

A double-branching model applied to long-term forecasting of Italian seismicity ($M_L \geq 5.0$) within the CSEP project

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

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

Article history

Received October 30, 2009; accepted April 22, 2010.

Subject classification:

Earthquake interactions and probability, Forecasts, Long-term, Italian seismicity, Clustering.

ABSTRACT

The purpose of this study was to apply a double-branching model to forecasting of moderate-to-large Italian seismicity within the Collaboratory Study for Earthquake Predictability project. This project is designed for statistical evaluations and comparisons of various forecasting models, on both the global and regional scales. This proposed double-branching model is time-dependent, as it assumes that each earthquake can generate, or is correlated with, other earthquakes through physical mechanisms that act on different spatio-temporal scales. Specifically, it consists of an application of two branching processes, in which any earthquake can trigger a family of later events on different space-time scales. In our recent study [Lombardi and Marzocchi 2009], we applied this model to a declustered historical database that included the strong Italian seismicity from over the last few centuries. This catalog only allowed us to describe the long-term time evolution of moderate-to-strong seismicity. Here, we have applied this double-branching model to a new database that has allowed us to describe both short-term clustering and long-term features at the same time. As the model can produce forecasting calculations of future seismicity, we provide some probability maps of occurrence of predicted events over different temporal windows.

Introduction

The modeling of the spatio-temporal distributions of moderate-to-large earthquakes is a major goal of geophysical investigations. The huge social impact of risk mitigation has promoted the formulation of both physically based and stochastic models, which are all directed towards providing better understanding of the driving features of earthquake occurrence and/or towards forecasting of future events [Dieterich 1994, Ogata 1998, Stein 1999, Kagan and Jackson 2000, Gerstenberg et al. 2005]. Most of the proposed models hypothesize a substantially stable temporal behavior of the seismic rate around a nearly constant mean, with short-term random fluctuations arising from aftershocks. This assumption explains the current use of the stationary Poisson paradigm in many practical applications that are related to long-term behavior of earthquake occurrence,

such as in the formulation of probabilistic seismic-hazard-assessment methodologies based on the Cornell method [Cornell 1968] and in the evaluation of earthquake prediction/forecasting models [e.g., Kagan and Jackson 1994, Frankel 1995, Varotsos et al. 1996, Gross and Rundle 1998, Kossobokov et al. 1999, Marzocchi et al. 2003].

Over the last two decades, considerable attention has been focused on long-term trends of seismic rates and on the multiyear stability of earthquake occurrence that has been called into question [Kagan and Jackson 1991a, Kagan and Jackson 1991b, Rhoades and Evison 2004, Lombardi and Marzocchi 2007, Marzocchi and Lombardi 2008]. The long-term memory of the seismogenetic potential of a region can be ascribed to different physical causes, such as possible stress perturbations on spatio-temporal scales a lot larger than the duration of the sequences [Stein 1999], or combined effects of fault recurrence and fault interactions [Marzocchi et al. 2009].

In Marzocchi and Lombardi [2008], we formulated a new model, known as a double-branching model (DBM). The DBM takes into account long-term modulation of earthquake occurrence, as well as short-term clustering of earthquakes. This model has shown better earthquake forecasting performances for large earthquakes on both worldwide [Marzocchi and Lombardi 2008] and regional [Lombardi and Marzocchi 2009] scales, with respect to models with time-independent background rates.

Following on from this, we applied the DBM to Italian seismicity using a declustered historical catalog that allowed recovery of about 400 years of seismic history above a moment magnitude (M_w) of 5.5 [Lombardi and Marzocchi 2009]. Here, we now apply this model to events above a local magnitude (M_L) of 4.5 that have occurred in Italy over about the last century. This last dataset also includes aftershock sequences, and it allows us to model the short-term features of the seismicity, which was ruled out in our previous study [Lombardi and Marzocchi 2009].

The DBM can provide earthquake forecasts for participation in the initiative known as the Collaboratory Study

for Earthquake Predictability (CSEP; www.cseptest.org) [Jordan 2006, Zechar et al. 2009]. Here, we present probability maps of future earthquakes with $M_L \geq 5.0$ in Italy over two different temporal windows: 5 years and 10 years. These forecasts have been submitted to the CSEP EU-Italy Testing Center, to be evaluated by suitable statistical tools.

The double branching model

In the present study, we apply the stochastic model proposed by Marzocchi and Lombardi [2008], which consists of the sequential application of two branching processes in which any earthquake can trigger a family of later events. The main goal of our model is to highlight the interactions between events that are due to different physical processes and that involve largely different spatio-temporal domains.

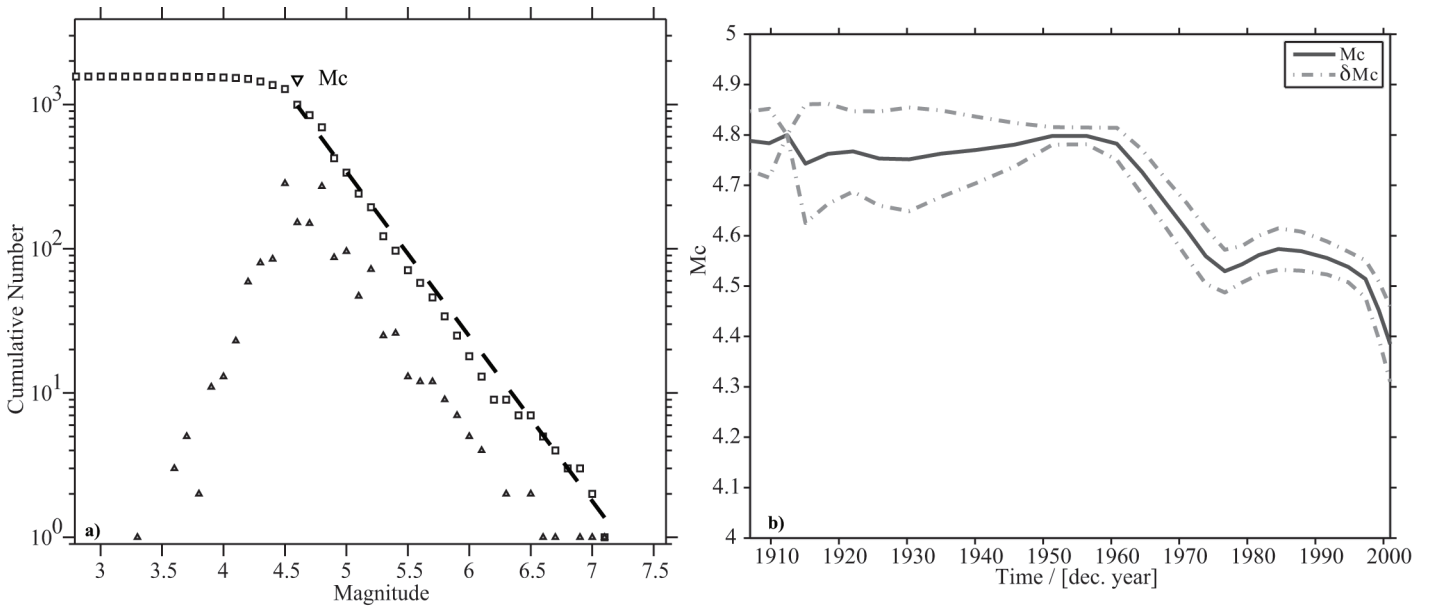
In the first step of our modeling, we applied a version of the well-known epidemic type-aftershocks sequences (ETAS) model [Ogata 1998] to describe the short-term clustering of seismic events in space and time. The intensity function of this first branching model is given by:

