THE ETAS MODEL FOR DAILY FORECASTING OF
ITALIAN SEISMICITY IN CSEP EXPERIMENT

Anna Maria Lombardi and Warner Marzocchi

Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata 605, 00143 Rome, Italy.

Revised version submitted to Annals of Geophysics

2010
Abstract

This paper investigates the basic properties of the recent shallow seismicity in Italy through stochastic modeling and statistical methods. Assuming that the earthquakes are the realization of a stochastic point process, we model the occurrence rate density in space, time and magnitude by means of an Epidemic Type Aftershock Sequence (ETAS) model. By applying the maximum likelihood procedure, we estimates the parameters of the model that best fit the Italian instrumental catalog, recorded by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) from April 16th 2005 to June 1st 2009. Then we apply the estimated model on a second independent dataset (June 1st 2009- Sep 1st 2009). We find that the model performs well on this second database, by using proper 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 (CSEP). Specifically we have submitted this model for the daily forecasting of Italian seismicity above Ml4.0.
1. Introduction

There is a growing consensus to accept the existence of an intrinsic stochasticity of the earthquake generating process (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 acting on different spatio-temporal scales (Kagan & Knopoff, 1981; Kagan & Jackson, 2000; Ogata, 1988; 1998; Helmstetter et al., 2006; Faenza et al., 2003; Rhoades & Evison, 2004; Gerstenberger et al., 2005; Marzocchi & Lombardi, 2008; Lombardi et al., 2006; 2007; 2010). Each model describes one or more different coexisting physical processes (tectonic loading, coseismic stress interactions, postseismic deformation, aseismic processes, and so on), which have more or less relevance for earthquake occurrence, depending on maturity in the seismic cycle. Here, we focus our attention on daily forecasts. For this class of forecasts, stochastic models describing the phenomenon of earthquake clustering are becoming widely accepted in the seismological community (e.g., Reasenberg & Jones, 1989, 1994; Gerstenberger et al., 2005; Marzocchi & Lombardi, 2009).

Specifically we describe a short-term earthquake forecasting model that we have submitted to the EU-Italy Collaboratory Studies for Earthquake Predictability (CSEP) experiment. The forecast method uses earthquake data only, with no explicit use of tectonic, geologic, 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, 1998; Console et al., 2003; Helmstetter et al., 2006; Lombardi & Marzocchi, 2007). We apply a version of the ETAS model to seismicity recorded in Italy in recent years. For a first retrospective test, we apply a well-know procedure that consists in fitting the model to the early part of the Italian earthquake catalog and then testing it on the most recent part of the data set. The real time forecasting performance of the model has been successfully checked on the occasion of the recent L’Aquila earthquake (Central Italy; April 6th 2009, Mw 6.3; see Marzocchi & Lombardi, 2009).

2. The Spatio-Temporal Epidemic Type Aftershock Sequences (ETAS) Model

single modified Omori function (Omori, 1894; Utsu, 1961) and that seismicity can include conspicuous secondary aftershock production. 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. Since the first time-magnitude formulation proposed by Ogata (1988), many others time-magnitude-space versions have been published in the literature, mostly based on empirical studies of past seismicity (Ogata, 1998; Zhuang et al., 2002; Console et al., 2003; Helmstetter et al., 2006; Lombardi & Marzocchi, 2007]. These approaches describe 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 seismicity not triggered by previous events in the catalog; the second is associated with stress perturbations caused by previous earthquakes of 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 past history) by equation:

$$\lambda(t,x,y,m/H_{t}) = \nu(x,y) + \sum_{i \in \mathcal{H}_{t}} \frac{K}{(t-t_{i}+c)^{p}} e^{\alpha(t-M_{i})} \frac{c_{dq}}{[r_{i}^{2} + d^{2}]^{q}} \beta e^{-\beta(m-M_{i})}$$

where $\mathcal{H}_{t} = \{(t_{i},x_{i},y_{i},M_{i}); t_{i} < t\}$ is the observation history up the time $t$; $M_{c}$ is the completeness magnitude of the catalog; $\nu$ is the rate of background seismicity for the whole area; $K$, $c$ and $p$ are the parameters of the modified Omori Law describing the decay in time of short-term triggering effects; $\alpha$ 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 triggered events and $c_{dq} = \frac{q-1}{\pi} [d^{2(q-1)}]$ is the relative normalization constant; $r_{i}$ is the distance between location $(x,y)$ and the epicenter of $i$-th event $(x_{i},y_{i})$; the function $u(x,y)$ is the spatial PDF of background events; finally, $\beta = b \ln(10)$ is the parameter of the well-known Gutenberg-Richer Law (Gutenberg & Richter, 1954), that is assumed to hold for all magnitudes and invariant in space. Specifically, the model assumes that large earthquakes are indistinguishable from the smaller ones, and therefore they have the same distribution.

