

The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment

Anna Maria Lombardi^{1,*} and Warner Marzocchi¹

¹ *Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy*

Article history

Received October 30, 2009; accepted August 10, 2010.

Subject classification:

Earthquake probability, Forecasting, Italian seismicity, Hypothesis test, Aftershocks.

ABSTRACT

This study investigates the basic properties of the recent shallow seismicity in Italy, through stochastic modeling and statistical methods. Assuming that earthquakes are the realization of a stochastic point process, we have modeled the occurrence rate density in space, time and magnitude using an epidemic-type aftershock sequence model. By applying the maximum likelihood procedure, we estimated the parameters of the model that best fit the Italian instrumental catalog, as recorded by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) from April 16, 2005, to June 1, 2009. Then we applied the estimated model to a second independent dataset (June 1, 2009, to September 1, 2009). We show that the model performed well on this second database, through the relevant statistical tests. The model proposed in the present study is suitable for computing earthquake occurrence probability in real time, and to take part in international initiatives such as the Collaboratory Study for Earthquake Predictability, where we have submitted this model for daily forecasting of Italian seismicity for $M_L > 4.0$.

Introduction

There is a growing consensus for the acceptance of the existence of an intrinsic stochasticity of earthquake generating processes [see Vere-Jones 2006, for a review on the use of stochastic models for earthquake occurrence]. This view has promoted the formulation of different stochastic models that have been applied on different spatio-temporal scales [Kagan and Knopoff 1981, Ogata 1988, Ogata 1998, Kagan and Jackson 2000, Faenza et al. 2003, Rhoades and Evison 2004, Gerstenberger et al. 2005, Helmstetter et al. 2006, Lombardi et al. 2006, Lombardi et al. 2007, Marzocchi and Lombardi 2008, Lombardi et al. 2010]. Each model describes one or more different coexisting physical processes (tectonic loading, co-seismic stress interactions, post-seismic deformation, aseismic processes, and others) that have greater or lesser relevance to earthquake occurrence, depending on the maturity of the seismic cycle. Here, we have focused our attention on daily forecasts. For this class of forecast, stochastic models that describe the phenomenon of earthquake clustering are

becoming widely accepted in the seismology community [e.g., Reasenbergs and Jones 1989, Reasenbergs and Jones 1994, Gerstenberger et al. 2005, Marzocchi and Lombardi 2009].

We describe here a short-term, earthquake-forecasting model that we have submitted to the European Union – Italy Collaboratory Studies for Earthquake Predictability (CSEP) experiment. This forecast method uses earthquake data only, with no explicit use of tectonic, geological or geodetic information. The method is based on the observed regularity of earthquake occurrence, rather than on any physical model. The basis underlying this earthquake-forecasting method is the popular concept of an «epidemic» process: every earthquake is a potential triggering event for subsequent earthquakes [Ogata 1988, Ogata 1998, Console et al. 2003, Helmstetter et al. 2006, Lombardi and Marzocchi 2007]. We have applied a version of the epidemic-type aftershock sequence (ETAS) model to the seismicity recorded in Italy over recent years. For a first retrospective test, we applied a well-know procedure that consists of fitting the model to the early part of the Italian earthquake catalog and then testing it on the most recent part of the dataset. The real-time forecasting performance of the model was successfully checked on the occasion of the recent L'Aquila earthquake in central Italy on April 6, 2009 (M_w 6.3) [see Marzocchi and Lombardi 2009].

The spatio-temporal ETAS model

The ETAS model [Kagan and Knopoff 1981, Ogata 1988, Kagan 1991, Ogata 1998] is a stochastic point process of particular relevance for modeling co-seismic, stress-triggered aftershock sequences. Its formulation followed on from the observation that aftershock activity cannot always be predicted by a single modified Omori function [Omori 1894, Utsu 1961], and that seismicity can include the production of conspicuous secondary aftershocks. Therefore, this model assumes that each aftershock has some magnitude-dependent «ability» to perturb the rate of earthquake production, and therefore to generate its own Omori-like

aftershock decay. Following the first time-magnitude formulation proposed by Ogata [1988], many other time-magnitude-space versions have been published, most of which have been based on empirical studies of past seismicity [Ogata 1998, Zhuang et al. 2002, Console et al. 2003, Helmstetter et al. 2006, Lombardi and Marzocchi 2007]. These approaches have described the seismicity rate of a specific area as the sum of two contributions: the «background rate» and the «rate of triggered events». The first refers to the seismicity that is not triggered by previous events in the catalog, and the second is associated with stress perturbations that have been caused by previous earthquakes in the catalog.

The ETAS model defines the total space-time conditional intensity $\lambda(t, x, y, m/H_t)$ (i.e., the probability of an earthquake occurring in the infinitesimal space-time volume conditioned to all of the past history), according to Equation (1):

$$\lambda(t, x, y, m/H_t) = \left[\nu u(x, y) + \sum_{t_i < t} \frac{K}{(t - t_i + c)^p} e^{\alpha(M_i - M_c)} \frac{c_{d,q}}{[r_i^2 + d^2]^q} \right] \beta e^{-\beta(m - M_c)} \quad (1)$$

where $H_t = \{(t_i, x_i, y_i, M_i); t_i < t\}$ is the observation history up to time t , M_c is the completeness magnitude of the catalog, ν is the rate of background seismicity for the whole area, K , c and p are the parameters of the modified Omori Law that describe the decay in time of short-term triggering effects, α determines how the triggering capability depends on the magnitude of an earthquake, the parameters d and q characterize the spatial probability density function (PDF) of the triggered events, $c_{d,q} = \frac{q-1}{\pi} [d^{2(q-1)}]$ is the relative normalization constant, r_i is the distance between the location (x, y) and the epicenter of the i -th event (x_i, y_i) , the function $u(x, y)$ is the spatial PDF of the background events, and finally, $\beta = b \ln(10)$ is the parameter of the well-known Gutenberg-Richter Law [Gutenberg and Richter 1954], which is assumed to hold for all magnitudes and to be invariant in space. Specifically, the model assumes that large earthquakes are indistinguishable from smaller earthquakes, and therefore they have the same distribution.

The most recent versions of the ETAS model [Helmstetter et al. 2006, Ogata and Zhuang 2006] have been characterized by the introduction of a further term that takes into account the correlation between the aftershock area and the magnitude of the triggered events. Some preliminary results have shown that this correlation might be negligible for Italy [see Marzocchi and Lombardi 2009]. We therefore decided to use the version of the ETAS model as described by Equation (1), in which the spatial decay of the triggered activity is independent of the magnitude of the triggering shock. A deeper analysis of this topic will be presented and discussed in future reports.

