AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Solid Earth

RESEARCH ARTICLE
10.1002/2015JB012398

Key Points:
- Magnitude conditional probability density function
- Increasing trend of the averages of triggered events' magnitude

Correspondence to:
I. Spassiani,
spassianiilaria@gmail.com

Citation:
Spassiani, I., and G. Sebastiani (2016), Exploring the relationship between the magnitudes of seismic events, *J. Geophys. Res. Solid Earth*, *121*, 903–916, doi:10.1002/2015JB012398.

Received 14 AUG 2015
Accepted 17 JAN 2016
Accepted article online 25 JAN 2016
Published online 18 FEB 2016

Exploring the relationship between the magnitudes of seismic events

Ilaria Spassiani[1] and Giovanni Sebastiani[1,2]

[1]Department of Mathematics Guido Castelnuovo, Sapienza University of Rome, Rome, Italy, [2]Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche, Rome, Italy

Abstract The distribution of the magnitudes of seismic events is generally assumed to be independent on past seismicity. However, by considering events in causal relation, for example, mother-daughter, it seems natural to assume that the magnitude of a daughter event is conditionally dependent on one of the corresponding mother events. In order to find experimental evidence supporting this hypothesis, we analyze different catalogs, both real and simulated, in two different ways. From each catalog, we obtain the law of the magnitude of the triggered events by kernel density. The results obtained show that the distribution density of the magnitude of the triggered events varies with the magnitude of their corresponding mother events. As the intuition suggests, an increase of the magnitude of the mother events induces an increase of the probability of having "high" values of the magnitude of the triggered events. In addition, we see a statistically significant increasing linear dependence of the magnitude means.

1. Introduction

The Epidemic Type Aftershock Sequences (ETAS) is a well-known model in seismology [*Ogata*, 1988, 1998, 1989, 1999]. It is a branching process in which, at each generation, each event may produce its own offsprings independently of the others. The magnitudes of events are mutually statistically independent and are distributed according to the *Gutenberg-Richter law* [*Gutenberg and Richter*, 1944]. The distribution density function of the magnitude m of an event is given by $p(m) = b \ln 10 \cdot 10^{-b(m-m_0)} = \beta e^{-\beta(m-m_0)}$, where $\beta = b \ln 10$ is a positive constant and the cutoff parameter m_0 is known as the *completeness magnitude*. This parameter is estimated from data in such a way that magnitudes exceeding it are statistically in good agreement with the Gutenberg-Richter law. This law is assumed to be valid both for the background events and for the triggered ones and is independent on the magnitude of the corresponding ancestors [*Zhuang et al.*, 2002]. Recall that the background seismicity is the component not triggered by precursory events and is usually connected to the regional tectonic strain rate; on the other hand, the triggered seismicity is the one associated with stress perturbations due to previous shocks [*Lombardi et al.*, 2010].

In this work we will use the term "triggering event" to indicate any mother event that produces its own progeny [*Saichev and Sornette*, 2008]. Both a triggered and a background event may be a triggering shock. We want to investigate the distribution of the magnitude of the triggered events, in order to assess its variation with the magnitude of the corresponding triggering ones. In fact, even if the correlation between subsequent events is very difficult to be detected and therefore is usually assumed to be absent [*Helmstetter et al.*, 2006; *Corral*, 2005], in some recent works it was found statistically different from zero [*Lippiello et al.*, 2007a, 2007b, 2008; *Sarlis et al.*, 2009, 2010]. We expect that the probability of having "high" values of the magnitude of the triggered events increases (decreases) with the increase (decrease) of the magnitude of the triggering shocks. Moreover, if this is true, the expected value of the magnitude of the triggered events should also be increasing in the same way.

2. Materials and Methods

In order to obtain experimental evidence to support the above mentioned hypothesis of magnitude correlation, we perform two different kinds of analysis of three Italian seismic catalogs, a Californian catalog, and some simulated ones. The real data catalogs used here are the following.

1. The first catalog includes events occurring from 16 April 2005 to 25 January 2012 in the whole Italy. The estimated value for the completeness magnitude is 2.5.

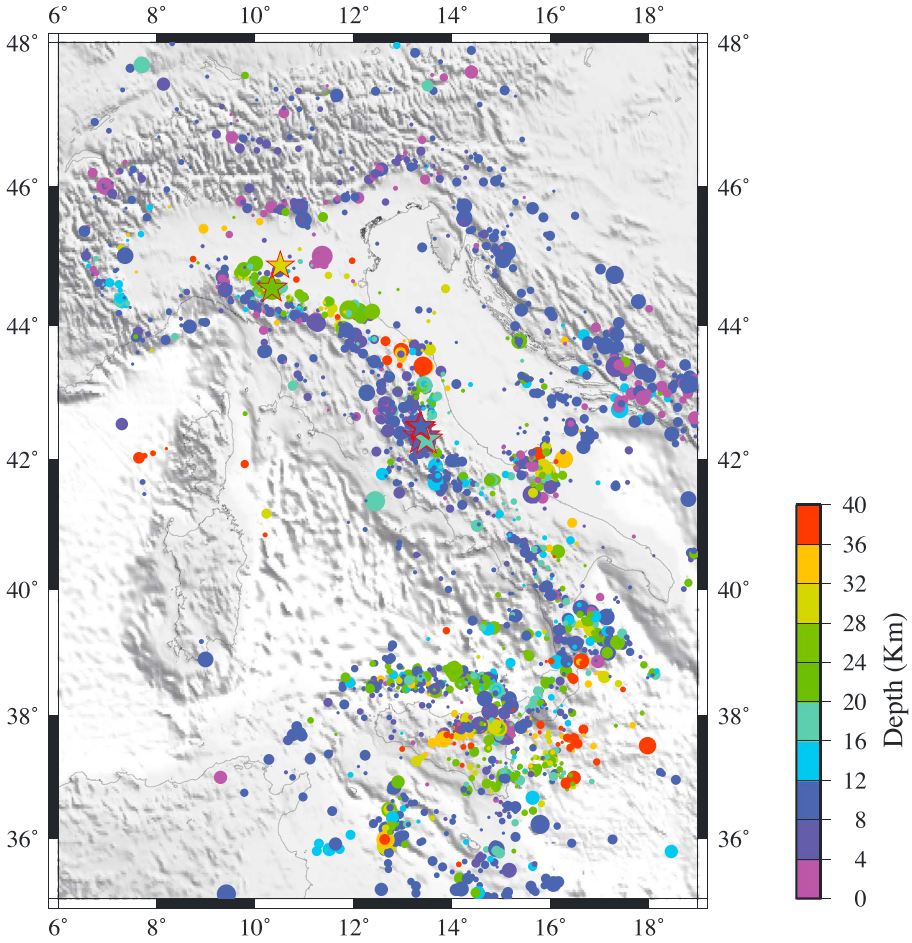

Figure 1. Seismicity map of the first catalog concerning the whole Italy.

2. The second catalog includes events occurring in the portion of Abruzzo region (Italy) corresponding to the square from latitude +41.866 to +42.866 and from longitude +12.8944 to +13.8944. This subregion includes L'Aquila. The temporal interval is the same as that for catalog one. The estimated value for the completeness magnitude is 1.8. This catalog includes the strong shock of magnitude 6.1 occurring in L'Aquila on 6 April 2009.
3. The third catalog differs from the previous one only for the temporal interval, which now goes from 16 April 2005 to 5 April 2009. The value of completeness magnitude is estimated to be equal to 1.5. This catalog does not include a strong shock.
4. The fourth catalog includes events occurring from 1 January 1984 to 31 December 1991, in the portion of the Southern California corresponding to the square from latitude +33.75 to +34.75 and from longitude −117.5 to −116.5. The estimated completeness magnitude is 2. This catalog is a portion of the waveform earthquake catalog relocated by Hauksson et al. in 2011 [*Hauksson et al.*, 2012; *Lin et al.*, 2007].

In all the previous cases the maximum depth considered is set equal to 40 km.

The values of the completeness magnitude have been computed with the ZMAP software, using Shi and Bolt uncertainty [*Shi and Bolt*, 1982].

The seismicity maps of the previous four catalogs are shown in Figures 1–4, respectively. They have been obtained by the Generic Mapping Tools.

