ANNALI DI GEOFISICA, VOL. XXXVI, N. 1, April 1993

A methodology for the falsification of local-seismic-hazard analysis

Giuseppe Grandori
Politecnico di Milano, Milano, Italia

Abstract
It is recognised that the results of the seismic-hazard analysis at a particular site are not directly falsifiable. In order to obtain indications about the reliability of hazard analysis, it is proposed to shift the attention from the final results to the procedures which lead to those results.

1. Introduction

The statistical elaboration of historical data is an important basis for the definition of the local seismic hazard. Many uncertainties affect this kind of elaboration, due to the limited time period for which reliable historical data (catalog and isoseismal maps) are available. The various procedures that are commonly used for hazard analysis, when applied to the same set of data, lead in general to significant differences in the results.

As an example, the results of the TERESA research project (Mayer-Rosa and Schenk, 1989) can be considered. Six European research groups evaluated the seismic hazard at five locations of Southern Italy using a common data base (table I).

As can be seen, the differences in the results are very large, and obviously nobody knows which one of the results represents the best approximation to «reality».

As far as the definition of «reality» is concerned, I assume here that the available data constitute a sample derived from an earthquake process (the «real» process) whose probabilistic characteristics are constant in time. In other words, we assume that at each location there is a «real» return period for a given intensity, so that we can speak about «errors» that affect the final calculated results, even if these errors are unknown.

In two previous papers (Grandori, 1991 a) and Grandori, *et al.* 1991b), I proposed a method for the evaluation of the order of magnitude of the above-mentioned errors, so that we can judge on a quantitative basis the reliability of any given procedure for the calculation of the seismic hazard.

The method consists, first, in the construction of a model of the seismic zone under consideration. We can say that the model defines a «synthetic» seismic zone, with given probabilistic distributions of earthquakes in space and time, with a given distribution of intensities and with a given attenuation law.

It must be stressed that I do not intend to discuss the errors of the results obtained from the syntethic zone, compared with reality. Now we forget reality and regard the synthetic zone as a «true» seismic zone.

Suppose that we have drawn a limited sample from the synthetic process and that we calculate the seismic hazard at a site on the basis of this limited amount of data. We may be wrong in the choice of the form of the various distributions (wrong modelling) either because the sample suggests distributions different from the «true» ones, or because we deliberately introduce simplified distributions. Moroever, even if we choose the «true» form of the various distributions, the parameters that numerically define

Reference	Method	Benevento	Boiano	Pomigliano	Lucera	Melfi
Barbano *et al.*	SRAMSC	52	73	122	125	75
	Gumbel-1	142	238	714	588	210
García and	Egozcue	51	51	80	77	64
Egozcue	counting	40	40	70	94	47
Lapajne *et al.*	SRAMSC	79	88	312		95
Mayer-Rosa	SRAMSC.mod	23	27	120	100	66
Shenk *et al.*	SRAMSC	34	65	83	120	49
Siro and	Cornell	24	24	333	77	77
Slejko	Gumbel-1	71	83	200	143	83
	Gumbel-3	67	91	333	167	83

Table I. Return period for intensity 8 at five locations of Southern Italy (SRAMSC is the program proposed by McGuire,1976).

them are affected by errors due to the scatter of the samples (wrong estimate of parameters).

The computation of the «true» seismic hazard at a site can be performed in an analytical way if the distributions that define the synthetic process allow such a treatment. However, it is practically impossible to evaluate in an analytic way the errors we may expect when only a limited sample of the synthetic process is used.

To overcome this problem, one can build up a synthetic catalog (respectful of the distributions that define the synthetic process) long enough to provide the «true» seismic hazard at a site by a merely statistical analysis. It is then easy to compare the «truth» with the results obtained (with any given procedure) by using only a limited portion of the catalog, chosen at random.

The error distribution obtained in this way can be considered as a good estimate of the order of magnitude of the expected errors when a given procedure is applied for the elaboration of the real data.

2. Modelling of an Italian seismic zone

Consider the seismic zone of Southern Italy represented in fig.1, and suppose to take into account only the events with epicentral intensity $I_0 \geq 6$.