The parameters $(\nu, K, c, p, \alpha, d, q, \beta)$ of the model for the

events within a time interval $[T_{start}, T_{end}]$ and a region R can be estimated by maximizing the log-likelihood function [Daley and Vere-Jones 2003], as given by Equation (2):

$$\log L(\nu, K, c, p, \alpha, d, q, \beta) = \sum_{i=1}^N \log \lambda(t_i, x_i, y_i, m/H_{t_i}) - \int_{T_{start}}^{T_{end}} \int_R \int_{M_c}^{M_{max}} \lambda(t, x, y, m/H_t) dt dx dy dm \quad (2)$$

where M_{max} is the expected maximum magnitude for the region R . The parameters of the model are estimated using the iteration algorithm developed by Zhuang et al. [2002]. By using a suitable kernel, in addition to the model parameters, this procedure provides an estimation of the PDF $u(x, y)$ for background events. The background rate is given by Equation (3):

$$\nu u(x, y) = \frac{1}{T} \sum_j p_j K_{d_j}(r_j) \quad (3)$$

where T is the length of time recovered by the dataset, p_j is the probability that the j -th event is not triggered by previous shocks in the catalog, and K_{d_j} is a Gaussian kernel function with a spatially variable bandwidth. Similarly, the rate of triggered events is given by Equation (4):

$$c(x, y) = \frac{1}{T} \sum_j (1 - p_j) K_{d_j}(r_j) \quad (4)$$

Several physical investigations have shown that static stress changes decrease with epicentral distance as r^{-3} [Hill et al. 1993, Antonioli et al. 2004]; therefore, in the present study we imposed $q = 1.5$. This choice was also justified by the trade-off between parameters q and d that can cause different pairs of q and d values to provide almost the same likelihood of the model [Kagan and Jackson 2000].

Testing the model

The gold standard for scientific evaluation of earthquake-forecasting models is through the comparison of forecasts and true values in prospective experiments [see, e.g., Field 2007, Schorlemmer et al. 2007, Luen and Stark 2008, Zechar et al. 2009]. However, a model can also be evaluated through retrospective experiments; e.g., by dividing an available dataset into two parts, the first of which can be used to set up the model, hereinafter the learning dataset, and the second of which can be used to check the reliability of the model, hereinafter the testing dataset [Kagan and Jackson 2000]. Verification of the forecasting capability of the model can be achieved by a comparison of the observations and the forecasts. This testing enables verification whether the model is performing significantly well, and eventually, identification of the features that will allow better forecasting. In the following subsections, we describe the statistical tests used in the present study to retrospectively check our model.

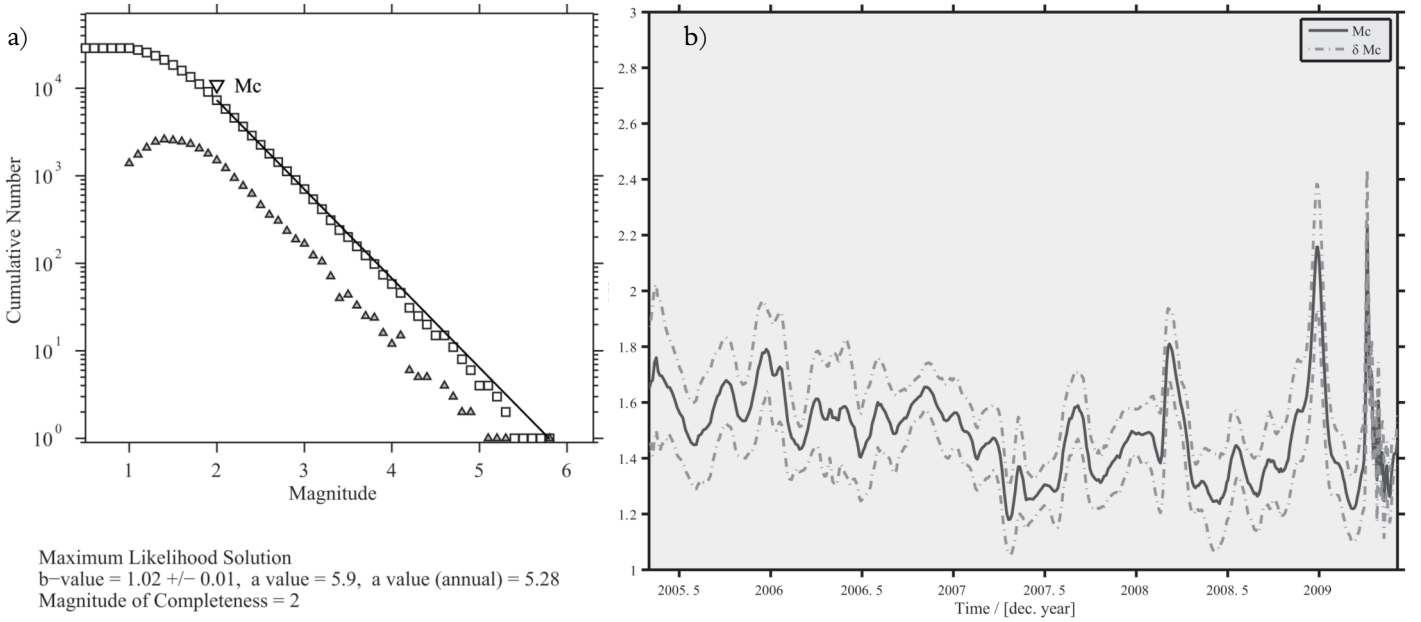


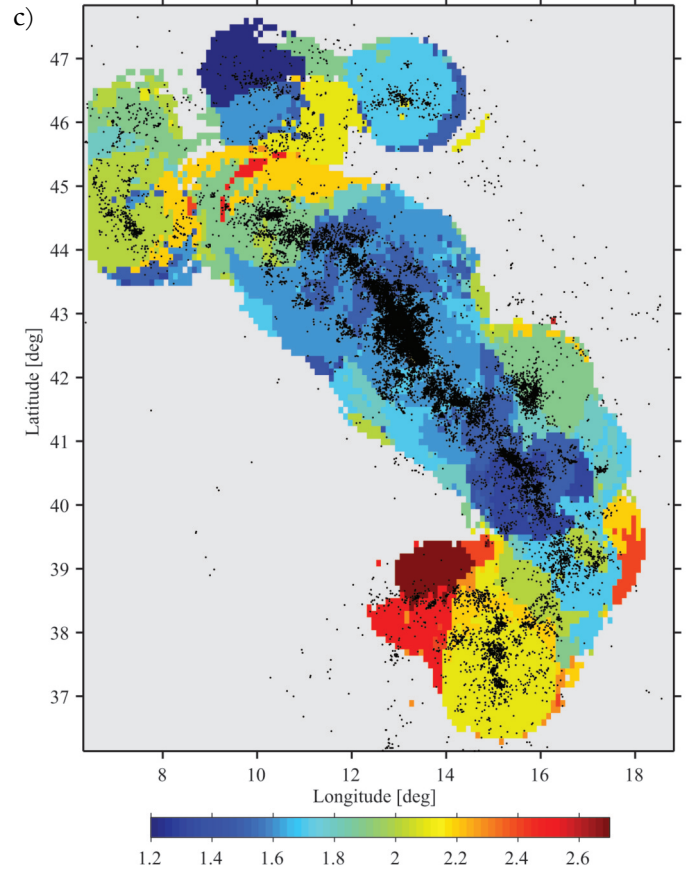
Figure 1. Completeness magnitude of the INGV bulletin (from April 16, 2005, to June 1, 2009) obtained by the maximum likelihood method. a) Frequency magnitude distribution for the whole dataset, giving $M_c = 2.0$. b) M_c as a function of time. c) M_c as a function of space.