The most recent versions of the ETAS model (Ogata & Zhuang, 2006; Helmstetter et al., 2006) are characterized by the introduction of a further term that takes into account the correlation between the aftershock area and the magnitude of triggered events. Some preliminary results show that this
correlation may be negligible for Italy (see Marzocchi & Lombardi, 2009). So we decide to use the version of ETAS model described by eq. (1), in which the spatial decay of triggered activity is independent of the magnitude of the triggering shock. A deeper analysis on this topic will be presented and discussed in future works.

The parameters \((v, K, c, p, \alpha, d, q, \beta)\) of the model, for the events within a time interval \([T_{\text{start}}, T_{\text{end}}]\) and a region \(R\) can be estimated by maximizing the Log-Likelihood function (Daley & Vere-Jones, 2003), given by

\[
\log L(v, K, c, p, \alpha, d, q, \beta) = \sum_{i=1}^{N} \log \lambda(t_i, x_i, y_i, m_i) - \int_{T_{\text{start}}}^{T_{\text{end}}} \int_{R} \int_{M_{\text{max}}} \lambda(t, x, y, m) \, dt \, dx \, dy \, dm \tag{2}
\]

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

\[
u u(x, y) = \frac{1}{T} \sum_j p_j K_d(r_j) \tag{3}
\]

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 \(K_d\) is a Gaussian kernel function with a spatially variable bandwidth. Similarly the rate of triggered events is given by

\[
c(x, y) = \frac{1}{T} \sum_j (1 - p_j) K_d(r_j) \tag{4}
\]

Several physical investigations show 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 impose \(q=1.5\). This choice is also justified by the trade-off between parameters \(q\) and \(d\) that may cause different pairs of \(q\) and \(d\) values to provide almost the same likelihood of the model (Kagan & Jackson, 2000).

3. Testing the Model

The gold standard for evaluating scientifically earthquake forecasting models is through the comparison of forecasts and true value in prospective experiments (see, e.g., Field, 2007; Schorlemmer et al., 2007; Luen & Stark, 2008; Zechar et al., 2009). Nevertheless, it may be conceivable to evaluate the model also through retrospective experiments, for instance, dividing the available dataset in two parts: a first part of dataset, hereinafter \textit{learning} dataset, can be used to set up the model and a second, the \textit{testing} dataset, to check its reliability (Kagan & Jackson, 2000).
verification of forecasting capability of the model can be achieved by a comparison of observations and forecasts. Such a testing enables us to verify if the model is significantly good performing, and, eventually, to identify the features allowing a better forecasting. In successive subsections we describe the statistical tests used in the present study to check our model retrospectively.

3.1 Residuals Analysis

A common diagnostic technique for stochastic point processes is based on transformation of the time axis $t$ into a new scale $\tau$ by the increasing function

$$\tau = \Lambda(t) = \int_{t_{\text{start}}}^{t} dt' \int dx dy \int_{M_c}^{M_{\text{max}}} dm \lambda(t', x, y, m/\mathcal{H}_e, \cdot) = \int_{t_{\text{start}}}^{t} \nu + \sum_{o \in \mathcal{T}} \frac{ke^{\nu(t'-t_c+c)}}{t'-t_c+c} \ dt'$$

