Bulletin of the Seismological Society of America, Vol. 98, No. 2, pp. 682–691, April 2008, doi: 10.1785/0120070020

The Attenuation of Seismic Intensity in Italy, Part I: Theoretical and Empirical Backgrounds

by C. Pasolini, P. Gasperini, D. Albarello, B. Lolli, and V. D'Amico*

Abstract We critically analyze the results on seismic intensity attenuation in Italy derived by Albarello and D'Amico (2004) and Gasperini (2001). We demonstrate that, due to the inadequacy of certain underlying assumptions, the empirical relationships determined in those studies did not best reproduce the decay of intensity as the distance from the source increases. We reconsidered some of the relevant concepts and assumptions used in these intensity-attenuation studies (macroseismic epicenter, epicentral intensity, data completeness) to suggest some useful recipes for obtaining unbiased estimates. In particular, we suggest that (1) data for distances from the source at which an intensity below the limit of diffuse perceptibility (≤IV) is expected should be excluded from attenuation computations because such data are clearly incomplete, (2) attenuation equations that include a term proportional to the epicentral intensity I_0 with a coefficient different from 1.0 must not be used because they imply a variable offset between I_0 and the intensity expected at the epicenter, and (3) epicentral intensities must be recomputed consistently with the attenuation equation because those reported by the Italian catalog do not generally correspond with the intensity predicted at the epicenter by the attenuation relationships so far proposed. Following these suggestions produces a significant reduction in the standard deviation of the model that might lead to a corresponding reduction of the estimates of seismic hazard.

Introduction

Despite its nonquantitative nature, the concept of seismic intensity is an irreplaceable tool for measuring the level of seismicity in many countries and thus for estimating seismic hazard. This is particularly true for Italy, where most information about large earthquakes comes from descriptive sources. In fact, due to the continuous presence of a relatively high level of civilization for many centuries in Italy, a huge volume of documentary sources dating back to the Roman age has been preserved. These sources can be (and, in many cases, have been) used to determine the characteristics of past earthquakes.

Intensity (macroseismic) data are, in many countries, usually employed only to locate and estimate the size of earthquakes and then to compute past seismicity rates to be used in probabilistic seismic hazard assessments made in terms of peak ground acceleration (Cornell, 1968; Bender and Perkins, 1987). However, seismic hazard assessment based on intensity as a shaking parameter has been proposed both in Italy (Slejko *et al.*, 1998, Albarello *et al.*, 2002) and in other countries (McGuire, 1993). In this case, a reliable definition of the attenuation law in terms of seismic intensity is also required.

In the last few years, a number of studies on macroseismic intensity attenuation in Italy have been published (Gasperini, 2001; Carletti and Gasperini, 2003; Albarello and D'Amico, 2004). However, we have found that although the studies were conducted correctly from the statistical point of view, the resulting attenuation relationships do not describe the behavior of intensity attenuation in Italy as well as they might. There are essentially three reasons for this: the inconsistent form of the attenuation relationship adopted to model the macroseismic field, the lack of a coherent definition of epicentral intensity, and the statistical biases induced by the incomplete recording of low-intensity data far from the source.

The existence of possible biases has been noted by Albarello and D'Amico (2004). They showed that the numbers of occurrences of observed intensities above given thresholds do not match those predicted by the attenuation relationships they proposed (columns N_{obs} and N_{pred} in Table 1). Such a relative imbalance can be reduced only by artificially increasing the value of the model standard obtained by the regression analysis by about 20% (Table 1). Albarello and D'Amico also indicated a more relevant overestimation

*Present address: Istituto Nazionale di Geofisica e Vulcanologia, Via della Faggiola, 32, I-56126 Pisa, Italy; damico@pi.ingv.it.

Table 1

Observed and Predicted Occurrences of Intensities

I_s	N_{obs}	σ_{obs}	N_{pred}	σ_{pred}	Difference (%)	Z
Albarello and D'Amico (2004) with model S.D. = 1.072						
VI	12822	18	13587	67	6	−10.97
VII	7786	22	7277	57	−7	8.33
VIII	3492	18	3086	43	−12	8.65
IX	1141	11	953	27	−17	6.34
X	332	6	188	13	−43	9.94
XI	46	4	20	4	−56	4.19
Carletti and Gasperini (2003)						
VI	11896	16	11206	54	−6	12.32
VII	6804	22	6772	50	−1	0.61
VIII	2258	18	3284	41	45	−22.86
IX	391	8	1189	29	204	−26.80
X	68	4	292	16	333	−13.83
XI	2	1	43	6	2778	−6.40
Albarello and D'Amico (2004) with model S.D. = 1.250						
VI	12822	18	13860	71	8	14.10
VII	7786	22	7664	61	−2	−1.89
VIII	3492	18	3437	46	−2	−1.10
IX	1141	11	1213	31	6	2.19
X	332	6	300	17	−9	−1.79
XI	46	4	44	7	−3	0.15
Carletti and Gasperini (2003) with corrected I_0						
VI	11896	16	11856	50	0	0.74
VII	6804	22	6620	48	−3	3.48
VIII	2258	18	2667	37	18	−10.02
IX	391	8	715	22	83	−13.69
X	68	4	108	10	60	−3.85
XI	2	1	7	3	358	−1.96
Carletti and Gasperini (2003) with free I_0 coefficient						
VI	11896	16	11643	57	−2	4.25
VII	6804	22	6309	53	−7	8.61
VIII	2258	18	2457	40	9	−4.61
IX	391	8	624	23	60	−9.67
X	68	4	92	9	37	−2.47
XI	2	1	7	3	369	−1.99
Albarello and D'Amico (2004) with data selection rule by Gasperini (2001)						
VI	12822	18	12568	60	−2	4.02
VII	7786	22	7393	52	−5	6.91
VIII	3492	18	3636	43	4	−3.10
IX	1141	11	1361	31	19	−6.74
X	332	6	341	17	3	−0.50
XI	46	4	48	7	5	0.26

Comparison between observed (obs) and predicted (pred) occurrences of intensities above given thresholds and corresponding standard deviations (σ) using different relationships and datasets. Difference % is the difference in percentage between predicted and observed numbers, while Z represents the standardized Gaussian statistics defined by Albarello and D'Amico (2005) as $Z = (N_{obs} - N_{comp})/\sqrt{\sigma^2_{obs} + \sigma^2_{comp}}$. If $|Z| > 1.97$, the resulting discrepancy can be considered statistically significant at the 5% level.

