11, 12, 13, 14 and 15.

Table 8 Simulated MSE of parameters estimates on the 100 simulated catalogues under the wrong Model11 and the parameters true values $\mu = 0.079, \kappa_0 = 0.003, c = 0.013, d = 0.5, \beta_1 = 0.39$, using a geometrical covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$
1.05	1.2	0.000032	0.000009	0.000636	0.0053	0.5272	0.0272	0.1181
1.10	1.2	0.000026	0.000056	0.000314	0.0095	0.3260	0.0192	0.1088
1.15	1.2	0.000031	0.000024	0.000178	0.0095	0.3424	0.0227	0.1369
1.05	1.3	0.000036	0.000007	0.000467	0.0081	0.3378	0.0352	0.1649
1.10	1.3	0.000024	0.000027	0.000240	0.0079	0.3097	0.0444	0.1974
1.15	1.3	0.000038	0.000017	0.000308	0.0102	0.3401	0.0408	0.1715
1.05	1.4	0.000022	0.010518	0.000231	0.0075	0.9461	0.2379	0.3331
1.10	1.4	0.000029	0.000970	0.000589	0.0132	0.7156	0.1781	0.1280
1.15	1.4	0.000028	0.000754	0.000616	0.0283	0.6075	0.1887	0.2127

Table 9 Simulated MSE of parameters estimates on the 100 simulated catalogues under the true Model2 and the parameters true values $\mu = 0.079, \kappa_0 = 0.003, c = 0.013, d = 0.5, \beta_1 = 0.39, \beta_2 = -0.03$, using a geometrical covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$	$\hat{\beta}_2$
1.05	1.2	0.000033	0.000112	0.000807	0.0058	0.5994	0.0276	0.1095	0.0001
1.10	1.2	0.000025	0.000070	0.000348	0.0103	0.2849	0.0174	0.1072	0.0001
1.15	1.2	0.000031	0.000384	0.000190	0.0095	0.4308	0.0261	0.1262	0.0001
1.05	1.3	0.000035	0.008829	0.000483	0.0080	0.7350	0.0576	0.1879	0.0001
1.10	1.3	0.000022	0.000110	0.000245	0.0071	0.2337	0.0370	0.1715	0.0001
1.15	1.3	0.000037	0.000119	0.000283	0.0086	0.2665	0.0327	0.1794	0.0001
1.05	1.4	0.000024	0.000614	0.000259	0.0079	0.4684	0.1285	0.2846	0.0001
1.10	1.4	0.000029	0.043971	0.000800	0.0132	1.1492	0.1847	0.1193	0.0001
1.15	1.4	0.000028	0.004354	0.000525	0.0219	0.5686	0.1733	0.2024	0.0001

Table 10 Simulated MSE of parameters estimates on the 100 simulated catalogues under the wrong Model11 and the parameters true values $\mu = 0.079, \kappa_0 = 0.003, c = 0.013, d = 0.5, \beta_1 = 0.39$, using a simulated covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$
1.05	1.2	0.000036	0.000005	0.000065	0.0016	0.0492	0.0052	0.0524
1.10	1.2	0.000044	0.000004	0.000064	0.0025	0.0708	0.0039	0.0545
1.15	1.2	0.000047	0.000004	0.000041	0.0030	0.0934	0.0043	0.0507
1.05	1.3	0.000039	0.000008	0.000214	0.0024	0.1843	0.0290	0.0665
1.10	1.3	0.000030	0.000006	0.000081	0.0037	0.1039	0.0143	0.0559
1.15	1.3	0.000036	0.000004	0.000065	0.0035	0.0686	0.0101	0.0452
1.05	1.4	0.000036	0.000005	0.000122	0.0028	0.1637	0.0211	0.0544
1.10	1.4	0.000030	0.000010	0.000271	0.0070	0.1374	0.0315	0.0909
1.15	1.4	0.000031	0.000008	0.000079	0.0043	0.1081	0.0185	0.0630

