

# The Proportional Hazard Model as applied to the CSEP forecasting area in Italy

Licia Faenza<sup>1,\*</sup> and Warner Marzocchi<sup>2</sup>

<sup>1</sup> Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale Terremoti, Roma, Italy

<sup>2</sup> Istituto Nazionale di Geofisica e Vulcanologia, sezione di Roma, Italy

## Article history

Received October 29, 2009; accepted April 23, 2010.

## Subject classification:

Earthquake forecasting, Hazard function, Statistical analysis, Italian seismicity.

## ABSTRACT

This study presents the strategies adopted to modify the Proportional Hazard Model to fit the requirements for forecasting testing within the Collaboratory Study for Earthquake Predictability (CSEP) experiment. The model was originally proposed to study the spatiotemporal distribution of  $M \geq 5.5$  seismicity in Italy, through two spatial models: a regular grid, and a seismotectonic zonation. A prospective 10-year-forecast test has already been ongoing since 2005, and the results are available on the internet (<http://earthquake.bo.ingv.it>). For that test, we have reported the probability maps of  $M \geq 5.5$  earthquakes for the next 10 years for the two spatial models. As the original model is time-dependent, it is updated every year, and also immediately after the occurrence of a target event, e.g.,  $M_w 5.5$ . Although this prospective test is continuing and the model updates probabilities that are different from those of the CSEP experiments, we argue that a full evaluation of the model can only be achieved through this CSEP testing, where the performances of different models are compared using the same rules and tests. The major modification we have introduced into our model is the simulation of the expected numbers of events in the exposure time  $\Delta t$ . This is performed considering the probability that an event occurs in  $\Delta t$ , and evaluating the change this will cause in the expected number of events. This procedure is also implemented for the first and second generation of aftershocks.

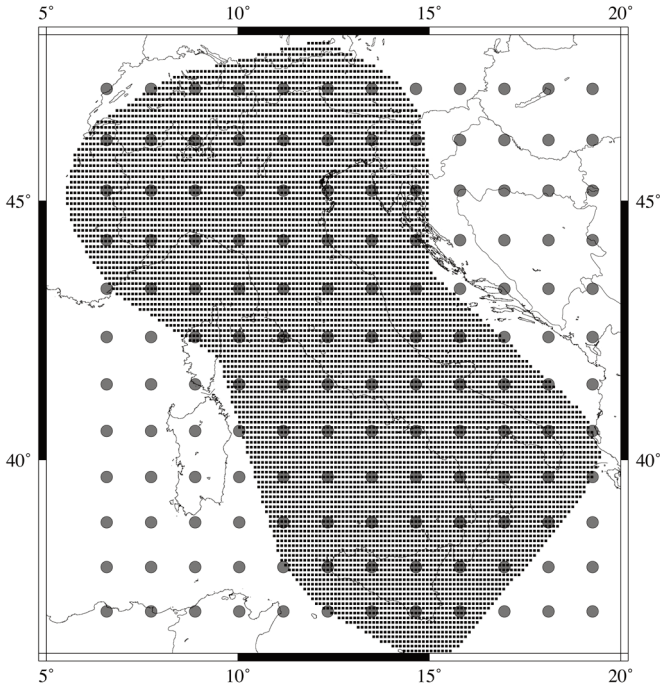
## 1. Introduction

In recent years, a multivariate non-parametric statistical method has been applied in different zones of the World to study spatiotemporal earthquake occurrence. The model is known as the Proportional Hazard Model (PHM), and it characterizes the temporal dependence of the hazard function, the empirical trend of which provides an immediate tool for the recognition of the main characteristics of the statistical distribution. The hazard function represents the instantaneous conditional probability of an occurrence at time  $t$  upon survivor until time  $t$ . By studying the empirical trend, it is possible to recognize the time behavior for earthquake occurrence. The valuable properties of this model are that it is non-parametric in the

temporal domain, and it can be used to simultaneously integrate different kinds of information. As a result, it does not assume any *a-priori* statistical distribution of the events, leaving the data to speak for themselves. This characteristic is particularly attractive, because it allows investigation of the physics of the earthquake occurrence process without forcing any *a-priori* model (i.e., Poisson, Seismic Gap, Seismic Cluster, and others). Also, it can account for tectonic/ physics factors that potentially influence the spatiotemporal distribution of the events, and it tests their relative importance.

The PHM has been used to provide a representation of the spatiotemporal distribution of destructive earthquakes in Italy ( $M_w \geq 5.5$ ) [Faenza et al. 2003, Cinti et al. 2004, Faenza and Pierdominici 2007], of medium-sized central European earthquakes ( $M_w \geq 4.0$ ) [Faenza et al. 2009], and of large earthquakes worldwide ( $M_w \geq 7.0$ ) [Faenza et al. 2008]. Sketching out the main results, as with other studies, we have seen that the time cluster of events is a feature that marks the occurrence of earthquakes within an area. The time length of the cluster might be a few years; after this time, the distribution of the earthquakes becomes a Poisson distribution. Of note, the time clustering appears to be longer than what might be expected for a typical aftershock sequence, as reported in Faenza et al. [2004] for Italian seismicity, and in Faenza et al. [2009] for central European seismicity. This suggests that the time cluster is a scale-independent feature, both for magnitude and for areal extension of the sample. We also note that all of the results have been checked through independent datasets.

