Probabilistic analysis of macroseismic fields: Iceland case study

R. Rotondi, C. Brambilla & E. Varini
CNR - Istituto di Matematica Applicata e Tecnologie Informatiche, Milano, Italy

R. Sigbjörnsson
Earthquake Engineering Research Centre (EERC), Selfoss, Iceland

SUMMARY:
In this work our objectives are, first, to describe a probabilistic model aimed at forecasting macroseismic fields and, second, to present the results and findings obtained by applying the model to two of recent damaging earthquakes in the South Iceland Seismic Zone, namely, the earthquakes on 17 June, 2000 and 29 May, 2008, both of assessed epicentral intensity MMI X. The model considers the intensity decay as a random variable having a binomial probability distribution with parameter *p*, which is in its turn assumed to be a beta random variable. Estimation is carried out according to the Bayesian paradigm; on the basis of a learning set of macroseismic fields, we express our belief on the phenomenon by assigning prior distributions to the model parameters and we then update them with current data. The model has already been applied to Italian earthquakes, in particular, earthquakes occurring in the Mt. Etna volcanic environment.

Keywords: Attenuation, damage scenario forecast, probability distribution of site intensity, UPStrat-MAFA

1. HISTORICAL OUTLINE OF AN IDEA

The ability to forecast seismic scenarios in terms of macroseismic intensity at a site is of great importance. This issue has largely been analysed applying the deterministic point of view, but in the last decade it has become increasingly clear that the intensity at a site, as well as the intensity decay, must be expressed in probabilistic terms in order to obtain a more complete treatment of its intrinsic uncertainty.

Following the probabilistic approach, Rotondi and Zonno (2004) proposed to estimate the probability distribution of the intensity at a site, conditioned on the epicentral intensity and the epicentre-to-site distance, by using a binomial-beta model. The estimation process was carried out according to the Bayesian paradigm, exploiting a learning set of macroseismic fields to assign prior distributions of the model parameters. The model was originally tested on the Colfiorito earthquake (1997/9/26, central Italy). In that study the learning set included macroseismic fields from the seismogenetic zones of the zonation ZS4 (Meletti et al., 2000) judged homogeneous, from the viewpoint of the kinematic context and expected rupture mechanism, to the zone to which the epicentre of the Colfiorito earthquake belongs.

Subsequently, Zonno et al. (2009) were the first to analyse Italian macroseismic fields through summaries of the spatial distribution of intensity decay in order to detect groups of earthquakes homogeneous from the attenuation point of view. The earthquakes considered in their study were 55 earthquakes of epicentral intensity MCS $\geq$ VII, selected from the DBMI04 Italian database (Stucchi et al., 2007) and judged to be representative of the temporal and spatial distribution of Italian seismicity. Each macroseismic field was characterized by location and dispersion measures computed for each set of distances from the epicentre to the sites where the same intensity was observed, and, on the basis of this information, the earthquakes were grouped by using an agglomerative clustering method (Kaufman and Rousseeuw, 1990). Three groups corresponding to three different decay trends were detected and employed as learning sets to reproduce the macroseismic fields by applying the binomial-beta model separately inside each group. Afterwards, the process was repeated with a much larger

number of earthquakes of epicentral intensity MCS $\geq$ VII and four groups of different decay trend were detected. At present, these are the groups available as possible learning sets for future studies.
The model has also already been applied to earthquakes occurring in the Mt. Etna volcanic district (Azzaro et al., submitted). Since these earthquakes were not used for the clustering procedure, this application can be seen as a test, both for the model and for the relevance of the learning sets derived as explained above. In that study the Italian macroseismic fields used as a learning set was the one corresponding to the fastest decay trend, since the decay trend of the earthquakes on the flank of Mt. Etna is very quick, due to the fractured ground and to very shallow seismicity activity (Azzaro et al., 2006).
The present work describes the application of the model to two recent damaging earthquakes in Iceland, in the South Seismic Iceland Zone. It is the first application of the model to earthquakes of a seismic region outside Italy. The work is part of a study of seismic regions of different European countries with the aim of implementing common strategies to forecast damage scenarios from macroseismic fields and to assess the seismic hazard.