In order to find empirical evidence for the hypothesis that the distribution of the magnitude of the triggered events depends on the magnitude of their own triggering shocks, it seems appropriate not to include in the catalog events those that are spatially "too" distant to each other. This is the reason for including catalogs 2–4 in the study. More precisely, we consider the third catalog to investigate the influence of a strong shock and

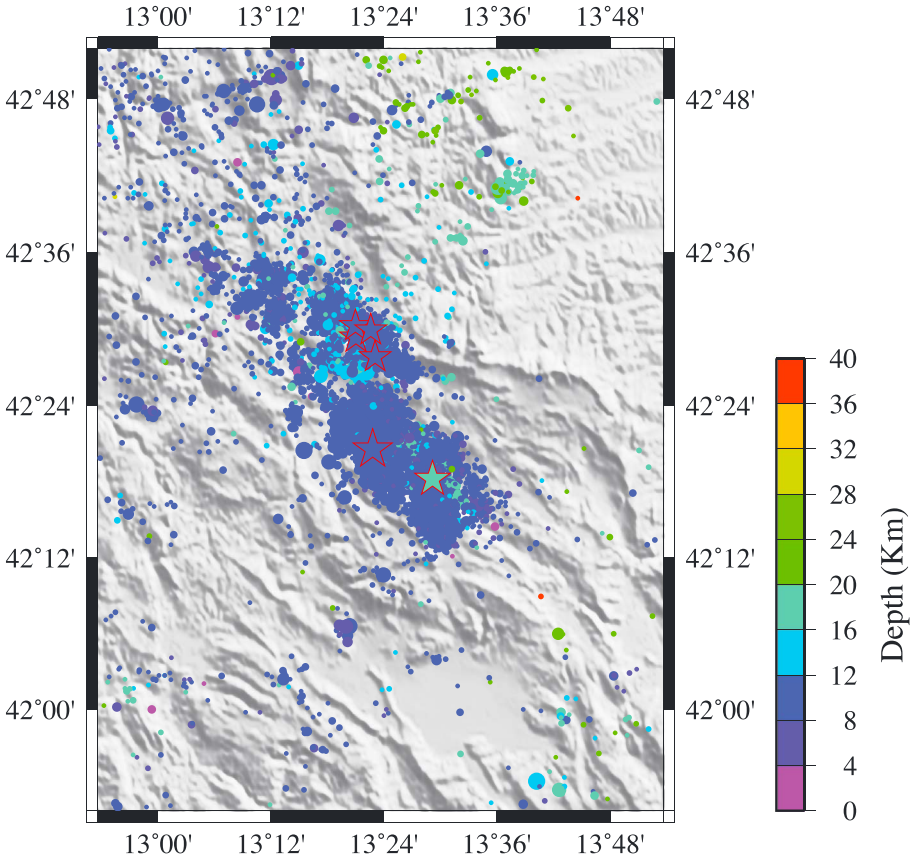

Figure 2. Seismicity map of the second catalog concerning the square centered at L'Aquila, till 2012.

the fourth one to verify the validity of our hypothesis also for a catalog relative to a region far away from the others and with a different seismicity. Instead, in the first catalog there is the presence of events that are very distant to each other, as can be seen from the mean values of the distances between the events considered by us in causal relation, shown below. We expect that in this case the above mentioned dependence becomes far less evident or absent when a maximum distance between events in causal relation is not included in the analysis.

The simulated catalogs have been considered, instead, to demonstrate that our hypothesis is true also for pure simulated catalogs not affected by any kind of "real effect" that may influence the analysis. The results reported here are referred to two simulated catalogs obtained as follows. The first one is simulated with the FORTRAN program [**etasim.f**] written by Ogata [*Ogata*, 1998, 1981, 2006], just modified by the fact that the number of events is random between the two input starting and ending times. The input parameters for the simulation in this catalog are $\beta = b \ln 10 = 1 \ln 10 = 2.3$ and $(\mu, \kappa, c, a, p) = (0.54945, 0.022324, 0.014172, 1.6943, 1.0868)$, where μ is the constant rate of the background component, c and p are the parameters of the *Omori-Utsu law* $\phi(t) = (t + c)^{-p}$, and κ and a are the ones of the *productivity law* $\rho(m) = \kappa e^{a(m-m_0)}$. Furthermore, we have chosen here the option of simulating the magnitudes with the Gutenberg-Richter law, instead of taking them from a given catalog. Then, we expect that there is no evidence of our hypothesis of conditioning.

On the other hand, we simulate the second synthetic catalog with a program written by us very similar to that of Ogata but adapted to our hypothesis of conditioning. More precisely, we simulate the magnitudes of background events again with the Gutenberg-Richter law, while the magnitudes of the triggered events with a new conditional probability density function with respect to the magnitudes of the triggering events. Since we expect that when the magnitude m' of the triggering events increases, the probability of having events with high magnitudes must increase, and at the same time, the ones with low magnitudes must decrease; we propose the following probability density function:

$$p(m|m') = \beta e^{-\beta(m-m_0)} \left[1 + C_1 \left(1 - 2e^{-(\beta-a)(m'-m_0)}\right)\left(1 - 2e^{-\beta(m-m_0)}\right)\right], \quad (1)$$

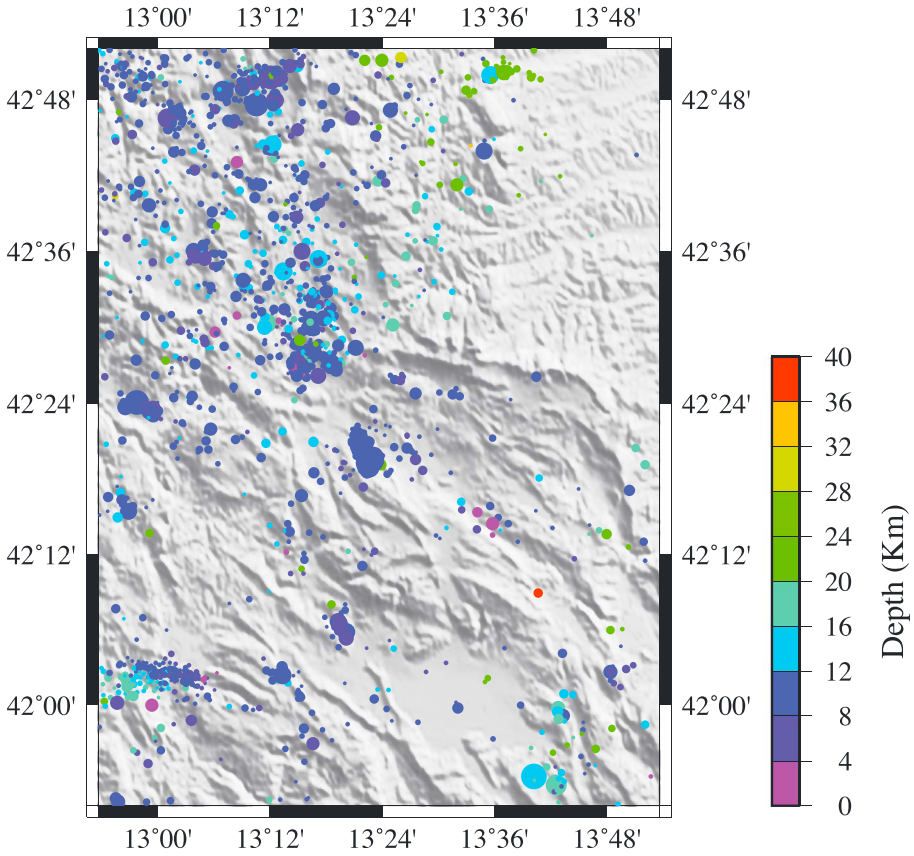

Figure 3. Seismicity map of the third catalog concerning the square centered at L'Aquila, till 5 April 2009.

where β and a are the parameters of the Gutenberg-Richter law and the productivity law, respectively, and $C_1 \in [0, 1]$. We notice that this model reduces to the Gutenberg-Richter one for $C_1 = 0$. For the theoretical motivation and properties about this formula, see *Spassiani and Sebastiani* [2015]. We just want to notice that if $m' \leq \frac{1}{\beta - a} \ln 2 + m_0$, the density function $p(m|m')$ is always decreasing in m. If instead $m' > \frac{1}{\beta - a} \ln 2 + m_0$, the above cited density increases in m till a certain maximum, reached in m_*, such that

$$4q(m')e^{-\beta(m_* - m_0)} = 1 + q(m') \quad \Leftrightarrow \quad m_* = \frac{1}{\beta} \ln \frac{4q(m')}{1 + q(m')} + m_0,$$

and for values larger than m_* it decreases. This trend is shown in Figure 5a. Instead, in Figures 5b and 5c, we can see how the conditional probability density function varies with the parameter C_1. Finally, the input parameters for the simulation of the second synthetic catalog are $\beta = b \ln 10 = 1 \ln 10 = 2.3$, $C_1 = 0.9$ and $(\mu, \kappa, c, a, p) = (0.54945, 0.022324, 0.014172, 0.8, 1.0868)$.