In the light of the results obtained by applying the model to medium to large Italian seismicity [Faenza et al. 2003, Cinti et al. 2004], we have built a forecasting test for  $M_w \geq 5.5$  earthquakes in Italy that can be found at: <http://earthquake.bo.ingv.it>. This webpage reports two probability maps for the occurrence of the next large



**Figure 1.** The 169 nodes of the regular grid (light gray circles), and the denser CSEP testing grid (dark gray shading).

events with  $M_w$  5.5+ in Italy (the target events) over the next 10 years, which were calculated using our statistical model, the Italian  $M_w$  5.5+ catalog for the last four centuries, and the two spatial grids. The test is updated every January 1 and after each target event, and it has been running since 2005. The test performed satisfactory for the last strong Italian earthquake, the April 6, 2010, L'Aquila event with a magnitude  $M_w$  6.3 [Pondrelli et al. 2010], and for its subsequent event on April 7, 2010, with a magnitude  $M_w$  5.6, as the probability ranking of the area in which these two earthquakes occurred is among the highest. In more detail, we updated our maps at the beginning of January 2009. The main L'Aquila event took place in an area with a very high probability ranking for both of the two maps. Specifically, the main event occurred in a seismotectonic zone that was ranked sixth out of the 34 highest rankings, and the closest node of the grid had the second highest probability density (see the webpage in the results section). As a result of the time cluster behavior shown by the seismicity, the probability in that area increased, and thus the second  $M_w$  5.5+ event occurred in the area with the highest probability ranking.

Although this prospective testing of the model is already ongoing, we recognize the importance of including the model in a larger experiment like the CSEP [Schorlemmer et al. 2010], together with other forecasting models. For this purpose, we have modified some of the constraints of the original model to fit the CSEP requirement for 5-year and 10-year forecasting. Thus, we describe here the adjustments that we have made, after a concise introduction to the statistical model.

## 2. The Proportional Hazard Model and its application to Italian seismicity

The PHM is a statistical model that is used to describe the spatiotemporal distribution of a point-process, such as the occurrence of an earthquake. PHM was first introduced by Cox [1972] and by Kalbfleisch and Prentice [1980]. In this study, we introduce the basic elements needed for an understanding of the PHM principles. Its mathematical formulation was described by Kalbfleisch and Prentice [1980], Faenza et al. [2003] and Faenza [2005], from which we have taken our references.

In the formulation, two types of random variables (RV) are considered: the inter-event time (IET), i.e., the time interval between two consecutive events; and the censoring time (CT), i.e., the time between the most recent event in the catalog and the end of the catalog itself. The crucial point of the model is to combine the RVs with other information that can influence the occurrence of the events, which can be qualitative or quantitative. This information is represented as explanatory variables, called covariates, which are linked to each RV; they can be related to the RV itself or to the area the RV belongs to. This is carried out in terms of a model in which the hazard function for a generic time  $t$  since the last event is:

$$\lambda(t; z) = \lambda_0(t) \exp(z\beta) \quad (1)$$

where  $\lambda_0(t)$  is an arbitrary unspecified baseline hazard function,  $z$  is the covariate vector and  $\beta$  is a column vector that provides the weight for each covariate. The hazard function  $\lambda(t; z)$  is therefore composed of two parts: one with temporal dependence ( $\lambda_0(t)$ ); and the other with information about the processes carried by other factors ( $\exp(z\beta)$ ).

The vector of coefficients  $\beta$  and  $\lambda_0(t)$  are the two unknown parameters, and these are estimated through a Maximum Likelihood Estimation strategy [Faenza et al. 2003: section 2; Faenza 2005]. There are some assumptions made behind this model. The baseline hazard function  $\lambda_0(t)$  is independent of the covariates  $z$ . From a physical point of view, this implies that the mechanisms of earthquake occurrence are the same for all  $z$  under study; the covariates  $z$  act as a multiplicative factor that can only rescale the baseline hazard function. The model also assumes that the spatiotemporal coverage of the catalog used (four centuries; see below) is representative of the spatiotemporal distribution of the events.

The evaluation of the hazard function is based on the empirical survivor function. There is a biunique relationship between the hazard function and the survivor function:

$$S(t) = \exp\left(-\int_0^t \lambda(u) du\right) \quad (2)$$

and for the PHM, this relationship becomes:

$$S(t; z) = \exp\left[-\int_0^t \lambda_0(u) \exp(z\beta) du\right] = S_0(t)^{\exp(z\beta)}. \quad (3)$$

A simple way to display the trend of  $\lambda_0(t)$  is through a comparison of the empirical survivor function  $S_0(t)$  and the survivor function of a Poisson process. We applied a double logarithmic transformation:

$$u(t) = \ln\{-\ln[S_0(t)]\} \quad (4)$$

to obtain the RV,  $u(t)$ , as asymptotically normally distributed [e.g. Kalbeisch and Prentice 1980, Faenza et al. 2003]. When the transformation of Equation (4) is applied to the survivor function of a Poisson process, this gives  $u_p(t) = \ln\lambda + \ln(t)$ , where  $\lambda$  is the mean of the distribution. We can then estimate the residuals  $\epsilon(t)$  as:

$$\epsilon(t) = u(t) - u_p(t); \quad (5)$$

The function  $\epsilon(t)$  versus  $t$  shows the departures of the empirical survivor function from the theoretical Poisson distribution as a function of elapsed time. By looking at Equation (4) and at the relationship between the survivor function and the hazard function in Equation (2), it is easy to see that the trend of  $\epsilon(t)$  has a shape that is comparable to the trend of  $\lambda(t)$  [Kalbfleisch and Prentice 1980].