$$\lambda_1(t, x, y, m/H_t) = \left[\nu_1 u_1(x, y) + \sum_{t_i < t} \frac{K_1}{(t - t_i + c)^p} e^{\alpha_1(M_i - M_c)} \frac{c_{d_1, q_1}}{[r_i^2 + d_1^2]^{q_1}} \right] \beta e^{-\beta(m - M_c)} \quad (1)$$

where $H_t = \{(t_i, x_i, y_i, M_i); t_i < t\}$ is the observation history up to time t , M_c is the minimum magnitude of the catalog, K_1 , c and p are parameters of the modified Omori Law that define the decay in time of short-term triggering effects, and α_1

defines the dependence of the triggering capability on the magnitude of an earthquake. The parameters d_1 and q_1 define the spatial distributions of triggered events as functions of the distance r_i between a general location (x, y) and the epicenter of the i -th earthquake (x_i, y_i) (C_{d_1, q_1} is a normalization constant). The function $u_1(x, y)$ is the probability density function of the locations of spontaneous events [Ogata 1998]. Finally, $\beta = b \cdot \ln(10)$ is a parameter of the well-known Gutenberg-Richter Law [Gutenberg and Richter 1954], which is assumed as the distribution for the magnitude of all events. Using ETAS modeling, it is possible to compute for each earthquake a probability of being a triggered event, and consequently to formulate a stochastic declustering algorithm that can remove the short-term triggered events from the original dataset [see Zhuang et al. 2002, for details].

The second step of our procedure consists of re-applying a branching process to the filtered database that is obtained using the ETAS-derived declustering procedure proposed by Marzocchi and Lombardi [2008]. This step is based on the stochastic technique of Zhuang et al. [2002], but unlike this stochastic technique, we apply an algorithm that produces a single declustered catalog. We have verified in the present and in previous analyses that the significance of the results does not depend on the adopted declustering procedure [see Marzocchi and Lombardi 2008, for details]. This second branching model works on larger space-time scales compared to the smaller domains involved in short-term clustering, which are removed after the first step. The time-dependent conditional rate of earthquake occurrences



Maximum Likelihood Solution
b-value = 1.14 +/- 0.03, a value = 8.25, a value (annual) = 6.22
Magnitude of Completeness = 4.6

Figure 1. Completeness magnitude of the CPTI08 catalog (January 1, 1901, to December 31, 2006; 1591 events) by the maximum curvature method (MAXC). (a) Frequency magnitude distribution for the whole dataset: the MAXC Method provides $M_c = 4.6$ (moment magnitude); (b) M_c as a function of time.