In both the two simulated catalogs, we consider the same minimum magnitude value 1.5 and a null learning (precursory) period.

For each catalog, in the two types of analysis we consider four magnitude subintervals contained in the magnitude range, from the completeness value m_0 to the maximum one $m_{\max}$. The first and the last subintervals considered are of the kinds $[m_0, \bar{m}_1]$ and $[\bar{m}_2, m_{\max}]$, for two certain values of $\bar{m}_1$ and $\bar{m}_2$; the two intermediate ones are instead opportunely chosen between the two above subintervals in order to contain a sufficient number of events. However, it is important to note that the results for the subintervals not shown are consistent with the ones illustrated here. The amplitude of each subinterval is opportunely chosen for each catalog and for each analysis in such a way to have a comparable number of triggered events in all the subintervals considered. The two kinds of analysis put events in causal relation in two different ways, and then, for each analysis, the opportunely chosen magnitude subintervals and amplitudes are different.

We then proceed differently for the two types of analysis. In the first approach, we group the events whose magnitude belongs to each of the four subintervals above. We now consider the events of any given

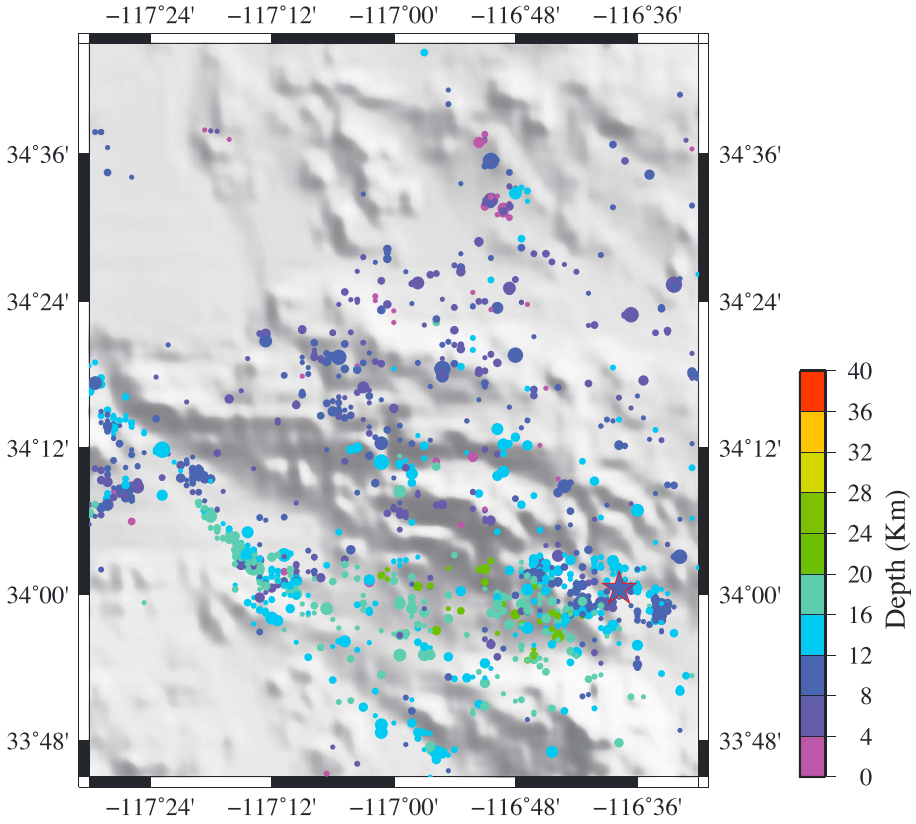

Figure 4. Seismicity map of the fourth catalog concerning the Southern California.

subinterval. For each event $Ev_{1,i}$ in the first subinterval, with magnitude $m_{1,i}$ and time of occurrence $t_{1,i}$, we find all the shocks occurring in the time interval $[t_{1,i}, t_{1,i}+\delta^*]$. The choice of the amplitude δ^* is explained later. Then, we group all the magnitudes $m_{1,i,j}$ of all the shocks belonging to all the previous time intervals and we repeat the reasoning for the other three magnitude subintervals considered. Based on the set of magnitudes $G_m = (m_{1,i,j}, m_{2,i,j}, m_{3,i,j}, m_{4,i,j})$, we estimate a probability density function using kernel density estimation [*Rosenblatt*, 1956]. This is a nonparametric method very often used in statistical analysis. More precisely, we consider the magnitudes $(m_1, m_2, \dots, m_n)$ and the frequencies $f = (f_1, f_2, \dots, f_n)$ of G_m. We then consider the set m of 1000 magnitudes equispaced from the completeness magnitude to the maximum one, and we compute the *kernel density estimator* of the empirical magnitude distribution M as

$$\hat{M}_\gamma(m) = \frac{1}{F}\sum_{i=1}^{n} f_i K\left(\frac{m-m_i}{\gamma}\right), \quad \text{with} \quad F = \sum_{i=1}^{n} K\left(\frac{m-m_i}{\gamma}\right), \tag{2}$$

where $K(\cdot)$ is known as *kernel* and the positive parameter γ is the *bandwidth* [*Parzen*, 1962]. The above formula is obtained by adapting to our case the *Nadaraya-Watson* kernel for kernel regression [*Scott*, 1992].

Let us be reminded that a kernel is a nonnegative, real-valued function such that

$$\int_{-\infty}^{\infty} K(x)\mathrm{d}x = 1, \quad \text{and} \quad K(x) = K(-x) \quad \forall x \in \mathbb{R}.$$

As very often done, here we use the Gaussian kernel

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

The value of the bandwidth is chosen using the leave-one-out cross-validation method, opportunely implemented by us [*Scott*, 1992]. More precisely, we consider the value that minimizes the quantity $\sum_{i=1}^{n} |\hat{f}_i - f_i|$, where

$$\hat{f}_i = \frac{1}{\bar{F}_i}\sum_{j\neq i} f_j K\left(\frac{m_i-m_j}{\gamma}\right), \quad \text{with} \quad \bar{F}_i = \sum_{j\neq i} K\left(\frac{m_i-m_j}{\gamma}\right).$$

We observe that differently from formula (2), the value f_i does not contribute to $\hat{f}_i$.

