

Refining earthquake clustering models

Rodolfo Console, Maura Murru, and Anna Maria Lombardi

Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

Received 1 August 2002; revised 3 January 2003; accepted 20 March 2003; published 9 October 2003.

[1] Assuming that earthquakes are the realization of a stochastic point process and that the magnitude distribution of all earthquakes is described by the Gutenberg-Richter law with a constant b value, we model the occurrence rate density of earthquakes in space and time by means of an epidemic model. The occurrence rate density is computed by the sum of two terms, one representing the independent, or spontaneous activity, and the other representing the activity induced by previous earthquakes. While the first term depends only on space, the second one is factored into three terms that include the magnitude, time, and location, respectively, of the past earthquakes. In this paper we use the modified Omori law for the time term, focusing our investigation on the magnitude and space terms. We formulate two different hypotheses for each of them, and we find the respective maximum likelihood parameters on the basis of the catalog of instrumental seismicity recorded in Italy from 1987 to 2000. The comparison of the respective likelihood computed for the seismicity recorded in 2001 provides a way for choosing the best model. The confidence level of our choice is then assessed by means of a Monte Carlo simulation on the various hypotheses. Our study shows that an inverse power density function is more reliable than a normal density function for the space distribution and that the hypothesis of scale invariance of aftershock productivity with respect to magnitude can be rejected with high confidence level. The final model is suitable for computing earthquake occurrence probability in real circumstances.

INDEX TERMS: 7223 Seismology: Seismic hazard assessment and prediction; 7230 Seismology: Seismicity and seismotectonics; 7260 Seismology: Theory and modeling; **KEYWORDS:** earthquake clustering, aftershocks, stochastic processes, hypothesis test, Italian seismicity

Citation: Console, R., M. Murru, and A. M. Lombardi, Refining earthquake clustering models, *J. Geophys. Res.*, 108(B10), 2468, doi:10.1029/2002JB002130, 2003.

1. Introduction

[2] Numerous probabilistic earthquake occurrence models have been developed to describe various aspects of seismic occurrence patterns [see, e.g., *Anagnos and Kiremidjian*, 1988]. At the present there is no simple answer to the question of which is the best model. Anyway an earthquake precursor or forecasting hypothesis should be defined in a quantitative way, so as to allow its testing by means of rigorous statistical procedures [Console, 2001]. These procedures are generally validated assessing the confidence level by which a previous “null” model should be rejected in comparison with a new one. Among several space-time statistical models based on earthquake clustering, *Ogata* [1998], *Jackson and Kagan* [1999], and *Console and Murru* [2001] showed that modeling earthquake clustering as a stochastic process characterized by a limited number of free parameters allows the computation of the occurrence rate density as a continuous function in space and time. Consequently, this formulation allows the estimate of the likelihood function of a seismic catalog under such hypothesis, and the assessment of the maximum likelihood

parameters of the hypothesis [Kagan, 1991]. Another possibility given by such formulation is the computation of the likelihood ratio (performance factor) of two models based on a given set of observations. *Console and Murru* [2001] concluded that the clustering hypothesis exhibits a much higher likelihood with respect to the time-independent Poisson hypothesis and that it can itself be regarded as a sort of more accurate null hypothesis against which different forecasting hypotheses should be tested. In this paper we refine the formulation of the clustering hypothesis and demonstrate the validity of this new formulation with respect to the previous one.

2. Earthquake Clustering Models

[3] Here we give a short outline of the method for modeling the interrelation of any earthquake with any other, referring the reader to the above mentioned papers for details.

[4] Earthquakes are regarded as the realization of a point process, known by statisticians as the Hawkes process. Each event is characterized by its location-time-magnitude parameters (x, y, z, t, m). In the following, we shall neglect the z (depth) coordinate only for sake of simplicity.

[5] For the magnitude density distribution, independent of any other parameter, we assume the validity of the Gutenberg-Richter (G-R) law:

$$\lambda_0(x, y, m) = \mu_0(x, y)\beta^{-(m-m_0)}, \quad (1)$$

where $\lambda_0(x, y, m)$ is the spatial magnitude density distribution of earthquakes; $\mu_0(x, y)$ is the spatial density of earthquakes of magnitude equal to or larger than m_0 , the minimum detectable magnitude; β is a characteristic parameter of each seismogenic area, supposed to be approximately independent of time and of the space coordinates; β is linked to the well-known b parameter by the relation $\beta = b \ln(10)$.

[6] Equation (1) treats aftershocks as if they were independent and identically distributed random events. The choice of m_0 is not critical, provided that the set of data is complete above it.

[7] In the computation of equation (1) we regard the space density distribution $\mu_0(x, y)$ as a continuous, smooth function of the geographical coordinates (x, y) . It can be obtained by a Gaussian smoothing of the catalog, following the method introduced by *Frankel* [1995]. Integrated over x, y and m , equation (1) gives the expected total number of earthquakes in a catalog.

[8] It is widely recognized that the occurrence of a seismic event has influence on the probability of occurrence of new ones. In this paper we assume that this influence is an effect of any earthquake upon all the subsequent ones, as in the epidemic model (ETAS) introduced by *Ogata* [1988, 1998]. The epidemic model represents the seismicity as a superposition of spontaneous and induced events. Hence the expected rate density of earthquakes, taking into account the influence of the previous ones, can be written as

$$\lambda(x, y, t, m) = f_r \lambda_0(x, y, m) + \sum_{i=1}^N H(t - t_i) \lambda_i(x, y, t, m), \quad (2)$$

where f_r is a factor called the “failure rate” (i.e., a measure of the proportion of events that can be considered truly independent and constitute the “spontaneous background seismicity” of a catalog), $\lambda_0(x, y, m)$ is expressed as in equation (1), t_i is the occurrence time of the earthquakes, the total number of which is N , $H(t)$ is the step function such as $H(t) = 0$ for $t \leq 0$ and $H(t) = 1$ for $t > 0$, and $\lambda_i(x, y, t, m)$ is the kernel of the previous earthquakes. The first and the second terms on the right side of equation (2) represent the “independent” and the “induced” seismicity, respectively, in probabilistic terms, without any means of distinguishing the real nature of each individual event. The rate density corresponding to any time-space point is, in general, constituted by the superposition of both components.

