

Short-term earthquake forecasting experiment before and during the L'Aquila (central Italy) seismic sequence of April 2009

Maura Murru¹, Jiancang Zhuang^{2,*}, Rodolfo Console^{1,3}, Giuseppe Falcone¹

¹ Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

² The Institute of Statistical Mathematics, Tokyo, Japan

³ Centre of Integrated Geomorphology for the Mediterranean Area, Potenza, Italy

Article history

Received May 5, 2014; accepted November 3, 2014.

Subject classification:

L'Aquila earthquake, ETAS model, PPE model, Probability forecast, Point process, Probability gain.

ABSTRACT

In this paper, we compare the forecasting performance of several statistical models, which are used to describe the occurrence process of earthquakes in forecasting the short-term earthquake probabilities during the L'Aquila earthquake sequence in central Italy in 2009. These models include the Proximity to Past Earthquakes (PPE) model and two versions of the Epidemic Type Aftershock Sequence (ETAS) model. We used the information gains corresponding to the Poisson and binomial scores to evaluate the performance of these models. It is shown that both ETAS models work better than the PPE model. However, in comparing the two types of ETAS models, the one with the same fixed exponent coefficient $\alpha = 2.3$ for both the productivity function and the scaling factor in the spatial response function (ETAS I), performs better in forecasting the active aftershock sequence than the model with different exponent coefficients (ETAS II), when the Poisson score is adopted. ETAS II performs better when a lower magnitude threshold of 2.0 and the binomial score are used. The reason is found to be that the catalog does not have an event of similar magnitude to the L'Aquila mainshock (M_w 6.3) in the training period (April 16, 2005 to March 15, 2009), and the α -value is underestimated, thus the forecast seismicity is underestimated when the productivity function is extrapolated to high magnitudes. We also investigate the effect of the inclusion of small events in forecasting larger events. These results suggest that the training catalog used for estimating the model parameters should include earthquakes of magnitudes similar to the mainshock when forecasting seismicity during an aftershock sequence.

1. Introduction

On April 6, 2009 at 01:32 (UTC), a magnitude M_w 6.3 earthquake struck the Abruzzi region in central Italy. Despite its moderate size, the earthquake caused more than 300 fatalities and partially destroyed the city of L'Aquila and many surrounding villages. The epicenter of the mainshock was located a few km WSW of the city. It occurred on one of the NW–SE trending

normal faults that form part of the 800 km long segmented normal fault system that accommodates crustal extension in the Apennine mountain range (e.g., Anderson and Jackson, 1987; Roberts et al., 2002). The mainshock had a hypocenter depth of 8.3 km, within a seismogenic layer that is thought to be between 2 and 10 km deep. It ruptured a northwest-trending, 18 km long fault, with a dip angle of 45° to the southwest. It was preceded by earthquake activity starting at the end of 2008, which occurred in a silent region at the southern end of the northern Apennine extensional belt (Chiarabba et al., 2009). This foreshock sequence was characterized by clustering around the mainshock nucleation area. The largest earthquakes of the foreshock sequence included an M_w 4.0 event, which occurred on March 30, 2009 at 13:38 (UTC), one week before the mainshock, and M_w 3.9 and M_w 3.5 events that occurred on April 5, 2009 at 20:48 and 22:39 (UTC), respectively. Many other destructive earthquakes are recorded in the historical catalog for this area (Gruppo di lavoro CPTI, 2004). The last dramatic event was the M_w 6.7, 1915 Avezzano earthquake, located about 30 km to the southeast of the 2009 L'Aquila earthquake epicenter, which caused about thirty thousand deaths (see Figure 1).

In light of the historical earthquake records, this region of the central Apennines, with one of the highest seismic hazards in Italy (Slejko et al., 1998; Albarello et al., 2001; Rebez et al., 2001; Akinici et al., 2009, and references therein), has been classified as a second category seismic zone in the Italian seismic resistant building code (Gruppo di lavoro MPS, 2004). By the end of November 2009, more than 16,000 aftershocks with $M_L \leq 5.5$ had been recorded by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) seismic network. Initially, this aftershock activity occurred close to L'Aquila

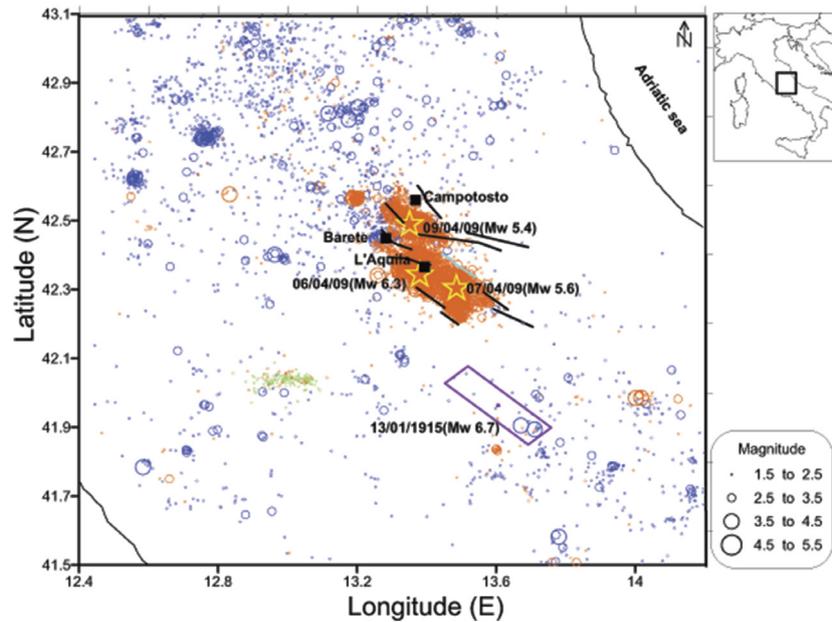


Figure 1. Epicentral distribution of earthquakes during the learning (April 16, 2005 – March 15, 2009) and test (March 16, 2009 – June 30, 2009) phases reported by the INGV data center in the area under analysis, affected by the 2009 L'Aquila earthquake. There were 2,588 ($M_L \geq 1.6$) events (blue circles) during the learning period. Green dots around (42.0°N, 13.0°E) indicate events (151) recognized as quarry blast activity, which were not considered in the analysis. Red circles indicate events with $M_L \geq 2.0$ (3007) analyzed in the testing phase. The L'Aquila sequence occurred in a silent region at the southern end of the northern Apennine extensional belt. Stars indicate the three $M_w \geq 5.4$ events that occurred in April 2009. Solid lines are mapped Quaternary faults. The Paganica Fault is indicated in blue. The 1915 Avezzano earthquake source is also shown in the map DISS Working Group, 2010 (DISS 3.1.1, 2010) as a rectangular box.

and towards the southeast of the town, but a couple of days later it migrated towards the northwest, to the Barete and Campotosto zones (Figure 1). In fact, after the mainshock of April 6, 2009, two other strong events occurred with their own aftershocks. The first one occurred about 40 hours after the mainshock, at 17:47 UTC on April 7, 2009 with a magnitude of M_w 5.6, to the southeast (in the Valle d'Aterno), and the second occurred at 00:53 UTC on April 9, 2009, with a magnitude of M_w 5.4, to the north of the mainshock near Campotosto, about 16 km north of L'Aquila (Figure 1). Large aftershocks can cause great damage to building structures that have already been weakened by the mainshock. Thus, it may be important, in such circumstances, to evaluate the aftershock probability in real time, to avoid losses from aftershocks. The principle of evaluating the probabilities of earthquake occurrence using point process models, which are mathematically formulated with conditional intensity functions, was framed by Vere-Jones (1998). Among different models for seismicity, the Epidemic-Type Aftershock Sequence (ETAS) model, proposed by Ogata (1988), which describes the features of earthquake clustering of foreshocks, main shocks, and aftershocks, has become a standard model for testing hypotheses, and a starting point for short-term earthquake forecasts (Console and Murru, 2001; Console et al., 2003; Murru et al., 2009; Zhuang et al., 2004; Helmstetter and Sor-

nette, 2003; Hainzl and Ogata, 2005; Zhuang et al., 2008; Lombardi et al., 2010; Marzocchi and Lombardi, 2009). In the last decades, this model has been greatly developed (see Helmstetter and Sornette, 2002; Ogata et al., 2003; Ogata, 2004; Zhuang et al., 2005; Console et al., 2006a,b, 2007). In particular, its space-time distribution now has several different mathematical forms, developed by different researchers such as Marzocchi and Lombardi (2008, 2009), Helmstetter et al. (2006), Werner et al. (2011), and Zhuang (2011).

During the L'Aquila sequence, the ETAS model was used to forecast aftershocks in real time after the April 6, 2009 L'Aquila earthquake (Marzocchi and Lombardi, 2009). For the month following the L'Aquila earthquake, there was a good fit between the forecasts and observations. Subsequently, in 2012, two versions of the ETAS model were applied in real time to the seismic sequence of the May–June 2012 Emilia earthquakes (Marzocchi et al., 2012) to track the daily and weekly evolution of aftershock sequences. The models used were by Falcone et al. (2010) and Lombardi and Marzocchi (2010). However, model performance was not evaluated in their work. Nanjo et al. (2012) applied the ETAS model to Japan, together with four other models, in a retrospective way, for the aftershock sequence following the March 11, 2011 magnitude (M) 9.0 Tohoku-Oki earthquake, and obtained reliable results. They found that the model by Falcone et al. (2010) was

better suited in the number forecasts compared to other models. This ETAS model is one of the models submitted to the various testing centers of the CSEP (Collaboratory for the Study of Earthquake Predictability) in Japan (Tsuruoka et al., 2012), Italy, New Zealand, and California. In this study, the primary objectives are to compare two versions of the ETAS model developed by different research groups, and to evaluate the influence of small earthquakes on predictive performance. The first version, named ETAS I, was developed by Ogata (1998) and Console et al. (2010a,b). The second, hereafter referred to as ETAS II, is used by Zhuang et al. (2005), Ogata and Zhuang (2006), and Zhuang (2011). Although these two versions belong to the same class of ETAS model, their details, such as how to determine background seismicity, how to deal with boundary conditions, and how to estimate parameters, are quite different. In this study, we hope to find which choices result in better forecasting performance.

To test their capability for short-term forecasting of moderate and large earthquakes in the area mentioned above, we apply these two epidemic models to the instrumental database of shallow seismicity (March 16, 2009 – June 30, 2009) collected by INGV. We use the probability gain to evaluate the forecasting performance, with the Proximity to Past Earthquakes (PPE) model, which is a forward kernel estimate model, as the reference model. We are also going to verify how much information is carried by small events, compared to larger events, in forecasting the occurrence of the mainshock. Two versions of the ETAS model are considered for the purpose of understanding the role of small events in forecasting large events, in order to verify the statement made by Helmstetter (2003) and Helmstetter et al. (2005) that small events are important in triggering large earthquakes.

