SASHA: A Computer Program to Assess Seismic Hazard from Intensity Data

Vera D'Amico
Istituto Nazionale di Geofisica e Vulcanologia

Dario Albarello
University of Siena

INTRODUCTION

The evaluation of seismic hazard over wide territories is a basic tool for planning activities aimed at earthquake damage mitigation. This is commonly performed through probabilistic approaches based on the statistical analysis of past seismicity. Among these, due to its wide application worldwide, the Cornell-McGuire approach (Cornell 1968; McGuire 1978) has become a kind of "standard" methodology for probabilistic seismic hazard assessment (PSHA). In Italy, several national seismic hazard maps were produced in recent years (Slejko *et al.* 1998; Albarello *et al.* 2000; MPS Working Group 2004) by following this procedure as implemented by Bender and Perkins (1987). Yet despite its widespread application, this standard methodology presents severe drawbacks due to its strong sensitivity to some ill-defined aspects, such as geometry of seismic sources, attenuation of ground motion with distance from the source, completeness of available seismic catalogs, etc. Moreover, this kind of approach does not allow the full exploitation of a large amount of documentary data available at the site about the seismic effects of past earthquakes (Albarello and Mucciarelli 2003). Another drawback is that the standard approach was developed with the assumption that the seismicity database used to feed the computational model is constituted by instrumental data (magnitude, epicentral locations, etc.). However, in many countries (first and foremost, Italy) the bulk of the seismic database is constituted by macroseismic data, and thus the application of the standard method requires a "forcing" of macroseismic information into a para-instrumental format. But macroseismic information is not isomorphic to instrumental data since intensity values are discrete, ordinal, and range-limited. This implies that, in principle, mathematical formalizations suitable to instrumental information cannot be used to manage macroseismic data (see, *e.g.*, Pasolini *et al.* 2008a, 2008b).

To overcome some of these difficulties and to better exploit available information, probabilistic hazard evaluations based on observed intensity data were performed in Europe (Monachesi *et al.* 1994; Papoulia and Slejko 1997; Azzaro *et al.* 1999; Albarello *et al.* 2002) and Japan (Bozkurt *et al.* 2007) using alternative numerical procedures. An apparent limitation of these studies is the fact that PSH estimates are provided in terms of intensity, and this conflicts with the fact that ground acceleration still remains the traditional output of PSHA devoted to seismic design. However, a new interest has recently grown around macroseismic intensity. In fact, when damage scenarios and post-earthquake emergency planning are of concern, hazard assessment in terms of intensity as ground-shaking measure may be more suitable than conventional estimates based on instrumental parameters (PGA, etc.). A further possible advantage of these kinds of approaches is that they provide hazard evaluations completely independent from the standard ones and more directly linked to empirical observations (local seismic history). Thus, they could represent a useful benchmark for a direct assessment of reliability of traditional PSH estimates (Mucciarelli *et al.* 2000, 2006, 2008; Bozkurt *et al.* 2007).

In this paper we present the computer program SASHA (Site Approach to Seismic Hazard Assessment), which implements the intensity-based PSHA procedure originally proposed by Magri *et al.* (1994) and then improved by Albarello and Mucciarelli (2002). It relies on the analysis of the site seismic history, *i.e.*, the dataset of seismic effects (macroseismic intensities) documented during past earthquakes at a given locality. This methodology (hereafter, site approach) has been specifically developed to handle macroseismic data, and thus both the peculiar nature of intensity values (discrete, ordinal, range-limited) and relevant uncertainty (ill-defined intensity values, completeness of site seismic history, etc.) are taken into account by a coherent statistical approach that does not require any assumption about earthquake recurrence model and seismic source geometry. Furthermore, no aftershock removal is required in advance and epicentral data are only considered to integrate (when necessary) felt data at the site. Several PSHA studies have been performed in the last decade in Italy using different versions of the site approach (Mucciarelli *et al.* 2000; Albarello *et al.* 2002; D'Amico and Albarello 2003; Albarello, Azzaro *et al.* 2007; Azzaro *et al.* 2008).

SASHA's theoretical background is briefly outlined in the next section of the paper. Then, we describe the most important features of SASHA along with a sample application to the Italian area.

doi: 10.1785/gssrl.79.5.663

THEORETICAL BACKGROUND

In the site approach (Albarello and Mucciarelli 2002), the assessment of seismic hazard (*i.e.*, probability that during a future time span of length Δt the site under study will be shaken by at least one earthquake with local effects $\geq I_s$) is performed in three steps: 1) building the local history of seismic effects (site seismic history) by considering uncertainty affecting available intensity data, 2) evaluating its overall reliability (completeness, representativeness), and 3) computing the seismic hazard at the site.

Building the Site Seismic History

The reconstruction of the effects produced by past earthquakes at a given locality is characterized by uncertainty. A way to express this uncertainty is through the exceedance probabilities $P_l(I_s)$, which represent the degree of belief in the hypothesis that the l-th seismic event has actually shaken the considered site with an intensity at least equal to I_s (Magri *et al.* 1994). The values of $P_l(I_s)$ are defined over the whole range of the adopted macroseismic scale (*e.g.*, I–XII for the Mercalli-Cancani-Sieberg [MCS] scale). In principle, these probabilities should be attributed after a critical analysis of documentary information about intensity data available at the site.

In the case of earthquakes lacking direct information about local effects, probability $P_l(I_s)$ can be derived from the relevant epicentral data by using the relation

$$P_l\left(I_s\right)=\sum_{I_e=1}^{12} g\left(I_e\right) S\left(I_s \middle| I_e, D\right) \tag{1}$$

where g is the probability that epicentral intensity is exactly I_e and S is the conditional probability that local seismic effects at least correspond to intensity I_s, given a causative earthquake having an epicentral intensity I_e and occurring at the distance D. Thus, the probability function P_l and the probability density g allow us to account for uncertainty on the intensity value at the site and at the epicenter, respectively. The probability function S plays the role of the attenuation function expressed in probabilistic form (Albarello and D'Amico 2005) and represents the uncertainty about the intensity value deduced at the site given epicentral intensity and source distance.