The Italian catalog for the period 1680-1980 leads to the mean annnual number of events $NI_0(i)$ shown in fig.2. The best fit curve has been obtained by assuming

$$NI_0(i) = \lambda_0 [1 - FI_0(i)] \quad (1)$$

$$1 - FI_0(i) = \exp(\exp 6\beta_0 - \exp i\beta_0) \quad (2)$$

For the numerical values of λ_0 and β_0 we obtained

$$\lambda_0 = 0.51, \beta_0 = 0.186 \quad (3)$$

The seismic zone has been subdivided into 96 cells. Figure 3a) shows the number of historical events that occurred in each cell. Figure 3b) shows a smoothed space distribution of epicenters (the epicenters of the events of one cell are supposed to coincide with the center of the cell).

The analysis of the isoseismal maps of 13 events of the zone led to the proposal of the following attenuation law, illustrated by Grandori *et al.* (1987):

$$I_0 - I = \frac{1}{\ln \psi} \ln\left[1 + \frac{\psi - 1}{\psi_0}\left(\frac{D_i}{D_0} - 1\right)\right] \quad (4)$$

$$\frac{D_0(I_0 = j)}{D_0(I_0 = j - i)} = \phi \quad (5)$$

$$\psi_0 = 1, \psi = 1.5, \phi = 1.3, D_0(I_0 = 10) = 9.5 \text{ km} \quad (6)$$

where D_0 is the equivalent radius of the isoseismic line of maximum intensity I_0 , and D_i is the