Another aim of this experiment is to see how the forecasting of seismicity works by using models with nearly blind parameters. The formal Italy catalog (ISIDE) begins in 2005, and the forecasting period begins in 2009. Thus, there are only four years of low seismicity with which to train the model parameters. We attempt to evaluate the forecasting performance of the models when applied to such a catalog.

2. Models for earthquake occurrences

2.1 The PPE model

The PPE (Proximity to Past Earthquakes) model was proposed/formulated by Jackson and Kagan (1999) and named by Rhoades and Evison (2004). Essentially, it is a Poisson model with a specific method of estimating the seismicity rate, i.e., it is a smoothed seismicity base-

line model. It can play the role of a spatially varying reference model against which the performance of time-varying models can be compared. For this reason, it is adopted in this study as the reference model. The Relative Intensity (RI) model by Nanjo (2010, 2011) and the Simple Smoothed Seismicity (Triple-S) model by Zechar and Jordan (2010), also belong to this category, the only difference being how the seismicity rate is estimated. The PPE model is used as a forward likelihood method to estimate the seismicity rate function (see also Chiodi and Adelfio, 2011), i.e., it has a conditional intensity (seismicity rate) of the form

$$\lambda(t, x, y, m) = \frac{s(m)}{t-t_0} \sum_{i < t} \left(\frac{a}{d^2 + (x-x_i)^2 + (y-y_i)^2} + \varepsilon \right), \quad m \geq m_c \quad (1)$$

where i includes all events before t (possibly including events outside the study region), t_0 is the starting time, a , d , and ε are model parameters, and $s(m)$ represents the probability density function (p.d.f.) form of the Gutenberg-Richter (G-R) magnitude-frequency relation for earthquake magnitudes larger than the magnitude threshold m_c , i.e.,

$$s(m) = \beta e^{-\beta(m-m_c)}, \quad m \geq m_c \quad (2)$$

with β linked to the b -value by $\beta = b \ln 10$. Such a formulation enables us to maximize the likelihood estimate of the seismicity rate.

Given an earthquake catalog, which is recorded as a list in the form $\{(t_i, x_i, y_i, m_i) : i = 1, \dots, N\}$ from spatial region S and time interval $[0, T]$, the likelihood has the standard form (see, e.g., Ogata, 1988; Daley and Vere-Jones, 2003; Zhuang et al., 2012)

$$\log L = \sum_{j:(t_j, x_j, y_j) \in S \times [0, T]} \log \lambda(t_j, x_j, y_j) - \int_0^T \int_S \lambda(t, x, y) dx dy dt + \sum_{j:(t_j, x_j, y_j) \in S \times [0, T]} \log s(m_j), \quad (3)$$

where $\lambda(t, x, y) = \lambda(t, x, y, m)/s(m)$ is the (t, x, y) -marginal conditional intensity.

2.2 Space-time ETAS models

The ETAS model is a stochastic point-process model, and is more complicated than the PPE model. It builds on two well-known empirical laws: (1) the aftershock rate generally decays according to the Omori-Utsu formula; (2) the aftershock area grows exponentially with mainshock magnitude (Utsu and Seki, 1955), as does the number of aftershocks (Kanamori and Anderson, 1975), leading to the productivity as an exponential function of the magnitude. It also takes stationary seismicity and secondary aftershocks into account. In the

model, we still assume that the magnitude distribution is separable from the other components, with the same density as specified by Equation (2). The expected number of earthquakes in the unit space-time window centered at (t, x, y) , given the observations before t , can be written as

$$\lambda(t, x, y, m) = s(m) \left[\mu(x, y) + \sum_{i: t_i < t} \xi(t, x, y; t_i, x_i, y_i, m_i) \right], \quad (4)$$

where $\mu(x, y)$ represents the spontaneous (background) seismicity rate, which is a function of spatial locations but is constant in time, and $\xi(t, x, y, m; t_i, x_i, y_i, m_i)$ is the contribution to the seismicity rate by the i th event occurring previously. In practice, the response function $\xi(t, x, y, m; t_i, x_i, y_i, m_i)$ is assumed to be separable and dependent on the differences in time and spatial locations of the initiating event, or explicitly,

$$\xi(t, x, y; t_i, x_i, y_i, m_i) = \chi(m_i) g(t - t_i) f(x - x_i, y - y_i; m_i), \quad (5)$$

where

$$\chi(m) = Ae^{(m-m_c)}, m \geq m_c \quad (6)$$

is the expectation of the number of ‘‘children’’ (productivity) from an event of magnitude m , with parameter α describing the scaling of the aftershock productivity with the magnitude of the triggering earthquake according to $e^{\alpha M}$,

$$g(u) = \frac{p-1}{c} \left(1 + \frac{u}{c} \right)^{-p}, \quad u > 0, \quad (7)$$

is the p.d.f. of the length of the time interval between a ‘‘child’’ and its ‘‘parent’’, c and p are constants, u is the elapsed time since the main event (Omori-Utsu formula), and $f(x, y; m)$ is the p.d.f. of the relative locations between the ‘‘parent’’ and ‘‘children’’. The spatial density distribution of aftershocks has been studied by many authors. Felzer and Brodsky (2006) examined the distance distribution of earthquakes occurring in the first five minutes after $2 \leq M < 3$ and $3 \leq M < 4$ mainshocks, and found that their magnitude $M > 2$ aftershocks showed a uniform power-law decay with slope of 1.35, up to 50 km away from the mainshocks. From this they argued that the decay with distance could be explained only by dynamic triggering. Richards-Dinger et al. (2010) found a similar pattern with a decay slope of 1.24 ± 0.09 under the same conditions, although they propose an alternative explanation for the decay. More recently, Gu et al. (2013) confirmed a power law decay in the spatial distribution of aftershocks with an exponent of less than two. However, it must be noted that these studies considered the density distribution of

the aftershocks as a whole, given the identification of a single mainshock in a sequence, whereas our models deal with the p.d.f. of the relative locations between each event and every direct offspring in the sequence, without including the effect of secondary triggering. Here we consider two forms for the spatial p.d.f.:

$$ETAS I. \quad f(v, w; m) = \frac{q-1}{\pi D e^{\alpha(m-m_c)}} \left[1 + \frac{v^2 + w^2}{D e^{\alpha(m-m_c)}} \right]^{-q}, \quad (8)$$

and

$$ETAS II. \quad f(v, w; m) = \frac{q-1}{\pi D e^{\gamma(m-m_c)}} \left[1 + \frac{v^2 + w^2}{D e^{\gamma(m-m_c)}} \right]^{-q}, \quad (9)$$

In the above formulations, the constant parameter vector is $\theta = (A, \alpha, c, p, D, q)$ for ETAS I and $\theta = (A, \alpha, c, p, D, q, \gamma)$ for ETAS II. A further free parameter for both ETAS I and ETAS II is the β value of the magnitude p.d.f. $s(m)$ introduced in Equation (2). This parameter is determined for the entire study area independently of the other ETAS model parameters, as will be explained later in this paper. In the ETAS II model, the parameter γ is independent of α . A detailed description of these ETAS I parameters is given in Appendix B. Consider that in ETAS I, the exponential term depending on magnitude in Equation (6) is divided by the right hand side denominator of Equation (8) in the product of Equation (5). The result is that the spatial p.d.f. is equal to one at zero distance from the parent event, regardless of its magnitude, as can be seen in Equation (23) of Appendix B. Therefore, the total number of triggered events from a triggering earthquake depends exponentially on its magnitude. In this case, assuming $\alpha = 1.0$, $\ln 10$ has the physical meaning that the number of triggered events is proportional to the rupture area of the triggering earthquake (see Console et al., 2006b; Hainzl et al., 2008).

Among the parameters, A is the expected number of direct offspring produced by an event of the threshold magnitude, α quantifies the difference in the productive efficiency from events of different magnitude, c and p are the parameters in the Omori-Utsu formula for aftershock frequencies, D is the characteristic triggering distance, q is the spatial decay coefficient, and γ is the spatial scaling factor. It is easy to see that both models described above are branching processes with ‘‘immigrants’’: the background (immigrant) process is a Poisson process; once an event occurs, whether it is from the background process or triggered by a previous event, it triggers a non-stationary Poisson process, specified by Equation (5), as its offspring process. This model is also a type of self-excitation process (Hawkes, 1971a,b). The difference between ETAS I and ETAS II is as follows. In ETAS I, the scale factor for the productivity function is the same as the scaling factor for the

spatial response function. ETAS I is used by Ogata (1998), Console et al. (2003, 2006a,b, 2010b,c), Murru et al. (2009), and Falcone et al. (2010). In this article, the conditional intensity of ETAS I given by Equation (8), with g and f being normalized to be integrated to one, is slightly different from its original form (See Appendix A). ETAS II was developed by Zhuang et al. (2005) and Ogata and Zhuang (2006), and is an improved version of the model in Ogata (1998). Zhuang et al. (2004) and Zhuang (2006) found that ETAS II fits the data better than ETAS I by using stochastic reconstruction and second-order residual analysis. The formulation and parameterization were slightly different from Console et al. (2003) and Ogata (1998), but were essentially the same: here we normalize the spatial response kernel, and explicitly separate $\kappa(m)$ from other components. If the background seismicity rate $\mu(x, y)$ is known, the model parameters, θ , can be estimated by maximizing the likelihood function, which takes the same form as Equation (3). When the background seismicity rate is unknown, we can still estimate the background and clustering parameters simultaneously by some iterative algorithms. We apply the following iterative procedure to both models to simultaneously estimate the model parameters and the spatially varying background seismicity rate.

(1) Set up an initial background seismicity rate $\mu(x, y)$.

(2) Find the maximum likelihood estimate (MLE) of model parameters, with assumed back-ground seismicity rate. For ETAS I, the free parameters are (A, c, p, D, q) , and α is fixed. For ETAS II, the free parameters are $(A, \alpha, c, p, D, q, \gamma)$.

(3) Compute the background probability ϕ_i as the ratio between the background seismicity component and the total seismicity rate for any event i .

(4) Update $\mu(x, y)$ with the new estimated $\{\phi_i; i = 1, 2, \dots, N\}$ using Equations (24) and (25) for ETAS I and ETAS II, respectively.

(5) Go to Step 2, and continue until the maximum likelihood parameters do not exhibit significant variations.