Residuals analysis

A common diagnostic technique that can be used for stochastic point processes is based on the transformation of the time axis t into a new scale, τ , using the increasing function given in Equation (5):

$$\begin{aligned} \tau = \Lambda(t) &= \int_{T_{start}}^t dt' \int_R dx dy \int_{M_c}^{M_{max}} dm \lambda(t', x, y, m/H_t) = \\ &= \int_{T_{start}}^t \left[\nu + \sum_{t_i < t'} \frac{ke^{\alpha(M_i - M_c)}}{(t' - t_i + c)^p} \right] dt' \end{aligned} \quad (5)$$

where T_{start} is the starting time of the observation history H_t [Ogata 1988]. The random variable τ represents the expected number of occurrences in the time period $[T_{start}, t]$ in whole region R and with a magnitude above the M_c . If a model with a conditional intensity $\lambda(t, x, y, m/H_t)$ describes the temporal evolution of the process well, the transformed data $\tau_i = \Lambda(t_i)$, which are known in statistical seismology as the residuals, are expected to behave like a stationary Poisson process with a unit rate [Ogata 1988]. Therefore, the values $\Delta\tau_i = \tau_{i+1} - \tau_i$ are random variables that are independent and exponentially distributed (with a mean of 1). We can check this hypothesis for the residuals of our ETAS model by means of two nonparametric tests: the Runs test, to verify the reliability of the independence property, and the one-sample Kolmogorov-Smirnov (KS1) test, to check the standard exponential distribution [Gibbons and Chakraborti 2003, Lombardi and Marzocchi 2007]. We used both of these tests because the KS1 test is not effective for checking the presence of memory in a time series. Hence, any discrepancies in the residuals of the Poisson hypothesis that are identified by either one or both of these tests will be a sign of the inadequacy of the ETAS



model to explain all of the basic features of the seismicity analyzed. This check on the analysis is similar to the N-Test, which is currently used by the Regional Earthquake Likelihood Model (RELM)/CSEP testing centers [Kagan and Jackson 1995, Schorlemmer et al. 2007], although it avoids the time binning that can lead to bias in the results of the testing phase [see, e.g., Lombardi and Marzocchi 2010].

Cumulative reliability diagram

The reliability diagram is a common diagnostic technique that is used to measure the consistency of a forecast model

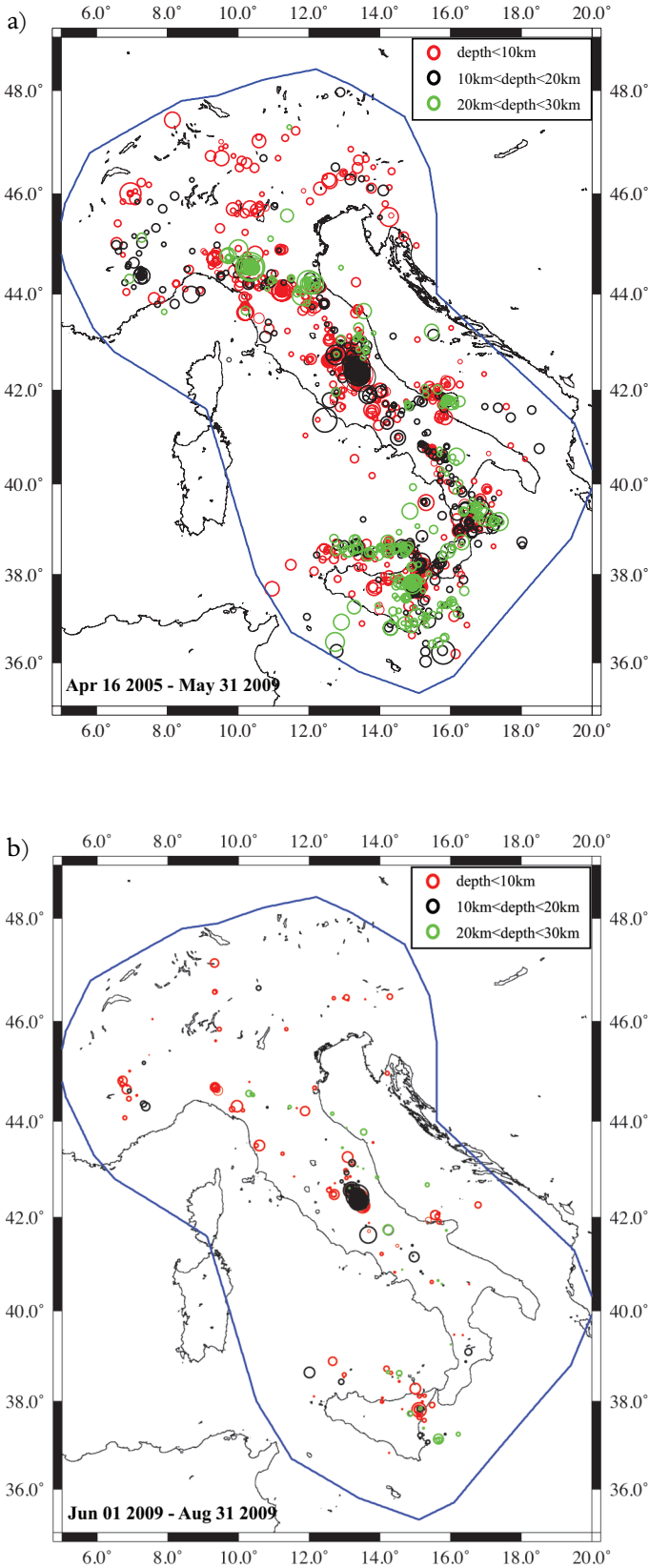


Figure 2. Maps of the seismic events with magnitudes >2.5 and depths <30 km that occurred in Italy inside the collection area identified by the CSEP experiment (blue solid line) [see Schorlemmer et al. 2010]. The symbol sizes are scaled with magnitude. a) Events of the learning dataset used to set-up the model (April 16, 2005, to June 1, 2009; 2,100 events). b) Events of the testing dataset used for the retrospective forecasting test of the model (June 1, 2009 to September 1, 2009; 179 events).

according to the observations. Roughly speaking, a probability forecast is reliable if the events actually happen with an observed frequency that is consistent with the forecast. More specifically, a reliability diagram consists of a plot of the observed relative frequencies against the predicted probabilities [Wilks 2005]. Reliability measures sort the forecast/observation pairs (F_j/O_i) into groups according to the values of the forecast variables, and they characterize the conditional distributions of the observations according to the forecasts. In particular, a way to visually identify departures from reliability is to plot the cumulative conditional observed frequency $p(O_i/F_j)$ against the cumulative predicted probability F_j ; this provides the cumulative reliability diagram; perfect reliability will be seen as the diagonal.

We have used this type of analysis to check the predicted spatial distribution of the observed seismicity. Specifically, we applied a case of dichotomous events, i.e., the observations are limited to two possible outcomes: the occurrence (O_1) or nonoccurrence (O_2) of an earthquake. To define the forecasting cumulative probabilities F_j , the area under analysis was partitioned in a nonoverlapping and exhaustive set of cells C_i ; for each cell, we computed the proportion of events f_i that were expected by the forecasting model. These values of f_i , which by definition were between 0 and 1, were sorted into ascending order and grouped into N bins B_j ($j = 1 \dots N$), which formed a partition of the unit interval composed of the overlapping increasing subintervals. These bins were characterized by a set of forecasting probabilities F_j that defined the probability to have at least one event in B_j , as shown by Equation (6):

$$I_j = \{i; f_i \in B_j\} \quad \sum_{i \in I_j} f_i \leq F_j \quad (6)$$

The most intuitive choice was to take F_j as equally spaced. If the distribution of the forecasts was nonuniform, then choosing the bins so that the sets I_j were equally populated (i.e. with the same number of events f_i) was a good alternative. The values F_j were compared with the cumulative observed frequencies, as shown in Equation (7):

$$P(O_1 | F_j) = \sum_{i \in I_j} \frac{N_i}{N} \quad (7)$$

where N_i is the observed number of shocks in the cell C_i , and N is the total number of events. In the case of perfect reliability, the conditional probability $p(O_i/F_j)$ is equal to F_j .