Table 11 Simulated MSE of parameters estimates on the 100 simulated catalogues under the true Model12 and the parameters true values $\mu = 0.079, \kappa_0 = 0.003, c = 0.013, d = 0.5, \beta_1 = 0.39, \beta_2 = -0.4$, using a simulated covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$	$\hat{\beta}_2$
1.05	1.2	0.000035	0.000004	0.000065	0.0014	0.0441	0.0046	0.0513	0.0191
1.10	1.2	0.000044	0.000004	0.000066	0.0027	0.0615	0.0036	0.0489	0.0194
1.15	1.2	0.000046	0.000004	0.000042	0.0030	0.0928	0.0043	0.0504	0.0224
1.05	1.3	0.000039	0.000014	0.000207	0.0023	0.1708	0.0260	0.0633	0.0274
1.10	1.3	0.000031	0.000006	0.000084	0.0036	0.0984	0.0134	0.0543	0.0315
1.15	1.3	0.000036	0.000003	0.000066	0.0033	0.0663	0.0097	0.0470	0.0194
1.05	1.4	0.000036	0.000004	0.000125	0.0029	0.1662	0.0216	0.0557	0.0332
1.10	1.4	0.000030	0.000015	0.000264	0.0069	0.1350	0.0296	0.0936	0.0276
1.15	1.4	0.000031	0.000017	0.000082	0.0042	0.1049	0.0184	0.0602	0.0236

Table 12 Simulated MSE of parameters estimates on the 100 simulated catalogues under the wrong Model11 and the other parameters true values $\mu = 0.079, \kappa_0 = 0.006, c = 0.013, d = 0.5, \beta_1 = 0.39$, using a geometrical covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$
1.05	1.2	0.000058	0.000020	0.000029	0.0007	0.0370	0.0022	0.0236
1.10	1.2	0.000048	0.000021	0.000016	0.0009	0.0293	0.0018	0.0267
1.15	1.2	0.000036	0.000022	0.000015	0.0013	0.0270	0.0019	0.0281
1.05	1.3	0.000027	0.000022	0.000078	0.0014	0.0499	0.0062	0.0361
1.10	1.3	0.000031	0.000021	0.000082	0.0021	0.0513	0.0048	0.0596
1.15	1.3	0.000038	0.000021	0.000048	0.0028	0.0423	0.0047	0.0438
1.05	1.4	0.000035	0.000023	0.000091	0.0022	0.0965	0.0185	0.0606
1.10	1.4	0.000031	0.000023	0.000098	0.0026	0.0873	0.0174	0.0758
1.15	1.4	0.000030	0.000020	0.000067	0.0027	0.0831	0.0209	0.0903

Table 13 Simulated MSE of parameters estimates on the 100 simulated catalogues under the true Model12 and the other parameters true values $\mu = 0.079, \kappa_0 = 0.006, c = 0.013, d = 0.5, \beta_1 = 0.39, \beta_2 = -0.03$, using a geometrical covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$	$\hat{\beta}_2$
1.05	1.2	0.000049	0.000012	0.000033	0.0007	0.0362	0.0023	0.0182	0.0000
1.10	1.2	0.000031	0.000013	0.000016	0.0007	0.0291	0.0019	0.0199	0.0000
1.15	1.2	0.000033	0.000014	0.000016	0.0010	0.0276	0.0019	0.0250	0.0000
1.05	1.3	0.000030	0.000013	0.000076	0.0013	0.0476	0.0060	0.0269	0.0000
1.10	1.3	0.000035	0.000012	0.000101	0.0023	0.0495	0.0049	0.0450	0.0000
1.15	1.3	0.000033	0.000014	0.000048	0.0025	0.0425	0.0046	0.0388	0.0000
1.05	1.4	0.000035	0.000017	0.000101	0.0020	0.0972	0.0179	0.0413	0.0000
1.10	1.4	0.000031	0.000020	0.000113	0.0026	0.1051	0.0186	0.0722	0.0000
1.15	1.4	0.000030	0.000020	0.000069	0.0027	0.0770	0.0192	0.0727	0.0000