2.3 Forecasting procedure

Given the observation up to time t , to forecast whether there will be one or more earthquakes of magnitude greater than M_f in the next time interval T_1, T_2 in region S , the probability that at least one event will occur is given by

$$\Pr\{N([T_1, T_2] \times S \times [m_f, \infty)) > 0\} = 1 - \exp\left\{-\int_{m_f}^{\infty} \int_{T_1}^{T_2} \iint_S \lambda(t, x, y, m) dx dy dt dm\right\}, \quad (10)$$

and the expectation of the number of events occurring in $[T_1, T_2] \times S$ is given by

$$\mathbb{E}[N([T_1, T_2] \times S \times [m_f, \infty))] = \int_{m_f}^{\infty} dm \int_{T_1}^{T_2} dt \iint_S \lambda(t, x, y, m) dx dy \quad (11)$$

In the equations above, $\lambda(t, x, y, m)$ in $[T_1, T_2] \times S \times [m_f, \infty)$ is, in principle, calculated based on the observation up to time t . However, since we have no observations in the forecasting period $[T_1, T_2]$, we use the approximate observations up to T_1 , and assume that there are no events occurring in between T_1 and t . As pointed out by Zhuang (2011), this procedure underestimates the number of event occurrences because the triggering of earthquakes in the forecasting time interval is not counted in this equation. This underestimation becomes worse during the activating triggering periods in the beginning stages of aftershock sequences. Such underestimation could be avoided by using simulations. However, in this study, we do not consider using simulations in order to make the comparison simpler, since simulation is not implemented in the forecasting procedure of ETAS I. In this article, we adopted a fast and neat computation method for both models:

$$\mathbb{E}\{N([T_1, T_2] \times S \times [m_f, \infty))\} = e^{-\beta(m_f - m)} (T_2 - T_1) \iint_S \lambda(T_1, x, y) dx dy. \quad (12)$$

As $\lambda(t, x, y)$ is a decreasing function of time t to the occurrence times of subsequent events, and $\lambda(t, x, y)$ increases after an event occurring in the forecasting time interval, Equation (14) underestimates the occurrence rate. In Equation (15), since $\lambda(T_1, x, y) > \lambda(t, x, y)$ for all t in $[T_1, T_2]$, the above approximation compensates to some degree the underestimation that is caused by the assumption that, while calculating $\lambda(t, x, y)$, there are no events occurring between T_1 and t .

3. Scoring procedures

Given forecasts from different models, it is important to know which model performs the best in forecasting. Many methods have been proposed for testing and evaluating the significance of forecasts (e.g., Zechar, 2010; Zhuang, 2010). In this study, we only consider the information score, also called the entropy score (e.g., Kagan and Knopoff, 1977; Vere-Jones, 1998; Harte and Vere-Jones, 2005), which is a natural way to evaluate the performance of probability forecasts. Suppose that the whole space-time-magnitude window for a forecast is divided into M cells of equal size and that the forecast gives a probability \hat{p}_i that at least one event occurs in the i th space-time-magnitude cell; the reference model, usually taken as the Poisson model, gives a probability of p_i . The binomial score for cell i against the reference model is then defined as the logarithm of the likelihood ratio of the forecasting and reference models

$$B_i^{(b)} = Y_i \log(p_i/\hat{p}_i) + (1 - Y_i) \log\left(\frac{1 - p_i}{1 - \hat{p}_i}\right), \quad (13)$$

where $Y_i = 1$ if there is at least one event occurrence in the i th cell, and, $Y_i = 0$ otherwise. Similarly, the Poisson score is defined by

$$B_i^{(P)} = \sum_k I(n_i = k) \log(p_{i,k}/\bar{p}_{i,k}), \quad (14)$$

where n_i is the number of events occurring in cell i , and $I(A)$ is the logical function, taking on a value of one if A is true, and zero if A is false, $p_{i,k}$ and $\bar{p}_{i,k}$ are the probabilities that the i th cell has k events occurring in it, given by the forecasting model and the reference model, respectively. The quantity $e^{B_i^{(b)}}$ or $e^{B_i^{(P)}}$ is the probability gain for the i th interval. The total information score over all cells is defined as

$$B_b = \sum_{i=1}^M B_i^{(b)} \quad \text{or} \quad B_p = \sum_{i=1}^M B_i^{(P)} \quad (15)$$

The information gain per unit space-time-magnitude volume is defined as

$$G = \frac{\Delta}{V} \sum_{i=1}^M B_i, \quad (16)$$

where V is the total volume of the space-time-magnitude range, B_i takes on the value of $B_i^{(b)}$ or $B_i^{(P)}$ and Δ is the size of each cell. When B_i takes on the value of $B_i^{(b)}$, G is called the binomial information gain. Similarly, G is the Poisson information gain when we use $B_i^{(P)}$. Such information gain varies when the division of the space-time-magnitude range of interest changes, but converges to a limit related to the forecasting potential

of the forecasting model when the volume of each cell becomes infinitesimally small (Vere-Jones, 1998).

4. Data analysis and results

4.1 Data

The data used for this study are drawn from the Italian Seismic Instrumental and Parametric Database (ISIDE) of INGV. Figure 1 shows the epicenter distribution of 2,739 shallow earth-quakes with magnitudes equal to or larger than 1.6 and depths ≤ 30 km, reported by INGV from April 16, 2005 to March 15, 2009, and 3,007 events from March 16, 2009 to June 30, 2009 with $M \geq 2.0$ in the region to be tested with all the forecast models (i.e., 12.4°E – 14.2°E , 41.5°N – 43.1°N). A preliminary analysis of this seismic catalog shows contamination during the first time period by man-made explosions from quarries, spatially concentrated around (42.0°N , 13.0°E), as shown in Figure 1. These 151 events are distinct, since their magnitudes are small (≤ 2.0) and they occur during local working hours (between 7 a.m. and 5 p.m., UTC time). Thus, the catalog used for the learning phase does not include the quarry blasts, contains 2,588 events, and is considered complete for $M \geq 1.6$.

4.2 Preliminary estimation results for the learning phase

PPE model. When fitting the PPE model to earthquake data, two cutoff magnitudes are used, 1.6 (2,588

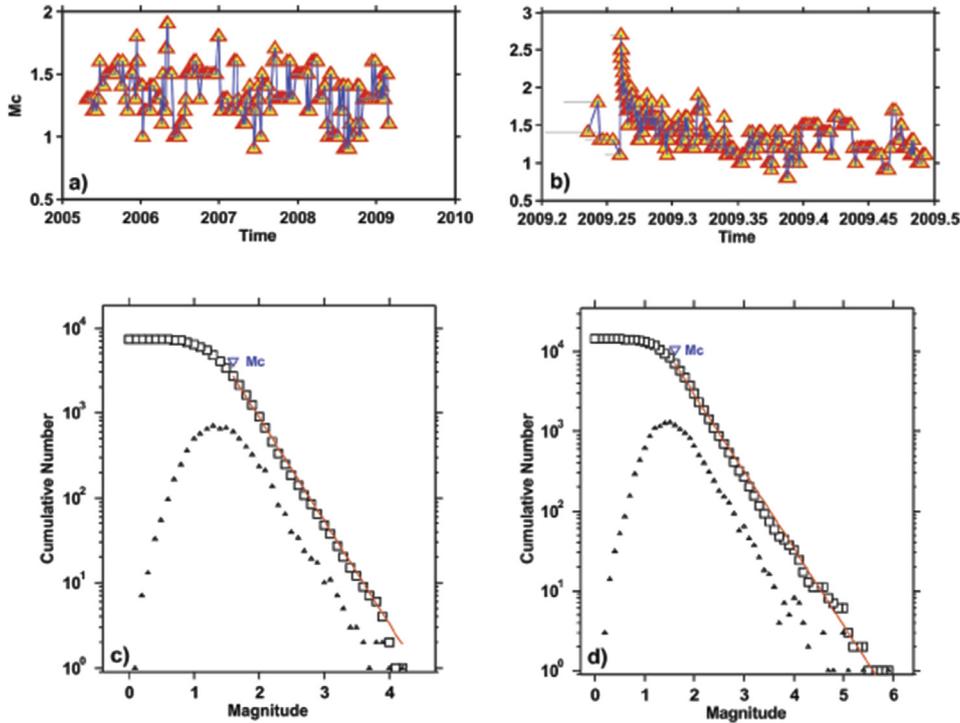


Figure 2. Magnitude versus time: (a) for the learning (April 16, 2005 to March 15, 2009) (upper left panel) and (b) test (March 16, 2009 to June 30, 2009) (upper panel, right side) periods. All magnitudes are considered in the plots. The area considered spans 41.5°N – 43.1°N and 12.4°E – 14.2°E . Magnitude-frequency plot: (c) for the learning (lower plot, left side) and (d) test (lower plot, right side) periods. The M_c is selected according to the best combination of the 95%–90% maximum curvatures where the observed data can be fit by a straight line.

Model	a # (events)	d deg	ϵ # (events)	$\log L$
PPE1.6+	$9.542 \cdot 10^{-3}$	$3.363 \cdot 10^{-3}$	$6.469 \cdot 10^{-8}$	-855.18
PPE2.0+	$1.497 \cdot 10^{-2}$	$3.858 \cdot 10^{-3}$	$3.360 \cdot 10^{-8}$	-875.14

Table 1. Parameters of PPE models estimated by the maximum likelihood method (learning phase).

Model	A # (events)	c day	α	p	D deg ²	q	γ	β
ETAS I	0.2015	0.0202	2.30 (fixed)	1.10	$2.13 \cdot 10^{-5}$	2.01	n.a.	2.58
ETAS II-1.6+	0.642	0.00382	0.521	1.04	$7.49 \cdot 10^{-5}$	1.73	0.910	2.60
ETAS II-2.0+	0.453	0.00511	0.588	1.12	$1.04 \cdot 10^{-4}$	1.73	0.755	2.60
ETAS II-2.0m	0.191	0.00396	1.140	1.09	$7.93 \cdot 10^{-5}$	1.70	0.736	2.60

Table 2. Parameters of the ETAS models estimated by the maximum likelihood method (learning phase).

events) and 2.0 (907 events), while the target events are still of magnitude 2.0 and higher. That is, in Equation (1), the value of i goes through all the events of the cut-off magnitude, whereas, in the likelihood Equation (3), i goes through all the events of magnitude 2.0 and higher. We denote the model with cutoff magnitudes 1.6 and 2.0 by PPE1.6+ and PPE2.0+, respectively. The parameters obtained from fitting PPE1.6+ and PPE2.0+ are listed in Table 1. Comparing between the likelihoods, we can see that PPE1.6+ is better than PPE2.0+, implying that using information from smaller earthquakes improves the fitting and forecasting performance. Figure 3a shows the map of the estimated seismicity rates from the PPE1.6+ model, using data from before March 15, 2009.

