Alma Mater Studiorum
Università di Bologna

Facoltà di Scienze Fisiche, Matematiche e Naturali
Dipartimento di Fisica

Dottorato di Ricerca in Geofisica

XX Ciclo

The attenuation of seismic intensity

Candidato:
Dr.ssa
Chiara Pasolini

Tutore:
Chiar.mo Prof.
Paolo Gasperini

Coordinatore:
Chiar.mo Prof.
Michele Dragoni

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Seismic intensity is an index that describes the effects on people, objects and buildings that are observed when seismic waves arrive at certain locality.

Despite its not quantitative nature, more than 60% of national hazard maps worldwide are expressed in terms of seismic intensity (McGuire, 1993). This choice can be explained considering that seismic intensity is irreplaceable for the representing of damage scenarios for future earthquakes, moreover in many countries (such as Italy), most of information about major past earthquakes comes from documentary sources that date back to the preinstrumental age.

A widely used method to estimate seismic hazard was proposed by Cornell (1968) and implemented by Bender and Perkins (1987). Such method needs an attenuation law that describes the decrease of the shaking parameter with distance from source. Moreover it requires the complete probabilistic form of the attenuation relationship. In particular Albarello et al. (2002) show that assuming a Normal distribution of residuals, the variance associated with attenuation law affects dramatically the hazard estimates. As variance increases the hazard estimates become less reliable. Every data set presents an intrinsic variability that could be estimated by subdividing data in epicentral distance bins and calculating the standard deviation for each of them. This represents a lower bound for standard regression associated to attenuation law.

Two intensity attenuation law have been recently proposed for Italy by Gasperini (2001) and Albarello and D'Amico (2004). By applying the procedure described above, a standard deviation significantly lower that those associated to the two attenuation laws is obtained. This indicates that their data fitting could be improved.

In chapter 2 of the present work a new intensity attenuation law for Italy is presented, formulated considering its entire probabilistic distribution, taking into account that the associated standard deviation has to be

reduced as more as possible. To achieve this objective some open issues characterizing the previous attenuation relation are considered. Epicentral intensity definition, which is not unique in literature, choice of the functional form of the attenuation law, which can be done empirically between various candidate models on the basis of their data fitting or can be based on assumptions related to the physics of seismic wave propagation, presence in the data base of observations that present the effect typical of a certain intensity degree but also some effect characterizing the higher degree (uncertain degrees, i.e. VII-VIII), are some examples of such problems.

A two step regression procedure that allow to separate propagation from sources effect on the observed intensity is applied.

The procedure followed in chapter 2 allow to obtain attenuation laws that represent on average how intensity decreases with distance. Considering the high structural complexity characterizing the Italian peninsula, strong deviations from this average behavior are expected at local scale.

To study the internal attenuation structure of a certain volume of material on the basis of measures taken above or near its surface, the tomographic method can be used. This method has been applied in a great number of discipline including geophysics. In particular, in seismology it allows to reconstruct the internal structure of the Earth in terms of physical properties related to seismic waves propagation and is based on the idea that seismic waves contain information about the materials that they crossed in their path between seismic source and station that records them.

Tomographic method had been widely applied in seismology to know the velocity structure of the Earth from travel time recorded at various seismic stations and also to study anelastic attenuation properties characterizing the material crossed by seismic waves using amplitude data (Iyer and Hirahara, 1993).

Also propagation properties of seismic intensity can be studied by tomographic method. Chapter 2 is focused on the log-linear law, that is physically grounded and widely used in hazard estimates. Comparing the log-linear attenuation law with the log-bilinear on the basis of Information Criteria, which premiate data fitting of the model but penalize models with an higher number of free parameter, the log-bilinear model results preferable. Chapter 3 concerns the estimation of lateral variations of attenuation law coefficients, taking as reference model the values of the coefficients of the log-bilinear law presented in chapter 2. Once such variations have been estimated, by assuming a linear relation between intensity and logarithm of PGA the anelastic attenuation structure in terms of seismic quality factor (Q) can be obtained.

When amplitude data are available, anelastic attenuation can be stud-

ied by estimating directly the seismic quality factor characterizing the materials involved.

Chapter 2

Intensity attenuation law

As evidenced in chapter 1, to estimate seismic hazard by Cornell (1968) method it is necessary to defining an attenuation law of the shaking parameter with distance. The attenuation law has to be described in its entire probabilistic form. It is desirable that the standard deviation characterizing such a relation is as low as possible. In the following sections some open issues regarding the use of intensity as ground shaking parameter are reconsidered.

A first aspect to consider in the statistical formalization of the attenuation law is that intensity is defined as a discrete quantity, in fact it can assume only integer values. A second aspect is that the data set includes "uncertain degrees" (i.e. VII-VIII). Following a maximum likelihood approach allows to treat in a formally correct way such aspects. Hence, it is necessary to defining a parameter that describes the strength of the seismic source. Epicentral intensity is a common choice but it definition in literature is not unique and has to be reconsidered. In particular it could be defined as the intensity expected at the epicenter by an attenuation law. In this case a two step procedure can be followed. The parameters describing the distance dependence of intensity are calculated in the first step and the epicentral intensity in the second one. As the relation between felt intensity and seismic wave propagation is not clearly known, the functional form of the attenuation law that describes the intensity decrease with distance from seismic source is not univocally defined. Moreover, various distance definitions are possible. A definition that substantially agrees with Gasperini (2001) and Albarello and D'Amico (2004) is chosen here, but instead of assuming an "average" hypocentral depth of 10 km, as they do, the depth is calculated directly from the data set. A log-linear functional form is assumed, but it is compared with other forms commonly presented in literature on the basis of their data fitting and also using Information Criteria, which penalize models with an higher number of parameters in order to avoid over fit-

ting. As a Normal distribution of residual is assumed, errors in parameter estimations could be evaluated from variance-covariance matrix. Bootstrap method could be applied to check their reliability. Fictitious data set are built by randomly resampling the original one, then the parameters values are evaluated for each of them. The frequency distribution of such values gives an empirical estimate of the parameters errors. The reliability of the presented attenuation law is also tested by observing the stability of the parameters of attenuation law when different values of some arbitrary parameters involved in the regression procedure are assumed.

A peculiarity of the regression procedure followed is that epicentral intensity is defined as the the intensity expected at the epicenter by the attenuation law. This is valid for earthquakes included in the computations. Regressions with other earthquake size measures are calculated to obtain epicentral intensity for each earthquake of the catalog. This makes the presented attenuation law suitable to be used in seismic hazard estimations.

2.1   Statistical formalization

2.1.1   Probability distribution of the attenuation law

To use an intensity attenuation law in seismic hazard estimates its complete probability distribution must be defined. This implies that a continuous distribution model (such as the Normal one) is applied to intensity data, as well as that intensity is defined as a discrete quantity. To consider this, the total probability of a given integer observation I is computed as the integral of the probability density function over the interval $[I-0.5, I+0.5]$ of the continuous variable that the integer function represents.

Gasperini (2001) and Albarello and D'Amico (2004) show that the distribution of intensity residuals for the Italian data set is close to Normal. Under the hypothesis that discrepancies from normality, albeit statistically significant, are not relevant in the practice for the evaluation of attenuation parameters, a Normal distribution of residuals can be assumed. In this case, the probability of observing an intensity I_S at a certain site, for an earthquake of size J occurring at distance D from the site, is

$$F(I_S) = prob(I = I_S | J, D) = \frac{1}{\sigma\sqrt{2\pi}} \int_{I_S-0.5}^{I_S+0.5} \exp\left[-\frac{(i - \mu(J, D))^2}{2\sigma^2}\right] \mathrm{d}i\,,$$

$$(2.1)$$

where $\mu(J, D)$ represents the mean of the probability distribution that is the intensity expected by the attenuation law at a distance D from the source, for an earthquake of size J, and σ is the standard deviation of data respect to the average value.

Eq. 2.1 will be used in section 2.1.3 to estimate by the maximum likelihood method the values of the parameters of the attenuation law.

2.1.2 Uncertain degrees

About 30% of the intensities included in the data set corresponds to "uncertain" intensity assessments (e.g. VII-VIII). This means that at the considered site the lower degree (e.g. VII) has certainly been observed, but there are also some evidences for the higher one (e.g.VIII). Such data have been treated in different ways, for example Gasperini (2001) used them as real intensity values positioned in the middle of the interval between the two degrees (e.g. VII-VIII=7.5) whereas Albarello and D'Amico (2004, 2005) discard all the uncertain degrees to fit the attenuation law and consider the two possible outcomes equally probable in its validation.

Magri et al. (1994) propose to assign to each intensity of the macroseismic scale a degree of belief expressed in terms of probability $p(I)$. According to this point of view a certain probability w_1 to the nearer higher degree (e.g. VIII) and the probability $1 - w_1$ to the nearer lower one could be attributed. The probabilities of the other intensity degrees are zeroes

$$p(I) = [0, 0, 0, 0, 0, 1 - w_1, w_1, 0, 0, 0, 0, 0] . \tag{2.2}$$

Initially $w_1 = 0.5$ is chosen. The effects of assuming different values for w_1 are considered in section 2.8. Also "well-defined" intensity degrees could be defined coherently with this representation. In this case the probability corresponding to the observed intensity degree is equal to 1 and the other are all equal to zero. For example, for I=VII

$$p(I) = [0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0] . \tag{2.3}$$

2.1.3 Maximum likelihood approach

The values of the attenuation law parameters are estimated by a maximum likelihood approach. This method gives the set of parameters values that maximize the probability of observing the data set (likelihood function). The likelihood function of a set of N observations could be written as the product of their independent probabilities

$$l = \prod_{j=1}^{N} g_j(\theta) , \tag{2.4}$$

where g_j is the probability of the j-th observation and θ is the set of model parameters. $g_j(\theta)$ could be calculated using the total probability theorem as

$$g_j(\theta) = \sum_{I=1}^{12} p_j(I) g_j(\theta|I)\,, \qquad (2.5)$$

where $g_j(\theta|I)$ is the probability of the j-th observation for the parameter set θ, given that intensity I actually occurred. In this case the likelihood function (eq. 2.4) becomes

$$l = \prod_{j=1}^{N} \sum_{I=1}^{12} p_j(I) g_j(\theta|I)\,. \qquad (2.6)$$

Following the approach described in section 2.1 the total probability of a given integer observation I is computed as the integral of the probability density function over the interval $[I - 0.5, I + 0.5]$ of the continuous variable i that the integer function represents

$$g_j(\theta) = \int_{I-0.5}^{I+0.5} h_j(\theta, i)\,\mathrm{d}i\,. \qquad (2.7)$$

Assuming a Normal distribution of residuals (see section 2.1) eq. 2.7 becomes

$$g_j(\theta|I) = \frac{1}{\sigma\sqrt{2\pi}} \int_{I-0.5}^{I+0.5} \exp -\frac{(i - \mu_j)^2}{2\sigma^2}\,\mathrm{d}i\,, \qquad (2.8)$$

where μ_j is the intensity expected by the attenuation law for the j-th observation (and depends on the set of model parameters θ) and σ is the corresponding standard deviation. Substituting eq. 2.8 in eq. 2.6 and taking the logarithm gives the log-likelihood function that can be maximized to find the best fitting model parameters values θ:

$$L = \sum_{j=1}^{N} \ln\left[\frac{1}{\sigma\sqrt{2\pi}} \sum_{I=1}^{12} p_j(I) \int_{I-0.5}^{I+0.5} \exp -\frac{(i - \mu_j)^2}{2\sigma^2}\,\mathrm{d}i \right] \qquad (2.9)$$

To maximize L, the quasi-Newton Method (Dennis and Schnabel, 1983) implemented by the Fortran routine UMINF/DUMINF of the IMSL Math Library (Visual Numerics, 1997) is used.

2.2 Attenuation model formulation

2.2.1 Source size definition

In most of intensity attenuation relations the source strength is expressed in terms of epicentral intensity. Even if it is commonly used, this parameter is not clearly defined in 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 the local amplification should be assessed. Furthermore, as largest intensities tend to reflect anomalously high levels of ground motion, if this definition is used the actual source strength probably would be overestimated.

An alternative definition of epicentral intensity can be given if this parameter is determined jointly with other attenuation parameters. In this case epicentral intensity may be defined as "the intensity expected at the epicenter" by the attenuation relationship. Defining epicentral intensity as the largest observed one, its value is constrained by only a subset of the data, whereas this second definition has the advantage of depending on the entire macroseismic data set. In this work I follow such approach that is described more in details in section 2.2.4.

2.2.2 Distance definition

The distance between seismic source and geographical position of intensity observation can be defined in several ways. The simplest way is to consider the distance R between the macroseismic epicenter and the site where the intensity datum is observed. This definition, however, neglects that generally the rupture that originates seismic waves occur at a certain dept from the Earth surface. Thus it is a common practice to obtain the distance as

$$D = \sqrt{R^2 + h^2}\,. \tag{2.10}$$

In this case h represents an empirical equivalent of the epicentral depth (Koveslighety, 1906; Blake, 1941). Even if it does not correspond to the actual distance traveled by seismic waves inside the Earth, this approximation is more realistic than assuming that waves radiate from the epicenter.

Gasperini (2001) and Albarello and D'Amico (2004), on the basis of a rough empirical evaluation of the average depth distribution of Italian in-

strumental hypocenters, assume a fixed source depth of 10 km. However, since the hypocentral depth could significantly differ from the centroid of seismic energy radiation, it would be preferable to estimate the depth of each earthquake directly from the statistical analysis of the data set. This approach has the drawback that the number of free parameters strongly increases and, as a strong interplay between them is expected, the result would be less stable. An alternative solution that follows the same philosophy but leads to more reliable results is to estimate a common depth for all earthquakes directly from the data. This could be viewed as the depth of the "apparent" radiating source. I choose this latter procedure and I calculate the distances relative to each intensity observation using eq. 2.10.

2.2.3   Previous attenuation laws

Seismic intensity is an index of ground shaking based on qualitative description of eartquake effects. Despite its not instrumental nature it may be related to ground motion amplitude. A linear relation between intensity and the logarithm of peak ground acceleration was proposed by Cancani (1904) and Sieberg compiled the Mercalli-Cancani-Sieberg (MCS) macroseismic scale aiming at a linear dependence between these two quantities. Evidences of such a linear dependence have been founded by empirical investigations in many parts of the world (Margottini et al., 1992; Boatwright et al., 1994; Wu et al., 2003; Kaka and Atkinson, 2004).

The mayor 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 large epicentral distances. While geometrical spreading and anelastic attenuation are quite well described by seismic ray theory (in terms of a power and an exponential laws, respectively) multipath scattering and near-source effects are not easily tractable as they depend on the particular local structure of the crust and on the geometry of the source (which is unknown for most historical earthquakes). Starting from this considerations, it is possible to formulate an empirical relation describing the intensity dependence on distance. Such relation would include a linear term that corresponds to anelastic and scattering induced dissipation and a logarithmic term that account for geometrical spreading (see Appendix A for details).

An initial attempt in this sense is represented by the relation proposed by Koveslighety (1906)

$$\Delta I = I_0 - I = 3 \log_{10} \frac{D}{h} + 3\alpha (\log_{10} e)(D - h) \,, \qquad (2.11)$$

where $D = \sqrt{R^2 + h^2}$, is the hypocentral distance in km, R is the epicentral distance, h is the source depth and α an empirical parameter. The first term represents geometrical spreading and the second one the anelastic attenuation. A similar approach, neglecting the linear term, was later proposed by Blake (1941) in his formula to compute hypocentral depth

$$\Delta I = s \log_{10} \frac{D}{h} \tag{2.12}$$

where s is a free parameter estimated from data. Both these relations give $I = I_0$ at $R = 0$ ($D = h$). However subsequent works introduced an intercept term to improve data fitting. One of such relations is the Gupta and Nuttli (1976) formula, based on physical consideration analogous to those of eq. 2.11

$$\Delta I = C_1 + C_2(\gamma\delta \log_{10} e + \log_{10} d) , \tag{2.13}$$

where γ is the coefficient of anelastic attenuation, determined independently from instrumental investigation, δ is the epicentral distance in degrees and C_1 and C_2 are empirical coefficients. The intercept term is also present in the bilinear model proposed by Gasperini (2001)

$$\Delta I = a + b \min(D, D_{cross}) + c \max(0, D - D_{cross}) \tag{2.14}$$

where $D_{cross} = 45\,km$. The physical justification for this model comes from the hypotheses that anelastic dissipation properties in the Earths 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 free parameters are $a = 0.445 \pm 0.019$ $b = 0.0590 \pm 0.0007$ $c = 0.0207 \pm 0.0003$ and they gives $\Delta I = I_0 - I \approx 1$ for $R = 0$, contradicting the implicit assumption that $I = I_0$ at the epicenter. A different point of view underlies the approach of Berardi et al. (1993). As the physical relation between intensity and seismic wave propagation is not known, the model could be chosen empirically between different options. In this case the major points to be considered are the ability of the model to fit the data but also its simplicity, expressed by a low number of free parameters, as the simpler the model the lower the possibility of over-fitting. In particular, they propose an empirical attenuation relation called Cubic Root Attenuation Model (CRAM) that use only two free parameters

$$\Delta I = \alpha + \beta D^{\frac{1}{3}} . \tag{2.15}$$

They found $\alpha = -0.79$ and $\beta = 1.22$ which gives $\Delta I = I_0 - I = -0.79$ for $R = 0$. Also in this case the intensity expected at the epicenter does not correspond to the epicentral intensity 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.

It is important to note that all of the above forms of attenuation relation implicitly 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. 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

$$I = a + bD + c\ln D + dI_0 \,. \tag{2.16}$$

Albarello and D'Amico (2004) obtain $a = 1.31$ and $d = 0.705$. Using these values intensity expected at the epicenter can be calculated as $I = 1.31 + 0.705I_0$. This implies that it is almost coincident with I_0 only for $I_0 = IV - V$. Other authors have made this approach more explicit including directly the magnitude in the attenuation equation (Chavez and Castro, 1988; Bakun and McGarr, 2002; Bakun et al., 2003; Musson, 2005).

2.2.4   Functional form of attenuation law and regression procedure

A first issue to be considered is the introduction of an intercept term. As evidenced in section 2.2.3, if its value differs from 0, this could cause a systematic deviation between epicentral intensity (irrespective of the way it has been defined) and intensity expected at the epicenter from attenuation law. So a formulation without intercept term is chosen. Also introducing a coefficient of proportionality multiplying the epicentral intensity would cause the same discrepancy. In fact a value different from 1 would lead to a deviation between epicentral intensity and that predicted at the epicenter, whose size and sign depends on the value of epicentral intensity. For this reason the coefficient of I_0 is fixed to 1.

With respect to the function describing the distance dependence, I choose a form analogous to the eq. 2.11 including two terms, the first linear and the second logarithmic, due to its physical justification and to its wide use in literature (see section 2.2.3). However, to check the appropriateness of such choice, in section 2.7 the same regression procedure used to estimate the log-linear model is applied to other functional relationships used in literature.

The formula adopted for the attenuation law is

$$\mu(I_E, D) = I_E + a(D - h) + b(\ln D - \ln h) \,, \tag{2.17}$$

where I_E is the intensity expected at the epicenter and h represents a common depth, for all earthquakes, from which seismic energy radiates. To reduce the tradeoff between propagation and source terms, a two step regression analysis is adopted. In the first step, the parameters describing the distance dependence (a, b, h) and the variance σ are estimated independently of the source term I_E. In the second step, I_E is determined from the results of the previous step or from other information.

first step

I_E can be eliminated from eq. 2.17 by considering first an empirical average $\overline{I}_m$ of the observed intensities relative to each m-th earthquake and located within a given epicentral distance R_{max}. It could be assumed reasonably that $\overline{I}_m$ coincides with the arithmetic average of the expectation given by eq. 2.17, relative to the m-th earthquake. This corresponds to

$$\overline{I}_m = I_E + a\frac{1}{M_m}\sum_{k=1}^{M_m} D_k^m - ah + b\frac{1}{M_m}\sum_{k=1}^{M_m}\ln(D_k^m) - b\ln h\,, \qquad (2.18)$$

where M_m is the number of available observations for the m-th events at distances shorter than R_{max}. Then, subtracting eq. 2.18 from eq. 2.17, gives

$$\mu(\overline{I}_m, D) = \overline{I}_m + a(D - \overline{D}_m) + b(\ln D - \overline{\ln D}_m)\,, \qquad (2.19)$$

where $\overline{D}_m = \frac{1}{M_m}\sum_{k=1}^{M_m} D_k^m$ and $\overline{\ln D}_m = \frac{1}{M_m}\sum_{k=1}^{M_m}\ln(D_k^m)$. Note that eq. 2.19 is independent of I_E and thus can be used to empirically fit the distance dependence without considering that parameter. According to the probabilistic approach described in section 2.1, $\overline{I}_m$ can be computed by maximizing, for each m-th earthquake, the likelihood function

$$L = \sum_{k=1}^{M_m}\ln\left[\frac{1}{\sigma_m\sqrt{2\pi}}\sum_{I=1}^{12} p_j(I)\int_{I-0.5}^{I+0.5}\exp-\frac{(i - \overline{I}_m)^2}{2\sigma_m^2}\,\mathrm{d}i\right] \qquad (2.20)$$

In this way, the average intensity $\overline{I}_m$ and the corresponding standard deviation σ_m are computed for each earthquake to be considered in the analysis.
$\overline{I}_m$ can be introduced in eq. 2.19 to calculate μ_j. This is used in eq. 2.9 to calculate the overall log-likelihood. Such function is maximized to determine the free parameters a, b, h and σ.

second step

It may be seen, from eq. 2.18, that I_E can be computed easily for all of the well-documented earthquakes, provided that parameters a, b, h

and σ are computed in the first step. This can be achieved by using the relationship

$$I_E = \mu(\overline{I}_m, h) = \overline{I}_m + a(h - \overline{D}_m) + b(\ln h - \overline{\ln D_m}) \qquad (2.21)$$

These I_E values can be used, in this second step, to fit empirical regressions with different source parameters (magnitude, I_0, I_{max}), which are generally available for all the earthquakes in the catalog. Such relationships allow I_E to be estimated even for earthquakes that lack a suitable macroseismic data set. As a result, eq. 2.17 can be used in place of eq. 2.19 to predict the intensity at a site for all of the earthquakes in the catalog. Of course, a different parametrization of regression uncertainty must be provided in the cases that I_E is obtained directly by the use of eq. 2.21 or is deduced from other epicentral parameters.

2.3 Quality of the attenuation model

2.3.1 Errors in parameter estimations

Under the hypothesis of Normal distribution, errors in parameters estimations can be evaluated from the diagonal terms of the variance-covariance matrix. This can be computed approximately as the inverse of the finite-difference Hessian of the log-likelihood $L(\hat{\theta}, \sigma)$ at the maximum (Guo and Ogata, 1997). Since this estimate represents a lower limit, errors are also evaluated by using a numerical resampling procedure (bootstrap).

The basic premise of the bootstrap approach (Efron and Tbishirani, 1986; Hall, 1992) is that the empirical frequency distribution of data provides an optimal empirical estimate (in the sense of maximum likelihood) of the probability distribution that characterizes the unknown parent population. This hypothesis implies that any new data sets (usually called bootstrap samples or paradata sets) that are obtained by randomly resampling (with replacements) the original set preserve the statistical features of the parent population. Paradata sets can be used to evaluate (via a distribution-free approach) the sampling properties of a given population parameter from the analysis of the empirical values of the parameter computed from each paradata set. In the present application, several paradata sets are obtained from the original data set. For each of them, a new set of parameters of the attenuation model is obtained by maximizing the relevant likelihood function (eq. 2.9). Then a distribution-free evaluation of the estimation errors is obtained from these samples by computing the empirical variance-covariance matrix.

2.3.2 Comparison of different models

An aspect to consider in comparing different attenuation models is their ability in data fitting. The value of σ obtained from the maximization of the likelihood function L of the data set (eq. 2.9) represents a first measure of the goodness of fit characterizing each model. The lower the σ the better its fit to the data.