[9] We hypothesize that the contribution of any previous earthquake (x_i, y_i, t_i, m_i) to the rate density of the subsequent earthquakes is decomposable (for $t > t_i$) into three terms, representing the time, space and magnitude density distribution, respectively, as

$$\lambda_i(x, y, t, m) = Kh(t - t_i)f(x - x_i, y - y_i)\beta^{\alpha m_i - \beta m}, \quad (3)$$

where K and α are constant parameters, while $h(t)$ and $f(x, y)$ are the time and space distributions, respectively.

Equation (3) implies that aftershocks are basically a nonstationary process.

[10] For the time dependence, we adopt in a general way the so-called modified Omori law [*Utsu*, 1961; *Utsu et al.*, 1995], which is to be applied to any of the subsequent earthquakes with respect to all the previous ones:

$$h(t) = (p - 1)c^{(p-1)}(t + c)^{-p} (p \neq 1), \quad (4)$$

where c and p are characteristic parameters of the process, and the expression is normalized so that $\int_0^\infty h(t)dt = 1$. In this context, equation (4) is used not only for first generation aftershocks but also for multiple generations.

[11] We model the spatial distribution of the induced seismicity by a function $f(x - x_i, y - y_i)$ that has circular symmetry around the point of coordinates (x_i, y_i) and is normalized to 1. We consider two different expressions of this kind. In polar coordinates (r, θ) , the first of these expressions can be written as

$$f(r, \theta) = \frac{1}{2\pi\sigma^2} e^{(-r^2/2\sigma^2)}, \quad (5a)$$

where r is the distance from the point (x_i, y_i) and σ is a free parameter determining the fall off of the induction effect with distance. The second distribution is

$$f(r, \theta) = \frac{(q - 1)}{\pi} \frac{d^{2(q-1)}}{(r^2 + d^2)^q}, \quad (5b)$$

where d and q are two other free parameters of the process. In the right side of both equations (5a) and (5b), θ does not appear explicitly because of the isotropy of the models, but the normalization has been done by integration of this parameter from 0 to 2π . In our study the spatial parameters σ and d are independent of the magnitude of the inducing earthquake. In this respect it can be considered that, in the context of this study, the typical location error of the epicenters (of the order of a few kilometers) is larger than the presumable linear size for most of the sources ($M < 5$) reported in the catalog.

[12] As to the dependence on the magnitude m_i of the previous earthquakes, in equation (3) we introduce a generalization of the *Reasenber and Jones* [1989] hypothesis that is coincident with our algorithm only if $\alpha = \beta$. In this particular case, parameter K has the meaning of the expected number of induced events of magnitude equal to or larger than m_i .

[13] In this study, we consider four formulations of equation (3), obtained by the combinations of the two different choices for the spatial distribution and the two hypotheses $\alpha = \beta$ or $\alpha \neq \beta$: (1) spatial distribution given by equation (5a) and $\alpha = \beta$ (free parameters: K, c, p, σ, β), (2) spatial distribution given by equation (5b) and $\alpha = \beta$ (free parameters K, c, p, d, q, β), (3) spatial distribution given by (5a) and $\alpha \neq \beta$ (free parameters $K, c, p, \sigma, \alpha, \beta$), and (4) spatial distribution given by (5b) and $\alpha \neq \beta$ (free parameters $K, c, p, d, q, \alpha, \beta$).

[14] Case 1 reproduces the hypothesis adopted by *Console and Murru* [2001]; cases 3 and 4 were already examined by *Ogata* [1998]. We have also considered and tested a version of the inverse power decay function [*Ogata*,

1998, equation 2.4], where the parameter d depends on the magnitude m_i of the inducing earthquakes through α . The results for this model are not reported here because they are similar, but not better, of those obtained from our model 4. In the set of free parameters we have not included the failure ratio f_r . This is because f_r is constrained by the condition that the integral over time of $\lambda(x, y, t, m)$ given by equation (2) must be equal to the time-independent density distribution $\lambda_0(x, y, m)$. This condition makes f_r dependent on the other parameters of the model.

[15] Each set of parameters has to be adjusted in order to reach the best fit of the respective model with real observations. The procedure adopted in this study reflects that already introduced by *Ogata* [1998] and used by *Console and Murru* [2001]. It consists in searching for the maximum of the log likelihood function of a realization of seismic events described by a catalog $\{x_j, y_j, t_j, m_j, j = 1, \dots, N\}$:

$$\ln L = \sum_{j=1}^N \ln[\lambda(x_j, y_j, t_j, m_j) V_0] - \int_X \int_Y \int_T \int_M \lambda(x, y, t, m) dx dy dt dm, \quad (6)$$

where V_0 is an arbitrary coefficient whose dimensions are equal to those of the inverse of the rate density $\lambda(x_j, y_j, t_j, m_j)$, and X, Y, T and M denote the range of integration of the four space, time and magnitude variables. Equation (6) is easily implemented in practice, given that the model allows the computation of the density function $\lambda(x, y, t, m)$ in space, time and magnitude. The first term on the right-side member of equation (6) is the sum of the logarithms of density functions computed at the locations of all the observed earthquakes. The second term is the total expected number of earthquakes for the given model. The expected number depends only on the model, not on particular events. Depending on the arbitrary choice of V_0 , the specific value of $\ln L$ does not have any absolute meaning. Only the difference of values computed for the same catalog with different models gives some information [*Kagan and Knopoff*, 1977; *Kagan*, 1991; *Vere-Jones*, 1998]. In the following, the results are given in terms of the ratio (performance factor) between the likelihood of the observations under a given model and the likelihood of the same observations under a reference model [see, e.g., *Evison and Rhoades*, 1997].

