Deriving numerical estimates from descriptive information: the computation of earthquake parameters

Paolo Gasperini (¹) and Graziano Ferrari (²)
(¹) Dipartimento di Fisica, Università di Bologna, Italy
(²) SGA, Storia Geofisica Ambiente, Bologna, Italy

Abstract
We describe here how the origin time, the epicenter location, the epicentral intensity and the magnitude of the earthquakes included in the Catalogue of Strong Italian Earthquakes (CFT13) have been determined from the documentary information available. As the literature lacks widely accepted formal definitions for these parameters in the case of macroseismic data, we developed ad hoc procedures that starting from generally shared qualitative statements give reasonable and stable results which can also be useful for statistical analysis. These procedures use all of the information available in the CFT13 database as much as possible and make use of «robust» estimators in order to minimize the effects of anomalous data. The resulting parametric catalogue has been examined statistically for completeness showing that the great majority of the Italian earthquakes with moment magnitude larger that 5.5-6.0 of the last four centuries are correctly reported in CFT13.

Key words macroseismic epicenter – epicentral intensity – macroseismic magnitude

1. Introduction

Summarising all the information available about an earthquake within the meagre numerical quantities of a catalogue record is certainly a limiting operation as regards the complexity of the results of historical and seismological studies. It is nevertheless necessary to supply a summary representation, both to provide an index of reference for access to the rest of the information in the Catalogue of Strong Italian Earthquakes (CFT13), and to create a tool which may be of immediate use for statistical analysis.

While in instrumental catalogues each single record always represents a single shock, it should be noted that here it may sometimes instead refer to the overall effects of a seismic period. This is because it is not always possible to distinguish the effects between a sequence of shocks, either because they are very close in time and space or simply because, particularly for those events furthest away in time, historic sources do not permit separation of the accumulated effects of the series of shocks.

In the scientific community, macroseismic data are often considered as purely qualitative information, not very reliable and in general unworthy of the qualification of experimental data. This belief has led many seismologists to reject this information, considering it to be non-scientific. Admittedly in the past the criteria of estimation of epicentral and energy parameters for historical earthquakes have rarely been given official standing and have often made use of quality techniques of assessment, also because in many cases the information was not available in a form which lent itself to analytical processing. This work describes the procedures used...
for the definition of the origin time, the epicentral coordinates, the epicentral intensity and the equivalent macroseismic magnitude reported in the concise records of the CFTI3 database. There is also a description of the criteria for assessing the reliability of these estimations, plus an estimation of the error associated with the magnitude value. No estimation has been made of the focal depth however, since the computation of this parameter on a purely macroseismic basis (e.g., Blake, 1941; Shebalin, 1973) was considered too uncertain to allow a reliable estimate to be made, also due to the absence of a reliable database of instrumental reference.

2. Origin time of earthquakes

The origin time reported by the catalogues is often considered an invariable factor of the earthquake, to such an extent that it is given the value of a code of identification. In principle this would be valid if the date and time of an earthquake had so far been determined with accuracy. In practice, however, this is not the case, since catalogues often do not take into account the conversions of date and time necessary to align all historical earthquakes with the same system of time reference—taken by convention as Greenwich Mean Time (GMT) in earthquake catalogues. In Italy, in fact, both historically and geographically a series of different styles of dating and chronometric scales have been used. For this reason the transposition of historic dates and times into GMT requires careful examination and specific methods of conversion which were unknown to some compilers of previous catalogues. In the most serious cases, this lack of knowledge has led to the duplication of earthquakes even of high intensity. The problem has been brought to light by few authors (Dominici and Marcelli, 1979; Ferrari and Marmo, 1985) and has not yet been fully absorbed into scientific practice. In the CFTI3, special attention to this aspect has led to a correction of the origin time of 35% of the 600 earthquakes analysed. For about 15% of the earthquakes of the CFTI3 the time was included, absent in previous catalogues, for about 19% of the events the correction concerned the hours, while only the minutes were adjusted in 8%. The greatest incidence of corrections concerned the period up to and including the eighteenth century.

3. Date

The exact definition of the date of occurrence of the events being studied, from the descriptions in the historical sources, has highlighted two problems: the difficulty in identifying a univocal date and that in converting to current dating usage.

The first problem emerges in those earthquakes in the remote past whose date may be identified within a chronological range of just a few years in some cases. This applies for example to the earthquake recorded in an epigraph about the restoration of damage to the Coliseum in Rome, considered to have occurred just before 484 or just before 508. The only element for dating is found in the epigraph which makes reference to the consul Decius Marius Venantius Basilius, who, perhaps due to a case of homonymy, figures in the consular lists of both years. Seismic events with these problems of chronology are shown with two date-fields in the database.

As regards conversion to current dating usage, without mentioning individual cases which are dealt with in the specific historical-critical comments, it is useful to summarise some types of problem here.

The main dating errors for earthquakes in the catalogues are caused by rough inaccuracies in converting the date given in the evidence.

Corrections made to the year, the month or the day have different origins and motivations. On the one hand there are the special cases of dating which are affected by errors of transcription by previous catalogue compilers, on the other there are all those cases in which the earthquakes have been dated without taking into account dating styles other than the current usage, to which sources implicitly refer. This aspect may in some cases cause considerable difficulty and has in the past led to a number of blunders.

