

Some insights into the time clustering of large earthquakes in Italy

Licia Faenza ⁽¹⁾, Warner Marzocchi ⁽¹⁾, Anna Maria Lombardi ⁽²⁾ and Rodolfo Console ⁽²⁾

⁽¹⁾ Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

⁽²⁾ Istituto Nazionale di Geofisica e Vulcanologia, Roma, Italy

Abstract

The aim of this work is to investigate the clustering properties of the large earthquakes which occurred in Italy in the last four centuries. In particular, we compare the results of a new multivariate nonparametric model applied to the catalog of large earthquakes in Italy, and to a synthetic catalog generated through a specific ETAS model, successfully applied to describe aftershock sequences. The results disclose a longer clustering time for real large earthquakes, suggesting that the physical process that governs aftershock sequences and the occurrence of large earthquakes may be different. Alternatively, the results can be explained by suggesting that the ETAS model, used to describe aftershock sequences, is not a suitable tool to model seismicity as a whole.

Key words earthquakes clustering – ETAS model – large earthquakes in Italy

1. Introduction

The spatio-temporal distribution of large earthquakes is a fundamental ingredient for seismic hazard assessment. Remarkably, despite the importance of the issue and many efforts made in the past, so far shared conclusions could not be reached and different statistical distributions are still used for large earthquake forecasting. The most striking recent example is the report concerning the seismic hazard assessment for the Northern California (Working Group on California Earthquake Probability, 1999), where quite different (and opposite) models were used for the calculations (*e.g.* Poisson, Brownian Passage Time, Time Predictable).

In a recent paper, Faenza *et al.* (2003) proposed a nonparametric and multivariate method (the Proportional Hazard Model; hereinafter PHM) to estimate the spatio-temporal distribution of large earthquakes, that drastically reduces the *a priori* assumptions on the temporal behavior. This method was applied to the Italian seismic catalog of the last four centuries. The results indicate the presence of a time clustering of large earthquakes. In spite of a first order similarity with the time behavior of the ETAS model (*e.g.*, Ogata, 1988) used to successfully model aftershock sequences (*e.g.*, Console *et al.*, 2003 for Italian seismicity), it is remarkable to note that the length of the time clustering for large earthquakes (few years) seems to be larger than the clustering time of aftershocks sequence. This difference, if confirmed, could have important theoretical implications. In particular, a difference in the clustering properties may suggest the existence of different physical mechanisms for aftershock and large earthquake occurrences, and/or that the specific ETAS model used to describe the aftershock sequences is not suitable to model some major features of seismicity.

Mailing address: Dr. Licia Faenza, Istituto Nazionale di Geofisica e Vulcanologia, Via Donato Creti 12, 40128 Bologna, Italy; e-mail: faenza@bo.ingv.it

Here we compare the clustering properties of large earthquakes found by Faenza *et al.* (2003), with those of the ETAS model used to describe seismicity of small magnitudes (Console *et al.*, 2003). In particular, we apply the PHM used for large earthquakes in Faenza *et al.* (2003) to a synthetic catalog of large earthquakes generated using the ETAS model. ETAS model parameters have been estimated from the spatio-temporal distribution of the events with $M \geq 2.0$ in Italy in the time interval 1987-2000 (Console *et al.*, 2003). This allows us to prove if the hazard function coming from the synthetic catalog is different from the one observed for real large earthquakes. In other words, we check the validity of the specific ETAS model used for aftershock sequences also to describe the spatio-temporal distribution of large events.

2. Proportional Hazard Model (PHM)

The model used here (PHM) is based on the study of the spatio-temporal occurrence of earthquakes through a non-parametric multidimensional fit of the hazard function. This method (Kalbfleisch and Prentice, 1980; Faenza *et al.*, 2003) presents several advantages compared to other more traditional approaches. In particular, it may account for tectonics/ physics parameters that can potentially influence spatio-temporal variability, and it tests their relative importance. Another important aspect is that this model is nonparametric; this means that no *a priori* distribution is assumed for the interevent time.

