

Assessing the completeness of Italian historical earthquake data

Massimiliano Stucchi ⁽¹⁾, Paola Albini ⁽¹⁾, Carmen Mirto ⁽¹⁾ and Alessandro Rebez ⁽²⁾

⁽¹⁾ *Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Milano, Italy*

⁽²⁾ *Istituto Nazionale di Oceanografia e Geofisica Sperimentale, Italy*

Abstract

The assessment of the completeness of historical earthquake data (such as, for instance, parametric earthquake catalogues) has usually been approached in seismology – and mainly in Probabilistic Seismic Hazard Assessment (PSHA) – by means of statistical procedures. Such procedures look «inside» the data set under investigation and compare it to seismicity models, which often require more or less explicitly that seismicity is stationary. They usually end up determining times (T_i), from which on the data set is considered as complete above a given magnitude (M_i); the part of the data set before T_i is considered as incomplete and, for that reason, not suitable for statistical analysis. As a consequence, significant portions of historical data sets are not used for PSHA. Dealing with historical data sets – which are incomplete by nature, although this does not mean that they are of low value – it seems more appropriate to estimate «how much incomplete» the data sets can be and to use them together with such estimates. In other words, it seems more appropriate to assess the completeness looking «outside» the data sets; that is, investigating the way historical records have been produced, preserved and retrieved. This paper presents the results of investigation carried out in Italy, according to historical methods. First, the completeness of eighteen site seismic histories has been investigated; then, from those results, the completeness of areal portions of the catalogue has been assessed and compared with similar results obtained by statistical methods. Finally, the impact of these results on PSHA is described.

Key words *completeness – historical earthquakes – Italy – seismic hazard*

1. Background

Among investigators working in historical seismology, or making use of this type of data, the assessment of data completeness is a typical problem that seismologists deal with – or, better, create – while historians find hard to understand. Of course, historical records are not complete; however, while historians can live in

peace with such an issue, seismologists – and, among them, those who make massive use of statistics – need to master it in some way.

Although obviously incomplete, Parametric Earthquake Catalogues (PEC), or historical earthquake datasets, such as those widely described in this volume, supply a picture which represents most of our knowledge of the past seismicity. Past seismicity is one of the keys for predicting future seismicity; therefore, the more reliable picture of the past we have, the more successful prediction of the future we expect to perform. This issue invites one to assess how representative the picture we get from the catalogue is with respect to the «true seismicity».

This question can be answered – in principle – in several ways. For instance, we are sure that the total energy release described by a

Mailing address: Dr. Massimiliano Stucchi, Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Milano, Via E. Bassini 15, 20133 Milano, Italy; e-mail: stucchi@mi.ingv.it

PEC over a certain time-interval and above a given threshold magnitude ($M \geq M_t$) represents a fraction of the «true» energy released in the same conditions; we might then be happy to estimate that this fraction is, say, 1/3 rather than 1/2, or 2/3. For certain purposes, on the other hand, it seems more useful to assess the time-interval for which this fraction approximates the value of 1. In this case we speak in terms of assessment of (time-intervals of) data completeness.

This paper mostly deals with the last issue, which is commonly addressed by statistical analysis performed in the perspective of Probabilistic Seismic Hazard Assessment (PSHA). It is to be stressed, however, that the problem of data representativeness is not frequently addressed by other types of analysis, such as the comparison between the strain release assessed by geodetic and seismological data over long time-intervals

(centuries). In this case the gap between the two values is often accounted for as due to non-seismic deformation, with little reference to possible seismological data incompleteness.

While addressing this issue, some aspects of the general problem will also be discussed.

2. The problem, common solution and some objections

The problem we want to solve is the following: *how can we assess the time-interval(s) during which seismological data derived from historical records are complete above a given M_t ?*

Historical-seismological data can be represented on a time scale in many ways. Figure 1a-d shows the seismic histories of some localities, that is, the sequences of earthquake effects

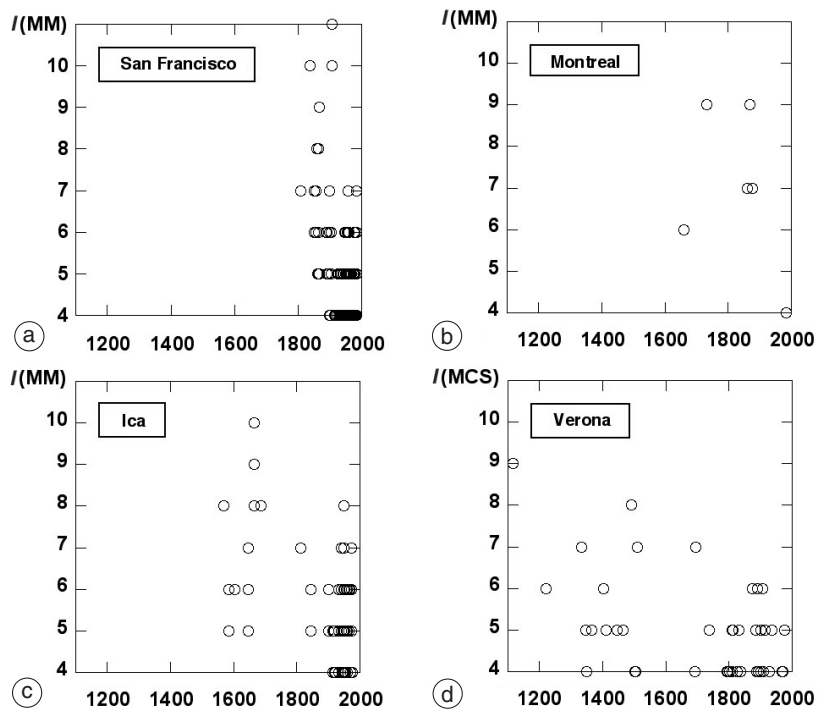


Fig. 1a-d. Seismic histories of (or macroseismic intensities reported at): a) San Francisco, CA, US, b) Montreal, Canada (from US Earthquake Intensity Database, <http://www.ngdc.noaa.gov/seg/hazard/eqint.html>); c) Ica, Peru (from CERESIS, 1985); d) Verona, Italy (from DOM4.1, Monachesi and Stucchi, 1997).

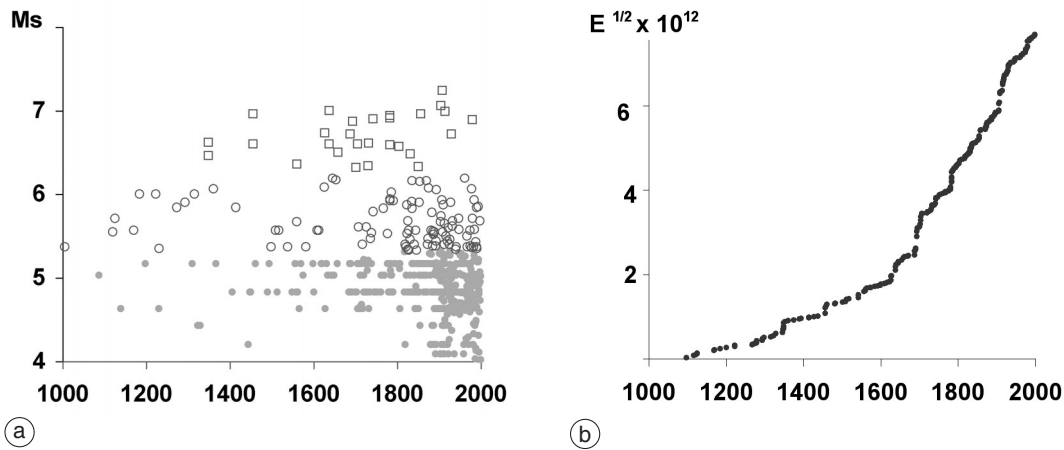


Fig. 2a,b. a) Seismic history of Southern Italy; b) cumulative strain release computed for events with $M_s \geq 5.5$ of the Italian CPTI catalogue.

reported for those localities in terms of macroseismic intensity, as derived from the relevant macroseismic databases. Figure 2a gives the seismic history of a large area (Southern Italy), in terms of magnitude; fig. 2b the cumulative energy release for the whole Italian PEC (Working Group CPTI, 1999).