Until the seventeenth century, the dating style tended to vary considerably depending on the
historical period, the geographical area and sometimes even the type of source (Cappelli, 1998). We may in fact calculate that 7 styles of dating have been used historically in Italy:

<table>
<thead>
<tr>
<th>Style</th>
<th>Effect on Year Start</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) current or Circumcision</td>
<td>1st January</td>
</tr>
<tr>
<td>2) more veneto</td>
<td>March 1st</td>
</tr>
<tr>
<td>3) Florentine style of the Incarnation</td>
<td>March 1st with which it coincides until 31st December</td>
</tr>
<tr>
<td>4) Pisan style of the Incarnation</td>
<td>March 25th</td>
</tr>
<tr>
<td>5) Easter or French style</td>
<td>Easter Day</td>
</tr>
<tr>
<td>6) Byzantine style</td>
<td>September 1st</td>
</tr>
<tr>
<td>7) style of the Nativity</td>
<td>December 25th</td>
</tr>
</tbody>
</table>

Different styles could therefore have been in use in a particular period of history, in different cities and areas of Italy, so that the effects of an earthquake in the different places, though practically synchronous, are not documented as such (fig. 1). In the first half of fifteenth century, for example, the Byzantine style was in use in Puglia and Calabria, in Venice the more veneto, in Florence the Florentine Incarnation style, in Rome the style of the Nativity, in Pisa the Pisan Incarnation style and in Turin the Nativity style.

Between the late eighteenth century and the early nineteenth century, an Italian daily time system was used alongside a French official usage. This led to frequent confusion between the two styles in the sources when identifying the day. Consider this source, for example: «in

![Fig. 1. Different styles of dating historically used in Italy.](image-url)
Bologna at 3 in the night of January 25th there was a violent tremor...». In Italian usage, January 25th began at eventide of the 24th, so 3.00 would correspond roughly to 20.00 GMT of the 24th. However, for various reasons some contemporary or later chroniclers might indicate the time in the Italian manner, but the day in the French style. In this case therefore the time would be 20.00 GMT of the 25th. Such situations can be dated with more accuracy where the sources provide more explicit time references: «in the night between the 24th and the 25th...» or «the night of the 24th going on for the 25th...».

4. Time

The measuring scales of time by which were regulated in the past human activities and which he has used to describe the natural phenomena observed have always made reference to the apparent movement of the sun with respect to the Earth. The intervals of time between dawn and sunset and vice versa have identified a diurnal and nocturnal cycle in the scansion of rhythms of human life. This distinction, which has no significance in the chronometric scale of today, was widely used in the past (fig. 2). The Romans, in fact, divided each of these periods into four parts of three hours each. These four parts were terminated respectively by the hours tertia, sexta, nona and duodecima (this latter in the diurnal cycle corresponds to sunset – solis occasus). Since the apparent diurnal cycle of the day is not regular throughout the year, the measurement (in hours) of the dies naturalis remains constant, while the unit of measurement varies (the hour). After the fall of the Western Roman Empire (late fifth century) this calculation of time was retained by the Christian church to mark its rituals, with the innovation of calculating the day from one sunset to the next.

The introduction at the start of the fourteenth century of bell-ringing mechanisms operated by large clocks in the towers and bell-towers of towns helped to spread a more accurate perception of the daily scansion of time, thus forming a greater synchronisation of human activity. Thus a usage of time named «Italian style» became more common, in which a day was taken as 24 h from eventide (half an hour after sunset). This time style was used to refer to the majority of earthquakes in this catalogue since with some exceptions it was in use at least from the fourteenth century up to the first decades of the nineteenth century, though in some areas even for some decades following that.

Fig. 2. Chronometric scales in use historically in Italy. In Italy, from the antiquity to the second half of the 19th century, different measuring scales of time succeeded and partly overlapped. Only starting from 1866 has Italy a modern measuring scale for the whole national territory. From this date to 1894 Rome Mean Time was therefore replaced by the Europe Mean Time (CEMT).
Other styles followed the Italian style, consolidating the passage from a local reference system of chronometric scales to a chronometric scale which was first national and then of international reference (GMT).

As we have mentioned, errors or inaccuracies in converting the date of the events have also led to earthquakes being duplicated, as is the case, for instance, of the earthquake in Slovenia on March 26th 1511 for which reference should be made to the catalogue.

Some conventions have been adopted in the CFTI with regard to the approximations of the origin times of the earthquakes. For earthquakes occurring before and during the nineteenth century, the approximation generally concerns the minutes, unless there are indications from observatories with precision clocks in constant correction. The GMT times from conversions from the Italian time style have generally been approximated to the nearest 5 min (5:18-5:20) because of the low accuracy of the original time determinations.

The need to indicate a time for the tremor, even though with a degree of approximation, has recommended the attribution where possible of the time also to general indications. Time indications reported in texts as «between 6 and 7» have been taken as 6:30. Indications like «before cockcrow», «just before dawn», and so on, have conventionally been attributed to the time of dawn of that day of the year.

5. Epicentral location

While the definition of the instrumental epicenter is well established in seismological literature the definition of macroseismic epicenter is still the matter of debate (see for example Cecić et al., 1996). This is mainly due to the lack of standardisation typical of macroseismic studies, and to the widely shared opinion that these are not in any case able to provide reliable estimates. Charles Richter in his famous textbook (Richter, 1958) stated «The practice, in the absence of seismographs, of drawing isoseisms and then locating an epicenter at the centre of the figure should be discontinued. In the majority of cases the instrumentally located epicenter proves to be at one side of the meizoseismal area».

Nowadays, it is widely accepted that at least for damaging earthquakes ($M > 5.5$) and due to the details of the rupture history of the earthquake causative fault, the macroseismic epicenter can be considerably distant from the instrumental epicenter. Indeed, while the instrumental epicenter corresponds to the projection onto the surface of the focus of the earthquake (the point where enucleation of the rupture occurs), the macroseismic epicenter represents the center of the distribution of major effects of the earthquake that in all likelihood can be related to the projection of the baricentre of the rupture zone.