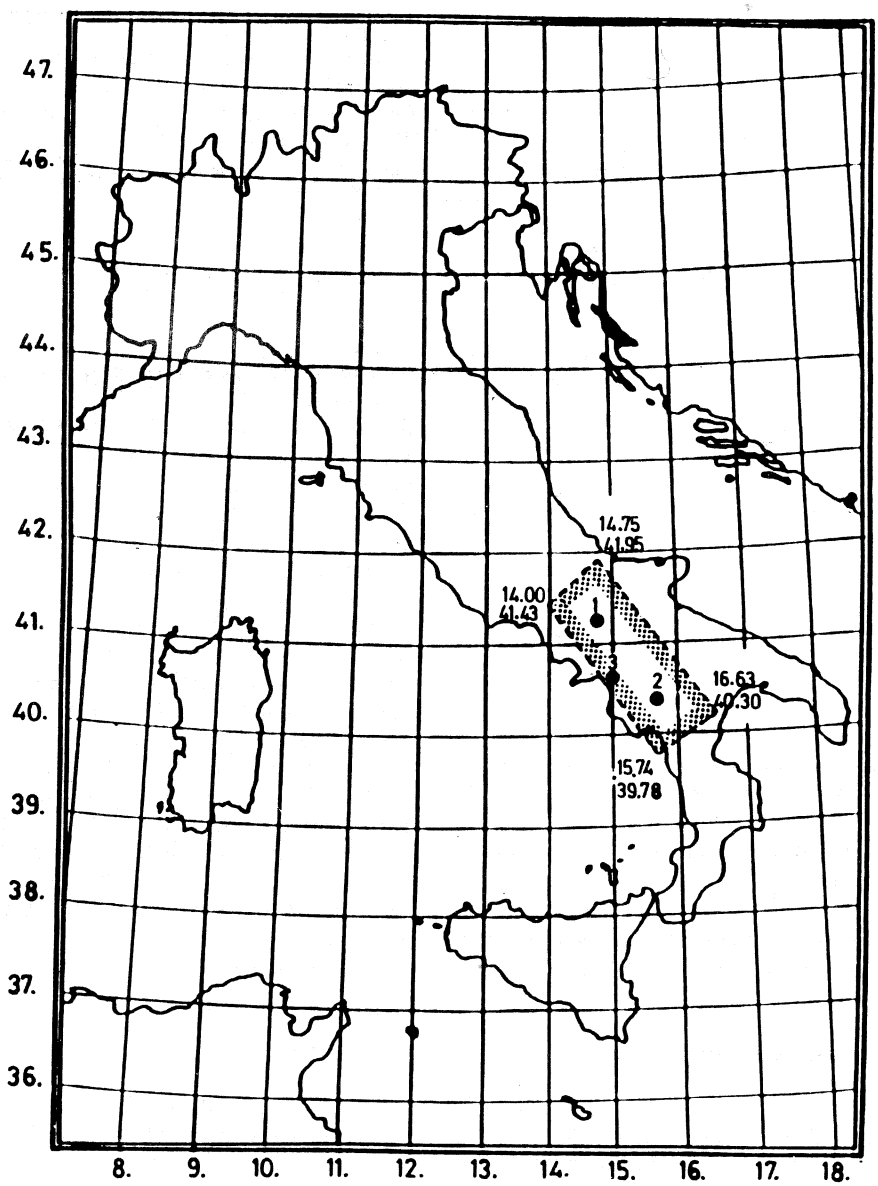

Fig. 1. Historical seismic zone in Southern Italy.

equivalent radius of the isoseismal line of intensity $I = I_0 - i$. Figure 4 shows the attenuation law together with the experimental data.

Formulae (1), (2), (3), (4), (5) and (6) and figure 3b), with the additional hypotheses that the earthquake occurence follows a stationary Poissonian process and that space, time and intensity distributions are independent of each other, completely define what we call the «synthetic» seismic zone. At any point, that can be influenced by

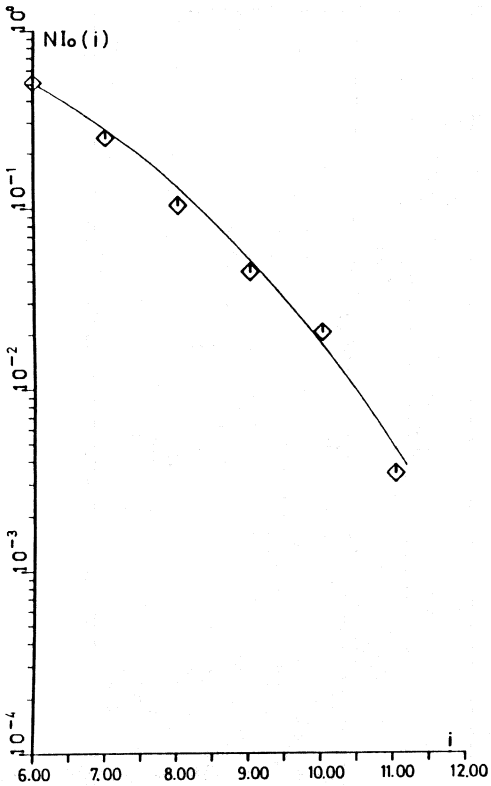

Fig. 2. $NI_0(i)$ *versus* i for historical earthquakes of the zone of fig. 1.

the earthquakes of the synthetic zone, it is possibile to evaluate the seismic hazard.

3. The synthetic catalog

For the synthetic zone we found $\lambda_0 = 0.51$. This means that the return period for $I_0 = 6$ is approximately 2 years. A catalog of 15 000 events (aproximately 30 000 years) has been considered long enough as to furnish reliable statistical results. The statistical analysis is based on an extension of the Bootstrap method (Efron, 1979).

By using a library program, the drawing of a random number, included between 1 and 15 000, has been repeated 15 000 times running. To each drawn number m, position, intensity I_0 and date have been assigned as follows.

The cells of the zone have been numbered in sequence. Let k be the number ($1 \leq k \leq 96$) assigned to a cell, and n_k the number of events (over 15 000) that the space distribution of fig. 3b) assigns to the cell k. The event m has been located in the cell k if

$$\sum_{0r}^{k-1} n_r < m \leq \sum_{0r}^{k} n_r \qquad (7)$$

with the obvious condition that $n_0 = 0$.

All the events of a cell have been located at the center of the cell. The distribution function $FI_0(i)$ obviously defines the range of the numbers of each cell that corresponds to a given intensity. As a consequence, an epicentral intensity I_0 has been assigned to each event m.