3. Application to the Italian Seismicity

[16] The database used in this study is the catalog of shallow seismicity collected by the Istituto Nazionale di Geofisica e Vulcanologia (INGV) (formerly ING) from 1 July 1987 to 31 December 2000. We considered the earthquakes contained in a polygonal area surrounding the Italian coasts and borders (Figure 1), within which the National Seismological Network has provided reliable locations. Our tests showed that this catalog is complete for magnitude equal to 2.0 and larger, including a total of 7851 events (Figure 2). The maximum likelihood value for the b parameter of the Gutenberg-Richter frequency-magnitude relation is 0.997 ± 0.011 , with the error computed using the formula suggested by *Shi and Bolt* [1982].

[17] The space density distribution $\mu_0(x, y)$ of earthquakes was obtained by interpolation from a grid of $100 \times$

120 cells, each having a size of 10×10 km, centered on the point 42°N , 13°E . A smoothed value is computed at each node by the average of the number of earthquakes observed in the neighbor cells, weighting each cell by a Gaussian function decreasing with the distance between its center and the node [*Frankel*, 1995]. The optimal parameter of 20 km for the Gaussian function was obtained by means of the criterion outlined by *Console and Murru* [2001]. It consists in a trial and error search for the maximum likelihood of a subcatalog containing about half of the total number of events, under the model obtained from the other half.

[18] The b (or β) parameter was not included in the sets of those obtained by the best fit procedure. We decided to keep its value fixed ($b = 0.997$), as it was obtained by the analysis of the entire catalog. This choice is rigorously justified only for models 3 and 4, but not exactly for models 1 and 2 because in these latter cases, b controls not only the overall magnitude distribution but also the proportion between spontaneous and induced events.

[19] The actual number of free parameters to be estimated is 4, 5, 5, and 6 for the different models described as cases 1, 2, 3, and 4 respectively. The maximum likelihood estimation of these parameters, was performed by a Newtonian estimation procedure described by *Console and Murru* [2001]. The final values of the parameters obtained by this procedures together with the respective values of the failure rate f_r and of the performance factors relative to the time-independent Poisson process, are reported in Table 1.

[20] The maximum likelihood deriving from the best fit denotes the better performance of models 2 and 3 (five free parameters) with respect to model 1 (four free parameters), and the overall best performance of model 4 (six free parameters).

4. Prospective Test of the Hypotheses

[21] The difference between the various values of the log likelihood ratios reported in Table 1 is of the order of hundreds. In light of the Akaike information criterion [*Akaike*, 1974; *Sakamoto et al.*, 1983], one could easily conclude that this difference is highly significant even taking into account the different number of free parameters in the models. However, we follow the opinion that an objective test of any forecasting hypothesis should be carried out with no free parameters and on a data set independent of that used for the formulation of the hypothesis itself [*Rhoades and Evison*, 1989]. In this work, we carry out a prospective test on the more recent data set collected by the Italian Seismological Network, from 1 January to 31 December 2001. This data set consists of 322 epicenters of shallow earthquakes with magnitude larger than or equal to 2.0, within the same geographical area selected for the best fit of the four models. Figure 3 shows the epicentral map of these earthquakes. It must be noted that the number of 322 earthquakes is smaller than the average rate of 582 earthquakes per year characterizing the whole 1987–2000 data set.

[22] Having fixed the parameters of the four models on the values given in Table 1, the likelihood of the 2001 observations was computed under each of the hypotheses