The INGV database

Italy is characterized by a generally high seismicity, with observed magnitudes up to about 7.5. The long tradition of seismological studies in Italy has produced many studies of seismic data collection, and therefore today Italy can boast of their careful seismic instrumental catalogs [Castello et al. 2005, Schorlemmer et al. 2010; <http://iside.rm.ingv.it/>], as well as of tried and tested experience in the compilation of historical

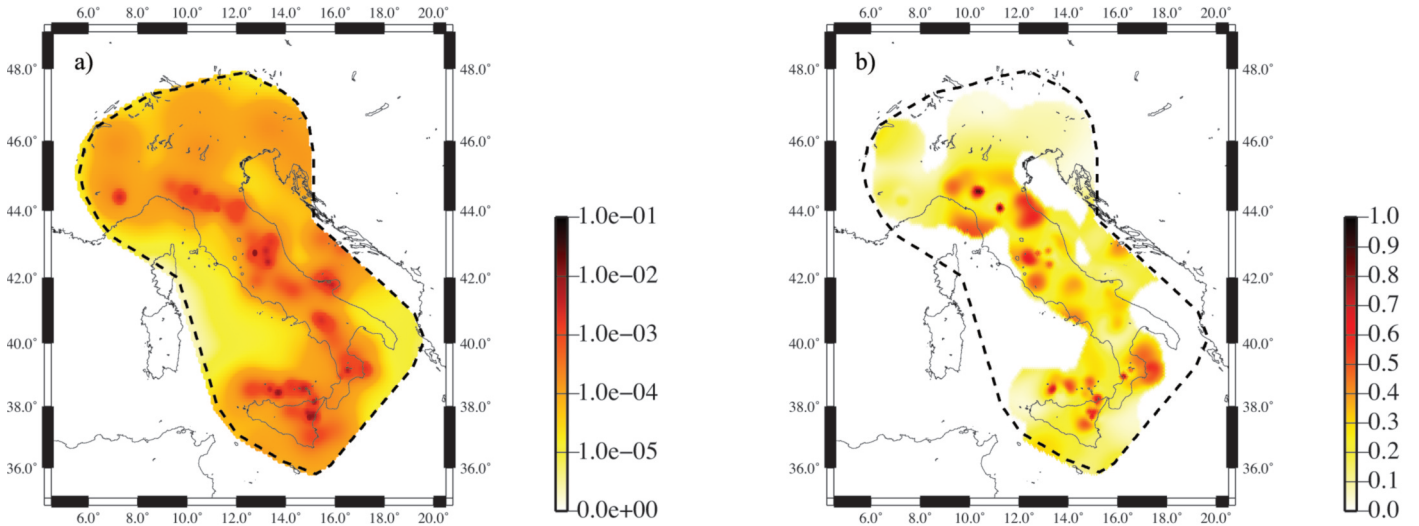


Figure 3. Maps of: a) the background seismicity rate $\nu u(x,y)$; and b) the ratio between the triggered rate and the total seismic rate of the INGV bulletin learning dataset (April 16, 2005, to June 1, 2009; 2,100 events).

databases [Boschi et al. 2000]. The most complete instrumental catalog of Italian seismicity is the seismic bulletin of the INGV, Istituto Nazionale di Geofisica e Vulcanologia (Bollettino Sismico Italiano; <http://bollettinosismico.rm.ingv.it/> and <http://iside.rm.ingv.it>). The Italian seismic network has changed significantly over the last few years. Indeed, April 16, 2005, marked the date when remarkable changes in the Italian seismic network [Bono and Badiali 2005; see also Schorlemmer et al. 2010] and the data processing began. Given the large differences seen for the INGV bulletin before and after this date, we set up our model using the parameters of the events collected from April 16, 2005, to June 1, 2009. The earthquakes from June 1, 2009, to September 1, 2009, were instead used for the first retrospective test of the model (the testing dataset). In agreement with the CSEP requirements, we selected events at <30 km in depth that occurred in the collection area, as defined by the CSEP experiment.

A correct understanding of the physical processes that control the rate of earthquake production depends on the quality of the seismic catalog that is available. Specifically, a critical issue that has to be addressed before performing any investigation is an assessment of the completeness of the dataset. Here, we determined the completeness magnitude (M_c) (the lowest magnitude at which a negligible number of events are not detected) and its variation with time. The algorithms used in this study are freely available together with the software package ZMAP [Wiemer 2001]. To estimate M_c , we used the method proposed by Cao and Gao [2002] that is based on the stability of the b -value as a function of the cut-off magnitude [see also Woessner and Wiemer 2005, for details on the method]. The numerical stabilization of the b -value was measured using the maximum likelihood method [Aki 1965] to estimate the b parameter, and the algorithm proposed by Shi and Bolt [1982] to quantify its uncertainty [see also Marzocchi and Sandri 2003]. These algorithms provided a value of M_c (as local magnitude) of 2.0 and a

b -value of 1.0. (Figure 1a). The analysis of the spatio-temporal variation of the M_c showed clear changes with time (Figure 1b) and space (Figure 1c). We performed these analyses using a minimum number of events of 100 and a radius of 50 km. In particular, M_c reached about 2.5 soon after the recent L'Aquila earthquake (April 6, 2009; M_w 6.3; Figure 2b). This value appeared to be a reliable completeness threshold for most of the national territory (Figure 2c). These data are also in agreement with Schorlemmer et al. [2010], who defined $M_c = 2.5$ as a reasonable magnitude threshold for most of the Italian territory. The only exception was for the southern part of Apulia and the western part of Sicily, which showed higher M_c [see also Schorlemmer et al. 2010, for details]. Considering the small size of these areas, for the present study we selected the events above a magnitude of 2.5 that were recorded in the INGV bulletin (2,100 events for learning, and 179 events for testing). Figure 2 shows the distributions of the selected seismic activities for both the learning (Figure 2a) and the testing (Figure 2b) databases, together with the boundaries of the collection area defined by the CSEP laboratory.

Parameter	Value
ν	237 ± 8 (year $^{-1}$)
K	0.011 ± 0.001 (year $^{p-1}$)
p	1.16 ± 0.02
c	0.00004 ± 0.00001 (year)
α	1.3 ± 0.1
d	1.10 ± 0.05 (km)
q	$\equiv 1.5$
Log-likelihood	-7808.1

Table 1. Parameters of the ETAS model for Italian seismicity ($M_c = 2.5$; April 16, 2005, to June 1, 2009; 2,100 events).