where $T_{\text{start}}$ is the starting time of the observation history $H_t$ (Ogata, 1988). The random variable $\tau$ represents the expected number of occurrences in time period $[T_{\text{start}}, t]$, into whole region R and with magnitude above $M_c$. If a model with conditional intensity $\lambda(t, x, y, m/\mathcal{H})$ describes well the temporal evolution of the process, the transformed data $\tau_i = \Lambda(t_i)$, known in statistical seismology with the name of residuals, are expected to behave like a stationary Poisson process with the unit rate (Ogata, 1988). Therefore the values $\Delta \tau_i = \tau_{i+1} - \tau_i$ are independent and exponentially distributed (with mean equal to 1) random variables. We check this hypothesis for residual 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 & Chakraborti, 2003; Lombardi & Marzocchi, 2007). We use both tests because the KS1 test is ineffective to check the presence of a memory in the time series. Hence, any discrepancy of residuals by Poisson hypothesis, identified by just one or both tests, is a sign of inadequacy of ETAS model to explain all basic features of analyzed seismicity. This check analysis is similar to the $N$-test, currently used by RELM/CSEP testing centers (Kagan & Jackson, 1995; Schorlemmer et al., 2007), but it avoids the time binning that may lead to biases in the results of the testing phase (see, e.g., Lombardi & Marzocchi, 2010).

3.2 Cumulative Reliability Diagram

The reliability diagram is a common diagnostic technique used to measure the consistency of a forecast model with the observations. Roughly speaking, a probability forecast is reliable if the event actually happens with an observed frequency that is consistent with the forecast. More specifically, a reliability diagram consists of a plot of observed relative frequencies against predicted probabilities (Wilks, 2005). Reliability measures sort the forecast/observations pairs ($F_j/O_i$) into groups, according to the value of forecast variable, and characterize the conditional
distributions of the observations given the forecasts. In particular a way to identify visually 193
departures from reliability is to plot the cumulative conditional observed frequency \( p(O_i|F_j) \) against 194
the cumulative predicted probability \( F_j \); this gives a Cumulative Reliability Diagram (CRD). The 195
perfect reliability is represented by the diagonal line.

We use this type of analysis to check the predicted spatial distribution on observed 198
seismicity. Specifically we apply a case of dichotomous events, i.e. observations are limited to 2 199
possible outcomes, the occurrence \((O_1)\) or nonoccurrence \((O_2)\) of an earthquake. To define the 200
forecasting cumulative probabilities \( F_j \), the area under analysis is partitioned in a non-overlapping 201
and exhaustive set of cells \( C_i \); for each cell we compute the proportion of events \( f_i \) expected by the 202
forecasting model. These values \( f_i \), by definition between 0 and 1, are sorted in ascending order and 203
are grouped into \( N \) bins \( B_j \) \((j=1...N)\), that form a partition of the unit interval composed by 204
overlapping increasing subintervals. These bins are characterized by a set of forecasting 205
probabilities \( F_j \) that define the probability to have at least one event in \( B_j \)

\[
I_j = \{i; f_i \in B_j\} \quad \sum_{i \in I_j} f_i \leq F_j
\]  

(6)

The most intuitive choice is to take \( F_j \) equally spaced. If the distribution of the forecasts is non-208
uniform, then choosing the bins so that the sets \( I_j \) are equally populated (i.e. with the same number 209
of events \( f_i \)) can be a good alternative. The values \( F_j \) are compared with the cumulative observed 210
frequencies

\[
P(O_1|F_j) = \sum_{i \in I_j} \frac{N_i}{N}
\]  

(7)

where \( N_i \) is the observed number of shocks into the cell \( C_i \) and \( N \) is the total number of events. In the 213
case of perfect reliability the conditional probability \( p(O_1|F_j) \) is equal to \( F_j \).

4. The INGV Database

Italy is characterized by a generally high seismicity, with observed magnitudes up to about 217
7.5. The long tradition of seismological studies in Italy produced many efforts for seismic data 218
collection, therefore today Italy can boast of careful seismic instrumental catalogs (Castello et al., 219
2005; Schorlemmer et al. 2010; http://iside.rm.ingv.it/), besides of a tested experience in compiling 220
historical databases (Boschi et al., 2000). The most complete instrumental catalog of italian 221
seismicity is the seismic bulletin of Istituto Nazionale di Geofisica e Vulcanologia (INGV) 222
(http://iside.rm.ingv.it). The Italian seismic network changed significantly in the last years. 223
Specifically the 16 April 2005 marks the date of remarkable changes of the seismic Italian network 224
(Bono & Badiali, 2005; see also Schorlemmer et al., 2010) and of data processing. Given the large
difference of INGV bulletin before and after this date, we decide to set up our model on parameters of events collected from April 16th 2005 to June 1st 2009. The earthquakes from June 1st 2009 and Sep 1st 2009 are instead used for a first retrospective test of the model (testing dataset). In agreement with CSEP requirements, we select events above 30 kms of depth occurred in the collection area, as defined by CSEP experiment.