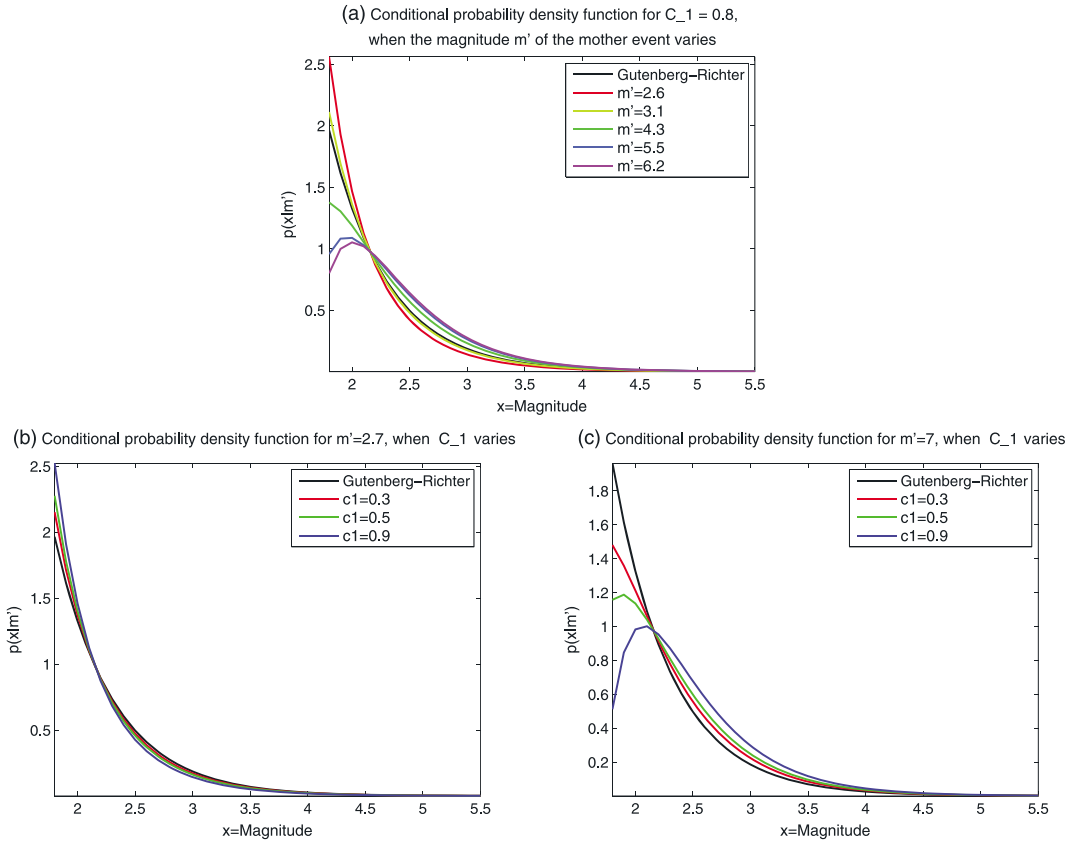

Figure 5. Conditional probability density function $p(x|m')$ relative to the magnitude of the triggered events. (a) The plot is obtained when the magnitude of the triggering events varies, with the set of parameters $(\beta, m_0, C_1, a) = (1.96, 1.8, 0.8, 1.5)$; in this case it holds $\frac{1}{\beta - a} \ln 2 + m_0 = 3.3$. Instead, the two plots are obtained for different values of the parameter C_1 and with fixed (b) $m' = 2.7$ and (c) $m' = 7$.

The time window δ^* is chosen here in such a way that two seismic events separated by a time larger than δ^* are not in causal relation. In order to determine this value, for each catalog we divide the whole time interval in daily subintervals. We then count the number of events that occur in each subinterval. Starting from this temporal sequence, denoted by X_t, we then compute an estimate $\widehat{R}(\delta)$ of the autocorrelation function at different integer values of the time lag δ:

$$\widehat{R}(\delta) = \frac{1}{(n-\delta)\widehat{V}} \sum_{t=1}^{n-\delta} (X_t - \widehat{\mu})(X_{t+\delta} - \widehat{\mu}),$$

where n is the dimension of the sample, $\delta = 0, 1, 2, \ldots$ and $\widehat{\mu} = \frac{1}{n} \sum_{t=1}^{n} X_t$ and $\widehat{V} = \frac{1}{n-1} \sum_{t=1}^{n} (X_t - \widehat{\mu})^2$ are the sample mean and variance, respectively [*Priestley*, 1982].

Then, we model $\widehat{R}(\delta)$ by a power law model containing two parameters. These parameters are estimated by least squares. Finally, we find the value δ^* such that for lag values larger than δ^*, the model is less than 5×10^{-2}. We notice that in the cases examined, this choice produces p values that are always smaller than 0.01.

Due to the strong shock on 6 April 2009, the second catalog shows a clear nonstationary pattern. Then, for this catalog we transform the original data set by considering the well-known random time change:

$$\int_0^t \lambda(x) \mathrm{d}x, \tag{3}$$

where $\lambda(\cdot)$ is the ETAS rate for seismic events. The parameters of the above formula have been computed with the FORTRAN program [**etas.f**] written by Ogata, by which statistical inference on the parameters of the ETAS model is performed [*Ogata*, 2006]. By this transformation, the process becomes stationary.

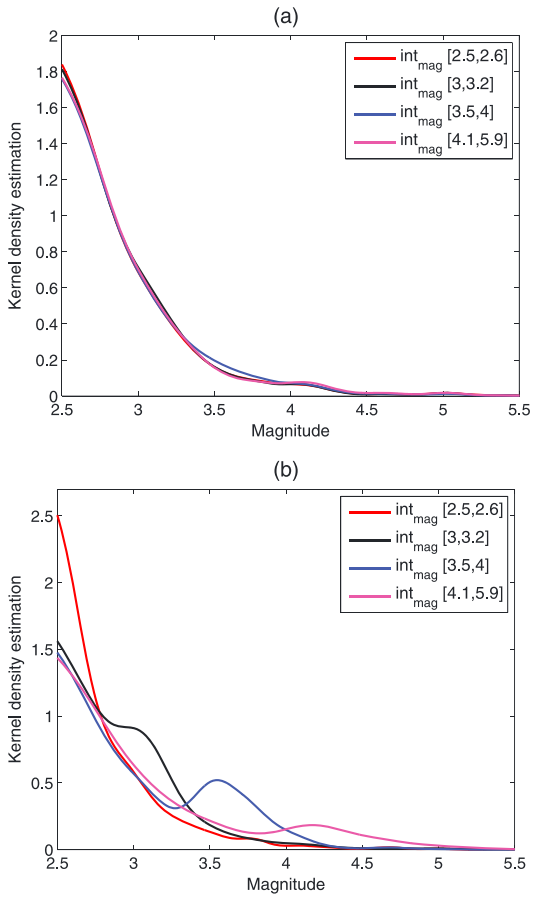

Figure 6. (a, b) Kernel density estimation of the magnitude of the triggered events concerning the first type of analysis of the whole Italian catalog. In both Figures 6a and 6b, the considered intervals in which the magnitudes of the triggering events fall are the following: [2.5, 2.6], [3, 3.2], [3.5, 4], and [4.1, 5.9] (the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 15 days for both the two plots. Figure 6a is obtained without considering a maximum distance between events in causal relation; instead, Figure 6b is the one resulting when we suppose that the generic event occurring in $(t, \text{lat}, \text{lon})$ is put in causal relation with the events belonging to the time-space window $[t, t+15] \times [\text{lat} - \gamma^*, \text{lat} + \gamma^*] \times [\text{lon} - \gamma^*, \text{lon} + \gamma^*]$, where $\gamma^* = 0.025$. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.12, 0.11, 0.12, and 0.12 in Figure 6a and 0.09, 0.13, 0.12, and 0.18 in Figure 6b.

In the second type of analysis we proceed as follows. At first we apply the same Ogata's program as above to estimate the parameters. For each event of the catalog, we then find the mother shock that most likely triggered it, by a variation of Ogata's criterion. More precisely, the latter considers as mother of the ith event the shock occurring in the smallest time t_j such that the ratio between the ETAS rate till J and the ETAS rate over all the previous $t_j < t_i$ is bigger than a uniform random number U generated in $(0, 1]$; instead, we consider as mother the preceding event that gives the highest contribution to the ETAS rate. After that, we consider the magnitude subintervals strategy as in the first type of analysis. For each of them, we group the triggered events that have a triggering shock with magnitude belonging to the considered subinterval. As in the first kind of analysis, for each subinterval we estimate the probability density function relative to the magnitude of the triggered events by using the Gaussian kernel density estimation method described above. The value of the bandwidth is determined as before.

We do not apply this analysis to the first catalog. In fact, we are performing a solely temporal analysis and it is not meaningful to use the pure temporal ETAS model for such a large region like the one in this catalog.

We implemented both methods of analysis in the MATLAB language.

3. Results

We present here some results obtained by the two types of the above mentioned analysis for the catalogs considered.

In Figure 6 we show the results obtained for the whole Italian catalog with the first type of analysis: as we have said, the second type of analysis is not performed in this case. By looking at the estimated densities of the magnitude of the triggered events in Figure 6a, corresponding to four magnitude subintervals opportunely chosen here, we can see no apparent differences among the densities. As said before, this can be explained by the fact that there are many pairs of events that are close to each other along time but spatially very separated. The elements of these pairs are erroneously put in relation in the analysis. In this case, the mean distance between the events of the pairs in causal relation is 140 km. Instead, if we include in the analysis a maximum distance parameter, the differences between the magnitude densities become evident, as one can see in Figure 6b. The latter has in fact been obtained by putting in causal relation the generic event $\text{Ev}_{k,i}$ in the kth magnitude subinterval (with magnitude $m_{k,i}$, time of occurrence $t_{k,i}$, latitude $\text{lat}_{k,i}$, and longitude $\text{lon}_{k,i}$), with all the events in the time-space window $[t_{k,i}, t_{k,i} + \delta^*] \times [\text{lat}_{k,i} - \gamma^*, \text{lat}_{k,i} + \gamma^*] \times [\text{lon}_{k,i} - \gamma^*, \text{lon}_{k,i} + \gamma^*]$, where we have arbitrary chosen $\gamma^* = 0.025$. Then, the first evidence of our hypothesis of magnitude correlation has been obtained.

