HAZGRIDX: earthquake forecasting model for $M_L \geq 5.0$ earthquakes in Italy based on spatially smoothed seismicity

Aybige Akinci[1]

[1] *Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy*

Article history
Received February 18, 2010; accepted June 25, 2010.
Subject classification:
Earthquake interactions and probability, Smoothed seismicity, Earthquake forecast, Italy.

ABSTRACT

We present a five-year, time-independent, earthquake-forecast model for earthquake magnitudes of 5.0 and greater in Italy using spatially smoothed seismicity data. The model is called HAZGRIDX, and it was developed based on the assumption that future earthquakes will occur near locations of historical earthquakes; it does not take into account any information from tectonic, geological, or geodetic data. Thus HAZGRIDX is based on observed earthquake occurrence from seismicity data, without considering any physical model. In the present study, we calculate earthquake rates on a spatial grid platform using two declustered catalogs: 1) the parametric catalog of Italian earthquakes (Catalogo Parametrico dei Terremoti Italiani, CPTI04) that contains the larger earthquakes from M_w 7.0 since 1100; and 2) the catalog of Italian seismicity (Catalogo della Sismicità Italiana, CSI 1.1) that contains the small earthquakes down to M_L 1.0, with a maximum of M_L 5.9, over the past 22 years (1981-2003). The model assumes that earthquake magnitudes follow the Gutenberg-Richter law, with a uniform b-value. The forecast rates are presented in terms of the expected numbers of $M_L > 5.0$ events per year for each grid cell of about 10 km × 10 km. The final map is derived by averaging the earthquake potentials that come from these two different catalogs: CPTI04 and CSI 1.1. We also describe the earthquake occurrences in terms of probabilities of occurrence of one event within a specified magnitude bin, ΔM0.1, in a five year time period. HAZGRIDX is one of several forecasting models, scaled to five and ten years, that have been submitted to the Collaboratory for the Study of Earthquake Probability (CSEP) forecasting center in ETH Zurich to be tested for Italy.

Introduction

Probabilistic seismic hazard analysis (PSHA) quantifies the probability of ground shaking at a specified site that exceeds a specified intensity level [Cornell 1968, SSHAC 1997]. It contains two basic ingredients: (1) the earthquake rate and rupture models, for specification of statistical distribution of earthquakes in time; and (2) the ground motion prediction equations, for estimation of ground shaking level at a site for each earthquake rupture [Field et al. 2005]. Moreover, in PSHA, it has become common practice to apply and develop seismic hazard maps to be used as input to various projects related to public and financial policies and for mitigating the seismic risk of future earthquakes [Petersen et al. 2008, Frankel et al. 1996, MPS Working Group 2004]. By now, it is also a common understanding among hazard practitioners that PSHA is affected by large uncertainties, which include both those of the estimation of the input parameters and those of the adopted sources and/or models [Cramer et al. 1996, Cramer 2001, Beauval and Scotti 2004, Bommer et al. 2005, Cao et al. 2005, Lombardi et al. 2005, Akinci et al. 2009, Akinci et al. 2010]. In this framework, in the present study, we focus on only one of the components of PSHA, the forecast earthquake rates for the probability of occurrences of earthquakes $M \geq 5.0$ in Italy, and do not deal with the relative ground motion produced by the earthquake ruptures.

Indeed, one of the important issues in PSHA has been recognized by the Regional Earthquake Likelihood Model (RELM) project [Field 2007]: the large uncertainties in earthquake hazard estimates and their impact on the seismic hazard maps [Petersen et al. 2007]. By comparing a wide range of independent, well-documented and well-defined forecasting models, the RELM working group and its members have developed more than a dozen five-year earthquake forecast models that are based on different earthquake rupture models and hypotheses in California [Field 2007]. Moreover, the Collaboratory for the Study of Earthquake Predictability (CSEP) testing experiments and infrastructures that are available in many regions of the World (e.g. California, Italy and Japan) have started gathering useful information that will improve our knowledge by providing a better understanding of earthquake complexity and earthquake predictability [Zechar et al. 2009]. In these experiments, candidate models are tested for their consistency with observations, and their performances are evaluated by comparing them to each other [Schorlemmer et al. 2007].

Therefore, one of the purposes of the present study was to provide a testable earthquake forecast model for Italy and to show the applicability of the spatially smoothed seismicity method [Frankel 1995]. The model, HAZGRIDX, is time-independent, and it gathers all of the necessary model parameters from historical and instrumental seismicity that

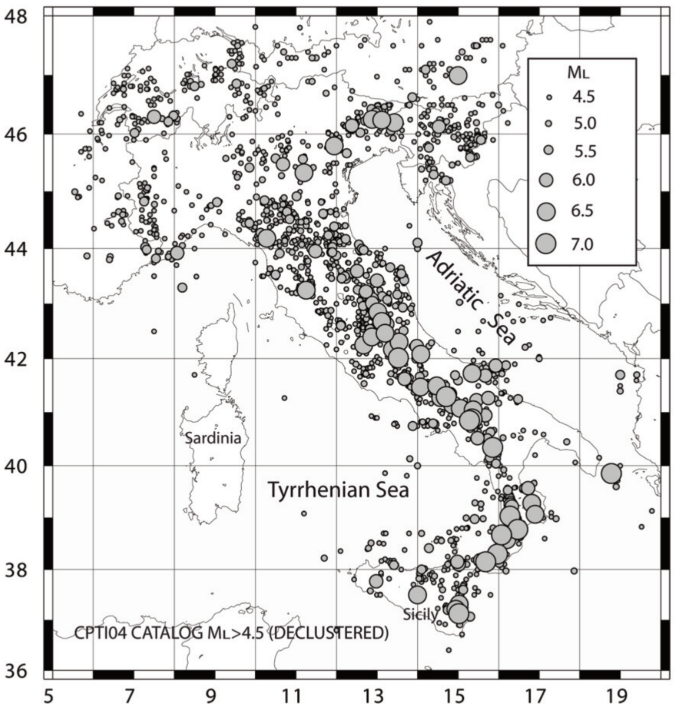


Figure 1. Epicentral distribution of earthquakes ($M_w \geq 4.65$) reported in the CPTI04 historical Italian seismic catalog [CPTI Working Group 2004].

has been spatially smoothed to different length scales. It is constructed on a spatially grid-based format so that each cell is assumed to be a potential seismogenic source. Seismicity rates in the model are determined by counting earthquakes in each grid cell with dimensions 0.1° longitude by 0.1° latitude. The model includes $M_L \geq 2.9$ events from the CSI 1.1 catalog of Italian seismicity (Catalogo della Sismicità Italiana), and $M_L \geq 4.3$ events from the CPTI04 parametric catalog of Italian earthquakes (Catalogo Parametrico dei Terremoti Italiani). These catalogs are both declustered to remove large fluctuations of seismicity rates in space and time due to aftershock sequences [Gardner and Knopoff 1974], and then the *b*-values were calculated separately from each catalog. Gridded 10^a-values are computed for each catalog using a maximum-likelihood method [Weichert 1980], and they are spatially smoothed using a two-dimensional Gaussian function with a 15-km correlation distance. For the recurrence behavior of the seismic activity, we assume a time-independent (Poisson) model, in which the occurrence of an earthquake does not change the probability of occurrence for following events. We also calculate the probability of occurrence of earthquakes of $M_L \geq 5.0$ for the five-year period, in 10 km × 10 km cells around each location. The five-year earthquake forecasts presented in the present study were obtained within the CSEP experiment in Italy, and will be tested in a forward perspective in the ETH Zurich Testing Centre of the Swiss Seismological Service.