2. THE PROBABILISTIC MODEL

The binomial-beta model is based on the hypothesis that, conditioned on the epicentral intensity I_0 and on a fixed epicentral distance, the intensity decay ΔI has a binomial distribution with parameter p, i.e.

$$Pr(\Delta I = i_0 - i \mid I_0 = i_0, p) = \binom{i_0}{i} p^i (1-p)^{(i_0 - i)}, \tag{2.1}$$

which is also equivalent to assuming that the intensity at a given site I_s has a binomial distribution with parameter p, since

$$Pr(I_s = i \mid I_0 = i_0, p) = Pr(\Delta I = i_0 - i \mid I_0 = i_0, p). \tag{2.2}$$

This choice is predicated on respecting insofar as far as possible the ordinal nature of the intensity scale applied. The parameter p, in its turn, is taken as a random variable in order to account for the variability in ground shaking even among sites with the same epicentral distance; it is assumed to have the beta distribution

$$Be(p;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^p x^{\alpha-1}(1-x)^{\beta-1} dx \tag{2.3}$$

with parameters α and β. Γ denotes the gamma function. The mean and the variance are given, respectively, as follows

$$E(p) = \frac{\alpha}{\alpha+\beta} \quad , \qquad \sigma^2(p) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad . \tag{2.4}$$

The beta distribution has been adopted due to its great flexibility and tractability within the Bayesian framework.
In this work we assume the isotropic model for the decay. According to this the model parameters are directionally independent. Because of the dependence of the attenuation from the epicentre-site distance, we consider J distance bins distributed around the epicentre and assume that in all the sites within each j th distance bin, ΔI has the same binomial distribution with parameter p_j. In turn, each p_j has a beta distribution with hyperparameters α_j and β_j. The width of the bins may vary, depending on the case considered.

Let us assume that we estimate macroseismic fields of earthquakes of given epicentral intensity I_0. Having selected a suitable learning set through which prior knowledge about the phenomenon can be expressed, the estimation algorithm proceeds as follows: In the first step we assign the hyperparameters $\alpha_{j,0}$ and $\beta_{j,0}$ of the prior beta distribution of the parameters p_j, on the basis of the information provided by the macroseismic fields of the earthquakes of epicentral intensity I_0 belonging to the selected learning set. We roughly estimate the probability of null decay $Pr(\Delta I = 0 | I_0 = i_0, j) = p_j^{i_0}$ in the j th distance bin by the relative frequency of null decay $N_j(i_0)/N_j$, where $N_j(i_0)$ is the number of sites in the j th distance bin where the intensity at site is not smaller than the epicentral intensity. Hence, the initial mean value for p_j will be $p_{j,0} = (N_j(i_0)/N_j)^{1/I_0}$. Fixed the variance of p_j we invert Eqns. 2.4 to obtain the values of the hyperparameters $\alpha_{j,0}$ and $\beta_{j,0}$.

In the second step, we compute the posterior beta distribution of the parameters p_j on the basis of the current macroseismic fields and estimate each p_j through its posterior mean

$$\hat{p}_j = \frac{\alpha_{j,0} + \sum_{n_j=1}^{N_j} i_s^{(n_j)}}{\alpha_{j,0} + \beta_{j,0} + I_0 \cdot N_j} \quad , \tag{2.5}$$

where $i_s^{(n_j)}$ is the intensity at the n_j-th site inside the j th bin and N_j is the total number of sites in that bin.

By smoothing the posterior mean of p in each bin through an inverse power function $g(d) = (\gamma_1 / d)^{\gamma_2}$, we can express this parameter as a continuous function of the epicentral distance d. We can then estimate I_s at any distance d from the epicentre by using what we call the smoothed binomial function:

$$\Pr_{smooth}(I_s = i \,|\, I_0 = i_0, d) = \binom{i_0}{i} g(d)^i \left(1 - g(d)\right)^{(i_0 - i)} \quad . \tag{2.6}$$

The mode of the smoothed binomial distribution, i_{smooth}, is taken as an estimate of the intensity at site I_s. Through the posterior distribution of the parameters, the Bayesian paradigm also provides rational measures of the parameter uncertainties.