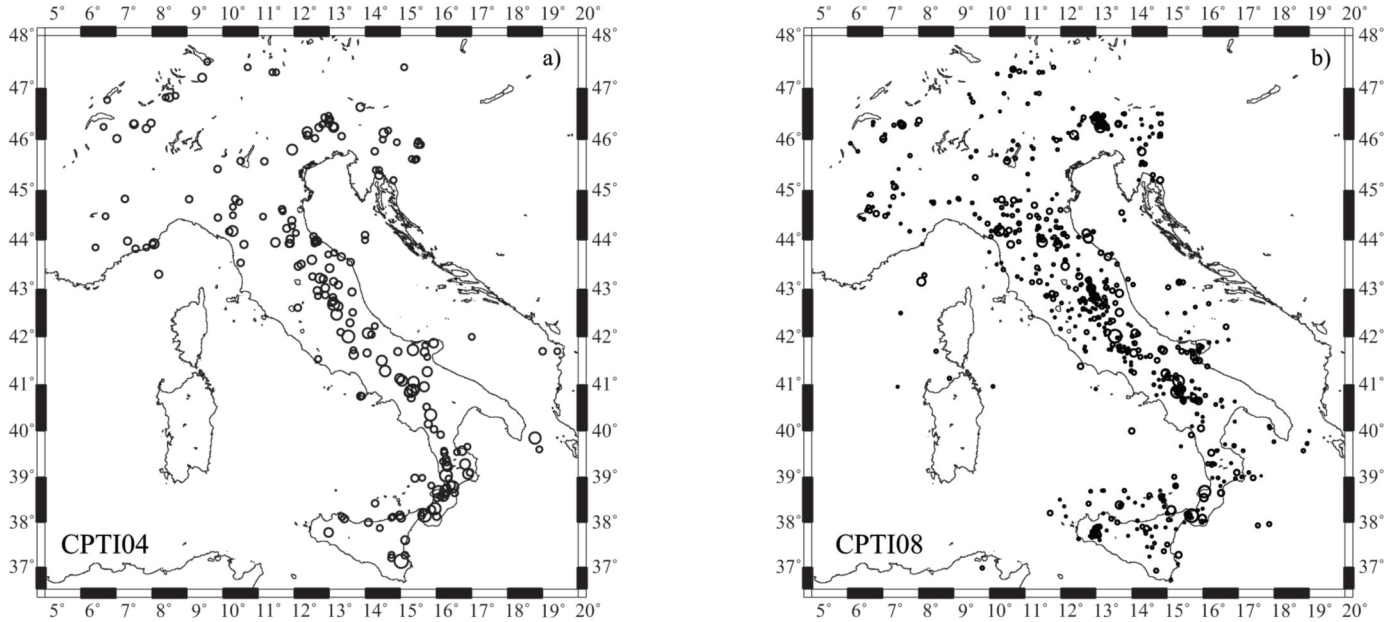


Figure 2. Maps of seismic events collected in the catalogs used in the present study. (a) Events of the CPTI04 dataset (January 1, 1600, to December 31, 2002; $M_w \geq 5.5$; 217 events); (b) Events of CPTI08 catalog (January 1, 1901 to December 31, 2006; $M_L \geq 4.5$; 699 events). The symbol sizes are scaled according to magnitude.

for the second step of our procedure is given by:

$$\lambda_2(t, x, y, m/H_i) = \left[\nu_2 u_2(x, y) + \sum_{t_i < t} K_2 e^{-\frac{(t-t_i)}{\tau}} e^{\alpha_2(M_i - M_c)} \frac{c_{d_2, q_2}}{[r_i^2 + d_2^2]^{q_2}} \right] \beta e^{-\beta(m - M_c)} \quad (2)$$

The residual seismicity that is obtained by filtering the original database of these short-term triggering effects is therefore ascribed to the superposition of two physical processes: the time-independent and spatially variable tectonic loading $\mu_2 = \nu_2 u_2(x, y)$, and the long-term coupling between events. This latter is described in time by an inverse exponential function with a characteristic time τ . Similar to the first branching step, the dependence of a hazard function with the magnitude of an exciting event is assumed to be of an exponential type, i.e. proportional to $e^{\alpha_2 M}$. Finally, the spatial decay of long-term stress variation is described by an inverse power law probability density function, the function of distance r_i from the epicenter of i -th perturbing event, and the parameters d_2 and q_2 (c_{d_2, q_2} is a normalization constant). To estimate the parameters of both of these branching processes, we used an iteration algorithm developed by Zhuang et al. [2002], which is based on the maximum likelihood method and on a kernel estimation of the total seismic rate. Further details on this model and on the estimation of its parameters can be found in Marzocchi and Lombardi [2008].

From a technical point of view, we chose to follow a step-wise procedure rather than to apply a single complex model, for two reasons: (1) the simultaneous estimation of a large number of parameters involves problems that do not have easy solutions; and (2) it has been shown that the description of additive complex processes through a step-

wise procedure can lead to significantly better modeling [e.g. the boosting approach of Bühlmann 2003].

The datasets: the CPTI04 and CPTI08 earthquake catalogs

Italy is characterized by well tested experience in compiling historical databases [Camassi and Stucchi 1997, Boschi et al. 1995, Boschi et al. 1997, Boschi et al. 2000]. The most recently compiled historical catalog of Italian seismicity is the parametric catalog of Italian earthquakes (Catalogo Parametrico dei Terremoti Italiani; hereinafter referred to as CPTI04); this catalog was published at 1999 [CPTI Working Group 1999] and revised in 2004 [CPTI Working Group 2004]. The CPTI04 catalog was obtained by merging some previously collected datasets and reports of earthquakes with intensities $I_0 \geq V/VI$ on the Mercalli-Cancani-Sieberg (MCS) scale that occurred in Italy (or in neighboring zones) from 217 B.C. to 2002. Given the strong hazard-oriented nature of this catalog, foreshocks and aftershocks were removed from the original datasets that were merged into the CPTI04 catalog. Specifically, the declustering procedure consisted of removing all of the events that occurred in a 30-km and 90-day spatio-temporal window with respect to a shock with a larger magnitude. Historical investigations relating to the completeness of the CPTI04 catalog have identify a substantially complete recording for events with magnitude $M_w \geq 5.5$ from 1600 (217 events) [Stucchi et al. 2004].

The CPTI04 catalog is suitable for hazard analysis and for studies of the long-term behavior of seismicity, although it cannot be used to model short-term clustering of events. Recently, a new catalog was proposed in its first version, known as CPTI08 [Rovida and the CPTI Working Group 2008]. This CPTI08 catalog includes aftershocks for which

Parameter	Value
ν_1	5.1 ± 0.2 (year ⁻¹)
K_1	0.005 ± 0.001 (year ^{$p-1$})
p	1.13 ± 0.04
c	0.00009 ± 0.00005 (year)
α_1	1.2 ± 0.1
d_1	0.6 ± 0.2 (km)
q_1	$\equiv 1.5$
Log-likelihood	-7298.3

Table 1. Parameters of ETAS model for CPTI08 catalog ($M_L = 4.5$; January 1, 1901, to December 31, 2006; 699 events).

macroseismic data are known. The first published version of CPTI08 collected the shocks that occurred from 1901 to 2006 (1591 events).

Given the strong uncertainty associated with magnitude values, we analyzed the completeness magnitude of the CPTI08 catalog using the simple maximum curvature (MAXC) method [Wiemer and Wyss 2000]. The overall estimate predicted a completeness magnitude M_w 4.6, although the analysis in time shows a decreasing completeness magnitude from M_w 4.8 to M_w 4.6. So we selected the events above a magnitude M_w 4.8 so as to have a homogeneous and complete catalog (see Figure 1).