A correct understanding of the physical processes controlling the rate of earthquake production depends on the quality of the available seismic catalog. Specifically, a critical issue that has to be addressed before performing any investigation is the assessment of completeness of dataset. Here we verify the completeness magnitude \( M_c \) (lowest magnitude at which a negligible number of the events are not detected) and its variations with time. The algorithms are freely available together with the software package ZMAP (Wiemer, 2001). The analysis of whole catalog by Maximum Likelihood method (Shi & Bolt, 1982) provides a value of \( M_c \) (local magnitude) equal to 2.0 (see Figure 1a). The analysis of the spatio-temporal variation of completeness magnitude shows clear changes of \( M_c \) with time (see Figure 1b) and space (see Figure 1c). We perform these analyses by using a minimum number of events equal to 100 and a radius equal to 50 km. In particular, \( M_c \) reaches about 2.5 soon after the occurrence of recent L’Aquila earthquake (April 6th 2009, Mw6.3; see Figure 2b). This value seems to be a reliable completeness threshold for most part of national territory (see Figure 2c). These results are also in agreement with Schorlemmer et al. (2010) which identify \( M_c=2.5 \) as a reasonable magnitude threshold for most of Italian territory. The only exception is for the southern part of Apulia and the western part of Sicily, showing a higher completeness magnitude (see also Schorlemmer et al., 2010 for details).

Considering the small size of these areas, we decide to select for the present study the events above magnitude 2.5 recorded into the INGV bulletin (2100 events for learning and 179 for testing databases). Figure 2 shows the distribution of selected seismic activity for both learning (Figure 2a) and testing (Figure 2b) databases, together with the boundaries of collection area defined by CSEP laboratory.

5. Application and testing of the ETAS model on Italian seismicity

We apply the ETAS model to Italian seismicity recorded into learning database, described in previous section. Following the procedure proposed by Zhuang et al. (2002) we estimate the model parameters together with the spatial distribution of background seismicity \( u(x,y) \). Table 1 lists the inferred values of model parameters together with their standard errors and the associated log-likelihood values. The total percentages of triggered and spontaneous events identified by the model
are 46% and 54% respectively. In Figure 3 we show two maps: the first represents the distribution of the time-independent background rate ($\nu(x,y)$, see eq. (3)), the second the distribution of the clustering ratio $r(x,y)$, i.e. the ratio between triggered and total rates, for the whole learning period. The clustering ratio is obtained by the formula

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

(8)

where $c(x,y)$ and $\lambda(t,x,y,m/H_t)$ are defined by eq. (1) and (4), respectively. By comparing the two maps shown in Figure 3, we find that the spatial distribution of triggering capability is not a proxy for the seismogenetic potential. For example, the southern part of peninsular Italy shows a lower triggering rate respect to other zones (see Figure 3b), whenever this area is one of most active of whole region (see Figures 2 and 3a). The estimated Omori Law decay predicts that the probability of triggering one or more events with magnitude above 2.5 for an earthquake of magnitude 3.0 is below 1% after about 5-6 hours. The corresponding times for a triggering event of magnitude 5.0 and 7.0 are 2-3 days and about 1 month, respectively (see Figure 4a). We stress that these probabilities refer to direct triggering effects. The secondary triggered events are not included in this calculation. As regards the spatial decay of the triggering capability, an event has a 50% of probability to trigger one or more events within 2km from its epicenter and about 40% at a distance larger than 10km, regardless its magnitude (see Figure 4b).

A preliminary check on the goodness of the inferred ETAS model is done by applying the residual analysis on the learning dataset used to set-up the ETAS model. We find that the residuals pass the KS1 test (p-value 0.8), but the Runs test rejects the hypothesis of no-correlation (p-value 0.007). The cumulative distribution of residuals (Figure 5a) shows a clear deviation from the expected Poisson behavior soon after the occurrence of $M_w$ 6.3 L’Aquila earthquake (April 6 2009). If we take out the L’Aquila sequence by the learning period, the ETAS model passes the Runs test (p-value 0.07). We argue that this result is probably due to the spatial variation of some parameters. In other words, at local scale the model could be significantly different with respect to the same model calibrated using the whole Italian territory. For example, Marzocchi & Lombardi (2009) reported an $\alpha$-value of 1.5 for the L’Aquila region that increases to 2.0 when $M_c = 2.5$ is considered; this value is certainly larger than the 1.3 found here for the whole Italian territory (see table 1).