Table 14 Simulated MSE of parameters estimates on the 100 simulated catalogues under the wrong Model11 and the other parameters true values $\mu = 0.079, \kappa_0 = 0.006, c = 0.013, d = 0.5, \beta_1 = 0.39$, using a simulated covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$
1.05	1.2	0.000103	0.000018	0.000008	0.0002	0.0062	0.0007	0.0057
1.10	1.2	0.000063	0.000017	0.000009	0.0003	0.0091	0.0006	0.0065
1.15	1.2	0.000071	0.000018	0.000006	0.0004	0.0076	0.0005	0.0071
1.05	1.3	0.000037	0.000017	0.000016	0.0006	0.0124	0.0018	0.0101
1.10	1.3	0.000036	0.000018	0.000015	0.0007	0.0130	0.0018	0.0199
1.15	1.3	0.000046	0.000016	0.000011	0.0008	0.0197	0.0020	0.0161
1.05	1.4	0.000042	0.000017	0.000035	0.0008	0.0210	0.0051	0.0250
1.10	1.4	0.000034	0.000017	0.000022	0.0010	0.0175	0.0054	0.0195
1.15	1.4	0.000032	0.000017	0.000025	0.0013	0.0161	0.0038	0.0230

Table 15 Simulated MSE of parameters estimates on the 100 simulated catalogues under the true Model12 and the other parameters true values $\mu = 0.079, \kappa_0 = 0.006, c = 0.013, d = 0.5, \beta_1 = 0.39, \beta_2 = -0.4$, using a simulated covariate

p	q	$\hat{\mu}$	$\hat{\kappa}_0$	\hat{c}	\hat{p}	\hat{d}	\hat{q}	$\hat{\beta}_1$	$\hat{\beta}_2$
1.05	1.2	0.000108	0.000014	0.000007	0.0002	0.0058	0.0006	0.0052	0.0031
1.10	1.2	0.000066	0.000013	0.000009	0.0003	0.0086	0.0006	0.0059	0.0029
1.15	1.2	0.000074	0.000014	0.000006	0.0004	0.0069	0.0005	0.0065	0.0022
1.05	1.3	0.000038	0.000013	0.000016	0.0005	0.0123	0.0018	0.0098	0.0055
1.10	1.3	0.000036	0.000014	0.000015	0.0006	0.0126	0.0018	0.0203	0.0062
1.15	1.3	0.000047	0.000012	0.000011	0.0008	0.0201	0.0021	0.0142	0.0063
1.05	1.4	0.000042	0.000013	0.000035	0.0008	0.0209	0.0053	0.0220	0.0105
1.10	1.4	0.000035	0.000013	0.000022	0.0010	0.0176	0.0054	0.0190	0.0081
1.15	1.4	0.000032	0.000012	0.000026	0.0014	0.0159	0.0038	0.0236	0.0098

References

Adelfio G, Chiodi M (2009) Second-order diagnostics for space-time point processes with application to seismic events. *Environmetrics* 20:895–911

Adelfio G, Schoenberg FP (2009) Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Ann Inst Stat Math* 61(4):929–948

Adelfio G, Ogata Y (2010) Hybrid kernel estimates of space-time earthquake occurrence rates using the ETAS model. *Ann Inst Stat Math* 62(1):127–143

Adelfio G, Chiodi M (2015a) Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stoch Env Res Risk Assess* 29:443–450

Adelfio G, Chiodi M (2015b) FLP estimation of semi-parametric models for space-time point processes and diagnostic tools. *Spat Stat* 14:119–132

Adelfio G, Chiodi M, Giudici P (2018) Credit risk contagion through space-time point processes (submitted)

Adelfio G, Siino M, Mateu J, Rodríguez-Cortés FJ (2020) Some properties of local weighted second-order statistics for spatio-temporal point processes. *Stoch Env Res Risk Assess* 34:149–168