Regularity of the macroseismic field can then be used to better constrain these "virtual" intensity values (*i.e.*, not actually observed at the site but derived from epicentral information). This is achieved by correcting estimates provided through Equation 1 by taking into account intensity observations available for the same earthquake at neighboring localities. This "correction" is performed via the Bayes's equation

$$p_l^*\left(I_s \middle| I_v\right)= p_l\left(I_s\right)\frac{q\left(I_v \middle| I_s\right)}{\sum_{I=1}^{12} p_l\left(I\right) q\left(I_v \middle| I\right)} \tag{2}$$

where p_l is the probability density associated to the probability function defined in Equation 1, *i.e.*,

$$p_l\left(I_s\right)=P_l\left(I_s\right)-P_l\left(I_s+1\right), \tag{3}$$

and $q\left(I_v \middle| I_s\right)$ represents the probability that, for a given site intensity I_s, intensity at the closest locality is I_v. Lastly, the probability vector $P_l\left(I_s \middle| I_v\right)$ to be used for hazard computations is derived from Equation 2 as follows:

$$P_l\left(I_s \middle| I_v\right)=\sum_{I=I_s}^{12} p_l^*\left(I \middle| I_v\right). \tag{4}$$

It is worth noting that the adopted formalization allows us to correctly manage relevant uncertainty (about information available at the site, on attenuation of energy from the source, and on regularity of the macroseismic field), accounting for the discrete and ordinal character of intensity. One should also be aware that the probability function S in Equation 1 and the probability density q in Equation 2 have to be determined empirically and are presumably region-dependent.

The set of P_l probability functions, each corresponding to a different earthquake (either obtained from direct site observation or derived from epicentral and neighboring information), represents the seismic history at the site under study, *i.e.*, the local seismic catalog.

In the probabilistic approach considered here, these probabilities can be used, as an example, to compute the expected number $\nu(I_s)$ of earthquakes that shook the site with $I \geq I_s$ during a given time span, that is

$$\nu\left(I_s\right)=\sum_{l=1}^{N} P_l(I_s), \tag{5}$$

where the summation is extended to all the N earthquakes that occurred during that period (Magri *et al.* 1994).

Overall Reliability of the Site Seismic History

Available information on past seismicity generally covers a time span much larger than that characterized by a satisfactory level of completeness. Usually, the longer this period, the more uncertain and incomplete our knowledge about past seismic effects. However, catalogs covering longer periods of time are more representative of the seismogenic process. Therefore, both these aspects (completeness and representativeness) should be considered to evaluate the reliability of hazard estimates deduced from the site seismic history covering the specific time interval ΔT_i. In terms of site approach, the completeness C of the local catalog used for hazard computation is expressed in terms of a probability $L_i(C)$.

The completeness of site seismic history is assessed through a statistical methodology (Albarello *et al.* 2001) based on three assumptions: 1) the seismogenic process is stationary, *i.e.*, it can be described in terms of time-independent statistical features; 2) the most recent part of the catalog is complete; and 3) the catalog is statistically representative of the long-term stationary seismogenic process (representativeness). In particular, the Cox and Stuart distribution-free procedure for stationarity assessment (*e.g.*, Rock 1988) is applied. The probability L_i is computed that the differences observed between apparent seismicity rates in the first and in the second half of time interval ΔT_i covered by the catalog can exist in the assumption of stationarity. If this probability is high, the relevant catalog is considered "complete" (for more details, see Albarello *et al.* 2001). One can see that the hypotheses of completeness and representativeness are not independent. This statistical procedure allows us to determine the conditional probability $L_i(C|R)$, which represents the degree of belief in the hypothesis that the catalog is complete (C) given it is representative (R).

The unconditional probability relative to the hypothesis that the catalog covering ΔT_i is complete for hazard computation is

$$L_i(C) = prob_i(R) L_i(C|R). \tag{6}$$

Concerning $prob_i(R)$, in the lack of information about long-term statistical properties of the seismogenic process, we *a priori* assume that it linearly increases with time extension of the site catalog. Equation 6 is thus rewritten as

$$L_i(C) = \frac{\Delta T_i}{\Delta T_{\max}} L_i(C|R) \tag{7}$$

with $\Delta T_{\max}$ corresponding to the longest time span covered by the available catalog.

In general, the estimate of hazard can be performed for M possible choices of the time span ΔT_i, included between t_f and t_i (respectively, most recent extreme and any starting point of the catalog explored), covered by the local seismic history. The probability $L_i(C)$ that the catalog is complete can be thus computed for each of the M periods ΔT_i that result from progressively displacing t_i back in time in the range $[t_0, t_f]$, with t_0 corresponding to the oldest extreme of the catalog under study.

Finally, to satisfy conditions of mutual exclusivity and exhaustivity, the "completeness function" $L_i^*(C)$ for the time interval ΔT_i results in

$$L_i^*(C) = \frac{L_i(C)}{\sum_{i=1}^{M} L_i(C)} \tag{8}$$

(the sum is extended to the M possible choices of the catalog starting at year t_i).

Computing the Seismic Hazard

Starting with the local seismic catalog, for each time interval Δt_j of length equal to Δt (exposure time), the probability Q_j can be assessed that at least one of the considered earthquakes has actually shaken the site with intensity $\geq I_s$. If N_j is the total number of events that occurred during Δt_j, it results in

$$Q_j = Q\left(\Delta t, I_s \middle| \Delta t_j\right) = 1 - \prod_{l=1}^{N_j} \left[1 - P_l(I_s)\right]. \tag{9}$$

If the available site catalog covers a time span ΔT_i, a number K_i of distinct time intervals (partially overlapping) can be individuated that cover the period ΔT_i. In particular,

$$K_i = \frac{t_f - \Delta t - t_i + 2}{dt}, \tag{10}$$

where t_i and t_f correspond, respectively, to starting and ending years of the relevant catalog and dt is the time unit.

The seismic hazard H_i for intensity threshold I_s and exposure time Δt can be computed in the form

$$H_i = H(\Delta t, I_s | \Delta T_i) = \sum_{j=1}^{K_i} s(\Delta t_j) Q(\Delta t, I_s | \Delta t_j), \tag{11}$$

where dependence of H on the specific subcatalog of length ΔT_i (among the M possible choices) is made explicit.