Model ETAS I. For the learning period of this model, we use events recorded by INGV from April 16, 2005 to March 15, 2009, with $M \geq 1.6$. The threshold magnitude, for both triggering and triggered events, was taken to be 1.6, and the test period is from March 16, 2009 to June 30, 2009, using events of $M \geq 2.0$ (see Figure 2). The b -value is assumed to be constant over the geographical area spanned by the catalog, and is estimated independently of the other parameters. We do not use a spatially varying b -value, since the study region is already quite small for justifying a subdivision. Instead, the b -value was determined for the whole investigated area, with the maximum likelihood method of Aki (1965), to be equal to 1.12 ± 0.02 . The standard deviation estimate is obtained following Shi and Bolt (1982).

Consequently, the β -value is also estimated independently of the other parameters, as it is equal to $b \ln 10$. As described in Section 2.2, ETAS I has five free parameters (A , α , c , p , q) that can be estimated by the maximum likelihood method in Equation (3). However, we fixed the value of α to $1.0 \ln 10$ to reduce the number

of free parameters, as done in Console et al. (2010a,c) and Falcone et al. (2010). Thus, the effective number of free parameters for ETAS I is four. Before determining the best fit of the above mentioned parameters, we determine the scaling factor d in Equation (24), modeling the seismicity as a pure Poisson process, so that no information about the parameters characterizing the induced seismicity is necessary at this stage. In this work, the optimal choice of d is made by trial and error. For the assessment of this parameter, we used the method of comparing the independent components of the seismicity distribution of two different periods containing roughly the same number of events.

We maximized the likelihood of the seismicity contained in half of the INGV earthquake catalog (from 2007 to 2009, $M \geq 1.6$) under the time-independent Poisson model obtained from the other half of the catalog (2005–2006). Here, d is modified until the maximum likelihood is obtained for one of the subcatalogs, using the smoothed seismicity derived from the other subcatalog. The correlation distance d was found to be 8 km, which is the value taken for the following analysis.

The maximum likelihood set of the four free parameters of ETAS I is found using the initial values of the smoothed seismicity $\mu(x, y)$ by interpolation of the gridded distribution. The initial smoothed background seismicity rate is then obtained for the whole data set. Parameter b is probably influenced by the location errors of the epicenters reported in the catalog analyzed, which is not taken into consideration in our simple algorithm. Figure 4(a) shows the map of the final smoothed estimate of the background rate, obtained after four iterations using the iterative procedures described in Step 3 of Section 2.2. The maximum log-likelihood parameters obtained after the last iteration, for the

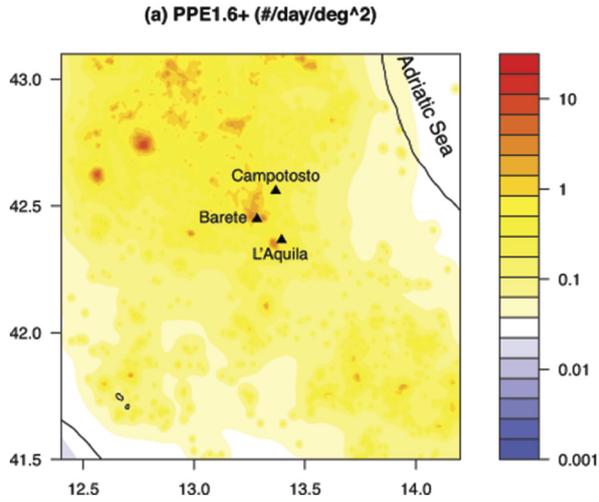


Figure 3. Seismicity rate estimated using data before $T_2 = 1430$ (at 00:00:00, March 16, 2009), using the PPE model (PPE1.6+). The learning period is from April 16, 2005 to March 15, 2009. The color scale represents the number of events of magnitude larger than $M_{2.0}$ occurring in a $1^\circ \cdot 1^\circ$ cell per day.

learning phase, are: $A = 0.2015$, $D = 2.132 \times 10^{-5}$, $q = 2.009$, $c = 2.002 \times 10^{-2}$, $p = 1.105$, and $\alpha = 2.30$ (fixed) (see also Table 2). The stability of our maximum likelihood algorithm was tested by starting from different initial sets of parameters, and checking that the same results are always obtained. One might notice that the parameters for ETAS I are super-critical, i.e., $\rho = A\beta/(\beta - \alpha) > 1$ (see, e.g., Helmstetter and Sornette, 2002; Zhuang et al., 2004 and Zhuang and Ogata, 2006, for details on the criticality of the ETAS model), unless a truncated G-R magnitude-frequency relation is used, or some other similar treatments are adopted. The direct consequence of supercriticality is that the total number of events grows, on average, exponentially with time. When the forecasting interval is long, over-forecasting of seismicity might happen. However, a supercritical process can still be used to forecast the aftershock production within a short period after a triggering event. In this case, the same parameters can be

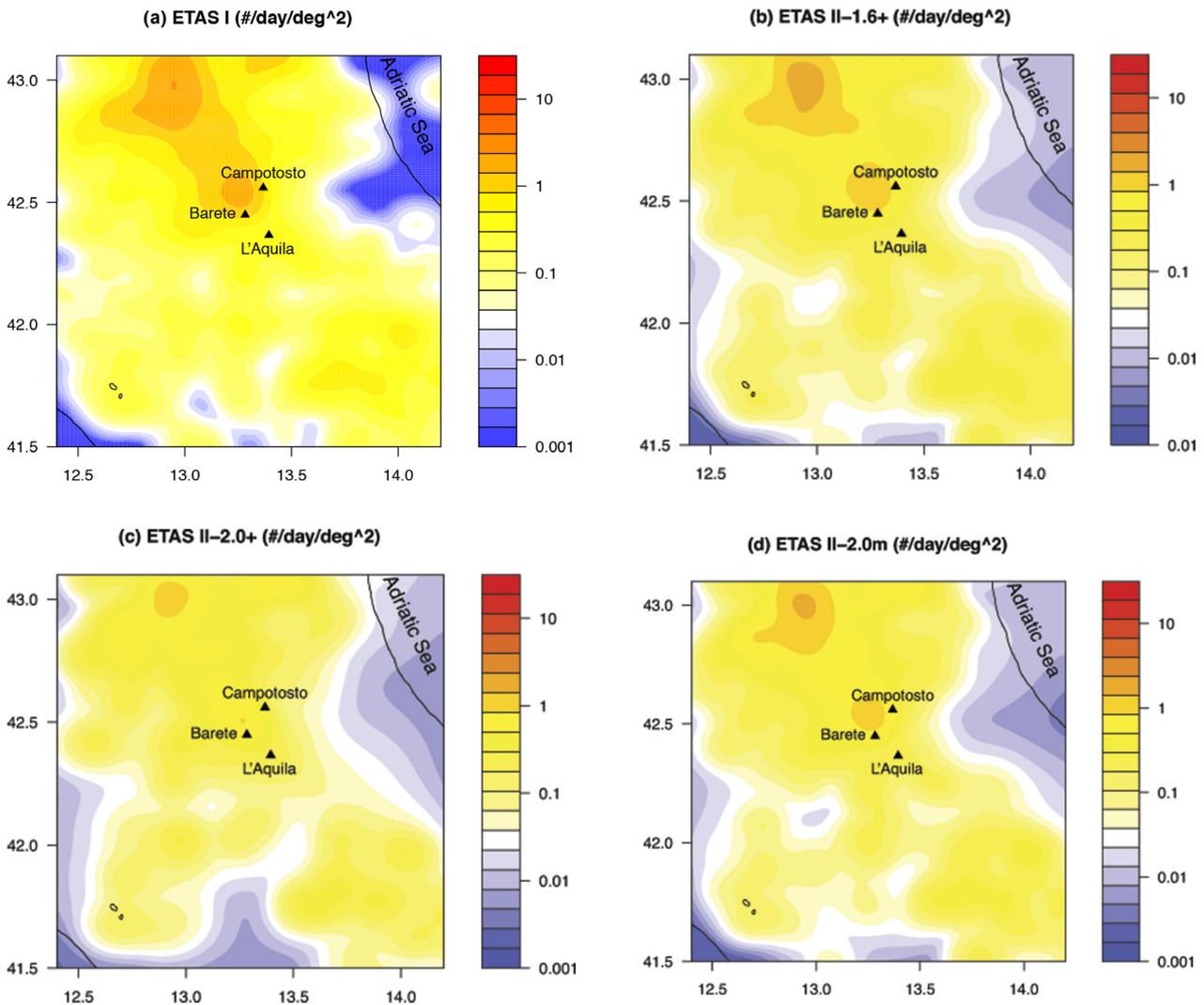


Figure 4. Final estimates of background rate $\mu_0(x, y)$ according to (a) ETAS I, (b) ETAS II-1.6+, (c) ETAS II-2.0+, and (d) ETAS II-2.0m using data before $T_2 = 1430$ (at 00:00:00, March 16, 2009). The learning period is from April 16, 2005 to March 15, 2009. The color scale represents the number of events of magnitude larger than $M_{2.0}$ occurring in cells $1^\circ \cdot 1^\circ$ per day.

used for simulation, since a supercritical ETAS process needs some time to enter the state of population explosion. In our tests, after each time step, the history of the process is updated by observation, but not by events simulated in previous steps. That is, events generated by the supercritical parameters are not put into the process history. With such a temporally piece wise scheme, forecasting by the ETAS model remains stable. Model ETAS II. In the estimation of ETAS II, earthquake data from the region (7.9°E–18.7°E) and (36.5°N–47.1°N) are used to obtain the seismicity rate. The events in this larger region with respect to the learning area of ETAS I (that is, outside of the L’Aquila area) are used as auxiliary events to incorporate the triggering effect from events outside into the contribution to the occurrence rate inside this area (see also Wang et al., 2010). That is, the events in this larger area are not considered in the likelihood computation. However, they are considered as events that could trigger other events inside the test area. This is a different point from ETAS I, arising from the independent development of

these two models. The target region (12.4°E–14.2°E, 41.5°N–43.1°N) is the same as in ETAS I. The learning period is April 16, 2005 – March 15, 2009.

