Pure appl. geophys. 159 (2002) 1289–1304
0033–4553/02/061289–16 $ 1.50 + 0.20/0

Pure and Applied Geophysics

Seismic Hazard Estimates Using Ill-defined Macroseismic Data at Site

D. ALBARELLO[1] and M. MUCCIARELLI[2]

Abstract—A new approach is proposed to the seismic hazard estimate based on documentary data concerning local history of seismic effects. The adopted methodology allows for the use of "poor" data, such as the macroseismic ones, within a formally coherent approach that permits overcoming a number of problems connected to the forcing of available information in the frame of "standard" methodologies calibrated on the use of instrumental data. The use of the proposed methodology allows full exploitation of all the available information (that for many towns in Italy covers several centuries) making possible a correct use of macroseismic data characterized by different levels of completeness and reliability. As an application of the proposed methodology, seismic hazard estimates are presented for two towns located in Northern Italy: Bologna and Carpi.

Key words: Seismic hazard, catalogue completeness, macroseismic data, intensity.

1. Introduction

Historical research carried out in Italy during recent years (see, e.g., BOSCHI *et al.*, 1997; MONACHESI and STUCCHI, 1997) has substantially improved present knowledge of past seismic history of the country. However, to date, this large data set has not been fully exploited for seismic hazard estimates in Italy. This is mostly due to the peculiar character of information produced by historical research that, due to its semi-qualitative character (in general documentary information is coded in terms of specific "macroseismic" scales), cannot be easily translated in quantitative terms suitable for the application of "standard" methodologies (e.g., BENDER and PERKINS, 1987) devoted to seismic hazard estimates. This problem may led some people to think that historical information is so "poor" with respect to instrumental data, that its actual contribution to reliable seismic hazard estimates is marginal. Others instead, tend to "force" seismic data deduced from documentary data (macroseismic information) into the frame of categories typical of instrumental data (earthquake

---

[1] Dip. di Scienze della Terra, Università di Siena, Siena, Italia. E-mail: albarello@unisi.it.
[2] Dip. di Strutture, Geotecnica e Geologia Applicata, Università della Basilicata, Potenza, Italia.

epicenter, macroseismic magnitude, etc.). In order to overcome both of these misleading attitudes, it is necessary to develop quantitative methodologies specifically devoted to the use of macroseismic data able to take into account both the specific character of basic information and the different levels of completeness and uncertainty associated with available data.

Some procedures have been presented in the literature to perform seismic hazard estimates by considering the time-varying quality of data (see, e.g., KIJKO and SOLLEVOLL, 1990; EGOZCUE and RUTTENER, 1997). A new attempt in this direction, specifically devoted to the analysis of macroseismic data and to a more extensive treatment of relevant uncertainties, has been performed in the frame of the National Group for Defence from Earthquakes (GNDT-CNR) and is outlined in the following. The methodology outlined below represents the integration in the frame of a unique coherent formulation of earlier preliminary attempts sparsely reported in the literature (MUCCIARELLI *et al.*, 1992, 1996; MAGRI *et al.*, 1994; ROTONDI *et al.*, 1994; GUIDOBONI and FERRARI, 1997; MARTINELLI and ALBARELLO, 1997).

The practical application of such a methodology is currently limited to the simplest case of the estimate of mean return period. Besides the case studies reported in the aforementioned papers, the methodology has been applied as:

- a comparison with standard probabilistic seismic hazard assessment (PSHA) techniques for 600 Italian towns (MUCCIARELLI *et al.*, 2000);
- a reference frame for deterministic scenarios for a large town (AZZARO *et al.*, 1999);
- a borrowing strength tool combined with horizontal vs. vertical spectral ratios (HVSR) measurement to identify site amplification (GALLIPOLI *et al.*, 2000).

In order to show performances of the proposed methodology and its capability, a sample application is shown concerning two sites, both located in Northern Italy: Bologna and Carpi. The results obtained in this way are then compared with analogous estimates obtained by the application of standard PSHA procedures.

2. Methodology

Hazard estimates are essentially an estimate of uncertainty in the form of the degree of belief (quantified in terms of a probability H) associated to statements such as: "During a future time span of Δt years (exposure time), the site under study will be shaken by at least one seismic event characterised by intensity at least equal to Io." This definition presumes an "epistemic" view of seismic hazard, which is considered as a measure of our worldwide knowledge (in particular about the future occurrence of earthquakes) and not a property of any stochastic process. Thus, the degree of belief represented by seismic hazard strictly depends on the reliability associated to available information pertinent to the seismic process, past seismicity in the area of interest and other relevant data. In this sense, seismic hazard estimates should be the result of a correct parameterisation of such uncertainties and of their coherent implementation in the hazard assessment procedure. The proposed approach has a phenomenological character, that is, it tries to obtain from the past history, information relative to future seismicity. Subsequently, uncertainties to be considered concern seismic effects of past earthquakes at the site (local seismic history), the reliability of the available catalogue of the site ("completeness") and the similarity of future seismogenic conditions during the exposure interval of interest with those which previously occurred. Each of these features will be characterised in terms of probability functions which express the respective degree of uncertainty. Their combination will result in the hazard estimate.