(ranging from 45% for intensity VIII to more than 2000% for intensity XI) for the bilinear model determined by Carletti and Gasperini (2003). In the present work, the inadequacy of these relationships is examined in depth by discussing in detail some theoretical and empirical issues regarding the concept of intensity and its attenuation as the distance from the epicenter increases. The result of this analysis is a set of recommendations that must be followed in order to determine unbiased attenuation relationships. It can be demonstrated that if these recommendations are followed, a substantial rebalance between observed and predicted intensity counts occurs for the relation used by Carletti and Gasperini (2003), as well as for that used by Albarello and D'Amico (2004), without the need to correct artificially the standard deviation of the model. In a companion article (Pasolini *et al.*, 2008), we follow these recommendations to deduce a reliable intensity attenuation relationship for Italy on the basis of the most recently released intensity database.

In what follows, we first discuss a physical interpretation of seismic intensity and of two parameters representative of the macroseismic field: the macroseismic epicenter and the epicentral intensity. Even though these parameters are commonly used in seismological practice, they still lack a clear and unequivocal definition. Then, we will review briefly the attenuation relationships proposed for the Italian region. In particular, in order to reveal possible biases, we will discuss in some detail a number of important hidden assumptions that underlie each of them (uniform scaling, consistency at the epicenter, and dataset completeness).

The Nature of Seismic Intensity and of Related Epicentral Parameters

Although seismic intensity is an index based on the qualitative descriptions of earthquake effects observed in a limited area, it can be related to the ground motion amplitude through a logarithmic transformation. In fact, a linear relation between intensity and the logarithm of ground peak acceleration (firstly proposed by Cancani, 1904) was an objective of Sieberg (1931) when he compiled the Mercalli–Cancani–Sieberg (MCS) macroseismic scale. The presence of such a logarithmic relationship has, to some extent, been confirmed by empirical investigations in many parts of the world (see Margottini *et al.*, 1992; Boatwright *et al.*, 2001; Wu *et al.*, 2003; Kaka and Atkinson, 2004).

On the other hand, it is well known that the major physical causes of seismic-wave amplitude attenuation are geometrical spreading and anelastic dissipation. In addition, multipath scattering across crustal discontinuities, and near-source effects may play a significant role, even at relatively long epicentral distances. While geometrical spreading and anelastic dissipation are quite well described by seismic-wave theory (in terms of power and exponential laws, respectively), multipath scattering and near-source effects are not easily tractable because they depend on the particular local structure of the crust and on the geometry of the source (which is unknown for most historical earthquakes). Moreover, it remains very difficult to determine the relative contribution of different seismic phases (direct, reflected, surface, etc.) to the maximum ground motion. For example, some instrumental investigations (Somerville and Yoshimura, 1990; McGarr *et al.*, 1991) have shown that Moho-reflected phases might play an important role in explaining the apparent lack of attenuation at epicentral distances in the range 30–100 km from the 1989 Loma Prieta earthquake. Then it is possible to formulate an empirical expression for the way in which macroseismic intensity depends on distance. Such an expression would include a linear term that corresponds to anelastic and/or scattering-induced dissipation and a logarithmic term that accounts for geometrical spreading. However, the values of the relevant coefficients cannot be physically constrained because the real path followed and the relative contributions of the different phases that control seismic intensity are unknown.

Since the pioneering work by von Kővesligethy (1906), most studies (Blake, 1941; Howell and Schultz, 1975; Gupta and Nuttli, 1976; Chandra, 1979; Chandra *et al.*, 1979; Tilford *et al.*, 1985; Papazachos and Papaioannou, 1997, 1998; Gasperini, 2001; Lee and Kim, 2002; García *et al.*, 2003; Boughacha *et al.*, 2004) have analyzed the attenuation of seismic intensity with distance in terms of the difference between epicentral and site intensity ($\Delta I = I_0 - I$) or vice versa. Such approach requires (1) the assumption that intensity scales with ground motion amplitude uniformly over the entire range of values and (2) a consistent definition of epicentral intensity I_0 as a level to which all the intensity values are referred.

In general, modeling intensity attenuation with the distance from the source requires the clear and reliable definition of the macroseismic epicenter (representative of the source location) and epicentral intensity (representative of the source strength). Some possible definitions are reviewed in brief in the following sections.

Macroseismic Epicenter

There are two main possible approaches (Cecic *et al.*, 1996; Cecic and Musson, 2004) to defining the macroseismic epicenter. The first considers the epicenter as the center of the entire intensity distribution and determines it as the point that minimizes the sum of squared residuals with respect to an attenuation relationship (Sirovich, 1996; Bakun and Wentworth, 1997). The second assumes only the highest intensities to be representative of the source location and determines the epicenter as the geometrical center (the barycenter) of the area that shows the largest effects. In both cases, the underlying assumption is that the macroseismic epicenter represents the surface projection of the seismic source centroid (from which the seismic wave field appears to radiate) rather than of the focus (where the rupture nucleates). Such assumption is consistent with the concept of the intensity center proposed by Bakun (1999).

Both approaches have their advantages and disadvantages. However, the second approach is able to separate the unknown parameters (source coordinates and attenuation coefficients). Furthermore, it has been demonstrated to be rather stable with respect to site misplacements and intensity assessment errors when robust estimators are used to compute the barycenter (Gasperini *et al.*, 1999; Gasperini and Ferrari, 2000). Because of these features, the second approach was used when compiling the Italian seismic catalogue (Catalogo Parametrico dei Terremoti Italiani [CPTI] Working Group, 1999, 2004).

Epicentral Intensity

Although quite commonly employed even in recent times (see the previous section), the concept of epicentral intensity I_0 is not clearly defined in the literature (Cecic *et al.*, 1996; Cecic and Musson, 2004). Generally speaking, it should correspond to the intensity observed at the epicenter and also should reflect a general feature of the macroseismic field in relation to the strength of the source. However, it is difficult to develop such definition into an objective and reproducible procedure. This difficulty is due not only to the absence of a site exactly located at the epicenter but also to the possible presence of amplification effects induced by local geostructural conditions at sites showing the largest intensities, which could make this parameter less representative of the source strength. If the epicenter is located in an inhabited inland area, one possible choice is to define epicentral intensity as the largest observed intensity in the absence of local amplification. This working definition is the one most commonly used in practice and was adopted for the Italian seismic catalog used for hazard assessment (CPTI Working Group, 1999, 2004). However, it is not clear how local amplification should be assessed. Furthermore, the use of this definition is very likely to result in the actual source strength being overestimated because the largest intensities tend to reflect anomalously high levels of ground motion. This is particularly true when largest intensities are few, located relatively far from each other, and intermixed with much lower intensities, as, for example, is the case for the two large Italian earthquakes shown in Figure 1a (occurred in Irpinia in 1980) and Figure 1b (occurred in Garfagnana in 1920).