Three versions of ETAS II are considered in this study: (1) we fit the ETAS model to earthquake events of $M \geq 2.0$ (named ETAS II-2.0+ in Table 2); (2) we fit the ETAS model to earthquake events of $M \leq 1.6$ (named ETAS II-1.6+ in Table 2) and scale the forecast seismicity rate for events of $M \geq 2.0$ using the G-R magnitude-frequency relation; and (3) we fit the ETAS model to earthquake events of $M \geq 2.0$, but with events of $1.6 \leq M < 2.0$ as auxiliary events (named Model II-2.0m in Table 2); that is to say, when calculating $\lambda(t)$ using Equation (4), i is taken over all the $M \geq 1.6$ events before t , and when computing the likelihood in Equation (3), i is only taken over the events of $M \geq 2.0$. These three versions are considered in order to understand the role of small events in forecasting large events, and to verify the statement made by Helmstetter (2003) and Helmstetter et al. (2005) that small events are important in triggering large earthquakes. The MLE of the

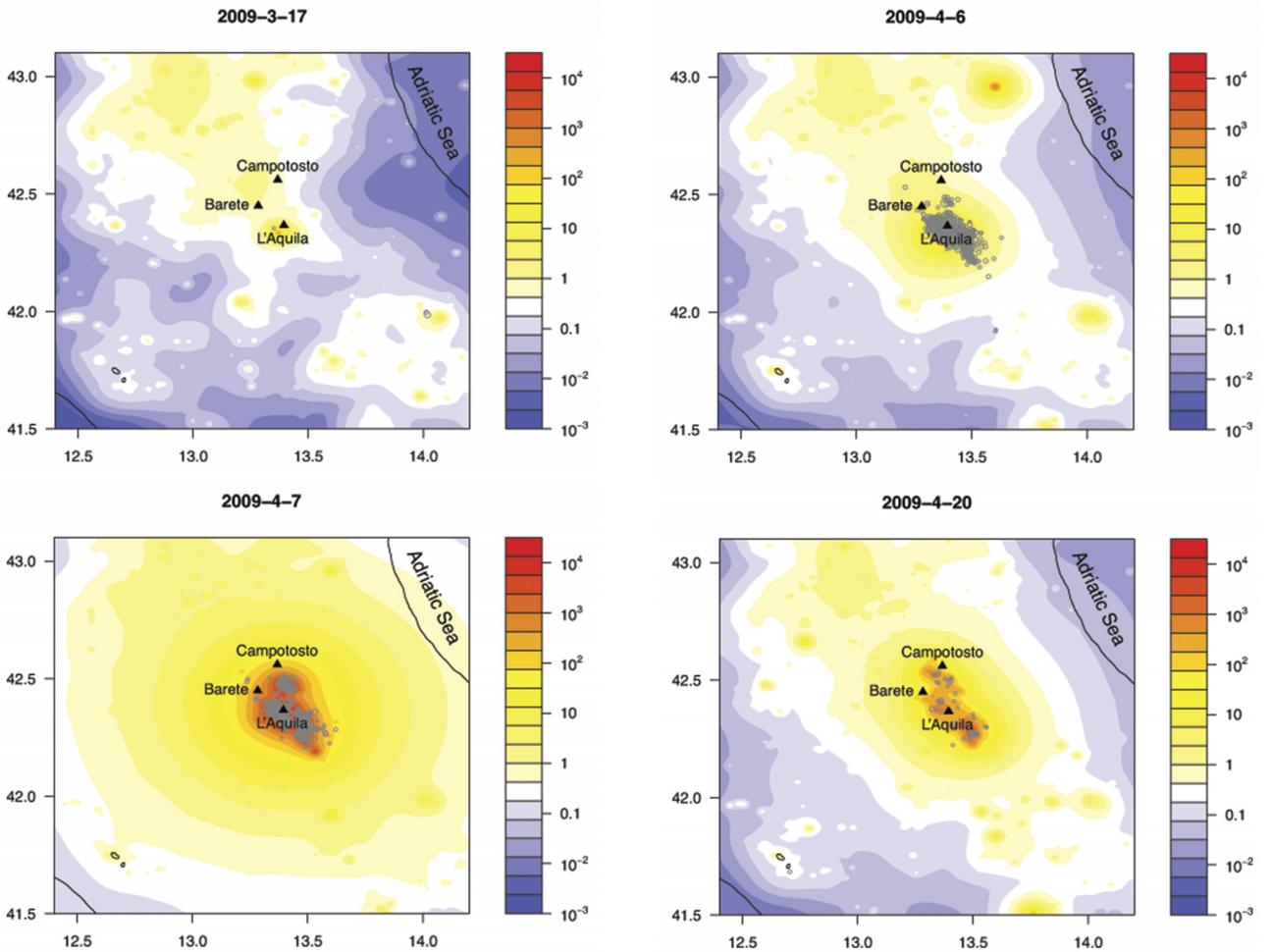


Figure 5. Examples of expected daily seismicity rate forecast by ETAS II-2.0, at 00:00 UTC on the days: (a) March 17, 2009, (b) April 6, 2009 (1 hour and 32 minutes before the L’Aquila mainshock $M_W 6.3$), (c) April 7, 2009 (the second-largest shock in the Abruzzi region occurred on April 7, 2009 at 17:47 UTC, with $M_W 5.6$), and (d) April 20, 2009. The color scale represents the number of events in units of magnitude larger than 2.0 in cells of $1^\circ \cdot 1^\circ$ per day.

model parameters are listed in Table 2. Figures 4(b) to 4(d) show maps of background rates estimated using ETAS II-1.6+, II-2.0+, and II-2.0m, respectively.

Daily forecasts. We assume that each forecast is for seismicity starting at 0:00 and ending at 24:00 (UTC) each day. Figure 5 gives the forecasting maps of expected intensity of events in units of $events/day/deg^2$ for four time windows, according to ETAS II-2.0m.

Figure 5(a) shows the expected daily seismicity rate forecast for March 17, 2009. Just before the occurrence of the mainshock (April 6, 2009), the density map is similar to the background seismicity rate (Figure 4d), taking into account the difference in the color scale for rate density.

After the occurrence of the mainshock, the seismicity in the L'Aquila region is dominated by aftershocks (Figure 5c). If the forecast is carried out shortly after the mainshock, it is possible that the forecast can be improved. In Figure 5(c), this feature has been well captured by ETAS II-2.0m. Two weeks after the mainshock has occurred, the rate of aftershocks decays to a low level (Figure 5d), but is still much higher than the average level (Figure 3).

5. Evaluation of forecast performance

The region to be tested for ETAS I and II for temporal and spatial occurrence is $(12.4^{\circ}E-14.2^{\circ}E) \cdot (41.5^{\circ}N-43.1^{\circ}N)$. The area has been divided into cells of $0.1^{\circ} \cdot 0.1^{\circ}$. The test period is from March 16 to June 30, 2009 for events of $M \geq 2$.

5.1 Predictability of the mainshock

An important feature is that a foreshock swarm, consisting of 50 earthquakes of $M \geq 2.0$, occurred before the mainshock (April 2009). This raises the question of how the mainshock can be forecast by the models considered in this study. On the day of the mainshock, PPE2.0+, PPE1.6+, ETAS I, ETAS II-1.6+, ETAS II-2.0+, and ETAS II-2.0m, give an expected occurrence rate on the order of a few events per day for the occurrence of events of $M \geq 2.0$ in the test region (Figure 6, top panel). Model ETAS I seems to give the highest estimate. However, when we consider its score for all the days before the mainshock, the probability gain becomes much lower, using PPE2.0+ as the reference model. The total conditional probability computed by ETAS I for an earthquake of magnitude $M \geq 5.0$ during the week preceding the April 6 mainshock was 0.39%. This probability was about 20 times larger than the background probability, because of the occurrence of some foreshock activity. The expected instantaneous occurrence rate density increased by several times in the few hours before the mainshock. However, this level still seems

low for justifying the implementation of risk mitigation measures. Here we refer to Van Stiphout et al. (2010) for further discussions on risk mitigation.

It is arguable whether a reliable prediction can be made based on these foreshocks. However, the answer seems to be that we cannot obtain more information than what can be forecast by a clustering model like the ETAS model. The same has been reached by many re-

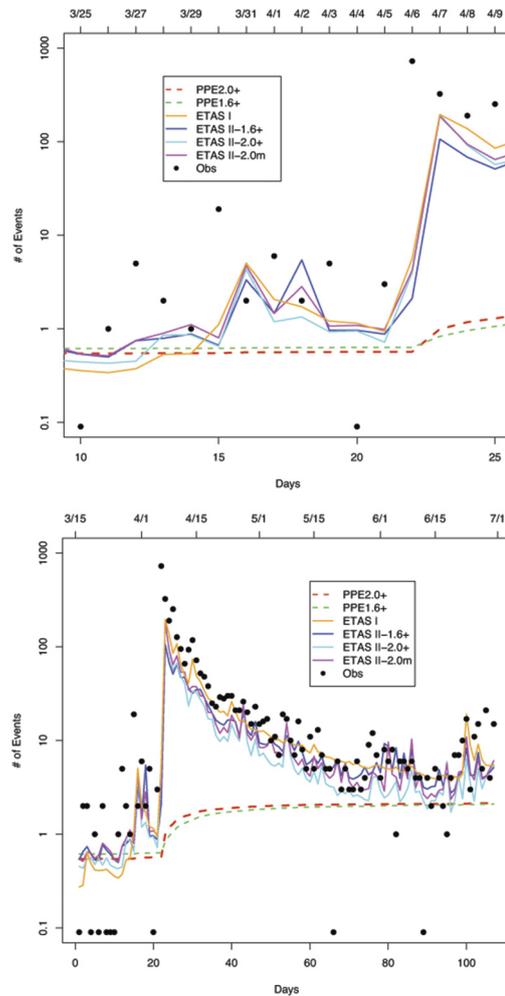


Figure 6. Comparison between forecast and observed rates during the test period (March 16 – June 30, 2009) for events of $M \geq 2.0$, per day. The scale of y-axis is logarithmic, with the value 0.1 considered to be equivalent to 0. The top and bottom panels show the periods March 25 to April 9, 2009, and March 25 to July 1, 2009, respectively.

searchers (see, e.g., Felzer et al., 2004; Helmstetter et al., 2003; Zhuang et al., 2008; Christophersen and Smith, 2008; Marzocchi and Zhuang, 2011).