There are few examples of a formal definition of the macroseismic epicenter reported in literature. Kondorskaya and Shebalin (1982) define it as «the weighted centre of the first isoseismal with error equal to its mean radius». In the Italian catalogue of the Progetto Finalizzato Geodinamica (PFG) (Postpischl, 1985) the macroseismic epicenter is defined as «the baricentre of the area enclosed within the higher isoseismal line. The area was defined considering at least three separate localities. If this was not possible, the approximated area is given by those localities which present the strongest observed effects». In the Introduction of the NT4.1 catalogue of Italian earthquakes Camassi and Stucchi (1997) wrote: «in most cases the epicentral coordinates are determined as the coordinates of the baricentre of the intensity data-point with intensity $I = I_{\text{max}}, I_{\text{max}} - 1$, computed without attributing differential weights». Bakun and Wentworth (1997) proposed a computational technique that, using individual intensity observations, determines the epicenter as the point where the standard deviation between the observed intensities and the ones computed by a circular attenuation function fitted over the same data is minimum.

Although not uniformly stated, the most widely shared definition of the macroseismic epicenter seems thus to be «the baricentre of the area featuring the maximum effects». Using isoseisms, this corresponds to taking the centre of the highest isoline, whereas if intensity data-points are used, the baricentre can be measured
by calculating the mean of the co-ordinates of the points with highest intensity.

In any case some further problems in determining the macroseismic epicenter, due to the particular nature of the intensity data, must also be taken into account. For example: if the maximum intensity is observed in one locality only, assuming it as the epicenter may be misleading if the site anomalously amplifies the earthquake effects. In this case the uncertainty associated with the intensity estimate may prejudice even the accuracy of the location. Then it may be reasonable to use for the location estimate not only the point with maximum intensity but also some other points with lower intensity. On the other hand only the values of the highest intensities seem to be in strict relation with the extension along the seismogenetic fault, while those of lower intensity seem instead to be linked rather to the pattern of propagation of the seismic waves. What is more, while the meizoseismic area could be attracted by possible geological or morphological structures (such as ridges and valleys) which often induce irregular effects of attenuation and amplification, the outermost areas, corresponding to the lowest degrees of the scale, may present azimuthal anisotropy due both to the peculiar geographical configuration of the Italian peninsula, and to possible gaps in the macroseismic information gathered. So it is necessary for the location procedure to provide the best balance of the relative contribution of the various areas of effects. The use of intensities in between classes, values given as an interval where it is not possible to discriminate accurately between two adjacent levels (e.g., VIII-IX), may also be a further source of inaccuracy. Lastly, for the majority of earthquakes, the spatial distribution of the data-points with respect to the epicenter is fairly irregular and it is very common to have highly anomalous data-points, due to material errors of intensity attribution and site location, and to other motivations of a physical, geological, demographic type, etc. It is therefore advisable to use a «robust» statistical estimator (such as the median or the trimmed mean) instead of the common arithmetic mean.

On the basis of all these considerations, we have elaborated a procedure which attempts to limit as far as possible the effects of the different sources of error described.

1) The data are divided into classes of intensity; data in between classes (uncertain) are assigned to the lower intensity class (e.g., an intensity VII-VIII is assigned to class VII).

2) Localities that reported the maximum observed intensity ($I_{\text{max}}$) are set aside for subsequent calculations.

3) If the maximum observed intensity falls in between full values (e.g., IX-X), the localities that experienced the lower full value are selected in addition to all those where $I_{\text{max}}$ was reported (e.g., all intensity IX-X plus all intensity IX data).

4) If the number of selected localities is less than given threshold N, the dataset is extended to include data from the lower intensity classes up to $I_{\text{max}} - K$ degrees (both N and K are parameters to be determined).

6) The epicentral coordinates are calculated by taking the mean of the corresponding values of the co-ordinates of the points selected, using the Arithmetic Mean (ME), the Trimmed Mean (TM) (that is, the average of all values included in the interval between the first and third interquartiles) and the Median (MD).

7) The values of the parameters N and K and the central tendency estimator CTE among ME, TM and MD are selected through comparative analysis.

For this comparative analysis the distances of the epicenter of each earthquake, calculated with given fixed values of the free parameters (N, K and CTE), from the resulting epicenters using all other different values varying within reasonable intervals have been computed. The parameter values for which the sum of these distances is minimum are those which generate epicenters found on average at the centre of the distribution of candidate epicenters and therefore represent the least questionable choice among all considered ones. These values are $N = 3$, $K = 1$, $CTE = TM$. In other words, this means that the least questionable epicenter can be calculated by computing the trimmed mean of the coordinates of the points belonging to the highest class of intensity and including the class immediately below if the number of these points is less than 3. It should be noted that this
6. Reliability and accuracy of the location

Like other procedures of data analysis, the macroseismic estimate of the epicenter is subject to uncertainty. Since there is no specific physical basis supporting the formulations we adopted and we cannot make a direct comparison with the instrumental datum, the associated error cannot be computed either theoretically or empirically. We can only give a rough estimate of it by computing the standard deviations with respect to the mean values of the coordinates of the localities used for the calculation. Although these cannot be taken directly as estimates of the location error since they also depend on the extension of the meizoseismic area, they still provide a criterion for checking the reliability of the estimate. Very high values may in fact indicate both that there are localities with highly anomalous intensity values, or that the information in the epicentral area is incomplete (e.g., epicenters located in the sea or in uninhabited areas).