Finally, the «observation period» of the whole synthetic catalog has been divided into 15 000 subperiods, and the events m have been distributed at random among these subperiods. If a subperiod received only one event, this event has been placed in the middle of the subperiod. In the case of many events in the same subperiod, they have been placed at regular time intervals.

The «Poissonian» catalog obtained in this way is respectful of the intensity distribution (2) and (3), and of the space distribution of fig. 3b). On the basis of the attenuation law (4), (5) and (6), the «true» seismic hazard at a few points of the zone has been then calculated through a merely statistical analysis.

4. Wrong modelling

For the sake of simplicity, suppose that a wrong form is chosen only for one of the distributions that define the synthetic earthquake process. For example, suppose to introduce the simplifying hypothesis that the space distribution is uniform, instead of the «true» distribution of fig. 3b) (hypothesis of «homogeneous zone»). The analysis of uncertainties due to other modelling errors can be carried out in an analogous way.

Under the hypothesis of homogeneous zone, on the basis of the «true» intensity distribution $FI_0(i)$ and of the «true» attenuation law, it is easy

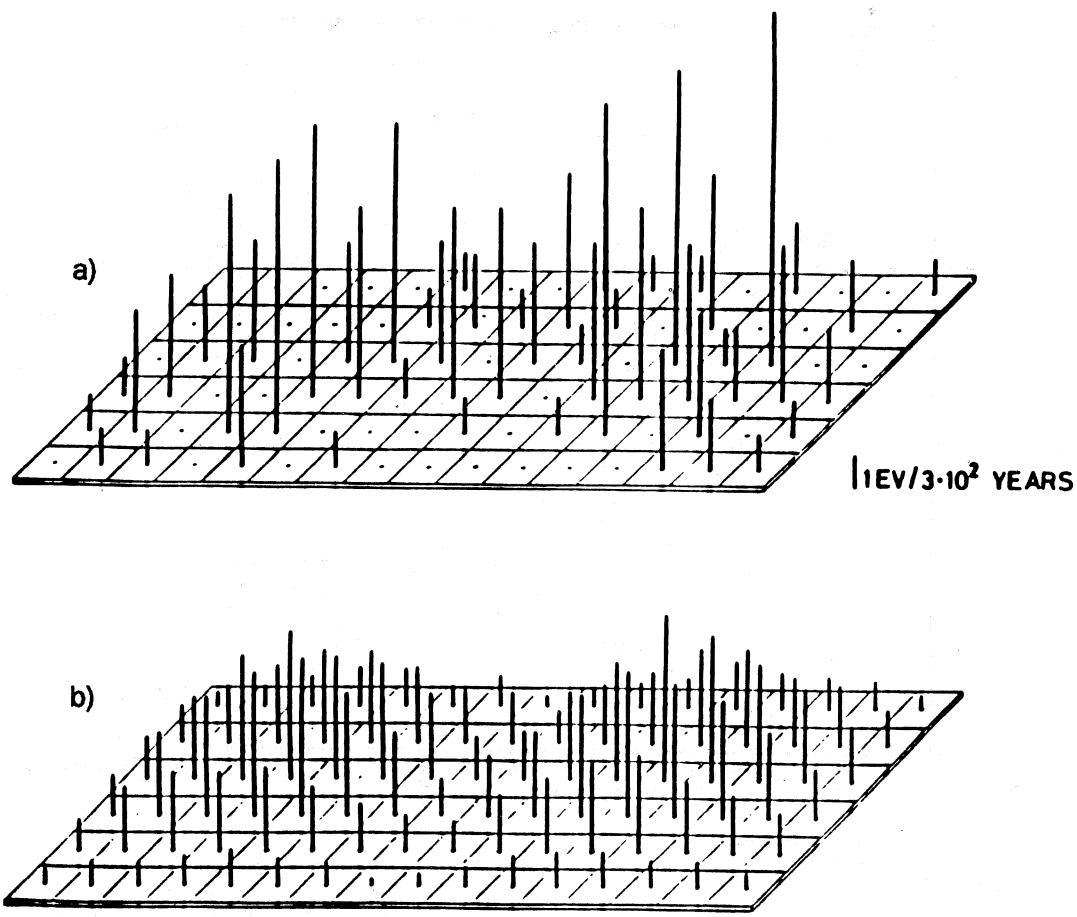

Fig. 3. a) Space distribution of historical earthquakes. b) Smoothed space distribution.

to calculate the seismic hazard at a site, expressed by the function

$$NI(i) = \lambda\,[1 - FI(i)] \qquad (8)$$

where λ and $FI(i)$ refers to local effects.

