Bayesian inference on earthquake size distribution: a case study in Italy

Licia Faenza[1], Carlo Meletti[2] and Laura Sandri[3]

[1] Istituto Nazionale di Geofisica e Vulcanologia, CNT; via di Vigna Murata 605, Roma, Italy

[2] Istituto Nazionale di Geofisica e Vulcanologia, sezione di Milano-Pavia, via Bassini 15, Milano, Italy

[3] Istituto Nazionale di Geofisica e Vulcanologia, sezione di Bologna, via Donato Creti 12, Bologna, Italy.

Accepted for publication at BSSA on 29th September 2009.

Abbreviated title: Bayesian Inference in Magnitude distribution

Corresponding author: Licia Faenza

Istituto Nazionale di Geofisica e Vulcanologia

Via di Vigna Murata 605

00143 Rome

Italy

Tel: +39 06 51860608

e-mail: licia.faenza@ingv.it

Abstract

This paper is focused on the study of earthquake size statistical distribution by using Bayesian inference. The strategy consists in the definition of an a priori distribution based on instrumental seismicity, and modeled as a power law distribution. By using the observed historical data, the power law is then modified in order to obtain the posterior distribution. The aim of this paper is to define the earthquake size distribution using all the seismic database available (i.e., instrumental and historical catalogs) and a robust statistical technique. We apply this methodology to the Italian seismicity, dividing the territory in source zones as done for the seismic hazard assessment, taken here as a reference model. The results suggest that each area has its own peculiar trend: while the power law is able to capture the mean aspect of the earthquake size distribution, the posterior emphasizes different slopes in different areas. Our results are in general agreement with the ones used in the seismic hazard assessment in Italy. However, there are areas in which a flattening in the curve is shown, meaning a significant departure from the power law behavior and implying that there are some local aspects that a power law distribution is not able to capture.

Online material Results of the analysis applying the statistical completeness.

Keywords: Earthquake size distribution, Bayesian inference, Italian seismicity, Seismic hazard

1 Introduction

In probabilistic seismic hazard analysis (see e.g., SSHAC, 1997), the study of the past seismicity is a fundamental element to model the occurrence of future ground shaking. Under the assumption of a stationary seismogenic process, in a standard approach (Cornell, 1968; McGuire, 1976) the seismic hazard is evaluated by studying seismicity in terms of time and magnitude recurrence laws from a probabilistic point of view, and by defining ground-motion attenuation relationships, after the identification of active seismic zones.

In Italy, seismic hazard analysis is usually carried out on historical seismicity (CPTI04, CPTI Working Group, 2004), like it has been done for the official seismic hazard map of Italy (MPS04, MPS Working Group, 2004), a work that represents the state-of-the-art in Italy for what concerns the seismic hazard. As each seismic zone has its own frequency of earthquake occurrence, in low frequency seismic zones this reflects a very small number of available historical events to be used for hazard model calibration, affecting the reliability of the model parameters. In this study, we would like to enhance the data set used for hazard model calibration regarding the earthquake size distribution, by properly taking into account instrumental seismicity, even if it contains a large number of small events, not suitable for classical seismic hazard models.

Bayesian inference is the process of fitting a probability model to a set of observations and summarizing the results with a probability distribution for the parameters of the model. Its applications span over several scientific and economical disciplines, showing some interesting aspects. First, according to Bayesian philosophy, Bayesian inference allows to simultaneously take into account, giving proper weights, heterogeneous sources of information on the process under study,

such as theoretical models, prior beliefs, and observations. In this view, Bayesian inference can provide an estimate of the quantity of interest even in the case of few observations available, if there are prior models. Secondly, Bayesian inference allows the estimation of the uncertainty on the results obtained. In dealing with a complex process (i.e., governed by several degrees of freedom of comparable weight) such as the seismogenic one, the treatment of aleatory and epistemic uncertainties is of primary importance. In particular, the aleatory uncertainty is associated to the intrinsic stochasticity of the process, resulting in an unavoidable impossibility of predicting deterministically its evolution. The epistemic uncertainty represents our limited knowledge of the system. An overview of this point can be found in Woo (1999), Field *et al.* (2003) and MacKay (2003). While it is possible to reduce this latter type of uncertainty (e.g., by increasing the number of data points or improving the physical knowledge of the phenomenon), the aleatory one is independent on our degree of knowledge and cannot be lowered.

Because of these features, Bayesian inference can be an interesting instrument to study the probabilistic law governing earthquake size. For example, Agostinetti and Rotondi (2003) by using Bayesian belief network investigated the dependence relationship between the size of the earthquakes in a sequence and their time of occurrence. Our study instead makes Bayesian inference on the parameters of the size distribution of events able to generate damage in Italy, i.e., earthquakes with magnitude $M_w \geq 4.65$ (MPS04). In detail, we use instrumental seismicity to calibrate a diffuse (i.e., with a large epistemic uncertainty) prior model, and we refine it into the posterior model by means of historical catalog. In this way we can provide more accurate estimates for model parameters.

In principle, the results obtained can be incorporated in a probabilistic seismic

hazard study for Italian damaging seismicity.

2 Data

The input data for this analysis are the seismic zonation ZS9 (Meletti *et al.*, 2008), the seismic catalog CPTI04 (CPTI Working Group, 2004) and the CSI catalog (Catalogo Strumentale Italiano, Chiarabba *et al.*, 2005; Castello *et al.*, 2007).

2.1 The seismic zones model ZS9

ZS9 (Figure 1) is the product of consensus from an expert team of researchers and it is an ingredient of the seismic hazard map for Italy (MPS04). The model was designed for the application in country-wide probabilistic seismic hazard assessment and it is largely based on data collected in the last 10 years, including historical earthquakes and instrumental seismicity, active faults and their seismogenic potential, and seismotectonic evidence from recent earthquakes. ZS9 is made out of 36 zones where earthquakes with $M_w \geq 5$ are expected and every zone is characterized also by its mean seismogenic depth and predominant focal mechanism. For a detailed description of the requirements the zonation model has to satisfy and for the input data used to build ZS9 up, we refer to Meletti *et al.* (2008).

2.2 The Parametric Catalog of Italian Earthquakes

The Parametric Catalog of Italian Earthquakes (CPTI04) is the most updated and complete catalog for the historical events above damage threshold in Italy. It contains events from 217 b.C. to 2002. Like ZS9, CPTI04 is one of the input element of the MPS04 seismic hazard map of Italy. During the MPS04 elaboration

the content of CPTI04 catalog was deeply analyzed in order to define the time-interval completeness and seismicity rates. The M_w magnitude ranges between 4.65 and 7.41; according to the MPS04, the magnitude is binned into 12 classes with increment equal to 0.23. The time-interval completeness of the CPTI04 catalog was estimated for each zone of ZS9 and for each class of magnitude according two different approaches: one approach is mainly historical, based on the analysis of the archive completeness; the second one is mainly statistical (for details, see MPS04). Tables 1 and 2 summarize the temporal windows of completeness and the number of events used in the analysis for all zones and magnitude classes, for the historical and statistical completeness, respectively.

2.3 The instrumental catalog

The CSI catalog (Chiarabba *et al.*, 2005; Castello *et al.*, 2007) is a collection of revised hypocentral locations of the earthquakes occurred in the Italian territory. Maintained by the INGV, it contains data collected by a network of seismic stations, active since 1981, spanning the time period 1981-2002. We check the completeness of the catalog with the cumulative of the events, finding that it is complete for $M_l \geq 2.9$ for events with shallow and intermediate hypocenters (depth $\leq$ 70km, to agree with CPTI04). We shall see in section 3.2.1 the reason for using this data set.

3 The method of Bayesian inference

3.1 General description

Different from the frequentist approach, in a Bayesian perspective there is not a "true" value for the variables of interest, but they are represented by a probability density function (pdf). In this view, Bayesian inference is the instrument used to best estimate the parameters of that pdf. Practically, this is accomplished by merging a prior model for the variables (constituting the so-called prior distribution, $[\theta]_{prior}$, where θ is the vector of the variables, and the square brackets denote a pdf) with a set of past observations (y), in order to maximize the probability of having observed those data with that prior model in mind. The output of Bayesian inference is thus a posterior distribution ($[\theta|y]$) for the variables of the model. The practical way used for combining prior distribution and data is the Bayes' theorem:

$$[\theta|y] = \frac{[\theta]_{prior}[y|\theta]}{[y]} \tag{1}$$

where $[y|\theta]$ is the so-called likelihood function (representing the probability of observing data y given parameters θ), and $[y]$ is a normalizing factor accounting for the total probability of observing the data y. With Bayesian inference the uncertainty on θ is also provided. In particular, while the mean of the posterior pdf represents an estimation of the intrinsic randomness of the process (aleatory uncertainty), its variance is an estimation of the epistemic uncertainty, due to our limited knowledge of the process. As mentioned above, the latter can be reduced by increasing the amount of observations available, if the prior model does not rule out the "true" model, (i.e., $\theta_{prior}(truemodel) = 0$). To avoid this extreme case, we will choose a prior model that does not rule out any possibility.