5.2 Predictability of the mainshock

In this subsection, we evaluate the temporal performance of the forecasts made using all six models. The results are shown in Figure 7. As discussed in Section 5.1, we first divide the target space-time range into cells of size $1(\text{day}) \times S \times [2.0, \infty)$, where S is the test re-

Binomial information scores									
Model	M 2.0+			M 3.0+			M 4.0+		
	EC	NEC	Tot.	EC	NEC	Tot.	EC	NEC	Tot.
PPE1.6+	0.83	0.31	1.14	2.96	-0.23	2.72	1.17	-0.06	1.1
ETAS I	16.15	-5.78	10.37	71.17	-22.21	48.96	35.90	-7.50	28.39
ETAS II-1.6+	18.12	-4.50	13.62	66.63	-17.27	49.36	31.00	-5.25	25.75
ETAS II-2.0+	14.09	-1.09	13.01	56.02	-9.35	46.67	30.11	-3.76	26.35
ETAS II-2.0m	17.87	-3.43	14.44	64.81	-14.71	50.10	32.13	-5.08	27.05

Poisson information scores									
Model	M 2.0+			M 3.0+			M 4.0+		
	EC	NEC	Tot.	EC	NEC	Tot.	EC	NEC	Tot.
PPE1.6+	215.79	0.31	216.10	12.82	-0.23	12.59	1.67	-0.06	1.60
ETAS I	7992.22	-5.78	7986.44	666.64	-22.21	644.44	95.18	-7.50	87.68
ETAS II-1.6+	6860.19	-4.50	6855.69	549.77	-17.27	532.50	75.67	-5.24	70.42
ETAS II-2.0+	7322.93	-1.08	7321.84	593.74	-9.35	584.39	81.88	-3.76	78.11
ETAS II-2.0m	7541.82	-3.43	7541.82	617.48	-14.71	602.77	85.19	-5.08	80.11

Table 3. Information gains compared to Model PPE2.0+ for each model. The boldface fonts show the best models in the corresponding category. EC stands for “event cell” and NEC stands for “non-event cell”.

gion under consideration.

The daily forecasts are given as the expected rate in each of the cells. Figure 6 shows the occurrence rates forecast by all these models, and the actual observations of daily number of earthquakes during the testing period starting on March 16, 2009. Among them, the PPE models start at a rate similar to that of the ETAS models, and with some increments in the forecast rates after the mainshock due to the dramatic increase of the number of events in the aftershock sequence. All the ETAS models catch the patterns of the temporal variation of the aftershock rates with slight differences, except that the curve of earthquake numbers forecast by ETAS I is smoother than those produced by ETAS II. A problem that arises when modeling the occurrence of relatively small events is that these small events are generally missing from the catalog after a large mainshock, so the completeness magnitude increases significantly after the large mainshock. Consequently, one effect of missing early aftershocks after the mainshock is that earthquake forecast models overestimate the observed

number of events during the first few hours or days following the mainshock.

Another effect of missing early aftershocks in the real catalogs is the underestimation of the expected subsequent seismicity rate (e.g., Helmstetter et al., 2006). However, none of these effects can be clearly noted for the catalog used for the L’Aquila mainshock, as can be seen in Figure 6.

Assuming the PPE2.0+ as the reference model, Figure 7 gives the information score for all the models for each day. By comparing the performance of the various models, and also making use of Table 3, we observe the following:

- (1) Inclusion of smaller events improves the forecasting results. As mentioned in a previous section, PPE1.6+ and ETAS II-20m both work better than the versions for events $M \geq 2.0$.
- (2) Both types of ETAS models are superior to the PPE models, with binomial or Poisson scores of one to two orders higher.
- (3) ETAS I is better at forecasting the number of

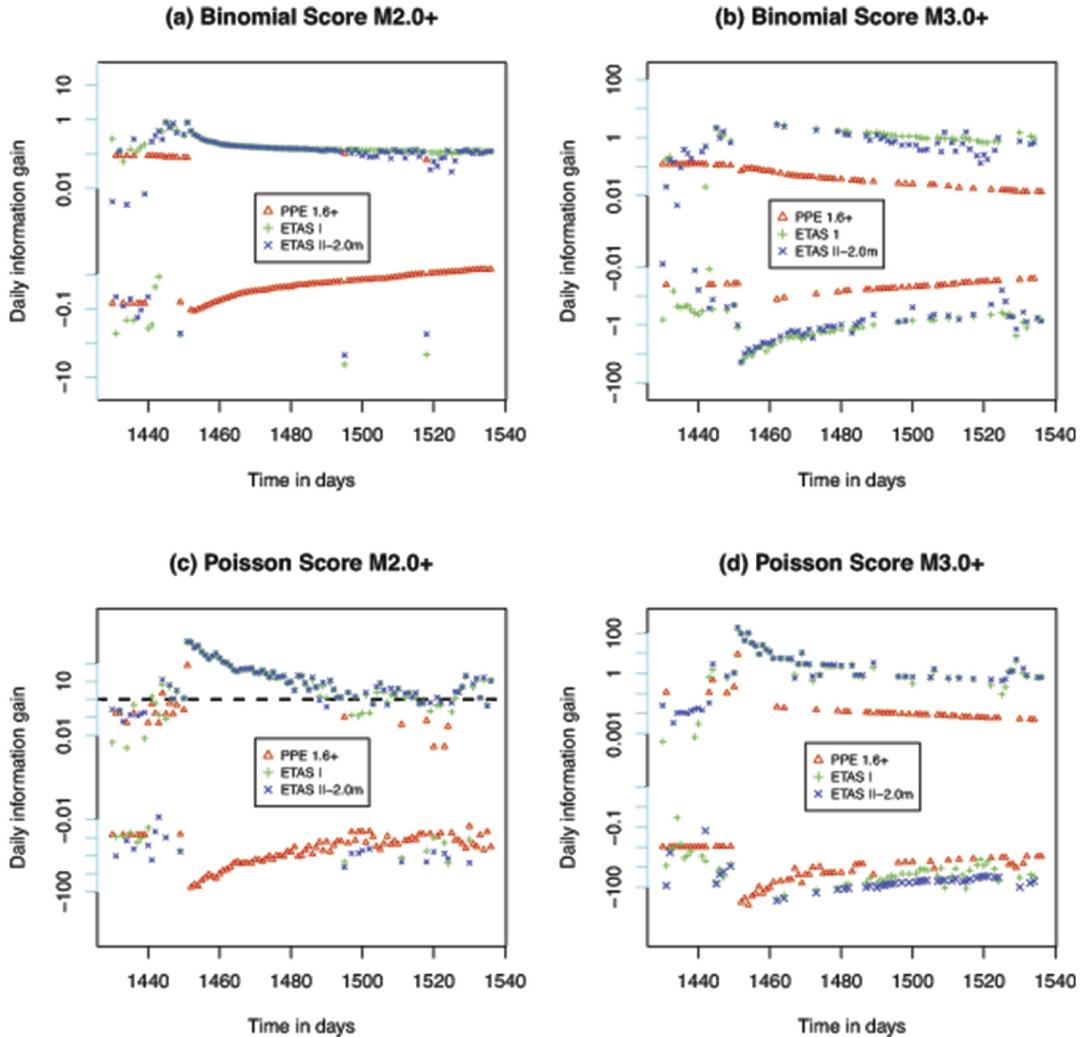


Figure 7. Temporal variations of daily information gains for each forecasting model, with PPE2.0+ as the reference model: (a) binomial information gains for earthquakes of magnitudes 2.0+, (b) binomial information gains for earthquakes of magnitudes 3.0+, (c) Poisson information gains for earthquakes of magnitudes 2.0+, and (d) Poisson information gains for earthquakes of magnitudes 3.0+.

earthquakes, while the ETAS II models are better at forecasting Yes/No for $M \geq 2.0$ or 3.0 .

(4) Direct fitting to $M \geq 2.0$ works better than using models that fit to all events of $M \geq 1.6$.

However, an adjusted ETAS model (II-2.0m), which fits earthquakes of $M \geq 2.0$ but considers the triggering effect from events of magnitudes between $M 1.6$ and $M 2.0$, yields better forecasts than its counterpart II-2.0+. Possible reasons for this are: (1) the triggering effect from small earthquakes is important and cannot be neglected, as shown by Helmstetter et al. (2005); (2) since events of magnitudes $M 1.6$ to $M 2.0$ comprise a major proportion of the data set, when fitting ETAS II-1.6+ to the data, the contributions to the occurrences of these events from events of similar and larger magnitudes are considered, while events smaller than $M 1.6$ are not, yielding a biased estimate.

6. Discussion and conclusions

In Section 5.2, we saw that all the ETAS models used in this study underestimate the occurrence rate

soon after the mainshock and during the aftershock sequence. One way to improve the performance of an ETAS model, especially in the case of a long aftershock sequence, is to update the maximum likelihood parameters at the end of each day, just before the computation of a new forecast for the following day.

Another possible cause of the systematic underestimation of occurrence rate during an aftershock sequence could be that our ETAS models were fit to a learning data set of relatively moderate seismic activity, since there are no earthquakes of magnitude 5.0+ in the catalog before March 15, 2009. In this case, our maximum likelihood (ML) parameters, applied to larger magnitude mainshocks, are in some sense the results of an extrapolation, which can easily cause large biases.

As to the issue of forecasting the April 6, $M_w 6.3$ mainshock, the results reported in Section 5.1 show that the occurrence of foreshock activity some days before the mainshock raised the probability gain up to an order of 20, using the ETAS I model. This does not appear to

be a great achievement towards the potential use of this information for civil protection purposes.

Considering that in this case, as often occurs before strong earthquakes, significant foreshock activity was observed just a few hours before the mainshock after midnight, a suitable strategy could be to issue forecasts in automatic and semi-real time whenever significant seismic activity is observed, instead of just once a day.

In this study, the most significant feature is that before the mainshock, all ETAS II-type models performed better than ETAS I, while in the aftershock period, ETAS I obtained a much better Poisson score than all the ETAS II models. We can see that ETAS I is a subclass of ETAS II. This means that the best model in I should not be able to outperform the best model in II. One possibility is that the seismicity behavior changes before and after the occurrence of the mainshock. To test whether the seismicity pattern changes before and after the mainshock, we fit the space-time ETAS II-2.0+ to the seismicity after the mainshock. In this case, the ML parameters are $A = 0.515$ event, $c = 0.0143$ day, $\alpha = 0.918$, $p = 1.20$, $D = 8.96 \times 10^{-5}$ deg², $q = 1.89$, and $\gamma = 0.508$. Clearly, the α -value is much larger than the one used for forecasting in the ETAS II models ($\alpha = 0.588$), which is obtained by fitting ETAS II-2.0+ to the data before March 15, 2009. This also explains why ETAS I with a high α -value can produce a better Poisson score. The above comparison covers only a very active period. To make an overall forecast comparison, we should carry out a comparison for much longer data set. However, during the quiescent period, ETAS I begins to over-perform other models.

One might ask: What is the authors' advice regarding the two information gain metrics? Is one preferred under certain conditions? In our opinion, there are no problems related to the gain metrics. For the Poisson score, the problem is caused by the assumption that the forecast number of events belong to a Poisson distribution, with the mean being the integral of the conditional intensity function. This is true for Poisson type models, such as the PPE models used in this study, but the ETAS model is a clustering model, with a stochastic conditional intensity, which increases once an event occurs. That is to say, the distribution of number of events in the forecasting interval has a variance much larger than that of a Poisson distribution with the same rate. This problem has been much discussed by Lombardi (2014). For the binomial score, such a problem does not exist. The zero and one probabilities are exactly given by Equation (10). In this respect, it seems the binomial score should be preferred.