For a generic time x^* since the most recent event, the hazard function of this model can be written as

$$\lambda(x^*, \mathbf{z}) = \lambda_0(x^*) \exp(\mathbf{z}\boldsymbol{\beta}) \quad (2.1)$$

where \mathbf{z} is a vector of covariate, $\boldsymbol{\beta}$ is a vector of coefficient and $\lambda_0(x^*)$ is an arbitrary and unspecified base-line hazard function. The Random Variables (from now on RVs) of the system are the InterEvent Time (from now on IET) and the Censoring Time (from now on CT), *i.e.*, the time interval between present time and the last earthquake which occurred. For a more accurate description of this model we refer to Kalbfleisch

and Prentice (1980) and Faenza *et al.* (2003); here we give a brief explanation of the model we proposed, stressing its main points.

First at all, this model is nonparametric because it does not assume any specific form for base-line hazard function $\lambda_0(\cdot)$; therefore we do not impose any *a priori* assumption on the earthquake occurrence process. In other words, we do not choose any arbitrary temporal distribution for fitting the events. The technique is robust because it allows us to consider at the same time all the available data coming from different regions. This is possible through the vector of covariate, \mathbf{z} , that is attached to each of the RVs and can contain any spatial/tectonic information on the subregion where IETs and/or CTs are sampled. There is a important assumption below this model. In eq. (2.1) the covariates act multiplicatively on the hazard function and they do not depend on time. This means that the shape of the base-line $\lambda_0(\cdot)$ *versus* time is always the same for each area apart for a multiplicative factor that depends on the covariates. So, from a physical point of view, the mechanism of earthquakes occurrence, described by $\lambda_0(\cdot)$, is the same for different areas; only the parameters of the system can vary (*i.e.*, $\exp(\mathbf{z}\boldsymbol{\beta})$).

The goal consists of estimating $\boldsymbol{\beta}$ and the nonparametric form of $\lambda_0(\cdot)$ in eq. (2.1). The vector of coefficients gives the relative importance of each covariates; $\lambda_0(\cdot)$ affords important insights into the physics of the process. A detailed explanation of such estimations can be found in Kalbfleisch and Prentice (1980) and Faenza *et al.* (2003).

2.1. Checking the model

We perform the statistical validation of the model using an independent dataset, *i.e.*, data that have not been considered at any step of the modeling. In order to do that the available dataset is divided into two parts, one used to set up the model (the *learning* phase), and the other to check the model (the *validation* phase). Each IET of the *validation* dataset is transformed in order to form residuals (see Kalbfleisch and Prentice, 1980; and Faenza *et al.*, 2003). This transformation is a sort of statistical standardization and it

changes the random non-negative point process into a Poissonian process with rate 1 (see, for instance, Ogata, 1988). Therefore, if the model is appropriate, the residuals are expected to behave like an exponential distribution with $\lambda=1$. The comparison between the cumulative of residuals and the theoretical exponential curve is checked through a one-sample Kolmogorov-Smirnov test (e.g., Gibbons, 1971). This provides a goodness-of-fit test of the model.

3. ETAS model and simulation of synthetic catalog

The Epidemic Type Aftershocks-Sequences (ETAS) model is a stochastic marked point process representing the occurrence of earthquakes of size larger than, or equal to, a threshold magnitude M_0 , in a region and in a period of time (Ogata, 1988). As in other triggering models, it is based on the principle that earthquakes are clustered in time and space because of occurrence of aftershocks; but, unlike those, it solves the debated problem to find the best way to identify clusters and to classify events (between mainshocks, aftershocks, foreshocks, ...). In fact, even if it considers the overall seismicity as the superposition of a background activity and of seismicity induced by previous earthquakes, its application to real data does not require any discrimination of events. We do not discuss the characteristics of this parametric model in details: a complete description of its formulation can be found in Ogata (1988, 1998).

An application of the ETAS model to Italian seismicity was provided by Console and Murru (2001) and by Console *et al.* (2003). Considering as marks magnitude (m_i) and epicentral coordinates (x_i, y_i), they inferred the following expression of conditional intensity function, based on history of occurrence $H_t = \{(t_i, m_i, x_i, y_i); t_i < t\}$

$$\lambda(t, m, x, y | H_t) = \left[\mu g(x, y) + \sum_{i: t_i < t} \frac{K e^{\alpha(m - M_0)}}{(t - t_i + c)^p} \frac{(q - 1) d^{2(q-1)}}{\pi [(x - x_i)^2 + (y - y_i)^2 + d^2]^q} \right] \cdot \beta e^{-\beta(m - M_0)}. \quad (3.1)$$

where $g(x, y)$ is the spatial density function of background events.