The risk of biased estimates of I_0 could be mitigated by considering average intensities observed at sites close to the epicenter (within a given distance) rather than the largest ones only. However, in that case, the estimated average would correspond approximately to the intensity expected at the average epicentral distance of sites considered for averaging intensity rather than at the epicenter. Thus, a further correction is required to obtain the intensity at the epicenter.

An alternative definition of epicentral intensity can be given if this parameter is determined jointly with attenuation parameters. Epicentral intensity may then be defined as the intensity expected at the epicenter or simply the largest expected intensity by the attenuation relationship.

These two definitions (largest observed and expected intensities) do not necessarily coincide because the latter is constrained by the entire macroseismic dataset while the former is constrained by only a subset of the data.

These problems in defining precisely and consistently epicentral intensity are probably the reason why many investigators, particularly in the United States (but not in other countries like Italy), abandoned completely I_0 in favor of magnitude in modeling the attenuation of seismic intensity. However, we will show that a procedure to redefine I_0 consistently with the attenuation relationship can be formulated and that such redefinition has the advantage to represent not only the strength of the seismic source but also a measure of the ground shacking at the epicenter in terms of the same parameter. We also note that most magnitudes that could be used in intensity attenuation relationships are empirical in nature because physically grounded estimates of moment magnitude from the inversion of complete seismograms are

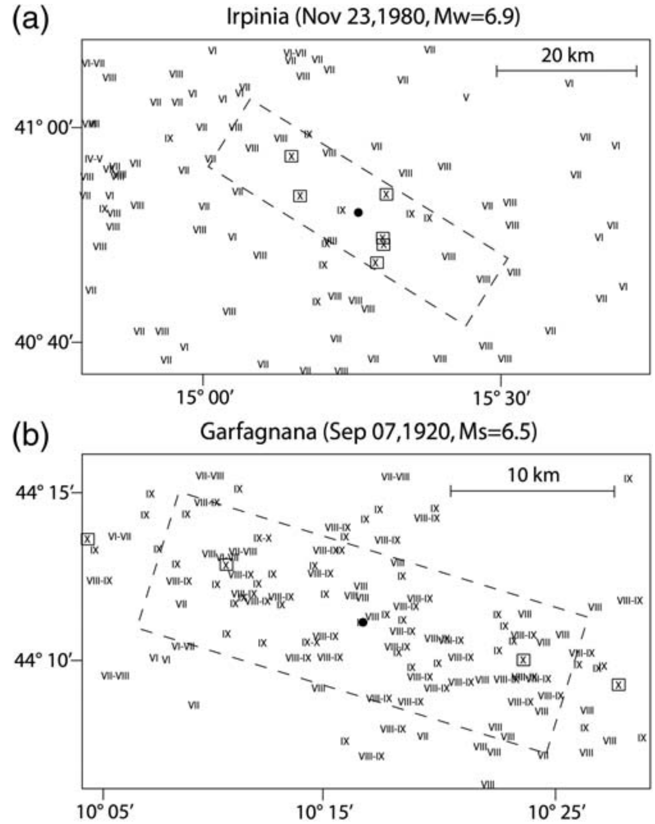

Figure 1. Maps showing the spatial distribution of intensity data for the 23 November 1980 Irpinia earthquake (M_w 6.9) and the 7 September 1920 Garfagnana earthquake (M_s 6.5). Black circles indicate the macroseismic epicenters (barycenter), while large dashed boxes indicate the surface projections of the faults estimated according to Gasperini *et al.* (1999). Small solid boxes highlight the sites where the maximum intensity (X) is observed. In both maps, these sites are intermixed with sites that have lower intensity.

not available for earthquakes that occurred before the 1970s. For such earthquakes, the magnitude is usually given in terms of empirical definitions (M_S, m_B, M_L, etc.) based on the maximum amplitude recorded or derived from empirical relationships with intensity data (i.e., Bakun and Wentworth, 1997; Gasperini *et al.*, 1999). In particular, in Italy, more than one-half of the earthquakes included in the Italian catalog have neither an instrumental magnitude nor a set of documented intensities that can be used to constrain the macroseismic magnitude by reliable methods; instead, they provide only an estimate of I_0. In these cases, the use of the magnitude to predict the intensity I at a site would imply a double empirical conversion from I_0 to M and then from M to I.

Attenuation Relationships

The previously mentioned assumption of a uniform scaling of intensity with the logarithm of ground motion amplitude is the basis of the von Kővesligethy (1906) relation

$$\Delta I = I_0 - I = 3\log_{10}\left(\frac{D}{h}\right) + 3\alpha(\log_{10} e)(D-h), \quad (1)$$

where $D = \sqrt{R^2 + h^2}$ is the hypocentral distance in kilometers, R is the epicentral distance, h is the source depth, and α is an empirical parameter. This formula reflects Cancani's (1904) hypothesis that seismic-wave amplitude approximately doubles for an increase of one intensity degree. This implies that the intensity is roughly proportional to the decimal logarithm of the seismic-wave amplitude with a coefficient close to $1/\log_{10}(2) \approx 3$. Koveslighety's decision to fix the coefficient of the logarithmic term at 3 is the result of his additional assumption that intensity is controlled mainly by the amplitude of direct body waves for which the geometrical spreading exponent is -1.0. The coefficient α in the linear term was left free to fit the different anelastic dissipation properties of the Earth's crust.

A similar approach, neglecting the linear term, was later proposed by Blake (1941) in his formula to compute the source depth

$$\Delta I = s \log_{10}\left(\frac{D}{h}\right), \quad (2)$$

where s is a free parameter to be estimated from data.

We note that both these relations (usually applied to the isoseismal radii estimated visually from hand-drawn isoseismals) consistently give $\Delta I = 0$ ($I = I_0$) at $R = 0$ ($D = h$). However, most subsequent work introduced an intercept term in order to improve the data fitting. This is the case, for example, with the Gupta and Nuttli (1976) formula, which is based on physical considerations analogous to those made by Koveslighety:

$$\Delta I = C_1 + C_2(\gamma\delta \log_{10} e + \log_{10} \delta), \quad (3)$$

where γ is the coefficient of anelastic attenuation (determined independently on the basis of instrumental investigations), δ is the epicentral (not hypocentral) distance in degrees, and C_1 and C_2 are empirical coefficients. The authors restricted the validity of this formula, which gives negative values at short distances and diverges at $\delta = 0$ to $\delta > 20$ km (where $\Delta I > 0$). However, this does not remove the inconsistency of predicting ΔI values significantly different from 0 at or near the geometrical epicenter.