Another useful parameter is the explained variance (R^2), that represent the fraction of data variability explained by the model and can be expressed in the form

$$R^2 = \frac{\sigma_{ave}^2 - \sigma^2}{\sigma_{ave}^2} \tag{2.22}$$

where σ_{ave} is the standard deviation of observed intensities with respect to the average value $\overline{I}_m$ for each m-the earthquake and is given by

$$\sigma_{ave}^2 = \frac{\sum\limits_{m=1}^{N} \sigma_m^2 M_m}{\sum\limits_{m=1}^{N} M_m} \tag{2.23}$$

where N is the number of earthquakes and σ_m is computed by maximizing eq. 2.20. In the case of a perfect data fitting R^2 would assume the value of 1. As the goodness of fit gets worse, the value of R^2 approaches to zero.

In comparing models with a different number of free parameters it is important to remind that even if the fit improves for models with an higher number of free parameters this does not necessarily imply that their predictive ability is better. Information Criteria take into account this point by a penalty term that depends on the number of free parameters. Bayesian Information Criterion (BIC) is defined as

$$BIC = L(\hat{\theta}) - \frac{k}{2} \ln \frac{n}{2\pi} \tag{2.24}$$

(Schwarz, 1978; Draper, 1995) and Akaike Information Criteria (AIC_C) is

$$AIC_C = L(\hat{\theta}) - k - \frac{k(k+1)}{n-k-1}, \tag{2.25}$$

(Akaike, 1974; Hurvich and Tsai, 1989). In these equations, $\hat{\theta}$ is the set of parameter values maximizing the log-likelihood function $L(\hat{\theta})$, k is the number of free parameters, and n is the number of independent data used in likelihood maximization. The better model in this case is that related to the higher score of Information Criteria. BIC and AIC_C represent two alternative approaches to information theory in model evaluation.

AIC_C is based on the frequentist approach, in which probability is derived from observed frequencies in defined distributions or proportions in populations, whereas BIC is related to Bayesian theory, that interprets probability as subjective degree of belief in a certain proposition.

In section 2.7 different attenuation models are compared on the basis of the above mentionated criteria.

2.4 The data set

2.4.1 Seismic compilations

The data set used for the analysis consists of the most recent version of the Parametric Catalog of Italian Earthquakes (CPTI04) (CPTI Working Group, 2004) and the related Data Base of Macroseismic intensity observations in Italy (DBMI04) (DBMI Working Group, 2007). These databases were constructed by combining and elaborating previous macrosesismic data collections (Boschi et al., 1995, 1997, 2000; Monachesi and Stucchi, 1997).

The CPTI04 catalog contains epicentral information for 2551 damaging (or potentially damaging) earthquakes that have occurred in Italy since 217 B.C. For 1042 of them, the DBMI04 macroseimic database reports a collection of intensity estimates at different localities (macroseismic field) based on documentary information. For most of the remaining earthquakes, the parametrization is also based mainly on macroseismic information, but documentation on effects at individual sites is lacking (no associated macroseismic field is available).

For earthquakes provided with a macroseismic field, the reported epicenter is computed as the barycenter of the localities at which the highest intensities were observed, according to the procedure described by Gasperini et al. (1999) and Gasperini and Ferrari (2000). Such an algorithm, using a robust estimator (trimmed mean), has proved to be fairly stable with respect to site misplacement and errors regarding the assessment of intensity.

For these earthquakes, the catalog also reports the epicentral intensity I_0. This value is computed according to another algorithm, also described by Gasperini et al. (1999) and Gasperini and Ferrari (2000). Generally, I_0 corresponds to the highest intensity observed for the relevant earthquakes. In cases for which local site amplification is likely to have occurred, this maximum intensity has been suitably reduced. In a few cases, the epicentral intensity has been manually adjusted to take into account specific situations (e.g., epicenter located offshore).

The macroseismic database includes 58926 macroseismic observations made at 14821 different sites. Some of these sites (about 100) are not

associated with a well-defined location but refer to very rough geographical definitions (e.g. Northern Italy). These sites cannot be considered in the analysis. About 10000 macroseismic observations are not expressed in terms of standard intensity values (e.g., felt, damages) defined by the Mercalli-Cancani-Sieberg (MCS) scale (Sieberg, 1931). Therefore, only about 48000 observations are suitable for use in the statistical analysis of intensity attenuation.

2.4.2 Data selection

Some data of CPTI04 have been excluded from computations to improve the data reliability or to avoid the introduction of biases a due to a peculiar geographical position of the epicenter.

Epicentral parameters based on few intensity data can be biased by the uneven spatial distribution of observing sites and by local amplification effects. Hence only earthquakes with a number of intensity data reported in the catalog greater than a minimum threshold N_{min} are considered. Initially $N_{min} = 10$ is chosen. The influence of different choices for the N_{min} value on estimated parameters values is evaluated in section 2.8.

The first step of the regression procedure (section 2.2.4) requires the definition of the threshold distance R_{max}. Only intensity observations situated at distances shorter than R_{max} are included in the computation of average intensities by eq. 2.18. The choice of the value R_{max} is quite arbitrary. Initially $R_{max} = 300$ km is used, that corresponds to using intensity points situated at any epicentral distance. Also the effects of the R_{max} value on the parameter estimation are evaluated in section 2.8.

Data relative to earthquakes that occurred before 1200 A.D. have been escluded due to their possible uncompleteness. Moreover, all intensity estimates that were deduced from the effects on single building rather than on settlements were not considered, as such data may depend on the particular structure of the building or may reflect amplification or attenuation effects due to site mechanical properties.

Another relevant issue concerns earthquakes whose epicenters are located offshore. Actually, in these cases, the asymmetric distribution of observations might induce a drift of the possible off-shore epicenter towards the coast and a bias of the relevant attenuation pattern. To detect earthquakes whose epicentral location was possibly biased in this way, for all of the events located close to the coastline, the intensity attenuation pattern as well as the aspect ratio of the area that showed the largest effects have been considered. Narrow areas along the coast are assumed to indicate an offshore epicenter, while almost circular ones have been associated with an epicenter located inland. Fortunately, there are few

offshore seismogenic sources in Italy that show significant activity. Table 2.1 reports 30 earthquakes that are not used in this study because their true epicenters are likely to be located offshore.

year	mo	da	ho	mi	se	lat	lon	epicentral area	No.	I_0
1690	12	23	0	20		43.546	13.593	Ancona area	17	8.0
1743	2	20	16	30		39.852	18.777	Southern Ionian Sea	77	9.0
1823	3	5	16	37		37.993	14.094	Northern Sicily	107	8.0
1831	5	26	10	30		43.855	7.849	West. Liguria	33	8.0
1848	1	11				37.543	15.174	Augusta	41	7.5
1854	12	29	1	45		43.813	7.543	West. Liguria-France	86	7.5
1875	3	17	23	51		44.062	12.547	South-east. Romagna	144	8.0
1882	8	16				42.979	13.875	Grottammare	13	7.0
1887	2	23	5	21	50	43.891	7.992	Western Liguria	1517	9.0
1889	12	8				41.830	15.692	Apricena	122	7.0
1896	10	16				43.909	7.872	Albenga	60	6.0
1905	9	8	1	43	11	38.670	16.068	Calabria	827	10.0
1916	5	17	12	50		44.010	12.623	North. Adriatic Sea	132	8.0
1916	8	16	7	6	14	43.961	12.671	North. Adriatic Sea	257	8.0
1917	11	5	22	47		43.506	13.586	Numana	26	6.0
1919	10	22	6	10		41.462	12.637	Anzio	138	6.5
1924	1	2	8	55	13	43.736	13.141	Central Adriatic Sea	76	7.5
1926	8	17	1	42		38.567	14.825	Salina island	44	7.0
1930	10	30	7	13		43.659	13.331	Senigallia	263	8.0
1972	1	18	23	26		44.203	8.163	West. Ligurian Coast	41	6.0
1972	1	25	20	25	11	43.614	13.355	Central Adriatic Sea	24	6.0
1972	2	4	2	42	53	43.590	13.295	Central Adriatic Sea	75	7.5
1972	2	4	9	19	4	43.589	13.283	Central Adriatic Sea	56	7.5
1972	2	29	20	54		41.841	15.459	South. Adriatic Sea	21	6.0
1972	6	14	18	55	46	43.580	13.416	Central Adriatic Sea	17	8.0
1978	4	15	23	33	47	38.125	15.022	Gulf of Patti	333	8.0
1982	3	21	9	44	2	40.008	15.766	Gulf of Policastro	126	7.5
1984	4	22	17	39	21	43.617	10.313	Leghorn	39	6.0
1990	12	13	0	24	28	37.266	15.121	South-Eastern Sicily	304	7.0
2002	9	6	1	21	29	38.081	13.422	Palermo	132	6.0

Table 2.1: Earthquakes excluded from computation because the epicenter is known or is supposed to be located in the sea.

As the aim of this part of the work is to define an attenuation law that represents the intensity attenuation properties for the whole Italian territory, the earthquakes that occurred at the volcanic areas of Mt. Etna and Ischia island have been excluded from the data base. Both areas in fact are known to be strongly attenuating, due to their volcanic nature (Carletti and Gasperini, 2003; Azzaro et al., 2006). Two events that occurred at Ischia island in 1881 and 1883, as well as all earthquakes located within a radius of 25 km from the summit of Mt. Etna (Lat=37.73, Lon=15.00) are excluded from the database.

The application of all these selection criteria further reduces the number of earthquakes to about 470 and the number of intensity data that are used in computations to 39000.

The intensity attenuation modeling assumes a spherical radiation from a point-form source of seismic energy. This could not be valid, especially for seismic sources of magnitude larger than 5, which have fault lengths larger than 5 km (Wells and Coppersmith, 1994) and could lead to a overestimation of the distance from the source. A possible choice is to discard intensity data that are closer than a minimum threshold D_{min}. Gasperini (2001) computes D_{min} as a function of the macroseismic moment magnitude while Albarello and D'Amico (2004) assumes $D_{min} = 15$ km for all earthquakes. Although the method by Gasperini (2001) might appear more accurate, it could bias the analysis since it selects the data points as a function of earthquake magnitude. As a result, only the data coming from small earthquakes would be selected at short distances, and thus the estimated attenuation properties in this range of distances would only reflect the attenuation behavior for such earthquakes. On the other hand, the assumption of a fixed threshold results in near-source data being discarded completely, while the prediction of intensity in the vicinity of the sources is crucial in hazard assessment. On the basis of these considerations, I assume initially that $R_{min} = 0$ (i.e., no data are discarded). I will discuss in section 2.8 the effects of applying different thresholds.

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 large 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 papers) Gasperini (2001) and by Carletti and Gasperini (2003) is to exclude all the data (independently of the ob-

(a) original   (b) selection
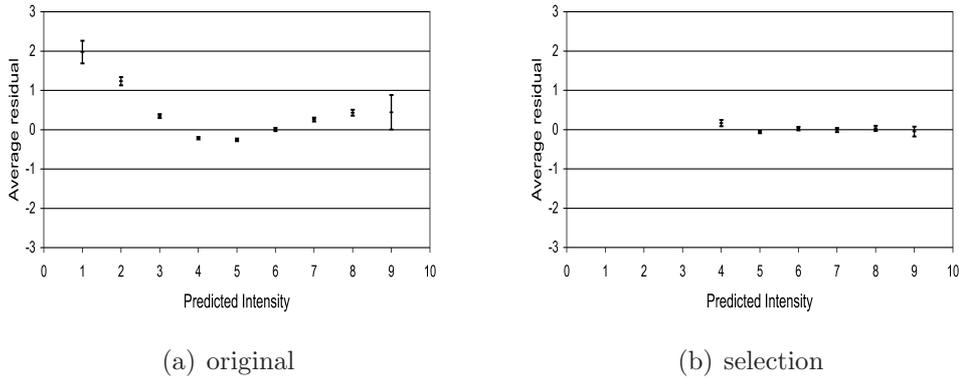
Figure 2.1: Intensity average residuals of the Albarello and D'Amico (2004) attenuation relationship for classes of predicted intensity. Error bars indicate 95% confidence intervals.(a) original data set.(b) Selection rule proposed by Gasperini (2001) applied to the data. Error bars indicate 95% confidence intervals.

served intensity) located at hypocentral distances D for which

$$I_0 - 0.53 - 0.055\min(D, 45) - 0.022\max(D - 45, 0) < 4 \qquad (2.26)$$

Note that this selection criterion does not introduce a bias in the attenuation regression because it applies to hypocentral distances (the independent variable) and not to observed intensities (the dependent variable).

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. These are plotted in figure 2.1. For the relation and dataset used by Albarello and D'Amico (2004), it can be noted in figure 2.1.(a) 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, the positive average residuals tend to increase outside this range. Moreover residuals are slightly negative 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 square 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 in eq. 2.26 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 (figure 2.1.(b)). This confirms that data for distances from the source at which intensity lower than IV are expected are incomplete and have to be excluded from the dataset to avoid an overestimation of predicted intensity at large distances.

The above mentioned selection criteria reduce the data set to about 21930 intensity values, referring to 313 earthquakes.

2.4.3 Intrinsic variability of data set

Felt intensity at a certain locality mainly depends on source distance and strenght. However, further variations could be induced by a number of additional factors, such as near-source radiation pattern, local amplification of seismic motion induced by the local stratigraphy and topography, and regional differences in the energy propagation pattern. This implies that whatever isotropic attenuation model is introduced, part of the variability in intensity remains unexplained and that a lower bound for the model variance exists. It does not depend on the specific attenuation model considered and could be called "intrinsic" variability. To estimate it, intensity observations relative to each m-th earthquake are grouped in contiguous 5 km bins of epicentral distance. For each l-th bin, standard deviation σ_l^m with respect to the average intensity μ_l^m is computed. To be consistent with the likelihood formulation described in section 2.1, μ_l^m and σ_l^m are computed by maximizing eq. 2.9, where μ_l^m and σ_l^m take the place of σ and μ_j respectively. Only bins characterized by at least 10 observations were considered. The intrinsic standard deviation relative to each distance bin for the entire intensity database was computed as

$$\sigma_l^{intr} = \sqrt{\frac{\sum_{m=1}^{N_j} (\sigma_l^m)^2 M_l^m}{\sum_{m=1}^{N_j} M_l^m}} \qquad (2.27)$$

where M_l^m is the number of intensity observations in the l-th distance bin for the m-th earthquake, and N_l is the number of earthquakes with at least 10 observations in the l-th distance bin. The results of these computations are reported in figure 2.2. The intrinsic standard deviation is nearly constant up to about 140 km from the source and its value
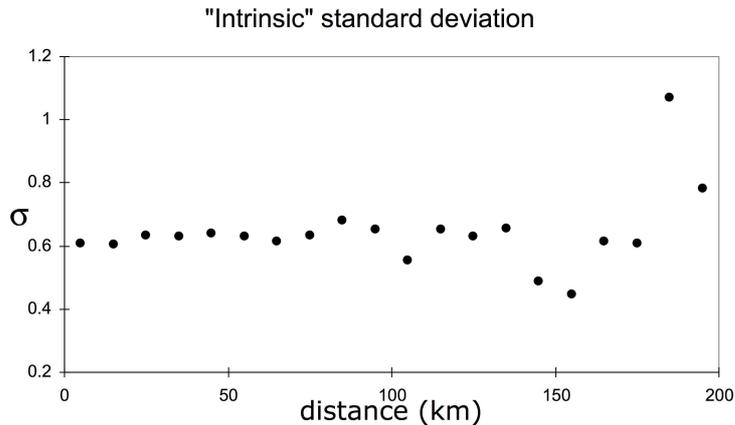
Figure 2.2: Standard deviation of observed intensities in each distance bin.

is close to about 0.62 intensity degrees. It is worth noting that such a value for the standard deviation is much lower than those relative to empirical intensity relationships so far computed (Gasperini, 2001; Albarello and D'Amico, 2004), which in all cases are greater than 1.0. This difference indicates that such models do not represent optimally the intensity attenuation in Italy.

2.5   Regression results

The data set is analyzed following the procedure described above and using the initial choices for arbitrary parameters $N_{min} = 10$, $R_{min} = 0$ km , $R_{max} = 300$ km, $w_1 = 0.5$. The maximization of eq. 2.9 with $\mu_j = \mu(\overline{I}_m, D)$ from eq. 2.17 gives $a = -0.0086 \pm 0.0005$ for the coefficient of the linear distance term, $b = -1.037 \pm 0.027$ for the coefficient of the natural logarithm of distance and $h = 3.91 \pm 0.27$ km for the average depth. Hence, the attenuation law (eq. 2.17) becomes

$$\mu(I_E, D) = I_E - 0.0086(D - 3.91) - 1.037(\ln D - \ln 3.91) \qquad (2.28)$$

with $D = \sqrt{R^2 + 3.91^2}$. The corresponding model standard deviation is $\sigma = 0.69$, close to the estimate of the intrinsic one determined above (see section 2.4.3).

The variance explained by the model (R^2) is 0.656. The errors reported above along with the parameters estimations are evaluated as the square root of the diagonal elements of the variance-covariance matrix, approximated as the inverse of the Hessian at the maximum of the log-likelihood function. Such matrix is compared with the variance-covariance matrix estimated by bootstrap (section 2.3.1) in table 2.2.

(a)	a	b	h	(b)	a	b	h
a	$2.35\,10^{-7}$	$-1.19\,10^{-5}$	$8.99\,10^{-5}$	a	$3.07\,10^{-7}$	$-1.54\,10^{-5}$	$12.2\,10^{-5}$
b	$-1.19\,10^{-5}$	$7.11\,10^{-4}$	$-6.04\,10^{-3}$	b	$-1.54\,10^{-5}$	$9.23\,10^{-4}$	$-8.25\,10^{-3}$
h	$8.99\,10^{-5}$	$-6.04\,10^{-3}$	$7.31\,10^{-2}$	h	$1.22\,10^{-4}$	$-8.25\,10^{-3}$	$10.0\,10^{-2}$

Table 2.2: Variance-covariance matrix estimated by means of the inverse of the Hessian (a) and bootstrap (b).

	a	b	h
a	0.77	0.77	0.74
b	0.77	0.77	0.73
h	0.74	0.73	0.72

Table 2.3: Ratios between the elements of the variance-covariance matrix showed in table 2.2 (a) and 2.2 (b).

The bootstrap estimates are obtained from 1000 paradata sets obtained by random resampling (with replacement) of 21932 intensity values from the original data set. In table 2.3 the two matrices are compared by computing the ratios between their corresponding elements. The values obtained indicate that the covariance elements from the Hessian underestimate those provided by the bootstrap analysis by a factor of 20-30%. This corresponds to an underestimate of parameter errors of the order of 10-15%.
The off-diagonal terms of the variance-covariance matrix indicate the presence of a significant multicollinearity. This can also be evaluated considering the distribution of parameter values obtained from bootstrap analysis (figure 2.3). A strong inverse correlation between the distance coefficients a and b as well as between b and h is evidenced, while the correlation between a and h is weaker. The high correlation among the parameters indicates that different combinations of their values might result in very similar attenuation equations. The high correlation could also indicate that small differences in the data set will result in apparently strong variations of the empirical parameters, despite the fact that the intensities expected by the attenuation laws corresponding to such parameters values will assume quite similar values.

2.6 Relations between I_E and other earthquake size measures

For earthquakes included in the data set, the value of I_E can be computed simply according to eq. 2.21. To compute it for any other earthquake
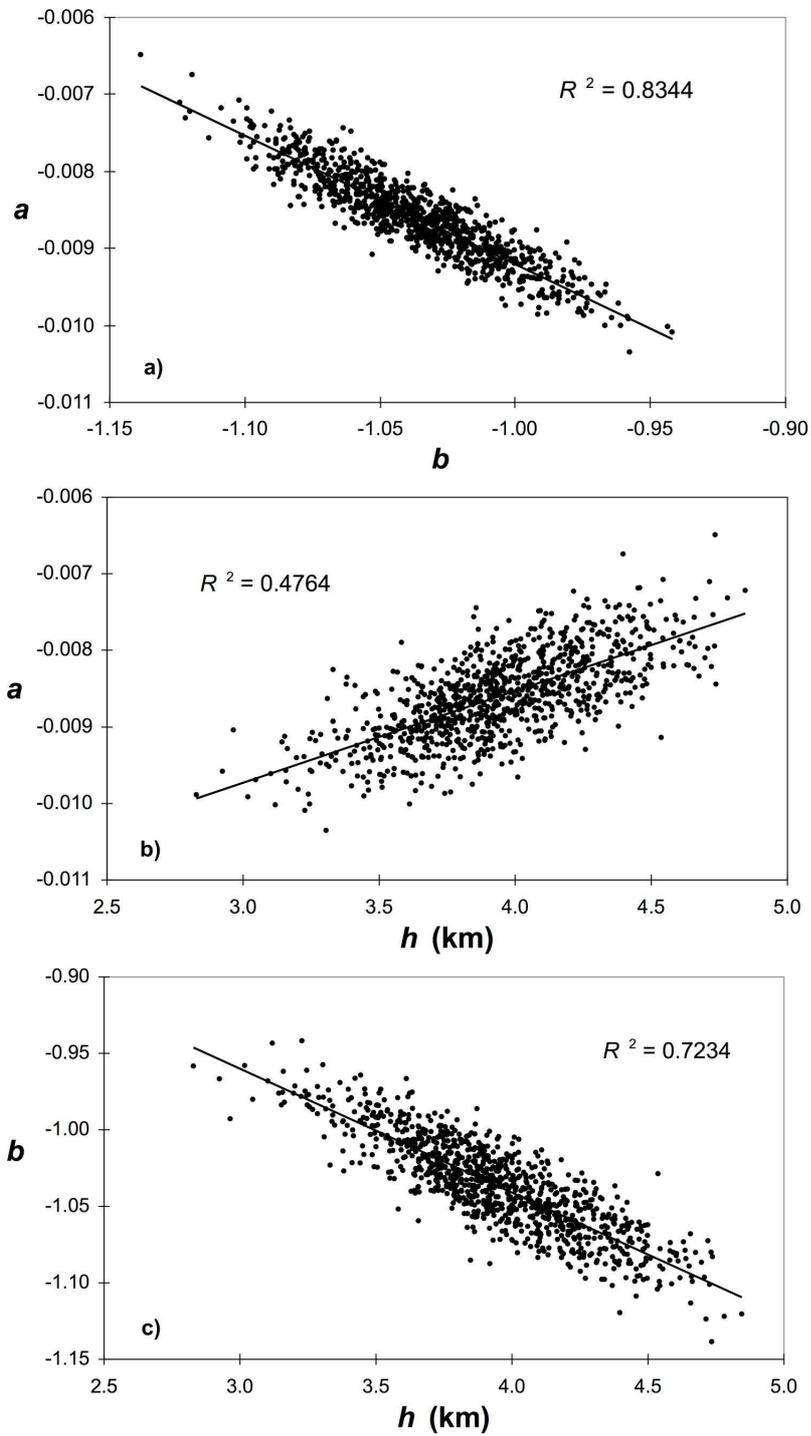
Figure 2.3: Distribution of parameter values obtained from 1000 bootstrap repetitions. This highlights the high correlation between parameters. The explained variance R^2 is reported in each plot.

in the catalog, empirical relation between I_E and other parameters that describe earthquake size could be calculated. For each earthquake of the CPTI04 catalog I_{max}, I_0 and M_{aw} are available. The first is the maximum intensity observed for a certain earthquake, the second is the epicentral intensity (see section 2.4.1) and the magnitude M_{aw} results from combining the instrumental and macroseismic information. Moreover a reduced set of moment magnitudes M_{sw}, which were either measured directly by moment tensor inversion or deduced from other instrumental magnitudes using empirical conversion rules (Gasperini, 2004) is considered. Figure 2.4 and tables 2.4 and 2.5 show the results of ordinary least squares (OLS) and general orthogonal regressions (GOR) of I_E vs I_{max}, I_0, M_{aw} and M_{sw}. For the GOR, the procedure described by Fuller (1987) and applied by Castellaro et al. (2006) is followed.

equation	η	c	d	σ
$I_E = c + dI_0$	0.09	-0.893 ± 0.254	1.118 ± 0.033	0.70
$I_E = c + dI_{max}$	0.09	-1.418 ± 0.289	1.154 ± 0.036	0.75
$I_E = c + dM_{aw}$	0.46	-5.862 ± 0.301	2.460 ± 0.055	0.53
$I_E = c + dM_{sw}$	0.31	-5.230 ± 0.645	2.210 ± 0.122	0.66

Table 2.4: Coefficients of regression of I_E with I_0, I_{max}, M_{aw} and M_{sw} from general orthogonal regression (GOR). η indicates the assumed variance ratio for each orthogonal regression (see Castellaro et al. (2006)).

equation	c	d	σ
$I_E = c + dI_0$	0.791 ± 0.444	0.897 ± 0.057	0.63
$I_E = c + dI_{max}$	0.591 ± 0.464	0.898 ± 0.058	0.67
$I_E = c + dM_{aw}$	-4.157 ± 0.723	2.150 ± 0.131	0.50
$I_E = c + dM_{sw}$	-1.343 ± 0.877	1.621 ± 0.166	0.58

Table 2.5: Coefficients of regression of I_E with I_0, I_{max}, M_{aw} and M_{sw} from ordinary least squares (OLS).