For details on formulation of function and on significance of parameters we refer to Console *et al.* (2003). The parameters of the model were estimated by the Maximum Likelihood Method on the national instrumental catalog, collected by INGV (Istituto Nazionale di Geofisica e Vulcanologia), for period 1987-2000. The values obtained are: $\beta=0.997 \cdot \ln(10)$, $\mu=0.0613$, $K=0.0014$, $c=0.0068$, $p=1.0580$, $\alpha=0.9740$, $d=3.07$, and $q=1.828$.

These parameters are used to simulate a synthetic catalog. It is developed following the thinning simulation procedure, outlined by Ogata (1981) for the Hawkes processes, of which the ETAS model is an application, and then adjusted by himself to the ETAS model (Ogata, 1998). In every realization events are simulated sequentially: first the time and then the magnitude and the epicentral coordinates are obtained. This method involves simulating the time to the next event, using a rate equal to an upper boundary of the intensity function, and calculating the intensity at this point. The ratio of this rate with the upper boundary is compared with a uniform random number to determine if the time is retained or not. Then epicentral coordinates and magnitude are simulated according with density functions chosen.

The synthetic catalog (from now on SC) is simulated with the same cutoff magnitude used to estimate parameters by real catalog ($M_0=3.0$); then only events with magnitude larger than, or equal to, 5.5 were selected. The interval time is the same of the historical real catalog (1600-2002) used in Faenza *et al.* (2003) to test the PHM model. The events obtained are 131.

4. PHM applied to ETAS catalog

We apply PHM to SC (see above) to compare these results with those obtained using the real catalog (from now on RC). As in the previous work (see Faenza *et al.*, 2003), the Italian territory has been divided into a 13×13 grid between the latitudes 36-48 N, and longitudes 5-20 E. Each node of the grid is the center of a circle. In order to cover the whole area the ra-

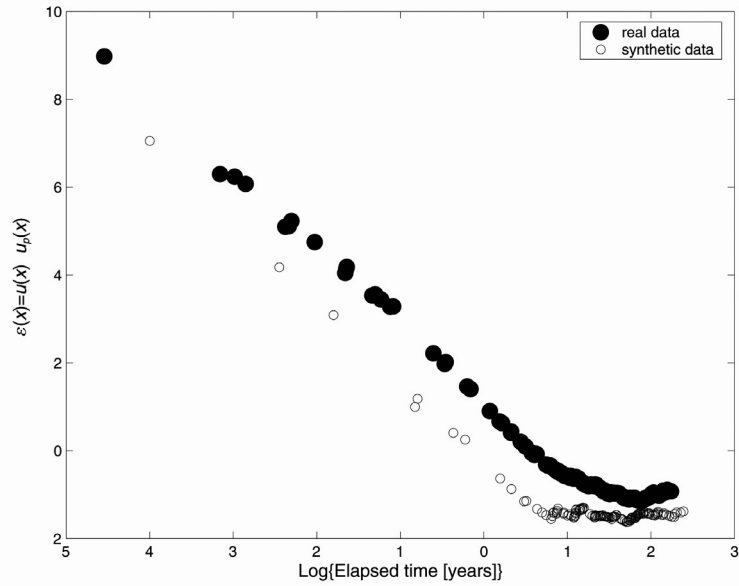


Fig. 1. Plot of the residuals $\varepsilon(x)$ (see equation 23 in Faenza *et al.*, 2003) for real and synthetic catalogs as a function of the time elapsed since the most recent event.

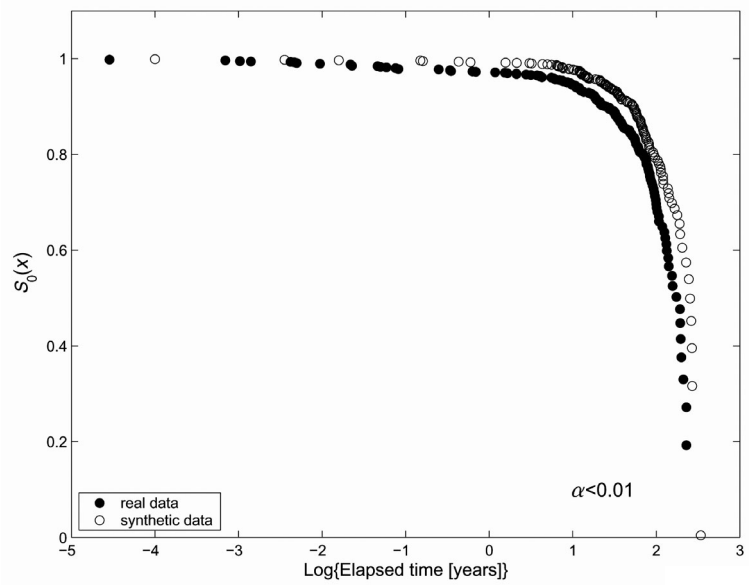


Fig. 2. Plot of empirical survivor functions for real and synthetic catalogs. The parameter α represents the significance level at which the null hypothesis of a common parent distribution can be rejected.