3. THE ICELAND CASE STUDY

All major damaging earthquakes in Iceland have originated within two fracture zones, one in the south, called the South Iceland Seismic Zone (SISZ), and one in the North, usually called the Tjornes Fracture Zone (TFZ). In this study we consider the damaging earthquakes striking in SISZ on June 17, 2000 and May 29, 2008 (Sigbjörnsson et al., 2007, Sigbjörnsson et al., 2009), of $6.5M_w$ and $6.3M_w$, respectively. They are among the most important seismic events in the area since 1896, when a sequence of moderate-to-strong earthquakes took place in SISZ over a period of two weeks. The macroseismic epicentre (63.97°N, 20.36°W) of the first earthquake was in the central part of SISZ, just north of the rural village of Hella, while the other (63.98°N, 21.13°W) was in the westernmost part of SISZ, precisely in the Ölfus District, between the towns of Selfoss and Hveragerdi. For this earthquake there are indications that the recorded earthquake waves were not generated by a single causative fault but by the almost simultaneous rupturing of two parallel faults.

For both the earthquakes $I_0 = X$ on the MMI intensity scale (Wood and Neuman, 1931, with the extensions referring to "Icelandic building tradition" by Tryggvason, 1979). Although nowadays in Iceland the most recent EMS scale has been adopted, the MMI scale is still used in an attempt to preserve continuity with earlier studies, in which the bulk of the intensity data was collected according to the MMI scale.
The data points of the earthquakes on June 17, 2000 and May 29, 2009 are 434 and 145, respectively.
For both earthquakes it is known that the intensity attenuates quite rapidly with increasing distance from the causative faults, as can clearly be seen by the exploratory analysis of the spatial distribution of the seismic decay pictured in Fig. 3.1.

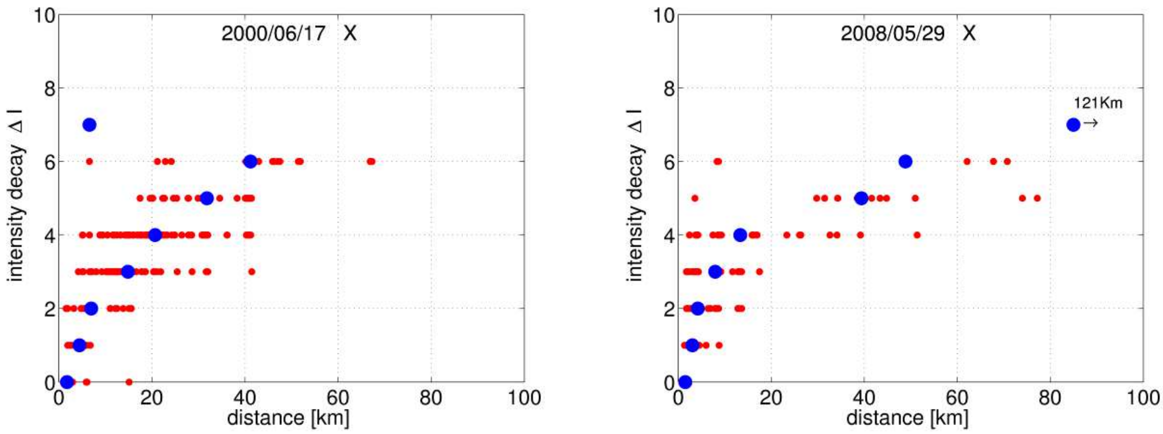

Figure 3.1. Intensity decay (red dots) vs. epicentral distance for the 2000/06/17 (left) and the 2008/05/29 (right) earthquakes. Blue dots denote the median of the distance subsets.