The major goal of the present study is to present a model that can forecast future seismicity in Italy, in the framework of the CSEP project. Thus, we converted the moment magnitudes reported in the CPTI08 catalog into local magnitudes (M_L). In this way, we conform to the CSEP in their tests of proposed models of the seismic bulletin of the INGV, Istituto Nazionale di Geofisica e Vulcanologia (Bollettino

Sismico Italiano; <http://bollettinosismico.rm.ingv.it/> and <http://iside.rm.ingv.it/>), where they have adopted M_L as their official magnitude scale. The magnitude conversion was obtained through the relationship proposed by the MPS Working Group [2004]:

$$M_L = (M_w - 1.145) / 0.812 \quad (3)$$

By using this relationship, the completeness moment magnitude M_w 4.8 is transformed into the completeness local magnitude M_L 4.5. In Figure 2, we show the locations of the CPTI04 catalog events ($M_w \geq 5.5$) and the CPTI08 catalog events ($M_L \geq 4.5$).

Application of the double-branching model to the CPTI08 catalog.

We applied the ETAS model, the first step of the DBM, to the Italian seismicity recorded in the CPTI08 database. Following the procedure proposed by Zhuang et al. [2002], we estimated the model parameters, together with the spatial distribution of the seismicity that was not co-seismically triggered ($u_1(x, y)$; see Equation 1). Table 1 lists the inferred values of the model parameters, together with their errors and the associated log-likelihood values. Figure 3 shows the ratios between the short-term triggering and the total rates for the whole period recovered from the CPTI08 database. These rates were obtained by integrating in time the intensity function $\lambda_1(t, x, y, m/H_i)$ of the first branching model (Equation 1) and the contribution provided only by the short-term triggering effects (sum in Equation 1). As can be seen, in most of the zones, the short-term triggered rate is below 30% of the total seismic rate. In some areas however (Friuli, central Apennines, Irpinia, Messina Straits, western Sicily), there were well-identified sequences that reach 90% of the total rate.

The application of the second-order branching model (DBM) to the CPTI08 dataset provided the parameters reported in Table 2, together with their relative errors and the maximum log-likelihood values. Given the short duration of the CPTI08 dataset, we set the τ parameter to a value (30 years) estimated from the longer CPTI04 catalog [see Lombardi and Marzocchi 2009]. Moreover we fixed $\alpha_2 = 0$, as seen in previous studies on both global and local scales [see Marzocchi and Lombardi 2008, Lombardi and Marzocchi 2009]. As discussed in Lombardi and Marzocchi [2009], this α_2 value might be due to the low number of events. In any case, we have verified that no different combinations of these parameters provides better log-likelihoods. To determine the reliability of the DBM, we also report here the mean rates and the maximum log-likelihoods obtained by Poisson modeling (see Table 2). The time-dependent model is significantly better than the Poisson process considering the Akaike information criterion [Akaike 1974]. This result was

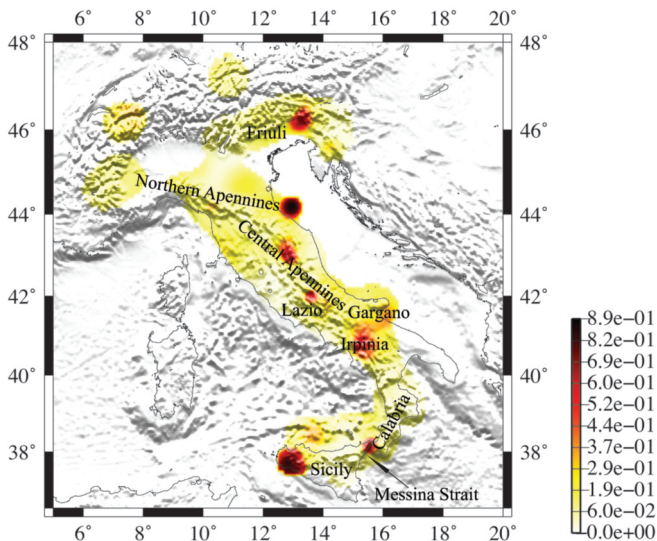


Figure 3. Map of ratios between short-term triggered and total seismic rates for the CPTI08 dataset.

Parameter	DBM	POISSON
ν_2	3.0 ± 0.3 (year ⁻¹)	5.1 ± 0.2 (year ⁻¹)
K_2	0.017 ± 0.002	
τ	$\equiv 30$ (years)	
α_2	$\equiv 0.0$	
d_2	15 ± 4.0 (km)	
q_2	2.3 ± 0.5	
Log-likelihood	-6799.2	-6905.5
Akaike Information Criterion	13606.5	13813.0

Table 2. Parameters of second branching for CPTI08 declustered catalog ($M_L = 4.5; 4.5$; January 1, 1901, to December 31, 2006; 545 events).

confirmed by applying residual analysis [Ogata 1998, Lombardi and Marzocchi 2009] to the CPTI08 dataset. This diagnostic technique confirms that if a point process model with intensity $\lambda(t)$ describes the temporal evolution of the seismicity well, the transformed data

$$\tau_i = \int_{T_{\text{start}}}^t \lambda(t) dt$$

are expected to behave like a stationary Poisson process with a unit rate [Ogata 1998]. We tested this hypothesis by the Runs and the Kolmogorov-Smirnov tests [Lombardi and Marzocchi 2007], for independence and exponential distribution of residuals, respectively. The P value of the Kolmogorov-Smirnov test leads to the rejection of the ETAS model ($P < 0.01$), but not the rejection of the DBM ($P = 0.2$); the Runs test does not reject the independence hypothesis of residuals for either ETAS or the DBM.

The most reliable cause of this finding is a probable long-term time-dependence of the occurrence of damaging shocks in Italy. Indeed, the temporal scale recovered by the

second model (with a characteristic time τ of 30 years) suggested that the interactions revealed by the second branching model (DBM) are not likely to be due to poor ETAS model formulation.