A look to these graphs invites a number of questions, among which – for instance – are the following: a) did San Francisco suffer intensity 11 MM only once, and Montreal intensity 9 MM two times, since their foundation?; b) is the gap of data of Ica (Peru) in the 18th century due to seismic quiescence or lack of historical records?; c) how would we feel if we neglect the seismicity around Verona before, say, 1600, because we think that the data are not complete?; d) which are the completeness time-intervals in Southern Italy for $M \geq$ given thresholds?; e) does the plot of fig. 2b suggest something?

Actually, it is a common practice to answer such questions looking at the graphs themselves assuming for instance that: i) completeness at Ica starts after the gap, because this is likely to be a gap in the historical records; ii) completeness for large events starts in Southern Italy before 1660 and shortly after

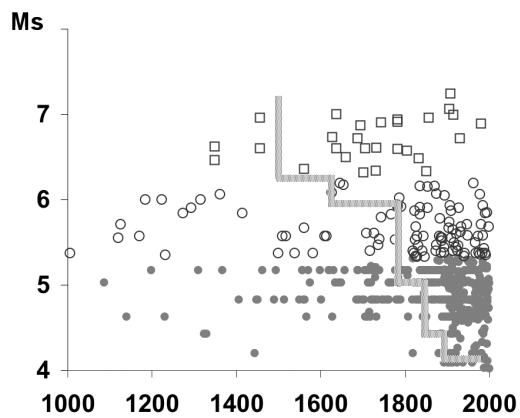


Fig. 3. Seismic history of Southern Italy and completeness time-intervals used by Slejko *et al.* (1998).

1600 for the whole Italy, as from the corresponding jump in the slope of the plot.

However performed, completeness assessment usually ends out with time-intervals which go from some T_i to present: often these

time-intervals are made very short in order to ensure completeness. Time-intervals can be plotted in the form of a broken line (fig. 3): in the common use the line becomes a true divide, which separates the «complete» world (right of it) from the «incomplete» one (left), usually bound for oblivion.

The approaches described above have been made more robust by the use of statistical methods, some of which will be recalled later. They have a common point: they want to infer how complete the data are looking inside the data themselves; and this point shows some inconsistencies, both from the historical and the seismological points of view. Factors which contribute to the wealth of historical-seismological data are mainly of external origin and concern how, where and for what purpose historical records have been produced, preserved and investigated; the evaluation of these factors is a typical historical problem which requires historians' expertise and methods. After all, the problem is similar to the completeness assessment of the data produced by a seismic network; to assess how complete they are above a given Mt one would mainly investigate whether and how long some seismic stations were out of order, and how this affected earthquake detectability and parameter determination. One would not normally infer it from the records of the network only.

On the other hand, when looking inside the data one has to compare – in a more or less open way – the pattern of one's data to some model in which one trusts. The above-mentioned approaches make reference – explicitly or not – to a stationary seismicity model, which may be representative over large areas and time-intervals but definitely fails over small areas and time-intervals. This is even truer in areas where seismicity is low or moderate and where faults are slow in comparison with the time-window spanned by the catalogue.

In order to avoid shortcomings or inconsistent assumptions, historical approaches should be used. Before moving to this, let us examine how the problem is dealt with in some cases.

3. Some current solutions

Following the former considerations, one would expect that most PEC or earthquake databases carry some completeness information so to assist users. In the reality this is hardly the case; this fact leaves users helpless in front of the task, and they usually get out of this situation adopting some statistical approach. This also implies that traces of completeness assessment – if any – are to be found in seismic hazard literature more than in catalogues. A complete review of the existing solutions is out of the scope of this paper: in the following, some examples are recalled, going from contributions dealing with the completeness of the parametric catalogues to those choosing a historical approach to evaluate the completeness of historical records for a region or a country.

It is a general issue that statistical approach to completeness assessment started with the work by Stepp (1972), which had a number of followers. The paper by Albarello *et al.* (2001) clearly describes basic assumptions, pros and cons of the approach and quotes the main literature on the subject, such as Kanamori and Abe (1979), Perez and Scholz (1984), Mulargia and Tinti (1985), Rotondi *et al.* (1994). It also offers the most recent and advanced attempt to assess the completeness with a statistical approach which analyses the seismicity patterns around a given site. It then applies it to one of the Italian PEC (NT4.1, Camassi and Stucchi, 1997) and performs a preliminary comparison of the results with some historical background considerations. For the first time results are associated to uncertainty estimates.

Although not directly suitable for catalogue completeness evaluation it is worth mentioning the analysis method by Kijko and Sellevoll (1989, 1992) to obtain Gutenberg-Richter relation parameters and maximum magnitude from an earthquake catalogue that can be partially incomplete. This method is applicable on a set of data composed by a first part of catalogue derived from historical macroseismic observations, a lack of information (no data) in the central part of catalogue and a final part obtained from recent instrumental observations.

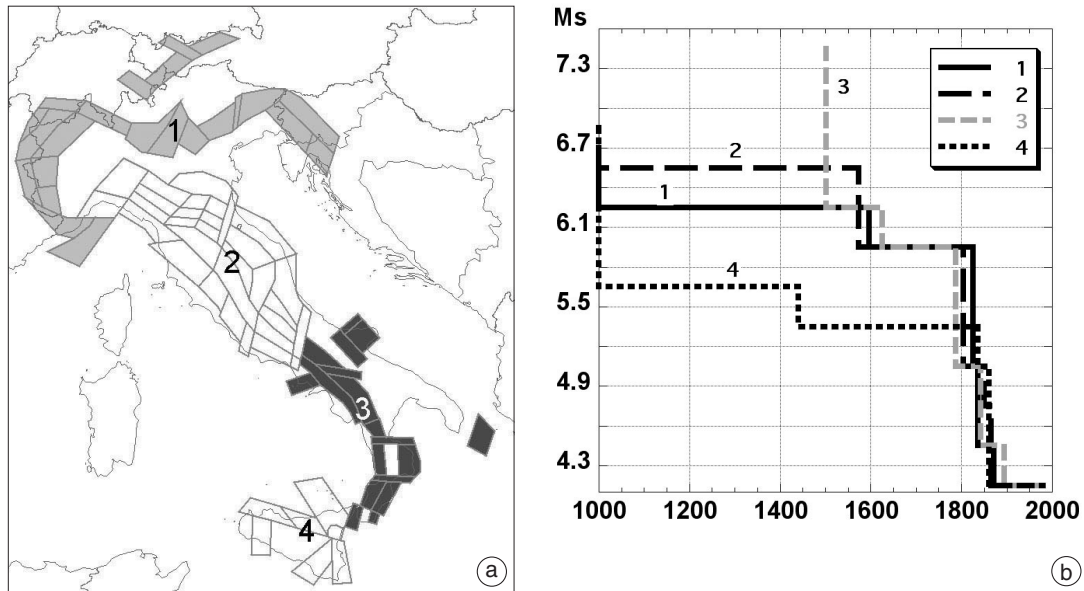


Fig. 4a,b. a) Homogeneous regions and (b) relevant completeness time-intervals (M_s versus time) used by Slejko *et al.* (1998).