This model was applied by Faenza et al. [2003] and Cinti et al. [2004] to Italian seismicity for the time period of 1600 to 2004. Here, we have integrated the catalog used by these two studies with the bulletin of the INGV, Istituto Nazionale di Geofisica e Vulcanologia (Bollettino Sismico Italiano; <http://bollettinosismico.rm.ingv.it/>), for the time domain of 2004 to 2009. In this study, we use the parametric seismic catalog published by the CPTI Working Group [1999, 2004], which we have integrated with the more recent data from the INGV bulletin. To have a more complete set of events, we considered the earthquakes with  $M_w$  5.5+ that have occurred since 1600; the catalog is considered to be complete for this time-magnitude window [Faenza et al. 2003].

### 2.1. The two spatial grids:

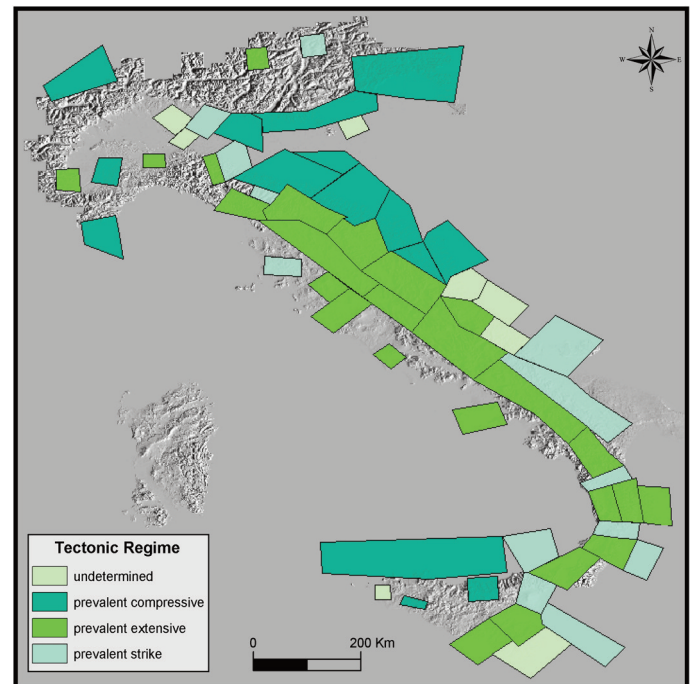
#### the geometrical grid and the seismotectonic zonation

The geometrical grid used in Faenza et al. [2003] is a  $13 \times 13$  grid between the latitudes of  $36^\circ\text{N}$  and  $48^\circ\text{N}$ , and the longitudes of  $5^\circ\text{E}$  and  $20^\circ\text{E}$  (Figure 1). Each node of the grid is the center of a circle. To cover the whole area, the radius  $R$  of the circle was set at about  $D/\sqrt{2}$ , where  $D$  is the distance between the nodes. For such a grid,  $D = 100$  km and  $R = 70$  km. We have considered all of the circles where at least one  $M_w$  5.5+ earthquake occurred over these last four centuries. For each area, we have calculated the IETs among the earthquakes inside the circle, and one CT relative to the time elapsed since the most recent event; the areas that have experienced only one earthquake have only the CT. The

vector of covariate  $z$  is a two-dimensional vector that comprises the logarithm of the rate of occurrence and the magnitude of the event, from which we calculate the IET and the CT.

The regionalization of Italy applied by Cinti et al. [2004] (Figure 2) has been used to account for the heterogeneities both in the tectonic domains and in the spatial distributions of the earthquakes, using the most up-to-date, complete and reviewed catalog of active stress indicators in Italy, such as borehole breakout data, earthquake fault-plane solutions, and seismogenic fault data. The description of this zonation is available in Cinti et al. [2004], and it includes 61 zones with different shapes and dimensions. Only 34 of these zones experienced at least one  $M_w$  5.5+ earthquake in these last four centuries. For each IET and CT, we attach a vector  $z$  with the seven components of: the logarithm of the rate of occurrence; the magnitude of the earthquake from which we calculated the IET and CT; the predominant tectonic regime; the known number of seismogenic faults; the homogeneity of the stress orientation; the homogeneity of the topography; and the extent of the area. The definition of these seven factors is a consequence of the definition of the seismotectonic zonation, and this is fully described in Cinti et al. [2004].

The data in these other two studies showed that only the component of the weight vector  $\beta$  associated to the rate of occurrence is significantly different from zero. In practice, this means that the rate of occurrence appears to be the only important covariate (of those considered) in modeling the spatiotemporal distribution of moderate to large earthquakes. Surprisingly, the magnitude is not



**Figure 2.** The seismotectonic zonation from Cinti et al. [2004], figure 1d. A total of 61 zones defines the hazard regionalization. The tectonic regimes are shown for each zone.

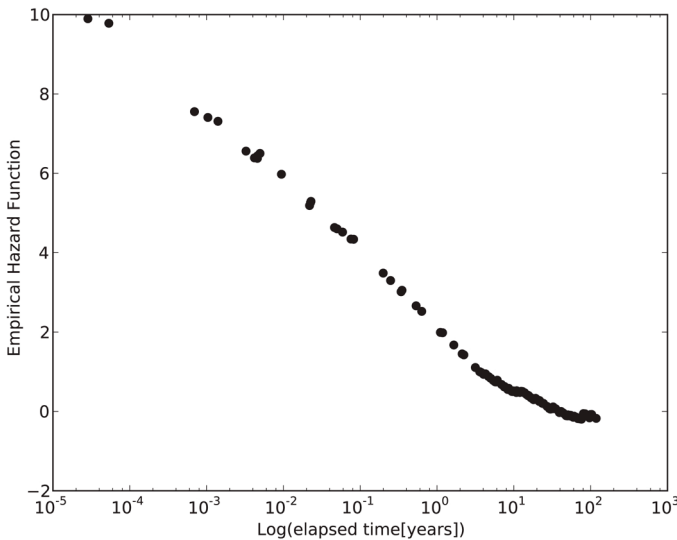
relevant in the determination of the time of occurrence of future events. Therefore, in the application of the PHM to the CSEP experiment, we used only one covariate, the logarithm of the rate of occurrence in the area. As previously mentioned, we applied the PHM to different datasets, from a local to a global scale. In all of these applications, we obtained similar results: the only relevant covariate was the rate of occurrence in the area, and the time-clustering of the seismicity. Moreover, in Faenza et al. [2008] we studied in detail the relevance of the tectonic regime in the earthquake occurrence for large earthquakes worldwide, and we found no statistical differences between different tectonic regimes with respect to the hazard functions.