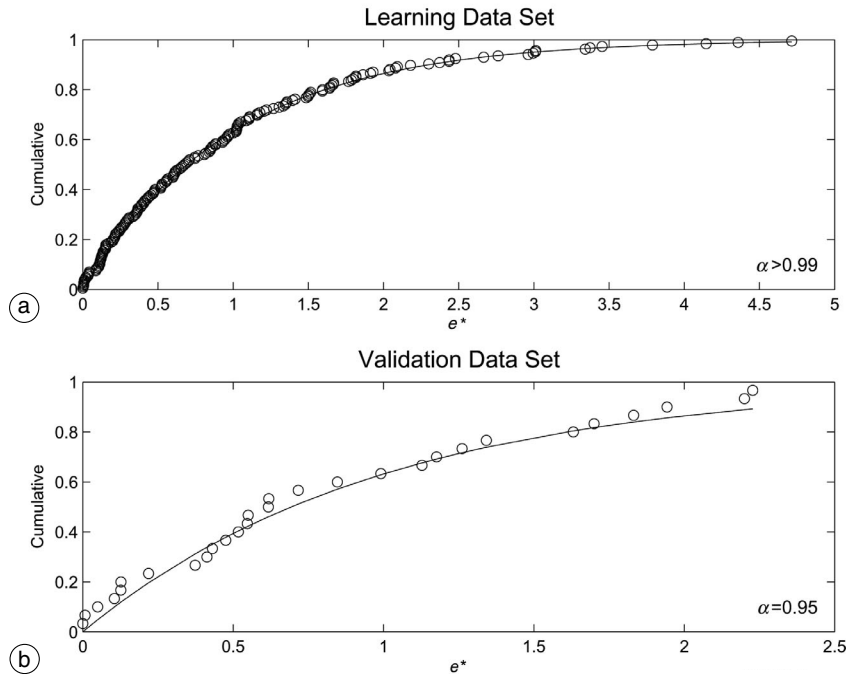


Fig. 3a,b. a) Empirical and theoretical (solid line) cumulative functions for the *learning* dataset. The parameter α is the significance level at which the null hypothesis (Poisson hypothesis) can be rejected (see text for more details). b) The same as (a), but relative to the *validation* dataset.

dius R of the circle is set about $D/\sqrt{2}$, where D is the distance between two nodes. For such grid, $D \approx 100$ km and $R = 70$ km.

In the following step, we select the circles that contain at least 3 earthquakes; 31 areas have been analyzed for SC, and 29 for RC. This choice makes the statistical analysis performed for the ETAS catalog homogeneous with that done to the real Italian catalog in Faenza *et al.* (2003). Moreover, we emphasize that this technique is robust because it considers all the spatially inhomogeneous data simultaneously to build the statistical model. For each one of these circles we calculate the IETs among the earthquakes inside the circle, one CT relative to the time elapsed since the most recent event, and a two-dimensional vector of covariates z , attached to each RV, bearing the information about the event itself and the area where this is

sampled. Specifically, z is composed of the logarithm of the rate of occurrence, *i.e.*, the total number of clusters which occurred inside the circle divided by 403 years (1600-2002), and by the magnitude of the event from which we calculate the IET and the CT. By using equations 10 and 11 in Faenza *et al.* (2003), we obtain $\beta_1 = 1.2 \pm 0.2$ and $\beta_2 = 0.2 \pm 0.2$. As for RC (see Faenza *et al.*, 2003) β_2 is not significantly different from zero, thus the magnitude of the events is not important for the distribution of the IETs. On the contrary, β_1 is positive and significantly different from zero.