In order to test the forecasting performance of the ETAS model, we analyze the residuals and plot the cumulative reliability diagram on testing dataset. By using the KS1 test we cannot reject the null
The hypothesis that values $\Delta \tau_i = \tau_{i+1} - \tau_i$ are exponentially distributed (with mean equal to 1) (the p-value is equal to 0.14). The Figure 5b show the cumulative number of residuals $\tau_i$ versus transformed time $\tau$ (solid line) together with the expected linear scaling predicted by a Poisson distribution (that is, the cumulative number of residuals should lie along the bisector). Similarly, the Runs test does not reject the independence hypothesis of $\Delta \tau_i$ (the p-value is equal to 0.81), implying that the hypothesis of uncorrelation of residuals cannot be rejected. This result is corroborated by Figure 5c, in which we plot 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 exponential random variables with unit mean, the statistics $U_k$ are i.i.d. uniform random variables on $[0,1)$. Assuming that a possible correlation is likely to show up in neighboring intervals, the plot of $U_{k+1}$ versus $U_k$ should recover uniformly the figure panel (Ogata, 1988).

The cumulative reliability diagram of spatial distribution on events collected by testing dataset shows a reliable forecasting (see Figure 6). To define the forecasting probabilities $F_j$ we compute the expected fraction of events $f_i$ by ETAS model, for each cell $C_i$ of the testing grid defined by CSEP laboratory. The values $f_i$ are computed as the ratio between the expected numbers of events in the cell $C_i$ and in whole region $R$. Specifically we use the formula

$$f_i = \frac{\int \int \int_{T, \mathcal{C}, \mathcal{M}} \lambda(t, x, y, m / \mathcal{H}_t) dt dx dy dm}{\int \int \int_{T, R, \mathcal{M}} \lambda(t, x, y, m / \mathcal{H}_t) dt dx dy dm}$$

(8)

where $T$ is the testing period, $R$ is the testing area defined by CSEP laboratory, $M$ is the magnitude range $[2.5, 9.0]$, and $\mathcal{H}_t$ is the occurrence history, starting by April 16 2005 (i.e. including the learning period). Then we regroup these values in 10 bins $B_j$, identified by increasing values of probabilities $F_j$. The error bars are defined so that the sets $I_j$ (see eq. 6) are equally populated. In Table 2 we report the values of probabilities $F_j$ and $p(O_i|F_j)$ (i.e. the observed frequencies of events in bin $B_j$), as defined in eq. (7). They are plotted in Figure 6. The error bars indicate the 95% confidence interval of values $p(O_i|F_j)$. These last are obtained by applying the reliability analysis on 1000 synthetic catalogs. These have the same duration of testing period of INGV bulletin and are simulated in agreement with ETAS model, including the real learning period into the past history. The reliability diagram shows that the pairs $[F_j, p(O_i|F_j)]$ are well fitted by diagonal that indicates a perfect reliability. Moreover they are in agreement with variation expected by the model. All these results show that the model estimated on learning dataset is in agreement with the following
seismicity. This result is also corroborated by the observation that the parameters estimated from the entire catalog are not statistically different by parameters listed in Table 1.

The model formulated and tested above allows us to compute forecasts in the framework of CSEP experiment. Predictions are in a form of daily probability of occurrence for at least one earthquake with $M_l \geq 4.0$, within a cell of 0.1°x0.1°, in Italy. These are obtained by integrating for each cell $C_i$ and for each forecasting period $T_j$ the intensity function of ETAS model (eq. (1)). The forecast rates above $M_l 4.0$ are obtained by rescaling the rate of earthquakes above $M_l 2.5$, in agreement with the Gutenberg-Richter relation. The eq. (1) shows that a time-dependent modeling as the ETAS model imposes to take into account also the triggering effect of seismicity occurred before and expected during the forecast interval. So we include in the past history all real seismicity with magnitude above $M_l 2.5$ and depth above 30 km, occurred up to the starting time of the forecasting time window. Moreover we simulate 1000 different stochastic realizations for the forecasting time window, by using the thinning method proposed by Ogata (1998) and the intensity function formulated in equation (1). Then we average predictions coming from each of these synthetic catalogs.