The intercept term also appears in the bilinear model proposed by Gasperini (2001) and later used by Carletti and Gasperini (2003) as a reference isotropic model for mapping the spatial variations of attenuation properties in Italy:

$$\Delta I = a + b \min(D, D_{\mathrm{cross}}) + c \max(0, D - D_{\mathrm{cross}}), \quad (4)$$

where $D_{\mathrm{cross}} = 45$ km. The physical justification for this model comes from the hypotheses that anelastic dissipation properties in the Earth's crust are depth dependent and that dissipation dominates over geometrical spreading in the considered interval of distances from the source (>10 km). The values of the empirical coefficients estimated by Carletti and Gasperini (2003) (Table 2) contradict the implicit assumption that $\Delta I = 0$ at the epicenter. In fact, as one can see in Figure 2, the relationship proposed by Carletti and Gasperini (2003) provides $\Delta I \approx -1$ at $R = 0$, irrespective of I_0. Other authors removed this inconsistency by various methods. Chandra *et al.* (1979) forced $\Delta I \approx 0$ at the epicenter by recomputing it by an iterative procedure, while Tilford *et al.* (1985) imposed explicit constraints to make $\Delta I = 0$ at the epicenter.

As already noted, all of these forms of attenuation relation assume that the scaling of intensity with distance is independent of the source strength. However, the spectral structure of seismic ground motion depends significantly on the size of the source, and this could be responsible for a different distribution of effects on the anthropic environment. In order to take this factor into account, López Casado *et al.* (2000) and Albarello and D'Amico (2004) adopted attenuation relationships that include a term proportional to I_0 in the form

$$I = a + bD + c \ln(D) + dI_0. \quad (5)$$

The value of coefficient $d = 0.705$ estimated by Albarello and D'Amico (2004) from Italian data (Table 2) differs significantly from 1.0. This result seems to support the necessity of introducing into the attenuation equation an explicit dependence of attenuation properties on the source strength. Other authors have made this approach more explicit by including the magnitude in the attenuation equation directly (see Chavez and Castro, 1988; Bakun and McGarr, 2002; Bakun *et al.*, 2003, Musson, 2005). This possible option, which could be useful in hazard analysis given that it allows the prediction of expected intensity even for instrumental earthquakes, will be considered in the companion article.

We note that even for the Albarello and D'Amico (2004) relationship, the expected intensity at the epicenter I_E underestimates the values of I_0 reported by the catalog by about 1° on average. Using the regression coefficients determined by Albarello and D'Amico (2004), we can obtain at the epicenter (even for such work, $h = 10$ km is assumed for all earthquakes):

$$I_E = I(D = h) = 1.31 + 0.705 I_0. \quad (6)$$

This implies that I_E and I_0 coincide among each other only for $I_E \approx I_0 = \mathrm{IV} - \mathrm{V}$, while, for example, for $I_0 = \mathrm{VIII}$, we have $I_E \approx \mathrm{VII}$ and for $I_0 = \mathrm{XI}$, $I_E \approx \mathrm{IX}$ (see Fig. 2). Note that, in general, an attenuation formula that includes a term proportional to I_0 with a fitted coefficient that differs significantly from 1 inevitably leads to an inconsistency between the intensity predicted at the epicenter and I_0, independently of the existence, as in this case, of an average offset between them. The reasons for this inconsistency may be found in the inadequacy of the formulation or in the incorrectness of the definition of I_0.

Table 2
Model Parameters for Attenuation Laws

Model and Dataset	a	b	c	d	S.D.	R^2
A&D original dataset (no large distances selection)	3.570 ± 0.108	-0.0030 ± 0.0004	-0.9840 ± 0.0320	0.705 ± 0.005	1.072	0.63
A&D corrected dataset (with large distances selection)	1.848 ± 0.158	-0.0137 ± 0.0009	-0.7084 ± 0.0534	0.847 ± 0.008	1.047	0.58
A&D Io averaged with $D_{max} = 50$ km	4.148 ± 0.086	-0.0038 ± 0.0003	-1.0597 ± 0.0261	0.915 ± 0.005	0.885	0.75
A&D Io averaged with $D_{max} = 50$ km and corrected to epicenter	3.310 ± 0.072	-0.0024 ± 0.0003	-1.1944 ± 0.0265	0.915 ± 0.005	0.873	0.76
A&D Io averaged with $D_{max} = 300$ km and corrected to epicenter	2.375 ± 0.068	-0.0060 ± 0.0003	-1.0126 ± 0.0260	0.978 ± 0.005	0.821	0.78
C&G original	0.445 ± 0.019	0.0590 ± 0.0007	0.0207 ± 0.0003		1.040	0.43
C&G free Io coefficient	-1.405 ± 0.046	0.0497 ± 0.0007	0.0165 ± 0.0003	0.739 ± 0.006	0.991	0.48
C&G Io averaged with $D_{max} = 300$ km and corrected to epicenter	-0.416 ± 0.007	0.0562 ± 0.0006	0.0206 ± 0.0003		0.829	0.64
C&G Io corrected and free Io coefficient	-0.830 ± 0.035	0.0580 ± 0.0006	0.0199 ± 0.0003	0.946 ± 0.005	0.826	0.64

A&D denotes the Albarello and D'Amico (2004) log-linear relation with free I_0 coefficient (equation 5) while C&G denotes the bilinear relation (equation 4), originally proposed by Gasperini (2001) with parameters computed by Carletti and Gasperini (2003), as well as for the modifications of them described in the text. Hypocentral distances in equations (4) and (5) are in kilometers. In all cases, the same dataset and the same computational procedures as in the original articles are used.

A Procedure for Redefining Epicentral Intensity Consistently with the Attenuation Relationship

Following the idea that I_0 must be representative of the average intensity observed in the vicinity of the epicenter, we can first compute it as the average of intensities observed at sites located within a distance D_{max} from the hypocenter. Using this definition with the same equation and dataset used by Albarello and D'Amico (2004) and $D_{max} = 50$ km, we obtain a significant improvement in fitting the attenuation equation, as indicated by a decrease of the standard deviation (S.D.) of model residuals from 1.072 to 0.885 (Table 2). However, this procedure is likely to underestimate I_0 because it corresponds roughly to the intensity observed at the average distance $\bar{D}$ of sites within D_{max}, rather than just at the epicenter. This is confirmed by the comparison of the intensity values expected at the epicenter (I_E) and I_0. This comparison shows that

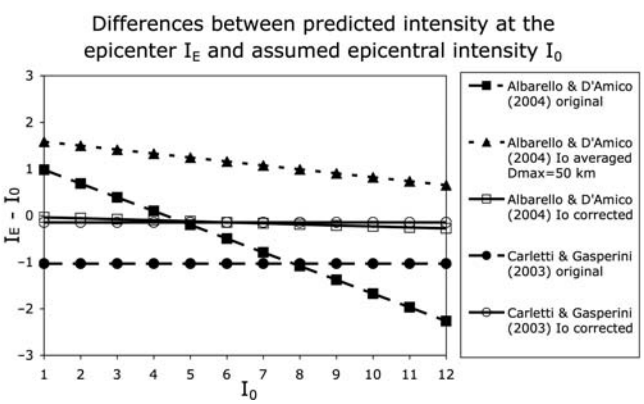

Figure 2. Differences between expected intensity at the epicenter I_E and epicentral intensity I_0 as a function of I_0 for various models that use the original I_0 reported in the catalog or recomputed according to the procedure described in the text.

$$I_E = 1.67 + 0.915 I_0. \tag{7}$$

Here, for significant intensities (> V), I_E overestimates I_0 by about 1° (for $I_0 = \mathrm{V}$, $I_E \approx \mathrm{VI}$; for $I_0 = \mathrm{VIII}$, $I_E \approx \mathrm{IX}$; for $I_0 = \mathrm{XI}$, $I_E \approx \mathrm{XI} - \mathrm{XII}$, etc.; see Fig. 2).