Uncertainty about Seismic Effects of Past Earthquakes

In general, reconstruction effects of older earthquakes is not an easy task and thus the assessment of local intensity for specific earthquakes can be affected by substantial uncertainty. In particular, this uncertainty could be expressed in terms of probability values $P_l(I_0)$ which represent the degree of belief in the hypothesis the l-th seismic event has actually shaken the considered site with effects of degree at least equal to I_0. A discussion regarding such formalisation is reported by MAGRI *et al.* (1994) and ROTONDI *et al.* (1994).

Values of $P_l(I_0)$, which fully represent the available seismic history for the site under study should be known for each documented event and constitute the keystone of the entire procedure. Expert judgements of historians could allow direct estimates of P_l. When direct documentation concerning felt effects is not available, P_l values can be estimated from epicentral data. In general, the following relationship holds

$$P_l(I_0) = \sum_{I=I_{\min}}^{I_{\max}} p_e(I)R(I_0|E) \ , \tag{1}$$

where p_e represents the probability that the maximum macroseismic intensity for the considered earthquake was equal to I. R represents the probability that the local effects are of intensity I_0 given that the earthquake responsible for local effects is characterised by a set of parameters E (maximum observed intensity, position of source with respect to the site, hypocentral depth, etc.). R plays the role of a probabilistic "attenuation" function.

The choice of R can be performed on the basis of theoretical considerations or from empirical estimates. On the basis of macroseismic data available in the Italian region (MONACHESI and STUCCHI, 1997), a possible simple form for R has been proposed by MAGRI *et al.* (1994) as a function of the maximum intensity I and of distance D from the site where effects of the considered earthquake were maximum. Since for most of the considered earthquakes very poor determination of hypocentral depths was available, this parameter has not been considered for the characterisation of R. Empirical distribution of conditional relative frequencies of felt intensities in the Italian region have been analysed to characterise the function R. These features developed well reproduced by the logistic form

$$R(I_0|D,I) = \frac{e^{a(A)+b(A)\ln D}}{1+e^{a(A)+b(A)\ln D}} \tag{2}$$

where a and b are empirical parameters and A is the attenuation ($A = I - I_0$). Estimates of a and b valid for the Italian area have been obtained by MAGRI *et al.* (1994) and result

$$\begin{array}{l} a(A) = 1.00 + 1.95A \\ b(A) = -1.15 - 0.16A \end{array} .$$

when the distance D is expressed in km. No regional differentiation has been attempted. Certainly, the choice of R in the simplified form (2) is debatable and depends on the specific statistical features of the data sample available for the region under study: different choices are possible. What is actually important and what differentiates the approach here proposed from the "standard" ones is the probabilistic character of the attenuation function R which allows for uncertainty of the actual presence of effects corresponding at least to the degree I_0 related with any specific earthquake which occurred at the time t_l.

By using this kind of information it is possible to estimate, for each specific time interval Δt_j, the probability Q_j that at least one of these events at the site has been actually felt with an intensity at least equal to I_0 during a specific time interval Δt_j of length Δt. If N_j is the total number of events which occurred during Δt_j it results that

$$Q_j = Q(\Delta t, I_0|\Delta t_j) = 1 - \prod_{l=1}^{N_j} [1 - P_l(I_0)] \tag{3}$$

(see MUCCIARELLI *et al.*, 1992). Furthermore, it is also possible to perform an estimate of the average seismicity rate n_j during the time interval Δt_j associated to events with local effects at least equal to I_0. This estimate can be obtained by using the relation

$$n_j = \frac{1}{\Delta t}\sum_{l=1}^{N_j} P_l(I) \tag{4}$$

(see MAGRI *et al.*, 1994).

Uncertainty about the Catalogue "Completeness"

In general, available information about past seismicity covers a time span much larger than the one characterised by a satisfactory level of reliability. Different opinions could exist about the interval ΔT_i to be considered as fully representative of actual seismicity. In general, the longer ΔT_i is, the better is our knowledge of the local seismicity pattern. However, to the same extent, the longer this interval is, the more unfavourable and incomplete is our knowledge about past seismic effects, and hazard estimates could be more misleading. This problem is known as "completeness" of seismic catalogues (see, e.g., MULARGIA and TINTI, 1987 for discussion).

A direct determination of completeness appears beyond our present possibilities due to the considerable information necessary for this purpose (e.g., evolution in time of socio/cultural environment, population density, survival of older documents, instrumental networks, etc.). A useful contribution to these studies can be supplied by statistical analyses of available catalogues. Since indications useful to address historical research and to perform preliminary seismic hazard estimates can be obtained by following such type of approaches, numerous efforts have been devoted in this direction (STEPP, 1971; CAPUTO and POSTPISCHL, 1974; BATH, 1983; TINTI and MULARGIA, 1985a,b; MULARGIA and TINTI, 1985; MULARGIA *et al.*, 1987; ROTONDI and PAGLIANO, 1993; ROTONDI *et al.*, 1994). In the present work a new "robust" statistical approach to the estimate of completeness has been adopted.