In Italy, the completeness of the NT4.1 catalogue was assessed by Slejko *et al.* (1998) with a statistical/pseudo-historical procedure which consisted of dividing the whole Italian territory in 4 regions, assumed as homogeneous from the historical side, and then performing stationarity tests in order to identify the periods of completeness (fig. 4a,b).

For Switzerland, the 1999 version of the MECOS catalogue (MECOS, 1999) available online, gives estimates of completeness for the whole country (epicentral intensity *versus* time), without explicitly mentioning the used approach (fig. 5a,b). The MECOS (1999) estimates are recalled in the introduction to MECOS-02 (Swiss Seismological Service, 2002), the updated version of the macroseismic catalogue. The «completeness» section of MECOS-02 contains historical considerations on the availability and retrieving of documents for the Swiss territory through time and space. However, these arguments are not clearly linked with the estimates attributed to eight Swiss regions, performed for varied time periods and epicentral intensity values.

By comparing the regional estimates of MECOS-02 with the previous of MECOS (1999) it emerges that for some of the eight regions the situation for $I_0=8$ is today more conservative. As an example, according to MECOS-02 in the Luzern region the catalogue is considered complete for $I_0=8$ from 1600 onwards (fig. 5a), and for the regions of Wallis/Valais (fig. 5b) and Tessin roughly from 1500, while for MECOS (1999) the completeness for $I_0=8$ and $I_0=9$ for the whole country started on 1300. It is worth noting that for some regions MECOS-02 states that completeness for all epicentral intensity values cannot be assessed for some time-intervals and epicentral intensities values, mostly on the basis that no primary sources have been found (an example is shown in fig. 5b for Wallis/Valais between 1600-1679).

The completeness of the catalogue for the U.K. territory is shortly discussed in the introduction to Musson (1994); more at length in Musson and Winter (1996). The completeness is assessed from both statistical and historiographical points of view. From the first type of con-

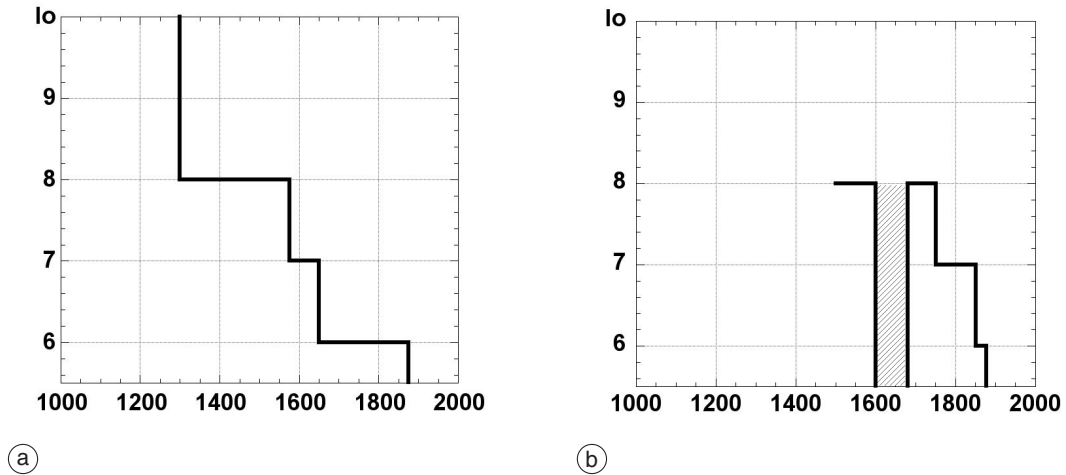


Fig. 5a,b. Completeness time-intervals (I_0 versus time) for MECOS-02: a) Luzern region; b) Wallis/Valais.

siderations it is concluded that the UK catalogue can be considered complete after about 1830 for $M_L \geq 4.0$, and after about 1720 for $M_L \geq 4.5$, although with regional variations. The authors also conclude that the statistical results are «generally consistent with what might be expected from a historiographical approach».

On the historical side, this aspect is being increasingly addressed, with attempts of the authors of giving an idea of how exhaustive the investigation can be considered or what more one can expect to find, and with growing attention to the investigation of unknown earthquakes.

The importance of being aware of the limitations of historical records in assessing «how complete and representative a sample of seismic activity has been recorded» is underlined by Ambraseys and Melville (1982) in their volume on Persian earthquakes. The evaluation of completeness is here supported by describing from which places and which types of historical sources one can expect to find for four unequal time periods, from 7th to 19th century.

In his «Earthquake History of Ethiopia and the Horn of Africa», Gouin (1979) deals with the completeness of the information by dividing the available observations into two periods: 1400-1874 and 1875-1974. He defines the break at

1875 a merely convenient choice, adding that prior to that year reporting of seismicity in the study area was random and sporadic. His evaluation of the 100 years of the second period are based on the comparison between the observed and computed frequency-magnitude curves.

Dealing with the seismicity of Quebec (Gouin, 2001) the author evidences some gaps in the reporting of events between late 17th century and early 19th century. The only reflection on completeness is made by comparing the number of events in the historical period with those reported in the 1990s.

In a very interesting and paradigmatic paper, which did not reach the great international circuit, Agnew (1991) investigates the completeness of the pre-instrumental record of earthquakes in Southern California, analysing the historical sources of the region, comparing the content of varied compilations and proposing some case-histories.

After a contribution with a very stimulating title (Musson, 2000), Musson *et al.* (2001) published some results of a research of seismicity of the Faroe Islands, an isolated area with low seismicity and about which no previous studies had been made. After a thorough investigation of historical sources, some considerations are supplied

on the capacity of records and the significance of their absence in seismological terms. In this case, time-intervals could not be assessed in terms of magnitude, but of peak intensity values at a significant inhabited area.

Some attempts at considering this problem from a historical perspective are present also in some other papers in this same volume, and especially in the contributions by Topozada and Branum (2004) on California, by Tatevosian (2004) on Russia and by Downes (2004) on New Zealand.

Summarising, the above recalled examples show that there is no common current solution. In defining the completeness time-intervals for assessing seismic hazard, the statistical approach has been recently used in combination with considerations coming from the historical side. On the other hand, a more properly defined historical approach has been increasingly used by investigators for selected regions and countries. However, the main goal of the latter contributions is still to explore how historical sources survival and their investigation is affecting the «incompleteness» of the existing catalogues, more than immediately using the results for PSHA purposes.