In light of the attenuation trend of the two earthquakes, among the four groups of Italian macroseismic fields which at present can be considered as learning sets, we selected one of the two characterized by a steep attenuation trend and used the information provided by the earthquakes of that set having $I_0 = X$ for assigning the hyperparameters $\alpha_{j,0}$ and $\beta_{j,0}$ of the prior beta distribution of the parameters p_j. Given the difference in the impact area of the Italian earthquakes and of the two Icelandic earthquakes, and, above all, given the differences in the distances at which the same decays were approximately observed, we decided to shrink the width of the bins (originally set up at 10 km) by suitable coefficients. The observed and estimated intensities are pictured in Fig. 3.2 and Fig. 3.3 on the left and the right, respectively.
To validate the results we used three criteria. The first one is the so-called logarithmic scoring rule, based on the logarithm of the likelihood function:

$$score_{smooth} = -\frac{1}{N}\log\prod_{n=1}^{N}\binom{i_0}{i_s^{(n)}} g(d_n)^{i_s^{(n)}}\left(1-g(d_n)\right)^{\left(i_0-i_s^{(n)}\right)} \quad , \tag{3.1}$$

where N is the total number of the observed intensities at the site; $i_s^{(n)}$ is the intensity at site n and d_n is the distance of site n from the epicentre.
The second criterion is based on the $p(O)/p(F)$ ratio between the probability that the fitted model assesses to an observation O and the probability of the forecast value F, that is, how much is gained from having predicted F when O occurs:

$$odds_{smooth} = -\frac{1}{N}\log\prod_{n=1}^{N}\frac{Pr_{smooth}\left(i_s^{(n)}\right)}{Pr_{smooth}\left(i_{smooth}^{(n)}\right)} \quad , \tag{3.2}$$

where $i_{smooth}^{(n)}$ is the estimate of the intensity at site n provided by the mode of the smoothed binomial distribution.
The third and last criterion is based on the absolute discrepancy between observed and estimated intensities at site:

$$diff_{smooth} = \frac{1}{N}\sum_{i=1}^{N} \left| i_s^{(n)} - i_{smooth}^{(n)} \right| \quad . \tag{3.3}$$

Table 3.1 shows the values we obtained for the two earthquakes.

Table 3.1. Results of the validation criteria applied to the 2000/06/17 and 2008/05/29 earthquakes.

earthquake	scoring	odds	discrepancy
2000/06/17	1.489	0.126	0.565
2008/05/29	1.576	0.307	0.779

The estimated macroseismic field for the 17 June 2000 event given in Fig. 3.2 reveals a fair overall prediction, confirmed by the results in Table 3.1.

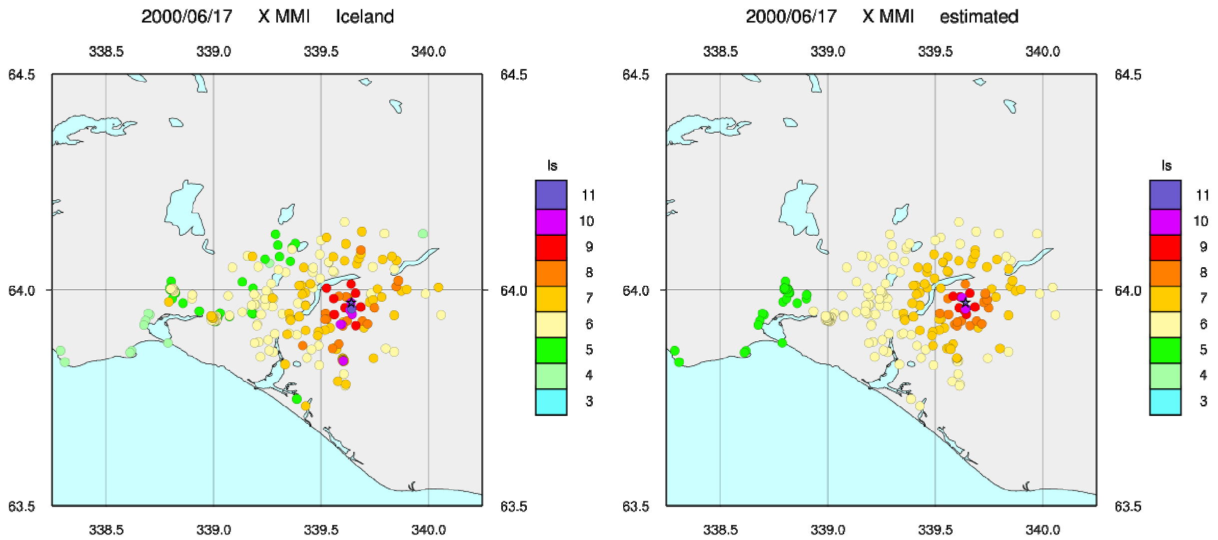

Figure 3.2. Macroseismic field of the 2000/06/17 earthquake in SISZ: observed intensities on the left, estimated intensities on the right. Stars denote the epicentre.