In the assumption that the seismic process is stationary, seismicity rates computed from (4) by taking into account a limited temporal portion of the catalogue can be considered as a sample estimate of the average number of earthquakes per time unit characteristic of the area under study. Due to the Central Limits theorem, such estimate of the population average is a random variate characterised by Gaussian probability distribution function. Symmetry of the Gauss distribution implies that, as n_i and n_j are two seismic rate estimates obtained from different time segments δt_i and δt_j of the catalogue, one has

$$q(n_i > n_j) = q(n_i < n_j) = 0.5, \tag{5}$$

where q indicates the relevant probability. This result is independent of specific features of the seismogenic process and can be used to assess the "completeness" of the available seismic catalogue for the interval ΔT_i. To this purpose, the time interval ΔT_i is subdivided in $2N$ elementary nonoverlapping subintervals of equal duration δt ($2N = \Delta T_i/\delta t$). For each elementary interval δt_i the empirical seismicity rate n_i is computed by (4) from seismic data in the catalogue. On the basis of these estimates, the binary random variate d is computed as

$$\begin{cases} d_i = 1 \leftrightarrow n_i > n_{i+N} \\ d_i = 0 \leftrightarrow n_i < n_{i+N} \end{cases},$$

where $i \in [1, N]$ and the cases $n_i = n_{i+N}$ are ignored. In practice, the variate d represents the comparison between seismicity rate estimates in the first part of the catalogue with the correspondent ones in the second part. In the assumption that "completeness" is monotonically not decrescent with time (see, e.g., MULARGIA and TINTI, 1985), a relatively large number of d values equal to 0 could suggest rarefaction of earthquakes receding in time due to decreased "completeness."

As N' ($\leq N$) is the number of effective comparisons of seismicity rates, m values of d result equal to 0. In the case that the catalogue is "complete," the probability F to obtain by chance a number $\geq m$ of d values equal to 0 as actually observed, can be computed by the binomial formula

$$F(\Delta T_i) = \sum_{j=m}^{N'} \frac{N'!}{j!(N'-j)!} q^j (1-q)^{N'-j} \tag{6}$$

and q is 0.5 by (5). Values of F near 0 imply a low reliability for the hypothesis of "completeness."

In general, a number L of possible choices exists for the actual extension of the time interval ΔT_{i} covered by the available catalogue to be considered for seismic hazard analyses. As an example, one can consider the time interval ΔT_{i} as comprised between t_f and t_i which respectively represent the most recent extreme and any starting point of the seismic catalogue to be considered. The completeness function F can be computed for each of the L possible choices of t_i in the range [t_0, t_{f}]with t_0 representative of the oldest possible starting point of the catalogue under study. For each possible choice of t_i a different value of the probability F can be computed. A probability function $r(\Delta T_i)$ which represents the catalogue completeness for the time span ΔT_i and also satisfies the conditions of mutual exclusivity and exhaustivity can be given in the form

$$r_i = r(\Delta T_i) = \frac{F(\Delta T_i)}{\sum_{i=1}^{L} F(\Delta T_i)}, \tag{7}$$

where the sum at the denominator is extended to the number L of possible choices for the starting point t_i of the seismic catalogue to be considered for hazard estimates. This position presumes that at least one of these possible choices actually allows individualization of a "complete" segment of the catalogue. The probability $r(\Delta T_i)$ represents the "completeness function" that is the degree of belief, expressed in terms of probability, in the hypothesis that for the specific time span ΔT_i the catalogue is actually complete.

The proposed statistical approach can only give a first-order rough estimate of completeness, useful when direct estimates of r based on historiographical considerations are not available. It in fact, presents a number of limitations related to the underlying assumptions. The most important one concerns the assumption that the most recent part of the catalogue considered on each turn, is in any case more "complete" with respect to the older one. This could not hold, for example, in the case of large "holes" in the available documentary sources due to long wars, etc.. Since the "completeness" analysis performed for hazard estimates covers relatively prolonged time spans, these "holes," in order to play a significant role, should span several decades. This possibility appears quite unrealistic for most recent times but cannot be ruled out when the most distant part is considered. Other approaches can probably be proposed to overcome such difficulties. Nonetheless, the most important methodological point is that the definition of a probabilistic function $r(\Delta T_i)$ is an integral part of the procedure for seismic hazard assessment.

Uncertainty about Future Seismogenic Conditions

The probabilities Q_j computed by (3) can be used for an empirical estimate of seismic hazard H at the site from available data concerning past seismic effects. In particular, interest is devoted to evaluate the probability $H(\Delta t, I_0)$ that within a future time interval of duration Δt at the site of interest at least one seismic event will occur which will be characterised by local effects corresponding at least to the degree I_0 of the adopted macroseismic scale.

Let us suppose that a seismic catalogue is available which reports seismic information about felt effects at the site of interest for a specific time span ΔT_i. A number K_i of distinct time intervals, each spanning Δt, can be individuated which cover (with partial overlapping allowed) the time interval ΔT_i. In order to compute this number, one can consider that for each possible instant t_j in the time interval ΔT_i comprised between t_i and $t_f - \Delta t$, it exists a time interval Δt_j beginning on t_j and characterised by a duration Δt. Thus, the value K_i can be simply computed as the ratio

$$K_i = \frac{t_f - \Delta t - t_i}{dt}, \tag{8}$$

where dt is the time unit.