In Equation 11, $s(\Delta t_j)$ represents the probability that during any future time span Δt, the same seismotectonic conditions that occurred during the specific past interval Δt_j will recur. In absence of reliable information about the seismogenic process responsible for local seismicity, the probability function $s(\Delta t_j)$ can be assumed uniform, that is, $s(\Delta t_j) = 1/K_i$. Otherwise, the values of $s(\Delta t_j)$ can be freely attributed on the basis of available knowledge about the seismotectonic pattern with the constraint that they sum up to unity.

Equation 11 is finally generalized to take into account the completeness of the site seismic catalog ($L_i^*(C)$ defined in Equation 8 for each of the M possible subcatalogs) and the unconditional seismic hazard results in

$$H\left(\Delta t, I_s\right) = \sum_{i=1}^{M} L_i^*(C) H\left(\Delta t, I_s \middle| \Delta T_i\right). \tag{12}$$

Resulting values of H, computed for each intensity degree I_s (*e.g.*, I–XII for the MCS scale), represent the hazard curve at the site under study. A "reference intensity" (I_{ref}) is then derived from this curve, corresponding to the lowest intensity that satisfies the relation

$$\int_{I_{ref}+1}^{12} H\left(\Delta t, I_s\right) dI_s < P_{exc}, \tag{13}$$

where P_{exc} is a fixed exceedance probability level (*e.g.*, 10%).

Likewise, a "reference PGA" (PGA_{ref}) value is also assessed corresponding to the highest PGA (peak ground acceleration in g) characterized by an exceedance probability $H(\Delta t, \mathrm{PGA})$ not lower than P_{exc} during the exposure time Δt. To this purpose, one has

$$H(\Delta t,\mathrm{PGA}) = \sum_{I_s=1}^{12} h(\Delta t, I_s)\, G\big(\log(\mathrm{PGA}) \,|\, I_s\big) \quad (14)$$

where

$$h(\Delta t, I_s) = H(\Delta t, I_s) - H(\Delta t, I_s + 1), \quad (15)$$

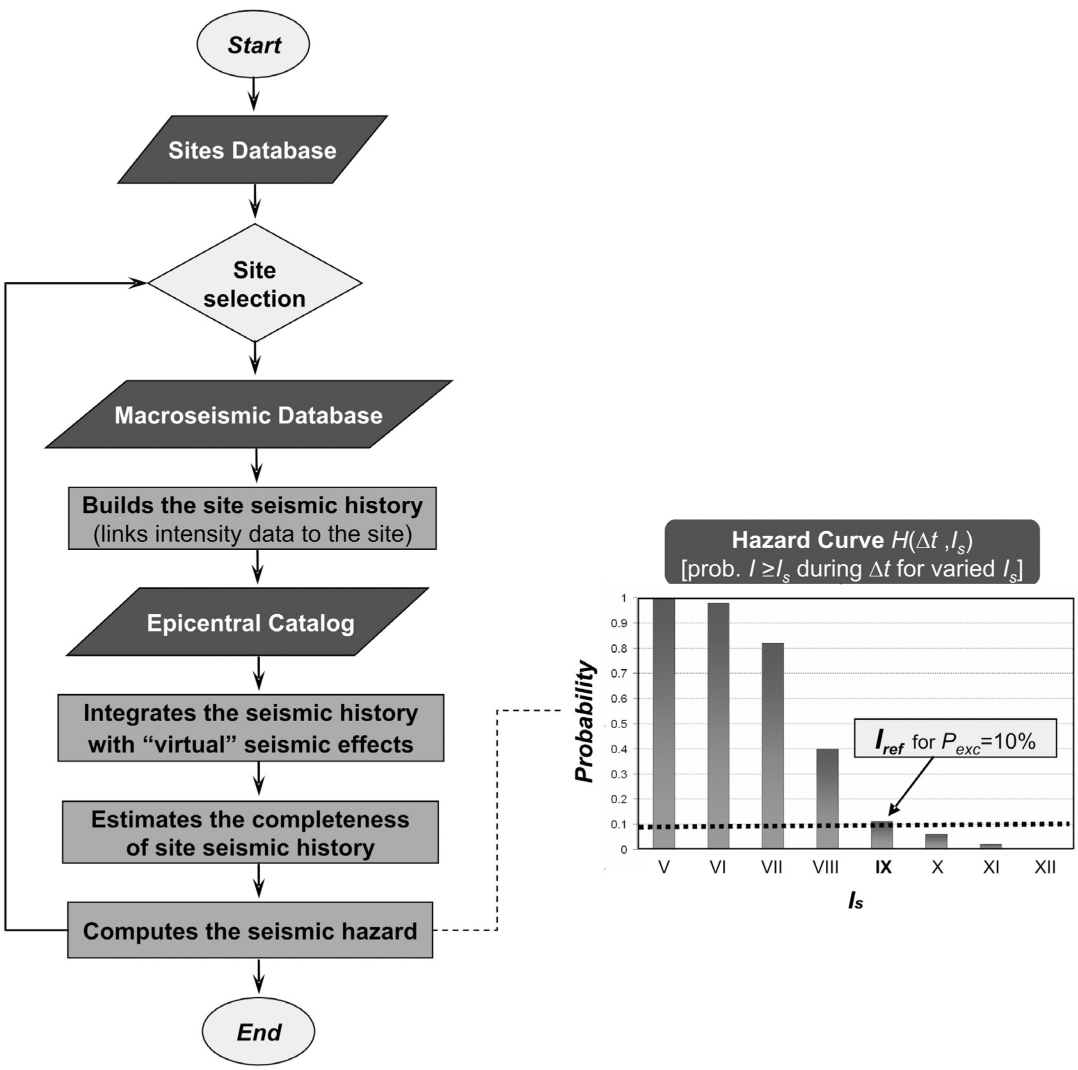
and G represents the probability that the logarithm of peak ground acceleration at the site is not less than PGA given that intensity is I_s.

THE SASHA PROGRAM

The SASHA program allows a number of choices concerning the reconstruction of site seismic history (criteria for data selection, intensity attenuation with distance, etc.) and the hazard computation (exposure time, exceedance probability, etc.). A flow chart displaying the working process of the program is shown in Figure 1.