4. The historical approach in Italy

The problem we are discussing is well known to Italian investigators since many years ago. Traces of this thread are found in many published or conference papers which, although not openly addressing it, went around varied aspects, such as the finding of unknown earthquakes, etc. The historical approach to completeness is briefly dealt with in Guidoboni and Stucchi (1993), a contribution in the frame of the GSHAP project, in a comprehensive perspective in Stucchi (1995) and, concerning the Italian town of Siena, in Castelli and Albarello (2000). The search for unknown earthquakes and the results obtained is reported by Albini and Rodriguez de la Torre (1993) for periodical press in 18th century Europe, by Castelli and Camassi (2000) with the support of some case histories, and from a comprehensive point of view by Valensise and Guidoboni (2000) and Mariotti *et al.* (2000); and surely many more could be mentioned.

All these contributions offer useful results; however, they remain on a qualitative side, which often prevents seismologists from trusting them completely.

4.1. Investigating the completeness of site seismic histories

In the view of overcoming such difficulty and of establishing a semi-quantitative approach, in the last few years a group of historians and seismologists turned their attention to the assessment of the completeness of site seismic histories. The basic idea was: a) to consider a locality, together with its set of historical sources, as a kind of «seismic recorder» and to assess the completeness of the seismic records concerning the locality; b) to infer from it, or from the results on a few localities, some conclusions on the completeness of the data concerning areas.

The investigation was performed in 2001 (Albini *et al.*, 2001), with the financial support of the National Seismic Survey of Italy, on fifteen localities (fig. 6a) selected according to the following criteria:

- i) To have suffered a site intensity $I_s \geq 8$ MCS at least once.
- ii) To be «important» enough, so that a good historical sources coverage could be expected in the considered time-window (1200-1870).
- iii) To be located in areas with preference to low-to-moderate seismicity.

Three further localities, Treviso, Asolo and Belluno, were investigated in 2002, in the frame of the research sponsored by the project «Damage scenario at Vittorio Veneto» of GNDT, the National Group for Defense against Earthquakes (Albini and Stucchi, 2002; Albini *et al.*, 2003).

The main goals of the investigation were:

- a) To find out new earthquake records leading to $I_s \geq 9$ MCS, unknown with respect to the seismic histories obtained from the available datasets (Monachesi and Stucchi, 1997; Boschi *et al.*, 2000).
- b) To understand the reasons why such records could have been lost.
- c) To assess potential information gaps.

It must be said that the available resources allowed a good and rigorous, although not exhaustive, investigation.

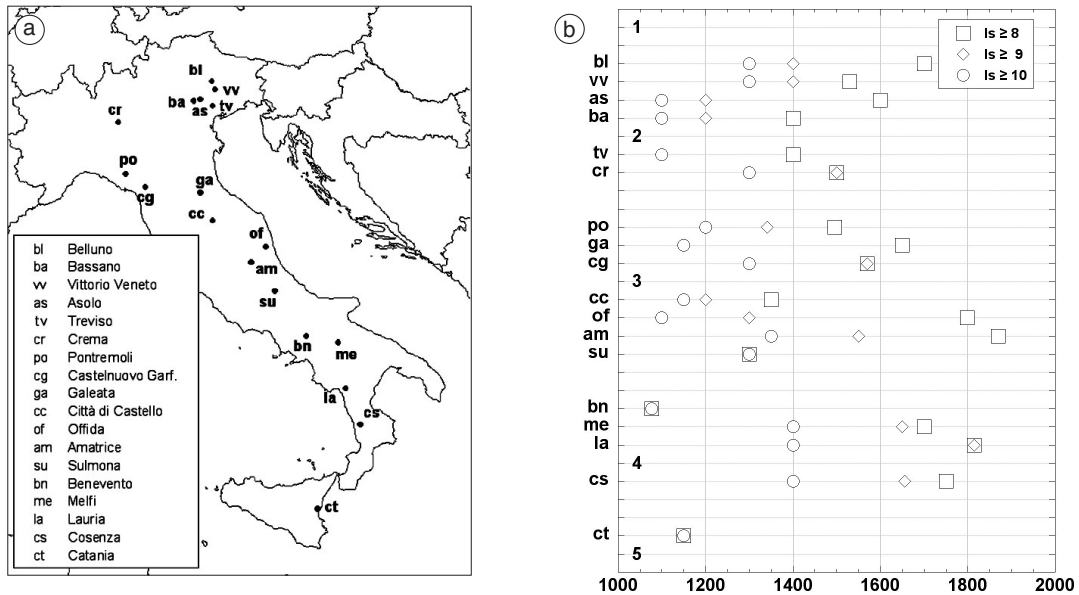


Fig. 6a,b. a) The 18 localities where the completeness of the site seismic histories was investigated. b) Completeness start times at the investigated localities for site intensity $I_s \geq 8$ MCS (squares), $I_s \geq 9$ MCS (diamonds), $I_s \geq 10$ MCS (circles).

Goals a) and b) were achieved by investigating historical sources and compilations which:

- 1) had contributed to the known seismic histories;
- 2) had not contributed to, so far.

The main result was that no unknown earthquake records leading to $I_s \geq 9$ MCS were found for the eighteen localities. In addition, a few cases (locality/events) were investigated, where no records were available for intensities of comparable size predicted by means of attenuation relationships. The results were: i) for a few cases (2) new earthquake records were found, leading to intensity assessments lower than predicted; ii) for another four cases no new records were found, but the predicted intensities were probably overestimated because of a bad fit of the attenuation relationship with the data of the corresponding earthquake; iii) in one case (Sulmona, 1315) no records were found; this fact contributed to fix that the completeness should start later than this date.

Goal c), the main one, to assess completeness and potential information gaps, was achieved by comparing category 1) (historical sources which contributed to the known seismic histories) and category 2) (newly investigated historical sources), and assessing conclusions from both of them. Briefly, the investigation was mainly devoted to exploring the time periods where no earthquakes were reported; the main goal was to assess whether in those periods the historical sources reported other events, either natural or political, or did not report any event at all. In the first case, the historian's expert judgement was able to assess when «no earthquake records» could be interpreted in terms of «no earthquake took place».

The investigation was initially foreseen with reference to $I_s \geq 9$ MCS, because it is believed that historical considerations useful for this purpose apply mainly to the range of heavily damaging to destructive effects. Actually, a heavily damaging earthquake is not an «instant» event, but something which leaves last-

ing traces in the life of the affected territory. It seems reasonable to affirm that in most cases such traces were recorded by the historical sources in the Italian region from about 1200 on, with a few exceptions only. Therefore, our point was to investigate where and how such sources were preserved and/or earthquake records transmitted to us by the compilers. At the end it was possible to supply results also for $I_s \geq 10$ MCS and $I_s \geq 8$ MCS.

The results are given in fig. 6b; they show how varied the completeness time-intervals are, even within short distances.

Summarising, it can be concluded that this approach appears promising and able to give reliable constraints to the completeness assessment. As for many scientific issues it can be performed according to either an expeditious or an intensive approach; in the first case one can obtain many clues in a shorter time, although with a lower reliability; in the second results are better but larger resources are needed.

4.2. From sites to areas, from I_s to M , from high to low M

At conferences the results presented above were observed with sympathy by seismologists, who soon asked for more localities/datapoints, and when the results in terms of catalogue completeness (time-intervals for M and areas) would be available.

More datapoints would cost money and time; therefore, to make our data usable for PSHA purposes we need to extend them. This requires performing three steps: *a* – from sites to areas; *b* – from site intensity to magnitudes; *c* – from high to low M .