To represent the significance of the long-term interactions revealed by the DBM, we show in Figure 4 the map of the proportion of seismicity due to long-term triggering effects, together with the map of the spatial distribution of tectonic-driven seismicity ($u_2(x, y)$). The model defines the maximum effects of long-term interactions in the Friuli region (see Figure 4b), and a larger tectonic loading along the northern-central Apennines range and in the Friuli region (see Figure 4a). In any case, as the CPTI08 dataset recovers a time period well below the reloading time of the faults in Italy, it will not be representative of the real seismogenetic potential of these different zones. By considering the model estimated by Lombardi and Marzocchi [2009] (the first application of DBM to Italian seismicity) with the CPTI04 catalog, we obtained the map of the probability distribution $u_2(x, y)$

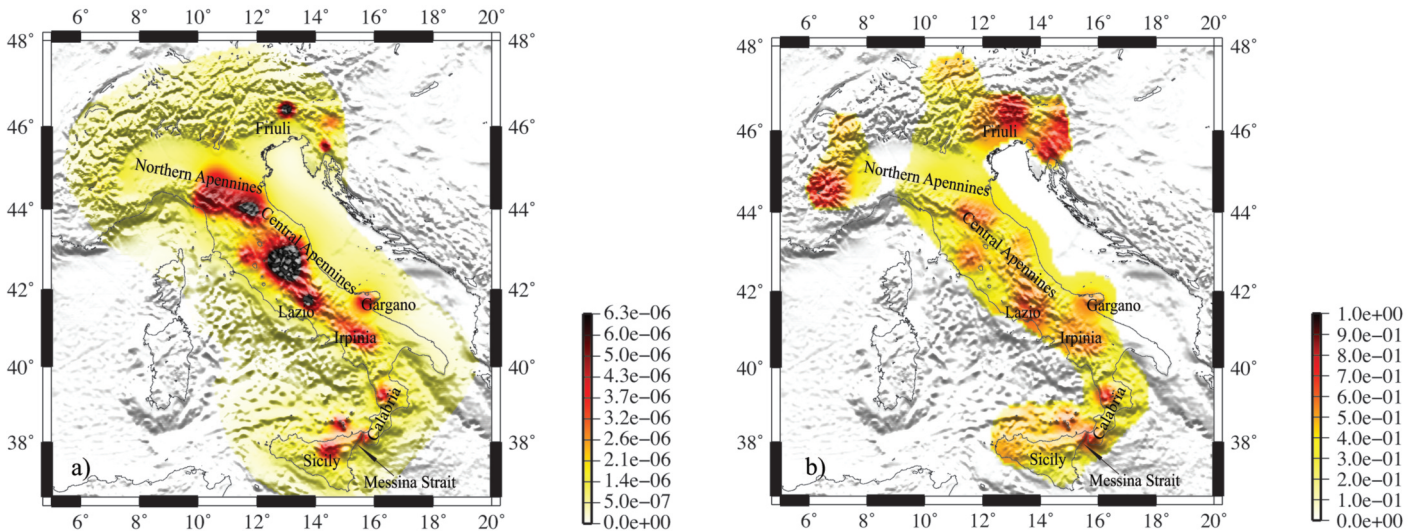


Figure 4. Maps of (a) spatial distribution of tectonic-driven seismicity $u_2(x, y)$; and (b) ratio between long-term triggered rate and total seismic rate for the CPTI08 declustered catalog.

	Location	Data	Lat	Lon	Depth (Km)	Mag (M_L)	$P_{\{M=4.5+\}}$	Q
1	Frignano	2008-12-23	44.544°N	10.345°E	22.9	5.2	$0.54 \cdot 10^{-2}$	7%
2	Frignano	2008-12-23	44.527°N	10.355°E	23.7	4.8	$0.54 \cdot 10^{-2}$	7%
3	Forlì	2009-04-05	44.236°N	11.999°E	28.2	4.6	$0.18 \cdot 10^{-1}$	0.4%
4	L'Aquila	2009-04-06	42.334°N	13.334°E	8.8	5.8	$0.70 \cdot 10^{-2}$	4.3%
5	L'Aquila	2009-04-06	42.355°N	13.342°E	9.7	4.7	$0.70 \cdot 10^{-2}$	4.3%
6	L'Aquila	2009-04-06	42.366°N	13.340°E	10.1	4.6	$0.70 \cdot 10^{-2}$	4.3%
7	L'Aquila	2009-04-06	42.451°N	13.364°E	8.6	4.8	$0.91 \cdot 10^{-2}$	2.5%
8	L'Aquila	2009-04-07	42.342°N	13.388°E	10.2	4.7	$0.70 \cdot 10^{-2}$	4.3%
9	L'Aquila	2009-04-07	42.275°N	13.464°E	15.1	5.3	$0.56 \cdot 10^{-2}$	6.7%
10	L'Aquila	2009-04-09	42.484°N	13.343°E	15.4	5.1	$0.91 \cdot 10^{-2}$	2.5%
11	L'Aquila	2009-04-09	42.501°N	13.356°E	17.2	4.9	$0.12 \cdot 10^{-1}$	1.1%
12	L'Aquila	2009-04-13	42.504°N	13.363°E	7.5	4.9	$0.12 \cdot 10^{-1}$	1.1%
13	L'Aquila	2009-06-22	42.446°N	13.356°E	14.2	4.5	$0.91 \cdot 10^{-2}$	2.5%

Table 3. Parameters of the 13 events with magnitude ≥ 4.5 occurred in Italy from January 1, 2007, to July 31, 2009.

shown in Figure 5. Here we see clear differences in the distribution obtained in the present study. Specifically the two studies have identified different tectonic loading in some regions of southern Italy (Calabria, Sicily, Irpinia), along the northern Apennines, and into the southern part of the Lazio region. To provide forecasts of future seismicity, we decided to use the parameters estimated in the present study and the spatial distribution of tectonic loading ($u_2(x, y)$) estimated by Lombardi and Marzocchi [2009] on the CPTI04 catalog (Figure 5), which we judged to be the more representative. This choice implies that the distribution of $M_w \geq 5.5$ earthquakes is the same as the $M_L \geq 4.5$ events.