The GOR allows to consider that even the independent variable is affected by errors. To apply this method the ratio between the variances of dependent and independent variables (η) has to be defined. To estimate such ratios the square of half a degree (corresponding to the resolution of the macroseismic scale) is assumed as variance of both I_{max} and I_0. The variance of I_E is assumed to be 0.0225 (corresponding to the average of the empirical variances computed for each earthquake in the data set). The variances of M_{aw} and M_{sw} (0.0484 and 0.0729 respectively) are computed by averaging the squares of the estimation errors reported in the CPTI04 catalog for each earthquake. Figures 2.4 shows that the correlation of I_E vs I_0 is slightly better than that vs I_{max}. The regression slopes are very similar, while there is a difference of about half of one degree
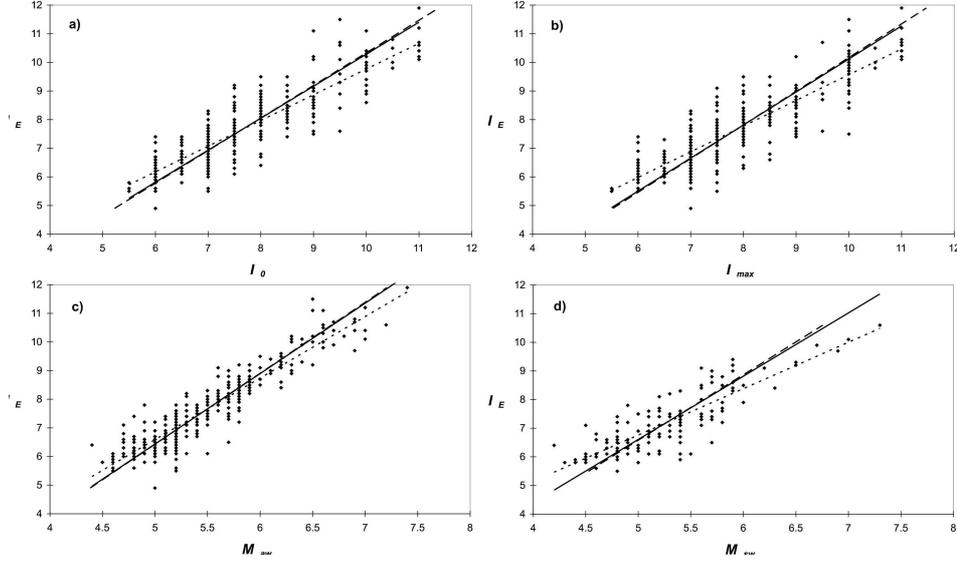
Figure 2.4: Empirical relations of I_E with I_0, I_{max}, M_{aw} and M_{sw}. Solid lines indicate the result of GOR regression, dotted and dashed lines indicate respectively results of the direct and inverse OLS regressions.

between the intercepts (Table 2.4 and 2.5). On average, I_E values result slightly larger than those of I_0 for high intensities and slightly lower than those of I_0 for low ones. In spite of this, there is a substantial coincidence among them, with the average difference being lower than a tenth of a degree. Concerning the relations involving magnitudes, in Table 2.4 and 2.5 the regression slopes for the two magnitudes M_{aw} and M_{sw} are quite different. This might appear surprising, given that M_{aw} (which is based both on instrumental and macroseismic information) has been calibrated on essentially the same set of M_{sw} instrumental magnitudes (Gasperini, 2004). However, it could be noted that such calibration has been made by weighting each instrumental magnitude estimate with the inverse of the respective squared error. This procedure gives higher weights to moment magnitudes computed by the inversion of complete seismograms and lower weights to instrumental magnitudes of the first half of the 20-th century (mainly M_s) computed by maximum amplitudes measured on historical seismograms. The calibration of mechanical instruments used in Italy up to about 1980 is doubtful. Hence the M_{aw} data set will be generally more homogeneous and reliable than the M_{sw}. This implies that M_{aw} values are to be preferred for computing I_E.

I also test the alternative option of computing magnitude coefficients at the same time as the distance terms by an OLS regression of the

equation

$$\mu(M, D) = c + dM + a(D - h) + b(\ln D - \ln h) \qquad (2.29)$$

Table 2.6 shows how the values of depth (h) and distance coefficients (a, b) are very close to those deduced from two-step regressions (section 2.5), while the magnitude coefficients (c, d) differ significantly from those reported in table 2.5 for the OLS method (the results from GOR analysis are not comparable, so they were not considered). Such differences are

	M_{aw}	M_{sw}
c	-3.062 ± 0.070	-1.147 ± 0.096
d	1.931 ± 0.010	1.567 ± 0.012
a	-0.0126 ± 0.0005	-0.0104 ± 0.0007
b	-0.900 ± 0.027	-0.912 ± 0.039
h	3.696 ± 0.302	4.155 ± 0.511
σ	0.81	0.79

Table 2.6: Simultaneous regression of distance dependence coefficients and magnitudes.

due to the different weighting of data. In fact, in the two-step regression, all earthquakes have the same weight while, in the regression of eq. 2.29, the weight of each earthquake is proportional to the number of observations. To validate this statement, I performed the regression of I_E vs. magnitude using independent estimates of I_E made from individual intensity observations (tables 2.7 and 2.8). These can be computed from equation

$$I_E = I_S - a(D - h) - b(\ln D - \ln h) \qquad (2.30)$$

where I_S is the individual intensity observed at a distance D from the macroseismic hypocenter located at a depth h and parameters a, b and h are those of eq. 2.28. The results of OLS regression (table 2.8) show values of c and d coefficients close to those found by simultaneous regression. This means that, when the data are similarly weighted, two-step and one-step standard regressions give consistent results.

Anyway, as GOR is more appropriate than OLS when the independent variable is affected by errors and a simultaneous GOR of distance and magnitude is not feasible, the procedure based on a two-step regression for distance and magnitude is preferable to one based on a one-step regression. Moreover regression coefficients in table 2.4 should be used, as they are not biased by the uneven distribution of intensity observations among different earthquakes (instead the coefficients of table 2.4 are characterized by this drawback). For the earthquakes in the CPTI04 catalog for which the M_{aw} magnitude is simply deduced from I_0 (according to regressions computed by Rebez and Stucchi (1999) and Gasperini

equation	c	d	σ
$I_E = c + dI_0$	0.068 ± 0.023	0.976 ± 0.003	0.56
$I_E = c + dI_{max}$	-0.451 ± 0.026	1.019 ± 0.003	0.59
$I_E = c + dM_{aw}$	-3.972 ± 0.028	2.095 ± 0.005	0.45
$I_E = c + dM_{sw}$	-2.129 ± 0.034	1.774 ± 0.006	0.46

Table 2.7: Coefficients of regression of I_E with I_0, I_{max}, M_{aw} and M_{sw} from general orthogonal regression (GOR), with one observation for each intensity datum.

equation	c	d	σ
$I_E = c + dI_0$	1.151 ± 0.055	0.851 ± 0.006	0.53
$I_E = c + dI_{max}$	0.796 ± 0.058	0.878 ± 0.006	0.55
$I_E = c + dM_{aw}$	-2.868 ± 0.082	1.910 ± 0.014	0.44
$I_E = c + dM_{sw}$	-1.055 ± 0.088	1.592 ± 0.015	0.44

Table 2.8: Coefficients of regression of I_E with I_0, I_{max}, M_{aw} and M_{sw} from ordinary least squares (OLS), with one observation for each intensity datum.

(2004)), to avoid the double conversion from I_0 to M_{aw} and from M_{aw} to I_E it is preferable to compute I_E directly from the regression with I_0.

The total standard errors of separate regressions can be obtained, following the usual error law, by summing the relevant variances of the involved empirical relationships. In the case that I_E is computed (by GOR regression) from I_0 and M_{aw} they are 0.98 and 0.87, respectively. These standard errors are larger than the one computed for the two-step procedure ($\sigma = 0.69$).

2.7   Other attenuation models

2.7.1   Alternative funcions to model the distance dependence

Even if the log-linear model is physically grounded and widely used in attenuation studies, it is interesting to test its performance in describing attenuation pattern in comparison with other functional forms of attenuation law. I apply the same two-step procedure and the same choices to fit the bilinear, logarithmic and cubic root model (section 2.2.3). Moreover, I consider a bilinear model with an added logarithmic term to account for geometrical spreading phenomenon. To follow the two step procedure, the attenuation law are rewritten as

bilinear (BIL)	**logarithmic (LOG)**
$\mu(I_E, D) = I_E + a[\min(D, 45) - \min(h, 45)] + a'[\max(D - 45, 0) - \max(h - 45, 0)]$	$\mu(I_E, D) = I_E + a[\ln D - \ln h]$
cubic root attenuation model (CRAM)	**log-bilinear(BIL-LOG)**
$\mu(I_E, D) = I_E + a[D^{\frac{1}{3}} - h^{\frac{1}{3}}]$	$\mu(I_E, D) = I_E + a[\min(D, 45) - \min(h, 45)] + a'[\max(D - 45, 0) - \max(h - 45, 0)] - b\ln D - lnh$

The attenuation laws obtained are compared using the criteria illustrated in section 2.3.2. The results are shown in table 2.9. The log-linear model

model	a	a'	b	h
BIL	-0.591 ± 0.0005	-0.0182 ± 0.0003	-	0.00 ± 0.24
LOG	-	-	-1.54 ± 0.15	7.77 ± 0.25
CRAM	-1.2432 ± 0.076	-	-	1.18 ± 0.25
LOG-LIN	-0.0086 ± 0.0005	-	-1.04 ± 0.27	3.91 ± 0.27
BIL-LOG	-0.0187 ± 0.0026	-0.0108 ± 0.0007	-0.80 ± 0.06	2.78 ± 0.40
model	σ	R^2	BIC	AIC_C
BIL	0.7023	0.6435	-26105.1	-26097.8
LOG	0.6943	0.6515	-25898.3	-25893.0
CRAM	0.6896	0.6562	-25785.0	-25779.8
LOG-LIN	0.6894	0.6565	-25781.2	-25773.8
BIL-LOG	0.6891	0.6567	-25780.7	-25771.3

Table 2.9: Comparison between different functional forms of attenuation relation, sorted for increasing BIC. The values of AIC_C agree with BIC.

performs better than other models previously used to describe intensity attenuation in Italy for all the considered quality factors (σ, R^2, BIC and AIC_C). This is in contrast with the results of a similar comparison made by Gasperini (2001) who, when considering the averages of intensity differences $\Delta I = I_0 - I$ over distance bins of 5 km (and not single-intensity observations), found a clear preference for the bilinear model with respect to the log-linear one. In fact, while such a discrepancy could be attributed to the use of binning, it could also be attributed to the near source portion of the intensity data set (excluded from computation by Gasperini (2001)) where the logarithmic term, which accounts for geometrical spreading (neglected in that intensity attenuation model), assumes a crucial role. In fact, if a logarithmic term is added to the bilinear model (BIL-LOG in table 2.9), the fit becomes slightly better than the log-linear one. This could confirm the inference made by Gasperini (2001) that anelastic dissipation properties are probably depth-dependent. However, the improvement of the fit of the log-bilinear model with respect to the simpler log-linear one is so small that in most cases, the complications in hazard computational procedures that its adoption

would seem to imply are not justified. For such reasons, the simple log-linear model is preferable to describe intensity attenuation in hazard assessment.

2.7.2   Alternative measure of earthquake size

In this work I choose to calculate epicentral intensity as the intensity expected at the epicenter by the attenuation law. In particular, the procedure followed assumes that the average of the intensities observed within a certain distance from the epicenter of an earthquake coincides with the intensity expected by the attenuation law at the corresponding average distance.

Even if this choice is suggested by reasonable considerations, it would be interesting to check the performance of the resulting attenuation law with respect to the law that would be obtained simply imposing that the intensity predicted at the epicenter is equal to the epicentral intensity reported by CPTI04. In this case the attenuation law becomes

$$\mu(I_0, D) = I_0 + a(D - h) + b(\ln D - \ln h), \qquad (2.31)$$

where I_0 is the epicentral intensity as reported in CPTI04. In table 2.10

a	b	h	σ	R^2	BIC	AIC_C
-0.0135 ± 0.0005	-0.87 ± 0.02	2.83 ± 0.13	0.89	0.43	-30428.4	-30421.0

Table 2.10: Attenuation law parameters and quality factors using epicentral intensity reported by CPTI04.

the parameters values, along with σ, BIC, AIC_C and R^2 are given. The values of all the four quality factors (higher σ, lower BIC, AIC_C and R^2) indicate that the data fitting is worst than in the case of eq. 2.28.

Another interesting aspect to consider is the possible presence of residuals trends with respect to dependent and independent variables (figure 2.5). Residuals almost horizontal and close to zero are desirable, whereas deviation from this pattern may indicate a poor data fitting or problems in modeling (Draper and Smith, 1981). Looking at the plot of residuals versus distances (figure 2.5, above) it can be noted that for eq. 2.28 they are almost horizontal whereas for the last attenuation law considered their deviations from zero are slightly greater. The plot of residuals versus intensity expected by attenuation law (figure 2.5, below) is almost horizontal until the 10-th degree for the preferred attenuation law, whereas departure from horizontal with underestimation at lower degrees and overestimation at higher degree are evident for the model described by eq. 2.31.
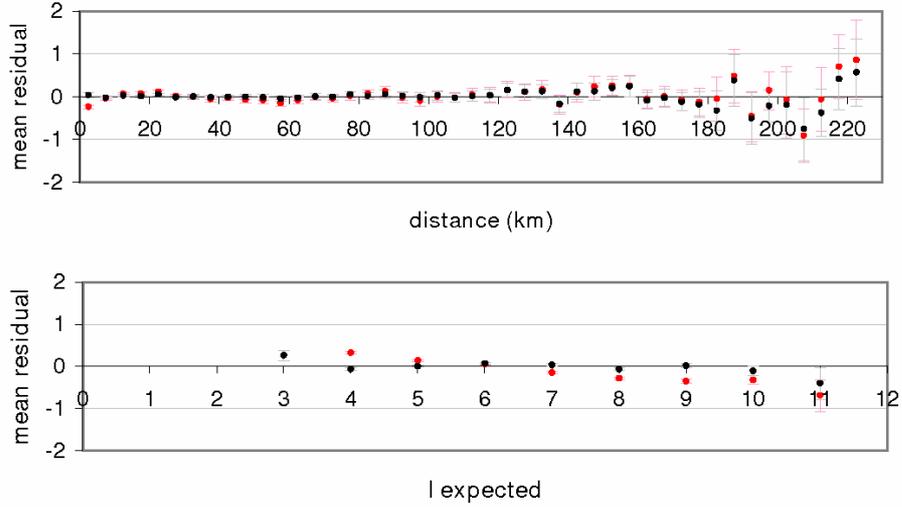
Figure 2.5: Residual versus distance (independent variable) and expected intensity (dependent variable) for attenuation law calculating epicentral intensity from the data set (black) or using epicentral intensity reported in CPTI04 (red).

These results confirm the choice of calculate the epicentral intensities directly from the data set, without using those reported in CPTI04.

2.8 Sensitivity analysis

The results shown in section 2.5 may depend on the choices of four arbitrary parameters. In fact, the values of $\overline{I}_m$, $\overline{D}_m$ and $\overline{\ln D}_m$, which are used in the first step of the inversion procedure, depend on the value of R_{max}, that is the maximum distance from the epicenter of intensity data considered for computing such averages. Moreover, the intensity observation that are actually used for computations are selected on the basis of the values of R_{min} and N_{min}, i.e., the minimum epicentral distance, and the minimum number of intensity data for each event to be considered in the analysis, respectively. Finally, different choices might be made for the weight w_1 that parametrize uncertainty on intensity data (see section 2.1.2). In order to quantify the effects of these choices, a sensitivity analysis is performed. In particular, different values of R_{max}, N_{min}, R_{min}, and weight w_1 are considered and for each of them the likelihood function (eq. 2.9) is maximized to obtain different parametrization of eq. 2.17.

Figure 2.6 shows that for R_{max} larger than 30 km, a and σ are almost insensitive to variations in R_{max}. Parameter b stabilizes when

Sensitivity of attenuation parameters to R_{max}
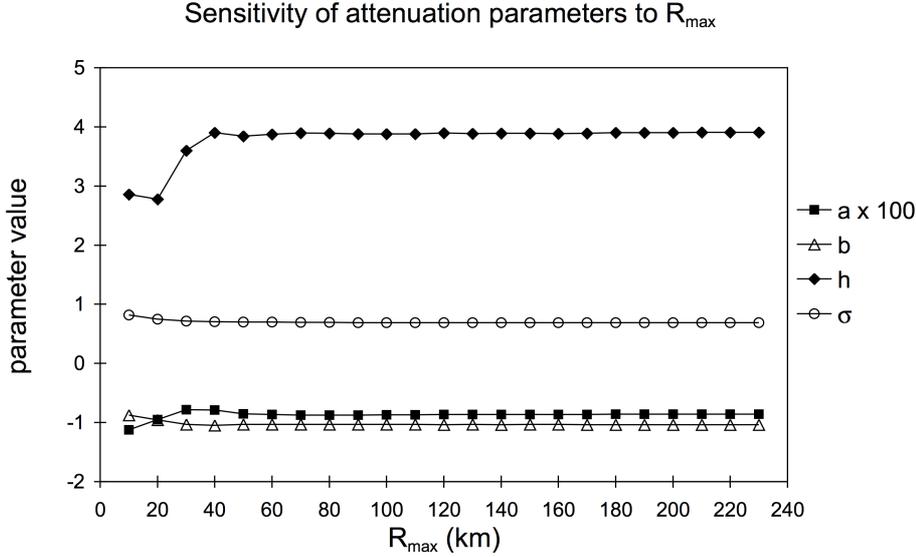
Figure 2.6: Sensitivity of attenuation parameters to R_{max} (the maximum epicentral distance of data used in the calculation of $\overline{I}_m$, $\overline{D}_m$ and $\overline{\ln D}_m$).

$R_{max} > 50$ and in general shows relatively small sensitivity to R_{max}. A larger sensitivity is shown for h. In particular, h increases monotonically for $R_{max} < 50$ km, then it assume a steady value of about 4 km. Similar results are obtained when different choices are considered for parameter N_{min} that is set to 10 in the analysis discussed in section 2.5 (figure 2.7). In this case also, h shows the largest sensitivity for $N_{min} < 60$. Whereas the other parameters values are quite stable for any considered choice of N_{min}.

In figure 2.8, we can see that h also shows the highest variability with respect to R_{min}. In fact, the value of h decreases strongly with increasing R_{min} up to 20 km. This means that the average depth is controlled by the data in the vicinity of the epicenter and that the exclusion of such data makes the depth estimate unrealistic (< 1 km). It is also interesting to note that if the data at distances shorter than 90 km are excluded from computations, the linear term (coefficient a) becomes dominant over the logarithmic one (coefficient b). It may be inferred that the data in the vicinity of the source, although they are possibly biased by near-source anisotropy for strong events, are still necessary to describe the overall attenuation behavior realistically. These results indicate that h is the parameter most sensitive to the data selection strategy, particularly with respect to data in the epicentral area. However, it is worth noting that the standard deviation of the model is almost unaffected by variations in h.

Figure 2.9 shows that the regression results are almost insensitive to

Sensitivity of attenuation parameters to N$_{min}$
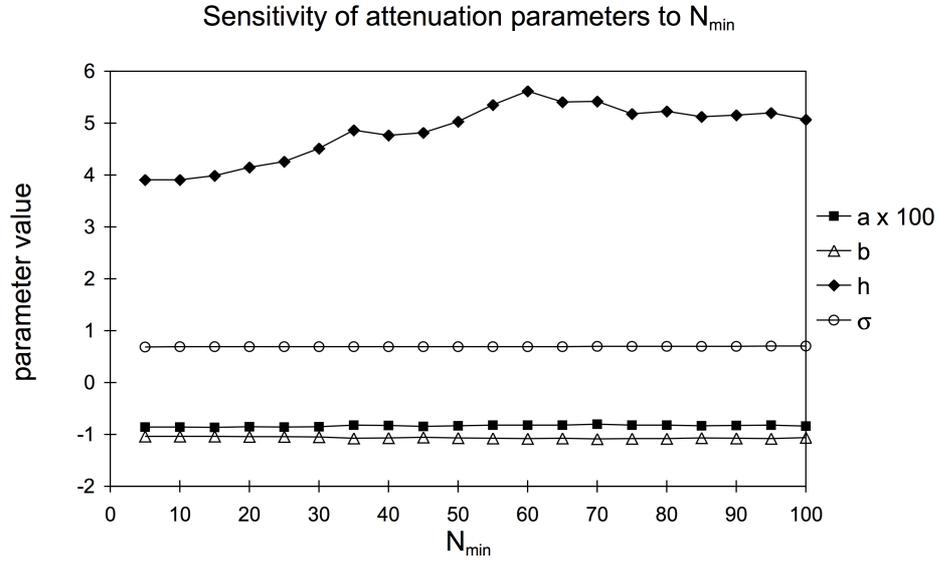
Figure 2.7: Sensitivity of attenuation parameters to N_{min} (the minimum number of intensity data for each earthquake included in the fit).