7. Epicentral intensity

The epicentral intensity of an earthquake $I_0$ is often the only parameter which may be used to estimate the energy released by an historical earthquake. In most cases it is assumed to coincide with the maximum observed intensity, $I_{\text{max}}$, thus implying that the macroseismic epicenter is located where the maximum effects took place. This is the only possible choice when the intensity has been observed in a single locality but when instead a number of felt reports is available for various locations a different approach can be followed in order to mitigate the effects of ground motion anomalous amplification, intensity evaluation errors, site mislocation and spatial incompleteness of the data. In fact if some of these cases occur, $I_0$ could significantly differ from $I_{\text{max}}$. The distinction between $I_{\text{max}}$ and $I_0$ is not univocally defined in literature, and has long been ignored completely. The PFG catalogue (Postpischl, 1985) gives the following definition of epicentral intensity: «the value of the closed isoseismic line which has the highest degree (and which includes at least 3 different points)». It is however added that the intensity reported in the catalogue «...is normally to be considered as $I_{\text{max}}$.» In literature different methods of determining $I_0$ have also been developed on the basis of the choice and the fit of an attenuation relation (also see the relative discussion in Cecié et al., 1996). Usually the result of these calculations is not an integer value, thus contradicting the nature of the macroseismic intensity. To overcome these difficulties, and in line with the criteria described above, we have chosen the simple procedure reported below:

1) If there is just one intensity estimate for the entire earthquake, $I_0$ is taken as equal to the intensity reported for that point.

2) If there are more than one intensity data-point with intensity equal to $I_{\text{max}}$, $I_0$ is taken as $= I_{\text{max}}$.

3) If a single data-point has intensity $I_{\text{max}}$ and at least one point has an intensity greater than $I_{\text{max}} - 1$, $I_0$ is taken as the intensity of this latter point.

4) In all other cases, $I_0$ is taken as $= I_{\text{max}} - 1$.

The $I_0$ value given by this procedure is consistent with the integer (or at most half integer) definition of intensity and, when used in place of $I_{\text{max}}$ to estimate the magnitude of the earthquake (see discussion below), has improved the correlation with the instrumental magnitude, in the set of events used to calibrate the empirical regression.
8. Equivalent magnitude

In this version of the CFTI3, we have chosen to take as a parameter of reference the moment magnitude \( M \) which is derived from the scalar seismic moment \( M_o \) according to the definition given by Hanks and Kanamori (1979)

\[
M = \frac{2}{3} \log(M_o) - 10.7
\]

where \( M_o \) measured in dyne·cm. It is well-known in fact that the scalar seismic moment \( M_o \) is a much more significant estimator of the «size» of an earthquake than any definition of magnitude. The main reason is that this parameter has a specific physical meaning linked to the dimensions of the source, the amplitude of the coseismic deformation and the shear elastic modulus of the rock material where the seismogenic rupture takes place. Moreover the value of the moment magnitude does not undergo effects of «saturation», as occurs instead with other magnitude definitions, since its seismological estimate is performed by making use of the whole spectrum of the elastic energy irradiated and not just a small portion of it. In actual fact, since the seismological determination of the scalar moment requires «modern» seismometric recordings, direct estimates of this parameter are available only for those earthquakes occurring as from the end of the Seventies. The main source of these estimates is currently the CMT database, available both in the printed version (Dziewonski et al., 1983 and subsequent quarterly papers published in Phys. Earth Planet. Int.), and in computerised form from the webserver of Harvard University. As regards Italy, an important source of data is also provided by the work of Ambraseys (1990) moreover two geological-geodetic estimates are also available regarding the earthquakes of Messina in 1908 (Boschi et al., 1992) and Avezzano in 1915 (Ward and Valensise, 1989).

Altogether about fifty evaluations of scalar moments of major Italian earthquakes from 1976 to the present day are available. Unfortunately, only for fourteen of these does the CFTI3 also give good quality macroseismic data which may be used to calibrate the empirical relations. This operation cannot therefore be performed effectively from the statistical point of view, on the basis of direct estimates of seismic moment alone, but must also make use of indirect moment estimates taken from traditional magnitude data.

For this reason, starting with the database of revised magnitudes of the Italian earthquakes of this century by Margottini et al. (1993), the estimate of the seismic moment of each event was initially done, on the basis of least squares regressions with the seismic moments measured directly. Admittedly an estimate of moment which uses a single type of magnitude shares the same limitations of the magnitude value itself. Where, instead, several evaluations of different magnitudes are available, the mean of the values obtained by different definitions of magnitude may overcome these limitations in part since it is the equivalent of performing the estimate on a broader spectrum.

9. Calculation of the moment magnitude for Italian earthquakes having instrumental magnitudes

The procedure suggested by Johnston (1996a) was followed performing polynomial regressions between the logarithm of the seismic moment and the available magnitudes \( M_i \), \( m_b \) and \( M_i^{'} \), weighted with the inverse of the combined standard deviation of the magnitude and the logarithm of the moment. The procedure differs from the work mentioned both as regards the criterion of choice of the form of the regression function, and for the fact that outliers are not removed. As regards the former point, we chose to use a simpler linear law even though for both \( M_i \) and \( m_b \) a slight improvement of the fit could be achieved by including a square term since the linear law is in theory the most appropriate in the range of magnitude below 6.5 (i.e., Johnston, 1996a) where the great majority of our data is included. On the other hand it must be pointed out that no strongly deviant data are actually present in the set examined.

Table I shows the regression coefficients and the parameters of goodness of fit for the various types of instrumental magnitudes. From these relations it is possible to estimate the value of
Table I. Parameters of linear regression between log($M_0$) and various instrumental magnitudes $M$. The form of the relation is $\log(M_0) = aM + b$. The standard deviation (std) is calculated between observed and computed values of moment magnitude estimated according to the relation of Hanks and Kanamori (1979).