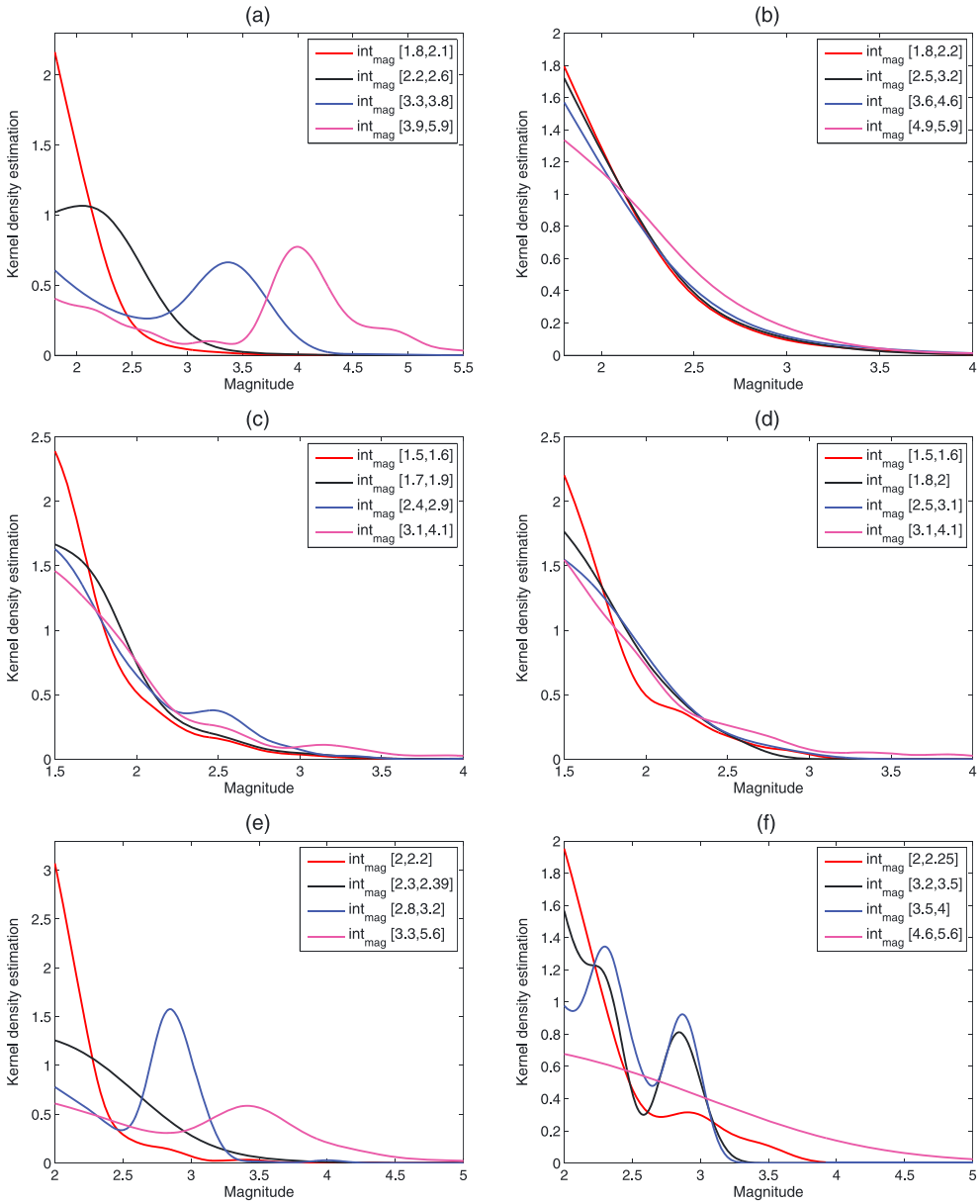

Figure 7. Kernel density estimation of the magnitude of the triggered events concerning the last three real catalogs in the case of the first (a, c, e) and the second (b, d, f) types of analysis. In Figures 7a and 7b, the results concern the second catalog (L'Aquila till 2012). The considered intervals in which the magnitudes of the triggering events fall are the following. First type of analysis, Figure 7a: [1.8, 2.1], [2.2, 2.6], [3.3, 3.8], and [3.9, 5.9]; second type of analysis, Figure 7b: [1.8, 2.2], [2.5, 3.2], [3.6, 4.6], and [4.9, 5.9] (in both cases, the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 1 day. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.22, 0.33, 0.28, and 0.19 in Figure 7a and 0.25, 0.26, 0.31, and 0.27 in Figure 7b. In Figures 7c and 7d, the results concern the third catalog (L'Aquila till 5 April 2009). The considered intervals in which the magnitudes of the triggering events fall are the following. First type of analysis, Figure 7c: [1.5, 1.6], [1.7, 1.9], [2.4, 2.9], and [3.1, 4.1]; second type of analysis, Figure 7d: [1.5, 1.6], [1.8, 2], [2.5, 3.1], and [3.1, 4.1] (in both cases, the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 4 days. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.11, 0.14, 0.11, and 0.15 in Figure 7c and 0.14, 0.16, 0.22, and 0.15 in Figure 7d. In Figures 7e and 7f, the results concern the fourth catalog (Southern California). The considered intervals in which the magnitudes of the triggering events fall are the following. First type of analysis, Figure 7e: [2, 2.2], [2.3, 2.39], [2.8, 3.2], and [3.3, 5.6]; second type of analysis, Figure 7f: [2, 2.25], [3.2, 3.5], [3.5, 4], and [4.6, 5.6] (in both cases, the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 1 day. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.11, 0.44, 0.12, and 0.25 in Figure 7e and 0.18, 0.13, 0.1, and 1 in Figure 7f.

Table 1. Results of the Tests Obtained With the Residual Analysis[a]

	Events Expected	Events Without Learning	Runs Test	Kolmogorov-Smirnov Test
Catalog 1	6138.27	6229	4.48E−001	5.02E−002
Catalog 2	1435.4	1466	4.64E−003	8.84E−001
Catalog 3	1418.21	1428	2.46E−001	2.99E−001
Catalog 4	3121.46	3144	6.03E−001	6.74E−001
Catalog 5	2390.14	2387	1.32E−001	8.7E−001

[a]Catalogs from one to five refer to L'Aquila till 2012, L'Aquila till 5 April 2009, Southern California, the catalog simulated with Ogata's model, and the one simulated with our new model, respectively.

In Figure 7, there are the estimated densities of the magnitude of the triggered events in the other real catalogs, obtained by the first (Figures 7a, 7c, and 7e) and the second (Figures 7b, 7d, and 7f) types of analysis. In Figures 7a and 7b, we show the results relative to the second catalog (L'Aquila till 2012). In this case, the spatial extension of the region analyzed is far smaller than the one of the whole Italian catalog and then we have not included a maximum distance between pairs in causal relation to be considered in the analysis. We recall that in this case, the first type of analysis has been applied to the data set transformed by the random time change (3), due to the nonstationary pattern of the process caused by the presence of the strong shock on 6 April 2009. The means of the distances between the events of the pairs are about 7 km and 13 km for the first and the second types of analysis, respectively. From the first type of analysis (Figure 7a), we notice that the increase of the referential magnitude corresponds to a qualitative variation of the density in agreement with our hypothesis. In fact, there is the increase (decrease) of the density for high (low) values of the magnitude. The results for the second type of analysis (Figure 7b) show the same qualitative variations. The learning period, chosen to estimate the parameters, ends at the time of the last event occurring on 5 April 2009. We get $(\mu, \kappa, c, a, p) = (0.304, 0.06, 0.104, 1.57, 1.39)$. In order to test the goodness of fit of the model with the set of parameters used, we have performed the residual analysis, according to *Ogata* [1998] (see Table 1). The results of the Kolmogorov-Smirnov and Runs tests show that the model slightly underestimates the number of target events. Furthermore, only the Runs test gives back a probability bigger than 5%. The fact that this is not true for the Kolmogorov-Smirnov test, even if the probability has a value very little lower than 5%, could be explained by the value of the completeness magnitude used for this catalog. As explained before, due to the very high number of events occurring during the days just after the strong shock on 6 April 2009, several events of magnitude 1.8 and a bit more have not been recorded. This means that the real completeness value for these days is higher than 1.8. Nevertheless, in order to use, for all catalogs, the same method for the estimation of the completeness magnitudes, we have decided to use the value estimated with ZMAP, which is exactly 1.8. Since the Runs test gives good results and considering that this real catalog has a very strong shock that implies an unusual amount of data to be recorded for the Italian case, we think that the set of parameters chosen can be considered good to be used in the analysis.