If we consider the third validation criterion, for instance, we see that the difference between observed and estimated intensities, on average, amounts only to about half a degree of the intensity scale. However, there are certain systematic discrepancies that can be traced back to the near-fault characteristics and the elongated shape of the epicentral region reflecting the finite surface fault trace. This is true, in particular, for underestimation of intensities southward from the epicentre.
Table 3.1 shows slightly worse results for the 29 May 2008 event, although the third validation criterion highlights a difference between observed and estimated intensities which, on average, is still less than one degree of the intensity scale. Inspecting Fig. 3.3 reveals some of the discrepancies between observed and estimated intensities. First of all, the modelling of the macroseismic epicentre as a point between the two causative faults leads to overestimation of intensities close to the virtual source representing the macroseismic epicentre. Moreover, there is an overall overestimation of the intensities at the Hveragerdi Village, whereas at the Town of Selfoss the highest intensities are underestimated. On the whole the results suggest that the forward directivity effects, which are missing in the current isotropic macroseismic field model, and some earthquakes peculiarities significantly influence the attenuation of the two shocks and suggest a need of further research to take these issues into account.

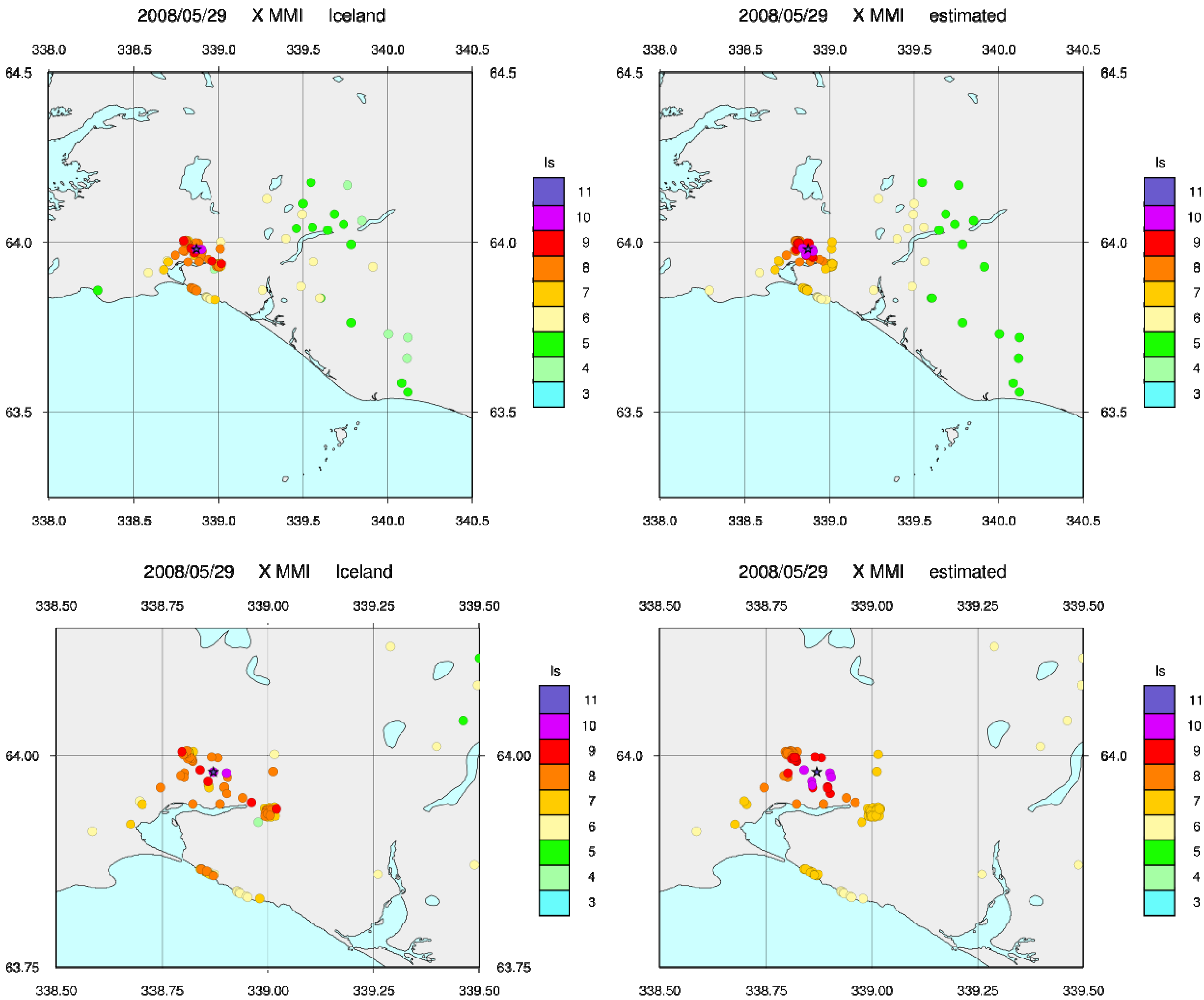

Figure 3.3. Macroseismic field of the 2008/05/29 earthquake in SISZ: observed intensities on the left, estimated intensities on the right. The bottom row provides a zoom of the area nearest to the epicentre. Stars denote the epicentre.

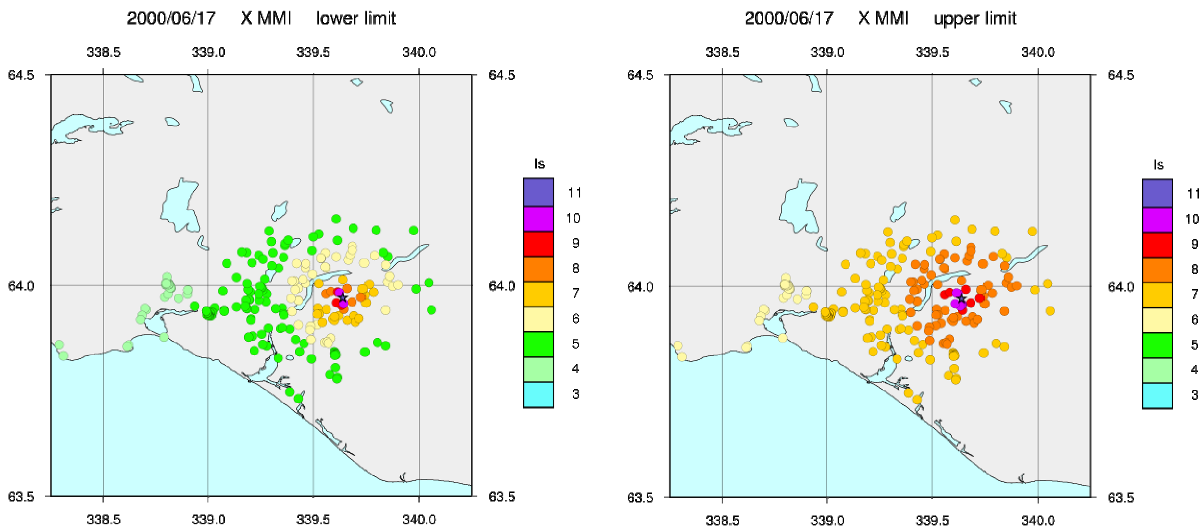

Figure 3.4. Plots of the shortest intervals with at least 50% mass probability for the 2000/06/17 earthquake.

To represent the dispersion of the distribution of I_s at each site given in Eqn. 2.6, in Fig. 3.4 and Fig. 3.5 we provide the extremes of the interval which, in the present study, is the smallest set of intensity values covering at least 50% probability. Figures are to be interpreted as follows: Consider a particular site on the map and look at the value of its intensity on the left and on the right; if, for example, the left value is V and the right value is VII, the intensity is between these two values with at least 50% of probability. Similarly, if the left value is VIII and the right value is IX, the intensity is either VIII or IX with at least 50% of probability, and so on.