Note that in this calculation the «true» parameters of all distributions have been used (only changing the form of the space distribution), so that no error in the evaluation of parameters has been introduced. In other words, the errors in the final results are solely due to the simplifying hypothesis of homogeneous zone.

Table II shows, as an example, the values of $NI(8)$ and $NI(10)$ at site 2 of fig.1.

5. Uncertainties due to wrong estimate of parameters

The «observation period» of the whole synthetic catalog (30 000 years) has been subdivided in a sequence of 50 subperiods of 600 years each. On the basis of the data contained in each single sample of 600 years, the seismic hazard at a few sites has been evaluated with the method of homogeneous zone, as follows.

From the data of each sample, the values of λ_0 and β_0 for distribution (2) have been derived. Note that, as we used the «true» distribution (2), no modelling errors have been introduced in this

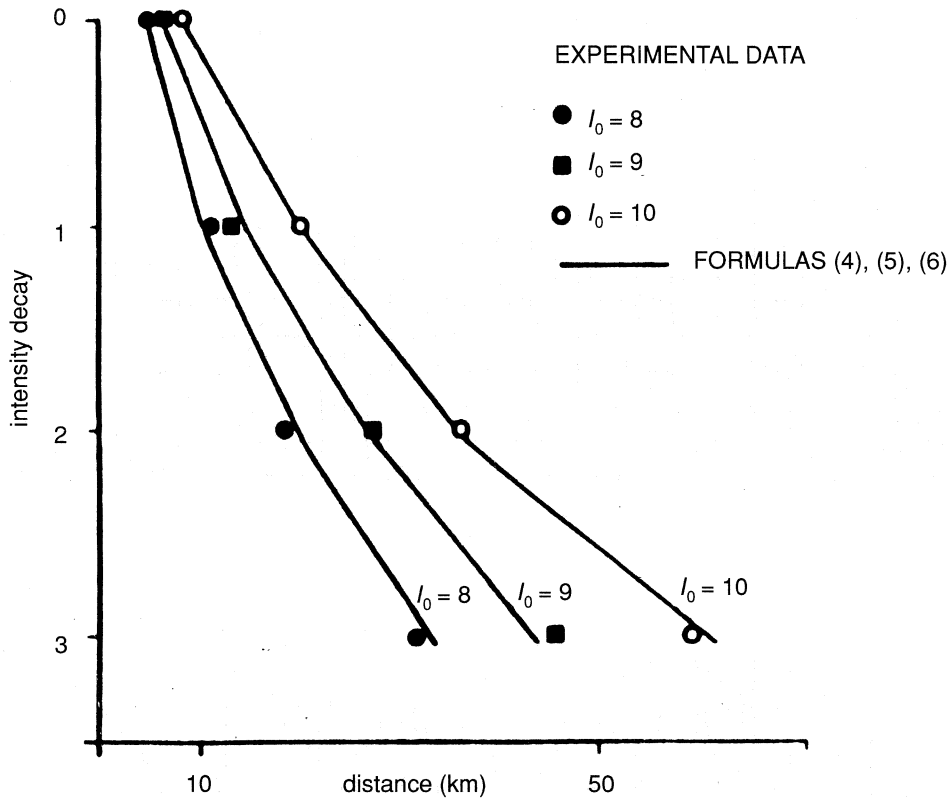

Fig. 4. Intensity decay *versus* epicentral distance.

Table II.

	NI(8)	*NI*(10)
«truth»	0.01316	0.00266
homog.zone	0.00840	0.00147
error %	−36	−45

phase, but only errors in the estimate of λ and β_0.