Sensitivity of attenuation parameters to R$_{min}$
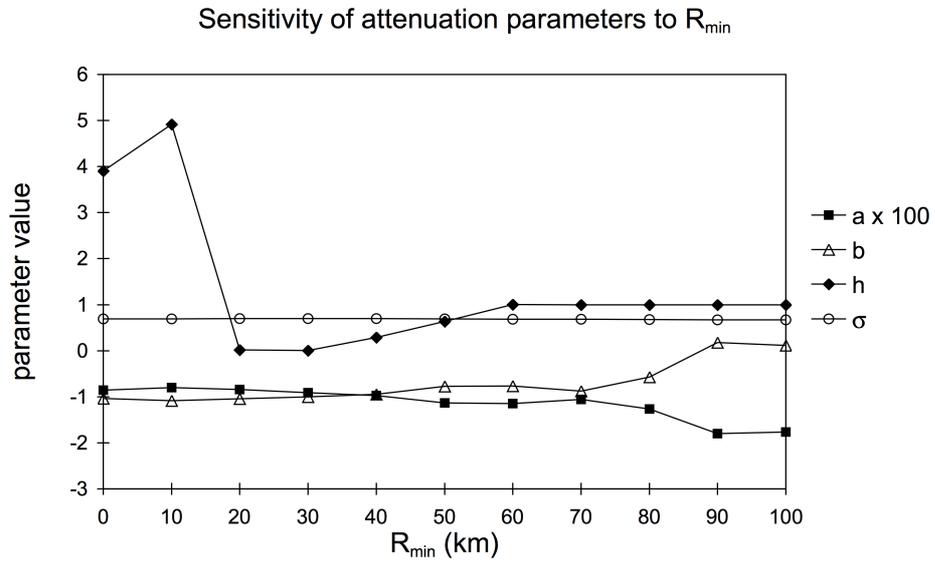
Figure 2.8: Sensitivity of attenuation parameters to R_{min} (the minimum epicentral distance of data used for the fit).
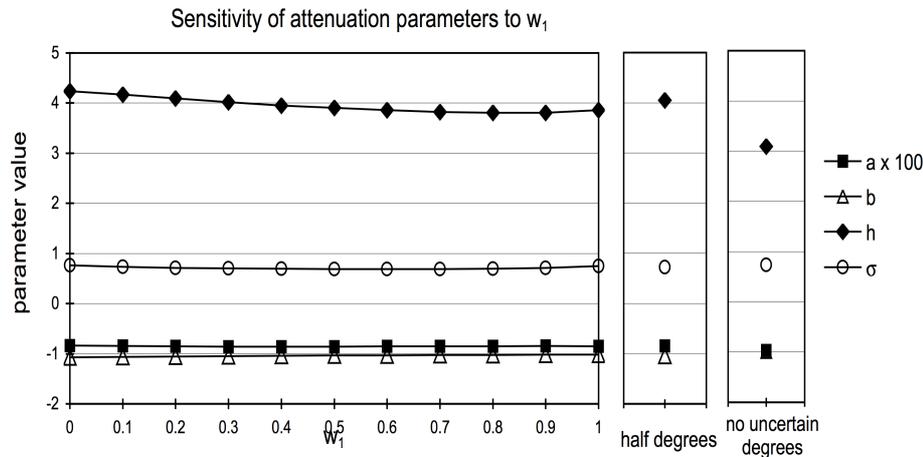
Figure 2.9: Sensitivity of attenuation parameters by varying w_1 (the probability assigned to the nearer lower degree) (left), attributing to uncertain degrees the corresponding real value (center) and excluding uncertain degrees from the data set (right).

the choice of the weight w_1. This quantity parameterizes the probability distribution of uncertain intensity values and represent respectively the probability of the nearer lower and higher integer degrees (see section 2.1.2). In figure 2.9 ("half degrees" inset), the parameter values obtained following Gasperini (2001) are also represented. In this case uncertain degree are considered simply as representative of real intermediate intensity values (i.e. VII-VIII equal to 7.5). The results are affected negligibly and are very similar to those obtained using equal weights ($w_1 = 0.5$). This might indicate that uncertain values are probably close to representing ground motion levels intermediate to those associated to adjacent integer intensities. In addition, the results are very similar ("no uncertain degrees" inset in figure 2.9) except with respect to h, even if the uncertain data are totally ignored in the fit. This similarity indicates a substantial coherence of attenuation properties for uncertain degrees with respect to standard ones.

2.9   Discussion

In the precedent sections the attenuation pattern of macroseismic intensity in Italy has been analyzed and modeled. The main objective was to develop an attenuation relationship for the probabilistic seismic hazard assessment in Italy in terms of macroseismic intensity. Thus, major attention has been paid to the characterization of the attenuation relationship in its complete probabilistic form. The statistical analysis of intensity was carried out by carefully considering several factors that have

generally been overlooked in previous analyses: the discrete character of intensity, uncertainty on original data, completeness, etc.

On the assumption that the general attenuation pattern reflects mainly the seismic energy radiation pattern, the local intensity value has been considered as being dependent on two variables: an expression of the energy radiated at the source and the hypocentral distance (computed assuming a unique hypocentral depth for all the events). The first variable appears to be the most problematic because it is not possible to directly estimate such a parameter from the available documentary data. In the present work this parameter is estimated from the whole macroseismic field available for each earthquake. In this regard, one should be aware that the number of parameters actually involved in the inversion procedure is much higher than that were included explicitly in the attenuation relationship (the coefficients relative to the source strength, the geometrical spreading, the anelastic/scattering dissipation, the average hypocentral depth, and the standard deviation) because a source term relative to each of the events has to be included in the computation. In most previous analyses, this role was played by the epicentral intensity I_0, which is usually determined a priori by catalog compilers. By contrast, in this work, a new estimate of radiated energy in terms of the expected intensity at the epicenter I_E for all earthquakes is introduced. It is worth noting that although this parameter plays the same formal role in all attenuation equations, it cannot be considered simply as a new estimate of the epicentral intensity I_0, however I_0 is defined. In principle, I_0 should be an integer value, characterized in each case by an additional uncertainty (e.g. VII-VIII or 7.5), which could be measured directly if a suitable settlement, unaffected by anomalous site amplification, exists at the epicenter. By contrast, the source term introduced here is not constrained to be an integer ordinal value and can be considered as a macroseismic equivalent of magnitude, determined without any direct reference to instrumental parameters. Rather, it is deduced by considering the entire macroseismic field and not from a single or few intensity observations as was done in the case of I_0. That being so, it is less influenced by local site effects or intensity assessment errors. Regression analyses corroborate the feasibility of this interpretation, since a satisfactory correlation exists between this energy term and magnitudes. This last result also allows the attenuation relationship here defined to be applied in cases in which knowledge about the relevant macroseismic field is poor, cases that will include earthquakes known only in terms of instrumental parameters. In these cases, I_E can be computed from moment magnitude by using the conversion relationships here determined. Of course, the standard deviation relative to intensities estimated by using such a change of variable must be increased appropriately (by the

factor here estimated) to take into account the uncertainties of I_E vs. moment magnitude relationships.

A stability analysis shows that the resulting attenuation relationship is robust with respect to arbitrary assumptions underlying the modeling approach. In particular, the only parameter seriously affected by these arbitrary choices is the average source depth h, which is (weakly) constrained by the intensity data located in the vicinity of the source only. However, its variability has little influence on the other attenuation parameters and the standard deviation of the regression and thus is not relevant for computing seismic hazard. The stability analysis also showed that uncertain intensities (e.g. III-IV) behave similarly to standard intensity estimates and are consistent with semi-integer intensity values (e.g. 3.5). This could justify the use of semi-integer values in simplified approaches.

The attenuation model determined here enables a number of difficulties relative to previous estimates to be overcome. It is almost optimal, because the relevant standard deviation is comparable to the intrinsic one related to the scattering of original data (aleatory uncertainty). The intrinsic standard deviation is almost independent of the epicentral distance and represents a lower bound for any empirical attenuation relationship that, like the present one, does not consider any regionalization of the area under study, source directionality, or possible local effects. In general, the standard deviation associated with the attenuation relationship here yielded results lower than those obtained in previous works (Gasperini, 2001; Albarello and D'Amico, 2004). This could have important consequences in hazard estimates in the Italian region, because the standard deviation of the attenuation relationship is well known to affect hazard estimates dramatically (Cornell, 1991; Brillinger, 1982; Albarello and D'Amico, 2005). The best-fit values and standard errors of attenuation coefficients (-0.0086 ± 0.0005 for the linear term, -1.037 ± 0.027 for the logarithmic one, and 3.91 ± 0.27 for the average depth) suggest some further general considerations:

1. The inferred average source depth is significantly shallower than the average hypocentral depth of strong Italian earthquakes (about 10 km) but quite similar to that (5.6 km) deduced by the attenuation of PGA (Sabetta and Pugliese, 1987). This apparent discrepancy could be explained by considering that the stability analysis demonstrates that the average depth is constrained mainly by the data in the vicinity of the epicenter ($D < 20 - 30$ km). It may be argued that for relatively large sources (whose width is of the order of some km), the shallower (and closer) portion of the seismogenic fault contributes more than the deeper one in determining the ground motion level at close sites. This might indicates that the

source depth estimated from macroseismic data is likely to reflect a shallower point from which seismic energy appears to be radiated at close sites, rather than the true (and deeper) hypocenter as was assumed by past procedures for determining macroseismic depth (Koveslighety, 1906; Blake, 1941; Musson, 1996).

2. The value of the linear term coefficient significantly differs from 0 (one order of magnitude larger than the associated standard error). This indicates that the contribution of anelastic dissipation to intensity attenuation is not negligible with respect to geometrical spreading (it becomes of the order of one intensity degree at distances of 100 km).

3. The coefficient of the logarithmic term, close to -1 for the log-linear model and around -0.8 for the log-bilinear model, implies that the geometrical spreading exponent seen by seismic intensity should range from -0.70 to -0.35. These estimates for the geometrical spreading exponent depend on the attenuation model (log-linear or log-bilinear) and on the assumed coefficient of the linear relation between intensity and the logarithm of PGA, which, for Italy, could vary (in terms of natural logarithms) between the empirical estimate of 0.44 by Margottini et al. (1992) and the value of 0.69 assumed by the MCS intensity scale (Cancani, 1904; Sieberg, 1931). This confirms that surface waves (and perhaps reflected and refracted phases), rather than body waves, are likely to have a dominant role to play in determining the seismic intensity observed at a site.

Chapter 3

Seismic tomography

The study of physical systems often involves an inverse problem. This consists of using the actual result of some measurements to infer the values of the parameters that characterize the system from a given point of view. The results of the measurements are called data. The parameters chosen to characterize the system are called model parameters, their values are the unknown of the inverse problem. Actually, the relation between data and model parameters may be known only approximately, moreover data may be affected by errors. These facts imply that to obtain reliable estimates of model parameter the problem has to be overdetermined, i.e. the number of data has to be greater than the number of unknown model parameters.

In some cases, the inverse problem could be solved by tomographic method. This technique is used in a large number of disciplines including geophysics. It is an imaging technique that allows to reconstruct the internal structure of an object using data recorded on its surface. The first application of this method was the X-ray tomography in medicine. The idea at the base of radiological tomography was expressed by the central slice theorem by Radon (1917). It states that it is possible reconstructing a 2-D image from a set of 1-D lines integrals, and a 3-D image from 2-D slices trough the object. Cormack (1963) proposed a method for solving the radiological imaging problem and Houndsfield (1973) got the method to work by combining a computer and a X-ray scanner. In this system an X-ray beam, of known intensity originated by a source, follows a linear path inside the object ad arrives to a receiver on the opposite side of the object. Due to the absorption, the intensity recorded by the receiver will be reduced with respect to that of the beam before crossing the object. To infer the absorption structure different paths inside the studied object are illuminated by changing sources and corresponding receivers positions. The greater the number of rays crossing a certain zone in the object interior the more reliable the absorption estimate.

In the geophysical field, seismic tomography allows to reconstruct an image of the Earth interior using data recorded on (or near) its surface. Seismic waves arriving at a certain receiver contain information about the materials encountered along the path they followed to arrive from the source to the recording position. The inverse problem involved in seismic tomography is less well constrained by data than the X-ray tomography one. Source position and strength are not known precisely, seismic ray path is not straight and often depends on the properties of the materials it crosses. A common example of application of tomography in the seismological field is the study of the velocity structure using travel time data. As ray path depends on velocity properties of the materials where seismic waves propagate, usually a starting model, based on precedent knowledge of the investigated area, is assumed and used to calculate ray paths. Velocity variations with respect to such model are calculated by solving the tomographic problem. On the basis of the results, new ray paths are calculated and a new tomographic inversion is carried on. The procedure is repeated iteratively until the results differ significantly from those obtained in the previous inversion. As well as travel times, amplitude data allow a better knowledge of Earth interior. In particular they are useful to study the attenuation properties (Iyer and Hirahara, 1993). Even the propagation pattern of macroseismic intensity with distance may vary depending on the geostructural characteristics of the underlying materials and lends itself to be studied by tomographic method. In particular, the Italian territory presents strong lateral variations of its geostructural features that could originate evident variations in the propagation properties of seismic intensity.

3.1   Inverse problem solution

Generally the inverse problem underlying the seismic tomography is described by a system of linear equations that can be expressed trough matrices as

$$\mathbf{d} = \mathbf{Gm}\,, \qquad\qquad (3.1)$$

where $\mathbf{d}$ is the data vector, $\mathbf{m}$ is the model parameter vector, that represents the unknown of the tomographic problem and $\mathbf{G}$ is a matrix describing the linear relation between data and model parameters. If the problem is overdetermined (the number of data is greater than the number of unknown) and the uncertainties are Gaussian, the model parameters values can be estimated by least squares

$$\mathbf{m} = (\mathbf{G^T G})^{-1}\mathbf{G^T d}\,. \qquad\qquad (3.2)$$

This corresponds to find the unknown values that minimize the sum of the squared residuals that quantifies the deviation between observations and model predictions.

Even if the problem is overdetermined, the rays coverage of the interest area might be not uniform hence the values of the unknown parameters situated in zones crossed by a low number of rays would be less reliable than those referred to zones with an high ray coverage. So it is preferable to obtain the model parameters values by the damped least squares method

$$\mathbf{m} = (\mathbf{G^T G} + \lambda^2 \mathbf{I})^{-1} \mathbf{G^T}(\mathbf{d} - \mathbf{G m_0}) \, , \qquad (3.3)$$

where $\mathbf{m_0}$ is an *a priori* model summarizing the precedent information available about the phenomenon described by eq. 3.1 and λ^2 is a constant called damping parameter that represents the ratio between data and model parameters variances. When uncertainties of data and model parameters are Gaussian, the damped least square solution (eq. 3.3) minimizes

$$\mathbf{\Phi(m)} = (\mathbf{d} - \mathbf{G m})^{\mathbf{T}}(\mathbf{d} - \mathbf{G m}) + \lambda^2 (\mathbf{m} - \mathbf{m_0})^{\mathbf{T}}(\mathbf{m} - \mathbf{m_0}) \, . \qquad (3.4)$$

The first term of eq. 3.4 is the sum of the squared data residuals and corresponds to a measure of the distance between model prediction and data. The lower this term the better the ability of the model in explaining the data. The second term is a measure of the distance between model parameter estimates and the *a priori* model. The damping parameter λ determines the weight of this second term in the model parameters estimation. A large damping value will force the solutions to assume values close to the reference model $\mathbf{m_0}$, whereas low values will allow large deviations between solutions and $\mathbf{m_0}$.

Errors on parameter values can be estimated as the square roots of the corresponding diagonal elements in the covariance matrix. The latter can be evaluated as

$$\mathbf{C} = \sigma^2 (\mathbf{G^T G} + \lambda^2 \mathbf{I})^{-1}(\mathbf{G^T G} + \lambda^2)^{-1} \mathbf{G^T G} \, . \qquad (3.5)$$

where σ^2 is the variance in the data error and could be estimated as the sum of the squares of the residuals divided by the number of degree of freedom of the inversion problem (Aki and Lee, 1977)

$$\sigma^2 = \frac{(\mathbf{d} - \mathbf{G m})^{\mathbf{T}}(\mathbf{d} - \mathbf{G m})}{N_{data} - N_{parameters}} \, . \qquad (3.6)$$

3.2   Intensity tomography

A first intensity tomography for Italian territory was presented by Carletti and Gasperini (2003). They assumed the bilinear attenuation law proposed by Gasperini (2001) (eq. 2.14) as reference model and studied the lateral variations of the slopes b and c with respect to the reference model values.

The procedure presented in Chapter 2 allows to solve some drawbacks characterizing the previous attenuation laws presented by Gasperini (2001) and Albarello and D'Amico (2004). In particular the data fitting significantly improves (section 2.5) with respect to such works.

In this section the results obtained in Chapter 2 are used to define the *a priori* model and develop a new tomographic study of the propagation pattern of seismic intensity in Italy.

3.2.1   Formulation of the tomographic problem

First of all, it is necessary to chose an *a priori* model as starting point to develop the tomographic study.

In section 2.7 different intensity attenuation models are compared. On the basis of all the considered criteria (BIC, AIC_C, R^2, σ) the log-bilinear model is chosen

$$\mu(I_E, D) = I_E + a(\min(D, 45) - \min(h, 45))$$
$$+a'(\max(D - 45, 0) - \max(h - 45, 0))$$
$$+b(\ln D - \ln h), \qquad (3.7)$$

where $D = \sqrt{R^2 + h^2}$, $a = -0.0187 \pm 0.0026$, $a' = -0.0108 \pm 0.0007$, $b = -0.80 \pm 0.06$ and $h = 2.78 \pm 0.40$. In this relationship a logarithmic term accounting for geometrical spreading is added to a bilinear attenuation law (section 2.2.3 and Appendix A).

A bilinear attenuation model for Italy was proposed for the first time by Gasperini (2001). Its functional form is described by eq. 2.14. Gasperini (2001) showed that, on the basis of the analysis of variance (ANOVA) statistical test, the bilinear model fits the data significantly better than other considered attenuation laws. The attenuation coefficients computed by Gasperini (2001) are $\beta = 0.056$ for distances shorter that 45 km and $\gamma = 0.022$ at longer distances. These values indicate that the intensity decrease at short distances is about two times faster than the decrease at large distances. Gasperini (2001) supposed that the attenuation coefficients decrease is due to subcrustal path of phases controlling the observed intensity at distances longer than 45 km.

The short distance attenuation coefficient of the log-bilinear law (eq. 3.7) is about two times the one at large distances (their values are respec-

tively ($a = -0.0187$ and $a' = -0.0108$). These results roughly agree with those presented by Gasperini (2001). Also in this case the short distance parameter a may be related to direct crustal paths between source and locality, whereas the lower value of the large distance parameter a' may indicate subcrustal paths crossing lower attenuation materials.

Because of the quite complex geostructural features characterizing the Italian territory, relevant lateral variations in the propagation of earthquake effects are expected. As geometrical spreading is due to the increase of the wavefront surface as the seismic waves propagate, the variation of the structural characteristics of the crust crossed by seismic waves will scarcely affect this phenomenon. It is instead reasonable to assume that such variations are mainly related to the anelastic and scattering dissipation, which is influenced by rocks properties, fracturing, discontinuities, temperature variations etc. In formulating the tomographic problem, the coefficient of the logarithmic term of eq. 3.7 is considered constant over the entire Italian territory. The tomographic method is applied to study the lateral variation of the dissipation terms.

Following the procedure proposed in section 2.2.4, I_E can be eliminated from the log-bilinear attenuation model

$$
\begin{aligned}
\mu(\overline{I}_m, D) \;=\; & a\left(\min(D, 45) - \frac{1}{M_m}\sum_{k=1}^{M_m}\min(D_k^m, 45)\right) + \\
& +\; a'\left(\max(D - 45, 0) - \frac{1}{M_m}\sum_{k=1}^{M_m}\max(D_k^m - 45, 0)\right) + \\
& +\; b\left(\ln D - \frac{1}{M_m}\sum_{k=1}^{M_m}\ln(D_k^m)\right) + \overline{I}_m \,.
\end{aligned}
\tag{3.8}
$$

where $\overline{I}_m$ is the average of the observed intensities relative to each m-th earthquake and located within a given epicentral distance R_{max}.

The distances traveled in cells at epicentral distances shorter and longer than 45 km can be defined respectively ΔD_k and $\Delta D'_k$, for $k = 1, N_{cells}$. The distance D between the source and a locality can be seen as sum of distances traveled in each cell

$$
D = \sum_{k=1}^{N_{cells}} \Delta D_k + \sum_{k=1}^{N_{cells}} \Delta D'_k \,.
\tag{3.9}
$$

By substituting eq. 3.9 in eq. 3.8 the average log-bilinear reference model, which assumes constant attenuation coefficient a and a' over the entire Italian territory, can be rewritten as

$$
\mu_0(\overline{I}_m, D) \;=\; \sum_{k=1}^{N_{cells}} a\left(\Delta D_k - \frac{1}{M_m}\sum_{l=1}^{M_m}\Delta D_l\right) + \sum_{k=1}^{N_{cells}} a'\left(\Delta D'_k - \frac{1}{M_m}\sum_{l=1}^{M_m}\Delta D'_l\right) +
$$

$$+ \quad b \left(D - \frac{1}{M_m} \sum_{l=1}^{M_m} D_l \right) + \bar{I}_m. \tag{3.10}$$

A tomographic attenuation model $(\mu_1(\bar{I}_m, D))$ that accounts for lateral variations of attenuation properties can be formulated similarly

$$\mu_1(\bar{I}_m, D) \quad = \quad \sum_{k=1}^{N_{cells}} a_k \left(\Delta D_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D_l \right) + \sum_{k=1}^{N_{cells}} a'_k \left(\Delta D'_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D'_l \right)$$

$$+ \quad b \left(D - \frac{1}{M_m} \sum_{l=1}^{M_m} D_l \right) + \bar{I}_m. \tag{3.11}$$

where a_k and a'_k are the attenuation coefficients characterizing the k-th cells. To make the problem similar to classic seismic wave tomography the system of equations $G_{np}m_p = d_n$ that describes the inverse problem can be defined in a differential form. Subtracting eq. 3.11 from eq. 3.10 gives

$$\Delta\mu(\bar{I}_m, D) = \mu_1(\bar{I}_m, D) - \mu_0(\bar{I}_m, D) =$$

$$= \sum_{k=1}^{N_{cells}} (a_k - a) \left(\Delta D_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D_l \right) + \sum_{k=1}^{N_{cells}} (a'_k - a') \left(\Delta D'_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D'_l \right) =$$

$$= \sum_{k=1}^{N_{cells}} \Delta a_k \left(\Delta D_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D_l \right) + \sum_{k=1}^{N_{cells}} \Delta a'_k \left(\Delta D'_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D'_l \right) \tag{3.12}$$

The tomographic problem can be formulated as a system of linear equations $G_{np}m_p = d_n$. The known terms vector d_n is composed of the differences between observed intensity and the intensity expected by the average reference model (eq. 3.7) at the same site, for an earthquake of the same strength as that originating the observed intensity. Assuming that such deviations are caused by discrepancies between reference and tomographic model ($d_n = \Delta\mu$), the Δa_k and $\Delta a'_k$ (with $k = 1, N_{cells}$) form the unknown parameter vector m_p (with $p = 1, 2N_{cells}$). The differences between the distance traveled by seismic rays in the different cells and the average of the distances traveled in the same cells, $\Delta D_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D_l$ and $\Delta D'_k - \frac{1}{M_m} \sum_{l=1}^{M_m} \Delta D'_l$ (with $k = 1, N_{cells}$), constitute the matrix G_{np} (with $p = 1, 2N_{cells}$).

Choosing the cell size is a particularly critical aspect of the problem, as it determines the resolution of the tomographic model. Such choice depends on the density of the source-site paths in the zone of interest. If a too wide grid spacing is chosen some local features could be lost, whereas a too thin grid spacing might cause data over-fitting so that the estimated unknown values could not reflect the physics of the studied phenomenon but could be determined by the randomness of the data. A possible

empirical criterion to compare different grid spacings is to make several inversions by moving the grid origin for each of them. When, using a certain grid spacing, the solutions obtained for different positions of the grid origin do not differ significantly, the maximum resolution allowed for the considered data set has been achieved. A grid spacing of 25 km was chosen initially. As moving grid origin caused not trascurable variation in the inversion solution, the spacing was decreases to 30. Also in this case the position of the grid influences the estimated parameter value. Then a value of 25 was considered. In this case model parameter estimation are quite independent on grid origin position, indicating that the maximum resolution configuration has been found.

As the spatial ray coverage is not uniform and the reliability of the parameter estimates depends on the number of rays that cross the correspondent cell, the damped least square method is applied (section 3.1). The damping parameter λ avoids that the solutions assume values too far from the reference model, and this is particularly useful when a unknown is constrained by a low number of data. Forcing the model values be close to the reference model corresponds to reduce the model variability. This can be quantified by the model length $L = \sum_{i=1}^{N_m} m_i^2$. On the other side, the aim of applying least square method is to obtain solutions that fits well to data. Data fitting of the model can be measured by the variance after inversion, σ, which is given by eq. 3.6.