Earthquake catalogs, declustering and completeness

Two of the important inputs for PSHA are the distribution of earthquake locations in space and their size, which can be retrieved easily from the seismic catalogs. Indeed, over the last 20 years, several research institutions in Italy have provided chronological lists of Italian earthquakes according to their epicenters, which have including their energy parameters and intensity measures, together with historical and geographical information. The record of the Italian historical earthquakes and the earthquake catalogs provide some of the most extensive details in the World. The history of earthquake investigations and studies related to earthquake cataloguing in Italy has been going on for a very long time, so that they cannot be described easily in a few words in this study. Therefore, the national catalogs of Italian seismicity over a full century can be obtained in detail from Camassi [2004] and Boschi et al. [1995]. In the following we briefly summarize the historical and the instrumental seismicity catalogs that are used in our study.

The historical earthquake catalog: CPTI04

In the late 1990s, the Italian Gruppo Nazionale per la Difesa dai Terremoti (GNDT; national group for defense against earthquakes) conducted a project to assess the probabilistic seismic hazard in Italy. Under the framework of the GNDT project, the Italian NT4.1 earthquake catalog was compiled and published by Camassi and Stucchi [1997] as a parametric catalog of damaging earthquakes from 1000-1992.

The NT4.1 catalog parameters are derived from intensity data points from about 1,000 earthquakes. The main magnitude provided in the catalog is M_S, which can be observed or calculated from the intensities, or from other kinds of magnitude measures (e.g. M_L, M_b) as well. Aftershocks and foreshocks have been removed from the catalog using a time-space window of 30 km and ±90 days. The authors pointed out that the declustering procedure adopted provides data that is quite similar to those obtained with more standard procedures, where the time-space window depends on the size of the main shock [in Gardner and Knopoff 1974].

Under the GNDT project (1996), the NT4.1 catalog was divided into four sub-catalogs (known as the PS4: northern Italy, central Italy, southern Italy and Sicily), and periods of completeness have been identified separately for each sub-catalog. An updated version of the catalog, NT4.1.1, is available from the website: http://emidius.mi.ingv.it/NT/.

An alternative catalog for large earthquakes ($M > 5.5$) was prepared and presented by the Istituto Nazionale di Geofisica e Vulcanologia (INGV): the CFTI2 catalog of strong earthquakes in Italy (Catalogo dei Forti Terremoti in Italia) by Boschi et al [1995]. Another catalog, the CPTI, was commissioned by the Italian Dipartimento della Protezione Civile (civil protection national service) and prepared by a working group in 1999 [CPTI Working Group 1999], which considered large amounts of information from the CFTI2 and NT4.1 catalogs.

The CPTI catalog adopted the Gasperini and Ferrari

Zone	Completeness according to magnitude class, M_w											
	4.76 ±0.23	4.99 ±0.23	5.22 ±0.23	5.45 ±0.23	5.68 ±0.23	5.91 ±0.23	6.14 ±0.23	6.37 ±0.23	6.60 ±0.23	6.83 ±0.23	7.06 ±0.23	7.29 ±0.23
ALPS	1871	1871	1700	1700	1530	1530	1300	1300	1300	1300	1300	1300
P. PADANA	1836	1836	1530	1530	1530	1300	1300	1100	1100	1100	1100	1100
CENTRAL	1871	1871	1650	1650	1530	1300	1300	1300	1300	1300	1300	1300
SOUTH	1871	1787	1787	1787	1787	1530	1530	1530	1400	1400	1400	1400
SICILY	1871	1871	1700	1700	1700	1530	1530	1300	1300	1300	1300	1300

Table 1. Time completeness intervals of the CPTI04 catalog used for analysis in the Italian territory [MSP04 Working Group 2004].

[2000] parameterization method, which represented a significant advance in the formalization of the procedures for the definition of parameters. It was an updated version of NT4.1, and in 1999, the CPTI catalog was compiled and published with the agreement of the entire Italian geological and seismological community.

In the present study, we have used the CPTI04 catalog, as compiled for PSHA maps in Italy [MSP Working Group 2004] and freely available for scientific purposes through the INGV web site: http://emidius.mi.ingv.it/CPTI04/. The CPTI04 catalog contains 2,550 records of earthquakes in the time window of 217 B.C. to 2003 (Figure 1). With the CPTI04 catalog designed for hazard purposes, a homogeneous magnitude scale was applied to each event. All of the magnitudes were recalculated, made homogeneous and given as two different types: moment magnitude, M_w, and instrumental magnitude, M_S. As it has become one of the "rules of the game" of the CSEP forecasting experiment that is being run for the Italian testing region, we have used only local magnitude, M_L, following the scaling relation proposed by Gasperini and Ferrari [2000]:

$$M_w = 0.812\ (\pm\, 0.032) * M_L + 1.145\,(\pm\, 0.154) \qquad (1)$$

Thus, the moment magnitudes, M_w, of the CPTI04 catalog were converted as the reverse, from M_w to local magnitudes, M_L.

The completeness of CPTI04

The completeness periods of the CPTI04 catalog were identified by the MSP Working Group [2004] for several magnitude classes in five zones, based on both historical and the statistical analyses. The completeness magnitude threshold was defined over 12 magnitude bins, and represents the centers of each magnitude class with a width of 0.23, starting from a minimum magnitude of M_w 4.76 (as the central value of the first magnitude class) to a maximum magnitude of M_w 7.39 for different periods of time. For example, the lower magnitude limit 4.76 in the complete catalog represents the center of the 4.65 to 4.87 class. Table 1 shows the completeness of the catalog in the five zones in terms of the variability at the beginning of this completeness time over twelve magnitude ranges. Figure 2 shows these five zones, which are indicated by different colors for different completeness time intervals and magnitude thresholds, as given in Table 1. The seismicity rates expressed as numbers of earthquakes per thousand years for the twelve magnitude bins ranging from M_w 4.76 to 7.29 are shown in Figure 3 [taken from the MSP Working Group 2004]. The cumulative number of earthquakes with $M_w > 4.76$ over a thousand-year period was calculated using both the activity rates and the straight Gutenberg-Richter law from the complete catalog.

The instrumental earthquake catalog: CSI 1.1

The CSI 1.1 catalog of Italian seismicity [Chiarabba et al. 2005, Castello et al. 2006] is the most recent instrumental

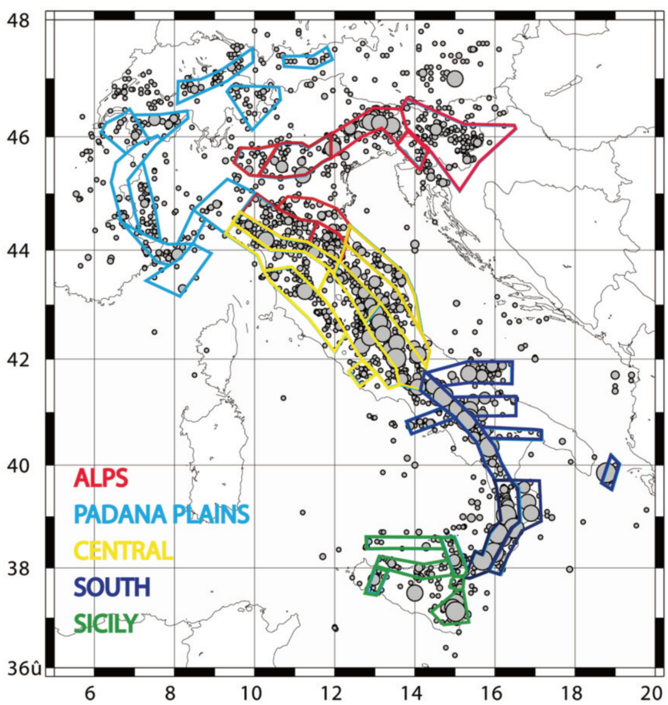
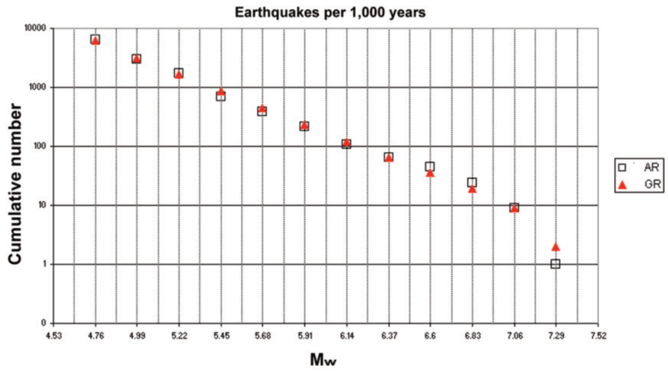


Figure 2. The completeness of the CPTI04 catalog for the five zones where the completeness magnitude threshold changed with time [MSP Working Group 2004].