Therefore, to make I_0 coincident with I_E, an additional correction, which accounts for the attenuation from $D = h$ to $D = \bar{D}$, is required. In a manner similar to that used by Chandra *et al.* (1979), it is possible to use an iterative procedure to correct I_0 on the basis of the attenuation coefficients previously computed:

$$I_0 = \bar{I} + b(h - \bar{D}) + c[\ln(h) - \ln(\bar{D})]. \tag{8}$$

Another possibility is to apply such correction simultaneously to the estimation of attenuation parameters, in the procedure of least-squares minimization. To do this, equation (8) must be substituted in the attenuation relationship (5):

$$I = a + bD + c\ln(D) + d\{\bar{I} + b(h - \bar{D}) + c[\ln(h) - \ln(\bar{D})]\}. \tag{9}$$

Because of the nonlinearity of the equation (the unknown parameter d multiplies both b and c), a numerical minimization algorithm must be used in this case. The result of such minimization, performed by using a quasi-Newton method (Dennis and Schnabel, 1983) implemented by the Fortran routine DUMINF of the International Mathematics and Statistics Library (IMSL) (Visual Numerics, 1997), shows that the fit further improves (S.D. = 0.873) and that the computed offsets between I_E and $\mathbf{I_0}$ become of the order of a few tenths of a degree, on average.

Varying D_{max} from 10 km to the maximum distance of 300 km, we also find that the residual standard deviation decreases almost monotonically down to about 0.82 intensity degree (Fig. 3) and that the corresponding estimated parameters of the attenuation equation (Table 2) for $D_{\max} =$ 300 km give at the epicenter ($D = 10$ km)

$$I_E = -0.017 + 0.978 I_0, \tag{10}$$

which indicates a very good coincidence of I_E with I_0 over the entire intensity range, with a maximum underestimation of about 0.3 intensity units for $I_0 =$ XII (Fig. 2). It is interesting to note how the best fit is obtained with a very large $D_{\max}$ (i.e., by including in the averages almost all the intensity data available). It should also be noted that, after such a redefinition of I_0, the coefficient d becomes very close to 1.0, thus eliminating the inconsistency noted previously. We may conclude that the value of this coefficient found by Albarello and D'Amico (2004), which differed significantly from 1.0, is likely to be an artifact induced by an inconsistent estimate of I_0.

One might object that the new definition of I_0 does not furnish an integer value and thus cannot be associated directly with a given degree of the intensity scale and with the corresponding set of macroseismic effects. However, the assumption that intensity behaves like a rational number with a uniform metric is implicit in the application of any attenuation relation given in functional form. From one standpoint, intensity can be seen as a physical quantity that is measured using a low-resolution instrument that has only integer notches (the intensity scale), but whose true value could be any real number (for other possible interpretations of noninteger intensity values, see also the introductory discussion in Gasperini, 2001). Alternatively, we might even not attribute a true meaning to the noninteger estimate of the intensity at the epicenter and might consider it simply as a parameter of the attenuation relationship. This procedure for recomputing I_0 can easily be adapted to any other formulation of the attenuation function. If g is the generic function that represents the distance dependence of seismic intensity, I_0 can be redefined as

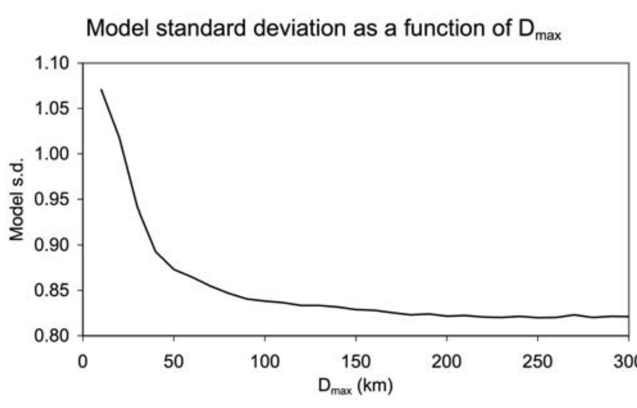

Figure 3. Standard deviation of model residuals as a function of $D_{\max}$ for the same attenuation relationship and dataset used by Albarello and D'Amico (2004) but with I_0 recomputed (equation 9; see text).

$$I_0 = \bar{I} - g(h) + g(\bar{D}). \tag{11}$$

A possible alternative to the log-linear equation is the previously mentioned bilinear equation (equation 4) proposed by Gasperini (2001) and used by Carletti and Gasperini (2003). We have verified that if the epicentral intensity in equation (4) is redefined according to equation (11):

$$I_0 = \bar{I} + bh - b\min(\bar{D}, D_{\text{cross}}) - c\max(0, \bar{D} - D_{\text{cross}}). \tag{12}$$

The estimated values of the parameters of the bilinear relationship (Table 2) indicate that the difference between the expected intensity I_E at the epicenter ($D = 10$ km) and I_0 is very close to 0 for any I_0 (Fig. 2).

Moreover, it is easy to verify that the overestimation of predicted numbers of intensities larger than VI for the bilinear relationship (Table 1) is due mainly to the inconsistency of I_0 or, conversely, to the assumption made by Gasperini (2001) and Carletti and Gasperini (2003) that the coefficient of I_0 (not recomputed according to the procedure described previously) is 1. If in fact the bilinear relationship is applied using I_0 recomputed according to equation (11), or if the coefficient of I_0 in the bilinear model is left to vary, we have a significant rebalance between the observed and predicted numbers of occurrences (Table 1).

The problem regarding the relation proposed by Albarello and D'Amico (2004) is more subtle. In the next section, we will show that it is related to the completeness of the macroseismic dataset and thus to the selection rules that Albarello and D'Amico (2004) applied to it.