The complete input consists of three files: 1) the list of sites where hazard has to be computed (sites database), 2) the set of available intensity observations (macroseismic database), and 3) the catalog with epicentral information relative to earthquakes potentially felt at the sites under study (epicentral catalog). A numerical code links each macroseismic observation to one earthquake in the epicentral catalog. Hazard can be computed automatically at a set of sites but each estimate is carried out independently. The location where hazard has to be assessed can be identified in geographic coordinates alone or by an additional numerical code (*e.g.*, the ZIP Code) that unequivocally identifies the site. In the latter case, the same coordinates and code should be reported in the macroseismic database for each site in order to correctly attribute the relevant observed intensities.

The first step of the program is to reconstruct the seismic history at each site considered. To this purpose, the operator can choose if the local catalog is composed only by documented intensities (listed in the macroseismic database) or if this has to be integrated by epicentral data (from the epicentral catalog) and neighboring macroseismic information (from the macro-



▲ **Figure 1.** Conceptual flow chart of the SASHA program. Dark gray boxes display the input/output files.

seismic database). The first possibility is useful when exhaustive documentation exists about a large number of intensity observations relative to past earthquakes. The association to the site of each intensity value in the database can be done through the numerical code of the site and/or its coordinates. In both cases, one has to fix a search distance (*e.g.*, 1–2 km) in order to envisage possible small coordinates differences. When multiple felt data exist for one earthquake within this distance, one can choose to consider either the nearest intensity observation to the site or the highest. Though the adopted formalization also allows other more general parameterizations, intensity values reported in the macroseismic databases (*e.g.*, Stucchi *et al.* 2007) are commonly classified in two categories: "certain" data (one single value of intensity I_{loc}', *e.g.*, VI) and "uncertain" data (two possible contiguous values of intensity $I_{loc}' - I_{loc}''$ *e.g.*, VI–VII). The probability vector $P_l(I_s)$ in Equation 1 related to each earthquake is defined as follows (Albarello and D'Amico 2005)

"certain" datum: $P_l(I_s) = 1$ for $I_s \leq I_{loc}'$ and $P_l(I_s) = 0$ for $I_s > I_{loc}'$

"uncertain" datum: $P_l(I_s) = 1$ for $I_s \leq I_{loc}'$, $P_l(I_s) = 0.5$ for $I_s = I_{loc}''$ and $P_l(I_s) = 0$ for $I_s > I_{loc}''$

(the latter position means that an equal probability is assigned to the hypotheses $I_s = I_{loc}'$ and $I_s = I_{loc}''$).

If epicentral data are also taken into account, one has to fix the maximum epicentral distance (*e.g.*, 200 km) to be considered. In the present version of the program, the probabilistic attenuation function S in Equation 1 is assumed in the form

$$S\left(I_s \middle| I_e, D\right) = 1 - \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{I_s - 0.5} e^{-\frac{\left[x - \mu\left(I_e, D\right)\right]^2}{2\sigma^2}} dx \tag{16}$$

(*e.g.*, Albarello and D'Amico 2005) derived by the Gaussian cumulative distribution. Relevant average μ and standard deviation σ are those recently derived for the Italian region (Pasolini *et al.* 2008b):

$$\mu(I_E, D) = I_E - 0.0086(D - h) - 1.037[\ln(D) - \ln(h)] \tag{17}$$

where I_E is the "average expected intensity at the epicenter" and $D = \sqrt{R^2 + h^2}$, with R being the epicentral distance in km and h = 3.91 km (for details see Pasolini *et al.* 2008a, 2008b). The standard deviation σ associated to Equation 17 and to be used in Equation 16 varies depending on the method followed for computing I_E: σ = 0.690 for earthquakes with I_E directly assessed from the relevant macroseismic field; σ = 0.87 or 0.98 for events with I_E derived through empirical relations from the relevant magnitude or epicentral intensity, respectively (Pasolini *et al.* 2008b). This parameterization can be changed by the user. Otherwise, the virtual seismic effects can be read from an external file listing the $P_l(I_s)$ values in Equation 1 computed for different bins of epicentral intensity and distance. The program also allows users to "combine" two different attenuation relations. This option could be of particular use when hazard is computed over a wide territory including regions with specific attenuation properties, such as volcanic districts (*e.g.*, Azzaro *et al.* 2008). Currently in the code, a log-linear attenuation model is provided for the regional relationship with relevant parameters to be entered by the user. However, this model and parameterization can be modified in the program.

In order to take into account intensity data eventually available at neighboring sites, the probability $q\left(I_v \middle| I_s\right)$ in Equation 2 has to be defined. In the present version of SASHA this has been derived from the Italian DBMI04 intensity database (Stucchi *et al.* 2007; http://emidius.mi.ingv.it/DBMI04) and is entered in a tabular form (Table 1) for individual values of $\Delta I = I_v - I_s$ (for details see Albarello, D'Amico *et al.* 2007). If I_v is uncertain between two contiguous intensity values I_v' and I_v'', Equation 2 becomes

$$p_l^*\left(I_s \middle| I_v\right) = 0.5\, p_l^*\left(I_s \middle| I_v'\right) + 0.5\, p_l^*\left(I_s \middle| I_v''\right). \tag{18}$$

In this case, too, the numerical values for the probability $q\left(I_v \middle| I_s\right)$ can be modified to account for a different situation (different database or study area).

After the elaboration of site seismic histories, SASHA produces the output in the form of a set of hazard curves, each representing the probability that effects greater or equal to a given intensity threshold will be observed at the site during an exposure time fixed by the operator. A reference intensity corresponding to a fixed value of hazard (*e.g.*, 10%) is then derived from this curve (see Equation 13 and Figure 1).

TABLE 1
Values of $q(I_v \mid I_s)$ expressed as probability that site intensity I_s differs by ΔI from the intensity I_v observed at the closest (within 20 km) locality (from Albarello, D'Amico *et al.* 2007).