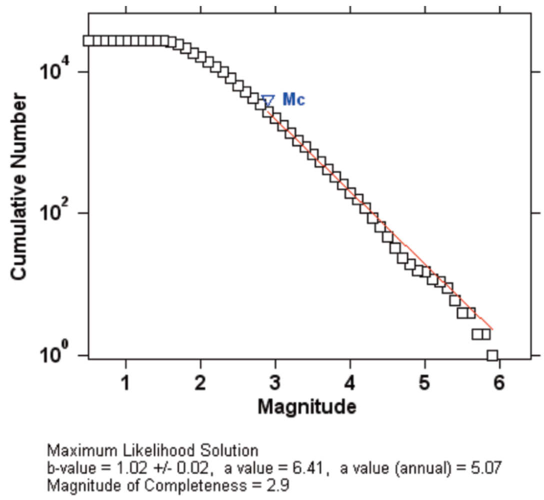


Figure 3 (left). Cumulative number of earthquakes with $M_w \geq 4.76$ calculated using activity rates (AR) and straight Gutenberg-Richter (GR) law by MSP04 [see MSP Working Group 2004, for details]. **Figure 4 (right).** Cumulative number of events per year versus magnitude from the CSI 1.1 declustered instrumental catalog, using events $M_L \geq 1.0$ (empty squares). The *b*-value is calculated using the maximum-likelihood solution (red line) for events greater than *M* 2.9 and using ZMAP computer codes [Wiemer 2001].

catalog to be compiled, in the framework of a 4-year project funded in 2004 by the Dipartimento della Protezione Civile. The catalog is freely available through the INGV web site (http://csi.rm.ingv.it/) and from the web site http://www.cseptesting.org/regions/italy.

The CSI 1.1 catalog contains located Italian earthquakes over the past 22 years (1981-2002); it includes 91,797 localized earthquakes from 136,850 recorded earthquakes, and 39,020 magnitude estimates greater than 1.5. The original database contains both the foreshocks and the aftershocks. Most of the earthquake magnitudes of $M_L < 4.0$ are located in the upper 35 km of the Earth crust [Chiarabba et al. 2005]. In the

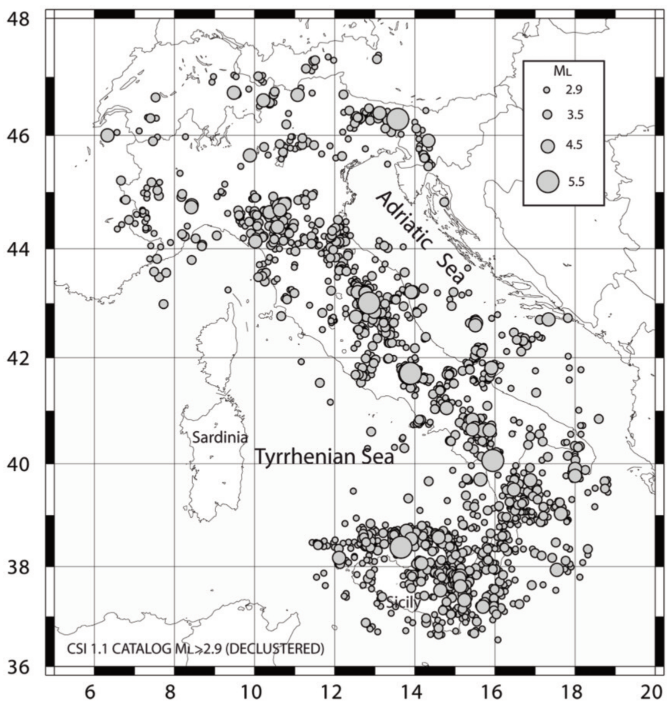


Figure 5. Epicentral distribution of $M_L \geq 2.9$ earthquakes from the CSI 1.1 declustered instrumental Italian seismic catalog [Castello et al. 2006], between 1984 and 2003.

data base, only 33 earthquakes exceeded a magnitude of 5.0 in the period between 1981 and 2002, and the largest event was the September 26, 1997, Umbria-Marche earthquake (M_w 6.0). The largest deep earthquake occurred at 200 km in depth, with a M_w of 5.8 (January 6, 1994) in the southern Tyrrhenian Sea, where the Ionian lithosphere subducts beneath the Calabrian arc.

In the present study, we removed the earthquakes deeper than 30 km from the catalog and then estimated the completeness magnitude threshold based on its departure from the linear magnitude-frequency relation, using the ZMAP computer codes [Wiemer 2001]. The magnitude-frequency distribution provided in Figure 4 suggests that the dataset may be strongly affected by the incompleteness of $M_L < 2.9$ events. This magnitude is in agreement with that observed by Schorlemmer et al. [2010], of 2.9, for the completeness at the desired detection probability level, P_E, of 0:999 for the entire mainland of Italy. Finally, we removed duplicates and dependent events from the catalog, assuming that our forecasting results will have been affected at a given location by numerous foreshock/aftershock sequences following moderate to large main shocks. Although there are many clustering/declustering methodologies in the literature [Gardner and Knopoff 1974, Reasenberg 1985, Lolli and Gasperini 2003, Console et al. 2010], we chose the declustering algorithm proposed by Gardner and Knopoff [1974],which uses the same method that was performed for declustering the CPTI04 catalog. In order to do so, the modified version of the ZMAP computer codes was used [Wiemer 2001, Murru et al. 2007]. We found 393 clusters of earthquakes as a total of 1,873 events (out of 3,674) with $M_L \geq 2.9$ since 1984 (Figure 5). The final CSI 1.1 catalog was judged to be complete down to M_L 2.9 from 1984 to 2003. The distribution of seismicity in the map (Figure 5) highlights the close relationship between seismicity and topography in the Apennines, and also represents high seismic release in active volcanoes (such as Mt. Etna) and offshore of northern Sicily.

Spatially smoothed seismicity method

The smoothed seismicity method is based on historical and instrumental seismicity generalized using exponential magnitude distributions with regionally determined *b*-values. It does not take into account any geological and tectonic observations. Due to its simplicity and performance, the smoothed seismicity approach developed by Frankel [1995] is widely used in many regions of the World to forecast future moderate and large earthquakes both over the short-term and over longer periods of time [Kagan and Jackson 1994, Helmstteter et al. 2006, Kagan et al. 2007].