Selection of Data

It may seem that Carletti and Gasperini (2003) and Albarello and D'Amico (2004) use the same dataset—the one resulting from the merging of intensity data coming from the Catalogo dei Forti Terremoti Italiani, Version 2 (CFTI2; Boschi *et al.*, 1997) and the Database delle Osservazioni Macrosismiche, Version 4.1 (DOM4.1; Monachesi and Stucchi, 1997) databases according to the choices made by the CPTI catalog (CPTI Working Group, 1999). However, the two articles apply different selection criteria to the data in order to reduce possible sources of bias. In particular, different from Carletti and Gasperini (2003), Albarello and D'Amico (2004) discard, in the regression analysis, all uncertain intensity observations (like VI–VII, IX–X, etc.) as well as all the data for earthquakes that have uncertain I_0. In the companion article, we will discuss the consequences of Albarello and D'Amico's choice and demonstrate that it actually affects the results very little. An additional selection criterion was applied by both articles, but with slightly different procedures, to eliminate the effects of near-source anisotropy: Albarello and D'Amico (2004) discard all intensity data at hypocentral distances less than 15 km, while Carletti and Gasperini (2003) adopt, according to Gasperini

(2001), a variable minimum distance depending on the earthquake's moment magnitude. These differences in selection criteria are the cause of the discrepancies among the observed number of occurrences of intensity data larger than IX for the two attenuation equations (Table 1). However, it will be shown in the companion article that even in this case, such choices alter significantly neither the values of attenuation parameters nor the ability to reproduce the data well. In what follows, we will concentrate on another difference between the datasets that were actually used by two articles, a difference that has significant consequences for both the values of coefficients and the predictive ability of the attenuation relation.

Different from Albarello and D'Amico (2004), Gasperini (2001) and Carletti and Gasperini (2003) noted that intensity estimates that are based on the feeling of shaking reported by few or very few people (below degree IV MCS) are likely to be missed by macroseismic reports that concern small settlements. Moreover, mild effects such as those characterizing low degrees may be neglected due to the low interest of macroseismic investigators. This incompleteness in reporting might result in the average intensity observed at relatively long distances being overestimated; in fact, intensities that are greater than average (perhaps exceeding the level of diffuse perceptibility) are more likely to be reported by documentary sources or noted by surveyors than those lower than average. Using these data in the fitting of the attenuation equation would result in predicted intensity being overestimated; hence, the attenuation would seem to be reduced at large distances. To remove this bias, Gasperini (2001) suggested discarding from the dataset all intensities observed at distances greater than those at which an intensity below IV is expected on the basis of a preliminary attenuation estimate. The precise selection criterion adopted (but not explicitly reported in the original articles) by Gasperini (2001) and by Carletti and Gasperini (2003) is to exclude all the data (independently of the observed intensity) located at hypocentral distances D for which

$$I_0 - 0.53 - 0.055\min(D, 45) - 0.022\max(D - 45, 0) < 4. \tag{13}$$

Note that this selection criterion does not introduce a bias in the attenuation regression because it applies to hypocentral distances (the independent variable) but not to observed intensities (the dependent variable). A consequence of applying this selection rule is that the maximum hypocentral distance of the data used by Carletti and Gasperini (2003) is reduced from about 300 to 220 km.

Albarello and D'Amico (2004) disagreed with this reasoning and the corresponding criterion for selection. They argued instead for the absence of the low-intensity bias by citing the lack of a statistically significant skewness of the intensity residual distribution at large distances. The overestimation of the predicted intensity at distances larger than about 120 km for the equation by Albarello and D'Amico (2004) with respect to Carletti and Gasperini (2003) can be observed in the superposed plot of the two attenuation functions displayed in figure 3 of Albarello and D'Amico (2004).

The nature of this discrepancy can be shown clearly by testing the behavior of the average residuals of the model as a function of the predicted intensity. In fact, for the relation and dataset used by Albarello and D'Amico (2004), we note from the data presented in Figure 4 that the observed intensities are significantly larger than those predicted (positive average residuals) in the range of values lower than IV and greater than VI. By contrast, we can note slightly negative residuals for predicted intensities IV and V (these latter two classes alone represent more than half of all the intensity data). Although the overall average of residuals is constrained to be zero by the least-squares procedure, the significant deviations (well outside the corresponding 95% confidence intervals) of single intensity classes indicate a biased fit that could be the cause of the overestimation of the expected frequencies of intensities larger than VI with respect to the observed ones actually verified by Albarello and D'Amico (2004).

In fact, if the selection rule (13) adopted by Gasperini (2001) and Carletti and Gasperini (2003) is applied to the dataset used by Albarello and D'Amico (2004), the average residuals of their attenuation relation become almost constant and very close to zero for all predicted intensities (Fig. 5). In addition, the application of these selection rules produces a substantial rebalance between observed and expected frequencies of occurrence without the need to introduce an artificial increase of the model standard deviation to 1.250 (Table 1). Although the rebalances of counts are similar, the former approach appears to be sounder and more physically grounded than the latter.

In our opinion, the evidence just presented indicates clearly that data for cases of low intensity are incomplete in the Albarello and D'Amico (2004) dataset and that the selection rule proposed by Gasperini (2001) is able to reduce its possible effect on the statistical analysis.

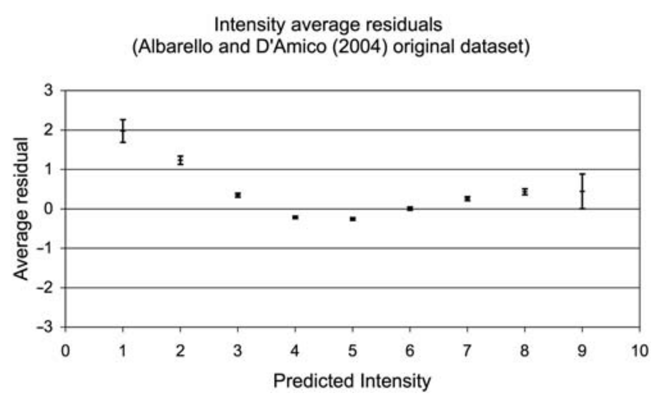

Figure 4. Intensity average residuals of the Albarello and D'Amico (2004) attenuation relationship for classes of predicted intensity. Error bars indicate 95% confidence intervals.