6. Discussion and Conclusions

In this paper we have adopted a version of ETAS model to describe the recent shallow seismicity occurred in Italy. The main motivation of this study was to submit our model to EU-Italy CSEP laboratory for 1-day forecasts. To achieve this goal we have proposed a model representing the main average properties of Italian seismicity. The reliability of this model has been successfully checked, at local scale, in a real-time forecasting experiment, on occasion of the occurrence of recent L’Aquila destructive earthquake (Marzocchi & Lombardi, 2009).

One finding of the present paper is that the generalization of local models to the whole Italian territory may be problematic for different reasons. First, the completeness magnitude varies with space (Schorlemmer et al., 2010); in this paper we have adopted $M_c=2.5$ that is probably optimistic for some zones. In fact, the $M_c$ for the whole territory is about 2.9 (see Figure 1c and Schorlemmer et al., 2010). We are conscious of this limit, but we preferred to adopt a value of $M_c$ that is reliable for most (not all) of Italian territory. The area with $M_c>2.5$ covers only a very small part of the whole region. The use of a larger completeness magnitude causes a strong reduction of dataset with a consequent increase of uncertainty of the model. Maybe more important, it has been recognized that smaller earthquakes have a decisive role in the triggering process (Helmstetter, 2003; Felzer et
al., 2002; Helmstetter et al., 2004); therefore, a too high value of $M_c$ might cause an erroneous identification of the triggered part of seismicity.

Second, some of the ETAS parameters may vary with space. This means that some parameters estimated for the whole territory and for a small region may be significantly different. Local variations may occur only as consequence of the occurrence of large earthquakes. For example, the model proposed here for the whole Italian territory is not able to reproduce correctly the time evolution of the first part of 2009 L’Aquila sequence (see Figure 3a). As anticipated before, we argue that this discrepancy is probably due to features of the local seismicity that cannot be extrapolated for the whole territory. In particular the seismicity of L'Aquila is characterized by a larger $\alpha$-value with respect to the whole Italian seismicity described by our ETAS model. The $\alpha$ parameter is crucial to quantify the dependence of triggering effect by magnitude of parent earthquake. The failure of the model to describe the starting phase of L'Aquila sequence suggests that possible inconsistencies could occur in forecasting future seismicity. This problem may call for the development of more complicated models that take into account local features of seismic activity.

We argue that other parts of the model could be improved in the future. In the following, we report only some possible hints in this direction. First, the model could be enhanced by adopting a modified magnitude distribution, to explicitly allow for the decrease of detection soon after a large earthquake (Kagan 1991, Helmstetter et al., 2006; Lennartz et al., 1998). Second, the background rate and the basic clustering proprieties of aftershocks sequences are assumed to be stationary in time. Such an assumption is mostly motivated by the short learning dataset adopted. Longer datasets may permit to capture departures from stationarity such as long-term time evolution of the seismicity (e.g., Lombardi & Marzocchi, 2007; Marzocchi & Lombardi, 2008). Moreover, other time-dependent processes acting on short time scales, like fluid injection, may have a significant impact on short-term spatio-temporal evolution of seismicity and therefore it may be necessary to include them into the ETAS model (Ogata & Hainzl 2005; Lombardi et al., 2006; 2010). Third, the ETAS model proposed here assumes that all earthquakes are equal. Possible distinctive precursory activity that anticipates large shocks is not considered in this parametrization. Finally, the present model does not incorporate tectonic/geologic information. Their inclusion may represent one possible future direction of investigation to improve the forecasting of large shocks. For example, the Gutenberg-Richter law is used everywhere indistinctively; this means that a magnitude 8 is considered possible everywhere. It is argued that geological information may provide in the future a more appropriate
frequency-magnitude law that varies in space.

References


Table Captions

Table 1: Maximum Likelihood parameters (with relative errors) and log-likelihood of ETAS model for the learning INGV bulletin (M_c = 2.5; Apr 16 2005 – Jun 1 2009; 2100 events).