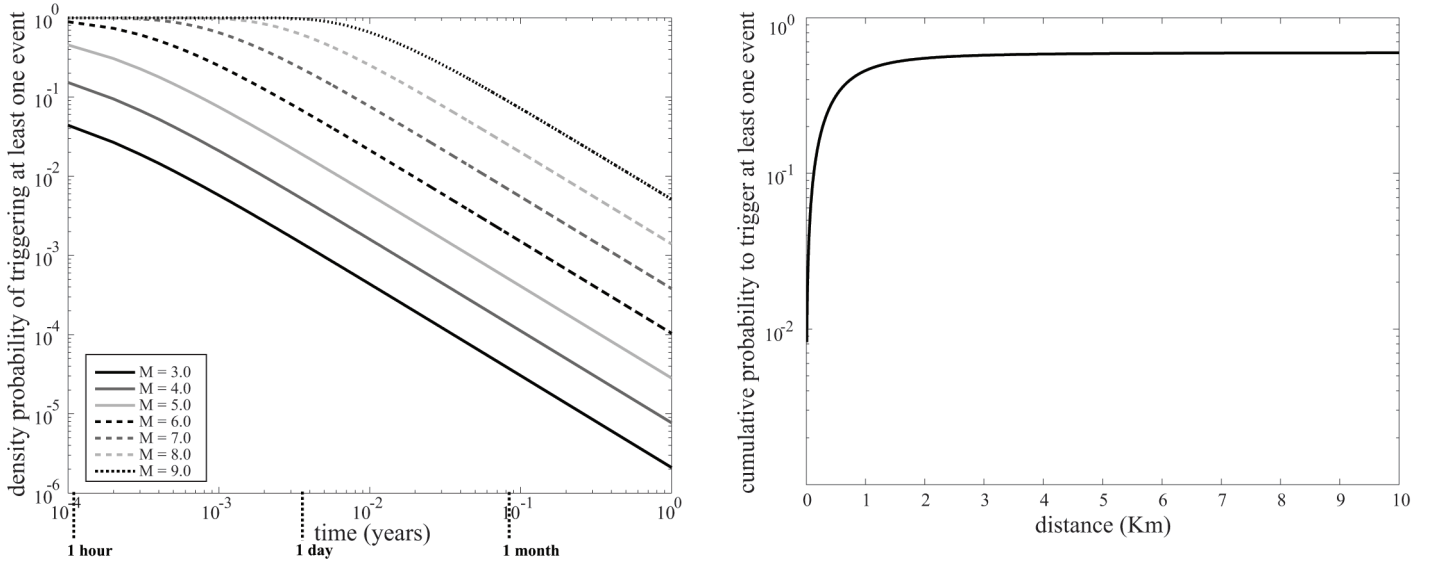


Figure 4. Spatio-temporal behavior of the triggering probability inferred by the ETAS model. a) Time decay (according to the Omori law) of the probability of generating at least one event for different magnitudes. b) Cumulative spatial probability distribution of triggering at least one event (see Equation 1).

Application and testing of the ETAS model on the Italian seismicity

We applied the ETAS model to this Italian seismicity recorded in the learning database described above. Following the procedure proposed by Zhuang et al. [2002], we estimated the model parameters together with the spatial distribution of the background seismicity ($u(x,y)$). Table 1 lists the inferred values of these model parameters, together with their standard errors and the associated log-likelihood values. The totals for the triggered and spontaneous events identified by the model were 46% and 54%, respectively. Figure 3 shows two maps: the first represents the distribution of the time-independent background rate ($\nu u(x,y)$; see Equation 3); and the second the distribution of the clustering ratio $r(x,y)$; i.e., the ratio between the triggered and the total rates, for the whole learning period. The clustering ratio was obtained according to Equation (8):

$$r(x,y) = \frac{c(x,y)}{\int_{T_1}^{T_2} \int_{M_c}^{M_{max}} \lambda(t,x,y,m/H_t) dt dm} \quad (8)$$

where $c(x,y)$ and $\lambda(t,x,y,m/H_t)$ are defined by Equations (1) and (4), respectively. By comparing the two maps shown in Figure 3, it can be seen that the spatial distribution of the triggering capability was not a proxy for the seismogenic potential. As an example, the southern part of the Italian peninsula showed a lower triggering rate with respect to the other areas (Figure 3b) whenever this area was one of most active of the whole region (Figures 2 and 3a). The estimated Omori law decay predicted that the probability of triggering one or more events with a magnitude above 2.5 for an earthquake of magnitude 3.0 was below 1% after about 5 to 6 hours. The corresponding times for a triggering an event of magnitude 5.0 and 7.0 were 2 to 3 days and about 1 month,

respectively (Figure 4a). We stress that these probabilities referred to direct triggering effects, as the secondary triggered events were not included in this calculation. For the spatial decay of the triggering capability, an event had a 50% probability to trigger one or more events within 2 km of its epicenter, and about a 40% probability at a distance >10 km, regardless of its magnitude (Figure 4b).

A preliminary check of the goodness of the inferred ETAS model was carried out by applying residual analysis to the learning dataset used to set up the ETAS model. The residuals passed the KS1 test (p -value, 0.8), but the Runs test rejected the hypothesis of no correlation (p -value, 0.007). The cumulative distribution of the residuals (Figure 5a) showed a clear deviation from the expected Poisson behavior soon after the occurrence of the M_w 6.3 L'Aquila earthquake (April 6, 2009). If we took the L'Aquila sequence out of the learning period, the ETAS model also passed the Runs test (p -value, 0.07). We would argue that this result is probably due to the spatial variation of some of the parameters. In other words, at a local scale, the model might be significantly different with respect to the same model calibrated using the whole Italian territory. For example, Marzocchi and Lombardi [2009] reported an α -value of 1.5 for the L'Aquila region that increased to 2.0 when $M_c = 2.5$ was considered; this value is certainly larger than the 1.3 seen here for the whole Italian territory (see Table 1).

To test the forecasting performance of the ETAS model, we analyzed the residuals and plotted the cumulative reliability diagram of the testing dataset. Using the KS1 test, we were not able to reject the null hypothesis that the values $\Delta\tau_i = \tau_{i+1} - \tau_i$ are exponentially distributed (with a mean of 1) (p -value, 0.14). Figure 5b shows the cumulative number of residuals τ_i versus the transformed time τ (solid line) together with the expected linear scaling that was predicted by a

Poisson distribution (i.e., the cumulative number of residuals should lie along the bisector). Similarly, the Runs test did not reject the independence hypothesis of $\Delta\tau_i$ (p -value, 0.81), which implied that the hypothesis of uncorrelation of the residuals cannot be rejected. This result is supported by Figure 5c, in which we plotted the variables $U_{k+1} = 1 - \exp(-\Delta\tau_{k+1})$ versus U_k for the testing dataset. If $\Delta\tau_k$ are i.i.d. random exponential variables with a unit mean, the statistics U_k are i.i.d. random uniform variables on (0,1). Assuming that a possible correlation was likely to show up in neighboring intervals, the plot of U_{k+1} versus U_k should have recovered the figure panel uniformly [Ogata 1988].

The cumulative reliability diagram of the spatial distribution of the events collected by the testing dataset showed reliable forecasting (Figure 6). To define the forecasting probabilities F_j , we computed the expected fraction of events f_i according to the ETAS model, for each cell C_i of the testing grid defined by the CSEP laboratory. The values f_i were computed as the ratios between the expected numbers of events in the cell C_i and in the whole region R . Specifically, we used the formula given in Equation (9):

$$f_i = \frac{\int_T \int_{C_i} \int_M \lambda(t, x, y, m / H_t) dt dx dy dm}{\int_T \int_R \int_M \lambda(t, x, y, m / H_t) dt dx dy dm} \quad (9)$$

where T was the testing period, R was the testing area defined by the CSEP laboratory, M was the magnitude range [2.5, 9.0], and H_t was the occurrence history, starting on April 16, 2005 (i.e., including the learning period).