$\Delta I = I_v - I_s$	$q(\Delta I)$
<–5	0.00000
–5	0.00001
–4	0.00053
–3	0.00396
–2	0.02823
–1	0.17920
0	0.55575
1	0.19115
2	0.03493
3	0.00539
4	0.00082
5	0.00002
>5	0.00000

Lastly, the probabilistic procedure in Equation 14 to "convert" hazard estimates from intensity to PGA and provide a reference value for the latter is applied. The relevant probability function G is assumed in the form

$$G\left(\log(\mathrm{PGA})\middle|I_s\right)=\frac{1}{\sigma\sqrt{2\pi}}\int_{\log(\mathrm{PGA})}^{+\infty} e^{-\frac{\left[x-\mu(I_s)\right]^2}{2\sigma^2}}\,dx, \qquad (19)$$

i.e., the Gaussian cumulative distribution, with average μ and standard deviation σ to be determined empirically through a specific relationship between log(PGA) and intensity I_s. In particular, the program currently implements the following relation derived from Italian data (Faccioli and Cauzzi 2006):

$$\mu(I_s)=\log(\mathrm{PGA})=-1.33+0.20I_s, \qquad (20)$$

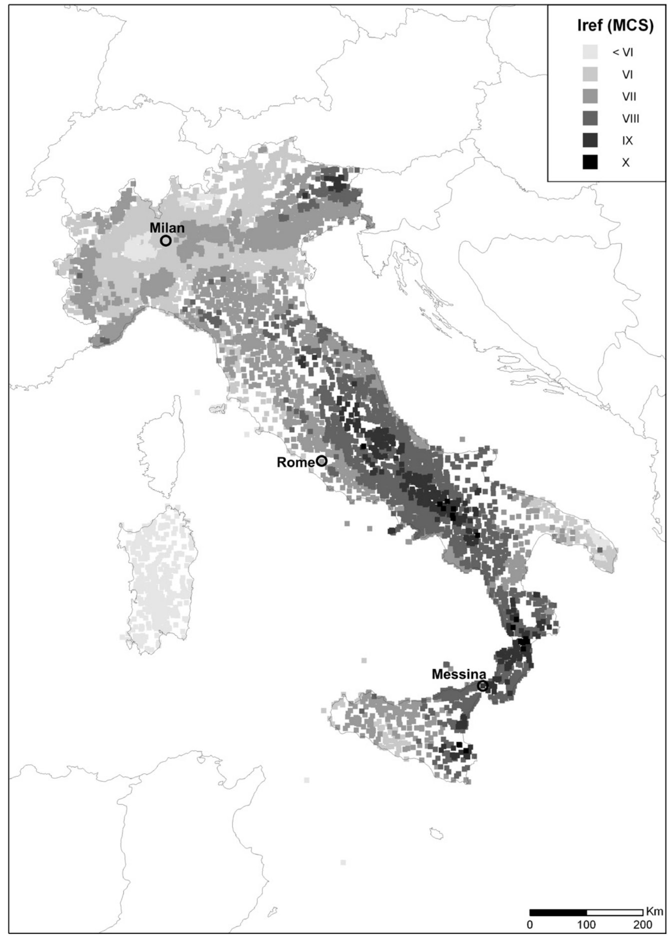
with $\mu(I_s)$ that represents the expectation of log(PGA) and $\sigma = 0.29$ (here PGA is in m/s^2). Also in this case, the empirical parameters of Equation 20 can be modified to account for the local situation.

APPLICATION TO PSHA IN ITALY

To provide an example of SASHA's performance, we assessed the seismic hazard over the Italian peninsula. Historical research carried out during the past 30 years led to about 58,000 macroseismic estimates of felt effects since 217 B.C. (DBMI04 database by Stucchi *et al.* 2007; http://emidius.mi.ingv.it/DBMI04). This large database constitutes the bulk of information used to build the CPTI04 epicentral catalog (CPTI Working Group 2004; http://emidius.mi.ingv.it/CPTI04): data on about 70% of destructive earthquakes derive in fact from documentary sources only.

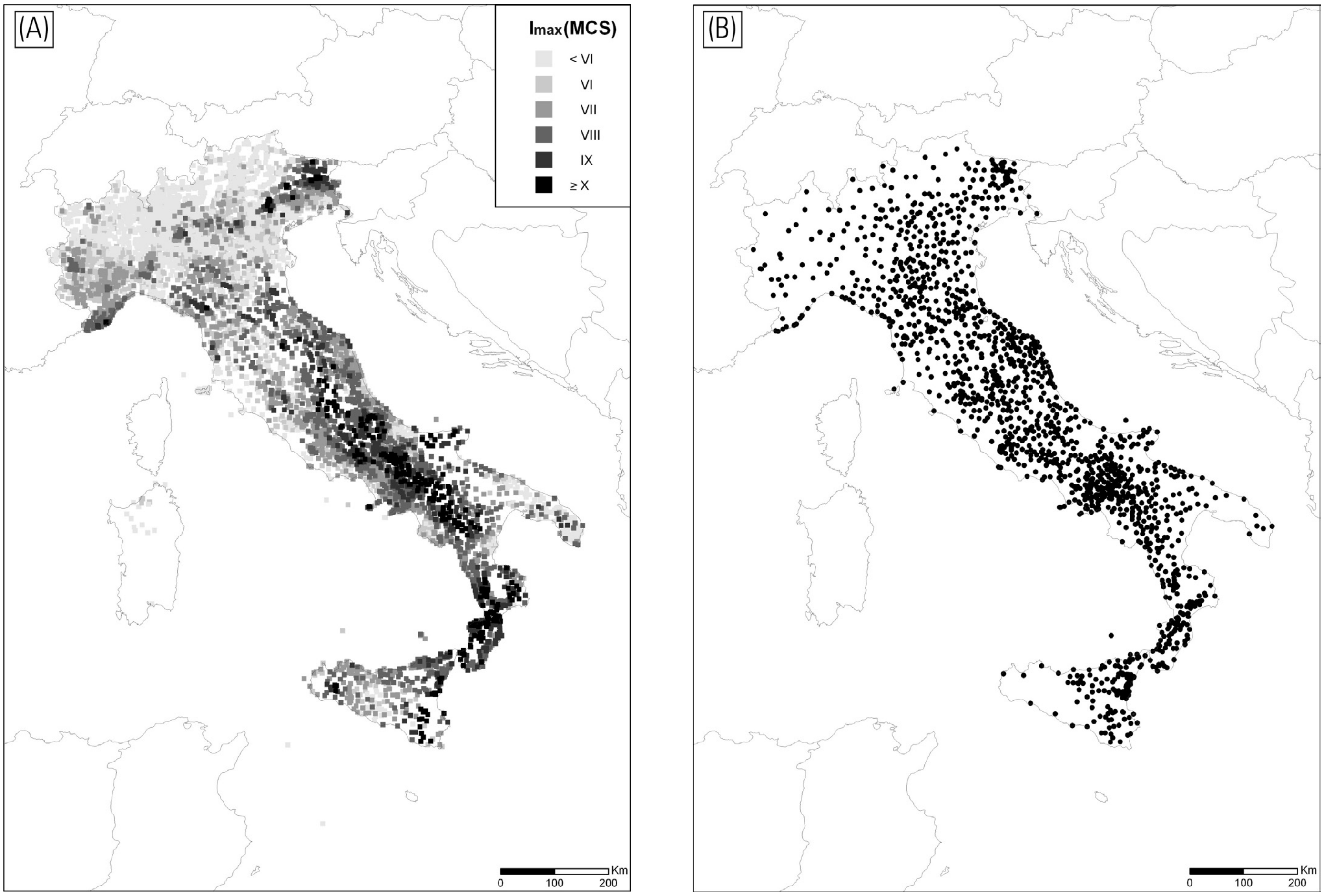
Computation was performed for the 8,101 municipalities of Italy. The seismic history of each locality (here the reference locality of the relevant municipality) was built from intensity data listed in the DBMI04 database. In particular, for each earthquake we considered the nearest macroseismic datum among those within a distance of 2 km from the investigated site and inside the same municipal territory. For those events lacking local felt data, virtual seismic effects were deduced from relevant epicentral data (from the CPTI04 catalog) and by taking into account possible macroseismic information available at neighboring (within 20 km of the site) localities (see Equations 1–4 and 16–18).