3.2 Our application

In this study, the random variable of our system θ represent the probability of a given earthquake size in the different zones in Italy, conditional to the hypothesis that a damaging event will occur in future. In particular,

$$\theta = (\theta_1, ...\theta_{12}) \tag{2}$$

where the 12 elements of vector θ represent the probabilities of 12 different possible classes of earthquake sizes, as binned by MPS04; the 12 classes derive from the consideration that magnitudes for the historical earthquakes come from a rough conversion of the epicentral intensity. The Italian catalogs normally adopt the MCS intensity scale, based on 12 degrees. For this reason, the range of magnitudes in CPTI04 catalog was divided in order to have in each class earthquakes with the same epicentral intensity (M_w class 1 corresponds to the intensity value of 5.5, M_w class 2 corresponds to the intensity value of 6, and so on). Since the basic hypothesis is that a damaging event will occur and the first class (with probability θ_1) represents the lowest class of magnitudes considered able to produce damage (i.e., $M_w \geq 4.76 \pm 0.115$), the 12 binned classes of magnitude are a complete set of mutually exclusive events, i.e.

$$\sum_{i=1}^{12} \theta_i = 1. \tag{3}$$

In our application, the available past observations are counts of past earthquake magnitudes in each of the 12 magnitude bins. Since each earthquake has a magnitude that is independent from the others, our past observations can be thought as independent trials producing one of 12 possible mutually exclusive outcomes, each characterized by a probability θ_j ($j = 1, ...12$). For this reason, we express the

likelihood of the available past observations through a multinomial distribution:

$$[y|\theta] = \begin{pmatrix} n \\ y_1 ... y_m \end{pmatrix} (\theta_1)^{y_1} ... (\theta_m)^{y_m} \tag{4}$$

where y_j $(j = 1, ...12)$ is the observed number of past earthquakes with magnitude falling in the j-th size class, and $n = \sum_{j=1}^{m} y_j$ is the total number of past data available.

In Bayesian inference, the choice of the shape of the prior pdf is subjective. In this application, similarly to Marzocchi *et al.* (2004, 2008), we choose a Dirichlet distribution for our prior model. Besides being the conjugate distribution of the multinomial, implying a simplification in the evaluation of the posterior distribution (which, for this choice, remains a Dirichlet distribution, see e.g. Gelman *et al.*, 2004), we would also like to highlight three important reasons for choosing a Dirichlet distribution. Firstly, we would like a unimodal prior for $[\theta]$; secondly, since our random variable $[\theta]$ is a probability, we need a pdf with domain [0,1]; lastly, the sample of the magnitude classes has to be a mutually exclusive and exhaustive sample, i.e. the intersection has to be null and the union equals 1.

In our specific application, the prior Dirichlet distribution reads:

$$[\theta] = Di(\alpha_1, ...\alpha_{12}) = \frac{\Gamma(\alpha_1 + \cdots + \alpha_{12})}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_{12})} [\theta_1]^{(\alpha_1 - 1)} \cdots [\theta_{12}]^{(\alpha_{12} - 1)} \tag{5}$$

where $\alpha_j (j = 1, \cdots, 12)$ are the parameters of the distribution and are greater then 0, and $\Gamma(\cdot)$ is the Gamma function.

For the j-th event, the first moment (i.e., the mean) is

$$E[\theta_j] = \frac{\alpha_j}{\alpha} \tag{6}$$

where $\alpha = \sum_{j=1}^{m} \alpha_j$, while the second moment (i.e., the variance)

$$Var[\theta_j] = \frac{\alpha_j(\alpha - \alpha_j)}{\alpha^2(\alpha + 1)} \tag{7}$$

(see Gelman *et al.*, 2004). In our application, $E[\theta_j]$ represents our initial best guess, for the probability of occurrence of an event in magnitude class j given the occurrence of a damage event, while $Var[\theta_j]$ represents our degree of confidence on the best guess for magnitude class j. A noteworthy case occurs when $\alpha = m$. In this case, the variance is comparable to that of the uniform distribution (often taken as the maximum ignorance distribution), and only one past observation is potentially able to modify the best guess of such a large-variance prior model. In such case, the epistemic uncertainty is very high (i.e., our knowledge is very scarce).

The marginal distribution for a specific parameter θ_j of a Dirichlet distribution is a Beta distribution

$$[\theta_j] = Beta(\alpha_j; \beta_j) = \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)} \theta_j^{\alpha_j - 1} (1 - \theta_j)^{\beta_j - 1} \tag{8}$$

where $\beta_j = \alpha - \alpha_j$.

As mentioned above, a Dirichlet distribution for the prior model conjugates with a multinomial distribution for past data. This means that the posterior density is still a Dirichlet, with updated posterior parameters:

$$[\theta|y] = Di(\alpha_1 + y_1, ...\alpha_m + y_m) \tag{9}$$

3.2.1 Prior distribution

For the definition of the prior distribution, we start by considering that the frequency-magnitude relationship for earthquake occurrence follows a power law distribution, known as the Gutenberg-Richter (G-R) relation (Gutenberg and Richter, 1954). We decide that the prior distribution is the same for all the 36 zones that compose the ZS9 seismic zonation and for both the two completeness

evaluations in tables 1 and 2. The calibration of the distribution is done using the events from the CSI catalog (Chiarabba *et al.*, 2005; Castello *et al.*, 2007).

This data set contains many small events and very few events of damaging sizes (however $M_w < 6$). In this view, it is not a very pertinent data set for our calibrating model, since it contains many events with a smaller magnitude than that considered in seismic hazard analysis. However, since it contains many records, it allows us to define a diffuse prior distribution even in the seismic zones with few historical events. We then used the CPTI04 catalog to estimate the likelihood; this is a more "accurate" data set for our purpose because it contains many damaging events. A problem we want to underline concerns the de-clustering of the catalogs. CPTI04 has been designed for the probabilistic seismic hazard analysis, using the approach of Cornell (McGuire,1976; Cornell, 1968), i.e. with a random temporal earthquake occurrence. Therefore, CPTI04 is composed only by independent events, while CSI catalog is not de-clustered. It is still debatable which dependent events and mainshocks have the same frequency-magnitude distribution. In our application, we assumed that the frequency-magnitude distribution of the whole seismicity (i.e. independent and dependent events) is the same as the one of independent events alone. We then convert M_l in M_w using the empirical relation in MPS04; the b-value of the frequency-magnitude relationship is calculated with the maximum likelihood method (Marzocchi and Sandri, 2003; Sandri and Marzocchi, 2007) and gives $b = 1.17 \pm 0.01$ (figure 2). Assuming universality of the G-R, we can extrapolate this b-value for greater events.

The relative probability of each binned class of magnitude is calculated by using the G-R law with the estimated b-value. Then these frequencies are normalized to 1. The meaning of this normalization is to impose that a damaging event (i.e.,

with $M_w > 4.65$) will occur in the future within the area covered by the seismic zonation. We therefore have 12 frequencies, f_i ($i = 1, \cdots 12$), with $\sum_{i=1}^{12} f_i = 1$. The parameters of the prior Dirichlet distribution (α_j, $j = 1, ...12$) are then found by imposing that:

1. for each of the 12 binned classes of magnitude, our best guess value is equal to the frequency given from the power law distribution, i.e., $E[\theta_j] = f_j$, for $j = 1, ...12$;

2. our degree of confidence on the best guess is the lowest possible, i.e., the prior distribution has the maximum allowed variance: $\alpha = 12$

By inverting equations 6 and 7 with these conditions, we obtain the 12 values (α_1,...α_{12}), provided in table 3.

We realize that the resulting prior has a strong assumption, therefore it has lowest degree of confidence. Other possible priors can be built fitting for each zone a different b parameter using the CSI catalog or the CPTI04 catalog. In both cases, the b parameter is fitted using a small data set for each zone and completeness; moreover, in the latter hypothesis the prior and the likelihood are based on the same data set. A point we want to stress is that the aim of this application is to build a final result, i.e. the posterior, which is accurate rather than precise.

3.2.2 Likelihood function

We use both historical and statistical completeness estimates to get two independent results. The values of table 3 are the starting point for both the two analysis.

For zone l ($l = 1, \cdots, 36$) and for binned class of magnitude i ($i = 1, \cdots, 12$), within the temporal window of completeness we count the number n_{il} of events with magnitude M_w in the class and occurring in the zone according to CPTI04

catalog. Since the length of temporal window of completeness is different for each binned class of magnitude and zone, we cannot simply use these numbers n_{il} as the occurrences y_l in equation 9. Because of this, we define the duration of the temporal window of completeness for zone l and magnitude class i as t_{il} (tables 1 and 2). Therefore, in each zone l there are N_l data available, i.e., $N_l = \sum_{i=1}^{12} n_{il}$. From here it is possible to define the rate of events in class i and zone l as $r_{il} = \frac{n_{il}}{t_{il}}$. We transform the rates into (fictitious) absolute numbers of occurrences as

$$y_{il} = \text{round}(\frac{N_l r_{il}}{\sum_{k=1}^{12} r_{kl}}) \tag{10}$$

where the y_{il} are rounded to the nearest integer, since they are fictitious counts of past magnitudes in the various bins.

The computation of y_{il} is based on two assumptions:

- the normalization of the rates to 1 (i.e., the division by $\sum_{k=1}^{12} r_{kl}$) implies that the final distribution is conditional to the occurrence of an event above the damage threshold, i.e., that an events with $M_w \geq 4.65$ will certainly occur

- the multiplication by N_l is based on the hypothesis that seismicity is a stationary feature in each specific zone. This implies that the parameters of the distribution do not change with time and that the data set used with the completeness interval is a representative sample of the distribution of the size of the earthquake. We recognize that the latter is a strong assumption; however, it is the basic assumption for every standard seismic hazard study.

3.2.3 Posterior distribution

The posterior distribution is obtained by letting the a priori distribution be modified by the available data using Bayes' theorem (equation 1). In particular, for zone l ($l = 1, \cdots, 36$), by using equation 9, with data y_{il} just computed (equation 10) we obtain

$$[\theta^l|y] = Di(\alpha_1 + y_{1l}, \cdots, \alpha_{12} + y_{12l}) \tag{11}$$

4 Results

For the sake of conciseness, only the results of the historical completeness are presented. For statistical completeness we obtain similar results, which we provide as electronic supplementary material.