We then regrouped these values into 10 bins B_j , which were identified by increasing values of probabilities F_j . The error bars were defined so that the sets I_j (see Equation 6) were equally populated. Table 2 gives the values of the probabilities F_j and $p(O_1 | F_j)$ (i.e., the observed frequencies of the events in bin B_j), as defined by Equation (7). These were plotted as Figure 6. The error bars indicated the 95% confidence interval of the $p(O_1 | F_j)$ values. These last were obtained by applying the reliability analysis to 1,000 synthetic catalogs. These had the same duration of the testing period as the INGV bulletin and were simulated in agreement with the ETAS model, including the real learning period in the past history. The reliability diagram showed that the pairs $[F_j, p(O_1 | F_j)]$ were well fitted by the diagonal, which indicated perfect reliability. Moreover, they were in agreement with the variation expected by the model. All of these data showed that the model estimated on the learning dataset was in agreement with the seismicity that followed. This result was also supported by the observation that the parameters estimated from the entire catalog were not statistically different from the parameters listed in Table 1.

This model formulated and tested above allowed us to compute forecasts in the framework of the CSEP

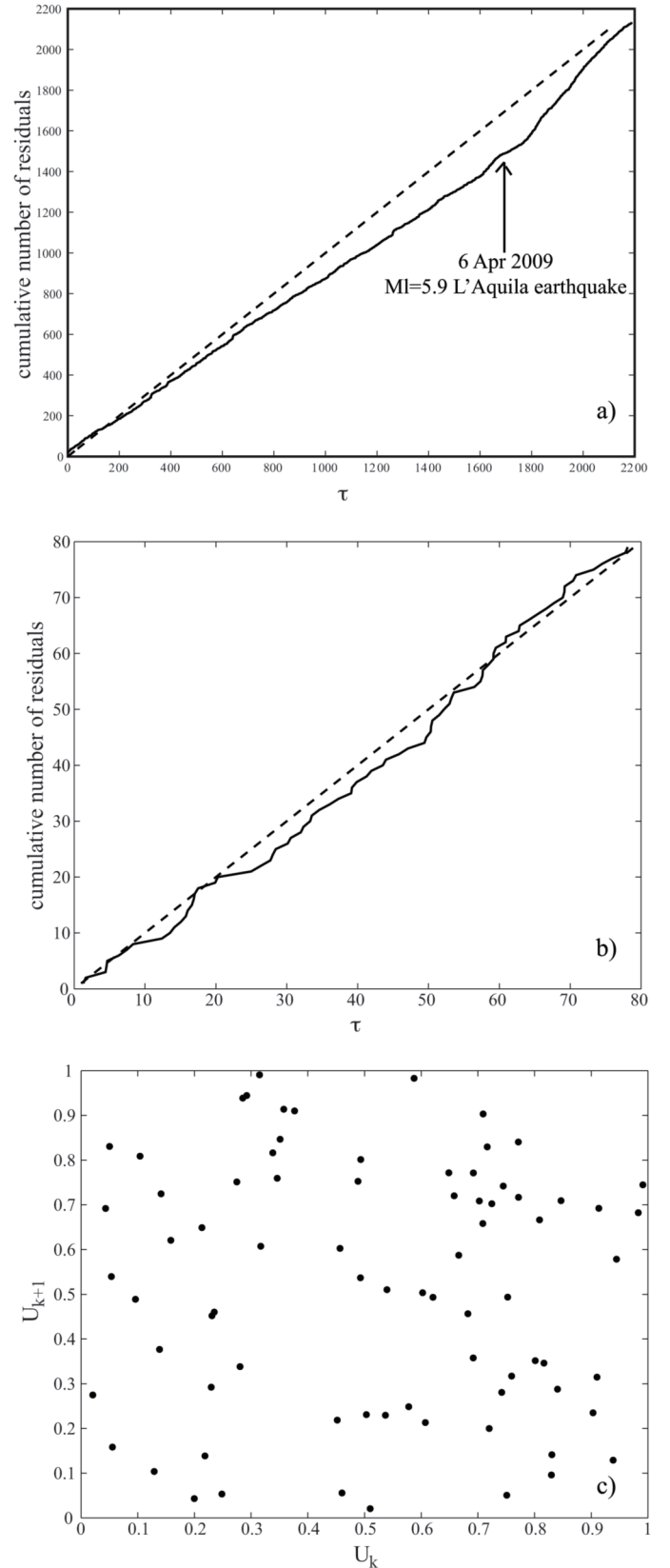


Figure 5. Residuals analysis of the ETAS model on the learning (April 16, 2005, to June 1, 2009; 2,100 events) and testing (June 1, 2009 to September 1, 2009; 179 events) INGV bulletin events. a) Cumulative number of transformed times τ_i (solid line) for the learning period together with the theoretical distribution (dotted line) predicted by a Poisson distribution. b) As for a) but for the testing period. c) Plot of the values $U_{k+1} = 1 - \exp(-\Delta\tau_{k+1})$ versus U_k for the testing period.

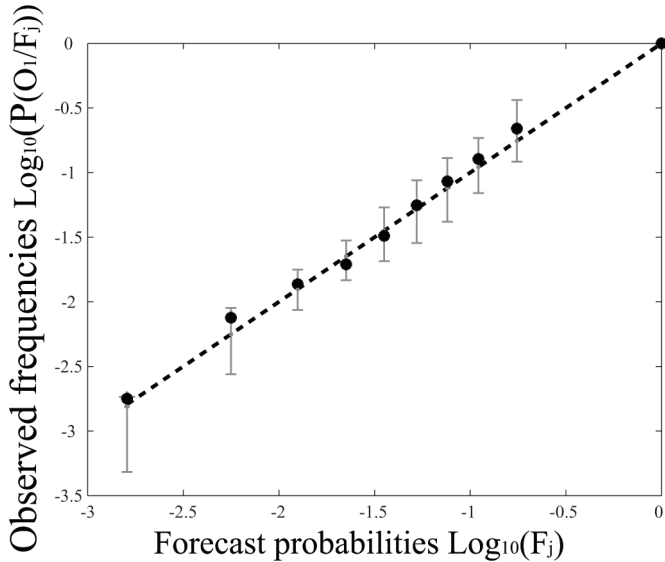


Figure 6. Cumulative reliability diagram of the spatial earthquake distribution predicted by the ETAS model for the testing INGV bulletin ($M_c = 2.5$; June 1, 2009 to September 1, 2009; 179 events). The stars mark the pairs $F_j/p(O_i|F_j)$, i.e., the forecasts and the observed spatial distributions. The dotted black line represents the perfect reliability. Error bars identify the 95% confidence intervals of the observed values $p(O_i|F_j)$. The forecast probabilities F_j identify equally populated bins B_j (see text for details).