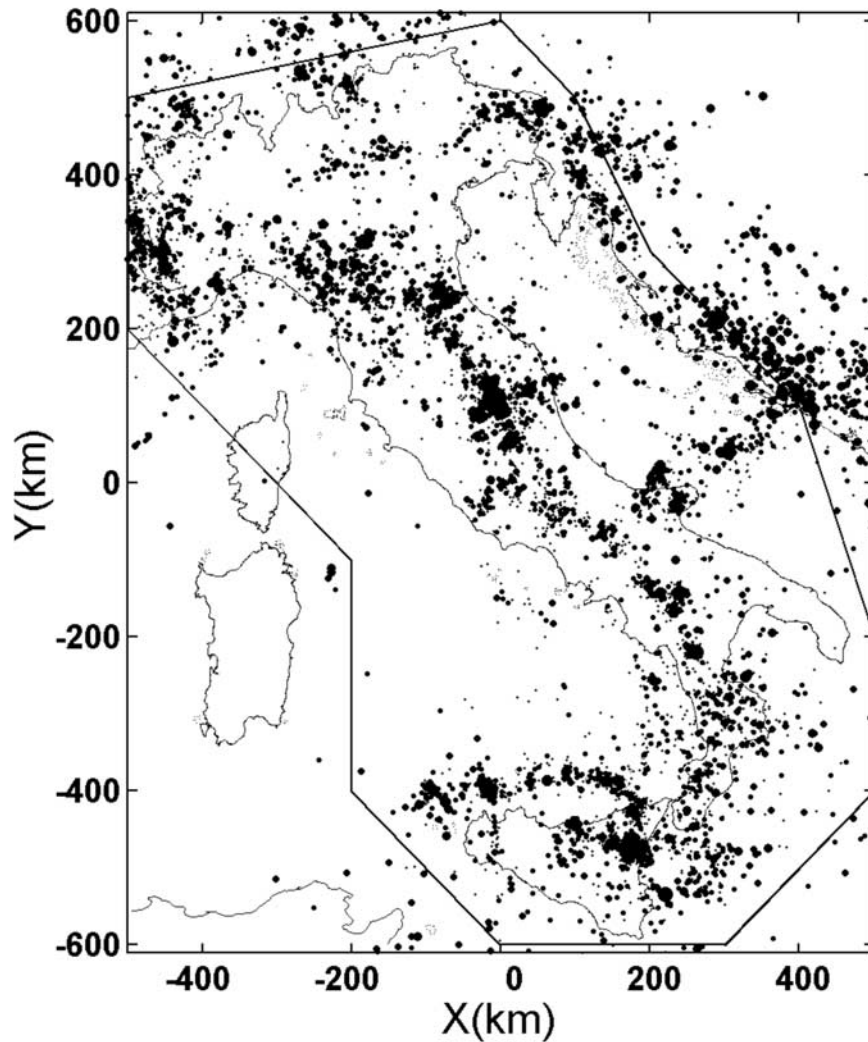


Figure 1. Epicenters of the earthquakes ($M \geq 2.0$ and $h \leq 70$ km) located by the National Seismological Network in Italy and surrounding areas from July 1987 to December 2000. The origin of the rectangular coordinates is in the point 42°N , 13°E . The size of the symbols is scaled with magnitude from 2.0 to 5.5. The rectangular coordinates (in km) of the corners of the polygonal area are $-500, 500$; $0, 600$; $100, 500$; $200, 300$; $400, 100$; $500, -200$; $500, -400$; $300, -600$; $0, -600$; $-200, -400$; $-200, -100$; $-400, 100$; $-500, 200$.

and divided by the likelihood obtained under the Poisson model. The results are given in Table 2.

[23] The likelihood ratios reported in Table 2 have been computed without taking into account the effect of earthquakes occurred prior to 2001. However, no earthquake larger than magnitude 4.5 occurred in the study area in 2000.

[24] The results of this test are in agreement with those obtained from the best fit procedure, because the log likelihood increases from model 1 to model 4. Table 3 shows a comparison within pairs of models in terms of their relative performance factors. The comparison indicates that models 2 and 4 (inverse power space distribution) compared respectively with models 1 and 3 (Gaussian space distribu-

tion) give a performance factor of the order of 5×10^6 and 1.5×10^9 . Moreover, models 3 and 4 ($\alpha \neq \beta$) give a performance factor of the order of 6×10^{30} and 2×10^{33} compared with models 1 and 2, respectively ($\alpha = \beta$).

[25] Are these results statistically significant? This can be answered by means of Monte Carlo simulations [Evison and Rhoades, 1997]. For instance, for the first case of Table 3, we built up 1000 synthetic catalogs under model 1 and computed the likelihood of both models 1 and 2 for each of them. The simulation procedure used in this test follows the steps proposed by Ogata [1998]. The result is that not even in one of these 1000 simulations the performance factor is larger than the observed one. As shown in Table 3, there is less than 1% of chances that we

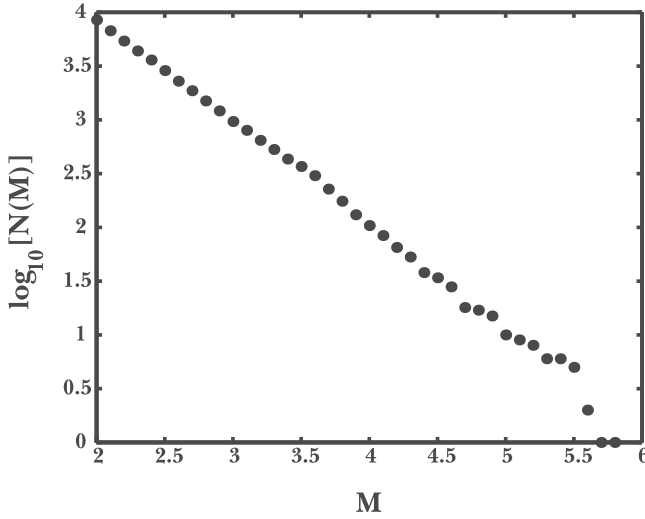


Figure 2. Cumulative frequency-magnitude plot of the earthquakes shown in the map of Figure 1.

make the mistake of rejecting the true model for the wrong one in any of these tests.

5. Discussion

[26] The hypothesis adopted in this study is centered on the assumption that all the earthquakes in the catalog belong to a unique family of events characterized by the same b value, and their magnitude is chosen randomly from the Gutenberg-Richter magnitude-frequency distribution [Felzer *et al.*, 2002]. The second important piece of the hypothesis is that all earthquakes interact through an epidemic process modeled by the magnitude of previous events and by the space and time distance from the previous earthquakes separately. In this way, no distinction is made between different modes of earthquake clustering. Definitions as foreshocks, aftershocks, main shocks, swarms, etc., are not necessary in this context. We mention some of them only for shortness of language. For instance, foreshocks can be regarded just as events that triggered a subsequent event that happened to be bigger than the triggering event [Shaw, 1993; Felzer *et al.*, 2002]. The observations carried out in Italy for 14 years provide the database used for the best fit of the free parameters characterizing four different models. One year of independent data (recorded in 2001) appears a suitable duration for validating one of the models with respect to the others.