For example, we first present the plot of the marginal posterior distributions for 4 zones, by using the data obtained by applying historical completeness estimates. Zone ZS901 is a zone with a few events (see figure 1). Figure 3a shows its marginal posterior distributions. The abscissa represents the probability θ_i ($i = 1, \cdots, 12$) that a certain magnitude will occur, given our a priori information and the past events; the y-axis represents the posterior probability distribution of each probability θ_i ($i = 1, \cdots, 12$) class. A remarkable aspect of ZS901 is that the marginal posterior distribution of class 3 is similar to the one of class 2. This means that the relative frequency of class 3 is higher than the one expected by a simple power-law distribution.

Zone ZS905 (figure 3b) is one of the areas with the highest number of events (figure 1). Classes 1, 2, 3 have classic bell-shaped distributions, these marginals are peaked (low epistemic uncertainty), since they are well populated (26,13,12

events in the time completeness, respectively). Remarkably, the occurrence of a class 9 event ($M_w = 6.6$) is more probable than a class 7 ($M_w = 6.14$) or 8 ($M_w = 6.37$) event. A quite similar situation can be found in ZS923 (figure 3c), also a very active seismic zone. There is a collapse of the marginal distributions of class 5 ($M_w = 5.68$) and 6 ($M_w = 5.91$) in an almost single curve, while the amount of data in class 1 (51 events) makes the bell shaped distribution more peaked (decreasing the epistemic uncertainty) with respect to the other 3 zones. We display the plot for zone ZS931 since it shows a particular case with more strong events than medium ones. This is shown in figure 3d, where the marginal posterior for class 10 is higher than expected from an a priori G-R based marginal distribution.

In order to visualize the relative influence of the prior model and the past observations in the posterior estimate, we then compare the prior – based on a G-R distribution with $b = 1.17$ – with the posterior. We compare the median of the Dirichlet prior and posterior distributions, and the 10th-90th percentile, for each single zone. These comparisons give an overall and reliable representation of the distributions, providing the central value and as well the confidence boundaries. In principle it would be possible to analytically compute the percentiles if the probability density functions were analytically integrable. However, in our case this is not an easy task because of distributions are 12-variate Dirichlet distributions. Because of this, for each zone we generate 10000 synthetic draws from the Dirichlet prior and posterior distributions with 12 variates, and compute the 10-50-90 percentiles of the synthetic draws. We use the percentiles and not the mean specifically, because the estimate of the mean from a finite number of draws is less stable than the median, especially when the mean is very low (if the mean is of

the order of 10^{-n}, we need at least 10^{n+1} draws for a stable estimate). In figure 4 we plot the median of the synthetic Dirichlet draws and the shaded grey areas of the 10th-90th percentile, for both prior and posterior distribution for each single zone.

The first comment is that as more and more data are available, the epistemic uncertainty decreases, if compared to the one of the prior model, as for example ZS905, which has a narrower dark-gray area (posterior confidence interval) than light-gray one (prior confidence interval).

The second point is that for almost half of the zones, the posterior probability of the medium and/or high magnitude classes seems to be increased, if compared to the prior. In more detail, there is an increase in the probability of medium size magnitude in the areas in the Northern Apennines (e.g., Zones ZS912, ZS914 and ZS917) with an increase of probability for classes 5 (Mw5.68) and 6 (Mw5.91). In Friuli region, Central Apennines, and Calabria Belt (e.g., respectively Zones ZS905, ZS918, ZS919, ZS923, ZS927, ZS929, and ZS930) the probability is increased for high classes of magnitude, like class 8 (Mw6.37) and above. This could reflect the difference in magnitude size distribution between the Northern Apennines and the Eastern Alps or Southern Apennines. There are some other areas in which the opposite seems to occur, i.e., the posterior distribution shows lower probability value compared to the prior ones, like Zones ZS913, ZS915 and ZS920 in North-Central Apennines. In these zones, in general, magnitude classes between 1 and 4 (Mw4.76 - 5.45) have more events than median to high classes. There are areas, as well, in which the posterior does not differ from the prior, like Zones ZS904, ZS908, ZS909, ZS911, ZS922 and ZS931.

A key point is that, in most of the areas in which the posterior distribution

seems significantly different from the prior, this shows a preferred magnitude or group of magnitudes, rather than showing a different power law. In other words, the analysis seems to reflect a change in the slope of the power law, rather than a deviation from the power-law distribution. This deviation is significant since it is outside the confidence interval. This tendency lead us to compare our posterior size distribution with the one applied in MPS04, based on a power-law, which in some ways represents a reference benchmark for the earthquake size distribution in Italy. Figure 5 shows the mean values of the prior, the posterior — plus and minus one standard deviation, grey area in the picture — and the values used in MPS04, with the b-values calculated in that elaboration for each zone and time-interval completeness estimate. In particular, to agree with our initial statement to have at least one $M_w \geq 4.65$ event in each zone, we normalized the frequency of each zone to 1, by imposing $\sum_{i=1}^{12} f_i = 1$, where f_i are the frequencies estimated from b-values of the G-R distributions (MPS04). We refer to these curves as MPS04-GR. To compare our results with the MPS04-GR we are forced to compare the mean values, rather than the median, since the average is the only value available from MPS04 elaboration. The main difference between our posterior and the MPS04-GR is that in the areas in which deviations from the power-law are shown, the MPS04-GR decreases the b-value, leading to a general increase in the probability of occurrence of the large magnitudes, without capturing the distribution shown by the data, based on the occurrence of a preferred magnitude size (see Zones ZS925, ZS929, and ZS935, for instance).

To check the stability of our results we have performed the same strategy considering only 6 classes of magnitude, i.e., grouping the events of classes 1 and 2 into a new class 1, the one of classes 3 and 4 into a new class 2 and so on. The

central magnitude values of the new classes are $M_{w\_1} = 4.88; M_{w\_2} = 5.34; M_{w\_3} = 5.80; M_{w\_4} = 6.23; M_{w\_5} = 6.72; M_{w\_6} = 7.18$ with an incremental step of 0.46. The results show the same patterns outlined in figure 4, when comparing the prior and posterior distribution (see figure 6), and in figure 5, when comparing the posterior and the MPS04-GR (see figure 7). This indicates that our results are not the consequence of incorrect magnitude evaluation, but reflect the behavior the earthquake size distribution.

5 Conclusions

This work is the first attempt studying the size distribution of earthquakes using a Bayesian approach. The methodology, developed in the context of natural hazard in Marzocchi *et al.* (2004, 2008), has been applied to the Italian seismicity, by using the most recent information regarding the seismic catalog (CPTI Working Group, 2004) and the seismic source zones model (Meletti *et al.*, 2008). The aim of this work is to quantify the statistical size distribution of earthquakes in Italy. Using this methodology, we are able to quantify how much the real data move away from the a priori distribution based on a power-law G-R distribution for the size of earthquakes. By using Bayes theory, we let the past data modify the a priori pdf. Our results are then compared to the current reference size distribution, that has been used in the set up of the official seismic hazard map MPS04 (MPS Working Group, 2004) by using the G-R distribution calibrated separately for each zone.

As a general comment, the MPS04-GR is able to capture the main trend in earthquake size distributions for different areas. In fact, zones with more strong events (e.g., ZS931) have lower b-value, in absolute value, than zones with more medium events (e.g., ZS920). But MPS04-GR is not able to capture other as-

pects of earthquake size distributions, as the flattening for medium-magnitude classes 5, 6 and 7, which occurs in many zones (e.g., ZS902, ZS904, ZS910, ZS914, ZS918, ZS919, ZS923 etc). Furthermore, calculating the b-value for each zone independently, as done in MPS04, implies working with few data, increasing the uncertainties on relevant estimates.

It is important to remark that our results are based on the hypothesis of stationarity in the seismicity, which is currently adopted in many papers on statistical seismology. In any case, there is a debate on the correctness of adopting catalog that covers different time spans, because this implies that the seismicity rate is constant with time. To partially check whether this bias affects our results, we re-run our analysis considering only the classes from 5 and above, and a common time span from 1500 to 2002. The result of this test confirms our previous conclusion, apparently excluding a bias due to the stationarity hypothesis This test, however, is not exhaustive, since we can compare only events with medium-high classes of magnitude. We can not investigate the possible changing rates in the medium size seismicity, since there is not historical data in ancient times.

A second aspect we want to emphasize in this discussion is the physical reasons for the mismatch between the posterior distributions and the G-R one. It seems, from figure 4 and 5, that for some zones there are some "preferred" magnitudes (like class 6, Mw=5.91, for zone ZS926 or class 9, Mw=6.60, for ZS905). This indicates that the "characteristic earthquake" (Schwartz and Coppersmith, 1984) behavior could result in these source zones, where a magnitude class, corresponding to a main seismogenic fault, is dominant and the other classes represent the seismicity of the minor fault systems in the zone. But, of course, this feature has still to be fully explored.

6 Data and Resources

All data used in this article came from published sources listed in the references. Some plots were made using the Generic Mapping Tools version 4.2.1 (www.soest.hawaii.edu/gmt; Wessel and Smith, 1998)

Acknowledgments: The author thank V. D'Amico for the useful comments to the preliminary version on the work. We would like to thank two anonymous reviewers and the Editor Jeanne Hardebeck for providing constructive criticism. We thank Dr. Karen Britten for improving the readability of the manuscript.

Bibliography

Agostinelli, C and R. Rotondi (2003).Using Bayesian belief networks to analyse the stochastic dependence between interevent time and size of earthquakes. *Journal of Seismology*, **7**, 281299.