Forecasts (F_j)	Observed frequencies [$p(O_i F_j)$]
1.6×10^{-3}	1.8×10^{-3}
5.6×10^{-3}	7.5×10^{-3}
1.2×10^{-2}	1.4×10^{-2}
2.2×10^{-2}	1.9×10^{-2}
3.5×10^{-2}	3.2×10^{-2}
5.2×10^{-2}	5.6×10^{-2}
7.6×10^{-2}	8.6×10^{-2}
1.1×10^{-1}	1.3×10^{-1}
1.7×10^{-1}	2.2×10^{-1}
1.0	1.0

Table 2. Values for the cumulative reliability diagram of the spatial distribution of earthquakes predicted by the ETAS model relative to the testing INGV bulletin. ($M_c = 2.5$; June 1, 2009, to September 1, 2009; 179 events).

experiment. These predictions were in a form of daily probabilities of occurrence for at least one earthquake with $M_L \geq 4.0$ within a cell of $0.1^\circ \times 0.1^\circ$ in Italy. These were obtained by integrating the intensity function of the ETAS model (Equation 1) for each cell C_i and for each forecasting period T_j . The forecast rates for $M_L > 4.0$ were obtained by rescaling the rates of earthquakes with $M_L > 2.5$, in agreement with the Gutenberg-Richter relation. Equation (1) shows that time-dependent modeling such as this ETAS model also takes

into account the triggering effects of the seismicity that occurred before and that is expected during the forecast interval. So we included in the past history all of the real seismicity with $M_L > 2.5$ and a depth > 30 km that occurred up to the start of the forecasting time window. Moreover, we simulated 1,000 different stochastic realizations for the forecasting time window using the thinning method proposed by Ogata [1998] and the intensity function formulated in Equation (1). Then we averaged the predictions coming from each of these synthetic catalogs.

Discussion and conclusions

In this study, we adopted a version of the ETAS model to describe the recent shallow seismicity that has occurred in Italy. The main motivation behind this study was to submit our model to the European Union – Italy CSEP laboratory for 1-day forecasts. To achieve this goal, we proposed a model that represented the main average properties of the Italian seismicity. The reliability of this model was successfully checked at the local scale in a real-time forecasting experiment, on the occasion of the occurrence of the recent L'Aquila destructive earthquake [Marzocchi and Lombardi 2009].

One finding of the present study is that the generalization of local models to the whole Italian territory can be problematic for different reasons. First, the M_c varied with space [Schorlemmer et al. 2010]; in this study, we adopted $M_c = 2.5$, which was probably optimistic for some zones. Indeed, the M_c for the whole of the Italian territory was about 2.9 (see Figure 1c, and Schorlemmer et al. 2010). We were aware of this limit, but we preferred to adopt a value of M_c that was reliable for most (if not all) of the Italian territory. The area with $M_c > 2.5$ covered only a relatively small part of the whole region. The use of a larger M_c caused a strong reduction in the size of the dataset, with the consequent increase in the uncertainty of our model. Of potentially greater importance, it has been recognized that smaller earthquakes have decisive roles in the triggering processes [Felzer et al. 2002, Helmstetter and Sornette 2003, Helmstetter et al. 2004]; therefore, a value of M_c that was set too high might have caused erroneous identification of the triggered section of the seismicity.

Secondly, some of the ETAS parameters can vary with space. This means that some parameters that were estimated for the whole territory or for a small region might be significantly different. Local variations can occur only as a consequence of the occurrence of large earthquakes. For example, the model proposed here for the whole Italian territory was not able to correctly reproduce the time evolution of the first part of the 2009 L'Aquila sequence (see Figure 3a). As indicated above, we argued that this discrepancy was probably due to features of the local seismicity that cannot be extrapolated to the whole territory. In particular, the seismicity of the 2009 L'Aquila sequence

was characterized by a larger α -value with respect to the whole Italian seismicity described by our ETAS model. The α parameter is crucial in the quantification of the dependence of the triggering effects according to the magnitude of the parent earthquake. The failure of our model to describe the starting phase of the 2009 L'Aquila sequence suggested that inconsistencies were possible in forecasts of the future seismicity. This problem might call for the development of more complicated models that can take into account the local features of seismic activity.

We would argue that other parts of this model can also be improved in the future. In the following, we report some of the possible indications in this direction. First, the model can be enhanced by adopting a modified magnitude distribution, to explicitly allow for the decrease in detection soon after a large earthquake [Kagan 1991, Helmstetter et al. 2006]. Secondly, the background rate and the basic clustering properties of the aftershock sequences were assumed to be stationary in time. This assumption was mainly motivated by the short learning dataset used; longer-term datasets might allow departures from stationarity to be captured, such as long-term time evolution of the seismicity [e.g., Lombardi and Marzocchi 2007, Marzocchi and Lombardi 2008]. Moreover, other time-dependent processes that act over short time scales, like fluid injection, might have significant impact on the short-term spatio-temporal evolution of the seismicity, and therefore it might be necessary to include these in the ETAS model [Hainzl and Ogata 2005, Lombardi et al. 2006, Lombardi et al. 2010]. Thirdly, the ETAS model proposed here assumed that all of the earthquakes were equal. Possible distinctive precursory activities that anticipate large shocks were not considered in this parametrization. Finally, the present model did not incorporate tectonic/geological information. Inclusion of such information might represent a possible future direction of investigation for the improvement of the forecasting of large shocks. For example, the Gutenberg-Richter law is used everywhere indistinctively; this means that a magnitude 8 earthquake is considered possible anywhere. It can thus be argued that in the future, geological information will provide a more appropriate frequency-magnitude law that varies in space.

References

- Aki, K. (1965). Maximum likelihood estimate of bin the formula $\log(N) = a - bM$ and its confidence limits, *Bull. Earthq. Res. Inst. Tokyo Univ.*, 43, 237-239.
- Antonioli, A., M.E. Belardinelli and M. Cocco (2004). Modeling dynamic stress changes caused by an extended rupture in an elastic stratified half space, *Geophys. J. Int.*, 157, 229-244.
- Bono, A. and L. Badiali (2005). Pwl personal wavelab 1.0, an object-oriented workbench for seismogram analysis on windows system, *Computers Geosci.*, 31, 55-64.
- Boschi, E., E. Guidoboni, G. Ferrari, D. Mariotti, G. Valensise and P. Gasperini (2000). Catalogue of Strong Italian Earthquakes from 461 A.C. to 1997, *Annali di Geofisica*, 43 (4), 609-868 (with database on CD-ROM).
- Cao, A.M. and S.S. Gao (2002). Temporal variation of seismic b -values beneath northeastern Japan island arc, *Geophys. Res. Lett.*, 29 (9), 1334; doi: 10.1029/2001GL013775.
- Castello, B., G. Selvaggi, C. Chiarabba and A. Amato (2006). CSI Catalogo della sismicit  italiana 1981-2002, versione 1.1. INGV-CNT, Roma; <http://csi.rm.ingv.it/>.
- Console, R., M. Murru and A.M. Lombardi (2003). Refining earthquake clustering models, *J. Geophys. Res.*, 108, 2468; doi: 10.1029/2002JB002130.
- Daley, D.J. and D. Vere-Jones (2003). *An Introduction to the Theory of Point Processes*, New York, 2nd ed., v. 1, 469 pp.
- Faenza, L., W. Marzocchi and E. Boschi (2003). A nonparametric hazard model to characterize the spatio-temporal occurrence of large earthquakes: An application to the Italian catalog, *Geophys. J. Int.*, 155, 521-531.
- Felzer, K.R., T.W. Becker, R.E. Abercrombie, G. Ekstrom and J.R. Rice (2002). Triggering of the 1999 M_w 7.1 Hector Mine earthquake by aftershocks of the 1992 M_w 7.3 Landers earthquake, *J. Geophys. Res.*, 107 (B9), 2190; doi: 10.1029/2001JB000911.
- Field, E.H. (2007). Overview of the working group for the development of regional earthquake likelihood models (reIm), *Seismol. Res. Lett.*, 78, 7-16.
- Gerstenberger, M.C., S. Wiemer, L.M. Jones and P.A. Reasenber (2005). Real-time forecasts of tomorrow's earthquakes in California, *Nature* 435, 328-331; doi: 10.1038/nature03622.
- Gibbons, J.D. and S. Chakraborti (2003). *Non-parametric Statistical Inference*, 4th ed., New York, 645 pp.
- Gutenberg, B. and C.F. Richter (1954). *Seismicity of the Earth and Associated Phenomena*, Princeton, 273 pp.
- Hainzl, S. and Y. Ogata (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling, *J. Geophys. Res.*, 110, B05S07; doi: 10.1029/2004JB003247.
- Helmstetter, A. and D. Sornette (2003). Foreshocks explained by cascades of triggered seismicity, *J. Geophys. Res.*, 108, 2237; doi: 10.1029/2001JB001580.
- Helmstetter, A., Y.Y. Kagan and D.D. Jackson (2004). Importance of small earthquakes for stress transfers and earthquake triggering, *J. Geophys. Res.*, 110, B05S08; doi: 10.1029/2004JB003286.
- Helmstetter, A., Y.Y. Kagan and D.D. Jackson (2006). Comparison of short-term and time-independent earthquake forecast models for Southern California, *Bull. Seism. Soc. Am.*, 96, 90-106; doi: 10.1785/0120050067.
- Hill, D.P., P.A. Reseanber, A. Michael, W.J. Arabaz, G. Beroza, D. Brumbaugh, J.N. Brune, R. Castro, S. Davis, D. Depolo, W.L. Ellsworth, J. Gomberg, S. Harmsen, L. House, S.M. Jackson, M.J.S. Johnston, L. Jones, R. Keller, S. Malone, L. Munguia, S. Nava, J.C. Pechmann, A. Sanford, R.W.