[27] First, we notice that, in agreement with Cao *et al.* [1996], Ogata [1998], and Kagan and Jackson [2000], the hyperbolic law expressed as in equation (5b) fits the data better than the normal distribution equation (5a) previously used by Console and Murru [2001]. For both distributions we obtained similar values of the characteristic parameters σ and d (ranging from 3 to 4 km: close to the order of magnitude of the typical location errors for the events reported in the catalog). The initial choice of the normal distribution had been made because it includes just one free parameter (σ). However, it can be recalled that the form of equation (5b) with $q = 1.5$ would be in agreement with the stress falloff by $1/r^3$ expected from a physical model of

stress change caused by a point source or a circular crack in an elastic medium [Shaw, 1993; Dieterich, 1994]. We may notice also how the two distribution equations (5a) and (5b), once the proper values of the free parameters of Table 1 are introduced, have a bell shape fairly similar to each other (Figure 4). Nevertheless, the adoption of the latter allows a significant improvement of the likelihood. As previously mentioned, the introduction of the alternative inverse power model [Ogata, 1998 equation 2.4], found the best fitting for Japanese data sets, does not produce any significant improvement in the case if the Italian data. A possible explanation of this apparent discrepancy could be found in the different magnitude thresholds used in the analysis for these two countries: ranging from 4.0 and 6.0 for Japan and only 2.0 for Italy. So, the location error dominates the short-range behavior in the Italian catalogue, while the real source size makes the dependence on magnitude relevant for the Japanese data.

[28] An even more substantial improvement of the model has been achieved by having let α and β assume different values in equation (3). It has been observed that α , mostly ranging between 0.2 and 3.0, is a measure of the efficiency of a shock with a certain magnitude in generating its induced activity [Utsu *et al.*, 1995]. A swarm-like activity is characterized by a small α value, while a typical main shock-aftershock activity is related to a large value of this parameter. As noted above, this is in contrast with the Reasenberg and Jones [1989] hypothesis of scale invariance, followed also by Felzer *et al.* [2002]. As the value of α fitting the Italian data is approximately half of the correspondent value of β , this difference is not negligible. It implies that the production of events induced by a shock of magnitude one unit larger than another, rather than about ten times larger, is only slightly more than three times (e^α) larger than those induced by the smaller earthquake. Taking into account the magnitude distribution of earthquakes, we may infer that small earthquakes have a substantial role in the triggering process. Therefore the late activity in an aftershock sequence, besides being produced by the main shock only, is largely due to aftershocks having occurred earlier. This fact is depicted in Table 4, showing the expected number of events induced by an earthquake of given magnitude occurred at time $t = 0$, within 15 km and different time intervals. The computation has been carried out using the parameters of model 4 in Table 4 with the

Table 1. Optimal Parameters of the Models for Earthquake Clustering^a

Parameter	Model 1	Model 2	Model 3	Model 4
K	0.0742	0.0768	0.0294	0.0316
c , days	0.0103	0.00994	0.00684	0.00680
p	1.102	1.103	1.054	1.058
σ , km	3.85	-	3.54	-
d , km	-	3.66	-	3.07
q	-	1.945	-	1.828
α	-	-	1.009	0.974
f_r	0.495	0.473	0.409	0.382
$\ln(L/L_0)$	15,138.45	15,600.35	16,826.15	17,398.85

^aINGV catalog 1 July 1987 to 31 December 2000, Nev.7851; average rate is 582 events per year, $M_c \geq 2.0$.