<table>
<thead>
<tr>
<th>$M$</th>
<th>n</th>
<th>a</th>
<th>b</th>
<th>$R^2$</th>
<th>std</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_s$</td>
<td>36</td>
<td>0.96 ± 0.06</td>
<td>19.3 ± 0.3</td>
<td>0.88</td>
<td>0.26</td>
</tr>
<tr>
<td>$m_b$</td>
<td>35</td>
<td>1.21 ± 0.15</td>
<td>17.9 ± 0.8</td>
<td>0.63</td>
<td>0.36</td>
</tr>
<tr>
<td>$M_I$</td>
<td>32</td>
<td>1.22 ± 0.08</td>
<td>17.7 ± 0.4</td>
<td>0.87</td>
<td>0.30</td>
</tr>
</tbody>
</table>

the decimal logarithm of the seismic moment for each independent magnitude evaluation given in the database, and therefore also the value of the seismic moment magnitude according to the definition given by Hanks and Kanamori (1979) shown above.

Through determination of the standard prediction intervals (Draper and Smith, 1980), it is also possible to compute the standard error of each seismic moment made using the above linear relations. On average, this ranges from about 0.34 moment magnitude units for $M_I$ to 0.36 for $M_s$ to 0.40 for $m_b$. When more than one magnitude is available for the same earthquake and it is thus possible to calculate the magnitude as the weighted mean (with the inverse of the squared errors) of different estimates, the error of the mean (computed as the square root of the inverse of the sum of the weights), can be reduced to about 0.2 units of magnitude.

10. Calculation of the moment magnitude for earthquakes with intensity data

As in the calculation of the epicenter, also in estimating the seismic moment and magnitude from macroseismic data, a fully consolidated method is not available in literature. In the absence of a well defined physical connection between intensity and ground motion amplitudes, the only acceptable criteria for establishing the validity of an empirical relation between intensity and magnitude is to resort to statistics. Tinti et al. (1987) have described the three main types of approach to the problem as used in literature: 1) calculation of the magnitude on the basis of epicentral intensity $I_o$ alone; 2) calculation of the magnitude as a function of the area of an isoseism, or of the Felt Area (FA); 3) mixed methods using both the epicentral intensity $I_o$ and the isoseisms.

10.1. Methods based on the epicentral intensity $I_o$

Due to the scarcity of available data, the procedure most commonly followed in the past to evaluate the magnitude of pre-instrumental events was to define empirical relations referring to the epicentral intensity alone. Many authors have determined epicentral intensity-magnitude relations for various seismic regions of the globe (e.g., Karmik, 1969; Toppozada, 1975). What is more, this method is practically the only one which has been used in Italy in past years (e.g., Basili et al., 1980; Peronaci, 1982; Tinti et al., 1986, 1987). In some of these works, the area of interest is divided on the basis of seismotectonic or statistical criteria into smaller areas which are analysed separately. More recently, some authors (Sibol et al., 1987; Albarelo and Mucciarelli, 1989) have suggested determining the magnitude separately for each degree of the scale of intensity, using robust estimators. The most recent and comprehensive evaluation of the intensity-magnitude table for Italy, is the one (table II), computed following Rebez and Stucchi (1996), used in the Catalogo Parametrico dei Terremoti Italiani (CPTI Working Group, 1999).

Besides the uncertainty implicit in the estimate of the intensity parameter, it is clear that in the relation between intensity and magnitude
Table II. Intensity magnitude table (computed following Rebez and Stucchi, 1996) used in the Catalogo Parametrico dei Terremoti Italiani (CPTI Working Group, 1999). Outside the intensity interval, from V-VI to XI, covered by the table the linear relation $M = 0.56 I_s + 0.94$, computed by least squares regression from the table itself, can be used.

<table>
<thead>
<tr>
<th>$I_s$</th>
<th>$M$</th>
<th>std</th>
</tr>
</thead>
<tbody>
<tr>
<td>XI</td>
<td>7.1</td>
<td>0.32</td>
</tr>
<tr>
<td>X/XI</td>
<td>6.8</td>
<td>0.31</td>
</tr>
<tr>
<td>X</td>
<td>6.6</td>
<td>0.30</td>
</tr>
<tr>
<td>IX/X</td>
<td>6.3</td>
<td>0.29</td>
</tr>
<tr>
<td>IX</td>
<td>6.0</td>
<td>0.26</td>
</tr>
<tr>
<td>VIII/IX</td>
<td>5.8</td>
<td>0.21</td>
</tr>
<tr>
<td>VIII</td>
<td>5.4</td>
<td>0.28</td>
</tr>
<tr>
<td>VII/VII</td>
<td>5.1</td>
<td>0.45</td>
</tr>
<tr>
<td>VII</td>
<td>4.8</td>
<td>0.45</td>
</tr>
<tr>
<td>VI/VII</td>
<td>4.6</td>
<td>0.49</td>
</tr>
<tr>
<td>VI</td>
<td>4.3</td>
<td>0.39</td>
</tr>
<tr>
<td>V/VI</td>
<td>4.0</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Another problem in using macroseismic data of this type is the arbitrariness and subjectivity inevitably present in the procedure of manual drawing of the isoseismic lines. It is in fact known that different researchers can produce very different results, though the very same data are used (Berardi et al., 1990). This occurs because the procedure of interpolation of the level curves is highly conditioned by the degree of experience of the operator and by his or her personal interpretative convictions. A solution could be to resort to the use of one of the many existing codes for automatic contouring. However, since in this case the entire list of intensity data must be made available, methods based on the direct analysis of the intensity points can be preferred since these do not require an arbitrary interpretation of the data. Examples of application of methods based on felt areas are fairly common for the United States (Street and Turcotte, 1977; Sibol et al., 1987, Bakun and Wentworth, 1997) but not for Italy, where only a few works have so far ventured in this direction (Westaway, 1992).

In any case, these methods are not the definitive solution to the problem of evaluating the magnitude on a macroseismic basis, because in this case too it is not possible to completely overlook the effects of the depth of the source. Indeed, because of the geometric divergence and the decrease with depth of the seismic waves attenuation coefficient in crust and mantle rocks, the deeper the source, the lesser the attenuation of the intensity with the distance. In this way the isoseisms are found to be wider for deeper earthquakes than for shallower ones. In principle, for a source located at the centre of the Earth (supposed solid), the whole surface of the Earth should experience the same level of shaking and thus belong to the same isoseism, in other words it would represent a single isoseismic area.