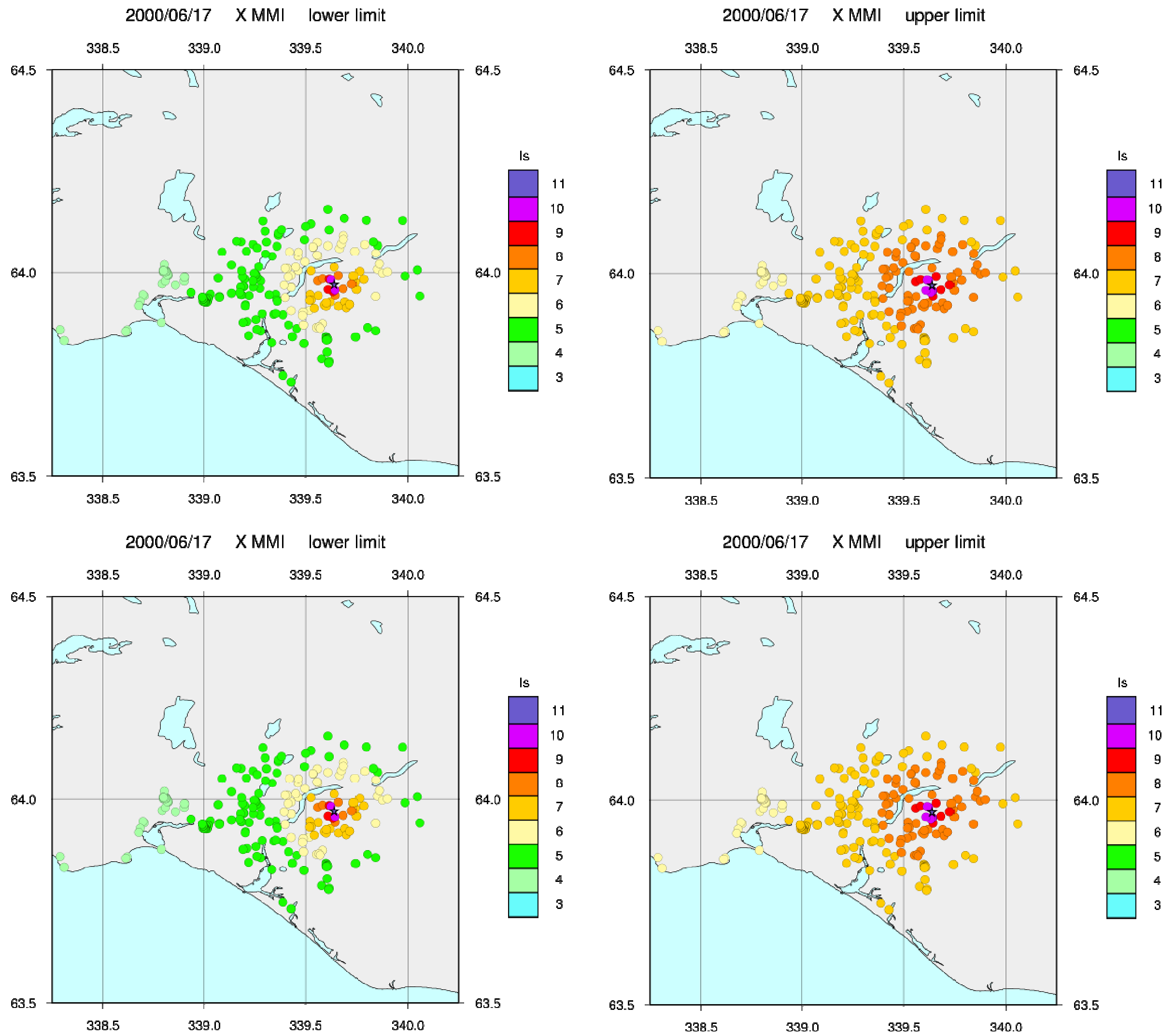

Figure 3.5. Plots of the shortest intervals with at least 50% mass probability for the 2008/05/29 earthquake. The bottom row provides a zoom of the area nearest to the epicentre. Stars denote the epicentre.