One could imagine that in each specific interval Δt_j, particular conditions occurred in terms of seismotectonic situations and buildings vulnerability. In some cases these conditions resulted in the occurrence of at least one earthquake that produced at the site effects corresponding at least to the degree I_0 of interest. In order to forecast the future occurrence of seismic effects at least equal to I_0 during any future time interval of duration Δt, we should know if conditions in that future time interval will correspond to those which occurred during that specific past interval Δt_j. Due to our incomplete knowledge of present and past conditions, we can only define a probability $s_j = s(\Delta t_j)$ that represents our degree of belief in the hypothesis that during any future interval Δt the same conditions active during the specific past interval of duration Δt_j will occur again. In the assumption that in the past interval ΔT_i all the possible geodynamic conditions occurred, the probability density distribution s_j is such that

$$\sum_{j=1}^{K_i} s_j = 1, \tag{9}$$

where the sum is extended to the whole set of K_i possible choices of Δt_j in the interval ΔT_i. This implies that the possible conditions are mutually exclusive (one and only one possible condition will occur during the future time interval Δt) and that, at least one of the possible conditions which occurred during the past time interval ΔT_i, will occur in the future time interval Δt. This position is equivalent to assuming that the seismogenic process is stationary and that considered seismic history is sufficient to characterise (in a probabilistic sense) future seismicity. In this assumption, conditional seismic hazard H_i correspondent to intensity I_0 and to the exposure time Δt can be computed in the form

$$H_i = H(\Delta t, I_0 | \Delta T_i) = \sum_{j=1}^{K_i} s(\Delta t_j) Q(\Delta t, I_0 | \Delta t_j), \tag{10}$$

where the dependence of H from the choice of specific sub-catalogue considered for the time interval ΔT_i has been made explicit. This position presumes that the joint occurrences of at least an earthquake with intensity not lower than I_0 during Δt_j and of analogous geodynamic conditions during the future time span Δt constitute a set of mutually exclusive events.

From (10) one can see that in order to compute H_i, it is necessary to estimate the probability density distribution $s_j = s(\Delta T_i)$. This should be deduced on the basis of available knowledge of present and past conditions. If sufficient information is lacking relative to the seismogenic processes responsible for local seismicity, the probability density distribution s_j can be considered uniform. Thus, K_i being the number of possible past time intervals in the time span ΔT_i each characterised by specific seismotectonic conditions, form (7) and (8) one has

$$s_j = \frac{1}{K_i} = \frac{dt}{t_f - \Delta t - t_i}, \tag{11}$$

where t_i and t_f respectively represent the beginning and the end of the considered time interval ΔT_i.

In order to consider the problem of completeness, equation (10) can be generalised. To this purpose the density probability distribution $r = r(\Delta T_i)$ defined by (7) can be used which represents the reliability of the catalogue considered for the time interval ΔT_i. In the assumption that L represents the whole number of possible choices for a sub-catalogue to be considered for seismic hazard analysis, the unconditional hazard H can be computed by equation

$$H(\Delta t, I_0) = \sum_{i=1}^{L} r(\Delta T_i) H(\Delta t, I_0 | \Delta T_i). \tag{12}$$

This position implies that each value r_i is interpreted as the probability that the catalogue available for the time span ΔT_i represents the "best" possible data set for hazard estimates. In this case, the different possible choices of ΔT_i constitute a set of mutually exclusive events and the formula (12) results as formally correct.

3. *A Theoretical Example*

In order to show how the present approach actually works, it could be useful to consider a simple theoretical example. One hypothetical seismic catalogue is supplied which covers a time span of 200 years. Four events occurred during such time span at the 25th, 75th, 125th and 175th year. For all the events the intensity has been estimated to be VII with no uncertainty ($P(I_0) = 1$ for I_0 less or equal to VII and $P(I_0) = 0$ otherwise). We aim at the estimate of seismic hazard for a future time span Δt of 50 years. The catalogue is assumed to be complete with degree of belief $F(\Delta T_1) = 1$ for $\Delta T_1 = [76–200]$ and with $F(\Delta T_2) = 0.5$ if $\Delta T_2 = [1–200]$. In terms of r this implies that $r(\Delta T_1) = 1/1.5 = 0.67$ and $r(\Delta T_2) = 0.5/1.5 = 0.33$.

The first sub-catalogue ($\Delta T_1 = [76–200]$) is considered initially. In the time span considered (125 years) there exist 76 possible choices (K_i) of Δt_j (one for each possible starting time t_j after the 76th is included and such that the relative ending time does not overcome the 200th). In the particular case considered here, all these intervals were characterised by the occurrence of at least one event with intensity equal to VII. In the assumption that unknown conditions occurred during each past 50 years, subinterval Δt_j has the same probability of occuring in the future interval Δt, it results from (10) and (11) that

$$H_i = H(50, \mathrm{VII} | \Delta T_1) = \sum_{j=75}^{151} \frac{1}{76} = 1.$$

This result can be interpretated as follows. If at each year of the considered time span ΔT_1 we would have performed the forecast in the next 50 years of at least one earthquake with an intensity of at least VII, the "success" rate would have been 100%. This implies a very high confidence associate to the forecast performed at the present time.