In the present study, following the smoothed seismicity procedure, we calculated a probabilistic earthquake forecast for $M_L > 5.0$ for Italy over a five-year duration. First, we divide the whole observational region into a grid with spacing of 0.1° in latitude and 0.1° in longitude (about 10 km × 10 km grid length). Then we counted the number of earthquakes n_i with magnitudes greater than M_{ref} in each cell *i* of a grid. The gridded 10^a values were computed for each catalog using a maximum-likelihood method [Weichert 1980] and they were spatially smoothed using a two-dimensional Gaussian function with a 15-km correlation distance [for details of the methodology, see Frankel et al. 1996, and Petersen et al. 2008]. For each cell *i*, the smoothed value $\check{n}_i$ was obtained from Equation (2), and normalized to preserve the total number of events:

$$\tilde{n}_i = \frac{\sum_{j:\Delta_{ij} \leq 3c} n_j e^{-\Delta_{ij}/c^2}}{\sum_{j:\Delta_{ij} \leq 3c} e^{-\Delta_{ij}/c^2}} \qquad (2)$$

where Δ_{ij} is distance between the centering of the grid cells *i* and *j*. The parameter *c* is the correlation distance. The sum is taken over cells *j* within a distance of 3*c* of cell *i*.

The Weichert maximum-likelihood method was used to obtain the 10^a values, with the completeness magnitude thresholds over different periods of time as given by the MSP Working Group [2004] (see Table 1), so that the rare large and the smaller earthquakes with short recording times were obtained more accurately. Based on the assumption that the seismicity rate is constant with time, the method included each earthquake that was counted as of equal weight in the rate calculations. Thus, the seismicity rates during time periods that have more countable earthquakes (for example; time periods with lower completeness thresholds) will be affected much more strongly than the seismicity rate at other times [for details of the methodology, see Felzer and Cao 2007, Felzer 2008].

The *a*-value specifies the seismicity rate as an exponential [Gutenberg-Richter 1949] magnitude-frequency distribution: $\log N = a - bM$, where *N* is the number of events with magnitude equal to or greater than *M*, and the *b*-value

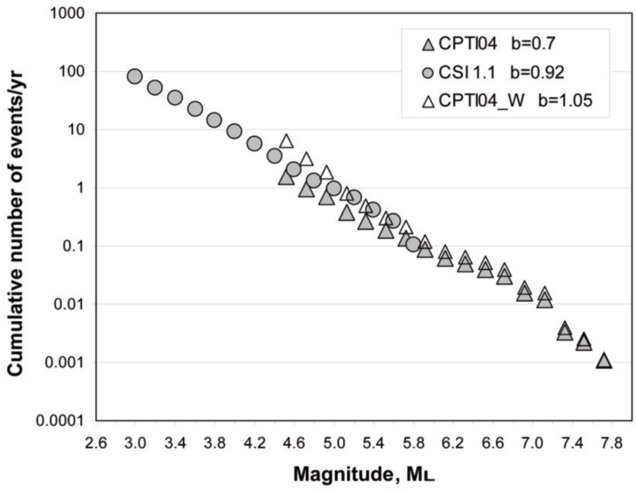


Figure 6. Cumulative number of events per year versus magnitude observed historically in Italy, CPTI04 (gray triangles), and the predicted rates using the Weicherst method, CPTI04_W (red triangles), and the cumulative number of events per year obtained from the CSI 1.1 instrumental catalog (gray circles).

is the slope of the distribution that describes the relative frequency of small and large magnitudes. The mentioned Gutenberg-Richter law is a accumulative magnitude-frequency relationship and it can be transferred to an incremental relationship [Herrmann 1977]; that is to say, the annual number of earthquakes, with a particular magnitude *M*: *N*(*M*). Incremental magnitude–frequency relationships still conform to the Gutenberg-Richter law, which has been described as $N(M) = 10^{a-bM}$. Therefore, the resulting "agrid" gives the annual rate of earthquakes in each grid cell as an incremental 10^a in the Gutenberg-Richter notation, between $M - 0.05$ and $M + 0.05$, or a 0.1 bin-width centered on $M = 0$.

For the recurrence time of the seismic activity, we assumed a time-independent (Poisson) model in which the occurrence of an earthquake does not change the probability of occurrence for following events. For a Poisson process, the probability *P* of occurrence of one or more events, in a time period *T* of interest, is given by Equation (3):

$$P = 1 - \exp(-\check{n}_i T) \qquad (3)$$

where ni is the rate of earthquakes and is the inverse of the average recurrence time.

Smoothed seismic activity rates and 5-year forecasts

The historical catalog: the CPTI04 case

The CPTI04 historical seismicity catalog has magnitudes of M_L from 4.3 to 7.8, and it is used to obtain the *a*-value distribution over the study area. In this case, we assumed that the probability of earthquake occurrence was greater where magnitude earthquakes $M_L \geq 4.3$ have occurred in the past. The smoothed activity rates for the gridded seismicity were determined using a maximum-likelihood method [Weichert 1980] and a Gaussian function filter with a 15-km correlation

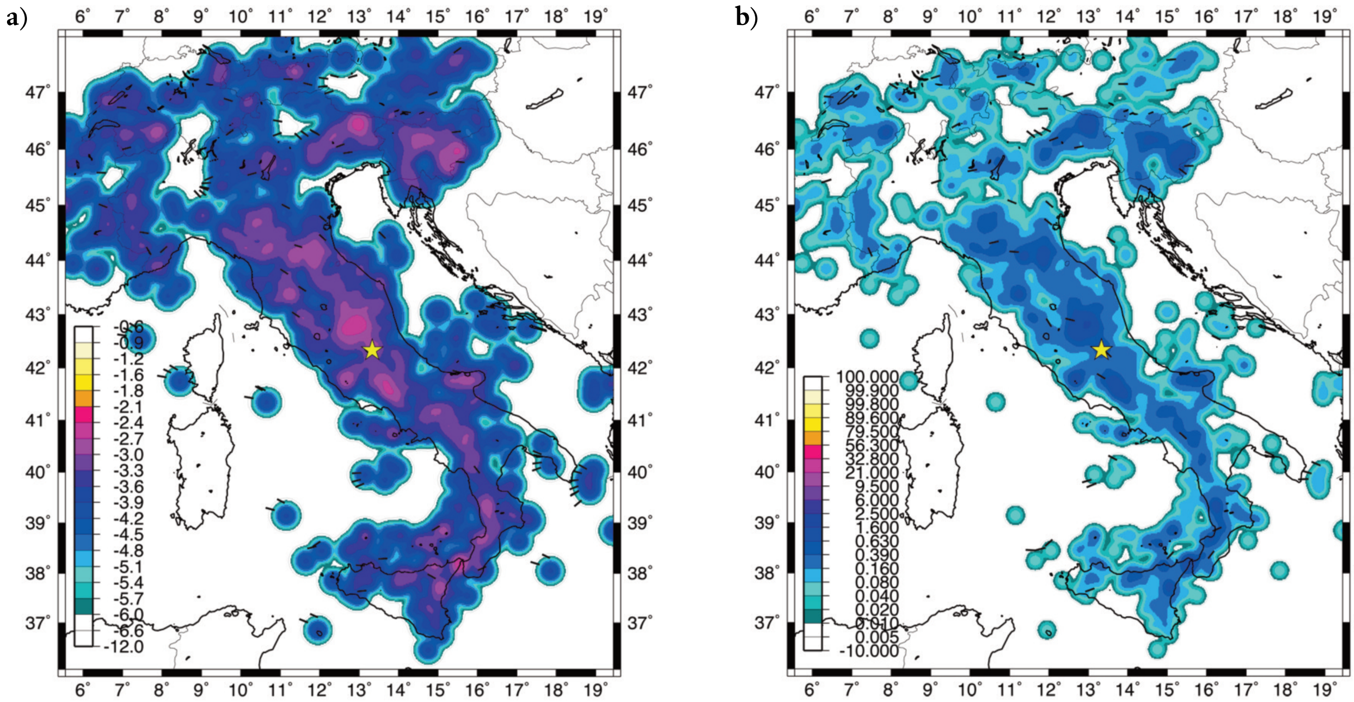


Figure 7. (a) Forecast seismicity rates obtained using a 15 km correlation distance (expected number of $M_L \geq 5.0$ events per year in each cell) for the CPTI04 case, using the spatially smoothed location of $M_L \geq 4.3$ earthquakes from 1100 to 2003. Yellow asterisk, main shock from the April 6, 2009, L'Aquila earthquake. (b) Five-year probabilities as $\log_{10}$ rates of events per year for $M_L \geq 5.0$ predicted in 10 km × 10 km cells around each location. Earthquake rates per km2 can be obtained by multiplying by 10^{-2}. For the recurrence time of the seismic activity we assumed a time-independent (Poisson) model, in which the occurrence of an earthquake does not change the probability of occurrence of following events.