The results, in terms of highest intensity values characterized by an exceedance probability not less than 10% in 50 years (I_{ref}), are shown in Figure 2. Except for a few areas, computed I_{ref} values exceed intensity VI MCS reaching degree IX in the northeast and up to X along the axial part of the peninsula and to the south. This pattern is consistent with that in Figure 3A, which shows the maximum intensity values documented during the past in Italy (Stucchi *et al.* 2007). In fact, as expected from the adopted PSHA methodology, which basically relies on observed intensity data, more hazardous areas correspond to those characterized by stronger past seismicity. Concerning low-hazard zones (*e.g.*, northern Italy), poor seismic histories are available (Figure 3B), and this makes the relevant values of hazard more strongly dependent on the attenuation relationship used to derive virtual local effects rather than on the few felt data at our disposal.

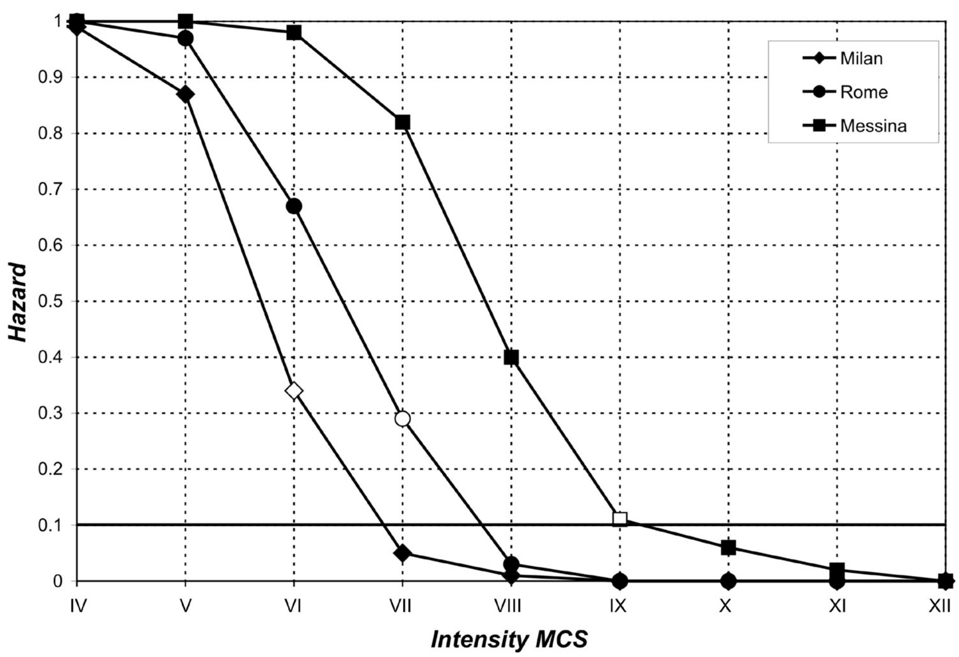


▲ **Figure 2.** I_{ref} estimates (highest intensity values with exceedance probability not less than 10% in 50 years) obtained at the 8,101 municipalities of Italy. Epicentral and macroseismic data are taken from the CPTI04 and DBMI04 databases respectively.

A more complete representation of the hazard level at a given site is provided by the hazard curves. Figure 4 displays the curves computed for the three cities of Milan, Rome, and Messina for an exposure time of 50 years. These cities—all with very rich and long seismic histories (> 80 felt data)—are located in areas characterized by very different levels of hazard (respectively, low, moderate, and high, as shown in Figure 2). As clearly emerges from Figure 4, hazard significantly increases as one moves from Milan to Messina (I_{ref} = VI, VII and IX, respectively). This increasing trend is also confirmed when hazard estimates are converted from intensity to PGA (see Equations 14–15 and 19–20), with the reference PGA value rising from 0.143 g at Milan to 0.204 g at Rome and up to 0.424 g at Messina. It is worth noting that these values are much higher than those (0.052, 0.120, and 0.248 g, respectively) computed