If the entire time span is considered (200 years) $K_i = 151$ and we have

$$H_2 = H(50, \mathrm{VII} | \Delta T_2) = \sum_{j=1}^{150} \frac{1}{150} = 1.$$

By taking into account equation (12) and completeness function r given in equation (7), one has the unconditional hazard H given by

$$H(50, \mathrm{VII}) = 0.75 H_1 + 0.34 H_2 = 1.$$

For the same catalogue, seismic hazard in the assumption that earthquake occurrence is a Poisson process would be much lower: 0.632. The methodology proposed here appears to correctly capture the regularity shown by the seismic occurrences (every 50 years at least one event occurs). It is interesting to note that the poissonian probability would be the same if the events considered were grouped as follows: 5th, 10th, 190th and 195th. In this case, by following the methodology proposed here, we would have

$$H_1 = H(50, \mathrm{VII}|\Delta T_1) = \sum_{j=75}^{140} \frac{1}{76} 0 + \sum_{j=141}^{151} \frac{1}{76} 1 = \frac{10}{76} = 0.13,$$

$$H_2 = H(50, \mathrm{VII}|\Delta T_2) = \sum_{j=1}^{60} \frac{1}{151} 1 + \sum_{j=61}^{149} \frac{1}{151} 0 + \sum_{j=141}^{151} \frac{1}{151} 1 = \frac{70}{151} = 0.46,$$

$$H(50, \mathrm{VII}) = 0.75 H_1 + 0.34 H_2 = 0.25,$$

which is lower than the estimate obtained from the Poissonian model. These examples also illustrate that the proposed approach is considerably more sensitive to the actual time distribution of events then standard approaches.

4. An Applicative Example

As an example of application of the proposed methodology to real data, the probability has been computed that seismic effects corresponding to intensities ranging from VII to X MCS degrees will affect within a time interval of 20 years, two localities located in North-Central Italy, both characterised by a good historical documentation of past earthquakes: Bologna and Carpi (see Fig. 1). Data used to reconstruct local seismic history have been deduced from the catalogue of epicentral data by CAMASSI and STUCCHI (1996) and from the macroseismic database by MONACHESI and STUCCHI (1997) which spans a time ranging from 1000 to 1980. For each site a probabilistic seismic history for each intensity threshold I_0 has been obtained by considering such data. The probability vectors $P_l(I_0)$ correspondent to each earthquake documented at the site have been computed in the assumption that uncertain intensities imply an equal density of probability for each possible value. As an example, in the case that I_0 is given by MONACHESI and STUCCHI (1997) in the form VII–VIII, the vector P_l is assumed to be equal to 1 for $I_0 \leq$ VII, 0.5 for $I_0 =$ VIII and 0 in the remaining cases. For earthquakes not documented at the site, probabilities correspondent to expected intensities have been computed from epicentral data by using equations (1) and (2). Thus, from these data, different evaluations of the completeness function r_i and conditional hazard H_i have been computed by equations (6), (7), (10) and (11) for each possible choice of the time interval ΔT_i beginning on t_i and ending on 1980 (see in Fig. 2). From these values the unconditional seismic hazard has been estimated by using equation (12). The results obtained for expected intensities in the range VII–IX MCS at the towns of Bologna and Carpi are summarised in Figs. 3 and 4.

For the same sites, seismic hazard estimates also have been obtained by SLEJKO *et al.* (1998) by using the same data set and the standard approach implemented in the

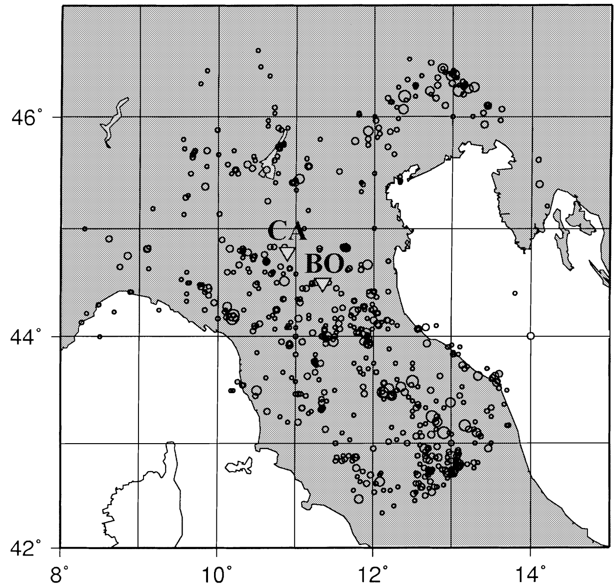

Figure 1

Epicenters of earthquakes considered in the seismic hazard analysis at Bologna and Carpi. Open circles indicate earthquakes epicentres and are proportional to the correspondent release of seismic energy. Triangles respectively indicate the sites of Bologna (BO) and Carpi (CA). In the figure, events occurred from 1000 to 1980, with epicentres located within 250 km from the two considered sites and with maximum intensity at least equal to VI degree MCS are reported (CAMASSI and STUCCHI, 1996).

computer code SEISRISK III (BENDER and PERKINS, 1987). These last results are also reported in Figs. 2 and 3 for comparison.