These previous studies also showed an increasing hazard function after the occurrence of an earthquake (see Figure 3: the plot of the hazard function using the geometrical grid and the updated catalog, as described above). This behavior implies that the probability increases immediately after a target event in the area, and it decreases until it reaches an approximately constant value, as expected in a Poisson process. In practice, this means that the seismic clustering is a prominent aspect of the time distribution of earthquakes, at least over the first few years following a moderate to large earthquake; afterwards, the process becomes almost time independent. These studies provided the probability map of the occurrence of at least one  $M_w$  5.5+ earthquake in the next 10 years. The starting point was the empirical survivor function. Figure 4 shows the survivor function for the Italian seismicity.

The conditional probability of an earthquake in the next  $\Delta\tau$  years, given the time  $t$  since the last event, can be approximated by:

$$P(t, \Delta\tau; z) = \frac{S(t; z) - S(t + \Delta\tau; z)}{S(t; z)} \quad (6)$$

where  $t$  is the CT, and  $S(t, z)$  is the survivor function estimated



**Figure 3.** Plot of the residuals of the model as a function of the time elapsed since the most recent event (CT). The residuals mimic the time behavior of the  $\lambda_0(t)$ .

from the data, i.e., the probability that the next earthquake will occur at a time longer than  $t$  for a vector of covariates  $z$ . Equation (6) is also used for the calculation of the maps on the webpage.

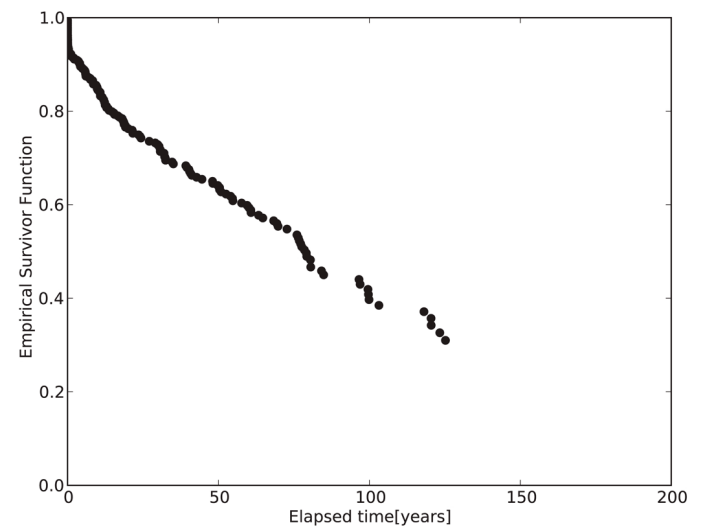
### 3. Adaptation of the Proportional Hazard Model for the CSEP Testing Center

The adaptation of the PHM to the CSEP experiment required some major modifications. Our starting point, as also shown on the webpage (<http://earthquake.bo.ingv.it>), was a spatial geometrical grid or a seismotectonic zonation (see above) for the probability of earthquake occurrence, with the catalog ending on December 31, 2009. The CSEP experiment is based on a  $0.1^\circ \times 0.1^\circ$  geometrical grid (see Figure 1). For each cell, the expected number of events on the surface is required, for the magnitude range  $M_w$  5.0 to 9.0, binned every 0.1 point of magnitude, in the forecasting time window  $\Delta\tau$ , i.e, 5 years and 10 years.

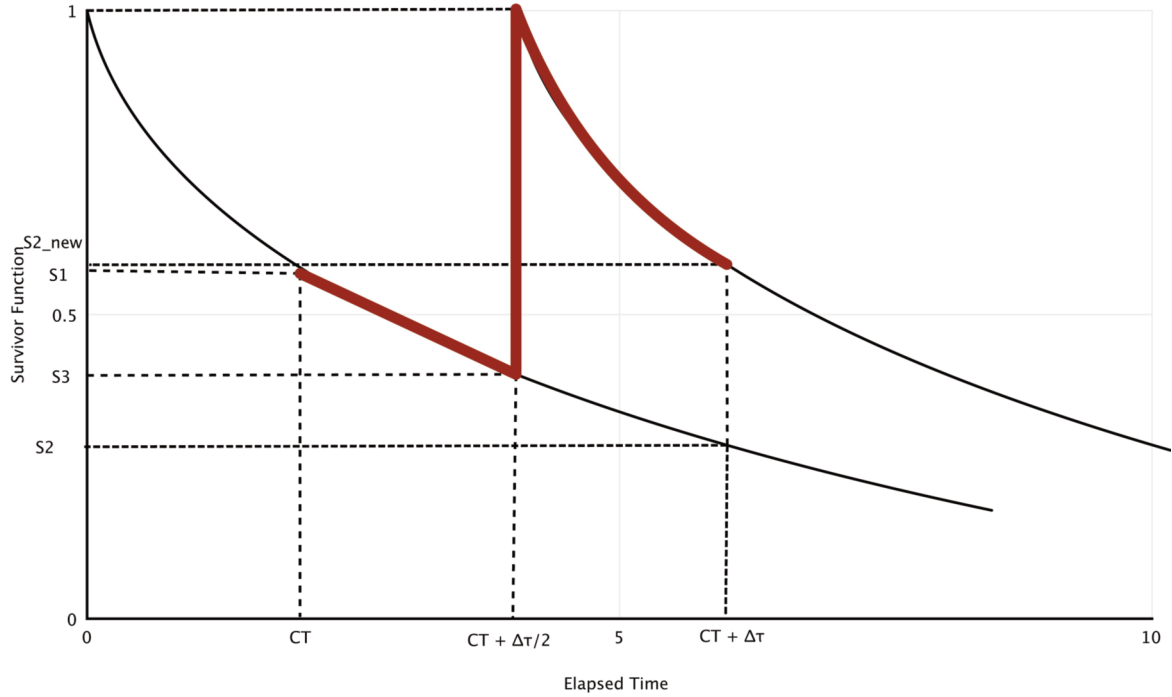
In the following, the points needed to convert our input into a suitable input for CSEP are listed:

- *Probability of occurrence for the forecasting time window  $\Delta\tau$*

While Equation (6) is suitable for our testing webpage (since we update the map every January 1 and after each  $M_w$  5.5 event, i.e., modifying the CT), it is expected to underestimate the number of events in the CSEP test. Indeed, the hazard function of Figure 3 leads to the definition of a time-dependent probability: the probability changes with the time elapsed since the most recent event. Thus, Equation (6) will not include the increased probability due to the possible occurrence of one or more earthquakes in  $\Delta\tau$ . To incorporate the events due to the clustering, we modified the model by simulating the effects on the probability of the occurrence of earthquakes in  $\Delta\tau$ . Specifically, we modified Equation (6)



**Figure 4.** Empirical baseline survival function (Equation 3) for the Italian seismicity.



**Figure 5.** Schematization of Equation (7). The red line shows the probability that accounts for the effects of the seismicity that occurred before  $\Delta\tau$  and its increase due to the occurrence of an event at  $CT + \Delta\tau/2$ .

to include the change in the survivor function due to the occurrence of an event during the forecasting time window  $\Delta\tau$ , such that:

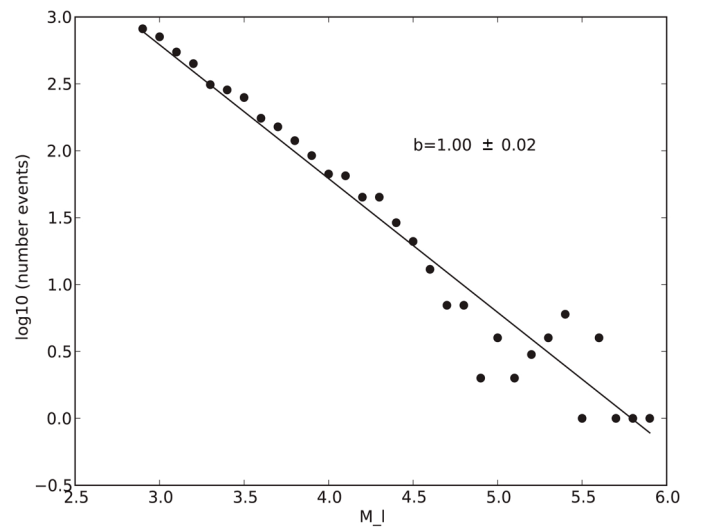
$$P(t, \Delta\tau; z) = \left[ \left( \frac{S_1 - S_2}{S_1} \right) (1 - P(1)) + \left( \frac{S_1 - S_3}{S_1} + 1 - S_2^{new} \right) P(1) \right] \quad (7)$$

where  $P(1)$  is the probability from Equation (6),  $S_1$  is the survivor function at the start time of the test, the  $CT$ ;  $S_2$  is the survivor function at the end time of the test, the  $CT + \Delta\tau$ , if no earthquake occurred in  $\Delta\tau$ . Assuming that one event occurs at  $CT + \Delta\tau/2$ ,  $S_3$  is the survivor function immediately before  $CT + \Delta\tau/2$ , and  $S_2^{new}$  is the survivor function at time  $CT + \Delta\tau$ . Figure 5 shows the representative explanation of Equation (7). We conditioned the probability in Equation (7) to the probability that at least one earthquake occurred in  $\Delta\tau$ , i.e.,  $P(1)$ . We arbitrarily chose the occurrence of the events at half of the forecasting time window, i.e.,  $\Delta\tau/2$ .

The first addendum in the right part of Equation (7) is the probability that takes into account the effects of the seismicity that occurred before  $CT$ . The second addendum mimics the increase in probability due to the occurrence of one event at  $CT + \Delta\tau/2$ . In this way, we account for the possibility of having exactly one event inside  $\Delta\tau$ . We can repeat the same calculations to take into account a second generation of earthquakes that occurs in the middle of the interval  $[CT + \Delta\tau/2, CT + \Delta\tau]$ . If we refer to these additional earthquakes as the «offspring» and the earthquakes that generated them (the previous ones) as the «fathers», we have seen through simulations that the number of offspring at

each step is about 20% of the number of fathers. This means that 100 expected earthquakes becomes about 120 when taking into account for the first generation of offspring, and 124 after the second generation of offspring. We have stopped at the second generation of offspring, assuming that the offspring of the third generation are negligible. In any case, we are aware that the CSEP testing classes downscale the actual potential forecasting capability of PHM, and that the most correct way to solve this point is to have the possibility of updating the model after an earthquake occurs.