In this study, we do not consider the performance of these models in forecasting spatial locations, since

the area is quite small, 1.1° by 1.5° , while the location error is several kilometers. Moreover, all the models use the G-R magnitude-frequency relation for the magnitude distribution, independent of the time and location components. Thus, there are no differences between them in forecasting the magnitude distribution of future events.

In summary, to achieve better forecasts, as well-known from the existing literature on the ETAS model, the training catalog should contain some typical examples of similar magnitudes. If not, it would be better to use a typical set of parameters from a nearby region or from a region of similar tectonic structure or equivalent seismicity levels. Other lessons from this study are as follows. (1) Making use of the earthquakes of lower magnitude works better than using only events larger than the target magnitude threshold. (2) Parameters should be updated during the forecast process, if possible. (3) Forecasts should be updated just after the occurrence of potential foreshock activity, whenever significant moderate earthquakes are observed.

Acknowledgements. The authors thank Warner Marzocchi for helpful discussions. Jancang Zuang was partially supported by JSPS grant-in-aid (C) 25330052.

References

- Aki, K. (1965). Maximum likelihood estimate of b in the formula $\log n = a - bm$ and its confidence limits. *Bulletin of Earthquake Research Institute*, 43:99–103.
- Akinci, A., Galadini, F., Pantosti, D., Petersen, M., Malagnini, L., and Perkins, D. (2009). Effect of time dependence on probabilistic seismic-hazard maps and deaggregation for the central Apennines, Italy. *Bulletin of the Seismological Society of America*, 99(2A):585–610.
- Albarelo, D., Bosi, V., Brammerini, F., Lucantoni, A., Naso, G., Peruzza, L., Rebez, A., Sabetta, F., and Slejko, D. (2001). New seismic hazard maps of the Italian territory. Gruppo Nazionale per la Difesa dai Terremoti - Servizio Sismico Nazionale .
- Anderson, H. and Jackson, J. (1987). Active tectonics of the adriatic region. *Geophysical Journal of the Royal Astronomical Society*, 91(3):937–983.
- Chiarabba, C. et al. (2009). The 2009 L'Aquila (central Italy) M_W 6.3 earthquake: Main shock and aftershocks. *Journal of Geophysical Research*, 36 (L18308).
- Chiodi, M. and Adelfio, G. (2011). Forward likelihood-based predictive approach for space-time point processes. *Environmetrics*, 22(6):749–757.
- Christophersen, A. and Smith, E. G. C. (2008). Foreshock Rates from Aftershock Abundance. *Bulletin of the Seismological Society of America*, 98(5):2133–2148.

- Console, R. and Murru, M. (2001). A simple and testable model for earthquake clustering. *Journal of Geophysical Research*, 106, B5:8699–8711.
- Console, R., Murru, M., and Lombardi, A. M. (2003). Refining earthquake clustering models. *Journal of Geophysical Research*, 108(B10):2468.
- Console, R., Murru, M., and Catalli, F. (2006a). Physical and stochastic models of earthquake clustering. *Tectonophysics*, 417:141–153.
- Console, R., Rhoades, D. A., Murru, M., Evison, F. F., Papadimitriou, E. E., and Karakostas, V. (2006b). Comparative performance of time-invariant, long-range and short-range forecasting models on the earthquake catalogue of Greece. *Journal of Geophysical Research*, 111:B09304.
- Console, R., Murru, M., Catalli, F., and Falcone, G. (2007). Real time forecasts through an earthquake clustering model constrained by the rate-and-state constitutive law: comparison with a purely stochastic ETAS model. *Seismological Research Letters*, 78(1):49–56.
- Console, R., Jackson, D. D., and Kagan, Y. Y. (2010a). Using the ETAS Model for Catalog Declustering and Seismic Background Assessment. *Pure and Applied Geophysics*, 167:819–830.
- Console, R., Murru, M., and Falcone, G. (2010b). Probability gains of an epidemic-type aftershock sequence model in retrospective forecasting of $m \geq 5$ earthquake in Italy. *Journal of Seismology*, 14(1):9–26.
- Console, R., Murru, M., and Falcone, G. (2010c). Retrospective forecasting of $m \geq 4.0$ earthquakes in New Zealand. *Pure and Applied Geophysics*, 167:693–707; doi:10.1007/s00024-010-0068-2.
- Daley, D. D. and Vere-Jones, D. (2003). *An Introduction to Theory of Point Processes – Volume 1: Elementary Theory and Methods* (2nd Edition). Springer, New York, NY.
- DISS Working Group (2010). Database of Individual Seismogenic Sources (DISS), Version 3.1.1: A compilation of potential sources for earthquakes larger than $m_{5.5}$ in Italy and surrounding areas. Technical report, Istituto Nazionale di Geofisica e Vulcanologia.
- Falcone, G., Console, R., and Murru, M. (2010). Short-term and long-term earthquake occurrence models for Italy: ETES, ERS and LTST. *Annals of Geophysics*, 53(3):41–50.
- Felzer, K. R., Abercrombie, R. E., and Ekstrom, G. (2004). A common origin for aftershocks, foreshocks, and multiplets. *Bulletin of the Seismological Society of America*, 94(1):88–98.
- Felzer, K. R. and Brodsky, E. E. (2006). Decay of aftershock density with distance indicates triggering by dynamic stress. *Nature*, 441(7094):735–738.
- Frankel, A. (1995). Mapping seismic hazard in the central and eastern united states. *Seismological Research Letters*, 66:8–21.
- Gruppo di lavoro CPTI (2004). Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04), Technical report, Istituto Nazionale di Geofisica e Vulcanologia, Bologna; <http://emidius.mi.ingv.it/CPTI04>.
- Gruppo di lavoro MPS (2004). Redazione della mappa di pericolosità sismica prevista dall’Ordinanza PCM del 20 marzo 2003, rapporto conclusivo per il Dipartimento della Protezione Civile, Technical report, Istituto Nazionale di Geofisica e Vulcanologia, Milano-Roma; <http://zonesismiche.mi.ingv.it/elaborazioni/>.
- Gu, C., Schumann, A. Y., Baiesi, M., and Davidsen, J. (2013). Triggering cascades and statistical properties of aftershocks. *Journal of Geophysical Research: Solid Earth*, 118(8):4278–4295.
- Hainzl, S. and Ogata, Y. (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling. *Journal of Geophysical Research*, 110(B05):B05S07.
- Hainzl, S., Christophersen, A., and Enescu, B. (2008). Impact of earthquake rupture extensions on parameter estimations of point-process models. *Bulletin of the Seismological Society of America*, 98(4):2066–2072.
- Harte, D. and Vere-Jones, D. (2005). The entropy score and its uses in earthquake forecasting. *Pure and Applied Geophysics*, 162(6):1229–1253.
- Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *J. Royal Stat. Soc. Series B (Meth.)*, 33(3):438–443.
- Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.
- Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research*, 107(B10):2237.
- Helmstetter, A. (2003). Is earthquake triggering driven by small earthquakes?, *Phys. Rev. Lett.*, 91(5):058501.
- Helmstetter, A. and Sornette, D. (2003). Foreshocks explained by cascades of triggered seismicity. *Journal of Geophysical Research*, 108(B10):2457.
- Helmstetter, A., Sornette, D., and Grasso, J.-R. (2003). Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws? *Journal of Geophysical Research*, 108:2046.
- Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. (2005). Importance of small earthquakes for stress transfers

- and earthquake triggering. *Journal of Geophysical Research*, 110.
- Helmstetter, A., Kagan, Y. Y., and Jackson, D. D. (2006). Comparison of short-term and time-independent earthquake forecast models for Southern California. *Bulletin of the Seismological Society of America*, 96(1):90–106.
- Jackson, D. D. and Kagan, Y. Y. (1999). Testable earthquake forecasts for 1999. *Seismological Research Letters*, 70:393–403.
- Kagan, Y. and Knopoff, L. (1977). Earthquake risk prediction as a stochastic process. *Physics of the Earth and Planetary Interiors*, 14:97–108.
- Kanamori, H. and Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. *Bulletin of the Seismological Society of America*, 65(5):1073–1095.
- Lombardi, A. M. and Marzocchi, W. (2010). The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment. *Annals of Geophysics*, 53(3):155–164.
- Lombardi, A. M., Cocco, M., and Marzocchi, W. (2010). On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, central Italy, sequence: apparent variation or fluid-driven triggering? *Bulletin of the Seismological Society of America*, 100(3):1138–1152.
- Lombardi, A. M. (2014). Some reasoning on the RELM-CSEP likelihood-based tests. *Earth, Planets and Space*, 66(1):4.
- Marzocchi, W. and Lombardi, A. (2008). A double branching model for earthquake occurrence. *Journal of Geophysical Research*, 113(B08317).
- Marzocchi, W. and Lombardi, A. M. (2009). Real-time forecasting following a damaging earthquake. *Geophysical Research Letters*, 36:L21302.
- Marzocchi, W. and Zhuang, J. (2011). Statistics between mainshocks and foreshocks in Italy and Southern California. *Geophysical Research Letters*, 38:L09310.
- Marzocchi, W., Murru, M., Lombardi, A. M., Falcone, G., and Console, R. (2012). Daily earthquake forecasts during the May-June 2012 Emilia earthquake sequence (northern Italy). *Annals of Geophysics*, 55(4):561–567.
- Murru, M., Console, R., and Falcone, G. (2009). Real time earthquake forecasting in Italy. *Tectonophysics*, 470:214–223.
- Musmeci, F. and Vere-Jones, D. (1986). A variable-grid algorithm for smoothing clustered data. *Biometrics*, 42:483–494.
- Nanjo, K. Z. (2010). Earthquake forecast models for Italy based on the RI algorithm, *Annals of Geophysics*. 53(3):117–127.
- Nanjo, K. Z. (2011). Earthquake forecasts for the CSEP Japan experiment based on the RI algorithm. *Earth Planets and Space*, 63(3):261–274.
- Nanjo, K. Z., Tsuruoka, H., Yokoi, S., Ogata, Y., Falcone, G., Hirata, N., Ishigaki, Y., Jordan, T. H., Kasahara, K., Obara, K., Schorlemmer, D., Shiomi, K., and Zhuang, J. (2012). Predictability study on the aftershock sequence following the 2011 tohoku-oki, japan, earthquake: first results. *Geophysical Journal International*, 191(2):653–658.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83:9–27.
- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50:379–402.
- Ogata, Y., Jones, L. M., and Toda, S. (2003). When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California. *Journal of Geophysical Research*, 108(B62318).
- Ogata, Y. (2004). Space-time model for regional seismicity and detection of crustal stress changes. *Journal of Geophysical Research*, 109(B3):B03308.
- Ogata, Y. and Zhuang, J. (2006). Space-time ETAS models and an improved extension. *Tectonophysics*, 413(1-2):13–23.
- Rebez, A., Sabetta, F., and Slejko, D. (2001). New seismic hazard maps of the Italian territory. Gruppo Nazionale per la Difesa dai Terremoti- Servizio Sismico Nazionale.
- Rhoades, D. A. and Evison, F. F. (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale. *Pure and Applied Geophysics*, 161:47–72.
- Richards-Dinger, K., Stein, R. S., and Toda, S. (2010). Decay of aftershock density with distance does not indicate triggering by dynamic stress. *Nature*, 467:583–586.
- Roberts, G., Michetti, A., Cowie, P., Morewood, N. C., and Papanikolaou, I. (2002). Fault slip-rate variations during crustal-scale strain localisation, central Italy. *Geophysical Research Letters*, 29(8).
- Shi, Y. and Bolt, B. A. (1982). The standard error of the magnitude-frequency b-value. *Bulletin of the Seismological Society of America*, 72:1677–1687.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1st edition.
- Slejko, D., Peruzza, L., and Rebez, A. (1998). Seismic hazards maps of Italy, *Annali di Geofisica*, 41(2):183–214.
- Tsuruoka, H., Hirata, N., Schorlemmer, D., Euchner, F., Nanjo, K. Z., and Jordan, T. H. (2012). CSEP testing