As concerns the Bologna site (Fig. 3), it arises that a probability slightly lower than 30% exists to observe, within 20 years, at least one seismic event with effects reaching VII MCS and that this probability reduces to 5% when effects reaching VIII MCS are considered. For the same site, estimates carried out by SLEJKO *et al.* (1998) result respectively equal to 14% and 0.3%.

As concerns the Carpi site (Fig. 4), hazard estimates deduced by the approach proposed here corresponding to expected effects reaching VII and VIII MCS, come out to be respectively 38% and 5%. For the same site, estimates carried out by using the standard methodology indicate for the same intensity thresholds the probabilities of 14% and 0.4%, respectively. These results show that, at the sites considered, standard methodologies tend to produce hazard values lower than those resulting from the approach proposed here. Furthermore, very similar estimates result for Bologna and Carpi on the basis of the standard approach. On the contrary, the approach proposed here suggests that seismic hazard is higher for the Carpi site.

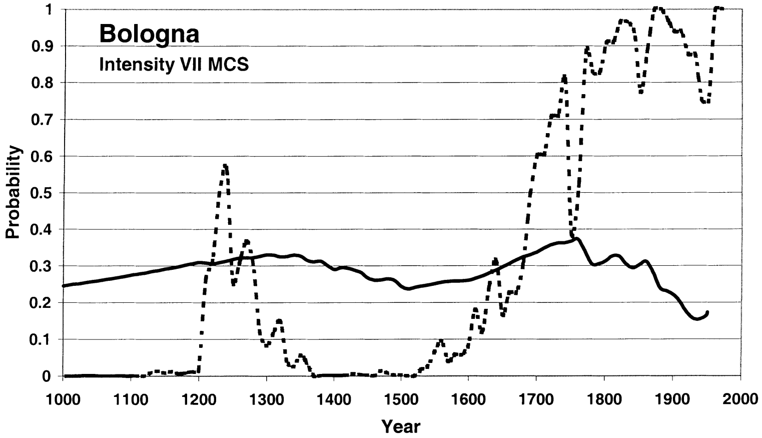

Figure 2
Dependence of completeness (dotted line) and conditional seismic hazard (solid line) on the starting time chosen for the considered seismic catalogue for the site of Bologna and for an intensity threshold of VII MCS. The termination for the catalogue considered for each computation is fixed at 1980. Completeness and conditional hazard have been obtained respectively by equation (6) and (10). These pieces of information are combined by equation (12) in order to obtain unconditional seismic hazard estimates (see text for details).

5. *Discussion and Conclusions*

A new methodology has been proposed for seismic hazard assessment from documentary data pertaining to past seismicity. It aims at the full and correct exploitation of data available concerning the effects of past earthquakes at the site of interest deduced from documentary data and parameterised in the form of macroseismic intensity whose discrete and ordinal character is explicitly taken into account. A basic aspect of the proposed methodology is the dominant role attributed to local seismic effects of past earthquakes with respect to epicentral data. This feature could prove to be very important for sites where significant local amplification of seismic effects is present due to peculiar morphological or geological conditions.

In the approach proposed here, hazard estimates directly result as a combination of uncertainties which affect relevant information. In particular, uncertainties regarding felt effects deduced from unexhaustive documentary data, intensity attenuation with distance, completeness of the available seismic history and future seismotectonic conditions are quantitatively provided for in the form of suitable probability functions (P, R, r and s, respectively). All these pieces of information can be deduced by statistical analysis of available data (as in the example considered here) or by expert judgement on the basis of historical or seismotectonic considerations.

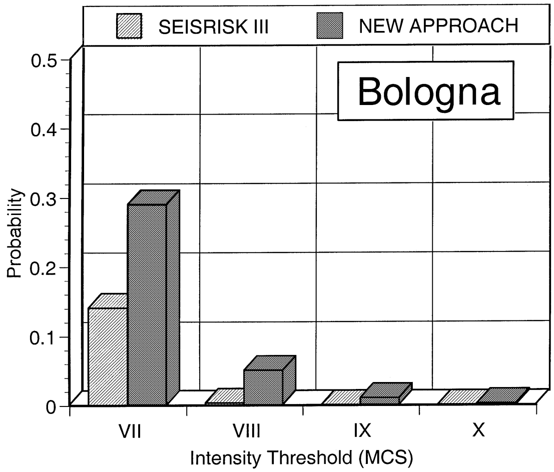

Figure 3
Seismic hazard estimates for the Bologna site as a function of the intensity threshold. Probability that each intensity threshold is reached within 20 years is reported. Data from the DOM 4.1 catalogue of seismic effects (Monachesi and Stucchi, 1996) and NT catalogue of epicentral data (Camassi and Stucchi, 1996) have been considered for the period 1000–1980. Hatched columns indicate results obtained by Slejko *et al.* (1998) using a standard approach implemented in the numerical code SEISRISK III (Bender and Perkins, 1987). Shaded columns indicate results obtained following the approach proposed here.