- Conversion of moment magnitude  $M_w$  into local magnitude  $M_l$
- On our webpages, we use the moment magnitude  $M_w$ ,



**Figure 6.** The magnitude versus frequency relationship of the  $M_w$  events since 1981 (data from the CSI catalog). The value of  $b$  of the GR relationship is plotted from a maximum likelihood estimation.

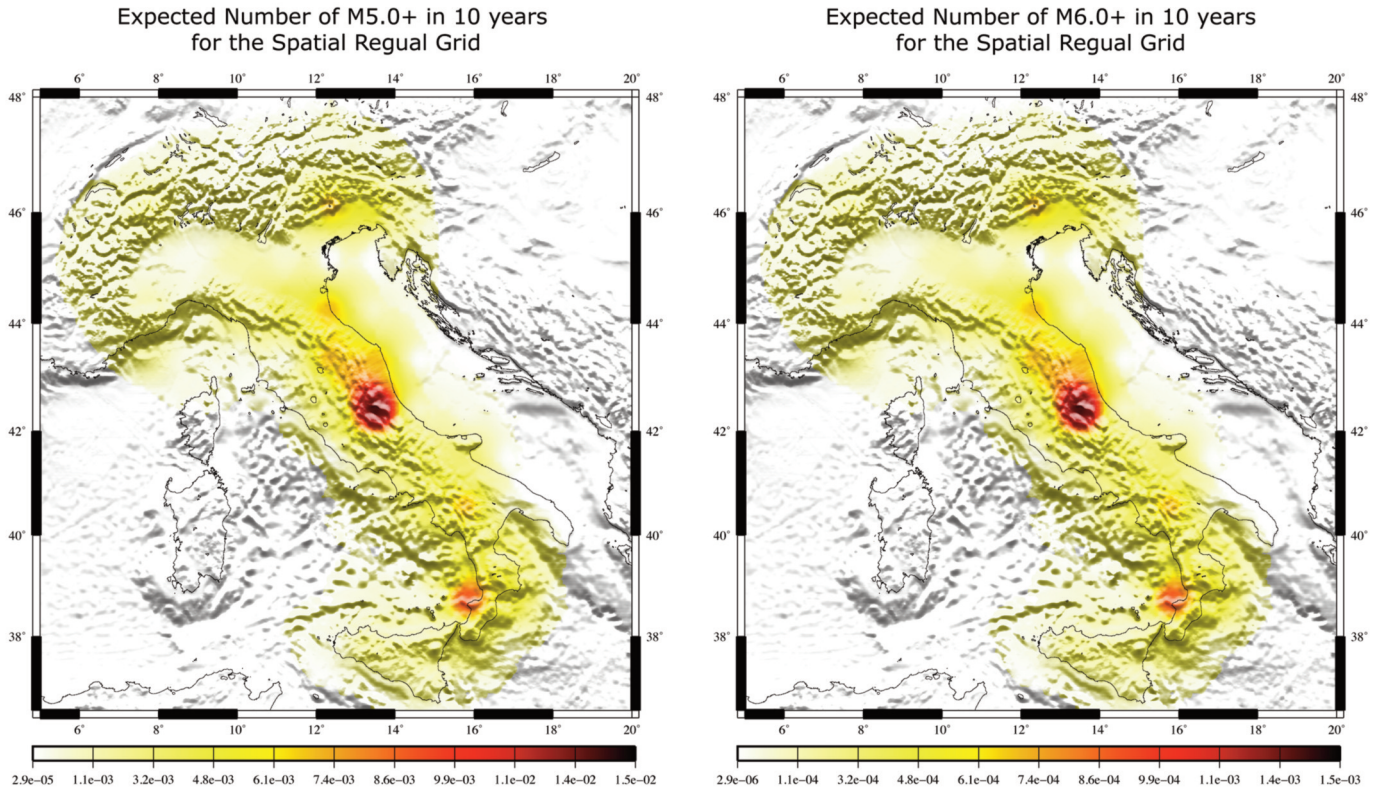


Figure 7. The number of expected events in the surface units of the CSEP testing grid for the spatial geometrical grid: left,  $M_1 5.0+$ ; right,  $M_1 6.0+$ .

while the CSEP testing is adopting the  $M_1$  from the INGV seismic bulletin. We adopted the relationship of Castello et al. [2007]. This leads to  $M_w 5.5 = M_1 5.4$ ; the geometrical grid table and the seismotectonic zonation represent the probability of earthquake occurrence with  $M_1 5.4$  in the next  $\Delta\tau$  years.

• *From probability to number of expected earthquakes*

The link between the probability of earthquake occurrence and number of expected events is carried out in three consecutive steps. Firstly, we use the formula:

$$P = 1 - \exp(-\mu t) \quad (8)$$

where  $P$  is the probability, and  $\mu t$  is the number of expected events in time  $t$  with magnitude  $M_1 5.4+$ ,  $N_{M_1 5.4+}$ .

Secondly, to extrapolate the number of events with  $M_1 5.0+$ , i.e.  $N_{M_1 5.0+}$ , the Gutenberg-Richter relation is applied, with the  $b$  parameter equal to  $1.00 \pm 0.02$  (see Figure 6) estimated from the CSI catalog [Chiarabba et al. 2005, Castello et al. 2007], using the maximum likelihood estimation [Marzocchi and Sandri 2003, Sandri and Marzocchi 2007]:

$$N_{M_1 5.0+} = N_{M_1 5.4+} 10^{b\Delta m} \quad (9)$$

where  $\Delta m = 5.4 - 5.0$ .