For a very preliminary check of the forecasting capability

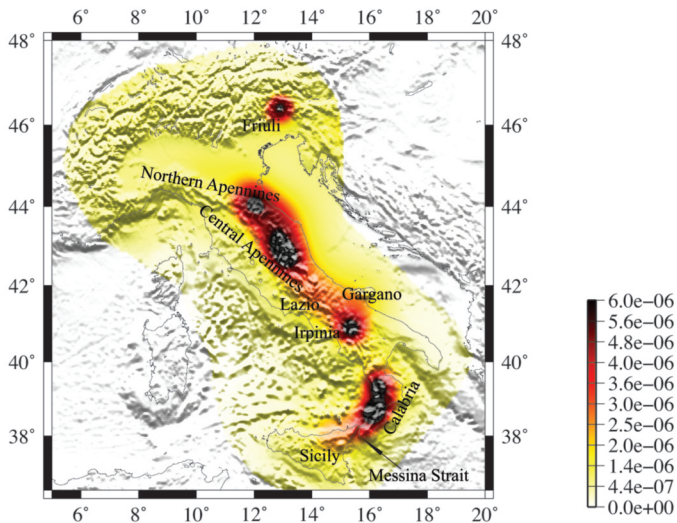


Figure 5. Map of spatial distribution of tectonic-driven seismicity $u_2(x, y)$ for the CPTI04 catalog.

of our model, we show in Figure 6 the map of the predicted number of events for the period January 1, 2007, to July 31, 2009. We also plot the locations of the events with $M_L \geq 4.5$ and greater than 30 km in depth that occurred over the same period in the CSEP testing region that were collected by the official bulletin of the INGV (<http://iside.rm.ingv.it>) (13 events). The basic parameters of these events are reported in Table 3. All of these events occurred in cells with relatively high forecast rates. For 10 of the 13 events, the proportion of cells with a higher forecast rate (with respect to the cell where the earthquake occurred) is lower than 5% (Table 3, last column). We note here that the forecasted rates were computed for the whole period, without taking into account the triggering effects of the real events that occurred within this period. In this way, we intended to reproduce the forecasts made for the CSEP, for which we have no information on future earthquakes and we can only simulate the triggering effects of these events. By including the real triggering effects for forecasting calculations, the results in Table 3 are further improved (at least for events 2 and 9).

Forecasting maps

The model formulated and tested in the previous sections allowed us to compute the forecasts for future seismicity in Italy. The predictions are in a form of the probabilities of the occurrence in Italy of at least one earthquake with a $M_L \geq 5.0$ within a cell of $0.1^\circ \times 0.1^\circ$ over the next 5 years (August 2009 to July 2014) and 10 years (August 2009 to July 2018). These are obtained by integrating in time and space a combination of the intensity functions of two branching models (Equations 1 and 2). Specifically for each cell C_i and for each forecasting

period T_j , we computed the relative rates of occurrence R_{ij} according to the following integration:

$$R_{ij} = \iint_{C_i} \int_{T_j} \left[\nu_2 u_2(x, y) + \sum_{t_i < t} \frac{K_1}{(t - t_i + c)^p} e^{\alpha_1(M_i - M_c)} \frac{c_{d_1, q_1}}{[r_i^2 + d_1^2]^{q_1}} + w_i K_2 e^{-\frac{(t - t_i)}{\tau}} e^{\alpha_2(M_i - M_c)} \frac{c_{d_2, q_2}}{[r_i^2 + d_2^2]^{q_2}} \right] dt dx dy \quad (4)$$

where w_i is the probability that the i -th event is not co-seismically triggered, computed by the first branching ETAS model [see Marzocchi and Lombardi 2008]. The forecasted rates above M_L 5.0 were obtained by rescaling the rate of earthquakes above M_L 4.5, in agreement with the Gutenberg-Richter relation.

Equation (4) shows that time-dependent modeling such as the DBM can also take into account the triggering effects of the seismicity that occurred before and that are expected during the forecast interval. So, in the past history, we included all of the seismicity with magnitudes greater than M_L 4.5 and depths greater than 30 km that occurred up to July 31, 2009, and were collected in the CPTI08 catalog (January 1, 1901 to December 31, 2006) and INGV catalog (January 1, 2007 to July 31, 2009). To generate the forecasting maps, we simulated 10,000 different stochastic realizations for each of two time windows, using the thinning method proposed by Ogata [1998] and the intensity function formulated in Equations (1) and (2). Here, no significant variations were found using a much lower number of simulations (up to 1,000). Then we averaged the predictions coming from each of these synthetic catalogs. The results relative to both of these time periods are shown in Figure 7. The expected numbers of events with $M_L \geq 5.0$ over the next 5 and 10 years are 9.4 and 18.4, respectively. The probabilities of one or more events over next 5 years with $M_L \geq 6.0$ and

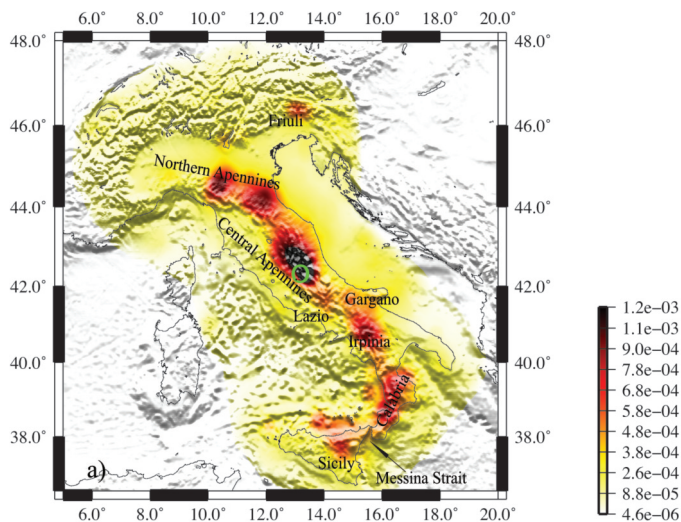


Figure 7. Maps of probability of occurrence of one or more events with magnitude above M_L 4.5 per cell of $0.1^\circ \times 0.1^\circ$ over the next 5 (a) and 10 (b) years. Green circles: spatial bin with highest probability.