An important feature of the methodology described here is the attempt to reduce to the minimum, assumptions about statistical properties of seismicity. The basic assumption is that the seismogenic process is stationary, i.e., it can be described in terms of time independent statistical features. This hypothesis can be corroborated by quite general thermodynamic arguments (Lomnitz, 1974) and is the basis for all the attempts to forecast basic aspects of future seismicity on the basis of past seismic history. All other hypotheses regarding statistical properties of the local seismogenic process are released. As an example, the common assumption of the Poissonian character of seismicity is not required. This makes the present methodology "robust" against possible "anomalies" in the time distribution of seismic events. On the other hand, a number of probability functions used to parameterise important features (local intensities of past earthquakes, attenuation of effects with distance, completeness, etc.) have been obtained under relatively restrictive assumptions or specific statistical models. These choices can strongly affect the final hazard estimates, and future improvements could reduce the dependence of hazard estimates on non-fully convincing assumptions. Nonetheless, beyond any particular choice of these functions, the basic point of the paper is that all the pieces of information considered for seismic hazard estimates must be parameterised in probabilistic terms since the related uncertainties directly contribute to the final hazard estimate.

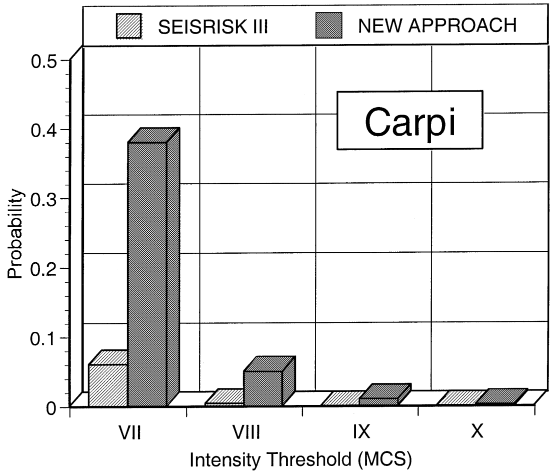

Figure 4
Seismic hazard for the Carpi site as a function of the intensity threshold. Probability that each intensity threshold is reached within 20 years is considered. Data from the DOM 4.1 catalogue of seismic effects (Monachesi and Stucchi, 1996) and NT catalogue of epicentral data (Camassi and Stucchi, 1996) have been considered for the period 1000–1980. Hatched columns indicate results obtained by Slejko *et al.* (1998) using a standard approach implemented in the numerical code SEISRISK III (Bender and Perkins, 1987). Shaded columns indicates results obtained following the approach proposed here.

The procedure proposed here has been applied to two sites in Northern Italy. The comparison with standard hazard evaluations shows that resulting estimates, though performed by using the same database, are quite different. The estimate of the statistical significance level associated to these differences has not been considered here. To this purpose, in fact, one must assume that a degree of uncertainty (statistically parameterised in some way) affects seismic hazard estimates. As a consequence, it would become possible to evaluate the hypothesis that the obtained differences are significant with respect to the level of relevant uncertainties. However, it is not easy to understand what such uncertainty means if attributed to seismic hazard which, on its turn, represents a degree of uncertainty about the occurrence of future earthquakes. Therefore, the "epistemic" view of seismic hazard adopted here does not allow attribution of any clear methodological meaning to estimates of statistical reliability relative to seismic hazard values (e.g., Musson, 2000).

It is worth noting that since Hazard estimates relative to Bologna and Carpi have been obtained on the same database covering the time interval 1000–1980, local seismic history encompassing the interval 1980–2000 can be used to "validate," in some sense, estimates of expected damage at the two sites. Resultingly since 1981 one earthquake has occurred in the area and shaken the Carpi site with intensity ranging VII–VIII MCS, while no effects greater or equal to VII MCS have been observed in Bologna. This result, of course, does not represent an effective validation of the proposed methodology against a "standard" approach but only represents a warning signal, suggesting that new approaches more sensitive to local data and relative uncertainty can be useful in avoiding possible misleading conclusions.

Acknowledgements

Many thanks are extending to S. Lasocki and to the other anonymous referee for their stimulating comments and fruitful suggestions. This work has been supported by the Gruppo Nazionale per la Difesa dai Terremoti.

REFERENCES

AZZARO, R., BARBANO, M. S., MORONI, A., MUCCIARELLI, M., and STUCCHI, M. (1999), *The Seismic History of Catania*, J. Seismol. *3*, 235–252.

BATH, M. (1983), *Earthquake Data Analysis: An Example from Sweden*, Earth Sci. Rev. *19*, 181–303.

BENDER, B. and PERKINS, D. M. (1987), *SeisriskIII: A Computer Program for Seismic Hazard Estimation*, USGS Survey Bulletin, 1772, 48 pp.

BOSCHI, E., GUIDOBONI, E., FERRARI, G., VALENSISE, G. and GASPERINI, P. (1997), *Catalogo dei forti terremoti in Italia dal 461a.C. al 1990*, 2nd edition (Istituto Nazionale di Geofisica – SGA, Roma), 644 pp.