10.3. Mixed method

This method, originally put forward by Galanopoulos (1961) and then almost forgotten for many years, combines the two criteria described above, which are thus considered together in evaluating the magnitude: the epicentral inten-
sity $I_o$ and the area of the isoseisms. The work of Galanopulos was based on the assumption that, for shocks with the same energy $E$, the epicentral intensity $I_o$ was «almost inversely proportional» to the focal depth $h$, while the Felt Area (FA) would be «approximately proportional to it». It follows that «the product of the Felt Area multiplied by the intensity, $FA I_o$, of earthquakes with the same energy $E$ remains practically constant for any focal depth (Toperczzer, 1953). This means that the quantity $\log(FA I_o)$, used to determine the magnitude of an earthquake has the advantage of being the same for two events of the same magnitude originating at different depths». Thus the magnitude could be obtained by excluding from the calculation the focal depth, the estimate of which is always difficult and often unreliable.

Unfortunately Galanopulos' argument cannot be confidently taken and accepted as it was originally formulated since it is based on formulae which have been deduced through heuristic considerations, with little physical justification. In particular, although it is reasonable to hypothesise that the dependence on depth is inverse for $I_o$ and direct for FA, the exact form of this functional relation (linear, exponential, power) is not known, and may only be deduced statistically. Nevertheless, the underlying principle of the method of Galanopulos, in other words the assumption that the influence of the depth on the estimate of the energy of earthquakes may be reduced by the combined use of the epicentral intensity and the felt area, is valid. Sibol et al. (1987) have applied a variation of this method to determine the magnitude of historical earthquakes in the United States. In this work the form of the functional relation between magnitude, intensity and felt area was chosen by the authors through an automatic procedure, using a selection technique (SAS, 1982) known as «the best of all possible regressions». The «best» model obtained for the estimate of the body wave magnitude $m_b$ for U.S.A. was the following:

$$m_b = a + b I_o^2 + c \log^2(FA)$$

where FA, expressed in km$^2$, corresponds approximately to the area where the earthquake was effectively felt by population, which coincides approximately with the area enclosed by the isoseism of intensity II to IV (depending on day hour and other conditions favouring or opposing the perceptibility of the earthquake).

It is clear that, when information on the spatial distribution of the intensity is available, mixed methods seem to be the best choice. In the case of the CFTI3, this technique may be used for many events and particularly for most of those of the last four centuries. In the practical usage of macroseismic information, however, some methodological problems emerge requiring further study.

First of all, intensity is by definition a quantity of the ordinal type with discrete values and, although Cancani (1904) had hypothesised a logarithmic relation between the degrees of the scale and progressive increases in ground acceleration, it is not however possible to establish a physical relation which guarantees an equidistance of the different degrees and therefore the possibility of using the intensity as a real numerical value. Thus the use of the intensity as a continuous parameter should where possible be limited.

Macroseismic parameters, such as epicentral intensity and the extension of felt areas, are also affected by strong variations due for example to the shape and dimensions of large sedimentary bodies, to different masses of lithological heterogeneity, to differences in building techniques and in the state of preservation of the buildings. This heterogeneity may considerably affect the intensity felt at single sites, whereas it is filtered when determinations for several different localities are used together. Therefore, the use of estimates based on only a few localities should be avoided. Lastly, for the geometry of the problem, the value of the felt area which can be estimated for the highest intensities, is generally not very reliable. In fact, close to the epicenter the surface of equal-shaking centered in the seismic source, of which the isoseism represents the surface section, is almost tangent to the earth surface.

Recently a proposal was made by Frankel (1994) and later picked up by Johnston (1996b) for a functional relation which includes two
terms, one proportional to the logarithm and the other to the square root of the felt area. This relation would in some way be deduced from the theoretic model of the decrease of the amplitude of the seismic waves with the distance due to the anelastic attenuation and the spherical spreading respectively. However some tests performed with the data of the CFTI3 have shown that Frankel’s relation has a decidedly worse fit with the data than that of Sibol’s law. The reason may be that, since Frankel’s formula does not include the epicentral intensity, it is not compensated for the effects of the source depth which instead may be considerably influential, especially on the areas corresponding to low intensities, in relation to the different excitation of surface waves.

On the basis of the above considerations, the form of the relation we used is as developed by Sibol et al. (1987), modified to take into account not only the Felt Area (FA) but all the areas corresponding to the different intensity isoseisms.

The intensity data corresponding to the uncertain degrees VI-VII, VII-VIII and VIII-IX are considered as independent classes. In all the other cases the uncertain degrees are merged to the degree immediately below, since a greater affinity of the mean distances has been found in this sense. The procedure used is as follows:

1) For each event in the database for which a sufficient number of felt reports are available, the macroseismic epicenter and the epicentral intensity \( I_0 \) are determined by using the procedures described above.

2) The distances of all the points from the epicenter are calculated, and from these the mean distances \( R_I \) (taking as estimator the trimmed mean) for each class of intensity.