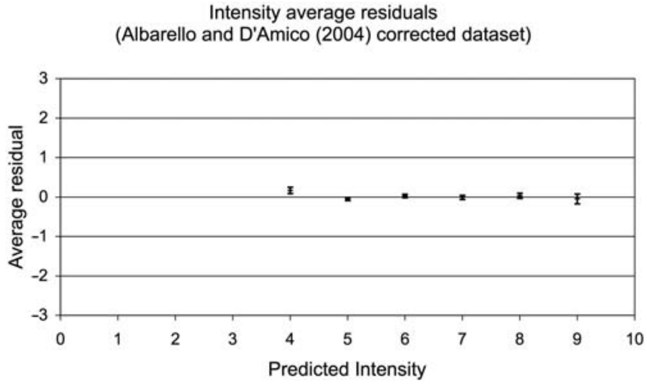

Figure 5. Intensity average residuals of the Albarello and D'Amico (2004) attenuation relationship for classes of predicted intensity when the selection rule proposed by Gasperini (2001) is applied to the data. Error bars indicate 95% confidence intervals.

Conclusions

We demonstrated that previous analyses by Gasperini (2001), Carletti and Gasperini (2003), and Albarello and D'Amico (2004) do not provide the best reproduction of the attenuation of seismic intensity in Italy. This is due to an inconsistent definition of epicentral intensity and to the incompleteness of intensity data at large distances from the source. In particular, we found that epicentral intensity I_0, commonly estimated as the highest observed one, differs significantly from the intensity predicted at the epicenter I_E by such attenuation models.

We also found that the use of intensity data located at distances from the source longer than those at which a degree IV is expected results in the intensity at large distances being overestimated. We also showed that these problems are the major cause of the imbalance between observed and predicted frequencies of occurrences of intensities already noted by Albarello and D'Amico (2004, 2005) for previous analyses.

Given the foregoing, for the future analysis of seismic intensity in Italy (and perhaps in other countries), we recommend the following:

1. The dataset used to fit the attenuation equation must exclude data located at distances from the source for which an intensity value below the limit of diffuse perceptibility (about IV) is expected. This can be done by, for example, following the procedure applied by Gasperini (2001) and Carletti and Gasperini (2003), which discards intensity data located at hypocentral distances D for which the intensity predicted by a preliminary attenuation law is lower than IV (equation 13).
2. The inclusion in the attenuation equation of a term proportional to the epicentral intensity I_0 with a coefficient different from 1 must, in general, be avoided. Although such a term might partially rebalance the overall counts of expected and observed intensities above given thresholds, it predicts a variable offset between I_0 and the intensity expected at the epicenter that is in contradiction with the meaning of I_0.
3. Epicentral intensity must be rendered consistent with the intensity predicted at the epicenter by the attenuation relationship. This can be done by, for example, redefining I_0 as a function of the average intensity $\bar{I}$ and the average hypocentral distance $\bar{D}$ of localities from which each earthquake is observed:

$$I_0 = \bar{I} - g(D = h) + g(D = \bar{D}),$$

where g is the function that represents the distance dependence of seismic intensity and h is the source depth.

If such recommendations are applied to the log-linear attenuation law proposed by Albarello and D'Amico (2004), the standard deviation of the model residual of about 30% will be reduced with respect to previous analyses (from 1.07 to 0.82). Because of the role played in probabilistic hazard assessment analyses by the standard deviation of the attenuation function model, this might imply a significant reduction of the hazard maxima assessed by previous analyses.

The aforementioned set of prescriptions must be considered as a first step. They do not exhaust the problems concerning the unbiased fit of intensity attenuation. Other issues, such as the role of source depth and geometry and the possible dependence of attenuation coefficients on earthquake energy and their spatial variability, still need to be addressed more deeply. The role of depth will be discussed in the companion article, while the other issues will be the subject of further studies currently in preparation.

Acknowledgments

We thank Bill Bakun and two anonymous reviewers for their thoughtful comments and suggestions. This work was supported by the Italian Civil Defence Department and the Istituto Nazionale di Geofisica and Vulcanologia (INGV), within the framework of their 2004–2006 agreement (Project S1). INGV also supported the Ph.D. grant of author C. Pasolini.

References

Albarello, D., and V. D'Amico (2004). Attenuation relationship of macroseismic intensity in Italy for probabilistic seismic hazard assessment, *Boll. Geofis. Teorica Appl.* **45,** 271–284.

Albarello, D., and V. D'Amico (2005). Validation of intensity attenuation relationships, *Bull. Seismol. Soc. Am.* **95,** 719–724.

Albarello, D., F. Bramerini, V. D'Amico, A. Lucantoni, and G. Naso (2002). Italian intensity hazard maps: a comparison of results from different methodologies, *Boll. Geofis. Teorica Appl.* **43,** 249–262.

Bakun, W. H. (1999). Seismic activity of the San Francisco Bay region, *Bull. Seismol. Soc. Am.* **89,** 764–784.

Bakun, W. H., and A. McGarr (2002). Differences in attenuation among the stable continental regions, *Geophys. Res. Lett.* **29,** 2121, doi 10.1029/2002GL015457.

Bakun, W. H., and C. M. Wentworth (1997). Estimating earthquake location and magnitude from seismic intensity data, *Bull. Seismol. Soc. Am.* **87,** 1502–1521.

Bakun, W. H., A. C. Johnston, and M. G. Hopper (2003). Estimating locations and magnitudes of earthquakes in eastern North America from modified Mercalli intensities, *Bull. Seismol. Soc. Am.* **93,** 190–202.

Bender, B., and D. M. Perkins (1987). *Seisrisk III: A Computer Program for Seismic Hazard Estimation*, U.S. Geol. Surv., Denver, Colorado, 1–48.

Blake, A. (1941). On the estimation of focal depth from macroseismic data, *Bull. Seismol. Soc. Am.* **31,** 225–231.

Boatwright, J., K. Thywissen, and L. C. Seekins (2001). Correlation of ground motion and intensity for the 17 January 1994 Northridge, California, earthquake, *Bull. Seismol. Soc. Am.* **91,** 739–752.

Boschi, E., E. Guidoboni, G. Ferrari, G. Valensise, and P. Gasperini (1997). *Catalogo dei Forti Terremoti in Italia dal 461 a.C. al 1990*, Istituto Nazionale di Geofisica/Storia Geofisica Ambiente Srl., Bologna, Italy, 644 pp. (and enclosed CD-ROM; in Italian).

Boughacha, M. S., M. Ouyedl, A. Ayadi, and H. Benhallou (2004). Seismicity and seismic hazard mapping of northern Algeria: map of maximum calculated intensities (MCI), *J. Seism.* **8,** 1–10.

Cancani, A. (1904). Sur l'emploi d'une double echelle sismique des intensités, empirique et absolute, *G. Beitr., Erganzungsband* **2,** 281–283 (in French).

Carletti, F., and P. Gasperini (2003). Lateral variations of seismic intensity attenuation in Italy, *Geophys. J. Int.* **155,** 839–856.