Finally, the expected number of events in each magnitude unit of 0.1 is obtained again from the Gutenberg-Richter relation, with the formula  $N_{M_1+\delta m} = N_{M_1} 10^{-b\delta m}$ , with  $\delta m = 0.1$  and  $M_1$  from 5.0 to 9.0.

The strategy adopted for the spatial distribution is

different for the regular grid and for the seismotectonic zonation.

### 3.1. The regular spatial grid

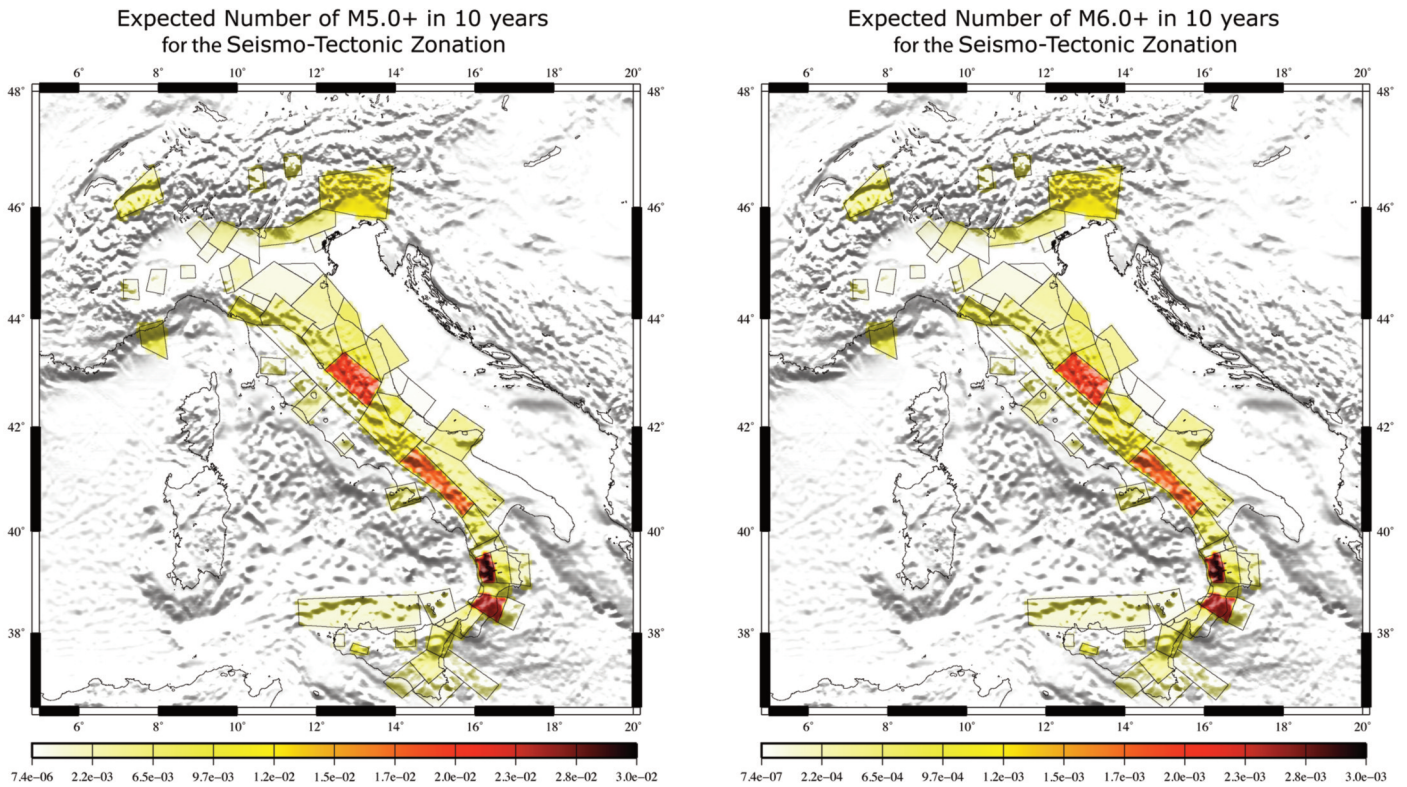
The input grid is a table that is 169 rows long and 3 columns wide, as the latitude, the longitude of the node of the circle, and the expected number of events per unit of surface ( $\text{km}^2$ ). To spread this Table on the  $0.1^\circ \times 0.1^\circ$  forecasting grid, we adopted the following strategies:

- the nodes of the grid with the number of events equal to 0 (i.e., a circle in which no  $M_w 5.5+$  has occurred over the last four centuries) have arbitrary values of an order of magnitude smaller than the smallest number estimated;
- for each node of the CSEP forecasting grid, the number of earthquakes per unit of surface is calculated as the weighted sum, with squared distance, of the values of the four closest nodes of the grid;
- then these numbers are multiplied by the dimension of the  $0.1^\circ \times 0.1^\circ$  cell.

### 3.2. The seismotectonic zonation

The adaption of the spatial distribution of the zonation for the CSEP testing grid was carried out as follows:

- the values are normalized for each area, to have values per unit of surface ( $\text{km}^2$ );
- the 27 zones with the number of events equal to 0 have an arbitrary value of an order of magnitude smaller than the smallest number estimated;
- the nodes of the CSEP testing grid that are inside the



**Figure 8.** The number of expected events in the surface units of the CSEP testing grid for the seismotectonic zonation: left,  $M_1$  5.0+; right,  $M_1$  6.0+. The zonation has been drawn as a proxy.

zonation take the value of the respective zone;

- the nodes of the CSEP testing grid that are outside the zonation take an arbitrary value of two orders of magnitude smaller than the smallest number estimated;
- then these numbers are multiplied by the dimension of the  $0.1^\circ \times 0.1^\circ$  cell.