In Figures 7c and 7d, there are the results concerning the third catalog. In this case the temporal period is shorter than that for the second one and ends the day before the strong shock on 6 April 2009. The means of the distances between the events of the pairs are 23 km and 17 km for the first and the second types of analysis, respectively. Due to the absence of a very strong shock in this catalog, the parameters used here are the averages over all the sets of parameters obtained by setting the learning period to 7%, 8%,… , 20%. We get $(\mu, \kappa, c, a, p) = (0.5893, 0.0219, 0.0151, 1.6521, 1.1186)$. The parameter values corresponding to the different precursory periods show small variations from the above means. The results are qualitatively very close to those of Figures 7a and 7b. This shows that our hypothesis of dependence is not related to the presence of a strong shock. Furthermore, recalling also that the completeness magnitude is smaller for the third catalog, we can conclude, according to *Lippiello et al.* [2012], that our hypothesis is not even connected to the incompleteness of the catalog, as instead proposed by *Corral* [2006]. From the residual analysis (see Table 1), we can see that the Runs test does not give a good result: the interevents between transformed times are not independent. This means that small values of interevent times follow small values. This could be due to the presence in the catalog of sequences near in time but spatially quite separated or of sequences not completely included in the area under examination, as can be seen in the seismicity map of Figure 3. Instead, the Kolmogorov-Smirnov test gives back a good result; in fact, in this case the probability is bigger that 5%. Concluding with the residual analysis, the number of expected events is very similar to that of the real target

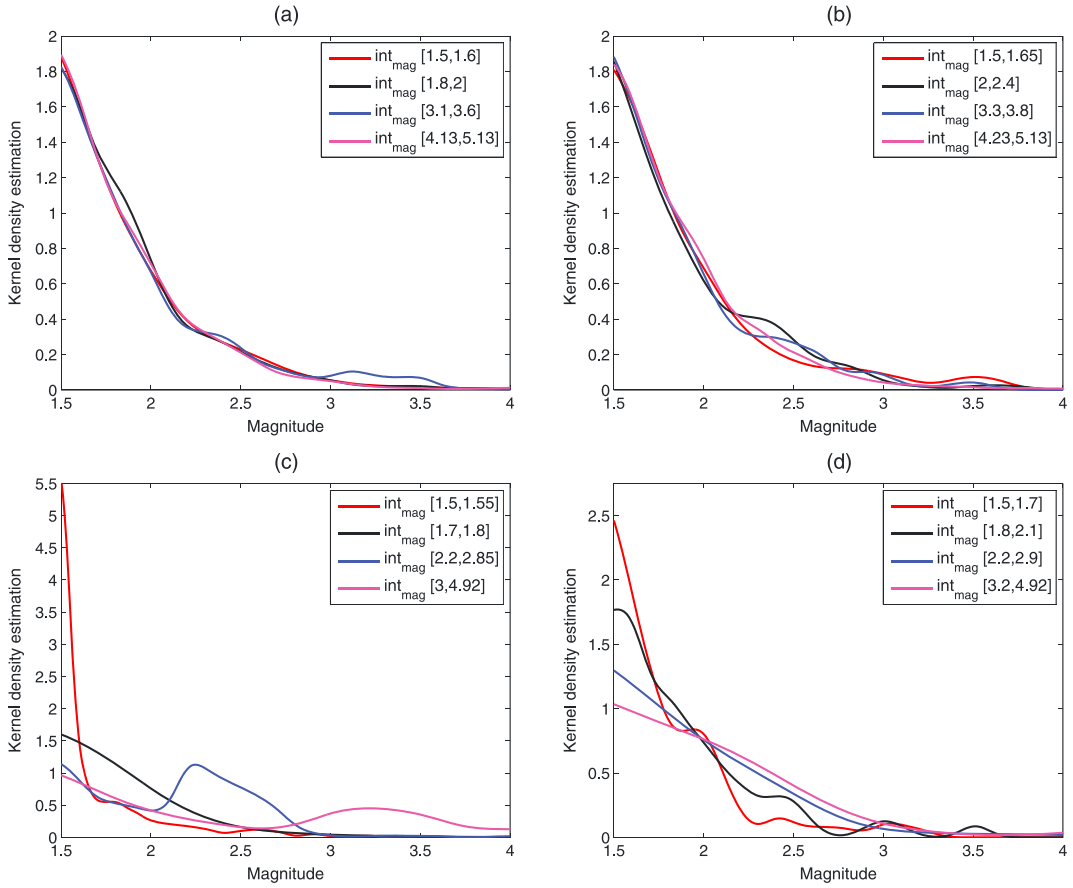

Figure 8. Kernel density estimation of the magnitude of the triggered events concerning the simulated catalogs in the case of the first (a,c) and the second (b,d) types of analysis. In Figures 8a and 8b, the results concern the catalog simulated with the classical Ogata's model. The considered intervals in which the magnitudes of the triggering events fall are the following. First type of analysis, Figure 8a: [1.5, 1.6], [1.8, 2], [3.1, 3.6], and [4.13, 5.13]; second type of analysis, Figure 8b: [1.5, 1.65], [2, 2.4], [3.3, 3.8], and [4.23, 5.13] (in both cases, the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 7 days. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.14, 0.09, 0.08, and 0.11 in Figure 8a and 0.12, 0.11, 0.1, and 0.1 in Figure 8b. In Figures 8c and 8d, the results concern the catalog simulated with our conditional model. The considered intervals in which the magnitudes of the triggering events fall are the following. First type of analysis, Figure 8c: [1.5, 1.55], [1.7, 1.8], [2.2, 2.85], and [3, 4.92]; second type of analysis, Figure 8d: [1.5, 1.7], [1.8, 2.1], [2.2, 2.9], and [3.2, 4.92] (in both cases, the curves are red, black, blue, and magenta, respectively). The δ^* value is equal to 1 day. The optimal bandwidth value for the Normal kernel density estimation is, respectively, for the four intervals considered, equal to 0.05, 0.34, 0.07, and 0.18 in Figure 8c and 0.08, 0.08, 0.26, and 0.28 in Figure 8d.

shocks. All things evaluated, we think that the model with the set of the mean parameters of above can be considered good to fit the real case.

In Figures 7e and 7f, we can see the results relative to the fourth catalog, that is, the Californian one. Here the mean of the distances is 13 km for both the two types of analysis. Both from Figures 7e and 7f, obtained respectively with the first and the second types of analysis, we get results in agreement with the above behaviors. It follows that even if we analyze the events of a region in another continent, the hypothesis is still supported by the results of the two types of analysis. The parameters are here again obtained by averaging over the sets estimated for a learning period fixed at 7%, 8%,..., 20%. We get $(\mu, \kappa, c, a, p) = (0.3729, 0.0116, 0.0002, 0.8579, 0.8879)$. Again, the values obtained for the different learning periods are close to these mean values. The results of the residual analysis are shown in Table 1. Both the Kolmogorov-Smirnov and the Runs tests give back a probability bigger than 5%. The number of expected events is also close to the real number of target events. We deduce the goodness of fit of the model with the set of the mean parameters above.