- Baddeley A, Turner R, Møller J, Hazelton M (2005) Residual analysis for spatial point processes. *J R Stat Soc B* 67(5):617–666
- Baddeley A, Gregori P, Mateu J, Stoica R, Stoyan D (2006) Case studies in spatial point process modeling. Springer, New York
- Balderama E, Schoenberg F, Murray E, Rundel P (2012) Application of branching point process models to the study of invasive red banana plants in Costa Rica. *J Am Stat Assoc* 107:467–476
- Becker N (1977) Estimation for discrete time branching processes with application to epidemics. *Biometrics* 33:515–522
- Bray A, Wong K, Barr CD, Schoenberg FP (2014) Voronoi residual analysis of spatial point process models with applications to California earthquake forecasts. *Ann Appl Stat* 8(4):2247–2267
- Caron-Lormier G, Masson J, Menard N, Pierre J (2006) A branching process, its application in biology: influence of demographic parameters on the social structure in mammal groups. *J Theor Biol* 238:564–574
- Chiodi M, Adelfio G (2011) Forward likelihood-based predictive approach for space-time processes. *Environmetrics* 22:749–757
- Chiodi M, Adelfio G (2017b) Mixed non-parametric and parametric estimation techniques in R package *etasFLP* for earthquakes' description. *J Stat Softw* 76(3):1–28
- Chiodi M, Adelfio G (2017a) *etasFLP*: estimation of an ETAS model. Mixed FLP (forward likelihood predictive) and ML estimation of non-parametric and parametric components of the ETAS model for earthquake description. R package version 1.4.1
- Chiodi M, Adelfio G (2020) *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 2.0
- Cox DR, Isham V (1980) Point processes. Chapman and Hall, London
- Daley DJ, Vere-Jones D (2003) An introduction to the theory of point processes, 2nd edn. Springer, New York
- Diggle PJ (2013) Statistical analysis of spatial and spatio-temporal point patterns. CRC Press, Boca Raton
- Giunta G, Luzio D, Agosta F, Calò M, Di Trapani F, Giorgianni A, Oliveri E, Orioli S, Perniciaro M, Vitale M, Chiodi M, Adelfio G (2009) An integrated approach to the relationships between tectonics and seismicity in northern Sicily and southern Tyrrhenian. *Tectonophysics* 476:13–21
- Hawkes A, Adamopoulos L (1973) Cluster models for earthquakes regional comparison. *Bull Int Stat Inst* 45(3):454–461
- Höhle M (2009) Additive-multiplicative regression models for spatio-temporal epidemics. *Biometric J* 51:961–978
- Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns, vol 70. Wiley, New York
- Johnson RA, Taylor JR (2008) Preservation of some life length classes for age distributions associated with age-dependent branching processes. *Stat Probab Lett* 78:2981–2987
- Lin D, Ying Z (1995) Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *Ann Stat* 23:1712–1734
- Martinussen T, Scheike T (2002) A flexible additive multiplicative hazard model. *Biometrika* 89:283–298
- Meyer P (1971) Démonstration simplifiée d'un théorème de knight. *Lect Not Math* 191:191–195
- Meyer S, Elias J, Höhle M (2012) A space-time conditional intensity model for invasive meningococcal disease occurrence. *Biometrics* 68(2):607–616
- Meyer S, Held L, Höhle M (2017) Spatio-temporal analysis of epidemic phenomena using the R package *surveillance*. *J Stat Softw* 77(11):605
- Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE (2011) Self-exciting point process modeling of crime. *J Am Stat Assoc* 106(493):100–108
- Ogata Y (1978) The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann Inst Stat Math* 30(2):243–261
- Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. *J Am Stat Assoc* 83(401):9–27
- Ogata Y (1998) Space-time point-process models for earthquake occurrences. *Ann Inst Stat Math* 50(2):379–402
- Ogata Y, Tanemura M (2003) Modelling heterogeneous space-time occurrences of earthquakes and its residual analysis. *Appl Stat* 52(4):499–509
- Paul M, Held L (2011) Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Stat Med* 30(10):1118–1136

- Paul M, Held L, Toschke A (2008) Multivariate modelling of infectious disease surveillance data. *Stat Med* 27(29):6250–6267
- Rathbun S (1996) Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes. *J Stat Plan Inference* 51(1):55–74
- Reinhart A, Greenhouse J (2018) Self-exciting point processes with spatial covariates: modelling the dynamics of crime. *J R Stat Soc Ser C (Appl Stat)* 67(5):1305–1329
- Sasieni PD (1996) Proportional excess hazards. *Biometrika* 83:127–141
- Schoenberg FP (2003) Multi-dimensional residual analysis of point process models for earthquake occurrences. *J Am Stat Assoc* 98(464):789–795
- Schoenberg FP (2016) A note on the consistent estimation of spatial-temporal point process parameters. *Stat Sin* 26:861–879
- Zhuang J, Ogata Y, Vere-Jones D (2002) Stochastic declustering of space-time earthquake occurrences. *J Am Stat Assoc* 97(458):369–379
- Zhuang J, Touat S (2015) Stochastic simulation of earthquake catalogs. Community online resource for statistical seismicity analysis (CORSSA) Version: 1.0

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Giada Adelfio^{1,2}  · Marcello Chiodi^{1,2}

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