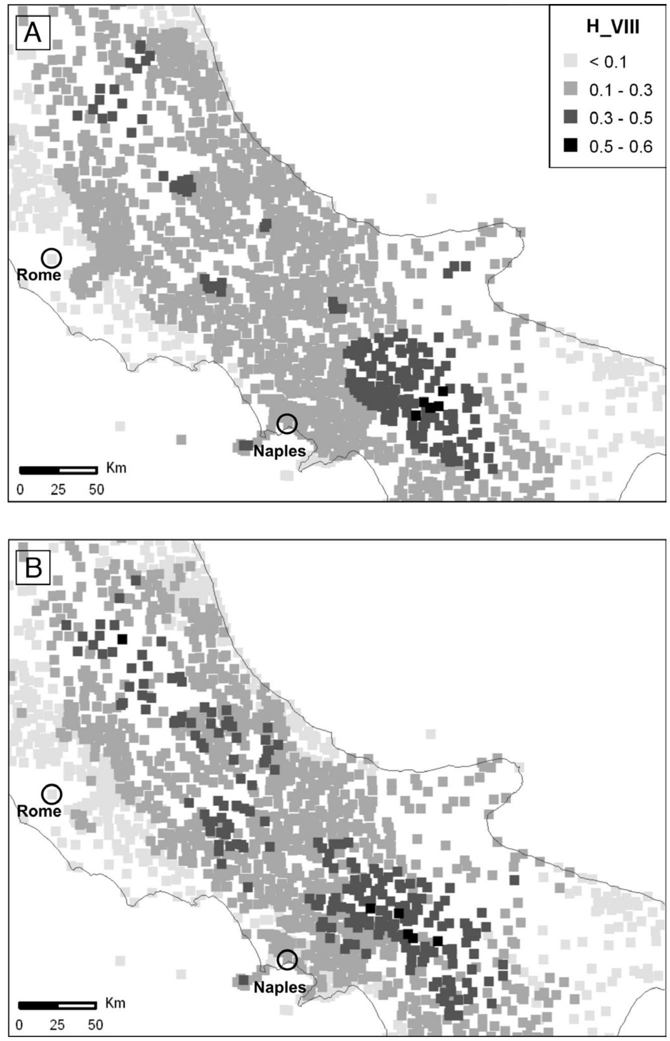
▲ **Figure 3.** (A) Maximum intensity values documented since 217 B.C. at each site examined (half degrees are rounded to upper integer intensity value); (B) Sites where at least 10 macroseismic data are available (data from Stucchi *et al.* 2007).



▲ **Figure 4.** Hazard curves computed at the cities of Milan, Rome, and Messina (for location see Figure 2). Open symbols indicate the relevant I_{ref} values (10% exceedance probability in 50 years).

at the same localities through the standard PSHA approach (MPS Working Group 2004).

SASHA also allows users to analyze seismic histories entirely made of virtual seismic effects deduced from epicentral data. This possibility can be useful for a preliminary detection of sites exposed to ground-motion amplifications strong enough to significantly enhance the seismic hazard level expected from radiation effects. Indications in this direction can be gained by comparing the hazard estimates deduced from virtual effects only, representative of average attenuation properties at regional scale, with those computed by considering observed intensity data and thus taking into account possible site amplifications during past earthquakes. In general, intensity values locally reduced from epicentral data are affected by a larger uncertainty, and this leads to hazard estimates higher than those derived from less uncertain felt data (*e.g.*, D'Amico and Albarello 2003). If the opposite occurs, this could be taken as a proxy for site effects (see D'Amico *et al.* 2002; Gallipoli *et al.* 2002). As an example, Figure 5 compares the two hazard levels assessed for central-southern Italy, expressed as the probability that during a future 50-year time span, each site will be shaken by at least one earthquake with local effects ≥ VIII MCS (the intensity threshold for structural damage to buildings). As expected, estimates derived by source data only (Figure 5A) show a more homogeneous pattern than those obtained by considering felt data (Figure 5B), which appear more spotlike and characterized by a wider spread of high-hazard sites over the region.



▲ **Figure 5.** Hazard estimates, expressed as probability H that each site will experience effects of intensity ≥VIII MCS in 50 years, computed at municipalities of central-southern Italy by considering: A) virtual effects only and B) felt data integrated with virtual effects.

CONCLUSIONS

The SASHA program for the estimate of seismic hazard from a probabilistic analysis of macroseismic data is presented and made freely available (together with the user guide and sample input files) upon request to the authors. The program allows the computation of typical hazard outputs (hazard curve for different intensity thresholds, intensity or PGA value characterized by a fixed exceedance probability in a given exposure time) for a generic set of sites by fully exploiting information locally available about seismic effects (intensity data) of past earthquakes. This information can be integrated with "virtual" intensity values deduced from epicentral data and macroseismic observations available at neighboring sites. Uncertainties about intensity data and the actual reliability of local seismic histories are taken into account.

SASHA has been developed in the framework of a recent research project devoted to seismic hazard assessment in Italy (INGV-DPC Project S1; `http://esse1.mi.ingv.it`), where the program has been extensively applied. For this reason, region-dependent coefficients of the empirical relations presently implemented in the code (*e.g.*, to compute intensity attenuation with source distance and to "convert" intensity hazard estimates to PGA) are the ones most recently proposed for the Italian territory. Relevant parameters can, however, be modified by the user to allow application to different countries.

ACKNOWLEDGMENTS

This work was supported by the Italian Civil Protection Department (DPC, Departimento della Protezione Civile) and the Istituto Nazionale di Geofisica e Vulcanologia (INGV), within the framework of their 2004–2006 Agreement (Project S1).

REFERENCES

Albarello, D., R. Azzaro, M. S. Barbano, S. D'Amico, V. D'Amico, R. Rotondi, T. Tuvè, and G. Zonno (2007). Valutazioni di pericolosità sismica in termini di intensità macrosismica utilizzando metodi di sito. INGV-DPC Project S1, Deliverable D9; `http://esse1.mi.ingv.it/d9.html`.

Albarello, D., V. Bosi, F. Bramerini, A. Lucantoni, G. Naso, L. Peruzza, A. Rebez, F. Sabetta, and D. Slejko (2000). Carte di pericolosità sismica del territorio nazionale. *Quaderni di Geofisica* **12**, 1–8.

Albarello, D., F. Bramerini, V. D'Amico, A. Lucantoni, and G. Naso (2002). Italian intensity hazard maps: A comparison of results from different methodologies. *Bollettino di Geofisica Teorica e Applicata* **43**, 249–262.

Albarello, D., R. Camassi, and A. Rebez (2001). Detection of space and time heterogeneity in the completeness of a seismic catalog by a statistical approach: An application to the Italian area. *Bulletin of the Seismological Society of America* **91**, 1,694–1,703.