3) For all of the earthquakes with instrumental magnitude, a bi-variate weighted regression of the law proposed by Sibol et al. (1987) is fitted (see coefficients in table III) for each intensity class

\[
M = a + b I_0^2 + c \log^2(AI)
\]

where \( AI \) is calculated as the area of a circle of radius \( R_I \). This was done using only the radii estimated through at least four data and exclud-

<table>
<thead>
<tr>
<th>Intensity</th>
<th>n</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>( R^2 )</th>
<th>std</th>
</tr>
</thead>
<tbody>
<tr>
<td>FELT</td>
<td>23</td>
<td>4.34 ± 0.29</td>
<td>0.015 ± 0.003</td>
<td>0.022 ± 0.014</td>
<td>0.69</td>
<td>0.21</td>
</tr>
<tr>
<td>II</td>
<td>43</td>
<td>3.55 ± 0.32</td>
<td>0.024 ± 0.002</td>
<td>0.025 ± 0.015</td>
<td>0.85</td>
<td>0.26</td>
</tr>
<tr>
<td>III</td>
<td>60</td>
<td>3.42 ± 0.22</td>
<td>0.023 ± 0.002</td>
<td>0.038 ± 0.012</td>
<td>0.85</td>
<td>0.24</td>
</tr>
<tr>
<td>IV</td>
<td>64</td>
<td>3.03 ± 0.18</td>
<td>0.019 ± 0.002</td>
<td>0.074 ± 0.012</td>
<td>0.88</td>
<td>0.20</td>
</tr>
<tr>
<td>V</td>
<td>65</td>
<td>3.28 ± 0.13</td>
<td>0.012 ± 0.002</td>
<td>0.103 ± 0.011</td>
<td>0.89</td>
<td>0.19</td>
</tr>
<tr>
<td>VI</td>
<td>56</td>
<td>3.82 ± 0.13</td>
<td>0.015 ± 0.002</td>
<td>0.070 ± 0.013</td>
<td>0.85</td>
<td>0.25</td>
</tr>
<tr>
<td>VI-VII</td>
<td>39</td>
<td>4.20 ± 0.16</td>
<td>0.009 ± 0.003</td>
<td>0.094 ± 0.017</td>
<td>0.80</td>
<td>0.24</td>
</tr>
<tr>
<td>VII</td>
<td>49</td>
<td>4.39 ± 0.18</td>
<td>0.009 ± 0.004</td>
<td>0.091 ± 0.018</td>
<td>0.79</td>
<td>0.28</td>
</tr>
<tr>
<td>VII-VIII</td>
<td>29</td>
<td>5.08 ± 0.10</td>
<td>—</td>
<td>0.110 ± 0.010</td>
<td>0.80</td>
<td>0.23</td>
</tr>
<tr>
<td>VIII</td>
<td>20</td>
<td>5.35 ± 0.15</td>
<td>—</td>
<td>0.116 ± 0.015</td>
<td>0.79</td>
<td>0.27</td>
</tr>
<tr>
<td>VIII-IX</td>
<td>7</td>
<td>5.88 ± 0.70</td>
<td>—</td>
<td>0.086 ± 0.068</td>
<td>0.25</td>
<td>0.39</td>
</tr>
<tr>
<td>IX</td>
<td>9</td>
<td>6.35 ± 0.51</td>
<td>—</td>
<td>0.048 ± 0.063</td>
<td>0.14</td>
<td>0.44</td>
</tr>
<tr>
<td>X</td>
<td>2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
Deriving numerical estimates from descriptive information: the computation of earthquake parameters

Fig. 3. Linear regression between instrumental and macroseismic moment magnitude on the full set of earthquakes used for the fit of empirical regressions.

Fig. 4. Linear regression between instrumental and macroseismic moment magnitude on the 14 earthquakes for which a direct scalar moment measurement is available.
ing those corresponding to the maximum intensity of each earthquake.

The weight is calculated as the product of three factors:

i) The inverse of the standard error of the instrumental magnitude estimate.

ii) A factor linked to the class of detail of the macroseismic study, with a value of 1.0 for expeditious studies, 2.0 for accurate studies and 1.5 for intermediate studies.

iii) The decimal logarithm of the number of observations used to estimate the mean distance.

4) For all the other events of the database, the magnitudes and relative errors are calculated, with the coefficient resulting from the previous point, using all the intensity classes for which the average radius may be computed. At this stage even the classes having at least two points are included in computations.

5) For each earthquake the mean of all the magnitude estimates performed on all available intensity classes, weighted with the inverse of the respective squared error, is calculated. To guarantee the reliability to the resulting magnitudes, all estimates based only on a single class having fewer than four points are discarded.

When the data do not allow magnitude to be computed with the above procedure, $M_s$ is computed from table II, as a function of $I_o$ only. Then $M_s$ is converted to $\log(M_s)$ using the linear relation of table I and to $M_s$ by the Hanks and Kanamori (1979) formula. Table III shows the values of the coefficients of the regressions and the corresponding parameters of goodness of fit obtained for the felt areas of different intensity classes.

As we can see, the data are not sufficient to calculate the regressions for the intensities greater than or equal to degree X. In addition, on the basis of the analysis of an analysis of variance test (Draper and Smith, 1980), for all of the intensities higher than degree VII it was found that the coefficient $b$ is basically indeterminate. This parameter can be therefore eliminated from the regression. It may also be seen that the regressions for the intensities VIII-IX and IX also provide very low $R^2$ values, besides being based on a fairly small number of observations. For this reason they have been excluded from the calculation procedure.

In order to give an overall evaluation of the correspondence between the moment magnitudes estimated macroseismically and instrumentally, a linear regression was performed (fig. 3) between the two values for the set of earthquakes used to calibrate the above relations. The percentage of total variance of the magnitude, expounded by the regression model $R^2$ is 74% while the total standard deviation (std) is 0.28 units. The same comparison (fig. 4) performed on the set of 14 earthquakes for which a direct scalar moment measurement is available, and therefore whose moment magnitude estimates may be considered more reliable, gives $R^2 = 91\%$ and std = 0.17.

A similar analysis has been performed for the magnitudes estimated by the $M-I_o$ table. For the whole set, $R^2 = 43\%$ and std = 0.42. If $I_{\text{max}}$ is used instead of $I_o$, the $R^2$ became 40% and the standard deviation 0.43. These results indicate both the great improvement in the accuracy of the magnitude estimate given by the isoseism radii method and the slight but significant improvement due to the use of this definition of $I_o$ instead of $I_{\text{max}}$.