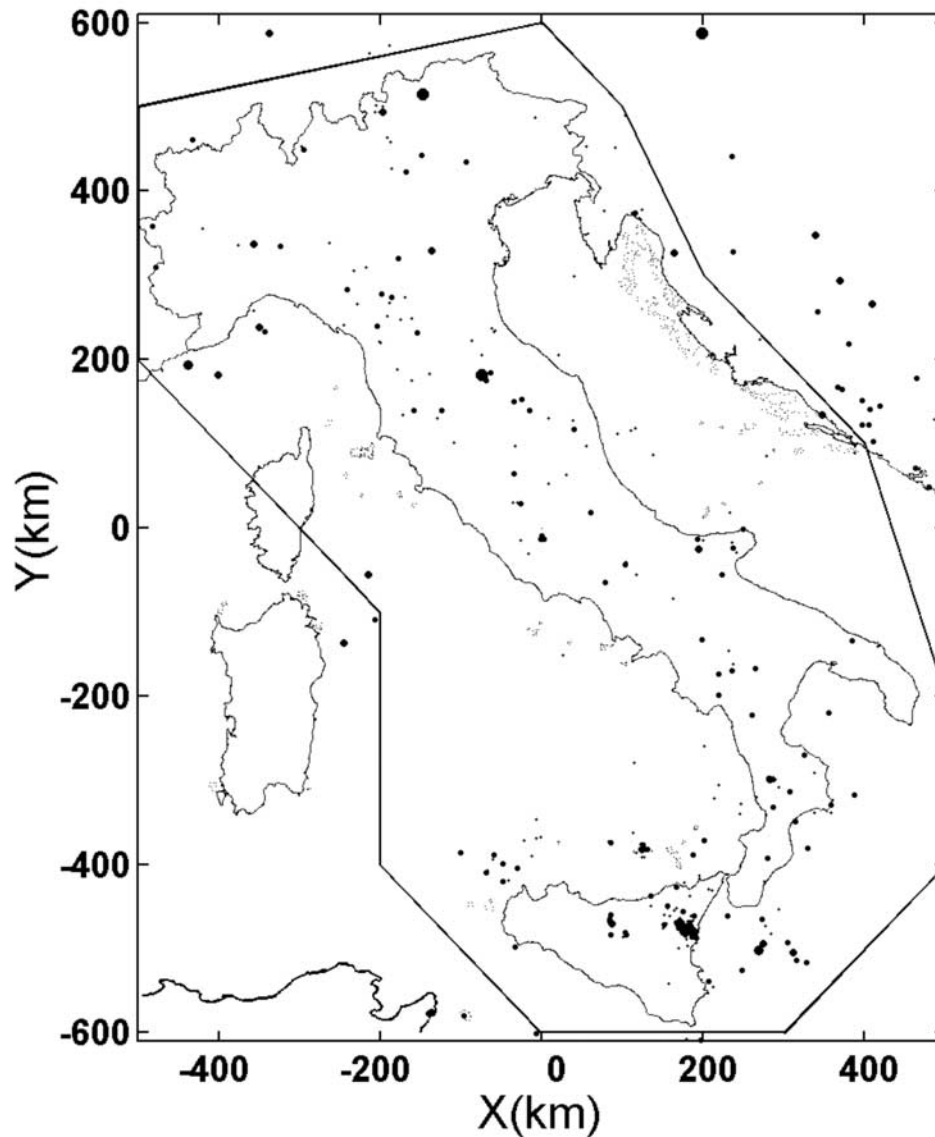


Figure 3. Epicenters of the earthquakes ($M \geq 2.0$ and $h \leq 70$ km) located by the National Seismological Network in Italy and surrounding areas from January 2001 to December 2001. The origin of the rectangular coordinates, the polygonal area, and the size of the symbols are the same as in Figure 1.

background seismicity estimated as the average of central Italy. The results represent the average of numerous independent simulations. The values reported in Table 4a have been estimated limiting the computation of the rate density only to the events induced by the main shock, modifying the models used, e.g., by *Reasenber and Jones* [1989] and *Console et al.* [1999] with the inclusion of the independent α parameter. This causes a decrease of probability of induced events with magnitude. This proba-

bility would be independent of magnitude for $\alpha = \beta$. Table 4b, obtained by a epidemic model with the influence of all the aftershocks on subsequent aftershocks, shows the substantial effect of secondary induced events on the resulting seismicity. The comparison of the results in Table 4a with the corresponding results in Table 4b shows that 100 days after the main shock, about 50% of the aftershock activity is imputable to a secondary effect of previous aftershocks [*Felzer et al.*, 2002].

Table 2. Results of the Test on the 2001 Data

Model	$\ln(L/L_0)$
1	510.83
2	526.21
3	581.71
4	602.86

Table 3. Logarithm of the Performance Factor for the Tests on the 2001 Data

	Observed Performance Factor	Significance Level, %
$\ln L_2 - \ln L_1$	15.38	<0.1
$\ln L_4 - \ln L_3$	21.15	<0.1
$\ln L_3 - \ln L_1$	70.88	0.6
$\ln L_4 - \ln L_2$	76.65	0.4

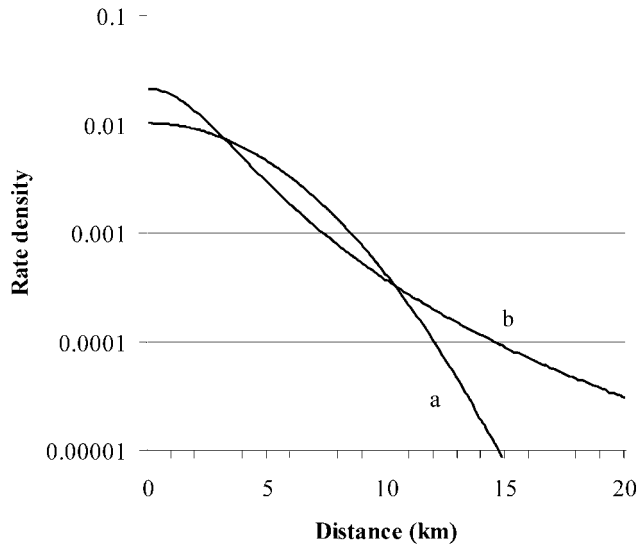


Figure 4. Dependence of the rate density of induced earthquakes versus the distance from the epicenter of the inducing event. Curve a is for normal distribution given in equation (5a), and curve b is for the inverse power law given in equation (5b).

[29] Our result is in agreement with those found by Ogata [1998] for moderate earthquakes in central and western Honshu Island (Japan). A physical interpretation might be that the productivity of induced shocks is proportional to the linear dimensions of the fault that ruptured during the inducing event, rather than to its total area. This is consistent with the idea that most aftershocks are produced in the high stress region surrounding a fault that is outside the edges of the fault, rather than inside it (A. Helmstetter, Stress triggering driven by small earthquakes?, submitted to *Geophysical Research Letters*, 2002).