Castello B. , M. Olivieri, and G. Selvaggi (2007). Local and duration magnitude determination for the Italian earthquake catalogue (1981-2002). *Bulletin of the Seismological Society of America* , **97**, 1B, 128-139.

Chiarabba C., L. Jovane, and R. Di Stefano (2005). A new view of Italian seismicity using 20 years of instrumental recordings, *Tectonophysics*, **395/3-4**, 251-268.

CPTI Working Group (2004). Catalogo Parametrico dei Terremoti Italiani, versione 2004 (CPTI04), INGV, Bologna.
http://emidius.mi.ingv.it/CPTI04/

Cornell, C.A. (1968). Engineering seismic risk analysis, *Bull. Seism. Soc. Am.*, **58**, 1583-1606.

Field, E.H., T.H. Jordan, and C.A. Cornell (2003). OpenSHA. A Developing Community-Modeling Environment for Seismic Hazard Analysis. *Seismological Research Letters*, **74**, 406-419.

Gelman, A., J.B. Carlin, H.S. Stren, and D.B. Rubin (2004). *Bayesian Data Analysis*, Chapman and Hall, New York, 2nd edition, pp 696.

Gutenberg, B., and C. Richter (1954). *Seismicity of the Earth and Associated Phenomena*, Princeton University Press, Princeton, New Jersey, 2nd edition, pp

310.

MacKay D.J.C. (2003) *Information Theory, Inference, and Learning Algorithms.* Cambridge University Press, pp 640.

Marzocchi, W. and L. Sandri (2003). A review and new insight on the estimation of the b-value and its uncertainty, *Ann. Geophys*, **46,** 6, 1271-1282.

Marzocchi, W., L. Sandri, P. Gasparini, C. Newhall, and E. Boschi (2004). Quantifying probabilities of volcanic events: the examples of volcanic hazard at Mount Vesuvius, *J. Geophys.Res*, **109**, B11201, doi:10.1029/2004JB003155.

Marzocchi, W., L. Sandri, and J. Selva (2008). BET_EF: a probabilistic tool for long- and short- term eruption forecasting, *Bull. Volcanol.*, **70**, 623-632, doi: 10.2007/s00445-007-0157-y.

McGuire, R.K. (1976). Fortran computer program for seismic risk analysis, *US Geological Survey open-File Report 76-67.*

Meletti C., F. Galadini, G. Valensise, M. Stucchi, R. Basili, S. Barba, G. Vannucci, and E. Boschi (2008). A seismic source model for the seismic hazard assessment of the Italian territory. *Tectonophysics*, **450(1)**, 85-108. doi:10.1016/j.tecto. 2008.01.003.

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, 2004 april, 65 pp. 5 appendixes. http://zonesismiche.mi.ingv.it

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, *Ann. Geophys*, **50**,3, 329-339

Senior Seismic Hazard Analysis Committee (SSHAC) (1997). Recommendations for probabilistic seismic hazard analysis-guidance on uncertainty and use of experts, U.S. Nuclear Regulatory Commission, NUREG/CR-6372.

Schwartz, D.P., and K.J. Coppersmith (1984). Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res.* **89** 5681-5698.

Woo, G. (1999). The mathematics of natural catastrophes. Imperial College Press, London.

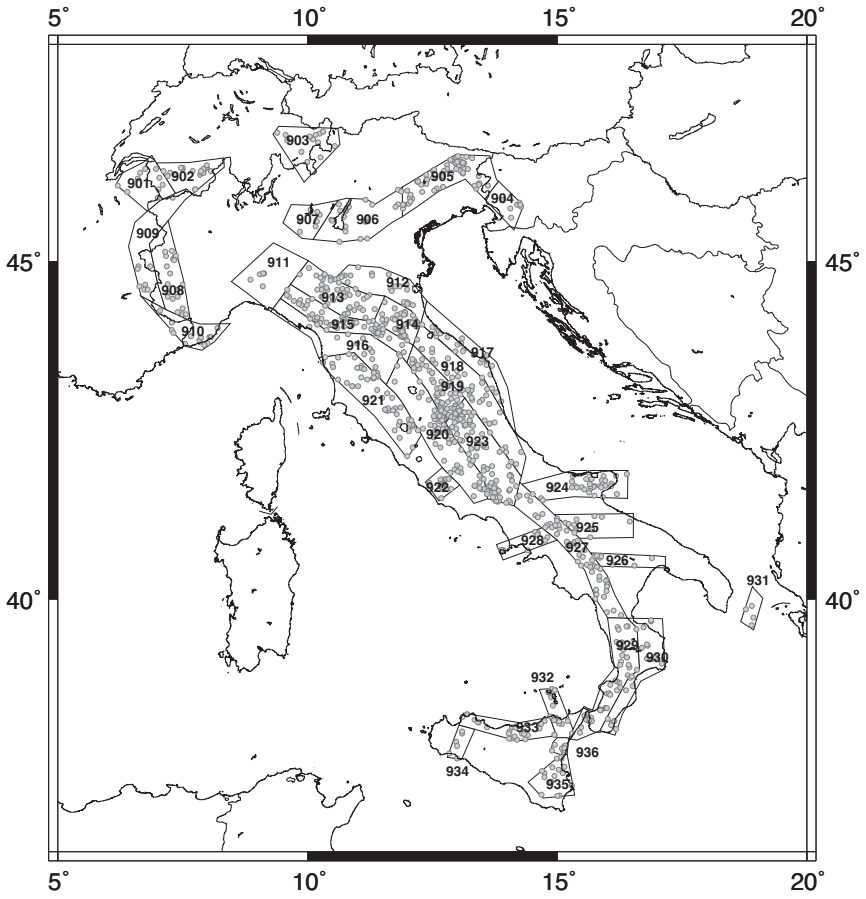


Figure 1: The seismic zonation ZS9 and distribution of the events in CPTI04 catalog that are in the complete portion according to the historical approach.

	Cl. 1 M4.76	Cl. 2 M4.99	Cl. 3 M5.22	Cl. 4 M5.45	Cl. 5 M5.68	Cl. 6 M5.91	Cl. 7 M6.14	Cl. 8 M6.37	Cl. 9 M6.60	Cl. 10 M6.83	Cl. 11 M7.06	Cl. 12 M7.29
ZS9001	1871(2)	1871(1)	1700(4)	1700(1)	1530(2)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9002	1871(7)	1871(2)	1700(5)	1700(3)	1530(1)	1530(2)	1300(1)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9003	1871(10)	1871(0)	1700(5)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9004	1871(4)	1871(2)	1700(0)	1700(0)	1530(1)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9005	1836(26)	1836(13)	1530(12)	1530(6)	1530(5)	1300(3)	1300(0)	1100(2)	1100(3)	1100(0)	1100(0)	1100(0)
ZS9006	1836(8)	1836(8)	1530(6)	1530(1)	1530(2)	1300(0)	1300(0)	1100(0)	1100(1)	1100(0)	1100(0)	1100(0)
ZS9007	1836(4)	1836(1)	1530(1)	1530(0)	1530(1)	1300(0)	1300(0)	1100(0)	1100(0)	1100(0)	1100(0)	1100(0)
ZS9008	1871(11)	1871(3)	1700(4)	1700(1)	1530(1)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9009	1871(4)	1871(1)	1700(2)	1700(1)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9010	1871(6)	1871(1)	1700(0)	1700(3)	1530(2)	1530(1)	1300(0)	1300(1)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9011	1871(3)	1871(0)	1700(1)	1700(0)	1530(1)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9012	1836(7)	1836(3)	1530(3)	1530(7)	1530(1)	1300(2)	1300(0)	1100(0)	1100(0)	1100(0)	1100(0)	1100(0)
ZS9013	1836(18)	1836(10)	1530(9)	1530(4)	1530(3)	1300(1)	1300(0)	1100(0)	1100(0)	1100(0)	1100(0)	1100(0)
ZS9014	1836(14)	1836(10)	1530(8)	1530(4)	1530(2)	1300(4)	1300(0)	1100(0)	1100(0)	1100(0)	1100(0)	1100(0)
ZS9015	1871(23)	1871(11)	1650(14)	1650(0)	1530(2)	1530(1)	1300(1)	1300(1)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9016	1871(5)	1871(4)	1530(6)	1530(1)	1530(1)	1300(0)	1300(0)	1200(0)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9017	1836(7)	1836(4)	1650(5)	1650(3)	1530(5)	1530(3)	1530(0)	1200(0)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9018	1871(21)	1871(3)	1650(6)	1650(2)	1650(3)	1530(3)	1530(2)	1300(0)	1300(1)	1300(0)	1300(0)	1300(0)
ZS9019	1871(15)	1871(7)	1530(14)	1530(6)	1530(2)	1300(7)	1300(1)	1200(2)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9020	1871(23)	1871(10)	1650(23)	1650(3)	1650(0)	1300(0)	1300(0)	1300(0)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9021	1871(21)	1871(11)	1530(11)	1530(4)	1530(2)	1300(1)	1300(0)	1200(0)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9022	1871(6)	1871(3)	1530(8)	1530(2)	1530(0)	1300(0)	1300(0)	1200(0)	1200(0)	1200(0)	1200(0)	1200(0)
ZS9023	1871(51)	1871(11)	1650(25)	1650(8)	1650(3)	1530(5)	1530(1)	1300(4)	1300(2)	1300(1)	1300(1)	1300(0)
ZS9024	1871(9)	1871(5)	1787(9)	1787(6)	1787(2)	1530(0)	1530(2)	1400(0)	1400(0)	1400(1)	1400(0)	1400(0)
ZS9025	1871(5)	1871(2)	1787(2)	1787(0)	1787(0)	1530(0)	1530(0)	1400(2)	1400(0)	1400(1)	1400(0)	1400(0)
ZS9026	1871(3)	1871(2)	1787(4)	1787(0)	1787(0)	1530(1)	1530(0)	1400(0)	1400(0)	1400(0)	1400(0)	1400(0)
ZS9027	1895(21)	1895(6)	1787(11)	1787(1)	1787(2)	1530(3)	1530(1)	1530(2)	1400(3)	1400(3)	1400(2)	1400(0)
ZS9028	1871(2)	1871(2)	1787(4)	1787(0)	1530(3)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9029	1895(23)	1895(4)	1787(14)	1787(6)	1787(1)	1787(3)	1530(4)	1530(0)	1400(2)	1400(1)	1400(2)	1400(1)
ZS9030	1895(7)	1895(3)	1787(1)	1787(2)	1787(2)	1787(1)	1530(1)	1530(1)	1400(1)	1400(1)	1400(0)	1400(0)
ZS9031	1950(1)	1950(1)	1787(0)	1787(0)	1787(1)	1787(0)	1530(0)	1530(0)	1530(0)	1530(1)	1530(0)	1530(0)
ZS9032	1895(8)	1895(2)	1700(5)	1700(1)	1700(0)	1530(1)	1530(1)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9033	1871(10)	1871(3)	1700(5)	1700(5)	1700(1)	1530(2)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9034	1895(2)	1895(0)	1700(2)	1700(1)	1700(0)	1530(0)	1530(1)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9035	1871(4)	1871(0)	1700(5)	1700(2)	1530(3)	1530(0)	1530(0)	1150(0)	1150(2)	1150(0)	1150(0)	1150(0)
ZS9036	1871(5)	1871(3)	1700(4)	1700(0)	1530(0)	1530(1)	1530(0)	1150(0)	1150(0)	1150(0)	1150(0)	1150(0)