The value of λ can be chosen empirically by plotting trade-off curves between data variance after inversion and squared model length (Evans and Aucher, 1993). The inversion is carried on for different values of damping parameter, then for each of the considered damping value L and σ are calculated. Each of the considered damping value corresponds to a points on the bi-dimensional plot, where one axis represents model variance and the other the data variance after regression. Damping values corresponding to both lower variance after regression and lower model length should be chosen. The trade-off plot is shown in figure 3.1. It can be noted that increasing the damping value the model variance decreases but the data variance increase. On the basis of this plot $\lambda = 20$ is chosen. A further way to improve the solution reliability could be to restrict the estimation of the parameter values only to cells crossed by a number of rays greater than a certain threshold. When the number of data referring to a cell is lower than the threshold, the attenuation parameter values for that cell are assumed to coincide with the reference model. A threshold value of $N_{min} = 30$ is chosen.
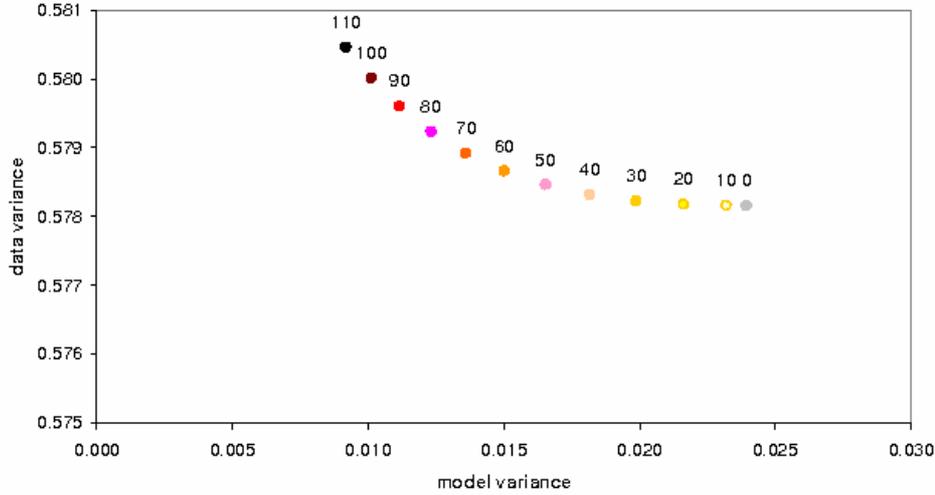
Figure 3.1: Data and model variance obtained for different damping values. Damping values are indicated by the black numbers near the corresponding data variance-model variance pair.

3.2.2 The data set

The data set used to study the lateral variations of intensity attenuation parameters is CPTI04 (see section 2.4.1).

Most of the selection criteria described in section 2.4.2 for the reference model regression are applied also for the tomographic study. In particular, earthquakes with a number of intensity observations lower than 10, occurred before 1200 A.D. or whose epicenter is likely to be located offshore (see table 2.1) are excluded from computation of tomography parameters values.

However, as the aim of the intensity tomography is just to study the lateral variations in the intensity attenuation properties (whatever is the cause of such deviations from the average attenuation pattern), the earthquakes occurred in volcanic areas are not discarded in this case.

Applying the above mentioned criteria a data set constituted of 23029 intensity observations, relative to 339 earthquakes, is obtained.

3.2.3 Tomography results

Figures 3.3 and 3.4 show the results of the inversion for the parameters Δa and $\Delta a'$, respectively. The areas and structures mentioned in the following paragraphs are shown in figure 3.2.

As explained above, the parameters Δa and $\Delta a'$ represent the deviation with respect to the parameters a and a' of the log-bilinear model
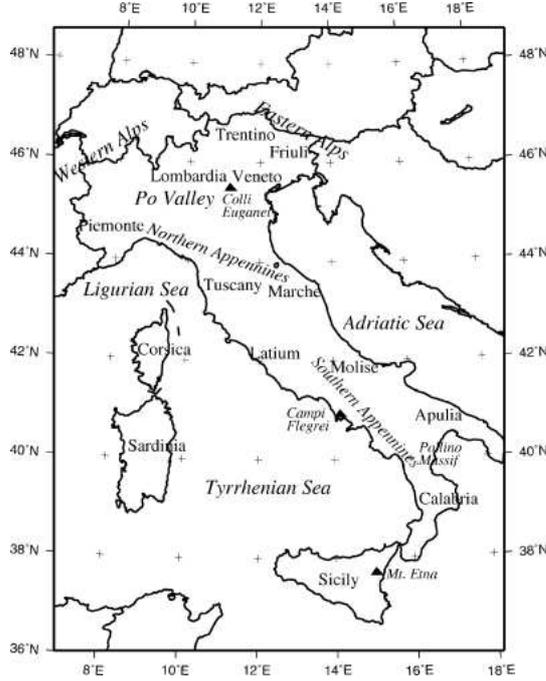
Figure 3.2: Map of Italy with the location of areas and structures mentioned in the text (from Carletti and Gasperini (2003)).

(eq. 3.7) that describes the average intensity attenuation pattern in Italy. Positive variations correspond to lower attenuation with respect to the reference model, whereas negative ones mean higher intensity attenuation. In cells excluded from the inversion because of a too low number of rays crossing them (see section 3.2.1) it is assumed that the deviations from the reference model are assumed to be zero.

The parameter Δa describes the lateral variations in the intensity attenuation pattern at short distances ($\leq$ 45 km). Northern Italy is characterized by low attenuation, with an evident spot of high attenuation in the Veneto-Friuli area (northeast of Italy). High attenuation characterizes the Tyrrhenian slope of northern and central Apennines, whereas the Adriatic slope is characterized by low attenuation in the northern section and high attenuation in the southern one. In southern Italy relatively small zones of positive and negative attenuation alternate. In particular the most evident ones are close to the Adriatic coast around latitude 42°, on the Tyrrhenian coast around latitude 39°, and in Sicily, around the Mt. Etna volcanic area.

The parameter $\Delta a'$, which describes the intensity attenuation at large distances ($>$ 45 km), shows an intensity attenuation pattern quite different from Δa. For this parameter in the North East of Italy there is a pronounced low attenuation zone (in contrast with the high attenuation
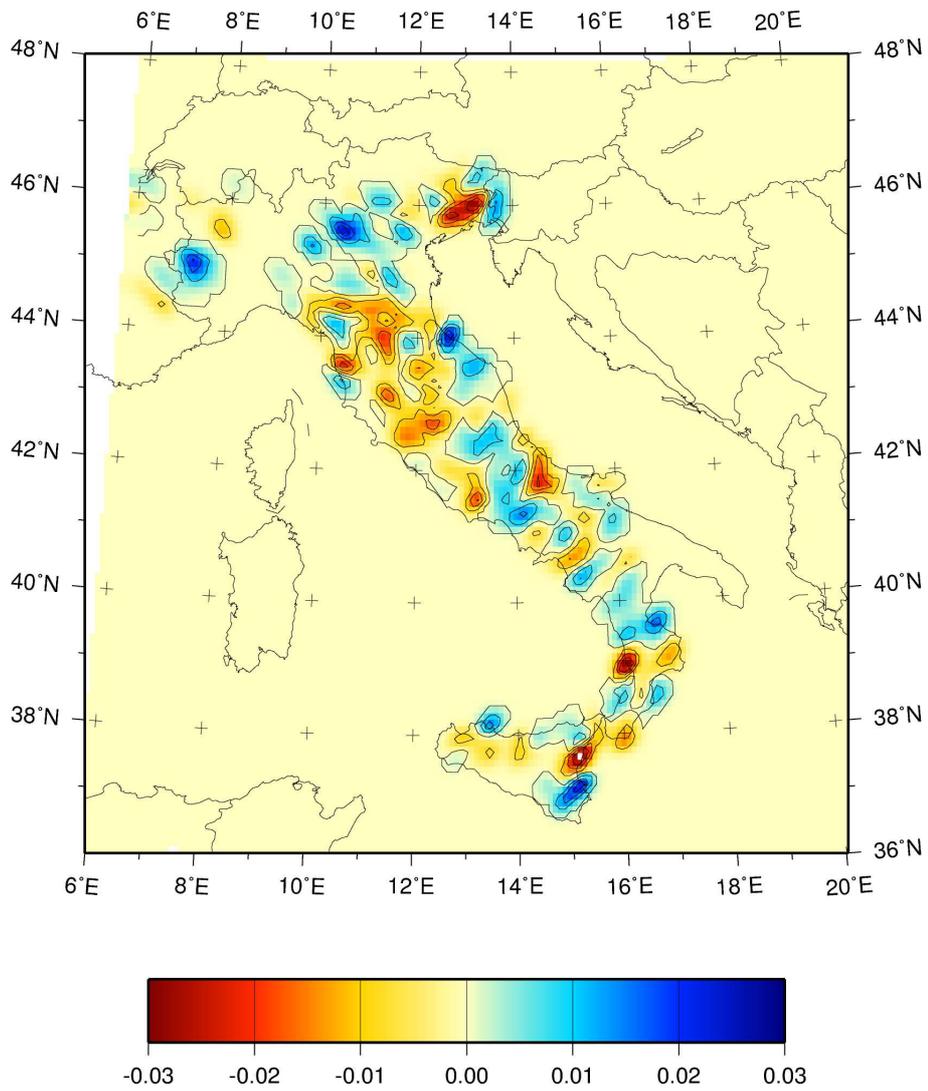
Figure 3.3: Results for attenuation coefficient Δa.
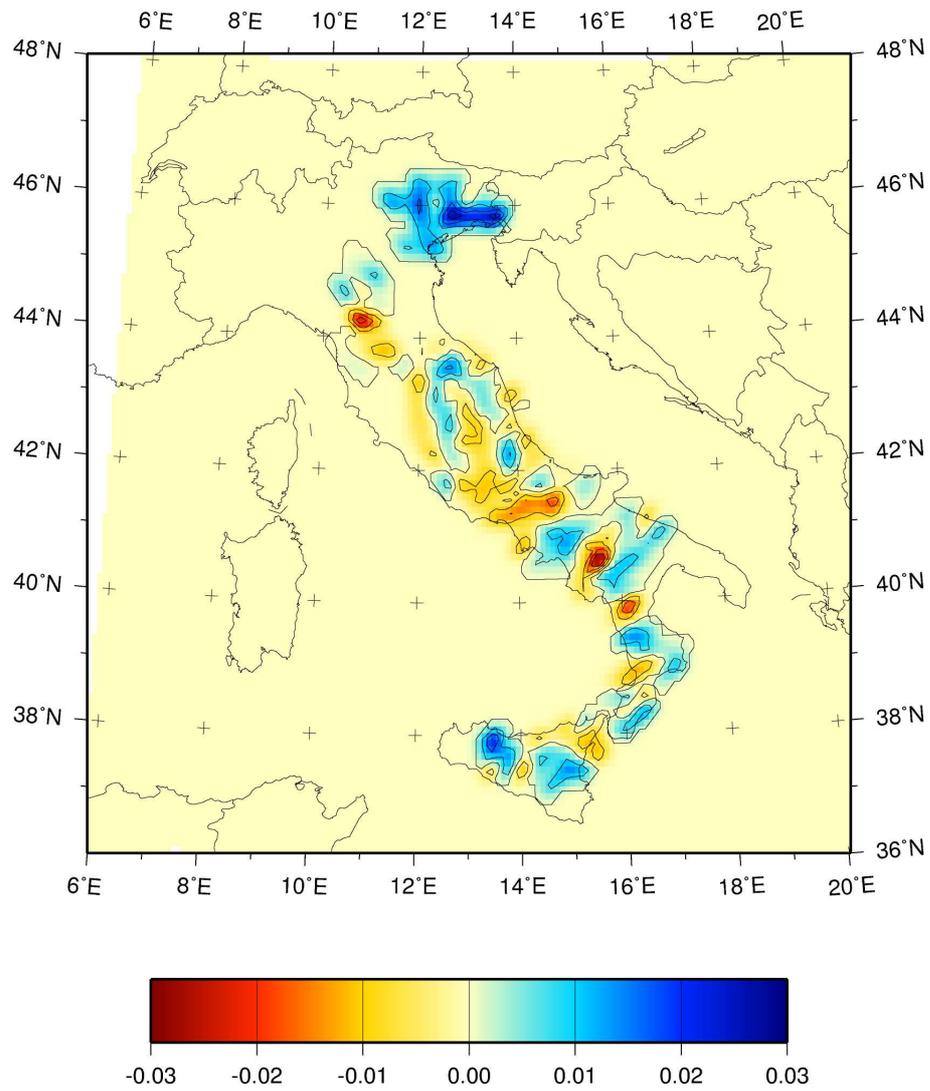
Figure 3.4: Results for attenuation coefficient $\Delta a'$.

spot evidenced by Δa parameter). High attenuation still characterizes
the Tyrrhenian slope of northern-central Apennines, and it is also present
on the Adriatic coast, whereas in the inner zone a low attenuation zone
appears. The Southern part of the Italian Peninsula is still character-
ized by alternating positive and negative attenuation zones. The positive
spots on the Tyrrhenian coast at about 39° of latitude and corresponding
to the Mt. Etna are still present but less pronounced. The spot located
at latitude 42° is substituted by a slight low attenuation zone. An high
attenuation spot appears at about lat=39.7° and lon=16.1°.

To check the reliability of these results the average site residual are
calculated for sites where more than 8 intensity observations are avail-
able. The existence of trends or regularities in the residuals distribution
would indicate a poor or biased fit. The absence of evident trends of
residuals (figure 3.6) confirms the reliability of the tomography results.
The tomography residuals can be compared with those obtained for the
same data set (section 3.2.2) using the reference log-bilinear model (fig-
ure 3.5). The tomographic residuals are generally lower, moreover their
distribution on the studied area is more homogeneous than those ob-
tained for the reference model. For example the Mt. Etna volcanic zone
is characterized in both cases by the prevalence of negative residuals,
that indicate an intensity decrease greater than the one expected by the
attenuation models, but some sites of the Etna area that show negative
residual for the reference model turn to a positive residual (observed in-
tensity higher than predicted one) for the tomographic one, indicating
that even if the intensity attenuation is still underestimated, the extent
of the average deviations between data and model predictions decreases.

Model parameters error are evaluated by eq. 3.5. Figure 3.7 and 3.8
show the errors on the short and large distance attenuation parameters
Δa_k and $\Delta a'_k$, respectively. In general higher errors are observed at the
boundary zones of the area interested by data, whereas the errors relative
to the inner zones assume quite low values. The minimum value of the
errors on the short distance attenuation parameters Δa_k is 0.0028 deg km
$^{-1}$ and the maximum is 0.0187 deg km $^{-1}$. The large distance coefficients
$\Delta a'_k$ errors vary between a minimum of 0.0040 deg km $^{-1}$ and 0.0186 deg
km $^{-1}$. The average values for both errors is about 0.008 deg km $^{-1}$. The
range of variation of parameters Δa and $\Delta a'$ is about 0.1 deg km^{-1} (from
-0.056 to 0.0482 and from -0.0537 to 0.0399 deg km^{-1}, respectively). On
average the errors affects for 10-13% their value, indicating that they
really reflect lateral variations in the attenuation pattern characterizing
the territory of interest and are not due to casual errors in intensity
observations. Model parameters error are evaluated by eq. 3.5.
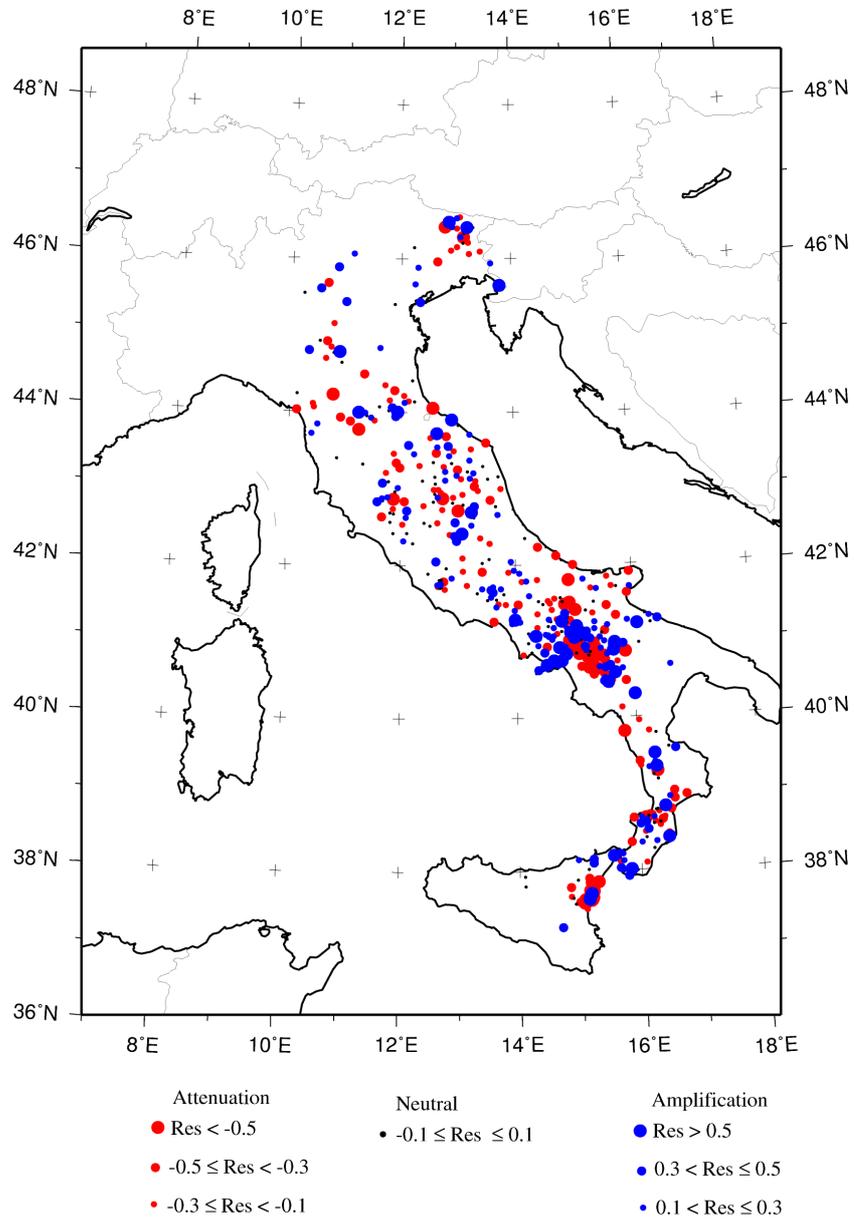
Figure 3.5: Local residuals resulting from the application of the reference intensity attenuation model.
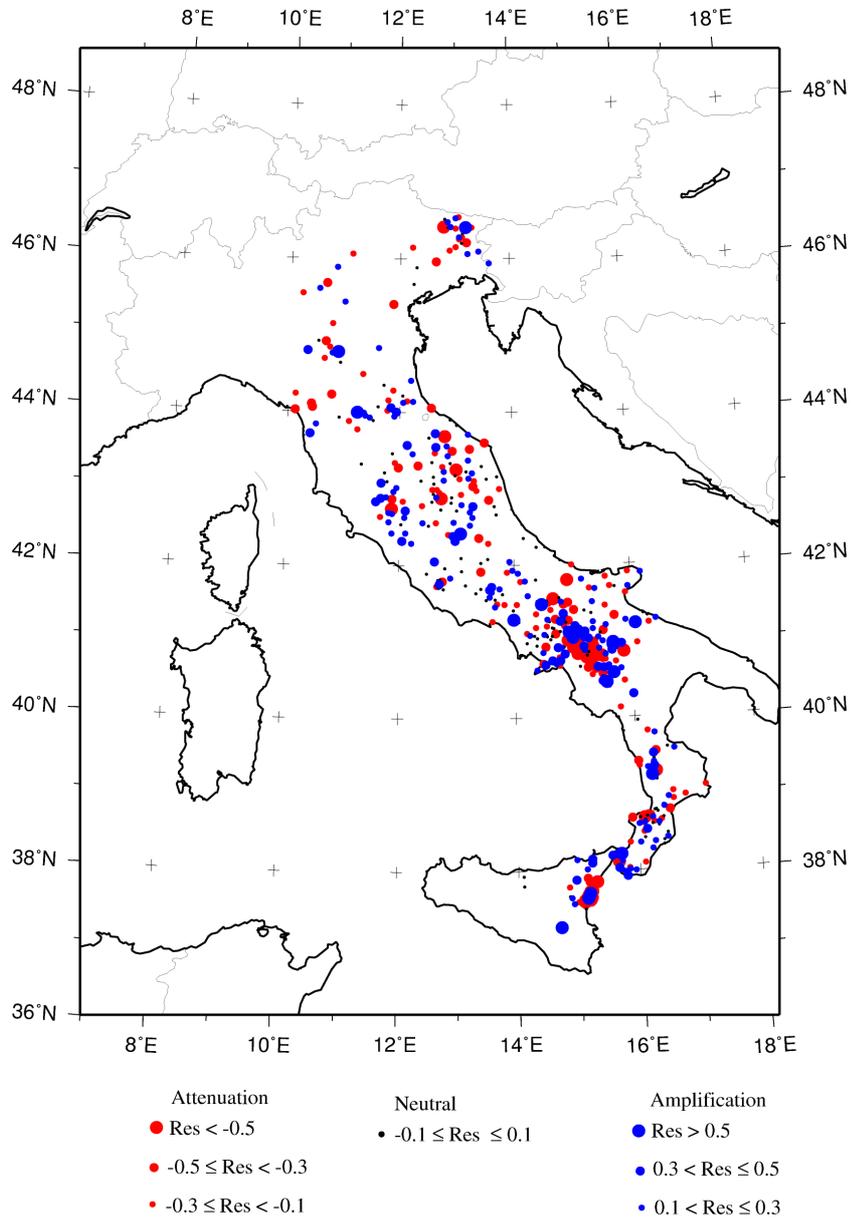
Figure 3.6: Local residuals resulting from the application of the tomographic intensity attenuation model.
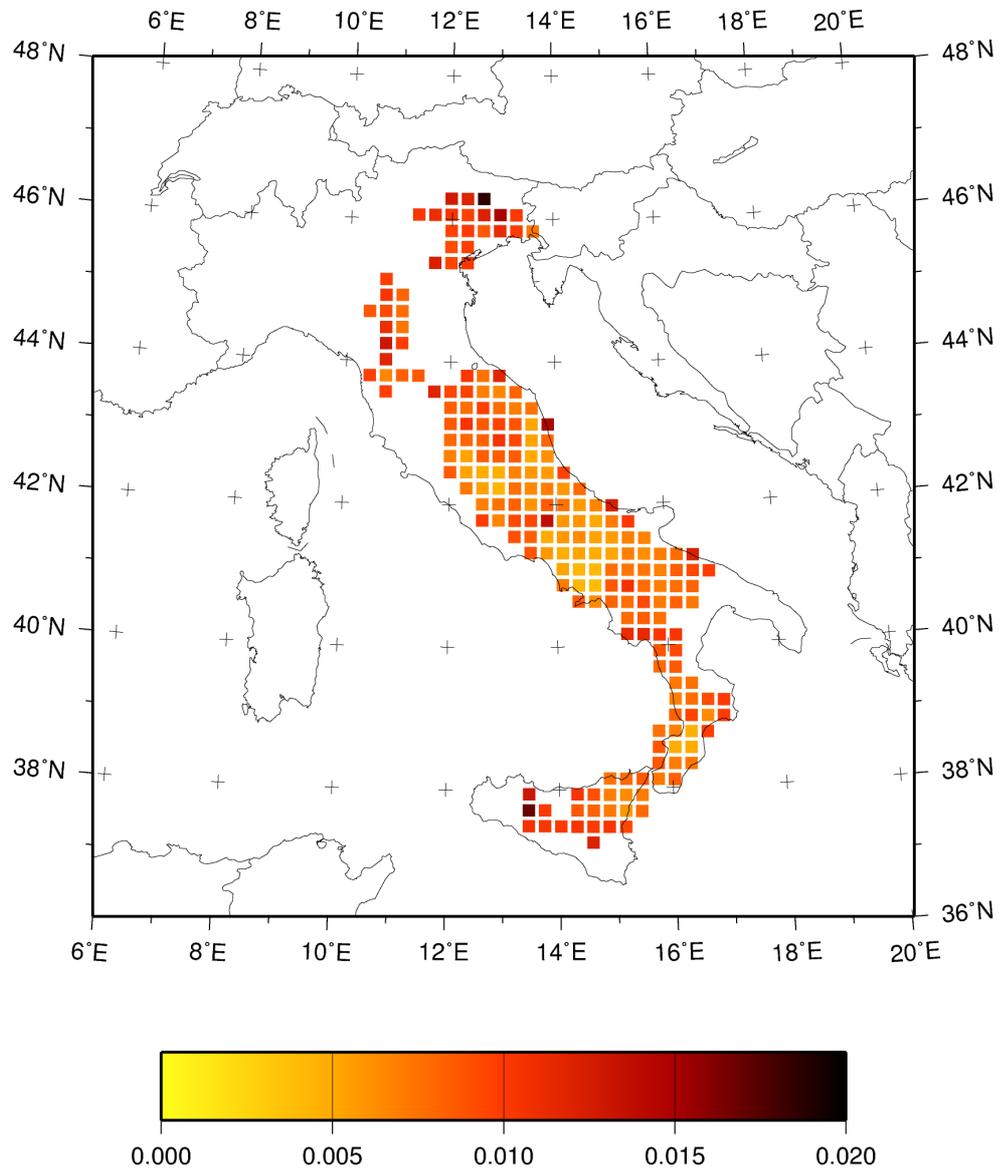
Figure 3.7: Errors on the short distance attenuation coefficient Δa.
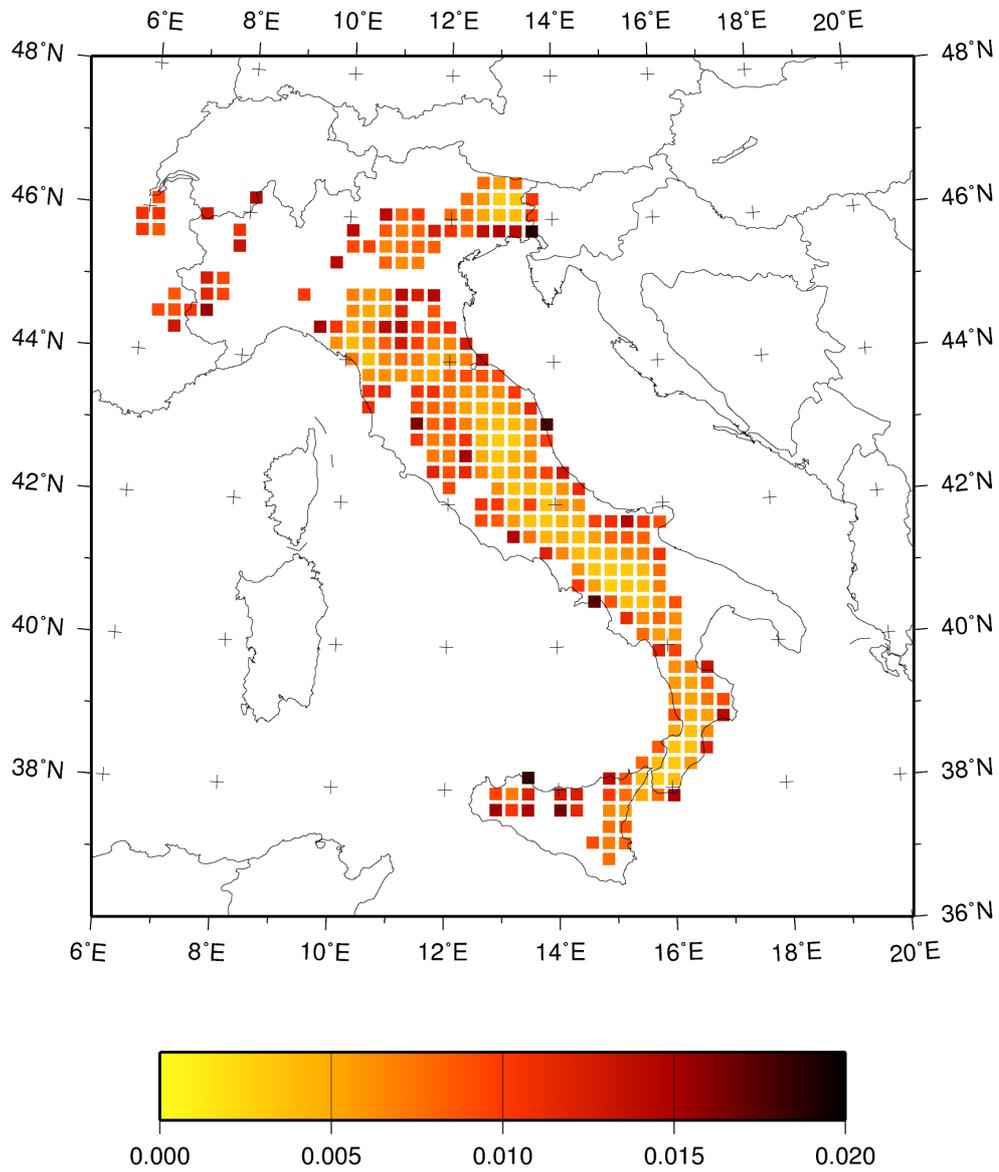
Figure 3.8: Errors on the large distance attenuation coefficient $\Delta a'$.