#### 4. Conclusions

This study presents the rearrangements that we have carried out to make the PHM model [Faenza et al. 2003, Cinti et al. 2004] fit the CSEP Testing Center requirements in the forecasting time windows of 5 years and 10 years. Although the optimal use of such a model is different from that for the CSEP requirements (i.e., the model should be updated every year and immediately after a target event), we recognize the paramount importance of these experiments for a full and unbiased verification of an earthquake forecasting model. In Figures 7 and 8, we show

four forecasting maps, two for the regular spatial grid, and two for the seismotectonic zonation. The forecasting time window is 10 years, and the maps show the expected number of events with  $M_1$  5.0+ and  $M_1$  6.0+. From Figures 7 and 8, we can see that the spatial model affects the results strongly, since in the regular spatial grid the maximum number of events is expected in the central Apennines, while the seismotectonic zonation produces a maximum of expected number of events in the Calabrian Arch. With this prospective, the submission of the PHM to the CSEP Testing Center is a good opportunity to appraise the effectiveness of the spatiotemporal PHM model. Table 1 also gives the overall number of earthquakes that are expected for the two testing classes (5 years and 10 years) and for the two zonations. The seismotectonic zonation predicts a smaller number of events compared to the geometrical zone for both of the two classes (e.i., 12.34 versus 13.01 for the zonation and the geometrical grid, respectively, for the 10-year class).

	Geometrical grid		Zonation	
	10 years	5 years	10 years	5 years
M 5.0+	13.01	6.48	12.34	6.16
M 6.0+	1.30	0.65	1.23	0.62

**Table 1.** Expected number of earthquakes for the 5-year and 10-year forecasting classes, over the whole territory under investigation, for the two spatial zones.

**Acknowledgements.** Some of the figures were constructed using the free software of Generic Mapping Tools [Wessel and Smith 1998].

#### References

Castello, B., M. Olivieri and G. Selvaggi (2007). Local and duration magnitude determination for the Italian earthquake catalog, 1981-2002, *Bull. Seismol. Soc. Am.*, 97, 128-139; doi:10.1785/0120050258.  
 Chiarabba, C., L. Jovane and R. Di Stefano (2005). A new

- view of Italian seismicity using 20 years of instrumental recordings, *Tectonophysics*, 395, 251-268.
- Cinti, F.R., L. Faenza, W. Marzocchi and P. Montone (2004). Probability map of the next M5.5 earthquakes in Italy, *Geochem. Geophys. Geosyst.*, 5, Q11003; doi: 10.1029/2004GC000724.
- Cox, D.R. (1972). Regression models and life tables (with discussion), *J. R. Stat. Soc., Series B.*, 34 (2), 187-220.
- 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/>.
- 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 catalogue, *Geophys. J. Int.*, 155, 521-531.
- Faenza, L., W. Marzocchi, A.M. Lombardi and R. Console (2004). Some insights into the time clustering of large earthquakes in Italy, *Annals of Geophysics*, 47, 5, 1635-1640.
- Faenza, L. (2005). Analysis of the spatio-temporal distribution of large earthquakes, Ph. D. Thesis, Università degli Studi di Bologna, Alma Mater Studiorum, Bologna, Italy (available at: <http://hdl.handle.net/2122/>).
- Faenza, L. and S. Pierdominici (2007). Statistical occurrence analysis and spatio-temporal distribution of earthquakes in the Apennines (Italy), *Tectonophysics*, 439, 13-31.
- Faenza, L., W. Marzocchi, P. Serretti and E. Boschi (2008). On the spatiotemporal distribution of M70.+ worldwide seismicity, *Tectonophysics*, 449, 97-104.
- Faenza, L., S. Hainzl and F. Scherbaum (2009). Statistical analysis of the central-Europe seismicity, *Tectonophysics*, 470, 195-204.
- Kalbeisch, J.D. and R.L. Prentice (1980). *The Statistical Analysis of Failure Time Data*, NewYork, 336 pp.
- Marzocchi, W. and L. Sandri (2003). A review and new insight on the estimation of the *b*-value and its uncertainty, *Annals of Geophysics*, 46, 6, 1271-1282.
- Pondrelli, S., S. Salimbeni, A. Morelli, G. Ekström, M. Olivieri and E. Boschi (2010). Seismic moment tensors of the April 2009, L'Aquila (central Italy), earthquake sequence, *Geophys. J. Intl.*, 180, 238-242.
- Sandri, L. and W. Marzocchi (2007). A technical note on the bias in the estimation of the *b*-value and its uncertainty through the least squares technique, *Annals of Geophysics*, 50 (3), 329-339.
- Schorlemmer, D., A. Christophersen, A. Rovida, F. Mele, M. Stucchi and W. Marzocchi (2010). Setting up an earthquake forecast experiment in Italy, *Annals of Geophysics*, 53, 3 (present issue).
- Wessel, P. and W. H. F. Smith (1998). New, improved version of the Generic Mapping Tools released, *Eos. Trans. AGU*, 79, 579.

---

\*Corresponding author: Licia Faenza,  
Istituto Nazionale di Geofisica e Vulcanologia, Centro Nazionale  
Terremoti, Roma, Italy; e-mail: [licia.faenza@ingv.it](mailto:licia.faenza@ingv.it)

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