Table 1: Starting date of the complete period in the CPTI04 catalog according to the historical approach for the ZS9 seismic source zones. In round brackets the number of events used in the analysis.

	Cl. 1 M4.76	Cl. 2 M4.99	Cl. 3 M5.22	Cl. 4 M5.45	Cl. 5 M5.68	Cl. 6 M5.91	Cl. 7 M6.14	Cl. 8 M6.37	Cl. 9 M6.60	Cl. 10 M6.83	Cl. 11 M7.06	Cl. 12 M7.29
ZS9001	1910(2)	1871(1)	1871(4)	1700(1)	1700(2)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9002	1910(3)	1871(2)	1871(1)	1700(3)	1700(1)	1530(2)	1530(1)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9003	1910(5)	1871(0)	1871(2)	1700(0)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9004	1910(4)	1871(2)	1871(0)	1700(0)	1700(1)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9005	1910(15)	1871(11)	1871(7)	1700(6)	1700(5)	1530(3)	1530(0)	1400(2)	1400(2)	1300(0)	1300(0)	1300(0)
ZS9006	1910(5)	1871(8)	1871(4)	1700(1)	1700(2)	1530(0)	1530(0)	1400(0)	1400(0)	1300(0)	1300(0)	1300(0)
ZS9007	1910(2)	1871(1)	1871(0)	1700(0)	1700(1)	1530(0)	1530(0)	1400(0)	1400(0)	1300(0)	1300(0)	1300(0)
ZS9008	1910(6)	1871(3)	1871(2)	1700(1)	1700(1)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9009	1910(3)	1871(1)	1871(2)	1700(1)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9010	1910(2)	1871(1)	1871(0)	1700(3)	1700(1)	1530(1)	1530(0)	1300(1)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9011	1910(2)	1871(0)	1871(1)	1700(0)	1700(1)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9012	1910(3)	1871(3)	1871(0)	1700(4)	1700(1)	1530(1)	1530(0)	1400(0)	1400(0)	1300(0)	1300(0)	1300(0)
ZS9013	1910(11)	1871(8)	1871(4)	1700(4)	1700(3)	1530(0)	1530(0)	1400(0)	1400(0)	1300(0)	1300(0)	1300(0)
ZS9014	1910(7)	1871(9)	1871(4)	1700(3)	1700(2)	1530(4)	1530(0)	1400(0)	1400(0)	1300(0)	1300(0)	1300(0)
ZS9015	1920(17)	1871(11)	1871(8)	1700(0)	1700(2)	1700(0)	1530(1)	1530(1)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9016	1920(2)	1871(4)	1871(4)	1700(1)	1700(1)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9017	1920(5)	1871(3)	1871(3)	1700(3)	1700(3)	1700(3)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9018	1920(13)	1871(3)	1871(4)	1700(2)	1700(3)	1700(3)	1530(2)	1530(0)	1300(1)	1300(0)	1300(0)	1300(0)
ZS9019	1920(11)	1871(7)	1871(5)	1700(5)	1700(2)	1530(4)	1530(1)	1530(1)	1900(0)	1300(0)	1900(0)	1300(0)
ZS9020	1920(7)	1871(10)	1871(12)	1700(3)	1700(0)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9021	1920(9)	1871(11)	1871(4)	1700(4)	1700(1)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9022	1920(0)	1871(3)	1871(4)	1700(2)	1700(0)	1700(0)	1530(0)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9023	1920(24)	1871(11)	1871(15)	1700(8)	1700(3)	1700(4)	1530(1)	1530(1)	1530(1)	1300(1)	1300(1)	1300(0)
ZS9024	1920(7)	1895(5)	1895(4)	1787(6)	1787(2)	1530(0)	1530(2)	1530(0)	1530(0)	1400(1)	1400(0)	1400(0)
ZS9025	1920(4)	1895(2)	1895(2)	1787(0)	1787(0)	1530(0)	1530(0)	1530(2)	1530(0)	1400(1)	1400(0)	1400(0)
ZS9026	1920(2)	1895(2)	1895(3)	1787(0)	1787(0)	1530(1)	1530(0)	1530(0)	1530(0)	1400(0)	1400(0)	1400(0)
ZS9027	1920(12)	1895(6)	1895(5)	1787(1)	1787(2)	1530(3)	1530(1)	1530(2)	1530(2)	1400(3)	1400(2)	1400(0)
ZS9028	1920(11)	1895(2)	1895(1)	1787(0)	1530(3)	1530(0)	1530(0)	1530(0)	1530(2)	1300(0)	1300(0)	1300(0)
ZS9029	1920(12)	1895(4)	1895(5)	1787(6)	1787(2)	1787(1)	1787(3)	1530(0)	1530(2)	1400(1)	1400(2)	1400(1)
ZS9030	1920(2)	1895(3)	1895(0)	1787(2)	1787(2)	1787(1)	1780(1)	1530(1)	1530(1)	1400(1)	1400(0)	1400(1)
ZS9031	1950(1)	1950(1)	1950(0)	1787(0)	1787(1)	1787(0)	1780(0)	1530(0)	1530(0)	1530(1)	1450(0)	1530(0)
ZS9032	1920(4)	1895(2)	1895(3)	1700(1)	1700(0)	1700(1)	1700(1)	1450(0)	1450(0)	1450(0)	1450(0)	1450(0)
ZS9033	1920(5)	1895(2)	1895(3)	1700(5)	1700(1)	1700(2)	1700(0)	1450(0)	1450(0)	1450(0)	1300(0)	1450(0)
ZS9034	1920(2)	1895(0)	1895(1)	1700(1)	1700(0)	1700(0)	1530(1)	1530(0)	1300(0)	1300(0)	1300(0)	1300(0)
ZS9035	1920(3)	1895(0)	1895(3)	1700(2)	1700(2)	1700(0)	1700(0)	1450(0)	1450(1)	1450(0)	1450(0)	1450(0)
ZS9036	1920(3)	1895(2)	1895(3)	1700(0)	1700(0)	1700(1)	1700(0)	1450(0)	1450(0)	1450(0)	1450(0)	1450(0)

Table 2: Starting date of the complete period in the CPTI04 catalog according to the statistical approach for the ZS9 seismic source zones. In round brackets the number of events used in the analysis.

a1 = 5.549	a2 = 2.987	a3 = 1.606	a4 = 0.864
a5 = 0.466	a6 = 0.250	a7 = 0.133	a8 = 0.074
a9 = 0.037	a10 = 0.024	a11 = 0.011	a12 = 0.004

Table 3: A priori α parameters. They are the normalized frequency of occurrence of the 12 magnitude classes, on the basis of INGV-CNT catalogue, with b-value of 1.17.