3.2.4 Previous intensity attenuation study

Even if the functional form of the reference model by Carletti and Gasperini (2003) is different from the one assumed in this work (it includes an intercept term and lacks a logarithmic term) the unknown of the inversion problem (the coefficients at short and large distances) have similar meaning. So the tomography results, that describe the lateral variation respect to the average values of such coefficients, are comparable.

In general the results of this work agree with those of Carletti and Gasperini (2003). The resolution is greater than those of the previous intensity tomography. In fact the cell size is 25 km, whereas in the previous work it was 50 km. Some local discrepancies between the two models may be due to the improvement in the resolution characterizing the second one and the different dataset used.

3.2.5 Comparison with surface heat flow

Attenuation of seismic waves in rocks depends strongly on temperature, hence a spatial correlation between this quantity and the coefficients of the linear terms in the intensity attenuation law is expected. As the temperature profile in the Italian crust and upper mantle is not known precisely, surface heat flow could be considered as an index of the thermal state in the underlying lithosphere. Assuming a constant conductivity and radioactive generation, the temperature would be roughly proportional to the surface heat flow. This assumption is not valid for geothermal areas, where also the heat transport by fluid circulation is important. However, at least at the first order, a direct correspondence between high heat flux and high subsurface temperature could be assumed. In this section the results of the intensity tomography are compared with the heat flow map of Italy (Della Vedova et al., 1991; Cataldi et al., 1995) shown in figure 3.9.

The spatial variations of intensity attenuation at short distances are described by the parameter Δa and show a very good correspondence with the surface heat flow. In particular the geothermal area along the Tuscany and Latium Tyrrhenian coast that presents q_s values up to some hundreds of mW m^{-2} corresponds to an high attenuation zone (negative Δa values). Also the high heat flow observed near the Mt. Etna and Campi Flegrei volcanic areas can be related to high attenuation zones showed in figure 3.3. Moreover the high attenuation spot in north-eastern Italy corresponds to a heat flow high ($q_s > 70$ mW m^{-2}). Low heat flow along the Adriatic coast corresponds, at least in the northern portion, to an high attenuation zone. Also in Northern Italy low heat flow generally reflects high attenuation although in its Western boundary the corre-
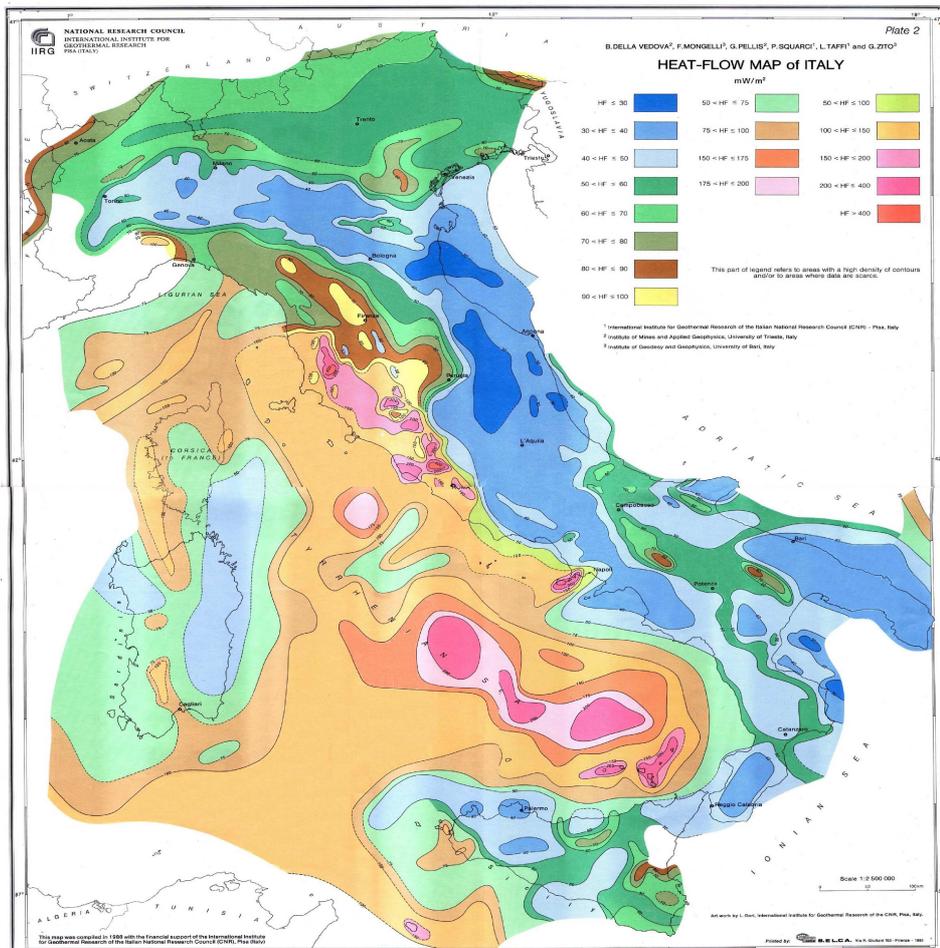
Figure 3.9: Heat flow map of Italy, (Della Vedova et al., 1991).

spondence is poor, probably due to the low number of rays crossing the boundary zones.

A correspondence between the results for the $\Delta a'$ parameter and the heat flow distribution is less evident, although high attenuation is still observed in the Tuscany-Latium Tyrrhenian coast.

3.2.6 Inferred attenuation structure

Lateral variations in the attenuation of seismic intensity with distance allow to infer the attenuation structure of the investigated area. Assuming a linear relation between intensity and the logarithm of ground motion amplitude, a log-linear intensity attenuation law is obtained (Appendix A)

$$
\begin{aligned}
\Delta I = I_0 - I &= \frac{1}{k}\left[\frac{n}{2}\log_{10}\frac{D}{h} + \log_{10}e\frac{\pi}{TVQ}(D-h)\right]\\
&= \frac{1}{k\ln 10}\left[\frac{n}{2}\ln\frac{D}{h} + \frac{\pi}{TVQ}(D-h)\right].
\end{aligned} \tag{3.13}
$$

Where the first term accounts for the geometrical spreading and the second one for the anelastic attenuation and scattering. When epicentral intensity is calculated by the two step procedure described in Chapter 2 ($I_0 = I_E$), the log-linear model described by eq. 2.17 is obtained, with $a = \pi/(kTVQ\ln 10)$ and $b = n/(2kln10)$.

In the tomographic study of intensity attenuation, the reference model is the log-bilinear (eq. 3.7). This could still be interpreted in terms of seismic energy decrease caused by anelastic attenuation and geometrical spreading. The values of attenuation parameters of the reference model are $a = -0.019$ deg/km for distances shorter than 45 km and $a' = -0.011$ deg/km at larger distances. These values indicate that the intensity decrease due to anelastic attenuation at distances shorter than 45 km is about two times that relative to longer distances. In fact the two estimated parameter values correspond to about 5.38 and 9.26 km/deg respectively.

The seismic quality factor Q describes the attenuation properties of rocks encountered by seismic waves as they propagates between the earthquake source and the site where intensity data are observed (Appendix A). From eq. 3.13 it can be written

$$
a = \left(\frac{\Delta I}{\Delta r}\right)_a \approx \frac{1}{k\ln 10}\frac{\pi}{TVQ_a}, \tag{3.14}
$$

and

$$
a' = \left(\frac{\Delta I}{\Delta r}\right)_{a'} \approx \frac{1}{k\ln 10}\frac{\pi}{TVQ_{a'}}. \tag{3.15}
$$

The Q values for shallower and deeper ray paths (Q_a and $Q_{a'}$, respectively) can be estimated from eq. 3.14 and eq. 3.15 by assuming that ground motion amplitude doubles for an increase of one intensity degree ($k = \log_{10} 2$), seismic wave velocity $V = 3.5$ km/s and $T = 0.4$ s (corresponding to a frequency of 2.5 Hz, generally causing the greater damages on buildings). From the reference model attenuation parameters average Q estimation for the entire Italian territory could be obtained. They are $Q_a = 173.2$ and $Q_{a'} = 298.9$. The intensity tomography results allow to calculate the lateral variation of seismic quality factor respect to the average values, that added to these last ones allow to obtain the attenuation structure characterizing Italian peninsula showed in figures 3.10 and 3.11.

3.2.7 Discussion of intensity tomography results

In the preceding sections a tomographic study of the lateral variation of seismic intensity attenuation properties in Italy has been developed. The resolution achieved by using the data from CPTI04 increases with respect to the intensity tomography by Carletti and Gasperini (2003). In fact, such study was carried on with a grid spacing of 50 km, while in the present one is possible to achieve a grid spacing of 25 km.

Both the marked reduction of locality residuals and their more uniform distribution with respect to the isotropic model, demonstrate that the lateral variations of attenuation coefficients give a meaningful description of the properties of seismic wave propagation in the Italian area. Also model parameter errors confirm the reliability of the results, as, on average, the errors affect for 10-13% only the parameters values (see section 3.2.3). Moreover by considering the spatial distribution of errors it can be noted that the parameters estimates corresponding to cells situated in the inner part of the area covered by data are more reliable than those situated at the boundaries.

Attenuation of seismic waves in rocks is strongly influenced by temperature, so a spatial correlation between tomography results and temperature is expected. As the temperature profile in the Italian crust and upper mantle is not known precisely, surface heat flow has be considered as an index of the thermal state in the underlying lithosphere. Intensity tomography results have been compared with the heat flow map of Italy. The clear correlation found between the spatial variations of the near-field ($D \leq 45$ km) coefficient with the heat flow in Italy supports the existence of a physical grounding for the presented approach.

Another interesting application of intensity tomography is to infer the attenuation structure in terms of seismic quality factor Q. By assuming a linear relation between the observed intensity and the logarithm of
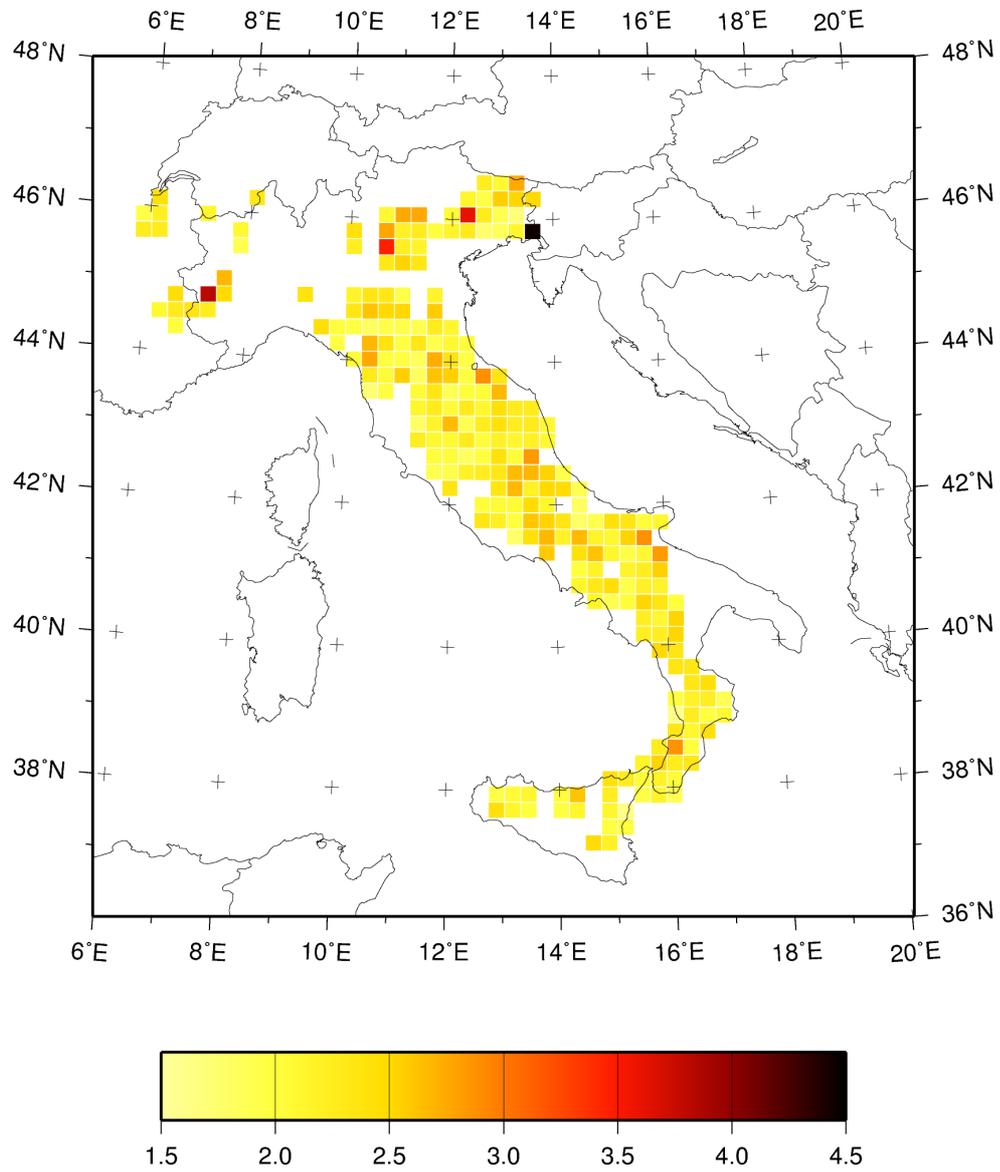
Figure 3.10: Decimal logarithm of seismic quality factor for shallower ray paths.
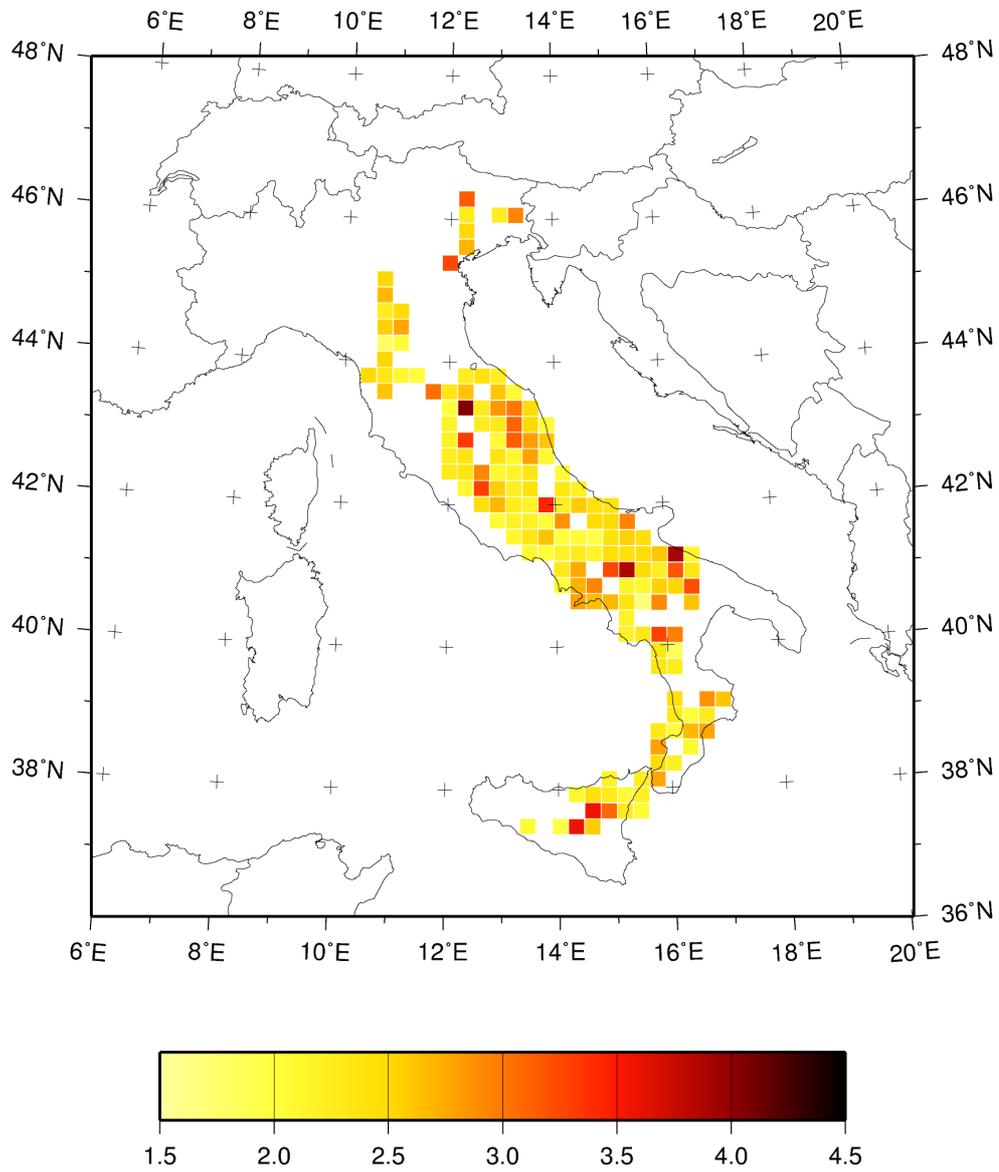
Figure 3.11: Decimal logarithm of seismic quality factor for deeper ray paths.

ground motion amplitude, average Q values for shallower and deeper wave path have been calculated. Tomography results have been used to obtain the lateral variations of the seismic quality factor with respect to these average values. This procedure potentially allows to obtain information about the attenuation structure for zones where instrumental data are not available.

3.3   Seismic quality factor tomography

Seismic waves arriving at a certain position on the Earth surface origi-
nate a motion of the medium constituting particles. The motion can be
converted into a signal by a seismometer or a geophone (in exploration
seismology). Depending on the kind of instrument used, ground motion
is recorded in terms of displacement, velocity or acceleration. Mechanical
seismometers were the first seismic recording systems. They are consti-
tuted by a frame and a mass connected to it by a spring and a damping
mechanism. Almost all seismometers are based on damped inertial pen-
dulum system. The frame of the seismometer is attached to the ground
and the pendulum is designed so that the movement of the internal mass
is delayed relative to the ground motion by the inertia of the mass. Each
pendulum system has an equilibrium position in which the mass is at rest
and to which it will return after small transitory disturbances. Ground
displacement originated by seismic waves is communicated trough the
spring to the mass, that in absence of damping would oscillate peri-
odically about its equilibrium position. The viscous damping tends to
restore the system to the equilibrium position. A mechanical or optical
system and a clock are used to produce the seismogram. If the ground
motion frequency is much lower than the resonant frequency of the un-
damped system (seismograph frequency), the seismograms record the
ground acceleration, whereas if it is much higher than the resonant fre-
quency, ground displacement is recorded. In electromagnetic instruments
the mass is attached to a coil and is subjected to a magnetic field. Its
motion produces an electrical current that is proportional to mass veloc-
ity. Also force feedback instruments are based on the same principle. In
this case a force proportional to the inertial mass displacement is applied
by a negative feedback loop to the mass to cancel its relative motion. An
electrical transducer converts the mass motion into an electrical signal to
asses how much feed back force has to be applied. The amount of force
required to hold the system at rest corresponds to ground acceleration.

Recorded ground motion is described in time domain by seismograms.
These are two dimension plots with one axis representing time and the
other the ground motion amplitude. A consequence of this representation
is that, historically, the attention of seismologists has been focused on
kinematic properties of wave motion (arrival time, travel time, velocities).
Because of its lower reliability of amplitude measures, it contributed less
to the early developments of seismology. With the introduction of mag-
nitude scale in 1930's greater attention was paid to amplitude measures.