Step *a* (extending point data to large areas) requires lot of care, as we are not dealing with measures of physical quantities; on the contrary, we deal with point data which have individual roots and might have only weak relationships one to another, over small areas. Having data from eighteen points only, the areal extension was performed by expert judgement, using the

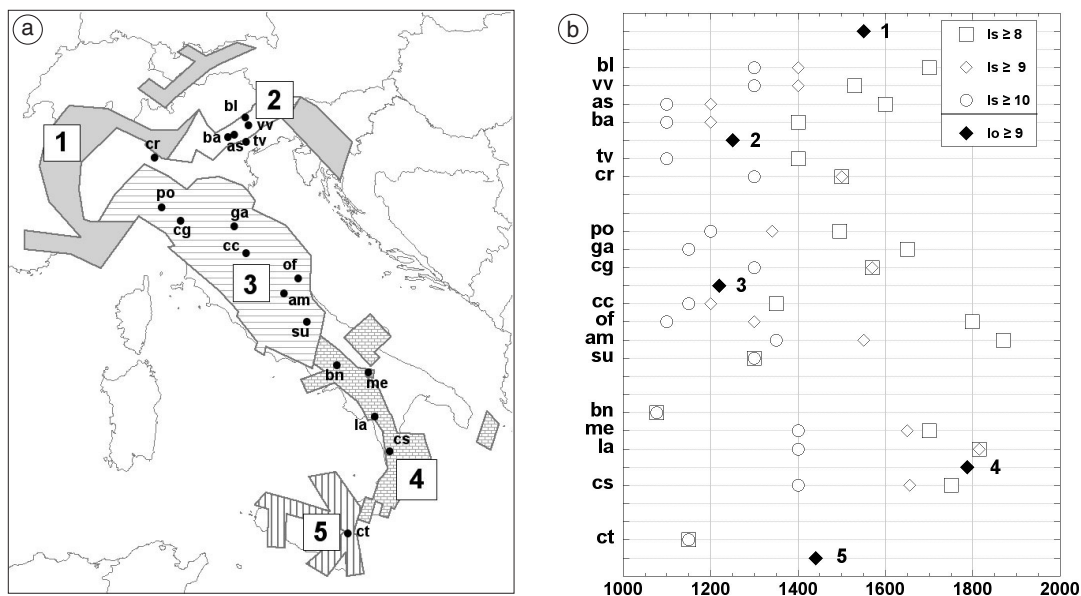


Fig. 7a,b. a) Homogeneous regions and (b) average relevant completeness start times for $I_0 \geq 9$ MCS or $M_s \geq 6.0$ (full diamonds), superimposed to the data of fig. 6b.

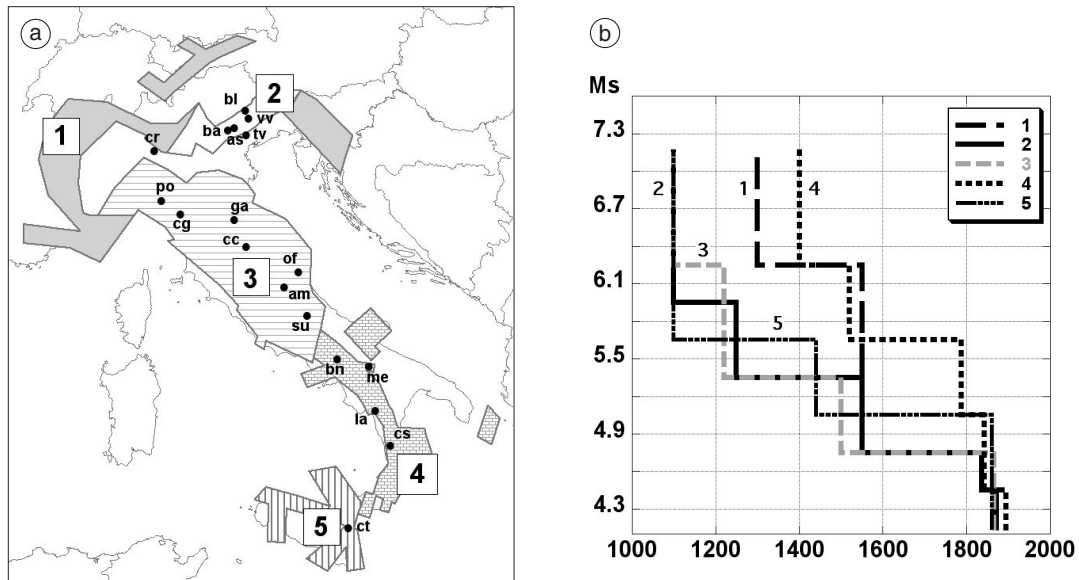


Fig. 8a,b. a) Homogeneous regions; b) broken lines representing the final completeness time-intervals (M_s versus time) in the five regions.

wealth of information and expertise acquired by historians during the last years, although not formalisable according to our schemes, yet. The expert judgement was first used to define five regions (fig. 7a) which, according to historical considerations, may be taken as sufficiently homogeneous from the point of view of historical record production and preservation. As an example, during the Middle Ages and the early Modern Age, the Alpine region in the north and the mountainous Calabria region in the south, with sparse settlements and a relatively small amount of local documents survived, can be compared to the regions of the central and eastern Po Valley, the history of whose Communes and Lordships is well documented by local and highly reliable chronicles. Then, the expert judgement was used to adopt a most reliable areal value for the beginning time of completeness, starting from the value obtained for the localities. It is to be stressed that the choice of the localities was performed for other reasons than this investigation; therefore, their distribution appears rather uneven with respect to the re-

gions (for instance, there was one locality only in Sicily and none in the Alps). Results are presented in fig. 7b and confirm that conservative solutions were adopted.

Step b (from site intensities to M) is performed as follows: the conclusion that $I_s \geq 9$ is complete at locality X since, say, 1400 means that no earthquake effects of that size at locality X are lost after 1400. Such effects can be produced by earthquakes with I_0 located at the same locality, or by earthquakes with larger I_0 located in a convenient area; therefore, using a I_0/M relationship we can conclude at least that no earthquakes with M corresponding to $I_0 = 9$ (roughly $M_s = 6.0$ using the CPTI I_0/M_s table), or larger, happened nearby that locality in the period considered as complete.

As for *step c* (from high M to low M), following the issues made above we must conclude that historical considerations cannot be used for assessing the completeness of slightly damaging effects ($I_0 \leq 7$ MCS and corresponding $M_s \leq 5.0$). Therefore, the values proposed by Slejko *et al.* (1998), assessed according to statistical considerations were adopted with

small adjustments. Figure 8a,b summarises the results in terms of completeness time-intervals for the five regions of fig. 7a.

5. Validation

As explained, these results are based on good although scanty data points and, therefore, may suffer from some uncertainty in relation with the areal extension and the assessment of the most representative value for each area. There is little hope to assess the reliability of these results with conventional methods; however, we can perform some rough estimates.

5.1. Comparison with statistical assessment

A first, although preliminary, comparison can be made with the results published by Albarello *et al.* (2001), both using the data for the

sites closest to our localities, and assessing a mean, weighted value of their results over the five regions adopted in this paper (fig. 9a,b). Results shows that, for $I_0 = 9$ MCS at least, statistical time-intervals are steadily shorter than historical ones, with the exception of two cases: i) Crema, because of a significant gap in the surviving local documentation between 1100 and 1450; ii) Melfi, which represents an anomaly of unclear origin in the statistical results.