Table 2: Cumulative Reliability Diagram of spatial distribution of earthquakes predicted by ETAS model relative to the testing INGV bulletin (M_c = 2.5; Jun 1 2009 – Sep 1 2009; 179 events). The values F_j and p(O_i|F_j) indicate the forecasts and the observed frequencies, respectively.
Figure Captions

Figure 1: Completeness magnitude of INGV bulletin (from April 16th 2005 up to June 1st 2009) obtained by the Maximum Likelihood Method (MLM). a) Frequency magnitude distribution for the whole dataset: the MLM provides $M_c=2.0$; b) $M_c$ as a function of time; c) $M_c$ as a function of space.

Figure 2: Map of seismic events with magnitude above 2.5 and depth smaller than 30 km that occurred in Italy inside the collection area identified by the CSEP experiment (blue solid line; see Schorlemmer et al., 2010b). The symbol sizes are scaled with magnitude. a) Map of events of the learning dataset (April 16th 2005-June 1st 2009; 2100 events) used to set-up the model; b) map of the testing dataset (June 1st 2009- Sep 1st 2009; 179 events) used for a retrospective forecasting test of the model.

Figure 3: Maps of a) the background seismicity rate $nu(x,y)$, and b) the ratio between the triggered rate and the total seismic rate of the INGV bulletin learning dataset (April 16th 2005-June 1st 2009; 2100 events).

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

Figure 5: Residuals Analysis of the ETAS model on the learning (April 16th 2005-June 1st 2009; 2100 events) and testing INGV bulletin (June 1st 2009- Sep 1st 2009; 179 events). a) Cumulative number of transformed times $\tau_i$ (solid line) for the learning period together with the theoretical distribution (dotted line) predicted by a Poisson distribution. b) The same as a), but for the testing period. c) Plot of values $U_{k+1}=1-exp(\tau_{k+1}-\tau_k)$ versus $U_k$ for the testing period.

Figure 6: Cumulative Reliability Diagram of the spatial earthquake distribution predicted by ETAS model for the testing INGV bulletin ($M_c = 2.5$; Jun 1 2009 – Sep 1 2009; 179 events). Stars mark the pairs $F_j / p(O_j|F_j)$, i.e., the forecasts and the observed spatial distributions. The dotted black line
represents the perfect reliability. Error bars identify the 95% confidence interval of the observed values $p(O_i|F_j)$. The forecast probabilities $F_j$ identify equally populated bins $B_j$ (see text for details).
Table 1: Parameters of ETAS model for Italian seismicity

(Mₑ = 2.5; Apr 16 2005 – Jun 1 2009; 2100 events)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ν</td>
<td>237 ± 8 (year⁻¹)</td>
</tr>
<tr>
<td>K</td>
<td>0.011 ± 0.001 (year⁻¹)</td>
</tr>
<tr>
<td>p</td>
<td>1.16 ± 0.02</td>
</tr>
<tr>
<td>c</td>
<td>0.00004 ± 0.00001 (year)</td>
</tr>
<tr>
<td>α</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>d</td>
<td>1.10 ± 0.05 (km)</td>
</tr>
<tr>
<td>q</td>
<td>≡ 1.5</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-7808.1</td>
</tr>
</tbody>
</table>
### Table 2: Values of Cumulative Reliability Diagram

| $F_j$         | $p(O_j|F_j)$ |
|---------------|-------------|
| $1.6 \times 10^{-3}$ | $1.8 \times 10^{-3}$ |
| $5.6 \times 10^{-3}$ | $7.5 \times 10^{-3}$ |
| $1.2 \times 10^{-2}$ | $1.4 \times 10^{-2}$ |
| $2.2 \times 10^{-2}$ | $1.9 \times 10^{-2}$ |
| $3.5 \times 10^{-2}$ | $3.2 \times 10^{-2}$ |
| $5.2 \times 10^{-2}$ | $5.6 \times 10^{-2}$ |
| $7.6 \times 10^{-2}$ | $8.6 \times 10^{-2}$ |
| $1.1 \times 10^{-1}$ | $1.3 \times 10^{-1}$ |
| $1.7 \times 10^{-1}$ | $2.2 \times 10^{-1}$ |
| $1.0$         | $1.0$       |
Maximum Likelihood Solution

b?value = 1.02 +/- 0.01,  a value = 5.9,  a value (annual) = 5.28
Magnitude of Completeness = 2
Figure 1b
Figure 1c
Figure 2a
Figure 2b
Figure 4a
Figure 5
Figure 6