From direct analysis of seismograms the existence of different wave
types and the evidence of common characteristics of the records relative
to the same source regions were noted. This leaded to the development

of waveform analysis, that works entirely in the time domain, generally aiming to associating certain observed wave characteristics with any given station, depending on the epicentral region.

Time domain approach is characterized by ease of measurements and interpretation, but is not very suitable for comparing data relative to the same event recorded at different positions. As seismic waves are a physical phenomenon which fluctuate in space and time, frequency (and wave number) would be more informative than time as independent variable. Unlike time domain representation, frequency domain allows to compare different records unambiguously. Besides to be more informative, the frequency-domain representation is often simpler to handle computationally. The representation of a function in frequency (or wave number) domain is called spectrum and will be defined in details in section 3.3.1.

3.3.1   Spectral analysis

Given an arbitrary function $f(t)$ representing a measured physical quantity, its Fourier transform (or Fourier spectrum) is defined as

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-i\omega t) \, \mathrm{d}t, \qquad (3.16)$$

where ω is the angular frequency, in radians per second, $F(\omega) = a(\omega) - ib(\omega) = |F(\omega)| \exp[i\Phi(\omega)]$, $a(\omega) = \int_{-\infty}^{\infty} f(\lambda) \cos\omega\lambda \, \mathrm{d}\lambda$, and $b(\omega) = \int_{-\infty}^{\infty} f(\lambda) \sin\omega\lambda \, \mathrm{d}\lambda$.

The inverse Fourier transform is given by

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp(i\omega t) \, \mathrm{d}\omega \,. \qquad (3.17)$$

Eq. 3.16 expresses the Fourier analysis of $f(t)$ and corresponds to a mapping of the considered function from time to frequency domain. Eq. 3.17 express the Fourier synthesis, i.e. synthesize the various spectral components $F(\omega)$ into the original function $f(t)$.

The function $F(\omega)$ corresponds to a mapping of the considered function from time to frequency domain. By virtue of the orthogonal properties of trigonometric functions, $\exp(-i\omega t)$ acts like an operator picking out from $f(t)$ only components with frequency ω. More intuitively, this means that the function $f(t)$ is represented by a sum of sine wave components characterized by a certain amplitude, phase and frequency (ω).

In particular, $|F(\omega)| = [a^2(\omega) + b^2(\omega)]^{\frac{1}{2}}$ is the amplitude spectrum and $\Phi(\omega) = \tan^{-1}[-\frac{b(\omega)}{a(\omega)}] + 2n\pi$ for $n = 0, \pm 1, \pm 2, ...$ is the phase spectrum. Amplitude and phase spectra describe respectively the amplitude and the

phase associated to the various frequencies characterizing the sinusoidal waves involved in the summation.

Fourier transform method assumes that the analyzed function assumes a null value outside the integration interval. This condition is satisfied by transient signals and in particular by seismic waves. In general a seismograms and its properties are the result of source action, medium properties and receiving seismograph characteristics and could be explained in terms of filtering. The input signal generated by seismic source is subjected to the action of a certain number of filters whose result is the recorded output signal. This could be described in frequency domain as

$$S(\omega)H_1(\omega)H_2(\omega)...H_n(\omega) = X(\omega) \,, \tag{3.18}$$

where $S(\omega)$, $H_1(\omega)$, $..H_n(\omega)$ describe source and medium properties and $X(\omega)$ is the spectrum of the recorded signal x(t). Eq. 3.18 can be split in two parts, i.e. the amplitude spectrum

$$|S(\omega)||H_1(\omega)||H_2(\omega)|...|H_n(\omega)| = |X(\omega)| \,, \tag{3.19}$$

and the phase spectrum

$$\Phi_S(\omega) + \Phi_{H_1}(\omega) + \Phi_{H_2}(\omega)... + \Phi_{H_n}(\omega) + 2m\pi = \Phi_X(\omega) \tag{3.20}$$
$$m = 0, \pm 1, \pm 2, ...$$

When a certain number of receivers record body waves produced by various sources, eq. 3.18 can be rewritten by considering more in detail the path effects affecting the seismic signal

$$A(f)^{ij} = S(f)^i K(f)^j I(f)^j G^{ij} \exp\left(-\pi f t_*^{ij}\right) \,, \tag{3.21}$$

where f is the frequency (in cycles per second) and the superscripts i and j identify respectively the source generating the seismic waves and the receiver that recorded them. The term $(f)^i$ is the source spectrum and is only influenced by source properties. $K(f)^j$ accounts for the site effect. $I(f)^j$ is the instrumental response and describes the action of the recording system on the signal. G^{ij} accounts for geometrical spreading, it is frequency independent and it is determined by the distance traveled by seismic waves to arrive from source to receiver. The exponential term $\exp(-\pi f t_*^{ij})$ accounts for anelastic attenuation along the path between the i-th source and the j-th receiver, with

$$t_*^{ij} = \int_{ray} \frac{1}{Q(r)v(r)} \, \mathrm{d}r \,. \tag{3.22}$$

$Q(r)$ is the seismic quality factor, $v(r)$ is the seismic velocity.

3.3.2 Spectral ratio method

Eq. 3.21 evidences the various factors that influence the recorded seismic signal. Usually both source and path effect are unknown, and cannot be determined unambiguously by a single observation $A(f)^{ij}$. To overcome this problem various measurements could be compared. These should be taken under similar conditions so that most of the factors of eq. 3.21 do not vary between the different data and a single factor can be isolated and studied. This procedure is referred to as equalization. Spectral ratio method is an example of equalization procedure.

Taking the natural logarithm of eq. 3.21 gives

$$\ln A(f)^{ij} = \ln S(f)^i + \ln K(f)^j + \ln I(f)^j + \ln G^{ij} - \pi f t_*^{ij} \qquad (3.23)$$

The spectral ratio could be defined as the difference between the logarithm of the observed amplitude spectrum ($\ln A(f)^{ij}$) and the average of the logarithms of the amplitude spectra recorded at the same receiver

$$
\begin{aligned}
\Delta \ln A(f)^{ij} &= \ln A(f)^{ij} - \frac{1}{N_j} \sum_{l=1}^{N_j} \ln A(f)^{lj} \qquad (3.24)\\
&= \ln S(f)^i + \ln K(f)^j + \ln I(f)^j + \ln G^{ij} - \pi f t_*^{ij} + \\
&- \frac{1}{N_j} \sum_{l=1}^{N_j} (\ln S(f)i + \ln K(f)^l + \ln I(f)^l + \\
&+ \ln G^{il} - \pi f t_*^{ij})
\end{aligned}
$$

As site effects and instrumental response are the same for all the spectra relative to the same receiver, the corresponding terms disappear from eq. 3.24. Spectral ratio method could be applied to seismic exploration data. In this case the sources are not earthquakes and generate almost the same signal for different shots. In this case the source spectrum is the same for all data and also the source terms in eq. 3.24 can be deleted, giving

$$\Delta \ln A(f)^{ij} = \ln G^{ij} - \frac{1}{N_j} \sum_{l=1}^{N_j} \ln G^{ij} - \pi f \left(t_*^{ij} + \frac{1}{N_j} \sum_{l=1}^{N_j} t_*^{lj} \right) \qquad (3.25)$$

As the first and the second term on the right side of eq. 3.25 are frequency independent, the spectral ratio follows a linear trend with frequency

$$\Delta \ln A(f)^{ij} = l^{ij} + m_1^{ij} f \,, \qquad (3.26)$$

with

$$l^{ij} = \ln G^{ij} - \frac{1}{N_j} \sum_{l=1}^{N_j} \ln G^{ij} \qquad (3.27)$$

$$m_1^{ij} = -\pi \left(t_*^{ij} + \frac{1}{N_j} \sum_{l=1}^{N_j} t_*^{lj} \right) \tag{3.28}$$

The slope m_{ij}^1 could be calculated (along with the intercept l_{ij}) by linear regression and by eq. 3.22 allows to obtain information about the velocity and attenuation structure of the investigated area.

3.3.3   Formulation of the tomographic problem

The area of interest could be parametrized with a grid constituted of N cells: each cell is characterized by a seismic velocity V_k and a seismic quality factor Q_k. Eq. 3.22 can be discretized as

$$t_*^{ij} = \sum_{k=1}^{N} \frac{r_k^{ij}}{V_k Q_k} \,, \tag{3.29}$$

where r_k^{ij} is the distance traveled in the k-th cell for the observation relative to the j-th receiver and the i-th source. From eq. 3.28 we can be obtain:

$$\begin{aligned}
m_1^{ij} &= -\pi \sum_{k=1}^{N} \frac{r_k^{ij}}{V_k Q_k} - \sum_{l=1}^{N_j} \sum_{k=1}^{N} \frac{r_k^{lj}}{V_k Q_k} \tag{3.30} \\
&= -\pi \sum_{k=1}^{N} \frac{1}{Q_k V_k} \left(r_k^{ij} - \sum_{l=1}^{N_j} r_k^{lj} \right)
\end{aligned}$$

$$\tag{3.31}$$

where N_j is the number of sources for the j-th receiver.

By inverting simultaneously for velocity and attenuation, the number of unknown could be too high in comparison with the number of data, and may lead to unreliable solutions. A fixed velocity model could be assumed and only the attenuation structure is calculated. In this case, the linear system representing the inverse problem (eq. 3.1) becomes

$$-\frac{m_1^{ij}}{\pi} = \sum_{k=1}^{N} \left[\left(\frac{r_k^{ij}}{V_k} - \sum_{l=1}^{N_j} \frac{r_k^{lj}}{V_k} \right) \frac{1}{Q_k} \right] \tag{3.32}$$

To use the same inversion scheme applied in Chapter 2, eq. 3.32 can be written in matrix form by defining

$$d_n = -\frac{m_1^{ij}}{\pi}$$

$$m_p = \frac{1}{Q_k}$$

$$G_{np} = \frac{r_k^{ij}}{V_k} - \sum_{l=1}^{N_j} \frac{r_k^{lj}}{V_k} \tag{3.33}$$

At this point, model parameters m_p can be estimated by solving the inverse problem $d_n = G_{np}m_p$ (see section 3.1).

3.3.4 The data set

The data have been recorded in the framework of the seismic reflection profile Alcudia. In particular, the data used to calculate the tomographic model cover a length of about 11 kilometers along the national road CM 403, in the province of Ciudad Real, (Castilla-La Mancha, Spain) between the town of Ventas con Peña Aguilera and the Torre Abraham reservoir.

The sources were four 22 ton Vibroseis originating a signal with frequency varying from 8 to 80 Hz and whose duration is 60 s. The four Vibroseis acted contemporary in each vibration point. The distance between vibration point was about 70-100 m. The station spacing was 35-50 m, and each station was composed of 10-12 geophones. In figure 3.12 the first and the last of the 17 vibration point are represented with stars whose label are respectively 01 and 17. The geophones were moved along the red line as the four Vibroseis moved from one shooting point to the following one.
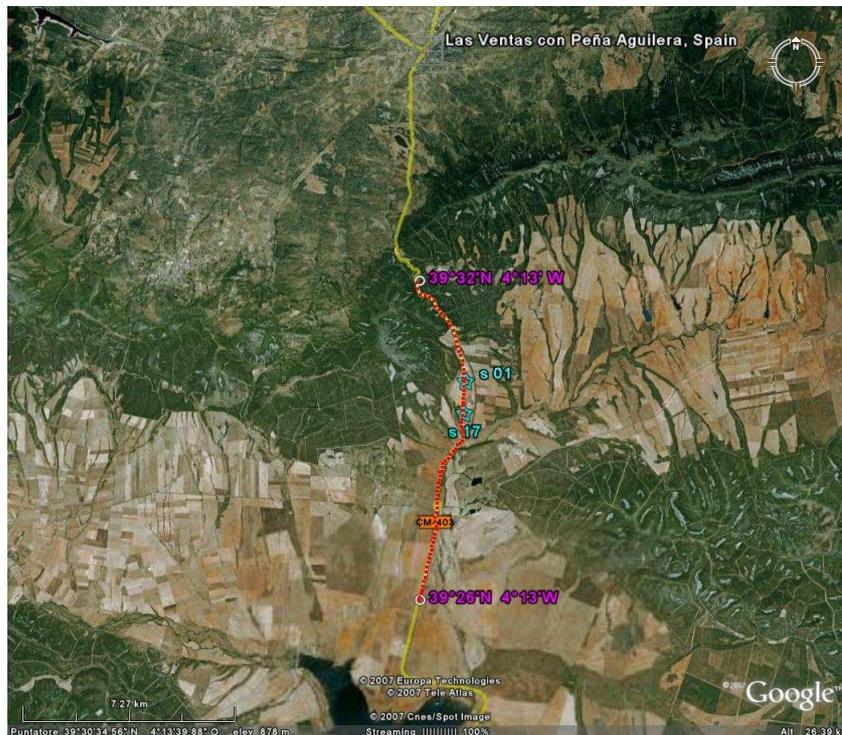
Figure 3.12: Geographic position of sources and receivers used to record the data (see text for more details).

The seismic profile crosses different rock types in particular conglomerates, sands and gravel layers whose thickness is about 10 m, which overlie shales, quarzites, tabular quarzites. Figure 3.13 represents a schematic geologic section of the zone interest by the seismic profile, deduced by geological information available for the zone of interest (IGME, 1986, 1970).
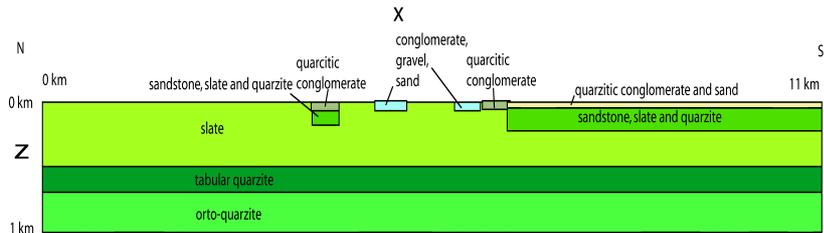
Figure 3.13: Rock types characterizing the interest area. Deduced from IGME (1986) and IGME (1970).

3.3.5   Data analysis

In order to apply spectral ratio method, data are processed using various utilities of the CWP Seismic Unix package (Cohen and Stockwell, 2002), which allow to calculate the amplitude spectra of each observation. Then the slopes m_1^{ij} of the regression line between spectra and frequency that are necessary for the seismic quality factor tomography are calculated.

To evaluate the attenuation properties of the investigated area in terms of inverse of P-waves seismic quality factor, travel times of first arrivals are necessary. Travel time picking is done by the utility Xpicker from the Seismic Unix package.

Observational data are necessary recorded in a limited time interval. The consequence of a limited data window is that a correct spectrum is impossible to obtain due to the finite integration interval. This causes a smoothing of the computed spectrum and a lack of spectral resolution that depends on time window length. This effects can be mitigated by modifying the window shape. To minimize the spectral distortion, the window $w(t)$ has to be broad and without sharp corners. Using not a particular window is equivalent to applying a rectangular (box-car) window, that does not cause signal distortion in time domain but leads to strong distortions in frequency domain. Windows that taper off gradually towards both ends of the record interval under investigation cause some distortion in time domain but mitigate the distortion in frequency

domain. There is no straight-forward procedure to derive the shape of the best window and generally it is derived by trial and error (Bath, 1974).

A linear taper that forces the amplitude values to get close to zero at the time window bounds is chosen. The taper is applied during a certain time interval following the window starting time and preceding the ending time. Tapering time interval and time window length are chosen empirically. A compromise time window length avoiding the inclusion of S-wave in the record but also providing the sufficient level of resolution is desired. The values of 0.10 , 0.15 , 0.20 for the window length and 0.015, 0.045, 0.055, 0.060, 0.065 and 0.075 have been considered. From the visual analysis of the obtained spectra a time window length of 0.015 s and a tapering time of 0.0055 were chosen.

If Fourier spectra are calculated for the raw traces, a pronounced pick at a frequency of about 50 Hz can be noted sometimes. This is probably related to anthropic activity, and can be eliminated by using the Seismic Unix utility Suramp to apply a linear filter in time domain, which becomes zero as frequency get close to 50 Hz.

After filtering, for each trace the logarithm of the Fourier spectrum $\ln A(f)_{ij}$ is recomputed and the corresponding spectral ratio $\Delta \ln A(f)_{ij}$ for each observed trace is calculated. Figure 3.14 illustrates the procedure for the trace referred to the first source recorded at the 1250-th receiver.

Once the amplitude spectra have been calculated, intercepts q_1^{ij} and slopes m_1^{ij} of the straight lines that describe the trend between spectral ratio and frequency for each j-th receiver and i-th source are estimated by least squares method (figure 3.15)

To describe the lateral variations of seismic waves attenuation, defining the geometry of the problem is necessary. As data concern a seismic profile, with sources and receivers disposed approximately along a line (figure 3.12), the tomographic problem is bi-dimensional. One dimension is the depth (z) and the other (x) is defined fitting a straight line from the sources and receiver geographic coordinates. The x coordinate of sources and receivers is calculated by projecting the points on the straight line. Sources and receivers elevation depends on topography. The maximum elevation point corresponds to the zero of z axis.

The area of interest is rectangular and measures about 1 km in dept and 11 km in length. It is parametrized trough a rectangular grid of 10 x 113 squared cells with side length of 0.1 km. Each cell of the grid is characterized by a velocity and a Q-value that are constant inside the cell.

A laterally homogeneous velocity model is assumed. It is constituted by 10 horizontal layers whose velocity increases with depth from 3.025 to
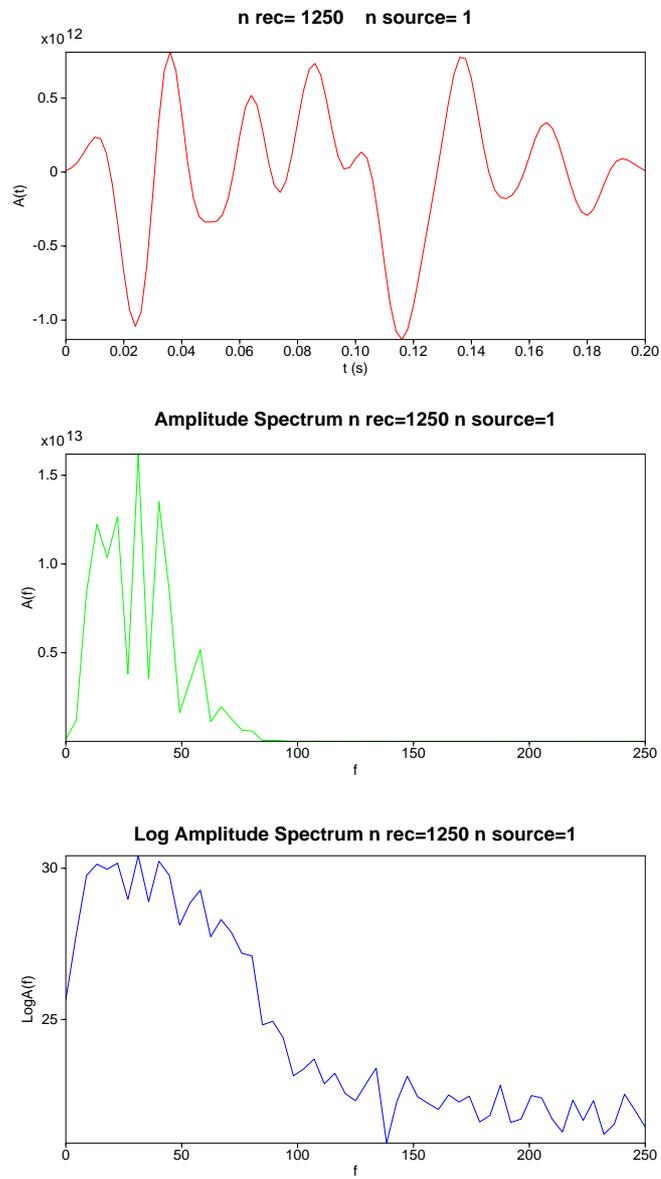
Figure 3.14: Amplitude spectrum computation for trace number 1250 and source 1. Seismic trace (red), with signal amplitude approaching to zero at time window bounds due to linear tapering. Amplitude spectrum (green) and logarithm of amplitude spectrum (blue).
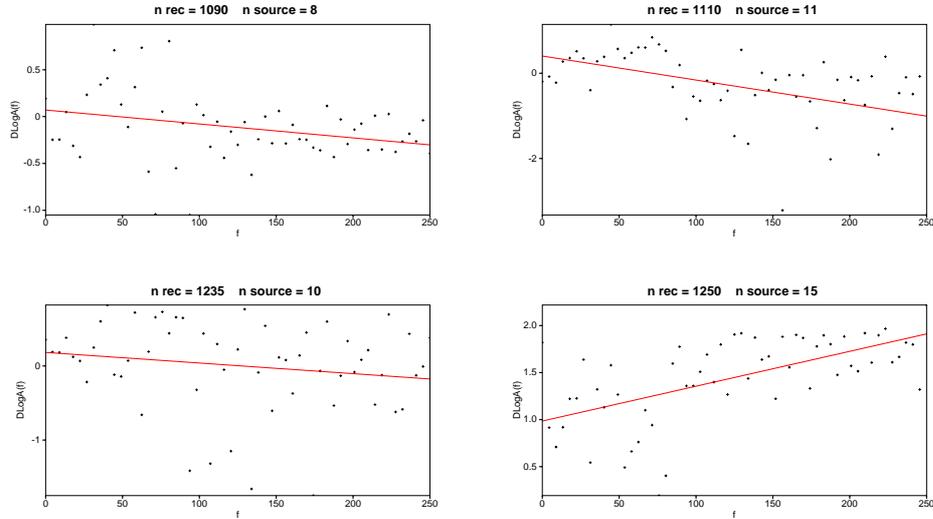
Figure 3.15: Examples of linear fit between spectral ratio and frequency.

3.475 km/s (figure 3.16). The distance traveled by the waves in each cell
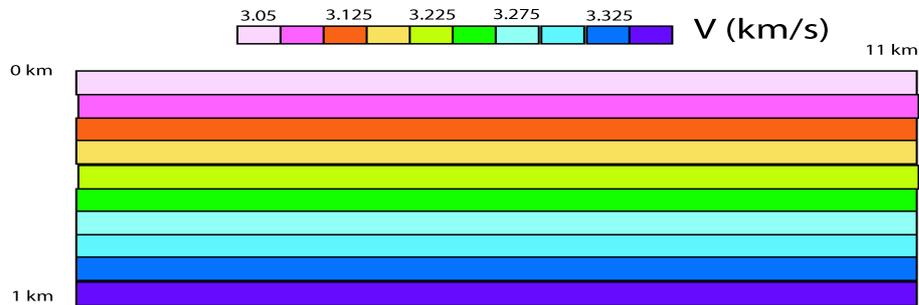
Figure 3.16: P-wave velocity model.

can be calculated by the program Ray_Trace (Carbonell, personal communication) if the geometry of the problem, the velocity model, the shots and receivers positions are known. Figure 3.17 shows the ray coverage of the zone of interest. The inversion problem is solved using the same algorithm applied to intensity tomography. The data refers to depths lower than 1 km and refer to a wide range of rock types (see figure 3.13) that would correspond to different Q values. However, on the basis of the average Q values for P-waves presented in literature (e.g. Cichowizc et al. (1990), Spottiswoode (1993), Feustel et al. (1993)) an average value of 300 is chosen for reference.