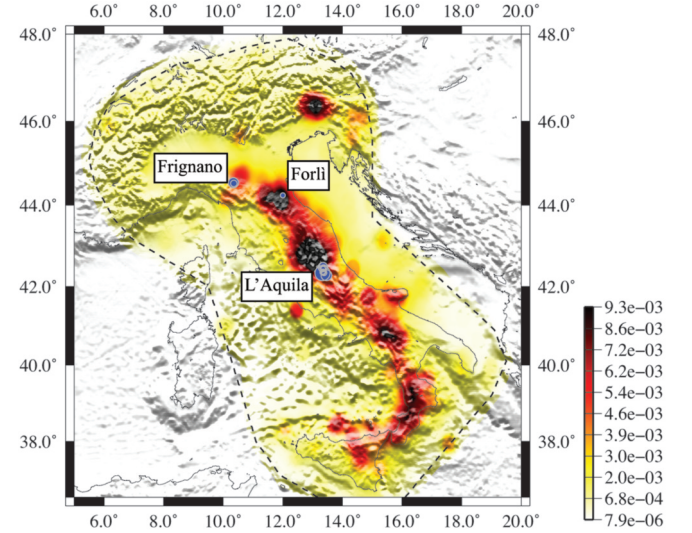
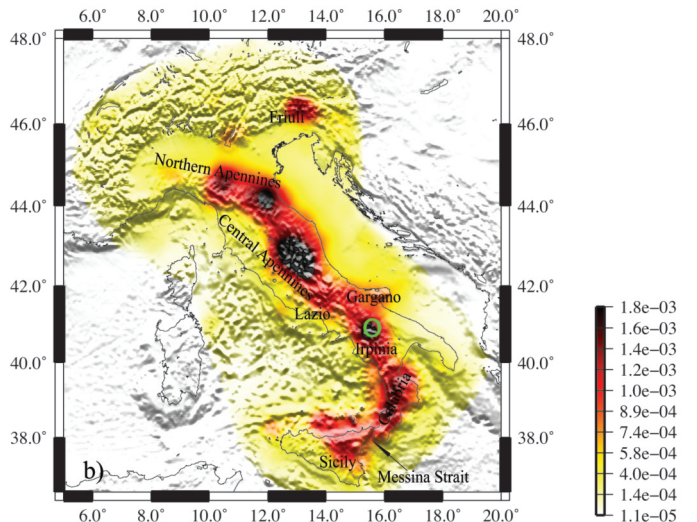


Figure 6. Map of seismic rates (number of events per cell of $0.1^\circ \times 0.1^\circ$) predicted by the DBM for the period January 1, 2007, to July 31, 2009, inside the CSEP testing region (dashed line). Blue circles: locations of 13 events that occurred over the same period (from the INGV dataset).

≥ 7.0 are 40% and 4%, respectively. The DBM identifies the most dangerous zones as the regions that have been affected by the most recent earthquakes. Above all, there are the Abruzzo region, which was hit by the recent M_w 6.3 L'Aquila earthquake (April 6, 2009; see Figure 5) and the Irpinia region that have the highest probabilities of experiencing a shock in the next 5 years and 10 years, respectively (see Figure 7). Smaller maxima of probability are seen in the northern-central Apennines, Friuli, Calabria and the eastern part of Sicily.

Conclusions

The main goal of the present study was to describe the DBM as applied to forecasting Italian seismicity for the CSEP project. This model represents an implementation of the model described in Lombardi and Marzocchi [2009]; the



main difference is in the seismic catalogs that were used to set-up the models. From a seismological point of view, the results obtained in the present study confirm the main findings of Lombardi and Marzocchi [2009]. First, large earthquakes in Italy tend to cluster in time and space, and on a temporal scale (from decades to a few centuries) that is a lot longer than what would be expected for a typical aftershock sequence [see also Faenza et al. 2003, Cinti et al. 2004]. Secondly, the time-dependent model proposed here is a significantly better model, in which the only time-dependent feature is short-term clustering.

The main improvements to this model, as compared to that described in Lombardi and Marzocchi [2009], is linked to the use of the first version of the CPTI08 catalog. This has allowed us to apply the whole DBM to Italy, including also the modeling of short-term behavior (Equation 1). In Lombardi and Marzocchi [2009], we applied only the second step of the DBM (see Equation 2) as the CPTI04 catalog, that we used in that study, is a declustered dataset.

The comparison of the forecasts proposed in the present study with the maps presented by Lombardi and Marzocchi [2009] highlights some differences. Specifically, the map of the probability distribution of occurrence over next 10 years (Figure 6b) proposed in the present study shows some hazardous regions that were not identified in our previous study: northern Apennines, the Gargano region and Eastern Sicily [see Figure 5b in Lombardi and Marzocchi 2009]. These findings are due to the larger proportions of moderate events that have occurred in these zones over last decades, with respect to the stronger seismicity of the last four centuries that was collected in the CPTI04 catalog. We stress here that the spatial distribution of the tectonic driven seismicity ($u_2(x,y)$) used in both of these studies is the same, and that the differences are due to long-term features that are modeled by the second step of our model (see Equation 2).

Finally we note that an intrinsic limitation of the DBM is that it is based on information coming from only a limited historical catalog. In this case, we argue that a few centuries could not be enough to get all of the time features to accurately describe the long-term variations; moreover, the inclusion of the relevant tectonic/geological information might represent a future direction for improvements to these models.