5.2. Constraining the estimation of the «true seismicity»

As said before, we do not have a method for estimating the fraction

$$F = CS/TS$$

where CS = seismicity in the catalogue and TS = «true» seismicity, the latter remaining undisclosed.

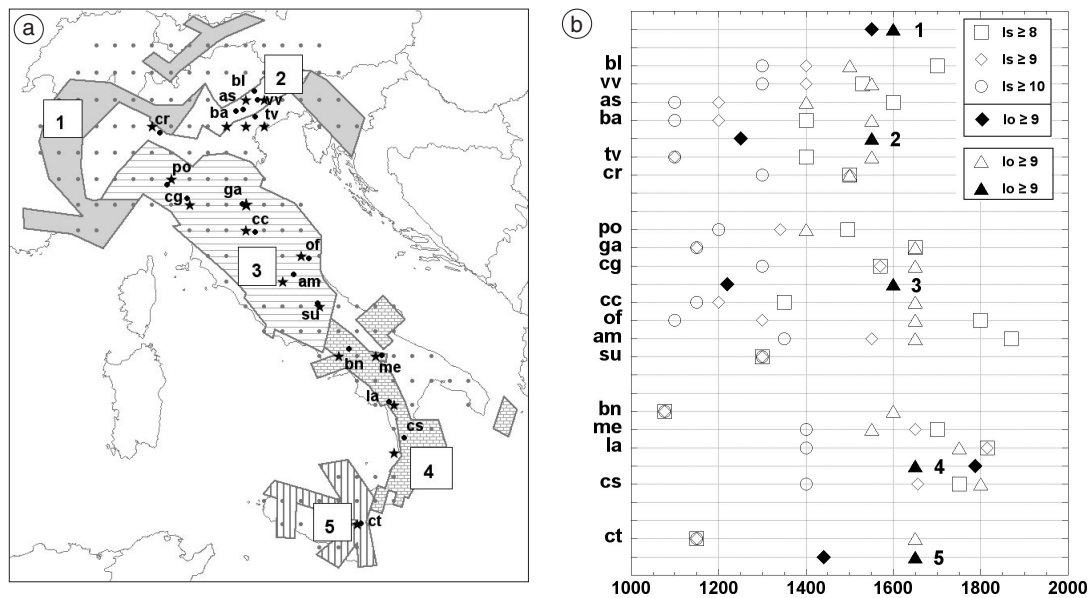


Fig. 9a,b. a) Sites (stars) investigated by Albarello *et al.* (2001) closest to the 18 investigated localities; b) completeness start times ($I_0 \geq 9$ MCS) at these sites, adapted from authors' results (empty triangles), and average start times in the five regions (full triangles), superimposed to the data of fig. 7b.

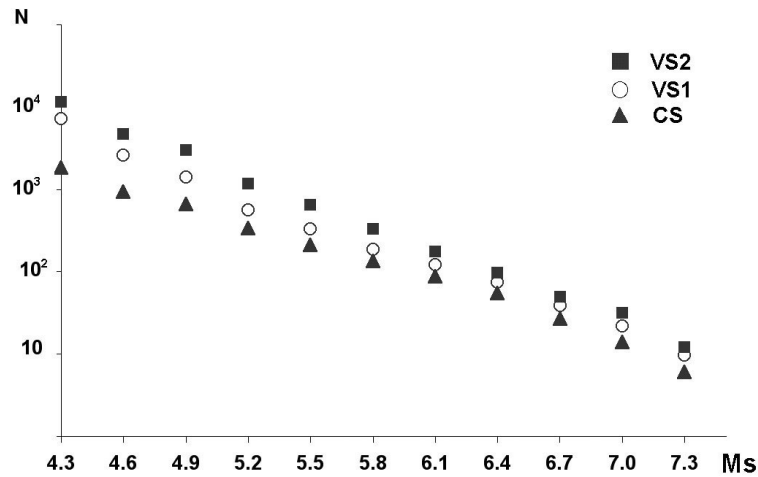


Fig. 10. Comparison of the cumulative NT4.1 catalogue (Camassi and Stucchi, 1997) (CS, full triangles) over 1000 years with the virtual, cumulative catalogues obtained by the historical completeness time intervals (this paper; VS1, open circles) and the statistical one (Slejko *et al.*, 1998; VS2, full squares).

Although TS is not known, as a matter of fact we estimate it in PSHA from the completeness time-intervals by assessing the seismicity rates (number of events of a given M divided by the time-interval). If we project the seismicity rates over – for instance – 1000 years, which is more or less the time-interval spanned by the Italian catalogues, we get the number of events which «should have happened» in 1000 years, under the common assumption that the seismicity rates are representative; in other words, that the seismicity is stationary, that is the current assumption adopted in most PSHA cases.

Figure 10 shows: a) the cumulative earthquake distribution of the NT catalogue (Camassi and Stucchi, 1997) in 1000 years, that is CS; b) the cumulative earthquake distribution in 1000 years which comes projecting over 1000 years the seismicity rates derived by sets of completeness time-intervals of fig. 8 (let us call it «virtual seismicity» VS1); c) the cumulative virtual seismicity VS2 obtained projecting the seismicity rates used by Slejko *et al.* (1998).

The gap between VS1, or VS2, and CS, represents at any M the number of earthquakes that should be missing from the catalogue if all the

basic assumptions are right; fraction F is given, at any M , by the values of CS divided by the values of VS1 or VS2.

Although most investigators could live in peace with almost any large gap or small fraction at any M (after all, catalogues are incomplete by definition!), the size of the gap between distribution CS and VS2 attracted the attention of Stucchi and Rebez (2000) and, later, of Stucchi and Albini (2000). Actually, this gap is about $97-55 = 48$ earthquakes with $M_s \geq 6.4$ ($F = 0.56$) and $329-134 = 195$ earthquakes for $M_s \geq 5.8$ ($F = 0.4$).

In other words, this would mean that the Italian catalogue knows only a half or less than the presumed, total number of earthquakes of the same size which should have happened in 1000 years. This rate, very poor for a country where the historical catalogue is widely considered as a very good one, is hard to believe from the historical point of view, too.

Considering distribution VS1, obtained from the completeness intervals determined in this investigation, *versus* CS, the figures go down to $19 = 74-55$ earthquakes with $M_s \geq 6.4$ and $53 = 187-134$ with $M_s \geq 5.8$; the correspon-

ding values of F are 0.74 and 0.71, which seem more reasonable although a bit high, still.

A more detailed analysis shows that the number of earthquakes which contribute to the gaps described above are not evenly distributed in space but show higher peaks in Southern Italy. This shows, in our opinion, that the completeness time-intervals determined for that region are very severe and/or that seismicity is there even less stationary than in the other areas. In general, gaps are higher in the areas where strong events took place recently. We do not go into details about this here: this analysis has the scope of exploring the possibility of constraining the completeness time-intervals by means of the mentioned gaps, the size of which can be roughly compared with evidence suggested by historical considerations.

5.3. Does it make any difference?

We believe that, in the frame of conventional PSHA, it may. First of all, more rigorous and semi-

quantitative methods must be welcome as a step of PSHA which is frequently overlooked (or casually dealt with). Next, it is clear that some considerations proposed above suffer from incorrect seismicity assumptions. However, such assumptions are behind PSHA conventional methods and they drive completeness assumptions and methods, too.