4. FINAL REMARKS

In this paper we have outlined the development of probabilistic macroseismic field modelling and mentioned some of its applications to Italian earthquake data. Furthermore, we have given a description of the model in mathematical settings and outlined rational validation procedures. Finally, we have applied the model to the macroseismic field of two Icelandic earthquakes.

The main result is that the current probabilistic model forecasts the two macroseismic fields reasonably well, and the use of three different validation criteria strengthens this finding. However, we observed certain systematic deviation that, as already pointed out, can be traced back to the specific nature of these two earthquakes and therefore suggests some modifications of the model. In particular, it is desirable to enhance representation of finite sources and to address the problems of directivity effects and of multiple causative faults.

ACKNOWLEDGEMENT

This work has been co-financed by the European project "Urban Prevention Strategies using Macroseismic and Fault sources" (UPStrat-MAFA - Num. 230301/2011/613486/SUB/A5), DG ECHO Unit A5. The maps were produced with the GMT software (Wessel and Smith, 1998).

REFERENCES

Azzaro, R., Barbano, M.S., D'Amico, S. and Tuvè, T. (2006). The attenuation of the seismic intensity in the Etna region and comparison with other Italian volcanic districts. *Annals of Geophysics*. **49:4/5**,1003–1020.

Azzaro, R., D'Amico, S., Rotondi, R., Tuvè, T. and Zonno G. Forecasting seismic scenarios on Etna volcano (Italy) through probabilistic intensity attenuation models: a Bayesian approach. *Journal of Volcanology and Geothermal Research*. submitted.

Halldórsson, B. and Sigbjörnsson, R. (2009). The Mw6.3 earthquake at 15:45 UTC on 29 May 2008 in South Iceland: Icearray strong-motion recordings. *Soil Dynamics and Earthquake Engineering*. **29**,1073–1083.

Kaufman, L. and Rousseeuw, P.J. (1990), Finding groups in data, Wiley, New York.

Meletti, C., Patacca, E. and Scandone, P. (2000). Construction of a seismotectonic model: the case of Italy. *Pageoph*. **157**,11–35.

Zonation ZS.4 available from http://emidius.mi.ingv.it/GNDT/P511/home.html.

Rotondi, R. and Zonno, G. (2004). Bayesian analysis of a probability distribution for local intensity attenuation. *Annals of Geophysics*. **47:5**,1521–1540.

Sigbjörnsson, R., Ólafsson, S. and Snaebjörnsson, J.T. (2007). Macoseismic effects related to strong ground motion; a study on the South Iceland earthquakes in June 2000. *Bullettin Earthquake Engineering*. **5**,591–608.

Sigbjörnsson, R., Snaebjörnsson, J.T., Higgins, S.M. and Ólafsson, S. (2009). A note on the Mw 6.3 earthquake in Iceland on 29 May 2008 at 15:45 UTC. *Bullettin Earthquake Engineering*. **7**,113–126.

Stucchi, M. et al. (2007). DBMI04, il database delle osservazioni macrosismiche dei terremoti italiani utilizzate per la compilazione del catalogo parametrico CPTI04. *Quaderni di Geofisica*. **49**,1–38.

Wessel, P. and Smith, W.H.F. (1998). New, improved version of Generic Mapping Tools released. *EOS Transactions of American Geophysical Union*. **79:49**,579.

Wood, H.O. and Neuman, F. (1931). Modified Mercalli intensity scale. *Bullettin of the Seismological Society of America*. **21:4**,277–283.

Zonno, G., Rotondi, R. and Brambilla, C. (2009). Mining macroseismic fields to estimate the probability distribution of the intensity at site. *Bulletin of Seismological Society of America*. **99:5**,2876–2892.

Tryggvason, E. (1979). Earthquakes in Iceland 1950–1959. *The Science Institute, University of Iceland, Report RH–79–06, Reykjavík*.