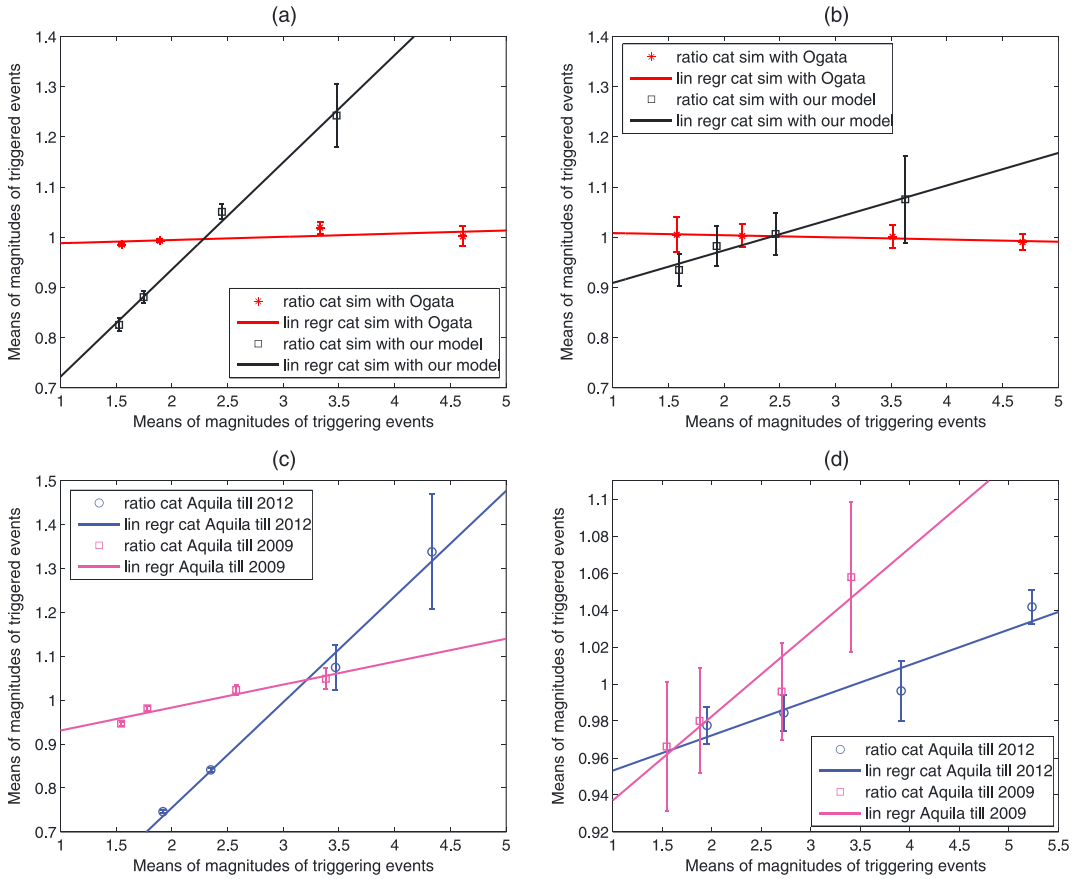

Figure 9. (a–d) Averages of the magnitudes of the triggered events. In Figures 9a and 9b, the results concern the catalog simulated with Ogata's model (red) and the one simulated with our model (black), obtained with the first and the second types of analysis (Figures 9a and 9b, respectively). Regarding Figure 9a, the percentage means for the four subintervals considered for the above two catalogs are the following. Catalog simulated with Ogata's model: 0.9856, 0.9932, 1.0194, and 1.0018 (corresponding to the magnitudes of the triggering events 1.5506, 1.8931, 3.3302, and 4.6113, respectively); catalog simulated with our model: 0.8256, 0.8808, 1.0508, and 1.2428 (corresponding to the magnitudes of the triggering events 1.5268, 1.7467, 2.4491, and 3.4779, respectively). Regarding Figure 9b, the normalized means for the four subintervals considered for the two catalogs are the following. Catalog simulated with Ogata's model: 1.0052, 1.0029, 1.0011, and 0.9907 (corresponding to the magnitudes of the triggering events 1.5742, 2.163, 3.5112, and 4.6786, respectively); catalog simulated with our model: 0.9349, 0.9827, 1.0067, and 1.0757 (corresponding to the magnitudes of the triggering events 1.595, 1.9324, 2.4614, and 3.6254, respectively). In Figures 9c and 9d, the results concern the second and the third catalogs, respectively, relative to L'Aquila till 2012 (blue) and L'Aquila till 5 April 2009 (magenta), obtained with the first and the second types of analysis (Figures 9c and 9d, respectively). Regarding Figure 9c, the percentage means for the four subintervals considered for the above two catalogs are the following. Catalog two: 0.7460, 0.8413, 1.0748, and 1.3379 (corresponding to the magnitudes of the triggering events 1.9199, 2.3491, 3.4687, and 4.3342, respectively); catalog three: 0.9466, 0.9814, 1.0233, and 1.0487 (corresponding to the magnitudes of the triggering events 1.5433, 1.7796, 2.5770, and 3.3818, respectively). Regarding Figure 9d, the normalized means for the four subintervals considered for the two catalogs are the following. Catalog two: 0.9776, 0.9844, 0.9963, and 1.0417 (corresponding to the magnitudes of the triggering events 1.9508, 2.7277, 3.9125, and 5.2333, respectively); catalog three: 0.9662, 0.9801, 0.9959, and 1.0578 (corresponding to the magnitudes of the triggering events 1.5433, 1.8763, 2.7056, and 3.4083, respectively). The continuous lines correspond to the results of the linear regression, and the semiamplitude of the error bars are the normalized mean standard errors.

We are now going to analyze the results obtained for the two synthetic catalogs.

Figure 8 contains the estimated densities of the magnitude of the triggered events obtained by the first (Figures 8a and 8c) and the second (Figures 8b and 8d) types of analysis, relative to the simulated catalogs. In Figures 8a and 8b, we plot the results concerning the catalog simulated with the classical Ogata's model. In this case, the magnitudes are computed randomly with the Gutenberg-Richter law. According to this law, the magnitudes of the triggered events are not correlated with the magnitudes of their respective mother shocks. This

Table 2. List of Correlation Coefficient R and p Values[a]

	First Type of Analysis		Second Type of Analysis	
	$R \simeq$	$p \simeq$	$R \simeq$	$p \simeq$
Catalog 1(a)	0.88	0.11	/	/
Catalog 1(b)	0.96	0.04	/	/
Catalog 2	0.99	0.004	0.95	0.05
Catalog 3	0.96	0.03	0.95	0.05
Catalog 4	0.99	0.0005	0.96	0.04
Catalog 5	0.61	0.38	−0.93	0.07
Catalog 6	0.99	0.002	0.98	0.01

[a]Catalogs from 1(a) to 6 refer to the whole Italy without maximum distance, the whole Italy with maximum distance, L'Aquila till 2012, L'Aquila till 5 April 2009, Southern California, the catalog simulated with Ogata's model, and the one simulated with our new model, respectively.

is reflected in the absence of variations of the densities in the four magnitude subintervals considered, both in Figures 8a and 8b. Let us notice that the results are qualitatively similar to those in Figure 6a. The parameters, estimated by setting the precursory at about 10%, are $(\mu, \kappa, c, a, p) = (0.62, 0.02, 0.013, 1.72, 1.11)$. The results of the residual analysis are good: the probabilities for the Kolmogorov-Smirnov and the Runs tests are bigger that 5% (see Table 1).

Instead, a result that strongly supports our hypothesis of correlation is the one shown in Figures 8c and 8d. Here we show the estimated densities of the magnitude of the triggered events for the catalog simulated with our model. That is, the catalog in which the magnitudes are computed with the conditional probability density function (1). Once this catalog has been simulated, we have then estimated its parameters with the classical Ogata's FORTRAN program [**etas.f**], fixing again the learning period at about 10%. We get $(\mu, \kappa, c, a, p) = (0.58, 0.022, 0.017, 0.83, 1.12)$. The behavior in these figures is exactly the same as the one we obtain for the real data. This strongly supports our hypothesis. The results of the residual analysis, shown in Table 1, highlight that the set of parameters used is good to fit the phenomenon. In fact, both for the Kolmogorov-Smirnov and the Runs tests, we obtain a probability bigger than 5%. The number of expected events is also very close to that of the events in the target period.

Finally, Figure 9 contains the plots of the averages of the magnitudes of the triggered events versus the magnitudes of the triggering events, for both types of analysis (respectively, left and right columns). In Figures 9a and 9b, we consider the catalog simulated with Ogata's model and the one simulated with our model (red and black lines, respectively). For each of the two catalogs considered, the four triggered magnitude averages are normalized by the means of the average of the four values. The results of the linear regression analysis and the error bars are also shown. The lengths of the latter are given by the normalized mean standard errors. Regarding both Figures 9a and 9b, one can see that there is almost no percentage variation of the magnitude of the triggered events in the catalog simulated with Ogata's model. Instead, a clear increasing trend is evident for the other simulated catalog considered.

Figures 9c and 9d contain the results concerning the second and the third catalogs, relative to L'Aquila till 2012 (blue line) and L'Aquila till 5 April 2009 (magenta line), respectively. The means have an increasing trend in both of these cases. The results for the Italian and the Californian catalogs are not shown here, but we can say that the Californian catalog exhibits the same behavior of Figures 9c and 9d here. On the other hand, the means obtained with the first type of analysis for the whole Italian catalog are similar to those obtained for the catalog simulated with Ogata's model when not considering a maximum distance between events in causal relation and are similar to the other real catalogs when considering the just-mentioned distance.

The results obtained are statistically significant, as one can see from the list of correlation coefficient R and p values in Table 2.

Concluding, for the catalogs whose results show different estimated densities, we have performed the Kolmogorov-Smirnov test in order to give a further proof, statistically consistent, of the differences shown. More precisely, we have tested the null hypothesis according to which the sample of the magnitudes of the triggered events obtained for the first magnitude subinterval has the cumulative distribution function (CDF)

estimated for the other three magnitude subintervals by the kernel method of above. We have then repeated the test for all the possible combinations of samples and estimated CDFs. The results show that the magnitudes are statistically different. For brevity, we report here the results of the tests obtained only for the real catalogs relative to L'Aquila till 2012 and L'Aquila till 5 April 2009. In the case of the first type of analysis, the test performed for the catalog concerning L'Aquila till 2012 always rejects the null hypothesis, with p values relative to all the combinations of the order 1.0E−015; the same holds for the other catalog considered, for which the p values are of the order 1.0E−009. In the case of the second type of analysis, the results of the K-S test relative to the two above catalogs are as before, with p values of the order 1.0E−049 and 1.0E−004, respectively.