- Simpson, R.B. Smith, M. Stark, M. Stickney, A. Vidal, S. Walter, V. Wong and J. Zollweg (1993). Seismicity Remotely Triggered by the Magnitude 7.3 Landers, California, Earthquake, *Science*, 260, 1617-1623.
- Kagan, Y.Y. and L. Knopoff (1981). Stochastic synthesis of earthquake catalogs, *J. Geophys. Res.*, 86, 2853-2862.
- Kagan, Y.Y. (1991). Likelihood analysis of earthquake catalogs, *Geophys. J. Int.*, 106, 135-148.
- Kagan, Y.Y. and D.D. Jackson (1995). New seismic gap hypothesis: Five years after, *J. Geophys. Res.*, 100, 3943-3959.
- Kagan, Y.Y. and D.D. Jackson (2000). Probabilistic forecasting of earthquakes, *Geophys. J. Int.*, 143, 438-453.
- Lombardi, A.M., W. Marzocchi and J. Selva (2006). Exploring the evolution of a volcanic seismic swarm: the case of the 2000 Izu Islands swarm, *Geophys. Res. Lett.*, 33, L07310; doi: 10.1029/2005GL025157.
- Lombardi, A.M. and W. Marzocchi (2007). Evidence of clustering and nonstationarity in the time distribution of large worldwide earthquakes, *J. Geophys. Res.*, 112, B02303; doi:10.1029/2006JB004568.
- Lombardi A.M., W. Marzocchi (2010). The assumption of Poisson seismic rate variability in CSEP/RELM experiments, *Bull. Seismol. Soc. Am.*, in press.
- Lombardi, A.M., M. Cocco and W. Marzocchi (2010). On the increase of background seismicity rate during the 1997-1998 Umbria-Marche (central Italy) sequence: apparent variation or fluid-driven triggering?, *Bull. Seismol. Soc. Am.*, 100, 1138-1152.
- Luen, B. and P.B. Stark (2008). Testing Earthquake Predictions, In: D. Nolan and T. Speed (eds.), *Probability and Statistics: Essays in Honor of David A. Freedman*, Beachwood (Ohio, USA), Institute of Mathematical Statistics, 302-315; doi: 10.1214/193940307000000509.
- Marzocchi, W. and L. Sandri (2003). A review and new insights on the estimation of the b -value and its uncertainty, *Annals of Geophysics*, 46 (6), 1271-1282.
- Marzocchi, W. and A.M. Lombardi (2008). A double branching model for earthquake occurrence, *J. Geophys. Res.*, 113, B08317; doi: 10.1029/2007JB005472.
- Marzocchi, W. and A.M. Lombardi (2009). Real-time forecasting following a damaging earthquake, *Geophys. Res. Lett.*, 36, L21302; doi: 10.1029/2009GL040233.
- Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes, *J. Amer. Statist. Assoc.*, 83, 9-27.
- Ogata, Y. (1998). Space-Time Point-Process Models for Earthquake Occurrences, *Ann. Inst. Statist. Math.*, 50 (2), 379-402.
- Ogata, Y. and J. Zhuang (2006). Space-time ETAS models and an improved extension, *Tectonophysics*, 413, 13-23.
- Omori, F. (1894). On the aftershocks of earthquakes, *J. Coll. Sci. Imp. Univ. Tokyo*, 7, 111-120.
- Reasenber, P.A. and L.M. Jones (1989): Earthquake hazard after a mainshock in California, *Science*, 243, 1173-1176.
- Reasenber, P.A. and L.M. Jones (1994). Earthquake aftershocks: Update, *Science*, 265, 1251-1252.
- Rhoades, D.A. and F.F. Evison (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pure Appl. Geophys.*, 161, 47-72.
- Schorlemmer D., M.C. Gerstenberger, S. Wiemer, D.D. Jackson and D.A. Rhoades (2007). Earthquake Likelihood Model Testing, *Seism. Res. Lett.*, 78, 17-29.
- Schorlemmer, D., F. Mele and W. Marzocchi (2010). A Completeness Analysis of the National Seismic Network of Italy, *J. Geophys. Res.*, 115, B04308; doi: 10.1029/2008JB006097.
- Shi, Y. and B.A. Bolt (1982). The standard error of the Magnitude-frequency b value, *Bull. Seism. Soc. Am.*, 72, 1677-1687.
- Utsu, T. (1961). A statistical study on the occurrence of aftershocks, *Geophys. Mag.*, 30, 521-605.
- Vere-Jones, D. (2006). The development of statistical seismology: my personal experience, *Tectonophysics*, 413, 5-12.
- Wiemer, S. (2001). A software package to analyze seismicity: ZMAP, *Seism. Res. Lett.*, 72, 373-382.
- Wilks, D.S. (2005). *Statistical Methods in the Atmospheric Sciences*, 2nd ed., New York, 627 pp.
- Wössner, J. and S. Wiemer (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty, *Bull. Seism. Soc. Am.*, 95, 684-698.
- Zechar, J.D., D. Schorlemmer, M. Liukis, J. Yu, F. Euchner, P.J. Maechling and T.H. Jordan (2009). The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science, *Concurrency and Computation: Practice and Experience*; doi: 10.1002/cpe.1519.
- Zhuang J., Y. Ogata and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrence, *J. Am. Stat. Assoc.*, 97, 369-380.

*Corresponding author: Anna Maria Lombardi,
Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy;
e-mail: annamaria.lombardi@ingv.it