We believe that completeness plays a significant role in areas with high seismic activity and moderate to long return periods of destructive earthquakes, such as Italy, the Iberian Peninsula and the Balkans. As an example, fig. 11a,b compares the hazard distributions obtained from data which give rise to distribution VS1 and VS2, respectively, keeping all the other input data and using the same code to compute seismic hazard.

It should play a minor role in areas with moderate seismicity (in Europe north of the Alps) or very high one (Aegean Arc). In the first case this is mainly due to the fact that seismicity is low and long-term fluctuations do not affect conventional PSHA. For instance, Musson (2002) showed that – in a typical moderate seismicity case – changing expert judgement

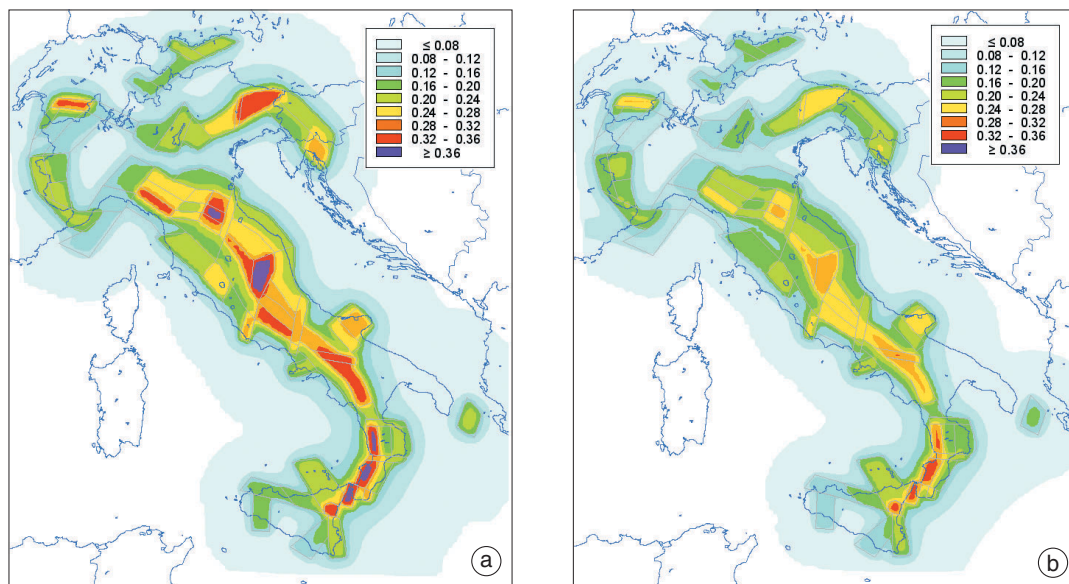


Fig. 11a,b. Comparison of the PSHA obtained using distributions (a) VS2 of fig. 10 and (b) VS1, this paper, leaving unchanged all the other input data and using the same code to compute seismic hazard.

about the completeness changed the final seismic hazard value by 10%.

In the second case, this is due to the fact that shorter time windows can be sufficient to approach the «true» seismicity pattern.

It is clear that we do not believe that distributions VS1 and VS2 really represent the «true» seismicity TS. However, at least we can assess that they overestimate it, may be by far, and that, therefore, using them we are on the safety side.

6. Conclusions

The problem of how good is the picture that our seismological data of historical origin supply with respect to the truth can be of varied importance, depending on the region and/or the final purpose of our research.

We believe that the problem should be mainly addressed with historical methods. This investigation should mainly look outside the earthquake records, in the realm of everyday life history; we are aware, however, that this research can be time-consuming and frustrating, although this is one of the few areas where to find a «credible nothing» may represent a highly valuable result.

We have proposed some examples; being aware that they represent a drop in the sea we are ready to use them, in agreement with the principle that little is better than nothing, for making our PSHA more robust. And we are confident that they will be.

Acknowledgements

This paper summarises the results of a number of ideas, attempts, discussions to which many investigators and friends have contributed, although not always in agreement, over a time-interval of many years. We want to acknowledge here the contribution of D. Albarello, N.N. Ambraseys, J. Bommer, F. Bramerini, R. Camassi, V. Castelli, G. Ferrari, D. Giardini, G. Grünthal, E. Guidoboni, I. Leschiutta, D. Mayer-Rosa, C. Meletti, P. Migliavacca, E. Migliorini, D. Molin, G. Monachesi, V. Montaldo, A.

Moroni, M. Mucciarelli, D. Perkins, L. Peruzza, R.M.W. Musson, R. Rotondi, F. Sabetta, D. Slejko, R. Tatevossian, G. Valensise, J. Vogt, G. Zonno. Without their encouragement this paper would have not come to completeness.

REFERENCES

- AGNEW, D.C. (1991): How complete is the pre-instrumental record for earthquakes in Southern California? *Environmental Perils – San Diego Region*, edited by P.L. ABBOTT and W.J. ELLIOTT (Association of Geologists, San Diego), pp. 14.
- ALBARELLO, D., R. CAMASSI and A. REBEZ (2001): Detection of space and time heterogeneity in the completeness level of a seismic catalogue by a «robust» statistical approach: an application to the Italian area, *Bull. Seismol. Soc. Am.*, **91** (6) 1694-1703.
- ALBINI, P. and F. RODRIGUEZ DE LA TORRE (1993): Known and unknown earthquakes through some European periodicals of the eighteenth century, *Terra Nova*, **5** (5), 481-487.
- ALBINI, P., M. STUCCHI and C. MIRTO (2003): Valutazione di completezza, su base storica, dei dati del catalogo sismico CPTI (1999) nell'area corrispondente alle ZS n. 5 e 6 in ZS4 (Belluno/Cansiglio e Asolo/Vittorio Veneto). Progetto GNDT «Scenari di danno in area veneto-friulana», *Rapporto Tecnico INGV-MI*, 2002, pp. 5.
- ALBINI, P., R. CAMASSI, V. CASTELLI and M. STUCCHI (Editors) (2001): Miglioramento della qualità delle informazioni macrosismiche per un loro utilizzo nella valutazione della pericolosità sismica, *Rapporto Tecnico INGV-MI per il Servizio Sismico Nazionale*, Milano, pp. 195 (Parte A, 43 pp.; Parte B, pp. 152.).
- ALBINI, P. and M. STUCCHI (2002): In search of missing earthquakes in south-eastern Alpine area: historical constraints, in *Special Session S03 – Historic Earthquakes, AGU Fall Meeting*, San Francisco (U.S.A.), 6-10 December 2002.
- AMBRASEYS, N.N. and C.P. MELVILLE (1982): *A History of Persian Earthquakes* (Cambridge University Press, Cambridge), pp. 219.
- BOSCHI, E., E. GUIDOBONI, G. FERRARI, D. MARIOTTI, G. VALENSISE and P. GASPERINI (Editors) (2000): Catalogue of Strong Italian Earthquakes from 461 B.C. to 1980, *Ann. Geofis.*, **43** (4), 609-868.
- CAMASSI, R. and M. STUCCHI (Editors) (1997): *NT4.1: un Catalogo Parametrico di Terremoti di Area Italiana al di Sopra della Soglia di Danno* (versione 4.1.1), Milano, pp. 93 (on line: <http://emidius.mi.ingv.it/NT/>).
- CASTELLI, V. and D. ALBARELLO (2000): Management of uncertainties about local seismic histories and hazard assessment: the case of Siena (Italy), in *27th ESC General Assembly, WSA-1*, Lisboa (Portugal), 10-17 September 2000, p. 116 (abstract).
- CASTELLI, V. and R. CAMASSI (2000): Per un'analisi di completezza del catalogo: spunti e case histories, in *Le ricerche del GNDT nel campo della pericolosità sismica (1996-1999)*, edited by F. GALADINI, C. MELETTI and A. REBEZ (GNDT, Roma), 321-331.
- CERESIS (Centro Regional de Sismología para América del