4. Conclusion

In order to find evidence to support the dependence of the distribution relative to the magnitude of the triggered events on the magnitude of the triggering events, we have applied two types of analysis to some catalogs, both real and simulated. In the well-known ETAS model the distribution of the magnitude of the triggered events is the same as that of the magnitude of the triggering events and is independent on past seismicity. Instead, our results support the intuitive and more realistic hypothesis of above. We notice that the probability density function for the magnitude of the triggered events seems to vary when the magnitude of the triggering events increases. In particular, we observe the increase (decrease) of the probability for high (low) values of the magnitude density. This is true for all the catalogs with the exception of the whole Italian one, when not considering a maximum distance between events in causal relation, and the one simulated by the classical Ogata's model. For the first of these two exceptions, the absence of the variation is due to the fact that it contains many pairs of events temporally close to each other but spatially very separated. For the simulated catalog just mentioned, the variation is instead absent since it is obtained using the standard Gutenberg-Richter law for the magnitudes. The fact that the catalog simulated by our magnitude model shows a clear evidence of the variations, qualitatively similar to those of the real catalogs, gives strong support to the validity of our hypothesis. Regarding the magnitude averages of the triggered events, we can see that they always have an increasing trend, again with the two exceptions of above. We interpret the results for the remaining catalogs as done before. Concluding, the law of the magnitude of the triggered events should be a conditional probability density function that changes in shape with the magnitude of the triggering events. More precisely, when the latter increases, it may have some relative maximum for higher values of the density. It should also have an increasing expected value.

The results obtained in this work could lead to the definition of a new version of the ETAS model, in which the probability density function of the magnitude of the triggered events is a conditional law with respect to the magnitude of the mother shock. More precisely, we can define a model with conditional intensity function of the kind

$$\lambda(t,x,y,m|\mathcal{H}_t) = p(m)\lambda_0(x,y) + \sum_{\{i|t_i<t\}} \lambda_{tr}\left(t-t_i, x-x_i, y-y_i; m_i\right) p(m|m_i),$$

where $\mathcal{H}_t = \{(t_i,x_i,y_i,m_i); t_i < t\}$ is the past history up to time t, $\lambda_0(x,y)$ is the rate of the background events, $\lambda_{tr}(t-t_i, x-x_i, y-y_i; m_i)$ is the rate of the triggered events, and $p(m), p(m|m_i)$ are, respectively, the magnitude distributions of triggering and triggered events, the latter conditioned to the magnitudes of the mother shocks.

We conclude by saying that the results are consistent with the ones obtained in *Nichols and Schoenberg* [2014]. In the latter work, the authors at first gave an extension of the algorithm for the declustering of a multidimensional Hawkes self-exciting process, proposed in *Marsan and Lengliné* [2008], in order to allow for a spatially dependent background rate; then, they perform an analysis to assess the correlations between the magnitude of triggering/triggered events. They come to the same conclusions as in this paper.

Acknowledgments
We would like to thank A. Govoni, W. Marzocchi, and specially A. Lombardi of the Istituto Nazionale di Geofisica e Vulcanologia for their very useful comments and suggestions. The Italian data are available through the online database ISIDe, Italian Seismological Instrumental and Parametric database (http://iside.rm.ingv.it/). The Californian data are available through the online database SCEDC, Southern California Earthquake Data Center (http://www.data.scec.org/research-tools/alt-2011-dd-hauksson-yang-shearer.html).

References

Corral, A. (2005), Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes, *Phys. Rev. Lett.*, *92*, 108501.
Corral, A. (2006), Dependence of earthquake recurrence times and independence of magnitudes on seismicity history, *Tectonophysics*, *424*(3–4), 177–193.
Gutenberg, B., and C. F. Richter (1944), Frequency of earthquakes in California, *Bull. Seismol. Soc. Am.*, *34*(8), 185–188.
Hauksson, E., W. Yang, and P. M. Shearer (2012), Waveform relocated earthquake catalog for Southern California (1981 to 2011), *Bull. Seismol. Soc. Am.*, *102*(5), 2239–2244.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2006), Comparison of short-term and time-independent earthquake forecast models for Southern California, *Bull. Seismol. Soc. Am.*, *96*(1), 90–106.
Lin, G., P. M. Shearer, and E. Hauksson (2007), Applying a three-dimensional velocity model, waveform cross correlation, and cluster analysis to locate Southern California seismicity from 1981 to 2005, *J. Geophys. Res.*, *112*, B12309, doi:10.1029/2007JB004986.
Lippiello, E., C. Godano, and L. de Arcangelis (2007a), Dynamical scaling in branching models for seismicity, *Phys. Rev. Lett.*, *98*, 098501.
Lippiello, E., M. Bottiglieri, C. Godano, and L. de Arcangelis (2007b), Dynamical scaling and generalized Omori law, *Geophys. Res. Lett.*, *34*, L23301, doi:10.1029/2007GL030963.
Lippiello, E., L. de Arcangelis, and C. Godano (2008), Influence of time and space correlations on earthquake magnitude, *Phys. Rev. Lett.*, *100*, 038501.
Lippiello, E., C. Godano, and L. de Arcangelis (2012), The earthquake magnitude is influenced by previous seismicity, *Geophys. Res. Lett.*, *39*, L05309, doi:10.1029/2012GL051083.
Lombardi, A. M., M. Cocco, and W. Marzocchi (2010), On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, central Italy, sequence: Apparent variation or fluid-driven triggering?, *Bull. Seismol. Soc. Am.*, *100*(3), 113–1152.
Marsan, D., and O. Lengliné (2008), Extending earthquakes' reach through cascading, *Science*, *319*(5866), 1076–1079.
Nichols, K., and F. P. Schoenberg (2014), Assessing the dependency between the magnitudes of earthquakes and the magnitudes of their aftershocks, *Environmetrics*, *25*(3), 143–151.
Ogata, Y. (1981), On Lewis' simulation method for point processes, *IEEE Trans. Inf. Theory*, *27*(1), 23–31.
Ogata, Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, *83*(401), 9–27.
Ogata, Y. (1989), Statistical model for standard seismicity and detection of anomalies by residual analysis, *Tectonophysics*, *169*(1–3), 159–174.
Ogata, Y. (1998), Space-time point process models for earthquake occurrences, *Ann. Inst. Stat. Math.*, *50*(2), 379–402.
Ogata, Y. (1999), Seismicity analysis through point-process modeling: A review, *Pure Appl. Geophys.*, *155*(2–4), 471–507.
Ogata, Y. (2006), *Statistical Analysis of Seismicity: Updated Version (SASeis2006), Comput. Sci. Monogr.*, vol. 33, Inst. of Stat. Math., Tokyo, Japan.
Parzen, E. (1962), On estimation of a probability density function and mode, *Ann. Math. Stat.*, *33*(3), 1065–1076.
Priestley, M. B. (1982), *Spectral Analysis and Time Series*, vol. 1, Academic Press, London, New York.
Rosenblatt, M. (1956), Remarks on some nonparametric estimates of a density function, *Ann. Math. Stat.*, *27*(3), 832–837.
Saichev, A., and D. Sornette (2008), Vere-Jones' self-similar branching model, *Phys. Rev. E*, *72*, 056122.
Sarlis, N. V., E. S. Skordas, and P. A. Varotsos (2009), Multiplicative cascades and seismicity in natural time, *Physi. Rev. E*, *80*, 022102.
Sarlis, N. V., E. S. Skordas, and P. A. Varotsos (2010), Nonextensivity and natural time: The case of seismicity, *Phys. Rev. E*, *82*, 021110.
Scott, D. W. (1992), *Multivariate Density Estimation: Theory, Practice and Visualization*, John Wiley, New York.
Shi, Y., and B. A. Bolt (1982), The standard error of the magnitude-frequency b value, *Bull. Seismol. Soc. Am.*, *72*(5), 1677–1687.
Spassiani, I., and G. Sebastiani (2015), A new magnitude-dependent ETAS model for earthquakes, arXiv:1504.05868 [math.PR].
Zhuang, J., Y. Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *J. Am. Stat. Assoc.*, *97*(458), 369–380.