The regional seismicity so defined has then been uniformly distributed over the zone. As a consequence, the modelling error that we analyzed in the preceding section has been introduced.

Finally, through the attenuation law (4), (5) and (6), the values of *NI*(8) and of *NI*(10) have been calculated.

As mentioned before, this has been done for each sample of the catalog. We obtained in this way 50 different values of *NI*(8) and as many of *NI*(10). Table III shows, as an example, what happens at site 2 in terms of mean values and standard deviation of these quantities. For comparison, the seismic hazard has been calculated by a simple «counting», as follows.

For each event of a catalog's sample, the local intensity at the considered site has been evaluated through the attenuation law. The «experimental» values of *NI*(*i*) have been derived in this way from each catalog's sample. Mean values and standard deviations are shown in table III for the example of site 2.

Finally, the same quantities have been evaluated with a «mixed method». This procedure has been proposed by Grandori *et al.* (1987) on the basis of the following observation. In some simple theoretical examples, the method of homogeneous zone gives a good estimate of the

	«truth»	homogenous zone	counting	mixed method
$NI(8)$	0.0131	0.0085 ± 0020	0.0130 ± 0.0045	0.0126 ± 0.0033
err.%	—	–35 ± 15	–1 ± 34	–4 ± 25
$NI(10)$	0.00267	0.00158 ± 0.00058	0.00270 ± 0.00182	0.00230± 0.00087
err.%	—	–41 ± 22	+1 ± 68	–14 ± 32

Table III. Mean values and standard deviation of $NI(i)$ over 50 samples of 600 years each (site 2).

distribution $1–FI(i)$, but may lead to large errors in λ. Counting, on the contrary, gives in general a good estimate of λ, but may be affected by large errors in the evaluation of the distribution $1–FI(i)$, especially for high intensities. The mixed method consists in taking λ from the counting method and $1–FI(i)$ from the method of homogeneous zone. The results obtained in this way for site 2 are shown in table III.

The method of homogenous zone leads to lowest standard deviations. However, the mean error may be much larger than in the case of the other two methods.

The counting method, as expected, leads to mean errors practically zero. However, standard deviations are rather high, especially for high intensities.

The mixed method seems, at least in the considered example, a reasonable compromise.

6. Conclusion

A methodology has been described for the evaluation of expected errors in the statistical calculation of local seismic hazard in the case of a «synthetic» seismic zone, both for wrong modelling and for wrong estimate of parameters.

The numerical results obtained here refer only to some examples of wrong modelling and of wrong estimate of parameters. However, a systematic application of the methodology would lead to more complete results that could be considered as a good estimate of the order of magnitude of expected errors in the case of a real seismic zone with similar characteristics.

More general conclusions, supposing they exist, could be reached by applying the proposed methodology to many synthetic zones, so that a large range of possible characteristics is covered.

It must be pointed out that the errors that have been discussed refer solely to the statistical elaboration of available data. The reliability of the data is out of question in this paper.

REFERENCES

EFRON, B. (1979): The 1977 Rietz Lecture, *Ann. Stat.*, **7**, 1-26.

GRANDORI, G. (1991a): Paradigms and falsification in earthquake engineering, *Meccanica*, **26**, 17-21.

GRANDORI, G., F. PEROTTI and A. TAGLIANI (1987): On the Attenuation of Macroseismic Intensity with Epicentral Distance, in *Ground Motion and Engineering Seismology*, edited by A.S. CAKMAK (Elsevier).

GRANDORI, G., E. GARAVAGLIA, and V. PETRINI (1991b): Analysis of some of the uncertainties in the statistical elaboration of historical data for local seismic hazard evaluation, *6th International Conference on Applications of Statistics and Probability in Civil Engineering, Mexico City.*

MAYER-ROSA, D. and V. SCHENK (Editors) (1989): *TERESA, Special Issue of Journal of the Int.Soc. for the Prevention and Mitigation of Natural Hazards.*

MCGUIRE, R. (1976): FORTRAN computer program for seismic risk analysis, U.S. Geological Survey, Open-File Report, pp. 76-77.