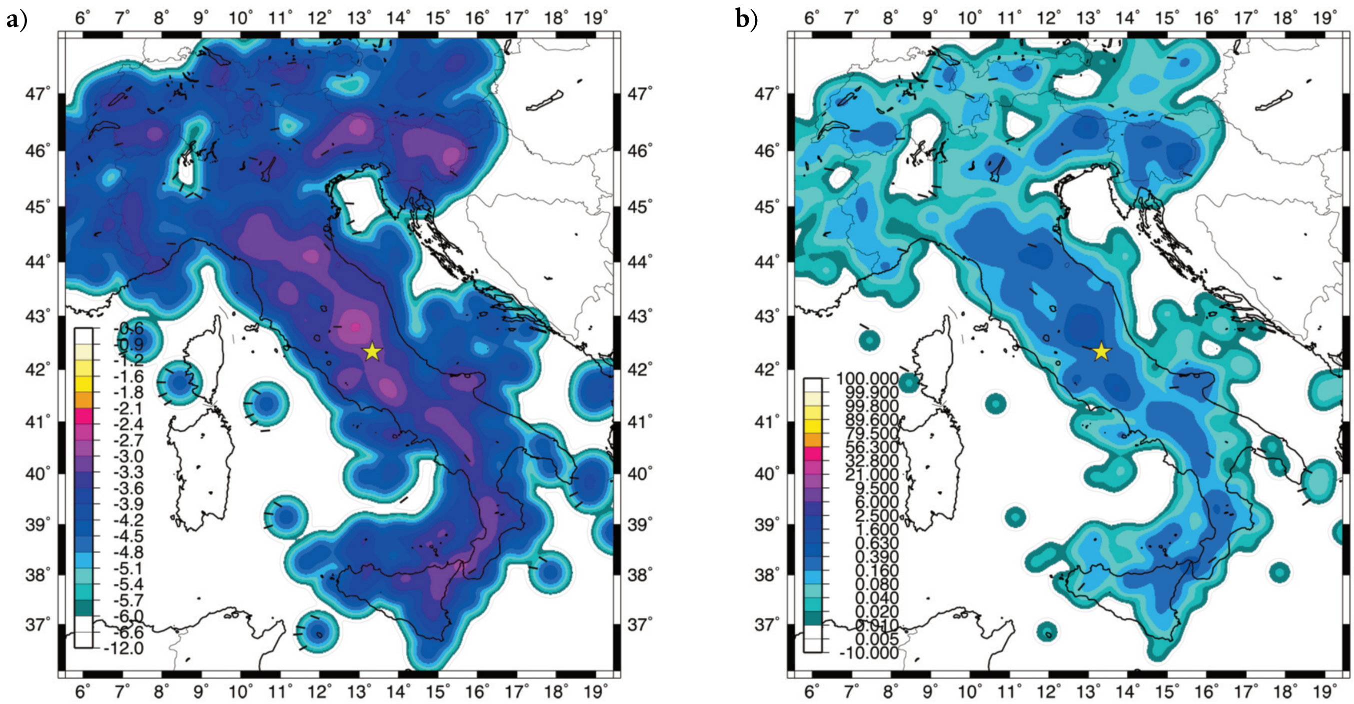


Figure 8. (a) Forecast seismicity rates obtained using a 25-km correlation distance (expected number of $M_L \geq 5.0$ events per year in each cell) for the CPTI04 case. (b) Five-year probabilities given as $\log_{10}$ rates of events per year for $M_L \geq 5.0$ predicted in 10 km × 10 km cells around each location.

distance. The *b*-value was assumed to be uniform and constant throughout the region, and $b = 1.05$ was obtained using the Weichert [1980] formulation. This is a maximum-likelihood estimate, but it allows unequal observation periods for different magnitudes. In Figure 6, the cumulative number of events in the CPTI04 catalog are compared to the number of events obtained by the Weichert method, CPTI04_W. The cumulative number of events, CPTI04_W, was much higher than those of the CTPI04 catalog for smaller magnitudes. It was closer at larger magnitudes, indicating that the larger magnitudes are more complete than the smaller magnitudes. This CPTI04_W is also in