CAMASSI, R. and STUCCHI, M. (1996), *NT4.1: un catalogo parametrico dei terremoti di area Italiana al di sopra della soglia di danno*, Internet Web site http:://emidius.itim.mi.cnr.it/GNDT.

CAPUTO, M. and POSTPISCHL, D. (1974), *Contour Mapping of Seismic Areas by Numerical Filtering and Geological Implications*, Ann. Geofis. *27*, 619–640.

GALLIPOLI, M. R., ALBARELLO, D., MUCCIARELLI, M., LAPENNA, E., CALVANO, G. (2000), *Correlation between Probabilistic Seismic Hazard Enhancement and in situ Amplification Measurements in the Val d'Agri Area, Italy*, submitted to Bull. Seism. Soc. Am.

EGOZCUE, J. J. and RUTTENER, E. (1997), *Bayesian Techniques for Seismic Hazard Assessment Using Imprecise Data*, Natural Hazards *14*, 91–112.

GUIDOBONI, E. and FERRARI, G. (1997), *Historical Cities and Earthquakes: Florence during the Last Nine Centuries and Evaluation of Seismic Hazard*, Ann. Geofis. *38*, 617–647.

KIJKO, A. and SOLLEVOLL, M. A. (1990), *Estimation of Earthquake Hazard Parameters from Incomplete and Uncertain Data Files*, Natural Hazards *3*, 1–13.

LEE, W. H. K. and BRILLINGER, D. R. (1979), *On Chinese Earthquake History – An Attempt to Model an Incomplete Data Set by Point process Analysis* Pure appl. geophys. *117*, 1229–1257.

LOMNITZ, C. (1974) *Global tectonics and earthquake risk*. In *Geotectonics*, 5 (Elsevier, Amsterdam 1974), 320 pp.

MAGRI, L., MUCCIARELLI, M., and ALBARELLO, D. (1994), *Estimates of Site Seismicity Rates Using Ill-Defined Macroseismic Data*, Pure appl. geophys. *143*(4), 617–632.

MARTINELLI, G. and ALBARELLO, D. (1997), *Main Constraints for Siting Monitoring Networks Devoted to the Study of Earthquake Related Hydrogeochemical Phenomena in Italy*, Ann. Geofis. *40*, 1505–1525.

MONACHESI, G. and STUCCHI, M. (1997), *D.O.M. 4.1 un database di osservazioni macrosismiche di terremoti italiani al di sopra della soglia di danno*, Internet Web site http:://emidius.itim.mi.cnr.it/GNDT.

MUCCIARELLI, M., MAGRI, L. and ALBARELLO, D. *For an adequate use of intensity data in site hazard estimates: mixing theoretical and observed intensities*. Proc. X World *Conference on Earthquake Engineering*, (Balkema, Rotterdam, 1992), pp. 345–350.

MUCCIARELLI, M., ALBARELLO, D., and STUCCHI, M., *Sensitivity of seismic hazard estimates to the use of historical site data*, In *Earthquake Hazard and Risk V. Schenk*, (ed.) (Kluwer 1996) pp. 141–151.

MUCCIARELLI, M., PERUZZA, L., and CAROLI, P. (2000), *Calibration of Seismic Hazard Estimates by Means of Observed Site Intensities*, *J. Earthq. Eng.* in press.

MULARGIA, F. and TINTI, S. (1985), *Seismic Sample Area Defined from Incomplete Catalogues: An Application to the Italian Territory*, Phys. Earth Planet. Int. *40*, 273–300.

MULARGIA, F., GASPERINI, P., and TINTI, S. (1987), *A Procedure to Identify Objectively Active Seismotectonic Structures* Boll. Geofis. Teor. Appl. *XXIX*, 114, 147–164.

MUSSON, R. M. W. (2000), *The Use of Monte-Carlo Simulations for Seismic Hazard Assessment in the U.K.*, Ann. Geofis. *43*(1), 1–9.

ROTONDI, R. and PAGLIANO, P. (1993), *The Problem of Completeness of a Seismic Catalogue: A Change-point Problem*, Proc. 2nd workshop of Statistical Models and Methods: Application on the prevention and Forecasting of earthquakes. Cephalonia, Greece, 170–182.

ROTONDI, R., MERONI, F., and ZONNO, G. (1994), *A Different Intensity Recording for Reducing the Uncertainty in its Assessment: An Application to the Completeness Analysis of Earthquake Catalogues.* Natural. Hazands. *10*, 45–58.

SLEJKO, D., PERUZZA, L., and REBEZ, A. (1998), *Seismic Hazard Maps of Italy*, Ann.Geofis. *41*(2), 183–214.

STEPP, J. C. (1971), *An Investigation of Earthquake Risk in the Puget Sound Area by the Use of the Type I Distribution of Largest Extreme*, Ph.D. Thesis, Pennsylvania State University, 131 pp.

TINTI, S. and MULARGIA, F. (1985a), *Completeness Analysis of a Seismic Catalogue*, Ann. Geophys. *3*, 407–414.

TINTI, S. and MULARGIA, F. (1985b), *An Improved Method for the Analysis of the Completeness of a Seismic Catalogue*, Lettere al Nuovo Cimento *42*, 21–27.

(Received March 28, 2000, accepted January 25, 2001)