[30] As to the time dependence of seismic activity, it seems that the modified Omori law expressed in equation (4) is well fitted by the data with c of the order of 0.01 days and p close to 1.1. The value of c is close to the minimum in the range of those reported by previous investigations [see, e.g., Utsu *et al.*, 1995]. In connection with this point, we make the comment that here c is computed from a model taking into account the influence of each single event on the following ones, while often it is derived by the analysis of the whole aftershock sequence related to a single main shock. Moreover, it has been noted that the determination of an accurate c value is made difficult by the problem of

Table 4a. Probability of Occurrence of an Earthquake of Magnitude Equal to or Exceeding the Magnitude of Another Earthquake of Magnitude M Within 15 km Distance and a Time Interval T After This Earthquake, Without Taking Into Account the Effect of Subsequent Events

T	$M = 3$	$M = 4$	$M = 5$
1	0.0440	0.0148	0.0049
10	0.0616	0.0204	0.0068
100	0.0940	0.0271	0.0086

Table 4b. Same as Table 4a but With the Probabilities Computed Also Taking Into Account the Induction by Subsequent Events

T	$M = 3$	$M = 4$	$M = 5$
1	0.0587	0.0202	0.0067
10	0.0928	0.0317	0.0109
100	0.1468	0.0476	0.0164

incomplete detection of small aftershocks shortly after the main shock [Utsu *et al.*, 1995].

[31] With regard to p , our determination is consistent with a number of previous observations. It usually falls in the range 0.9–1.8 [Utsu *et al.*, 1995]. A value close to 1 is justified by a theoretical approach [Shaw, 1993; Dieterich, 1994]. In our model $p < 1$ is unrealistic, because it would lead to the instability of the process. However, Sornette and Sornette [1999] demonstrated that it is possible to observe $p < 1$ (over short periods of time) as a result of looking at combined secondary aftershocks (with all of the direct sequences following Omori's law with $p > 1$).

6. Conclusions

[32] We found that the stochastic model of earthquake clustering expressed by equations (3), (4), and (5b), partly supported by theoretical arguments, provides a fit of the data better than other models considered here. A notable result concerns the dependence of the aftershock productivity on the magnitude of the inducing event that is expressed through an exponential law in which the magnitude is multiplied by a coefficient α significantly smaller than the β parameter of the G-R law. With the acceptance of model 4 against the other three considered in this study (Tables 2 and 3), still maintaining the concept of earthquake self-similarity in magnitude distribution, we abandon the simple idea that the rate of aftershocks is determined only by the magnitude difference between the main shock and its aftershocks [e.g., Kagan and Knopoff, 1987; Reasenber and Jones, 1989; Davis and Frohlich, 1991; Console and Murru, 2001; Felzer *et al.*, 2002].

[33] In order to compare how the model fits the experimental data, we show in Figure 5 a comparison made on the real seismic sequence occurred in central Italy few years ago. This sequence has been characterized by multiple main shocks and migration of activity. Figure 5a shows the number of earthquakes with magnitude equal to or larger than 2.0 observed in a circular area of 15 km radius centered on the epicenter of the $M5.6$ main shock occurred on 26 September 1997, in intervals of 12 hours. Figure 5b shows the expected number of earthquakes with the same magnitude and within the same radius and time intervals. It has been computed by model 4 with the maximum likelihood parameters of Table 1. The computation has been carried out dividing each time interval of 12 hours in 10 steps of 1.2 hours, and estimating the theoretical occurrence rate density at the beginning of each time step, making use of the data included in the previous steps. The total number of observed events is larger than the theoretical values. This can be ascribed to the fact that our

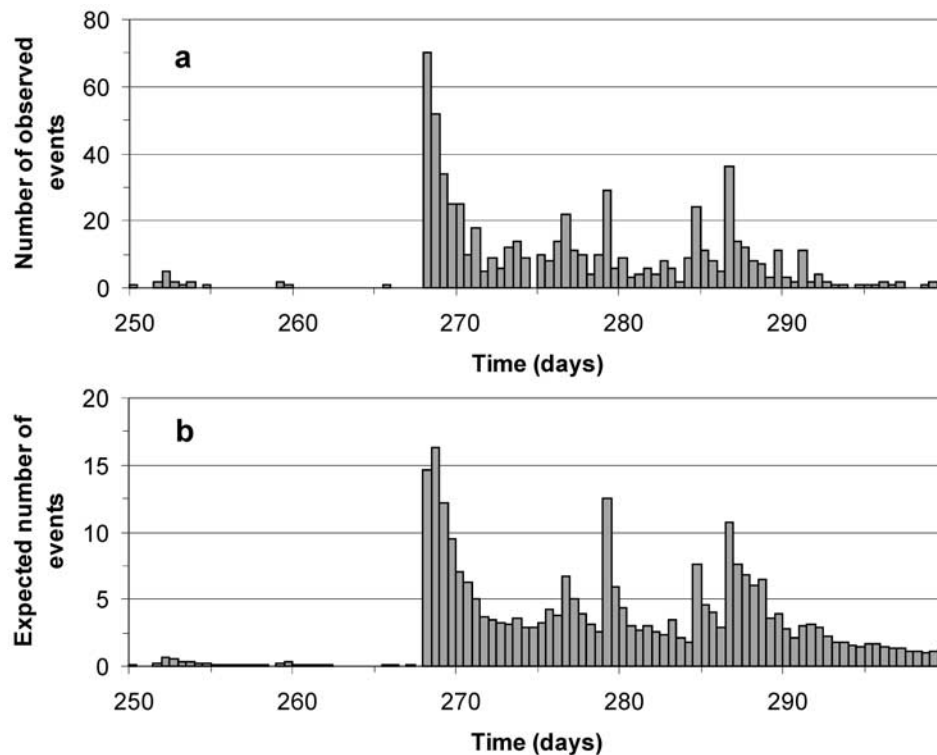


Figure 5. (a) Histogram of the number of events with magnitude equal to or larger than 2.0 observed in a circle of 15 km radius centered on the epicenter of the $M_{5.6}$, 26 September 1997 (day 268) earthquake in central Apennines, Italy (43.020°N , 12.910°E); time is counted in days after 1 January 1997; the duration of each time interval is 12 hours. (b) Histogram of the expected number of events according to model 4 in the same magnitude range, place, and time interval of Figure 5a.

model (and in particular the K value), fitted on the whole seismicity of Italy, underestimates the real number of induced events for single episodes characterized by an explosive aftershock activity. Taking into account this circumstance, the qualitative trends of the two plots can be judged quite similar.