References

- Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 19, 716-722.
- Boschi, E., G. Ferrari, P. Gasperini, E. Guidoboni, G. Smriglio and G. Valensise (1995). *Catalogo dei forti terremoti in Italia dal 461 a.C. al 1980*, vol. 1, ING-SGA Bologna, 973 pp.
- Boschi, E., E. Guidoboni, G. Ferrari, G. Valensise and P. Gasperini (1997). *Catalogo dei forti terremoti in Italia dal 461 a.C. al 1990*, vol. 2, ING-SGA Bologna, 644 pp.
- Boschi, E., E. Guidoboni, G. Ferrari, D. Mariotti, G. Valensise and P. Gasperini (2000). *Catalogue of Strong Italian Earthquakes from 461 A.C. to 1997*, *Annals of Geophysics*, 43 (4), 609-868 (with database on CD-ROM).
- Bühlmann, P. (2003). Boosting Methods: Why They Can Be Useful for High-Dimensional Data, in: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003; March 20-22, Vienna, Austria)*, edited by K. Hornik, F. Leisch and A. Zeileis; available at <http://pubs.usgs.gov/of/2003/of03-214>.
- Camassi, R. and M. Stucchi (1997). NT4.1, un catalogo parametrico di terremoti di area italiana al di sopra della soglia di danno/NT4.1, a parametric catalog of damaging earthquakes in the Italian area, CNR-GNDT Milano, 95 pp.; <http://emidius.mi.ingv.it/NT/>.
- Cinti, F., L. Faenza, W. Marzocchi and P. Montone (2004). Probability map of the next $M \geq 5.5$ earthquakes in Italy, *Geochem. Geophys. Geosyst.*, 5, Q11003; doi: 10.1029/2004GC000724.
- Cornell, C.A. (1968). Engineering seismic risk analysis, *Bull. Seism. Soc. Am.*, 58, 1583-1606.
- CPTI Working Group (1999). *Catalogo Parametrico dei Terremoti Italiani, versione 1 (CPTI99, luglio 1999)*, ING-GNDT-SGA-SSN, Bologna, 88 pp. ; <http://emidius.mi.ingv.it/CPTI99/>.
- CPTI Working Group (2004). *Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04)*, ING-V, Bologna; <http://emidius.mi.ingv.it/CPTI04/>.
- Dieterich, J.H. (1994). A constitutive law for rate of earthquake production and its application to earthquake clustering, *J. Geophys. Res.*, 99, 2601-2618.
- Faenza, L., W. Marzocchi and E. Boschi (2003). A non-parametric hazard model to characterize the spatio-temporal occurrence of large earthquakes; an application to the Italian catalog, *Geophys. J. Int.*, 155, 521-531.
- Frankel, A. (1995). Mapping seismic hazard in the central and eastern United States, *Seismol. Res. Lett.*, 66, 8-21.
- Gerstenberger, M.C., S. Wiemer, L.M. Jones and P.A. Reasenberg (2005). Real-time forecasts of tomorrow's earthquakes in California, *Nature*, 435, 328-331.
- Gross, S. and A. Rundle (1998). A systematic test of time-to-failure analysis, *Geophys. J. Int.*, 133, 57-64.
- Gutenberg, B. and C.F. Richter (1954). *Seismicity of the Earth and Associated Phenomena*, Princeton.
- Jordan, T.H. (2006). Earthquake predictability: Brick by brick, *Seismol. Res. Lett.*, 77, 3-6.
- Kagan, Y.Y. and D.D. Jackson (1991a). Long-term earthquake clustering, *Geophys. J. Int.*, 104, 117-133.
- Kagan, Y.Y. and D.D. Jackson (1991b). Seismic gap hypothesis, ten years after, *J. Geophys. Res.*, 96, 21419-21431.
- Kagan, Y.Y. and D.D. Jackson (1994). Long-term probabilistic forecasting of earthquakes, *J. Geophys. Res.*, 99, 13685-13700.
- Kagan, Y.Y. and D.D. Jackson (2000). Probabilistic forecasting of earthquakes, *Geophys. J. Int.*, 143, 438-453.

- Kossobokov, V.G., V.I. Keilis-Borok, L.L. Romashkova and J.H. Healy (1999). Testing earthquake prediction algorithms: Statistically significant real-time prediction of the largest earthquakes in the Circum-Pacific, 1992-1997, *Phys. Earth Planet. Inter.*, 111, 187-196.
- Lombardi, A.M. and W. Marzocchi (2007). Evidence of clustering and non-stationarity in the time distribution of large worldwide earthquakes, *J. Geophys. Res.*, 112, B02303; doi: 10.1029/2006JB004568.
- Lombardi, A.M. and W. Marzocchi (2009). Double-branching model to forecast the next $M > 5.5$ earthquakes in Italy, *Tectonophysics* (in press).
- Marzocchi, W., J. Selva, A. Piersanti and E. Boschi (2003). On the long-term interaction among earthquakes: Some insight from a model simulation, *J. Geophys. Res.*, 108, 2538; doi: 10.1029/2003JB002390.
- Marzocchi, W. and A.M. Lombardi (2008). A double-branching model for earthquake occurrence, *J. Geophys. Res.*, 113, B08317; doi: 10.1029/2007JB005472.
- Marzocchi, W., J. Selva, F.R. Cinti, P. Montone, S. Pierdominici, R. Schivardi and E. Boschi (2009). On the occurrence of large earthquakes: New insights from a model based on interacting faults embedded in a realistic tectonic setting, *J. Geophys. Res.*, 114, B01307; doi: 10.1029/2008JB005822.
- MPS Working Group (2004). Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003, Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004 (MPS04), 65 pp. + 5 appendices; <http://zonesismiche.mi.ingv.it>.
- Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, *Ann. Inst. Statist. Math.*, 50, 379-402.
- Rhoades, D.A. and F.F. Evison (2004). Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pageoph*, 161, 47-72.
- Rovida, A. and the CPTI Working Group (2008). Catalogo Parametrico dei Terremoti Italiani, 1901-2006, versione 2008 (CPTI08), INGV, Milano-Pavia; <http://www.cseptesting.org/regions/italy>.
- Stein, R.S. (1999). The role of stress transfer in earthquake occurrence, *Nature*, 402, 605-609.
- Stucchi, M., P. Albini, C. Mirto and A. Rebez (2004). Assessing the completeness of Italian historical earthquake data, *Annals of Geophysics*, 47 (2-3), 659-673.
- Varotsos, P., K. Eftaxias, F. Vallianatos and M. Lazaridou (1996). Basic principles for evaluating an earthquake prediction method, *Geophys. Res. Lett.*, 23, 1295-1298.
- Zechar, J.D., D. Schorlemmer, M. Liukis, J. Yu, F. Euchner, P.J. Maechling and T.H. Jordan (2009). The Collaboratory for the Study of Earthquake Predictability perspective on computational earthquake science, *Concurrency and Computation: Practice and Experience*; doi: 10.1002/cpe.1519.
- Zhuang, J., Y. Ogata and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrence, *J. Am. Stat. Assoc.*, 97, 369-380.
- Wiemer, S. and M. Wyss (2000). Minimum magnitude of complete reporting in earthquake catalogs: examples from Alaska, the western United States and Japan, *Bull. Seism. Soc. Am.*, 90, 859-869.

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

© 2010 by the Istituto Nazionale di Geofisica e Vulcanologia. All rights reserved.