11. Statistical evaluation of the completeness of the CFTI3

The use of a seismic catalogue for different models of quantitative analysis must be accompanied by an evaluation of its degree of completeness: in other words the degree of correspondence of the apparent seismicity, i.e. as indicated by the catalogue, with actual seismicity, i.e. what actually took place. The problem may be described on the basis of the rates of apparent $A(x,t)$ and actual $R(x,t)$ seismicity (both a function of space and time) as

$$A(x,t) = R(x,t) \times C(x,t)$$

where $C(x,t)$ represents the rate of completeness of the catalogue (also varying in principle in space and time). Since according to the Gutenberg and Richter (1956) law, the number of events is an inverse exponential function of the magnitude, the evaluation of the completeness function must be conducted separately by class of energy released.
The causes of the possible divergence between actual and apparent seismicity are many, ranging from omission of events which actually took place, due to missing information in the historical records or gaps in the surveys performed, to the mistaken attribution of an earthquake to a specific class of energy (Tinti and Mulargia, 1985).

Considering the problem from a deterministic point of view, the completeness of a catalogue may be deduced only through the evaluation of the global efficiency of the system of identification of earthquakes. For an instrumental catalogue, for example, we may presume that, for a given area where a seismic network is operating and for a certain interval of time in which this network has remained unaltered, the ability to identify all the earthquakes with a magnitude greater than a certain threshold (therefore known as «completeness threshold») remains constant. Only variations in the configuration of the network, such as the addition of stations, or the change of its detection characteristics, such as alteration in the sensitivity of the instruments, can alter this threshold. In this case the temporal extremes of the interval of completeness are known, and only the magnitude threshold must therefore be determined. On the contrary, for historical catalogues the interval of completeness is generally not known beforehand. In other words, at the current state of detail of the historical research, we do not know the exact trends in time of the efficiency and determining power of the detection «network» (here represented by the localities in which the macroseismic intensity is determined) which are in turn functions of a series of factors of a social, economic, geographical, cultural and political nature. So, at least given the current state of knowledge, it is not possible to estimate the completeness in a deterministic way. A number of procedures of statistical type have instead been developed which, on the basis of some simplifying assumptions, still permit an estimate to be made of the level of completeness. Because of the way in which the events included in the CFTI3 were chosen (in the first version of this catalogue, all those which in the PFG catalogue had been reported with a maximum intensity greater than the VIII degree) the magnitude completeness lower threshold must range between magnitudes 5.0 and 6.0. In this case therefore the minimum threshold is approximately known while the temporal interval has to be identified.

The two hypothesis underlying almost all the statistical methods used in the past (Knopoff and Gardner, 1969; Stepp, 1971; Tinti and Mulargia, 1985; Mulargia et al., 1987) for determining the completeness of a historical catalogue are the stationariness of the process of seismic generation, in other words the fact that the seismicity of a sufficiently extensive area remains almost constant in time, and the spatial homogeneity of the function of completeness on the territory considered. That is, in the formula relating the apparent to real seismicity rates, the time dependence of $A(x,t)$ is due to completeness and the spatial dependence to the real rate

$$A(x,t) = R(x) C(t).$$

In addition, the majority of these techniques require, as a preliminary condition, the statistical independence of the data. Since the problem of removal of dependent events has not yet been solved satisfactorily, and we are interested here in evaluating the incompleteness of the entire set of data and not just the independent data, we have chosen to follow the simplified approach proposed by Mulargia et al. (1987), which involves a visual inspection of the cumulative plot of the number of events reported as a function of time (CUVI). This method appears to be very efficient and accurate even when applied to small sets of data.

Figure 5 shows the plot relative to the classes of magnitude $M \geq 5.0$, $M \geq 5.5$ and $M \geq 6.0$ from year 1500 up to present. For the lowest threshold of magnitude we may clearly observe, disregarding the non-stationarinesses limited to brief periods of time, at least two fairly distinct increases of the slope of the cumulative curve corresponding to the beginning of the seventeenth century and the end of the eighteenth century and perhaps a third around 1870. These are evidence of a significant increase of the detection capacity of the system of observation with respect to the previous periods. For the
Fig. 5. Cumulative number of seismic events included in the Catalogue of Strong Italian Earthquakes, with equivalent magnitude greater or equal to 5.0, 5.5 and 6.0; the arrows indicate the change points.
intermediate magnitude threshold the first two «change points» may still be identified while the third disappears completely. Lastly, for the highest threshold of magnitude, only the first turning point may be identified. Given the stationarity of seismic generation, the completeness, relative to each magnitude threshold, may be taken from the last observable increment in the apparent seismicity level that is about year 1870 for $M \geq 5.0$, year 1780 for $M \geq 5.5$ and year 1625 for $M \geq 6.0$.

The earliest turning point could be explained by the fact that as from the seventeenth century, historical investigations were able to rely, for various reasons, on a wealth of sources which were decidedly superior both in quantity and quality to those of the previous centuries. This was due to the new bureaucratic and financial organisation of the Italian states during that period and the standards of control of the population (such as the institution of parish registers, for example) established by the Church after the Council of Trent (1542-1564). As regards the other points, we may observe, for the second variation of slope, that since this corresponds roughly to the Calabrian earthquakes of 1783, it could in some way have been linked to the greater interest towards such phenomena on the part of the government of the Kingdom of Naples. In the same way, the turning point at the end of 1800 could be due to the development of the evolved scientific environment which led to the publication of Mario Baratta's great work, «I terremoti d'Italia», at the start of the new century, and the formulation first by Forel and De Rossi and then by Mercalli of modern macroseismic scales.

REFERENCES