Figure 1 compares between the residuals $\varepsilon(x)$ (see equation 23 in Faenza *et al.*, 2003) for SC and RC, respectively. The trend of $\varepsilon(x)$ is comparable to the trend of $\lambda_0(\cdot)$ (see Faenza *et al.*, 2003); in particular, the use of cumulatives to calculate $\varepsilon(x)$ make it a filtered version

of $\lambda_0(\cdot)$. In both cases a negative trend is shown. In other words, earthquakes tend to be more clustered than in a Poisson process, where $\lambda_0(\cdot)$ is a constant. This is what we expect to obtain from the ETAS model, since it is constructed imposing a cluster of events.

An important result is that for SC the time clustering seems to be shorter than for RC. Both catalogs show a negative trend (indicating clustering) that becomes almost flat for larger times as in a Poisson process; fig. 1 shows that the flattening occurs before for SC than for RC. We suggest that this longer clustering for RC may be due to the fact that the interaction between earthquakes lasts longer than that imposed by the specific ETAS model used. Note that in both cases there is no evidence of a positive trend as expected for the gap seismic hypothesis.

A quantitative test of the difference between RC and SC is performed through a two-sample Kolmogorov-Smirnov of the empirical survivor functions of the PHM reported in fig. 2. The null hypothesis of equal distribution is rejected at a significance level $\alpha < 0.01$.

Figure 3a,b reports a graph representing the goodness-of-fit of the PHM model applied to SC. As for the study of RC (Faenza *et al.*, 2003), the time interval 1600-1950 (109 earthquakes) is used for the *learning* phase, *i.e.*, to set up the model. The time interval 1951-2002 (22 earthquakes), instead, is used for the *validation* phase, *i.e.*, to verify the model with an independent dataset. The goodness-of-fit is quantitatively evaluated through a one-sample Kolmogorov-Smirnov test (*e.g.*, Gibbons, 1971). The significance levels at which the null hypothesis of equal distributions is rejected are reported in figure. They are both > 0.95 .

5. Concluding remarks

The main goal of this paper was to investigate on the capability of a specific ETAS model, successfully used to describe aftershock sequences, to model also the spatio-temporal distribution of large earthquakes ($M \geq 5.5$) which occurred in Italy in the last four centuries. To this purpose, we compared the hazard functions

obtained by the real catalog and by a synthetic catalog generated by an ETAS model, the parameters of which have been estimated by Console *et al.* (2003) by analyzing aftershock sequences. The two catalogs are both clustered in time, but the length of the clustering seems to be significantly different. For the real catalog it reaches a few years and after this time the distribution of the earthquakes appears to follow a Poisson distribution (see also Faenza *et al.*, 2003). In the synthetic catalog the cluster is shorter: the negative trend lasts for a few months after the event, and then it becomes flat as in a Poisson distribution. Thus, the cluster imposed by the specific ETAS model used (described in Section 3) seems to be shorter than the one in the real catalog. A plausible reason might be that the interaction between earthquakes may last for longer than the typical characteristic time of aftershock sequences.

REFERENCES

- CONSOLE, R. and M. MURRU (2001): A simple testable model for earthquake clustering, *J. Geophys. Res.*, **106** (5), 8699-8711.
- CONSOLE, R., M. MURRU and A.M. LOMBARDI (2003): Refining earthquake clustering models, *J. Geophys. Res.*, **108** (B10), 2468, doi:10.1029/2002JB002130.
- FAENZA, L., W. MARZOCCHI and E. BOSCHI (2003): A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the Italian catalogue, *Geophys. J. Int.*, **155**, 521-531.
- GIBBONS, J.D. (1971): *Non-parametric Statistical Inference* (McGraw-Hill, New York), pp. 306.
- KALBFLEISCH, J.D. and R.L. PRENTICE (1980): *The Statistical Analysis of Failure Time Data* (John Wiley and Sons, New York), pp. 336.
- OGATA, Y. (1981): On Lewis' simulation method for point processes, *IEEE Trans. Inf. Theory*, IT-30, 23-31.
- OGATA, Y. (1988): Statistical model for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Ass.*, **83**, 9-27.
- OGATA, Y. (1998): Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, **50**, 379-402.
- WORKING GROUP ON CALIFORNIA EARTHQUAKE PROBABILITY (1999): Earthquake probability in the San Francisco Bay region: 2000-2030 – A summary of findings, *U.S. Geol. Surv. Open-File Rep.* 99-517.

(received March 5, 2004;
accepted May 27, 2004)