[34] The model considered in this study is subject to improvements that should take into account regional variations of the free parameters and should eventually allow the adjustment of the parameters on the basis of real time observations. Nevertheless, we retain that the methodology described in this paper is practicable for a possible application in earthquake risk mitigation.

[35] **Acknowledgments.** The authors are grateful to Joan Gomberg, Karen Felzer, and Frank Evison for reading a preliminary version of this paper and providing valuable comments and suggestions. We wish also to thank Yan Kagan and an anonymous referee for their thorough and helpful reviews. David Jackson has provided useful suggestions in the preparation of the revised version of the paper.

References

- Akaike, H., A new look at the statistical model identification, *IEEE Trans. Automat. Control*, **AC-19**, 716–722, 1974.
- Anagnos, T., and A. S. Kiremidjian, A review of earthquake occurrence models for seismic hazard analysis, *Probab. Eng. Mech.*, **3**, 3–11, 1988.
- Cao, T. Q., M. D. Petersen, and M. S. Reichle, Seismic hazard estimate from background seismicity in southern California, *Bull. Seismol. Soc. Am.*, **86**, 1372–1381, 1996.
- Console, R., Testing earthquake forecasting hypotheses, *Tectonophysics*, **338**, 261–268, 2001.
- Console, R., and M. Murru, A simple and testable model for earthquake clustering, *J. Geophys. Res.*, **106**, 8699–8711, 2001.
- Console, R., F. Di Luccio, M. Murru, M. Imoto, and G. Stavrakakis, Short term and short range seismicity patterns in different seismic areas of the world, *Nat. Hazards*, **19**, 107–121, 1999.
- Davis, S. D., and C. Frohlich, Single-link cluster analysis, synthetic earthquake catalogues, and aftershock identification, *Geophys. J. Int.*, **104**, 289–306, 1991.
- Dieterich, J., A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, **99**, 2001–2618, 1994.
- Evison, F. F., and D. A. Rhoades, The precursory earthquake swarm in New Zealand: Hypothesis test II, *N. Z. J. Geol. Geophys.*, **40**, 537–547, 1997.
- Felzer, K. R., T. W. Becker, R. E. Abercrombie, G. Ekstrom, and J. R. Rice, Triggering of the 1999 M_w 7.1 Hector Mine earthquake by aftershocks of the 1992 M_w 7.3 Landers earthquake, *J. Geophys. Res.*, **107**(B9), 2190, doi:10.1029/2001JB000911, 2002.
- Frankel, A., Mapping seismic hazard in the central and eastern United States, *Seismol. Res. Lett.*, **66**, 8–21, 1995.
- Jackson, D. D., and Y. Y. Kagan, Testable earthquake forecasts for 1999, *Seismol. Res. Lett.*, **70**, 393–403, 1999.
- Kagan, Y. Y., Likelihood analysis of earthquake catalogues, *Geophys. J. Int.*, **106**, 135–148, 1991.
- Kagan, Y. Y., and D. D. Jackson, Probabilistic forecasting of earthquakes, *Geophys. J. Int.*, **143**, 438–453, 2000.
- Kagan, Y. Y., and L. Knopoff, Earthquake risk prediction as a stochastic process, *Phys. Earth Planet. Inter.*, **14**, 97–108, 1977.
- Kagan, Y. Y., and L. Knopoff, Statistical short-term earthquake prediction, *Science*, **236**, 1467–1563, 1987.
- Ogata, Y., Statistical models for earthquake occurrence and residual analysis for point processes, *J. Am. Stat. Assoc.*, **83**, 9–27, 1988.
- Ogata, Y., Space-time point-process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, **50**, 379–402, 1998.
- Reasenber, P. A., and L. M. Jones, Earthquake hazard after a mainshock in California, *Science*, **243**, 1173–1176, 1989.
- Rhoades, D. A., and F. F. Evison, On the reliability of precursors, *Phys. Earth Planet. Inter.*, **58**, 137–140, 1989.

- Sakamoto, Y., M. Ishiguro, and G. Kitagawa, *Akaike Information Criterion Statistics*, 290 pp., D. Reidel, Norwell, Mass., 1983.
- Shaw, B. E., Generalized Omori law for aftershocks and foreshocks from a simple dynamics, *Geophys. Res. Lett.*, 20, 907–910, 1993.
- Shi, Y., and A. Bolt, The standard error of the magnitude-frequency b value, *Bull. Seismol. Soc. Am.*, 72, 1677–1687, 1982.
- Sornette, D., and A. Sornette, Renormalization of earthquake aftershocks, *Geophys. Res. Lett.*, 26, 1981–1984, 1999.
- Utsu, T., A statistical study on the occurrence of aftershocks, *Geophys. Mag.*, 30, 521–605, 1961.
- Utsu, T., Y. Ogata, and R. S. Matsu'ura, The centenary of the Omori formula for a decay law of aftershock activity, *J. Phys. Earth*, 43, 1–33, 1995.
- Vere-Jones, D., Probabilities and information gain for earthquake forecasting, *Comput. Seismol.*, 30, 248–263, 1998.
-
- R. Console, A. M. Lombardi, and M. Murru, Istituto Nazionale di Geofisica e Vulcanologia, Via di Vigna Murata, 605, I-00143 Rome, Italy. (console@ingv.it; lombardi@ingv.it; murru@ingv.it)