Albarello, D., and V. D'Amico (2005). Validation of intensity attenuation relationships. *Bulletin of the Seismological Society of America* **95**, 719–724.

Albarello, D., V. D'Amico, P. Gasperini, F. Pettenati, R. Rotondi, and G. Zonno (2007). Nuova formulazione delle procedure per la stima dell'intensità macrosismica da dati epicentrali o da risentimenti in zone vicine. INGV-DPC Project S1, Deliverable D10; http://esse1.mi.ingv.it/d10.html.

Albarello, D., and M. Mucciarelli (2002). Seismic hazard estimates using ill-defined macroseismic data at site. *Pure and Applied Geophysics* **159**, 1,289–1,304.

Albarello, D., and M. Mucciarelli (2003). Seismic hazard assessment and site effects evaluation at regional scale. In *Earthquake Science and Seismic Risk Reduction*, ed. F. Mulargia and R. J. Geller, 148–158. NATO Science Series IV, vol. 32. Boston: Kluwer Academic Publishers.

Azzaro, R., M. S. Barbano, S. D'Amico, T. Tuvè, D. Albarello, and V. D'Amico (2008). First studies of probabilistic seismic hazard assessment in the volcanic region of Mt. Etna (southern Italy) by means of macroseismic intensities. *Bollettino di Geofisica Teorica e Applicata* **49**, 77–91.

Azzaro, R., M. S. Barbano, A. Moroni, M. Mucciarelli, and M. Stucchi (1999). The seismic history of Catania. *Journal of Seismology* **3**, 235–252.

Bender, B., and D. M. Perkins (1987). *SEISRISK III: A Computer Program for Seismic Hazard Estimation*. USGS Bulletin 1772, 48 pps.

Bozkurt, S. B., R. S. Stein, and S. Toda (2007). Forecasting probabilistic seismic shaking for greater Tokyo from 400 years of intensity observations. *Earthquake Spectra* **23**, 525–546.

Cornell, C.A. (1968). Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* **58**, 1,583–1,606.

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

D'Amico, V., and D. Albarello (2003). The role of data processing and uncertainty management in seismic hazard evaluations: Insights from estimates in the Garfagnana-Lunigiana area (northern Italy). *Natural Hazards* **29**, 77–95.

D'Amico, V., D. Albarello, and M. Mucciarelli (2002). Validation through HVSR measurements of a method for the quick detection of site amplification from intensity data: An application to a seismic area in northern Italy. *Soil Dynamics and Earthquake Engineering* **22**, 475–483.

Faccioli, E., and C. Cauzzi (2006). Macroseismic intensities for seismic scenarios estimated from instrumentally based correlations. Proceedings, First European Conference on Earthquake Engineering and Seismology, paper number 569, 3–8 September 2006. Geneva Switzerland. Published on CD.

Gallipoli, M. R., D. Albarello, M. Mucciarelli, V. Lapenna, M. Schiattarella, and G. Calvano (2002). Hints about site amplification effects comparing macroseismic hazard estimate with microtremor measurements: The Agri Valley (Italy) example. *Journal of Earthquake Engineering* **7**, 51–72.

Magri, L., M. Mucciarelli, and D. Albarello (1994). Estimates of site seismicity rates using ill-defined macroseismic data. *Pure and Applied Geophysics* **143**, 617–632.

McGuire, R. K. (1978). *FRISK: Computer Program for Seismic Risk Analysis Using Faults as Earthquake Sources*. USGS Open File Report 78-1007.

Monachesi, G., L. Peruzza, D. Slejko, and M. Stucchi (1994). Seismic hazard assessment using intensity point data. *Soil Dynamics and Earthquake Engineering* **13**, 219–226.

MPS Working Group (2004). Redazione della mappa di pericolosità sismica prevista dall'Ordinanza PCM 3274 del 20 marzo 2003. Rapporto conclusivo per il Dipartimento della Protezione Civile, INGV, Milano-Roma, April 2004, 65 pps. + 5 appendices; http://zonesismiche.mi.ingv.it.

Mucciarelli, M., D. Albarello, and V. D'Amico (2006). Comparison between the Italian seismic hazard map (PRSTN04) and alternative PSHA estimates. Proceedings, First European Conference on Earthquake Engineering and Seismology, paper number 595.

Mucciarelli, M., D. Albarello, and V. D'Amico (2008). Comparison of probabilistic seismic hazard estimates in Italy. Submitted to the *Bulletin of the Seismological Society of America*.

Mucciarelli, M., L. Peruzza, and P. Caroli (2000). Tuning of seismic hazard estimates by means of observed site intensities. *Journal of Earthquake Engineering* **4**, 141–159.

Papoulia, J., and D. Slejko (1997). Seismic hazard assessment in the Ionian islands based on observed macroseismic intensities. *Natural Hazards* **14**, 179–187.

Pasolini, C., P. Gasperini, D. Albarello, B. Lolli, and V. D'Amico (2008a). The attenuation of seismic intensity in Italy, part I: Theoretical and empirical backgrounds. *Bulletin of the Seismological Society of America* **98**, 2, 682–691.

Pasolini C., D. Albarello, P. Gasperini, V. D'Amico, and B. Lolli (2008b). The attenuation of seismic intensity in Italy, part II: Modeling and validation. *Bulletin of the Seismological Society of America* **98** (2), 692–708.

Rock, N. M. S. (1988). *Numerical Geology*. Berlin: Springer-Verlag, 427 pps.

Slejko, D., L. Peruzza, and A. Rebez (1998). Seismic hazard maps of Italy. *Annali di Geofisica* **41**, 183–214.

Stucchi, M., R. Camassi, A. Rovida, M. Locati, E. Ercolani, C. Meletti, P. Migliavacca, F. Bernardini, and R. Azzaro (2007). DBMI04, il database delle osservazioni macrosismiche dei terremoti italiani utilizzate per la compilazione del catalogo parametrico CPTI04; http://emidius.mi.ingv.it/DBMI04/, *Quaderni di Geofisica* **49**, 1–38.

Istituto Nazionale di Geofisica e Vulcanologia
via della Faggiola, 32
56126 Pisa, Italy
damico@pi.ingv.it
(V.D.)

University of Siena, Department of Earth Sciences
via Laterina, 8
53100 Siena, Italy
albarello@unisi.it
(D.A.)