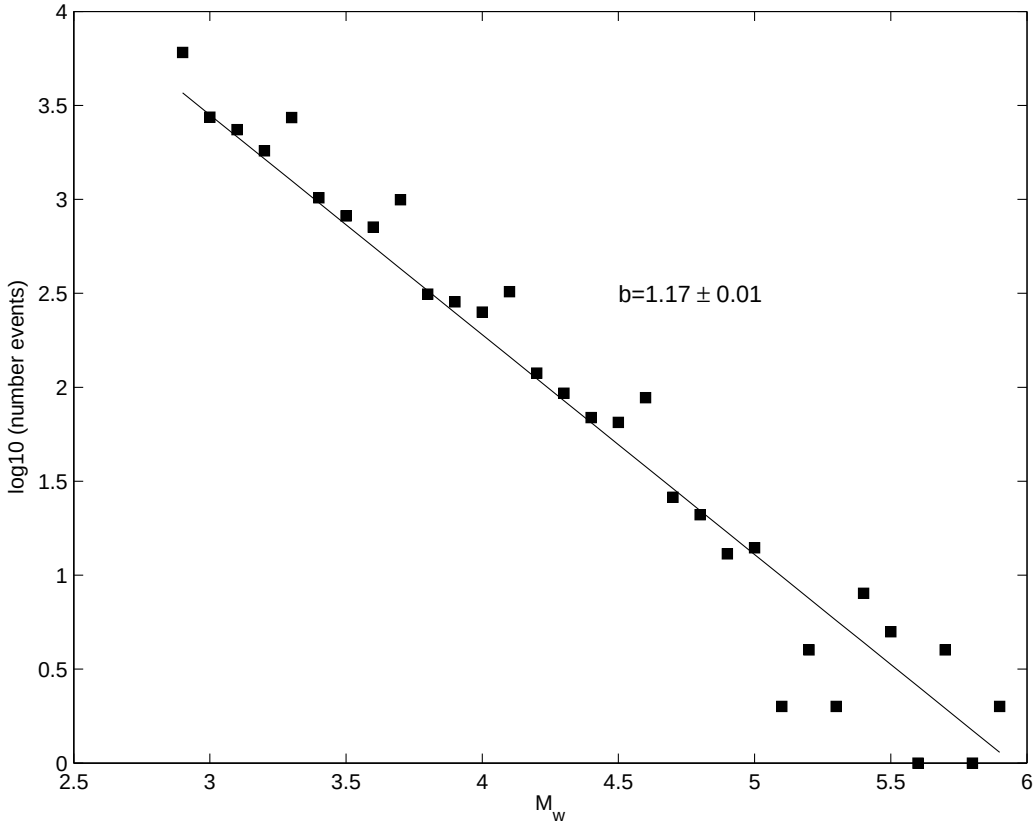


Figure 2: The magnitude-frequency relation of $M_w \geq 2.9$ events since 1981 (data from INGV-CNT catalog). We plot the value of b of the Gutenberg-Richter relation from a Maximum Likelihood Estimation.

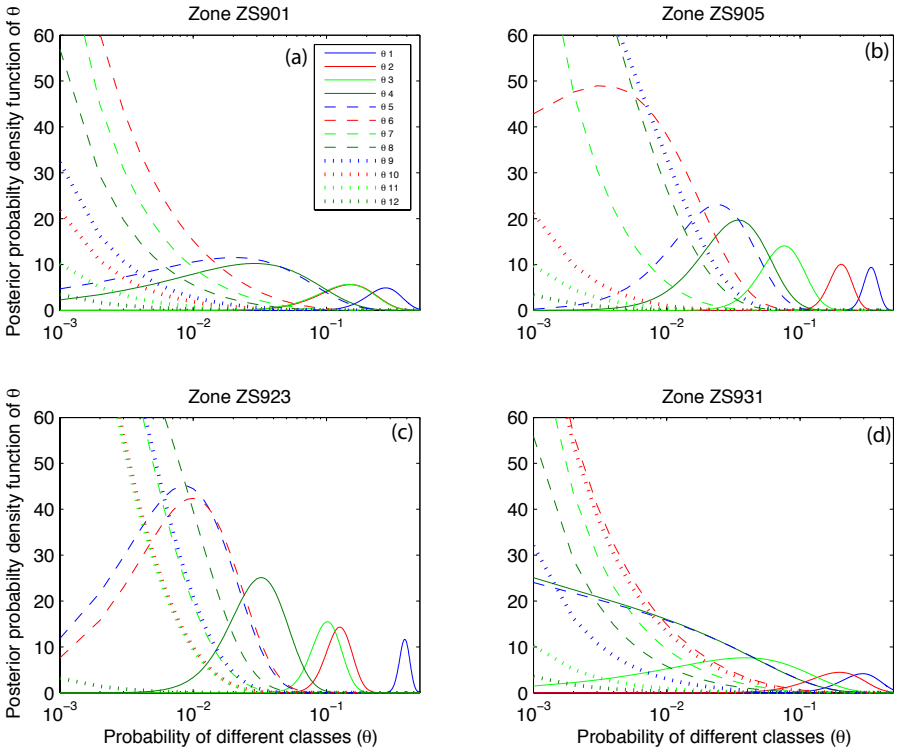
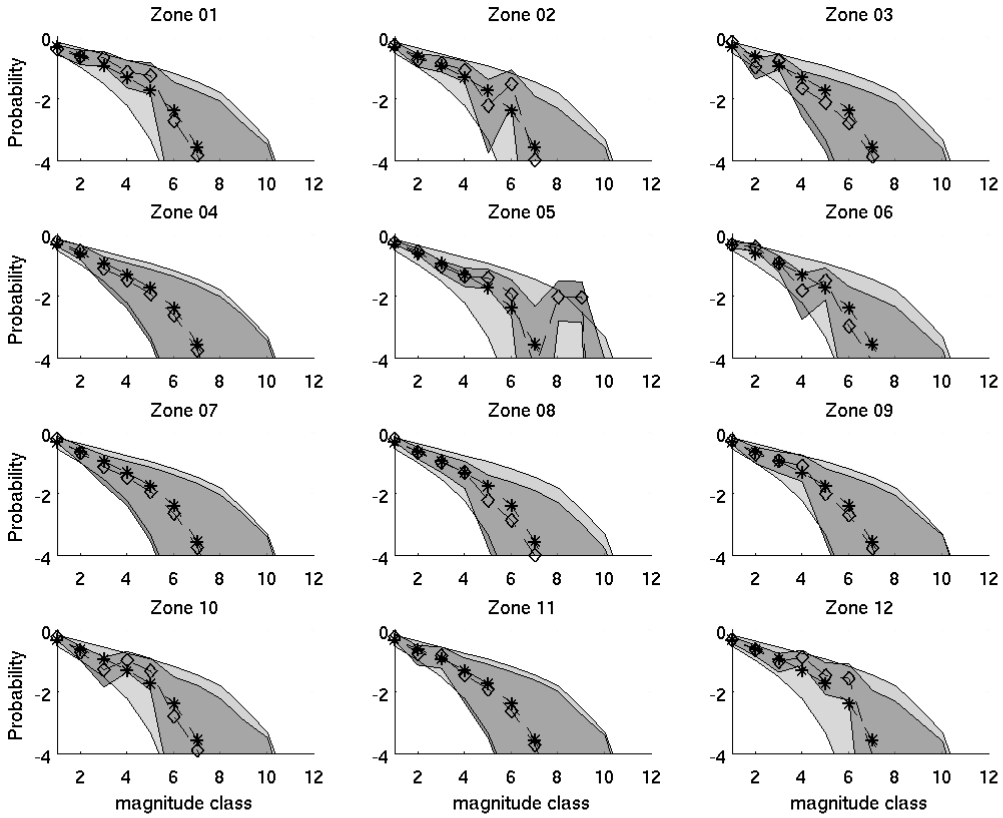
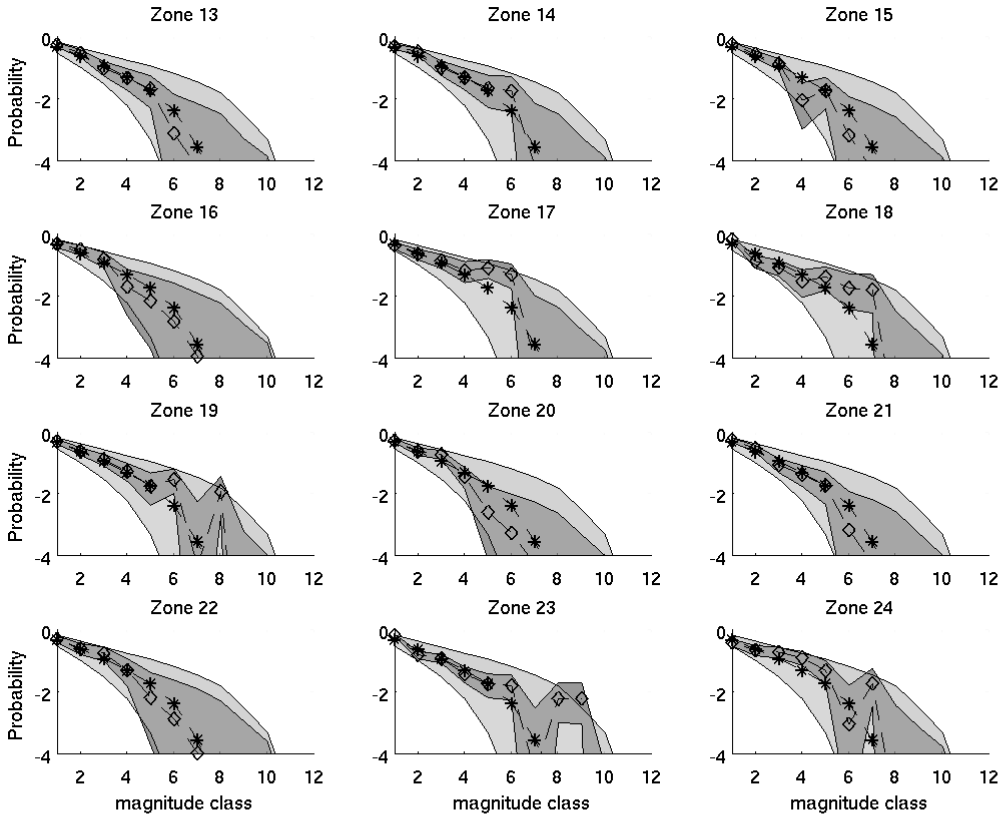


Figure 3: Marginals posterior distributions for the 12 states for zone ZS901(a), ZS905 (b), ZS923 (c) and ZS931(d)