The choice for the damping value influences the weight of the reference
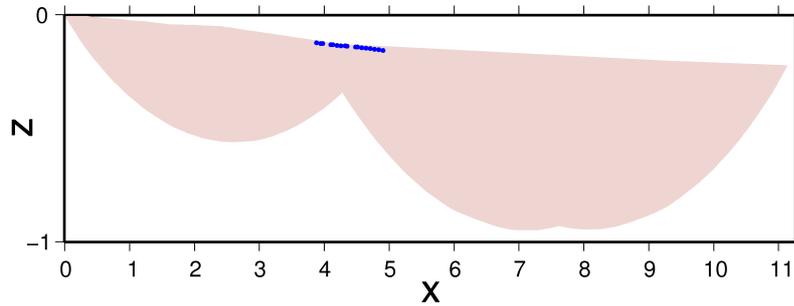
Figure 3.17: Ray coverage. Cells crossed by at least one ray are in pink. Sources are represented with blue points.

model in the evaluation of model parameters. Different damping values have been considered, ranging from 0.01 to 20 and a value of $\lambda = 0.5$ is chosen. Figure 3.18 shows that as the damping value increases, the solutions are constrained to assume values closer to the reference model.
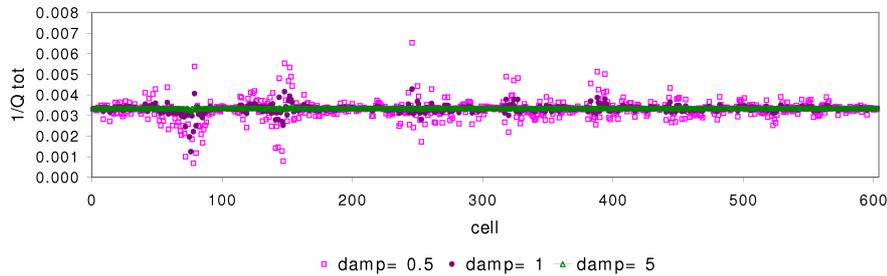
Figure 3.18: Effect of different choices for the damping value.

3.3.6 Q-tomography results and discussion

Figure 3.19 shows the seismic quality factor obtained. Close to the sources position, strong variations of the seismic quality factor are observed. Moving away from such zone, the range of variations decreases but their spatial frequency is high and does not allow to describe clearly the attenuation structure. Lateral variations in attenuation properties do not seems to be related to the rock type spatial distribution shown in figure 3.13.

Differently from the intensity tomography case, for the Q tomography
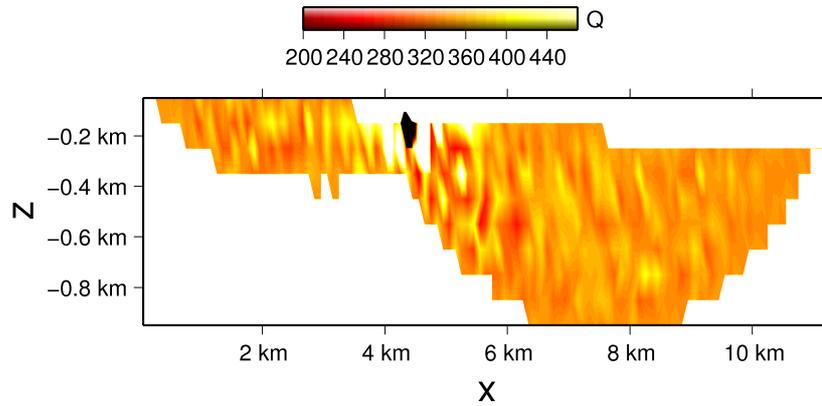
Figure 3.19: *Q*-tomography results.

it is not possible calculate the errors by eq. 3.5 because of negative components in the matrix G_{np} of eq. 3.33. A possible explanation of the lack of a clear attenuation structure is that the noise in recorded seismic trace causes strong fluctuations of the spectral ratio with frequency, hiding its expected linear trend with frequency.

The spectral ratio method is based on the assumption that attenuation is frequency independent. This is a common assumption for earthquakes frequency range (Aki and Richards, 2002), but it could not be valid for the frequency band involved in exploration seismology and in particular for the 8-80 Hz range covered by the sources related to the data used in the present work.

Another possible explanation is that, for the range of distances considered in this study, the geometrical spreading effect prevails on the anelastic attenuation in causing the seismic wave attenuation.

Chapter 4

CONCLUSIONS

In chapter 2 the attenuation pattern of macroseismic intensity in Italy has been analyzed to develop an attenuation relationship for the probabilistic seismic hazard assessment in Italy. Hence, major attention has been paid to the characterization of the attenuation relationship in its complete probabilistic form. Moreover, several open issues concerning seismic intensity and its use as shaking parameter have been reconsidered.

The intensity that is observed at a certain site is assumed to depend on source strength and on its distance from source, which is calculated assuming a unique hypocentral depth for all the events. Differently from previous intensity attenuation studies (Gasperini, 2001; Albarello and D'Amico, 2004), the average hypocentral depth is deduced directly from data. The value obrained is $h = 3.91 \pm 0.27$, which is significantly shallower than the average hypocentral depth of strong Italian earthquakes (about 10 km) but quite similar to that (5.6 km) deduced by the attenuation of PGA (Sabetta and Pugliese, 1987). Probably, it reflects a shallower point from which seismic energy appears to be radiated at close sites, rather than the true (and deeper) hypocenter.

In the present work the parameter that describes the source energy is estimated from the whole macroseismic field available for each earthquake. In most previous analyses, this role was played by the epicentral intensity I_0, which is usually determined a priori by catalog compilers. By contrast, in this work, a new estimate of radiated energy in terms of the expected intensity at the epicenter I_E for all earthquakes is introduced. As it is deduced by considering the entire macroseismic field and not from a single or few intensity observations, it is less influenced than I_0 by local site effects or intensity assessment errors.

A stability analysis shows that the only parameter seriously affected by arbitrary assumptions underlying the modeling approach is the average source depth h. However, its variability has little influence on the

other attenuation parameters and the standard deviation of the regression and thus is not relevant for computing seismic hazard. Moreover, the stability analysis showed that uncertain intensities (e.g. III-IV) behave similarly to standard intensity estimates and are consistent with semi-integer intensity values (e.g. 3.5). This could justify the use of semi-integer values in simplified approaches.

The attenuation model determined in this chapter enables a number of difficulties relative to previous estimates to be overcome. As the relevant standard deviation is lower than those obtained in previous works (Gasperini, 2001; Albarello and D'Amico, 2004) and in particular is comparable to the intrinsic standard deviation related to the scattering of original data (aleatory uncertainty), overestimation of seismic hazard is prevented.

The best-fit values and standard errors of attenuation coefficients are $a = -0.0086 \pm 0.0005$ for the linear term, $b = -1.037 \pm 0.027$ for the logarithmic one. The linear terms significantly differs from 0. This indicates that the contribution of anelastic dissipation to intensity attenuation is not negligible with respect to geometrical spreading. The coefficient of the logarithmic term, close to -1 for the log-linear model and around -0.8 for the log-bilinear model, implies that the geometrical spreading exponent should range from -0.70 to -0.35, depending on the attenuation model (log-linear or log-bilinear) and on the assumed coefficient of the linear relation between intensity and the logarithm of PGA. This confirms that surface waves (and perhaps reflected and refracted phases), rather than body waves, are likely to have a dominant role to play in determining the seismic intensity observed at a site.

The decision to consider a unique attenuation pattern for the whole Italian area could be considered a basic limitation of the previous analysis. In fact, the peculiar geostructural setting of the Italian peninsula should reflect differentiated attenuation patterns and, in this respect, the attenuation relationship here obtained cannot be considered as a definitive characterization of macroseismic fields in the study area. However, the definition of such a reference relationship is a basic preliminary step towards a regionalization of the area under study by using objective quantitative criteria. In chapter 3 the lateral variations in the attenuation of seismic intensity are studied by applying the tomographic method. The unknown parameter values are estimated by damped least squares method. Although the functional form of the attenuation law is chosen on the basis of empirical criteria between different options, the good correlation of the attenuation coefficients estimates with the surface heat flow map of Italy demontrate that also the log-bilinear model is physically grounded. The decrease of the absolute value of the coefficient of the linear term in the refrence model from -0.0187 to -0.0108 for dis-

tances longer than 45 km may be due to subcrustal path of the waves that influence the intensity observed at those distances. The tomography results describe the lateral variations of the coefficients of the linear term respect to the reference model values. The term proportional to the logarithm of the distance is considered instead constant, as it can be related to geometrical spreading, which is assumed independent on lateral variations in the geostructural features of the investigated area. The reliability of the tomography result is confirmed by the estimations errors, which affects on average for 10-13% the parameters values.

An additional confirmation of good reliability of results is given by the analysis of residuals. Their values decrease with respect to those associated to the average reference model, moreover their spatial distribution is more uniform in comparison with those resulting from the isotropic model. The remainig discrepancies between observations and prediction of the tomographic model may be caused by topography and local site characteristics. In fact in the considered case the resolution of the tomography is 25 km, whereas the scale of such local variations may be smaller than the cell area.

Intensity data temporal coverage for Italy spans from 1200 A.D. to present days, whereas the time interval covered by instrumental data is of the order of few tens of years. Tomography results can be used to integrate attenuation studies based on instrumental data, expecially for zones characterized by a low seismicity rate. Assuming a linear dependence between intensity and logarithm of ground motion, lateral variation of the seismic quality factor value have been considered.

Appendix A

seismic attenuation law and ground motion amplitude

Seismic wave amplitude generally tends to decrease with distance from source. In particular, geometrical spreading and anelastic dissipation are relevant in determining seismic waves attenuation. Moreover, multipath scattering across crustal discontinuities and near-source effects may affect significantly the observed wave amplitude.

Geometrical spreading is the seismic wave energy decrease related to the wavefront surface increase. In fact as wavefront expands with time, the total energy on the surface remains constant, whereas the energy per unit surface area decreases. Wavefront of body and surface waves could be assumed respectively spherical and cylindrical, at least at the first order. This implies that wave energy decreases as seismic wave propagates proportionally to R^{-2} in the first case and to R^{-1} in the second one. As wave amplitude is proportional to the square of seismic energy, it depends on distance as R^{-1} for body waves and $R^{-0.5}$ for surface waves.

Another possible cause of attenuation of a seismic waves is the interaction with boundaries between materials characterized by different seismic velocity. For example, when compressional (P) or shear waves that vibrates in the vertical plane (SV) encounter a velocity boundary, they split in four derivative waves (the refracted P, refracted SV, reflected P and reflected SV) whose geometry depends on the Snells law. The splitting of a single incident wave in two or more reflected and refracted ones implies that its energy will be partitioned between the derived waves.

Geometrical spreading and reflection and transmission of energy at boundary would be the only mechanisms controlling the amplitude of seismic pulse if the Earth would be perfectly elastic. This is not the case, in fact part of the seismic wave energy it is lost due a variety of phenomena collectively described as internal friction or anelastic attenuation.

A phenomenological model for seismic attenuation is represented by a spring attached to a mass sliding on a surface. The anelastic attenuation of amplitude as function of distance x is given by

$$A(x) = A_0 \exp\left(-\frac{f\pi}{Q_{\mathrm{an}}V}x\right) ,$$

where A_0 is the amplitude at distance $x = 0$, V is the velocity, f is the frequency and

$$\frac{1}{Q_{\mathrm{an}}} = -\frac{\Delta E}{2\pi E} .$$

Q_{an} is the seismic quality factor and represents the fractional energy loss per cycle of oscillation.

Earth interior presents small scale heterogeneities. When a seismic wave interact with such irregularities, high frequency wave field is partitioned in a sequence of arrival called coda waves. This could decrease the energy of direct waves, shifting the energy back in to the coda arrivals. The wave interaction with small scale heterogeneities involve the conventional effects of refraction, reflection, conversion and diffraction that characterize the propagation of seismic waves in an elastic Earth, but the resulting overall wave field is so complex that individual arrivals can not be associated to singles velocity discontinuities, so they are described by an exponential attenuation quality factor Q_{sc}. As it is difficult distinguish between scattering and anelastic dissipation both phenomena could be summarized by a unique seismic quality factor Q (Lay and Wallace, 1995).

Assuming that geometrical spreading, anelastic attenuation and scattering are the main causes of seismic wave attenuation, the amplitude decrease with distance R can be described by

$$A(R) = A_0 R^{-n/2} \exp\left(-\frac{\pi R}{TQV}\right) \qquad (A.1)$$

where the term $R^{-n/2}$ accounts for geometrical spreading with $n = 2$ for body waves and $n = 1$ for surface waves and the exponential term summarizes the effects of anelastic attenuation and scattering.

Although seismic intensity is an index based on qualitative description of seismic effects, it may be related to ground motion amplitude. Cancani (1904) proposed a linear relation between intensity and the logarithm peak ground acceleration, and Sieberg compiled the Mercalli-Cancani-Sieberg (MCS) macroseismic scale aiming at a linear dependence between these two quantities. Empirical investigations in many parts of the world seem to confirm the existence of such a linear relation (Margottini et al., 1992; Boatwright et al., 1994; Wu et al., 2003; Kaka and Atkinson, 2004).

Assuming that the source depth is h, I_0 could be defined as the intensity at the epicenter $(D = h)$. In this case the intensity attenuation law could be formulated in terms of deviation ΔI between epicentral intensity I_0 and intensity observed at a certain distance from the source $I(D)$. Assuming that

$$\log_{10} A = l + kI \,, \tag{A.2}$$

and substituting eq. A.1 gives

$$
\begin{aligned}
\Delta I &= I(h) - I(D) \\
&= \frac{1}{k}\left[\log_{10} A(h) - \log_{10} A(D)\right] \\
&= \frac{1}{k}\left[\frac{n}{2}\log_{10}\frac{D}{h} + \log_{10} e \frac{\pi}{TVQ}(D - h)\right]
\end{aligned} \tag{A.3}
$$

Assuming that one intensity degree corresponds to a doubling of the amplitude (Cancani, 1904), gives

$$\frac{1}{k} = \frac{1}{\log_{10} 2} \approx 3 \,.$$

Another possible assumption is that intensity is mainly related to body waves, whose spreading coefficient is $n = 2$. Substituting in A.3 gives the attenuation law proposed by Koveslighety (1906)

$$\Delta I = I_0 - I = 3\log_{10}\frac{D}{h} + 3\alpha(\log_{10} e)(D - h) \tag{A.4}$$

with $\alpha = \pi/(TVQ)$.

Bibliography

H. Akaike. A new look at the statistical model identification. *IEEE Trans. On Automatic Control AC*, 19:716–723, 1974.

K. Aki and W.H.T. Lee. Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes. 1. a homogeneous initial model. *J. Geophys. Res.*, 81: 4381–4399, 1977.

K. Aki and P. Richards. *Quantitative seismology*. University Science Book, California, second edition, 2002.

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

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

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

R. Azzaro, M.S. Barbano, S. D'Amico, and T. Tuvè. The attenuation of seismic intensity in the Etna region and comparison with other Italian volcanic districts. *Annals of Geophysics*, 2006. in press.

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

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

M. Bath. *Spectral Analysis in Geophysics*. Elsevier, Amsterdam, 1974.

B. Bender and D.M. Perkins. Seisrisk III: a computer program for seismic hazard estimation. *U.S Geological Survey Bullettin, Washington*, 1772: 48, 1987.

R. Berardi, C. Petrungaro, L. Zonetti, L. Magri, and M. Mucciarelli. *Mappe di sismicit per l'area italiana*. ISMES, Bergamo, Italy, 1993.

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

J. Boatwright, K. Thywissen, and L.C. Seekins. Correlations of ground motion and intensity for the 17 January 1994 Northridge, California, earthquake. *Pageoph*, 143:617–632, 1994.

E. Boschi, E. Guidoboni, G. Ferrari, G. Valensise, and P. Gasperini. *Catalogo dei Forti Terremoti in Italia dal 461 a.C. al 1980*. ING/SGA, Bologna, Italy, 1995. enclosed CDROM.

E. Boschi, E. Guidoboni, G. Ferrari, G. Valensise, and P. Gasperini. *Catalogo dei Forti Terremoti in Italia dal 461 a.C. al 1990*. ING/SGA, Bologna, Italy, 1997. enclosed CDROM.

E. Boschi, E. Guidoboni, G. Ferrari, D. Mariotti, G. Valensise, and P. Gasperini. Catalogue of strong Italian earthquake from 461 B.C. to 1997. *Annali di Geofisica*, 43:843–868, 2000. enclosed CDROM.

D.R. Brillinger. Seismic risk assessment: some statistical aspects. *Earthq. Predict. Res.*, 1:183–195, 1982.

A. Cancani. Sur l'emploi d'une double echelle sismique des intensits, empirique et absolute. *G. Beitr., Erganzungsband*, 2:281, 1904.

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

S. Castellaro, F. Mulargia, and Y.Y. Kagan. Regression problems for magnitudes. *Geophys.J.Int.*, 165:913–930, 2006.

C. Cataldi, F. Mongelli, P. Squarci, L. Taffi, G. Zito, and C. Calore. Geothermal ranking of italian territory. *Geothermics*, 24:115–129, 1995.

I. Cecic and R. Musson. Macroseismic surveys in theory and practice. *Natural Hazards*, 31:39–61, 2004.

I. Cecic, R. Musson, and M. Stucchi. Do seismologists agree upon epicentre determination from macroseismic data? A survey of the ESC 'Macroseismology' Working Group. *Annali di Geofisica*, 39(5):1013–1027, 1996.

U. Chandra, J.C. McWhorter, and A.A.Nowroozi. Attenuation of intensities in Iran. *Bull. Seism. Soc. Am.*, 69:237–250, 1979.

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

A. Cichowizc, W.E. Green, A. Van Zyl Brink, P. Grobler, and P.I. Montfort. The space and time variation of microevent parameters occurring in front of an active slope. In C. Fairhurst, editor, *Rockbursts and Seismicity in Mines*, pages 171–175. A.A. Balkema, Rotterdam, 1990.

J.K. Cohen and Jr.J.W Stockwell. Cwp/su: Seismic un*x release no 37 an open source software package for seismic research and processing. avaible from http://www.cwp.mines.edu/cwpcodes/index.html, 2002.

A.M. Cormack. Representation of a function by its line integrals, with some radiological applications. *J. Appl. Phys.*, 34:2722–7, 1963.

C.A. Cornell. Engeneering seismic risk analysis. *Bull. Seism. Soc. Am.*, 58:1583–1606, 1968.

C.A. Cornell. Probabilistic analysis of damage to structures under seismic loads. In D.A. Howells, I.P. Haigh, and C. Taylor, editors, *Dynamic Waves in Civil Engineering*. Wiley, London, 1991.

CPTI Working Group. *Catalogo Parametrico dei Terremoti Italiani CPTI)*. Ed. Compositori, Bologna, Italy, 1999.

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

DBMI Working Group. *Database macrosismico utilizzato nella compilazione di CPTI04)*. INGV, Milano, Italy, 2007. also avaible from http://emidius.mi.ingv.it/DBMI04/.

B. Della Vedova, F. Mongelli, G. Pellis, P. Squarci, L. Taffi, and G. Zito. Heat flow map of italy. Technical report, International Institute for Geothermal research, CNR-Pisa, 1991.

J.E. Dennis and R.B. Schnabel. *Numerical method for unconstrained optimization and nonlinear equation*. Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

D. Draper. Assessment and propagation of model uncertainty (with discussion). *J. Royal Stat. Soc., Series B*, 57:45–97, 1995.

R. Draper and H. Smith. *Applied regression analysis*. Wiley, New York, second edition, 1981.

B. Efron and R.J. Tbishirani. Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Stat. Sci.*, 1:54–77, 1986.

R.J. Evans and U. Aucher. Teleseismic velocity tomography using the ACH method: theory and application to continental-scale studies. In H.M. Iyer and K. Hirahara, editors, *Seismic Tomography: theory and practice*. Chapman & Hall, London, 1993.

A.J. Feustel, T.I. Urbancic, and R.P. Young. Estimates of q using the spectral decay technique for seismic events with m ¡ -1. In R.P. Young, editor, *Rockbursts and Seismicity in Mines*, pages 337–341. A.A. Balkema, Rotterdam, 1993.

W.A. Fuller. *Measurement error models*. John Wiley, New York, 1987.

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

P. Gasperini. Appendice 1. In M. Stucchi, editor, *Redazione della mappa di pericolosit sismica prevista dallordinanza PCM del 20 marzo 2003, n. 3274, Rapporto conclusivo*, page 29. INGV, Milan, 2004. available at http://zonesismiche.mi.ingv.it/documenti/App1.pdf.

P. Gasperini and G. Ferrari. Deriving numerical estimates from descriptive information: the computation of earthquake parameters. *Annali di Geofisica*, 43:729–746, 2000.

P. Gasperini, F. Bernardini, G. Valensise, and E. Boschi. Defining seismogenic sources from historical earthquakes felt reports. *Bull. Seism. Soc. Am.*, 89:94–110, 1999.

Z. Guo and Y. Ogata. Statistical relations between the parameters of aftershock in time, space and magnitude. *J. Geophys. Res.*, 102:2857–2873, 1997.

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

P. Hall. *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York, 1992.

G.N. Houndsfield. Computerized transvers axial scanning tomography: Part I. description of the system. *Br.J.Radiol.*, 46:1016–22, 1973.

C.M. Hurvich and C.L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76:297–307, 1989.

IGME. Mapa geológico de españa. 1:50000. hoja 710, retuerta del bullaque, 1986.

IGME. Mapa geológico de españa. 1:200000. hoja 53, toledo, 1970.

H.M. Iyer and K. Hirahara, editors. *Seismic tomography: theory and practice*. Chapman & Hall, London, 1993.

S.I. Kaka and G.M. Atkinson. Relationship between instrumental grouund-motion parameters and Modified Mercalli intensity in Eastern North America. *Bull. Seism. Soc. Am.*, 94:1728–1736, 2004.

R. Koveslighety. Seismonomia. *Boll. Soc. Sism. Ital.*, 13, 1906.

T. Lay and T.C. Wallace, editors. *Modern Global Seismology*. Academic Press, San Diego, California, 1995.

C. López-Casado, S. Molina Palacios, J. Delgado, and J.A. Peleaez. Attenuation of Intensity with Epicentral Distance in the Iberian Peninsula. *Bull. Seism. Soc. Am.*, 90:34–47, 2000.

L. Magri, M. Mucciarelli, and D. Albarello. Estimates of site seismicity rates using ill-defined macroseismic data. *Pageoph.*, 143:617–632, 1994.

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

R.K. McGuire, editor. *The practice of earthquake hazard assessment*. IASPEI, Denver, 1993.

G. Monachesi and M. Stucchi. DOM4.1, un database di osservazioni macrosismiche di terremoti di area italiana al di sopra della soglia del danno. Technical report, GNDT, Milano-Macerata, 1997. Open file report, also avaible from http://emidius.mi.ingv.it/DOM/home.html.

R.M.V. Musson. Determination of parameters for historical british earthquakes. *Annali di Geofisica*, 39:1041,1047, 1996.

R.M.V. Musson. Intensity attenuation in the U.K. *J. Seism.*, 9:73–86, 2005.

J. Radon. Uber die bestimmung von funktionen durch ihre integralwerte langs gewisser mannigfaltigkeiten. *Ber. Verh. Saechs. Akad. Wiss. Leipzig. Math. Phys. Kl*, 69:262–77, 1917.

A. Rebez and M. Stucchi. Determinazione dei coefficienti della relazione tabellare Io-Ms. In *Catalogo Parametrico dei Terremoti Italiani*, pages 20–21. Ed. Compositori, Bologna, 1999.

F. Sabetta and A. Pugliese. Attenuation of peak horizontal acceleration and velocity from Italian strong-motion records. *Bull. Seism. Soc. Am.*, 77:1491–1513, 1987.

G. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.

A. Sieberg. Erdebeben. In B. Gutengerg, editor, *Handbuch der Geophysik*, volume 4, pages 552–554. 1931.

S.M. Spottiswoode. Seismic attenuation in deep-level mines. In R.P. Young, editor, *Rockbursts and Seismicity in Mines*, pages 409–414. A.A. Balkema, Rotterdam, 1993.

N.R. Tilford, U. Chandra, D.C. Amick, R. Moran, and F. Snider. 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. Seism. Soc. Am.*, 75:923–937, 1985.

Visual Numerics. *IMSL Math/Library Volumes 1 and 2*. Houston, 1997.

D.L. Wells and K.J. Coppersmith. New empirical relationships among magnitude, rupture length, rupture width, rupture area and surface displacement. *Bull. Seism. Soc. Am.*, 84:974–1002, 1994.

Y.M. Wu, T. Teng, and N.C. Hisiao. Relationships betwees Peak Ground Acceleration, Peak Ground Velocity and Intensity in Taiwan. *Bull. Seism. Soc. Am.*, 93:386–396, 2003.