- center and the first results of the earthquake forecast testing experiment in Japan. *Earth Planets Space*.
- Utsu, T. and Seki, A. (1955). A relation between the area of aftershock region and the energy of the mainshock. *Zisin*, 7:233–240.
- Van Stiphout, T., Wiemer, S., and Marzocchi, W. (2010). Are short-term evacuations warranted? case of the 2009 L'Aquila earthquake. *Geophysical Research Letters*, 37:L06306.
- Vere-Jones, D. (1998). Probability and information gain for earthquake forecasting. *Computational Seismology*, 30:248–263.
- Wang, Q., Jackson, D. D., and Zhuang, J. (2010). Missing links in earthquake clustering models. *Geophysical Research Letters*, 37(21):L21307.
- Werner, M. J., Helmstetter, A., Jackson, D. D., and Kagan, Y. Y. (2011). High-resolution long-term and short-term earthquake forecasts for California. *Bulletin of the Seismological Society of America*, 101(4):1630–1648.
- Zechar, J. D. (2010). Evaluating earthquake predictions and earthquake forecasts: a guide for students and new researchers. Community Online Resource for Statistical Seismicity Analysis.
- Zechar, J. D. and Jordan, T. (2010). Simple smoothed seismicity earthquake forecasts for Italy. *Annals of Geophysics*, 53(3):99–105.
- Zhuang, J., Ogata, Y., and Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97(3):369–380.
- Zhuang, J., Ogata, Y., and Vere-Jones, D. (2004). Analyzing earthquake clustering features by using stochastic reconstruction. *Journal of Geophysical Research*, 109(3):B05301.
- Zhuang, J., Chang, C.-P., Ogata, Y., and Chen, Y.-I. (2005). A study on the background and clustering seismicity in the Taiwan region by using a point process model. *Journal of Geophysical Research*, 110:B05S13.
- Zhuang, J. (2006). Second-order residual analysis of spatiotemporal point processes and applications in model evaluation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(4): 635–653.
- Zhuang, J. and Ogata, Y. (2006). Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. *Physical Review E*, 73:046134.
- Zhuang, J., Christophersen, A., Savage, M. K., Vere-Jones, D., Ogata, Y., and Jackson, D. D. (2008). Differences between spontaneous and triggered earthquakes: Their influences on foreshock probabilities. *Journal of Geophysical Research*, 113(B12).
- Zhuang, J. (2010). Gambling scores for earthquake predictions and forecasts. *Geophysical Journal International*, 181:382–390.
- Zhuang, J. (2011). Next-day earthquake forecasts by using the ETAS model. *Earth, Planet, and Space*, 63:207–216.
- Zhuang, J., Harte, D., Werner, M.J. and Hainzl, S., and Zhou, S. (2012). Basic models of seismicity: temporal models. Community Online Resource for Statistical Seismicity Analysis.

*Corresponding author: Jiancang Zhuang,
The Institute of Statistical Mathematics, Tokyo, Japan;
email: zhuangjc@ism.ac.jp.

Appendix A: Original form of ETAS I

Although essentially the same, the conditional intensity function for ETAS I takes a slightly different form from Console et al. (2010a) and Falcone et al. (2010). The original form is

$$\lambda(t, x, y, m) = f_r \cdot \lambda_0(x, y, m) + \sum_{j: t_j < t} \lambda_j(t, x, y, m), \quad (17)$$

where f_r is the fraction of background events over the total number of events. The background seismicity rate is

$$\lambda_0(x, y, m) = \lambda_0(x, y) \beta e^{\beta(m-m_c)}, \quad (18)$$

where $\beta = b \ln 10$, and m_c is the threshold magnitude of the analysis. The triggering kernel is factorized into three terms depending on time, space, and magnitude, as:

$$\lambda_j(t, x, y, m) = K f(x - x_j, y - y_j, m_j) h(t - t_j), \quad (19)$$

where K is a constant parameter, and $f(x, y, m)$ and $h(t)$ represent the space and time distributions, respectively.

The spatial distribution of the triggered seismicity is modeled by a function with circular symmetry around the point of coordinates (x_j, y_j) . This function can be written in polar coordinates (r, θ) as:

$$f(r, \Theta) = \left(\frac{d_0^2 10^{\alpha'(m_j-m_c)}}{r^2 + d_0^2 10^{\alpha'(m_j-m_c)}} \right)^q, \quad (20)$$

where r is the distance from point (x_j, y_j) , q is a free parameter modeling the decay with distance, d_0 is the characteristic triggering distance of an earthquake of magnitude m_0 , and α' is a free parameter describing the productivity for triggering earthquakes. For the time dependence, the Omori-Utsu formula is adopted:

$$h(t) = c^{p-1} (p-1) (t+c)^{-p}, \quad p > 1 \quad (21)$$

where c and p are characteristic parameters of the process.

The set of free parameters for ETAS I actually estimated by Falcone et al. (2010) are:

- (1) K (productivity coefficient),
- (2) d_0 (characteristic triggering distance),
- (3) q (exponent of the spatial distribution of triggered events),
- (4) c (time constant of the Omori-Utsu formula), and
- (5) p (exponent of the Omori-Utsu formula).

In the present study, as well as in Falcone et al. (2010), the value of the α' parameter is fixed to one to reduce the number of free parameters, taking into account that in Falcone et al. (2010) $\alpha' = 1$ because decimal exponentials are used in their algorithm (see Equation (25)). This parameter is the coefficient of the exponential relation between the magnitude of triggering earthquakes and their average triggering distance. The

b -value is assumed to be constant over the geographical area spanned by the catalog, and is estimated independently of the other parameters. In the estimation procedures, an iterative process is applied to ETAS I by Falcone et al. (2010), done in five steps, in order to have a spatial distribution that does not include the triggered component of the seismicity and still preserves the total seismic moment released by the seismicity, which is proportional to the total number of earthquakes.

Comparing Equation (17) to Equation (8), the parameters c and p are the same, $D = d_0^2$, $\alpha' = \alpha \ln 10$, and

$$K = A \frac{q-1}{\pi d_0^2}, \quad A = \frac{K \pi d_0^2}{q-1}. \quad (22)$$

Please note that, since $K d_0^2$ and A are both dimensionless quantities, it is not necessary to change the units of c and d_0 when converting K to A using the above equation.

Appendix B: Parameters and background rate estimation

Zhuang (2011) compiled a summary of the approaches for assessing background rates, and classified them as follows: (1) using a rate proportional to the total seismicity rate of all events (Console and Murru, 2001; Console et al., 2003) or only of the large events in the catalog (Musmeci and Vere-Jones, 1986); (2) using a declustering method to decluster the catalog and using the total rate in the declustered catalog as the background rate (Ogata, 1998; Helmstetter et al., 2006; Werner et al., 2011); (3) weighting each event by the probability that it is a background event (Zhuang et al., 2002, 2004; Console et al., 2010a); and (4) the method introduced by Ogata (2004), which is a Bayesian smoothing procedure on tessellation grid to simultaneously estimate the spatial variation of the background and the model parameter. In this study, the third category of methods was used for models ETAS I and II, but with slight differences.

First, we estimate the background probability through the relative contribution of the background rate to the total seismicity rate, i.e.,

$$\varphi_i = \frac{\mu(x_i, y_i)}{\lambda(t_i, x_i, y_i)} \quad (23)$$

The background rate is estimated by applying a smoothing technique to the whole catalog, with all the events weighted by their corresponding background probabilities.

In ETAS I, a smoothed grid kernel method is used to estimate the background intensity, i.e., the value of $\mu(x, y)$ is computed at each node k of a regular grid using the method introduced by Frankel (1995) and adopted by Console and Murru (2001) and

subsequent works:

$$\tilde{N}_k = \frac{\sum_l \sum_i [\varphi_i I(i \in C_l)] \exp(-\Delta_{kl}^2/d^2)}{\sum_i \exp(-\Delta_{kl}^2/d^2)} \quad (24)$$

where φ_i is the background probability for the i th event, $I(i \in C_l)$ is one when event i is in cell C_l , and zero otherwise, Δ_{kl} is the distance between the centers of cells k and l , and d is the bandwidth. In order to implement Equation (24), one has to select a d value; larger d values will lead to a smoother distribution (at each grid node) and background rate, and to more contamination between different areas. The rate density $\mu(x, y)$ at any point is obtained by linear interpolation among the nearest grid nodes. The bandwidth d is determined using cross-validation procedures, i.e., maximizing the likelihood of the seismicity contained in half of the catalog under the model obtained from the other half.

In ETAS II, we use a variable kernel estimate,

$$\hat{\mu}(x, y) = \frac{1}{T} \sum_i \varphi_i Z_{d_i}(x - x_i, y - y_i), \quad (25)$$

where Z_d is the Gaussian kernel with d as the bandwidth, i goes through all of the events in the whole process, T is the length of the time period of the process, and d_i represents the varying bandwidth calculated for each event i in the following way. Given a suitable integer n_p , find the smallest disk centered at the location of the i -th event which includes at least n_p other earthquakes, and is larger than some minimum value (e.g., a distance within 0.02° , which is of the order of the location error) and let this minimum radius be d_i (e.g., Silverman (1986), Chapter 5).