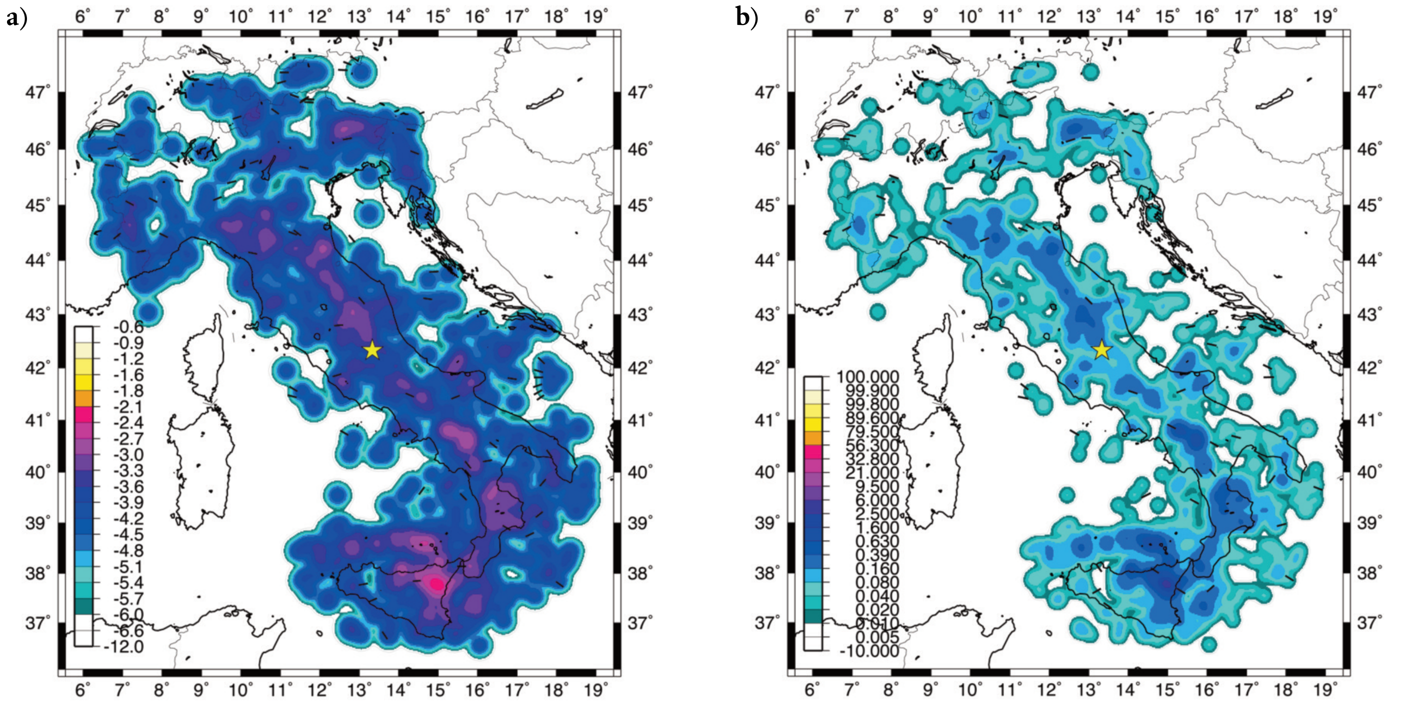


Figure 9. (a) Forecast seismicity rates obtained using 15 km correlation distance (expected number of $M_L \geq 5.0$ events per year in each cell) for the CSI 1.1 case, using the spatially smoothed locations of $M_L \geq 2.9$ earthquakes from 1984. Yellow asterisk, main shock of the April 6, 2009, L'Aquila earthquake. (b) Five-year probabilities given as $\log_{10}$ rate of events per year for $M_L \geq 5.0$ predicted in 10 km × 10 km cells around each location. Earthquake rates per km^2 can be obtained by multiplying by 10^{-2}. For the recurrence time of the seismic activity we assumed a time-independent (Poisson) model, in which the occurrence of an earthquake does not change the probability of occurrence of following events.

agreement with that obtained by the MSP Working Group [2004]. The expected total annual number of earthquakes for $M \geq 5.0$ events is 1.46. It is 7.30 and 14.6 for the five-year and 10-year forecasts, which are obtained by multiplying the annual rates by the number of years. Figure 7a shows the seismic earthquake potential model for $M_L \geq 5.0$, which was calculated using the spatially smoothed locations of $M_L > 4.3$ earthquakes with a 15-km correlation distance. Figure 7b shows the five-year probabilities represented as log10 rates of $M_L > 5.0$ events per year predicted in a 10 km × 10 km cell around each location. To get earthquake rates per km^2, we simply multiply the number in each cell by 10^{-2}.

The correlation distance of the Gaussian function was chosen as $c = 15$ km for the case of CPTI04, assuming that the earthquakes can occur anywhere in the fault region at distances larger than tens of kilometers, and we decided to define the shortest smoothing radius as not to be less than 10 km. The error in the epicenter location was assumed to be around 50 km, where the correlation distance becomes 1/3 of this (around 15 km, as chosen). We believe that it suits larger location uncertainties of older events, and partly larger rupture areas of larger events. One problem with the smoothing method is apparent in some parts of Italy, especially in the central belt of the Apennines, where the seismicity that occurs in narrow linear zones is over-smoothed.

We have also produced several trail forecast maps using correlation distances of 5, 10 15, 20, 25 and 50 km for $M_L \geq 4.3$ events. Each earthquake was smoothed into a circle with radius equal to $3c$. The 5 km case presented very grainy-looking maps and the 50 km case produced much more smoothed ones, showing how the seismicity rates can be affected by this subjective choice of the correlation distance. The Gaussian correlation distance considered was 25 km, as estimated by Console and Murru [2001]. Figure 8a, b shows the seismic earthquake potential model for $M_L \geq 5.0$, and the five-year forecast that was calculated using a 25-km correlation distance, which spatially smoothes the location of $M_L \geq 4.3$ earthquakes.

As can be seen from Figures 7a and 8a, the 25-km correlation distance spread out and smoothed the seismicity too much. The higher seismicity rates observed using a smaller correlation distance disappeared in the same regions, e.g. the southern part of the Apennines (Figure 7a). Therefore, we used a 15-km correlation distance in our calculations. The higher rates in Figure 7a mainly concentrate along the central Apennines, for the Friuli, Campania and Messina areas at about 4×10^{-3} per 10 × 10 km^2 area. The mean occurrence probabilities to have one earthquake at least with $M_L \geq 5.0$ in the next five years range from 1.5% to 2.5% in those regions under the Poisson model.

Instrumental catalog: the CSI case:

The declustered instrumental earthquake catalog is assumed to be sufficiently complete for the entire country, with a magnitude range of $2.9 \leq M_L < 5.9$ since 1984. Therefore, the spatially smoothed seismic activity rates were derived from the $M_L \geq 2.9$ events, to estimate the probabilities of future earthquakes of $M_L \geq 5.0$ in Italy in the next five-year period.

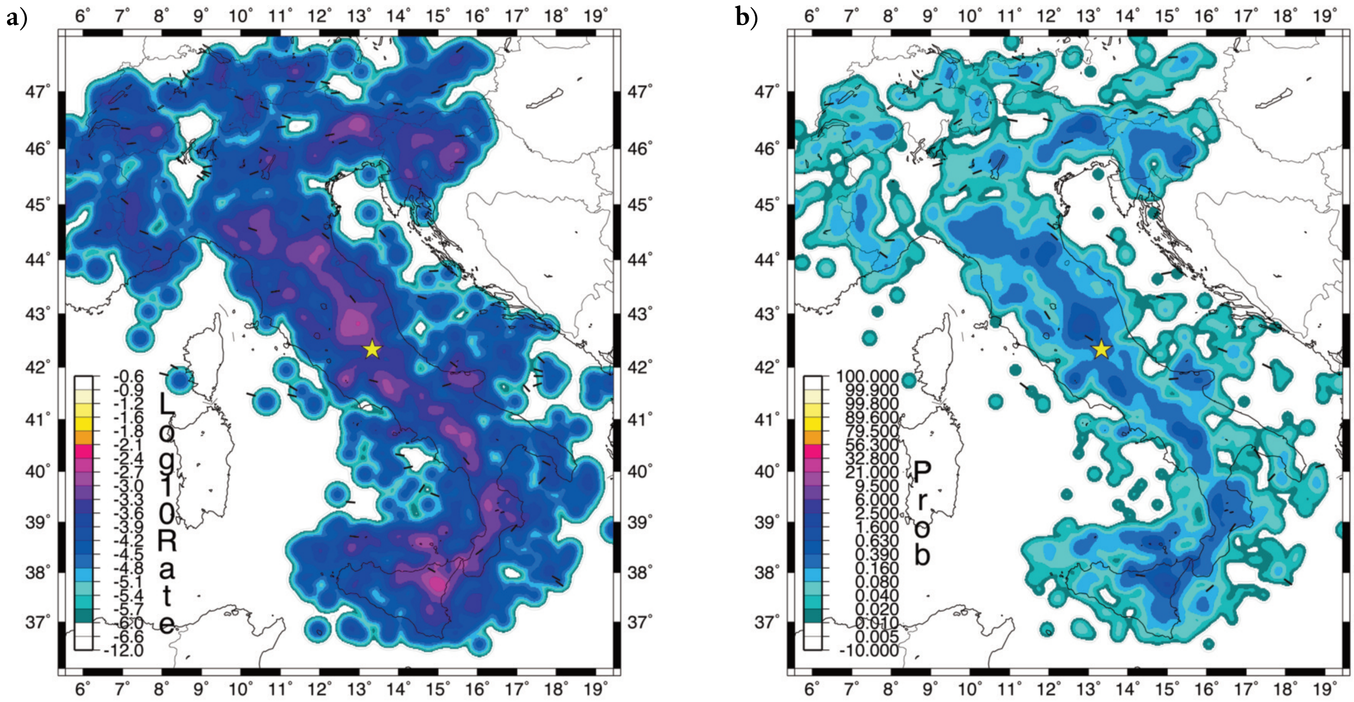


Figure 10. (a) Forecast seismicity rates (expected number of $M_L \geq 5.0$ events per year in each cell) for the combined case, for submission to the CSEP Testing Center in ETH Zurich. (b) Five-year probabilities given as $\log_{10}$ rates of events per year for $M_L \geq 5.0$ predicted in 10 km × 10 km cells around each location. Earthquake rates per km² can be obtained by multiplying by 10^{-2}. For the recurrence time of the seismic activity we assumed a time-independent (Poisson) model, in which the occurrence of an earthquake does not change the probability of occurrence of following events.