- Sur) (1985): *Catálogo de Terremotos Para América del Sur*, edited by B. ASKEW and S.T. ALGERMISSEN, Programa para la mitigación de los efectos de los terremotos en la región andina (Proyecto SISRA), vols. 1-9.
- DOWNES, G. (2004): Procedures and tools used in the investigation of New Zealand's historical earthquakes, *Ann. Geophysics*, **47** (2-3), 399-419 (this volume).
- GOUIN, P. (1979): *Earthquake History of Ethiopia and the Horn of Africa* (International Development Research Centre, Ottawa), pp. 259.
- GOUIN, P. (2001): *Tréblemens de terre «historiques» au Québec. «Historical» earthquakes felt in Québec ...* (Guérin, Montreal), pp. 1491 + numerous maps.
- GUIDOBONI, E. and M. STUCCHI (1993): The contribution of historical records of earthquakes to evaluation of seismic hazard, *Ann. Geofis.*, **XXXVI** (3-4), 201-215.
- KANAMORI, H. and K. ABE (1979): Reevaluation of the turn-of-the-century seismicity peak, *J. Geophys. Res.*, **84** (B11), 6131-6139.
- KIJKO, A. and M.A. SELLEVOL (1989): Estimation of the earthquake hazard parameters from incomplete data files. Part I, Utilization of extreme and complete catalogues with different threshold magnitudes, *Bull. Seismol. Soc. Am.*, **79**, 645-654.
- KIJKO, A. and M.A. SELLEVOL (1992): Estimation of the earthquake hazard parameters from incomplete data files. Part II, Incorporation of magnitude heterogeneity, *Bull. Seismol. Soc. Am.*, **82**, 120-134.
- MARIOTTI, D., A. COMASTRI and E. GUIDOBONI (2000): «Unknown» earthquakes: a growing contribution to the Catalogue of Strong Italian Earthquakes, *Ann. Geofis.*, **43** (4), 787-795.
- MECOS (1999), on line: http://histserver.ethz.ch/old_catalogs/mecos/index.html
- MONACHESI, G. and M. STUCCHI (1997): DOM4.1, un database di osservazioni macrosismiche di terremoti di area italiana al di sopra della soglia del danno, *Rapporto Tecnico, GNDT*, Milano-Macerata, pp. 1052 (on line: <http://emidius.mi.ingv.it/DOM/home.html>)
- MULARGIA, F. and S. TINTI (1985): Seismic sample areas defined from incomplete catalogues; an application to the Italian territory, *Phys. Earth Planet. Interiors*, **40** (4), 273-300.
- MUSSON, R.M.W. (1994): A catalogue of British earthquake, *Br. Geol. Surv., Tech. Rep. WL/94/04*, Edinburgh, pp. 99.
- MUSSON, R.M.W. (2000): The historical seismicity of places with no earthquakes, in *27th ESC General Assembly, WSA-1*, Lisboa (Portugal), 10-17 September 2000, p. 116 (abstract).
- MUSSON, R.M.W. (2002): The Significance of Absent Earthquakes: What is the Weight of a Historian's Opinion?, in *Special Session S03 – Historic Earthquakes, AGU Fall Meeting*, San Francisco (U.S.A.), 6-10 December 2002.
- MUSSON, R.M.W. and P.W. WINTER (1996): Seismic Hazard of the U.K. A report produced for the Department of Trade and Industry, *Rep. GNSR(DTI) / P(96)196*.
- MUSSON, R.M.W., T. MIKKELSEN and H. ZISKA (2001): Historical seismicity of the Faroe Islands, *Ann. Geofis.*, **44** (5-6), 1031-1047.
- PEREZ, O.J. and C.H. SCHOLZ (1984): Heterogeneities of the instrumental seismicity catalog (1904-1980) for strong shallow earthquakes, *Bull. Seismol. Soc. Am.*, **74** (2), 669-686.
- ROTONDI, R., F. MERONI and G. ZONNO (1994): A different intensity recording for reducing the uncertainty in its assessment; an application to the completeness analysis of earthquake catalogues, *Nat. Hazards*, **10** (1-2), 47-58.
- SLEJKO, D., L. PERUZZA and A. REBEZ (1998): Seismic hazard maps of Italy, *Ann. Geofis.*, **41** (2), 183-214.
- STEPP, J.C. (1972): Analysis of completeness of the earthquake sample in the Puget Sound area and its effect on statistical estimates of earthquake hazard, in *Proceedings First Microzonation Conference*, Seattle, U.S.A., 897-909.
- STUCCHI, M. and A. REBEZ (2000): How many earthquakes have been lost? Some consequences of theseismicity rates evaluation in PSHA, in *Workshop on Seismicity Modeling in Seismic Hazard Mapping*, Polce (Slovenia), 22-24 May 2000.
- STUCCHI M. and P. ALBINI (2000): Quanti terremoti distruttivi abbiamo perso nell'ultimo millennio? Spunti per la definizione di un approccio storico alla valutazione della completezza, in *Le ricerche del GNDT nel campo della pericolosità sismica (1996-1999)*, edited by F. GALADINI, C. MELETTI and A. REBEZ (GNDT, Roma), 333-344.
- STUCCHI, M. (1995): Short notes from a journey through earthquakes and historical records, in *Earthquakes in the Past. Multidisciplinary Approaches*, edited by E. BOSCHI, R. FUNICIELLO, E. GUIDOBONI and A. ROVELLI, *Ann. Geofis.*, **38** (5-6), 567-571.
- SWISS SEISMOLOGICAL SERVICE (2002): *ECOS – Earthquake Catalog Of Switzerland, ECOS Report to PEGASOS, Version 31.3.2002, Appendix A: ECOS Database* (SED, Zürich), (on line: <http://histserver.ethz.ch/>).
- TATEVOSSIAN, R. (2004): History of earthquake studies in Russia, *Ann. Geophysics*, **47** (2-3), 811-830 (this volume).
- TOPPOZADA, T. and D. BRANUM (2004): California earthquake history, *Ann. Geophysics*, **47** (2-3), 509-522 (this volume).
- US EARTHQUAKE INTENSITY DATABASE, on line <http://www.ngdc.noaa.gov/seg/hazard/eqint.html>
- VALENSISE, G. and E. GUIDOBONI (2000): Towards new research strategies: silent seismogenic areas or silent sources?, *Ann. Geofis.*, **43** (4), 797-812.
- WORKING GROUP CPTI (1999): *Catálogo Parametrico dei Terremoti Italiani* (ING, GNDT, SGA, Bologna), pp. 92 (on line: <http://emidius.mi.ingv.it/CPTI/>).