Zone 01
Zone 02
Zone 03
Zone 04
Zone 05
Zone 06
Zone 07
Zone 08
Zone 09
Zone 10
Zone 11
Zone 12
Probability
0
-2
-4
2
4
6
8
10
12
magnitude class


Zone 13
Zone 14
Zone 15
Zone 16
Zone 17
Zone 18
Zone 19
Zone 20
Zone 21
Zone 22
Zone 23
Zone 24
Probability
0
-2
-4
2
4
6
8
10
12
magnitude class

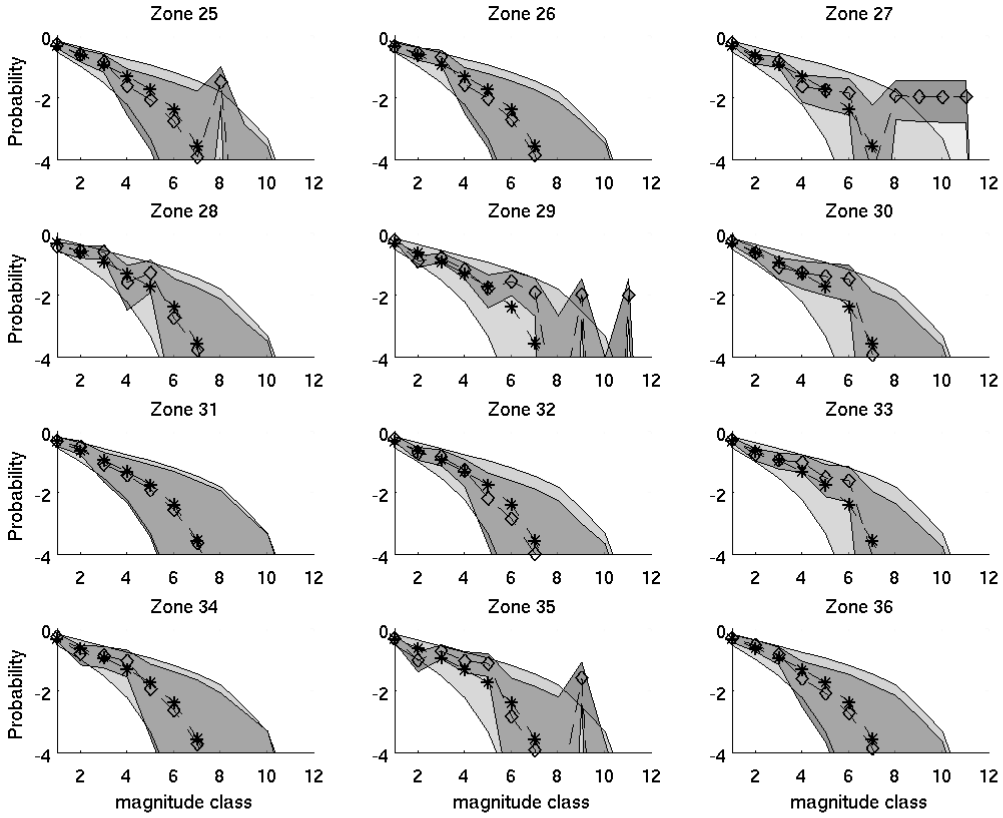
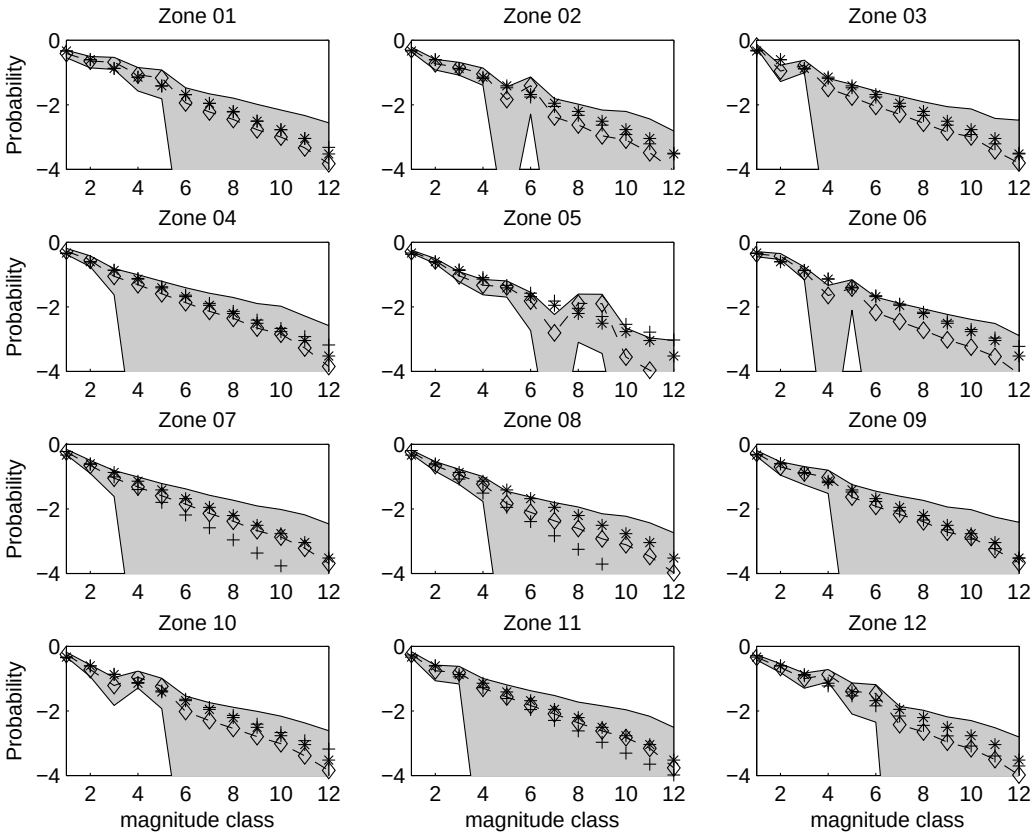
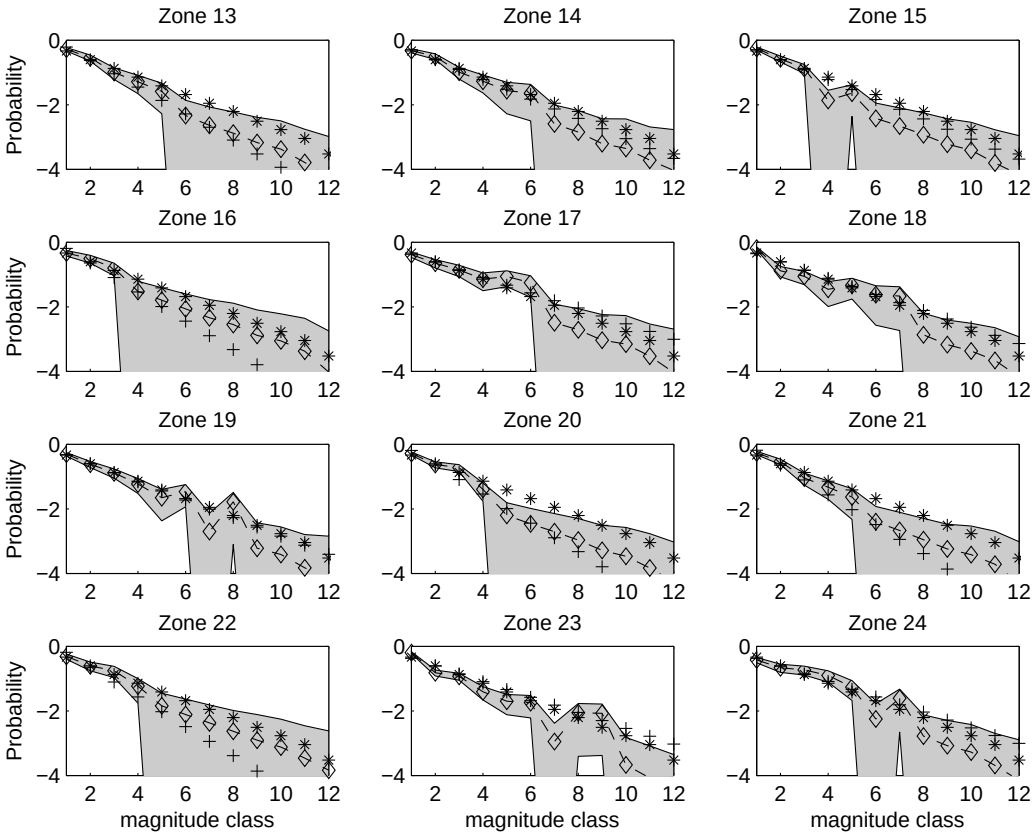


Figure 4: Marginal probability distribution of earthquake size for the different zones based on the historical completeness estimates, for the prior and posterior distributions. Diamonds represent the 50 percentile for the posterior Dirichlet distribution; Stars represent the 50 percentile for the prior Dirichlet distribution. Dark-gray area covers the area of the 10-90 percentile of the posterior Dirichlet distribution; light-gray area the same but for the prior Dirichlet distribution. Top panel: zones 1 to 12; Middle panel: zones 13 to 24; Lower pannel: zones 25 to 36.