Catalogo Parametrico dei Terremoti Italiani (CPTI) Working Group (1999). Catalogo Parametrico dei Terremoti Italiani (CPTI), http://emidius.mi.ingv.it/CPTI/home.html (last accessed January 2008; in Italian).

Catalogo Parametrico dei Terremoti Italiani (CPTI) Working Group (2004). Catalogo parametrico dei terremoti Italiani, (CPTI04), http://emidius.mi.ingv.it/CPTI04/ (last accessed January 2008).

Cecic, I., and R. Musson (2004). Macroseismic surveys in theory and practice, *Nat. Hazards* **31,** 39–61.

Cecic, I., R. Musson, and M. Stucchi (1996). Do seismologists agree upon epicentre determination from macroseismic data? A survey of the ESC "macroseismology" working group, *Ann. Geofis.* **39,** 5, 1013–1027.

Chandra, U. (1979). Attenuation of intensities in the United States, *Bull. Seismol. Soc. Am.* **69,** 2003–2024.

Chandra, U., J. G. McWhorter, and A. A. Nowroozi (1979). Attenuation of intensities in Iran, *Bull. Seismol. Soc. Am.* **69,** 237–250.

Chavez, M., and R. Castro (1988). Attenuation of modified Mercalli intensity with distance in Mexico, *Bull. Seismol. Soc. Am.* **78,** 1875–1884.

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

Dennis, J. E., and R. B. Schnabel (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice Hall, Englewood Cliffs, New Jersey.

García, J., D. Slejko, L. Alvarez, L. Peruzza, and A. Rebez (2003). Seismic hazard maps for Cuba and surrounding areas, *Bull. Seismol. Soc. Am.* **93,** 2563–2590.

Gasperini, P. (2001). The attenuation of seismic intensity in Italy: a bilinear shape indicates dominance of deep phases at epicentral distances longer than 45 km, *Bull. Seismol. Soc. Am.* **91,** 826–841.

Gasperini, P., and G. Ferrari (2000). Deriving numerical estimates from descriptive information: the computation of earthquake parameters, *Ann. Geofis.* **43,** 729–746.

Gasperini, P., F. Bernardini, G. Valensise, and E. Boschi (1999). Defining seismogenic sources from historical earthquake felt reports, *Bull. Seismol. Soc. Am.* **89,** 94–110.

Gupta, I. N., and O. W. Nuttli (1976). Spatial attenuation of intensities for central U.S. earthquakes, *Bull. Seismol. Soc. Am.* **66,** 743–751.

Howell, B. F., and T. T. Schultz (1975). Attenuation of modified Mercalli intensity with distance from the epicenter, *Bull. Seismol. Soc. Am.* **65,** 651–665.

Kaka, S. I., and G. M. Atkinson (2004). Relationship between instrumental ground-motion parameters and modified Mercalli intensity in Eastern North America, *Bull. Seismol. Soc. Am.* **94,** 1728–1736.

Lee, K., and J. Kim (2002). Intensity attenuation in the Sino-Korean craton, *Bull. Seismol. Soc. Am.* **92,** 783–793.

López Casado, C., S. Molina Palacios, J. Delgado, and J. A. Peláez (2000). Attenuation of intensity with epicentral distance in the Iberian peninsula, *Bull. Seismol. Soc. Am.* **90,** 34–47.

Margottini, C., D. Molin, and L. Serva (1992). Intensity versus ground motion: a new approach using Italian data, *Eng. Geol.* **33,** 45–58.

McGarr, A., M. Çelebi, E. Sembera, T. Noce, and C. Mueller (1991). Ground motion at the San Francisco international airport from the Loma Prieta earthquake sequence, 1989, *Bull. Seismol. Soc. Am.* **81,** 1923–1944.

McGuire, R. K. (Editor) (1993). *The Practice of Earthquake Hazard Assessment*, Int. Assoc. Seism. Phys. Earth Interiors (IASPEI), Denver, Colorado, 284 pp.

Monachesi, G., and M. Stucchi (1997). GNDT Open-File Rept. DOM4.1, Un database di osservazioni macrosismiche di terremoti di area italiana al di sopra della soglia del danno, Gruppo Nazionale per la Difesa dai Terremoti (GNDT), Milano-Macerata, http://emidius.mi.ingv.it/DOM/home.html (last accessed January 2008; in Italian).

Musson, R. (2005). Intensity attenuation in the U.K., *J. Seism.* **9,** 73–86.

Papazachos, C., and Ch. Papaioannou (1997). The macroseismic field of the Balkan area, *J. Seism.* **1,** 181–201.

Papazachos, C., and Ch. Papaioannou (1998). Further information on the macroseismic field in the Balkan area, *J. Seism.* **2,** 363–375.

Pasolini, C., D. Albarello, P. Gasperini, V. D'Amico, and B. Lolli (2008). The attenuation of seismic intensity in Italy, part II: Modeling and validation, *Bull. Seismol. Soc. Am.* **98,** no. 2, 692–708.

Sieberg, A. (1931). Erdebeben, in *Handbuch der Geophysik* B. Gutenberg (Editor) Vol. **4**, 552–554 (in German).

Sirovich, L. (1996). A simple algorithm for tracing out synthetic isoseismals, *Bull. Seismol. Soc. Am.* **86,** 1019–1027.

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

Somerville, P., and J. Yoshimura (1990). The influence of critical Moho reflections on strong ground motion recorded in San Francisco and Oakland during the 1989 Loma Prieta earthquake, *Geophys. Res. Lett.* **17,** 1203–1206.

Tilford, N. R., U. Chandra, D. C. Amick, R. Moran, and F. Snider (1985). Attenuation of intensities and effect of local site conditions on observed intensities during the Corinth, Greece, earthquakes of 24 and 25 February and 4 March 1981, *Bull. Seismol. Soc. Am.* **75,** 923–937.

Visual Numerics (1997). *International Mathematics and Statistics Library (IMSL) Volumes l and 2* Visual Numerics, Inc., Houston.

von Kővesligethy , R. (1906). Seismonomia, *Boll. Soc. Sismol. Italy* **11,** 113–250 (in Latin).

Wu, Y. M., T. Teng, T. C. Shin, and N. C. Hisiao (2003). Relationship between peak ground acceleration, peak ground velocity and intensity in Taiwan, *Bull. Seismol. Soc. Am.* **93,** 386–396.

Dipartimento di Fisica
Università di Bologna
Viale Berti Pichat 8
I-40127 Bologna, Italy
chiara.pasolini@gmail.com
paolo.gasperini@unibo.it
barbara.lolli@unibo.it
(C.P., P.G., B.L.)

Dipartimento di Scienze della Terra
Università di Siena
Via Laterina 8
I-53100 Siena, Italy
albarello@unisi.it
(D.A., V.D.)

Manuscript received 2 February 2007