The *b*-value was calculated from the complete part of the catalog as $b = 0.92$ (Figure 6). After visual inspection, the correlation distance of 15 km was chosen, which smoothes the earthquake activity rates of the CSI 1.1 catalog events.

Figure 9a shows the seismic earthquake potential model for $M_L \geq 5.0$ that was calculated using the spatially smoothed location of $M_L \geq 2.9$ earthquakes. Earthquake occurrence is again modeled according to a time-independent (Poisson) process. The expected total annual number of earthquakes for $M \geq 5.0$ events is 1.04. It is 5.20 and 10.4 for five-year and 10-year forecasts, which are obtained by simply multiplying the annual rate by the number of years. Figure 9b shows the five-year probabilities given as log10 rates of events per year for $M_L \geq 5.0$, as predicted from each cell.

In the CSI 1.1 case, the higher seismicity rates are also mainly concentrated along the Apennines and offshore of northern Sicily, as 4×10^{-3} within a zone of 100 km². Significantly high rates are observed around Mount Etna, at 1.5×10^{-2} per 100 km², as the catalog contains many earthquakes with magnitudes >4.0 in the area.

Combined case:

In the combined case, as the final forecasting model for submission to the CSEP Testing Center, we averaged the rates calculated from the two catalogs (CPTI04 and CSI 1.1) without normalizing the occurrence rates one to the other, as the two catalogs were judged to be complete after the procedures followed in the previous sections (Figure 6). We obtained slightly higher seismicity rates from the historical catalog than the instrumental one. Therefore, by averaging the rates, we allow the historical part of the record to have influence despite its high completeness threshold.

Figure 10a, b shows the forecast seismicty rate (expected number of $M_L \geq 5.0$ events per year in each cell) for the combined case, and the predicted five-year probabilities in each grid cell, presented as $\log_{10}$ rates of events per year for $M_L \geq 5.0$, respectively. Because the HAZGRIDX model is time independent, the most hazardous zones are the areas that have been more affected over the last centuries by important earthquakes (like the central Apennine, Friuli, Calabria and Irpinia regions, and the eastern part of Sicily). The highest rate on the map corresponds to the Mount Etna area, at 8×10^{-3}, with 6% probability of occurrence of events with $M_L \geq 5.0$ over the next five-years.

Conclusions and discussions

We have here used a spatially smoothed seismicity method to estimate the probability of earthquakes, as a function of space and magnitude, and from both the historical and instrumental seismicity. We estimated the spatial density of seismicity in each cell, smoothing with an isotropic Gaussian filter the locations of $M_L \geq 4.3$ and $M_L \geq 2.9$ earthquakes of the CPTI04 and CSI 1.1 catalogs, respectively. Our final forecast is given as an average number of events per year in each cell within a zone of 100 km². For the instrumental catalog CSI 1.1, the average number of events $M \geq 5.0$ is 1.04 per year, whereas 1.46 is the yearly rate for the CPTI04 historical catalog. Results indicate some probable potential zones for future earthquakes of higher magnitudes in Italy, like the central Apennines, Friuli, the Irpinia region, and Mount Etna, with about 6% probability of occurrence within 100 km^2 areal zone in the next five-year.

The data derived in this study are mainly driven by the earthquake catalogs and by their distributions in space and size. Seismic rate changes or earthquake-sized distributions are quite sensitive to earthquake catalogs and their completeness. Therefore, catalog completeness is one of the crucial topics in earthquake forecasting, and cannot be disregarded because the reliability and the quality of any statistical analysis are strongly affected by earthquake catalog data. Indeed, Marzocchi et al. [2003] showed that the forecasting ability of any model can easily be disturbed by the use of incomplete catalogs. Moreover, there are many more sophisticated declustering methods that are different from the one used in this study. The application of these approaches might result in different forecasting rates in the region, and they present improvements of the present study to be investigated further in the future.

The importance of the *b*-value for seismic hazard assessment is commonly accepted and it is easily seen from the hazard formulation [Cornell 1968]. A low *b*-value increases the hazard, while a high one may underestimate it. Ideally, an adequate density of seismic data would allow us to obtain a *b*-value distribution with smaller variance over the grid cells. Unfortunately, we did not have sufficient data to robustly estimate *b*-values over the whole study area in about a 100 km^2 dimension for magnitudes equal to 4.3 and greater. However, we assume a uniform *b*-value distribution over the Italian region.

In the present study, the probabilities of future large earthquake occurrence in the next five years is calculated based on the assumption that earthquake processes have no memory, i.e., the occurrence of a future earthquake is independent of the occurrence of previous earthquakes from the same source. Although this is the most widely used hypothesis in probabilistic seismic hazard analysis, its assumption is not physically valid for individual fault sources, given that the process of stress builds up and its release is inherently time dependent. Moreover, other factors such as clustering, static-elastic fault interactions, dynamic-stress changes and viscoelastic stress transferred from earthquakes on nearby faults can also influence the short-term and long-term probabilities for earthquake recurrence. Even though it is still a challenge to answer what probability distribution and occurrence model best describes earthquake behavior, in recent years, time-dependent earthquake occurrence models have become increasingly a part of probabilistic seismic hazard analysis [Pace et al. 2006, Petersen et al. 2007, Akinci et al. 2009]. Presently, it is difficult to specify a suitable time-dependent model for large-earthquake occurrence because of insufficient data, but it is important to discriminate the candidate models that differ significantly from one another, and to evaluate differences in a rigorous and quantitative way.

Data and sharing resources

In this study, we used both the CPTI04 historical catalog (http://emidius.mi.ingv.it/CPTI04), which is prepared by the CPTI Working Group [2004], and the CSI 1.1 catalog of Italian seismicity [Chiarabba et al. 2005, Castello et al. 2006], which is the most recent instrumental catalog, and which is freely available through the INGV web site (http://csi.rm.ingv.it/) and from the web site http://www.cseptesting.org/regions/italy. Many of the plots were prepared using the Generic Mapping Tools, version 4.2.1 (www.soest.hawaii.edu/gmt) [Wessel and Smith 1998].

Acknowledgements. This study was performed in the framework of the activities of project S2, coordinated by Warner Marzocchi and Ezio Faccioli, and as one of the INGV-DPC 2006-09 applied research programs with the financial support of the Italian Dipartimento della Protezione Civile (DPC). Our thanks to Giuseppe Falcone and Anna Maria Lombardi of INGV for their availability and support at each step of this study. The authors are also grateful to two anonymous reviewers for their comments and suggestions that improved the manuscript.

References

Akinci, A, F. Galadini, D. Pantosti, M. Petersen, L. Malagnini and D. Perkins (2009). Effect of time dependence on probabilistic seismic-hazard maps and deaggregation for the central Apennines, Italy, Bull. Seism. Soc. Am., 99, 585-610.

Akinci, A., D. Perkins, A.M. Lombardi and R. Basili (2010). Uncertainties in the estimation of the probability of occurrence of strong earthquakes from individual seismological sources in the Apennines, Italy, J. Seismol., 14, 95-117.

Beauval, C. and O. Scotti (2004). Quantifying sensitivities of PSHA for France to earthquake catalog uncertainties, truncation of ground-motion variability, and magnitude limits, Bull. Seismol. Soc. Am., 94, 1579-1594.