Zone 01
Zone 02
Zone 03
Zone 04
Zone 05
Zone 06
Zone 07
Zone 08
Zone 09
Zone 10
Zone 11
Zone 12
Probability
magnitude class
0
−2
−4
2
4
6
8
10
12


Zone 13
Zone 14
Zone 15
Zone 16
Zone 17
Zone 18
Zone 19
Zone 20
Zone 21
Zone 22
Zone 23
Zone 24
Probability
magnitude class
0
−2
−4
2
4
6
8
10
12

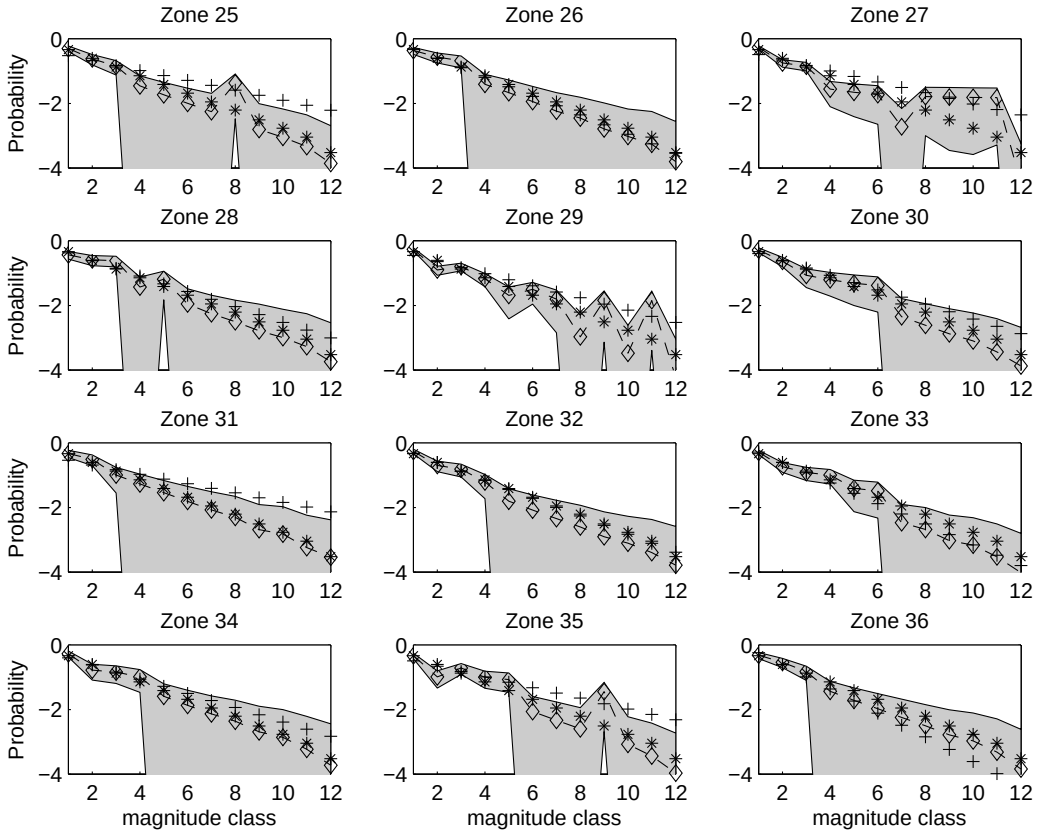


Figure 5: Comparison of the mean values of the probability distribution of earthquake size for the different zones based on the historical completeness estimates. Diamonds represents the mean for the posterior Dirichlet distribution. The gray area is the $\pm\sigma$ value. Stars: the mean of the prior distribution; Plus the MPS04-GR distribution. Top panel: zones 1 to 12; Middle pannel: zones 13 to 24; Lower pannel: zones 25 to 36.

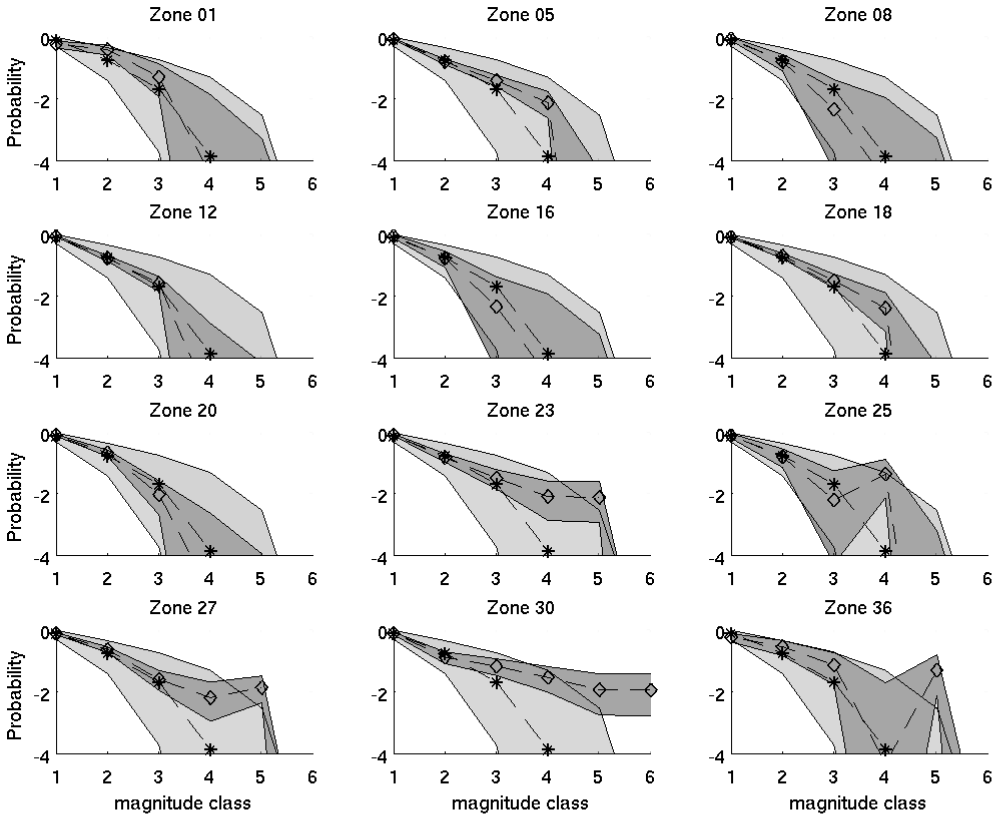


Figure 6: Marginal probability distribution of earthquake size - now for 6 larger classes - for the different zones based on the historical completeness estimates, for the prior and posterior distributions. Diamonds represent the 50 percentile for the posterior Dirichlet distribution; Stars represent the 50 percentile for the prior Dirichlet distribution. Dark-gray area covers the area of the 10-90 percentile of the posterior Dirichlet distribution; light-gray area the same but for the prior Dirichlet distribution. In picture, only 12 out of 36 zones are shown, as sake of comparison.

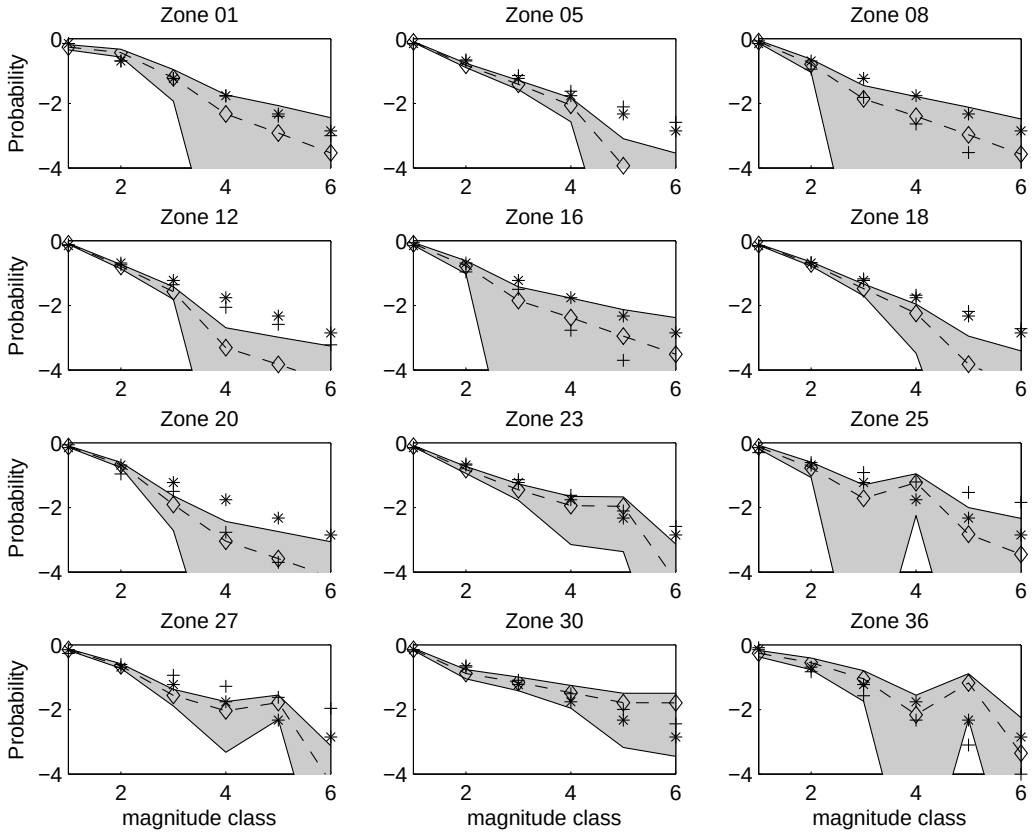


Figure 7: Comparison of the mean values of the probability distribution of earthquake size - now for 6 larger classes - for the different zones based on the historical completeness estimates. Diamonds represents the mean for the posterior Dirichlet distribution. The gray area is the $\pm\sigma$ value. Stars: the mean of the prior distribution; Plus the MPS04-GR distribution. In picture, only 12 out of 36 zones are shown, as sake of comparison.