Bommer, J., F. Scherbaum, H. Bungum, F. Cotton, F. Sabetta and N.A. Abrahamson (2005). On the use of logic trees for ground-motion prediction equations in seismic hazard analysis, Bull. Seism. Soc. Am., 95, 377-389.

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.

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/.

Camassi, R. (2004). Catalogs of historical earthquakes in Italy, Annals of Geophysics, 47 (2-3), 644-657.

Cao, T., M. Petersen and A. Frankel (2005). Model uncertainties of the 2002 update of California seismic hazard maps,

Bull Seism Soc Am., 95, 2040-2057.
Castello, B., G. Selvaggi, C. Chiarabba and A. Amato (2006). CSI Catalogo della sismicità italiana 1981-2002, versione 1.1, INGV-CNT, Roma; http://csi.rm.ingv.it/.
Chiarabba, C., L. Jovane and R.D. Stefano (2005). A new view of Italian seismicity using 20 years of instrumental recordings, Tectonophysics, 395, 251-268.
Console, R. and M. Murru (2001). A simple and testable model for earthquake. clustering, Journal of Geophysical Research, 106, 8699-8711.
Console, R., D.D. Jackson and Y.Y. Kagan (2010). Using ETAS model for Catalog Declustering and Background Assessment, Pure and Applied Geophysics, 167 (6-7), 819-830; doi: 10.1007/s00024-010-0065-5.
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), INGV, Bologna; http://emidius.mi.ingv.it/CPTI04/.
Cramer, C.H., M.D. Petersen and M.S. Reichle (1996). A Monte Carlo approach in estimating uncertainty for a seismic hazard assessment of Los Angeles, Ventura and Orange Counties, California, Bull Seism. Soc Am., 86, 1681-1691.
Cramer, C.H. (2001). A seismic hazard uncertainty analysis for the New Madrid seismic zone, Eng. Geol., 62, 251-266.
Gardner, J.K. and L. Knopoff (1974). Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian?, Bull Seism. Soc Am., 64, 1363-1367.
Gasperini, P. and G. Ferrari (2000). Deriving numerical estimates from descriptive information: the computation of earthquake parameters, in CFTI 3, Annals of Geophysics, 43, 729-746.
Gutenberg, B. and C.F. Richter (1944). Frequency of earthquakes in California, Bull Seism. Soc Am., 142, 185-188.
Felzer, K.R. and T. Cao (2007). Appendix H - WGCEP historical California earthquake catalog, in Earthquake Rate Model 2.2 of the Working Group on California Earthquake Probabilities, http://www.wgcep.org/, USGS Open File Report 2007-1437, CGS Special Report 203.
Felzer, K.R. (2008), Calculating California seismicity rates, Appendix I in The UniformCalifornia Earthquake Rupture Forecast, version 2 (UCERF 2), U.S. Geological Survey Open-File Report 2007-1437, and California Geological Survey Special Report 203, 42 pp.
Field, E.H., V. Gupta, N. Gupta, P. Maechling and T.H Jordan (2005). Hazard map calculations using grid computing, Seismological Research Letters, 76, 565-573.
Field, E.H. (2007). Overview of the Working Group for the Development of Regional Earthquake Likelihood Models (RELM), Seismological Research Letters, 78, 7-16.
Frankel, A. (1995). Mapping seismic hazard in the central and eastern United States, Seismological Research Letters, 66, 8-21.
Frankel, A., C. Mueller, T. Barnard, D. Perkins, E.V. Leyendecker, N. Dickman, S. Hanson and M. Hopper (1996). National seismic-hazard maps; documentation June 1996: U.S. Geological Survey Open-File Report 96-532, 110 pp.
Helmstetter, A., Y.Y. Kagan and D.D. Jackson (2006). Comparison of short-term and time-independent earthquake forecast models for southern California, Bull Seism. Soc Am., 96, 90-106.
Herrmann, R.B. (1977). Recurrence relations, Earthquake Notes, 48, 47-49.
Kagan, Y.Y. and D.D. Jackson (1994). Long-term probabilistic forecasting of earthquakes, Journal of Geophysical Research, 99, 13685-13700.
Kagan, Y.Y., D.D. Jackson and Y. Rong (2007). A Testable Five-Year Forecast of Moderate and Large Earthquakes in Southern California Based on Smoothed Seismicity, Seismological Research Letters, 78, 94-98.
Lolli, B. and P. Gasperini (2003). Aftershocks hazard in Italy, Part I: Estimation of time-magnitude distribution model parameters and computation of probabilities of occurrence, J. Seism., 7, 235-257.
Lombardi, A.M, A. Akinci, L. Malagnini and C.H. Mueller (2005). Uncertainty analysis for seismic hazard in Northern and Central Italy, Annals of Geophysics, 48, 853-865.
Marzocchi, W., L. Sandri and E. Boschi (2003). On the validation of earthquake-forecasting models: The case of pattern recognition algorithms, Bull. Seism. Soc.Am., 93, 1994-2004.
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.
Murru, M., R. Console, G. Falcone, C. Montuori and T. Sgroi (2007). Spatial mapping of the *b*-value at Mt. Etna, Italy, using earthquake data recorded from 1999 to 2005, J. Geophys. Res., 12, B12303; doi: 10.1029/2006JB004791.
Pace, B., L. Peruzza, G. Lavecchia and P. Boncio (2006). Layered seismogenic source model and probabilistic seismic hazard analyses in Central Italy, Bull Seism. Soc Am., 96, 107-132.
Petersen, M.D., T. Cao, K.W. Campbell and A.D. Frankel (2007). Time-independent and time-dependent seismic hazard assessment for the State of California: uniform California earthquake rupture forecast model 1.0, Seismol Res Lett., 78, 99-109.
Petersen, M.D., A.D. Frankel, S.C. Harmsen, C.S. Mueller, K.M. Haller, R.L. Wheeler, R.L. Wesson, Y. Zeng, O.S.

Boyd, D.M. Perkins, N. Luco, E.H. Field, C.J. Wills and K.S. Rukstales (2008). Documentation for the 2008 Update of the United States National Seismic Hazard Maps, U.S. Geological Survey Open-File Report 1128, 61 pp.

Reasenberg, P. (1985). Second-order moment of Central California seismicity, 1969-82, Journal of Geophysical Research, 90 (B7), 5479-5496; doi: 10.1029/JB090iB07p05479.

Schorlemmer, D., M.C. Gerstenberger, S. Wiemer, D.D. Jackson and D.A. Rhoades (2007). Earthquake Likelihood Model Testing, Seismological Research Letters, 78, 17-29.

Schorlemmer, D., F. Mele and W. Marzocchi (2010). A completeness analysis of the National Seismic Network of Italy, J. Geophys. Res., 115, B04308; doi: 10.1029/2008JB006097.

SSHAC, Senior Seismic Hazard Analysis Committee (1997). Recommendations for Probabilistic Seismic Hazard Analysis:Guidance on Uncertainty and Use of Experts, U.S. Nuclear Regulatory Commission, U.S. Dept. of Energy, Electric Power Research Institute; NUREG/CR-6372, UCRL-ID-122160, vol. 1-2. Also a review of the document by National Academy Press, Washington, DC, 73 pp.

Weichert, D.H. (1980). Estimation of the earthquake recurrence parameters for unequal observation periods for different magnitudes, Bull. Seism. Soc. Am., 70, 1337-1346.

Wessel, P. and W.H.F. Smith (1998). New, improved version of the Generic Mapping Tools Released, EOS Trans. AGU, 79, 579.

Wiemer, S. (2001). A software package to analyze seismicity: ZMAP, Seismol. Res. Letts., 72, 373-382.

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, 22 (12), 1836-1847; doi: 10.1002/cpe.1519.

*Corresponding author: Aybige Akinci,
Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy;
e-mail: akinci@ingv.it