Probabilistic approach to earthquake prediction

Rodolfo Console, Daniela Pantosti and Giuliana D'Addezio
Istituto Nazionale di Geofisica e Vulcanologia, Dept. of Seismology and Tectonophysics, Roma, Italy

Abstract
The evaluation of any earthquake forecast hypothesis requires the application of rigorous statistical methods. It implies a univocal definition of the model characterising the concerned anomaly or precursor, so as it can be objectively recognised in any circumstance and by any observer. A valid forecast hypothesis is expected to maximise successes and minimise false alarms. The *probability gain* associated to a precursor is also a popular way to estimate the quality of the predictions based on such precursor. Some scientists make use of a statistical approach based on the computation of the likelihood of an observed realisation of seismic events, and on the comparison of the likelihood obtained under different hypotheses. This method can be extended to algorithms that allow the computation of the density distribution of the conditional probability of earthquake occurrence in space, time and magnitude. Whatever method is chosen for building up a new hypothesis, the final assessment of its validity should be carried out by a test on a new and independent set of observations. The implementation of this test could, however, be problematic for seismicity characterised by long-term recurrence intervals. Even using the historical record, that may span time windows extremely variable between a few centuries to a few millennia, we have a low probability to catch more than one or two events on the same fault. Extending the record of earthquakes of the past back in time up to several millennia, paleoseismology represents a great opportunity to study how earthquakes recur through time and thus provide innovative contributions to time-dependent seismic hazard assessment. Sets of paleoseimologically dated earthquakes have been established for some faults in the Mediterranean area: the Irpinia fault in Southern Italy, the Fucino fault in Central Italy, the El Asnam fault in Algeria and the Skinos fault in Central Greece. By using the age of the paleoearthquakes with their associated uncertainty we have computed, through a Montecarlo procedure, the probability that the observed inter-event times come from a uniform random distribution (null hypothesis). This probability is estimated approximately equal to 8.4% for the Irpinia fault, 0.5% for the Fucino fault, 49% for the El Asnam fault and 42% for the Skinos fault. So, the null Poisson hypothesis can be rejected with a confidence level of 99.5% for the Fucino fault, but it can be rejected only with a confidence level between 90% and 95% for the Irpinia fault, while it cannot be rejected for the other two cases. As discussed in the last section of this paper, whatever the scientific value of any prediction hypothesis, it should be considered effective only after evaluation of the balance between the costs and benefits introduced by its practical implementation.

Mailing address: Dr. Rodolfo Console, Istituto Nazionale di Geofisica e Vulcanologia, Dept. of Seismology and Tectonophysics, Via di Vigna Murata 605, 00143, Roma, Italy; e-mail: console@ingv.it

1. Introduction

Throughout the history of seismology, the problem of earthquake prediction has been characterised by conflicting opinions. These difficulties arise mainly from the lack of a common understanding on how a prediction should be defined. According to the suggestion of the IASPEI Sub-Commission for Earthquake Prediction, an earthquake precursor is

defined as «a quantitatively measurable change in an environmental parameter that occurs before mainshocks, and that is thought to be linked to the preparation process for this mainshock» (Wyss, 1991). The set of ideas that forms the basis and lead to the quantitative definition of a precursor is recognisable as an «hypothesis», or a «model». It should be stressed that the hypothesis, or the model, characterising the concerned anomaly or precursor, must be defined in a univocal way, so that it could be objectively recognised and evaluated in any circumstance and by any observer. We refer to Rhoades and Evison (1989), for a description of the steps of the process leading to the acceptance of a given precursor by the application of a rigorous statistical approach. These steps include a phase during which the parameters of the model are assessed by the analysis of past cases (learning phase), and a subsequent phase where the hypothesis is tested against a new and independent data set (testing phase).

Non-random time-space distribution of seismicity can be recognised itself as a precursory phenomenon of forthcoming earthquakes. In seismology, two different aspects of non-random occurrence are commonly taken into consideration: earthquake clustering and long-term quasi-periodicity. The first kind of behaviour is frequently observed as a short or medium term interaction among earthquakes, that tend to occur in groups, such as foreshock-mainshock-aftershock sequences, multiplets, swarms, etc. The second kind of behaviour is supposed to characterise strong earthquakes, that appear separated by fairly regular time intervals, when they occur on the same seismic sources. Both these non-random behaviours, though they appear of opposite nature, can be properly modelled and used for earthquake forecasting. As we shall see later in this paper, while the clustering hypothesis can be straightforwardly tested, because of the abundant data sets collected by the modern seismological networks, testing the recurrence hypothesis requires periods of observations that can exceed even the longest historical records. In this context, data coming from paleoseismology appear a very useful tool.

2. Statistical analysis of predictions

A simple definition of an earthquake forecasting hypothesis could consist of the identification of particular sub-volumes in the total time-space volume (usually named *alarm volumes*) within which the probability of occurrence of strong earthquakes is higher than the overall average. The prediction related to the occurrence of a precursor (or a set of precursors) is successful if an earthquake of magnitude equal to or larger than a given threshold magnitude (*target event*) occurs in the concerned alarm volume. If the observation of a significant number of past cases is available, the number of success (N_S), the number of alarms (N_A) and the total number of earthquakes which occurred in the total time-space volume (N_E) can be respectively counted.

A good prediction hypothesis should have the capability of maximising both the *success rate* (the rate at which precursors are followed by target events in the related alarm volume $= N_S /N_A$) and the *alarm rate* (the rate at which target events are preceded by precursors $= N_S /N_E$). A similar result can be obtained by the trivial way of giving few alarms over very large alarm volumes. This kind of apparent success neglects the fact that non occurrence of events in non-alarm volumes has also to be accounted for as a positive factor (see *e.g.*, Rhoades and Evison, 1979). In this respect, another parameter, such as the *probability gain* (the ratio between the rate at which target events occur in the alarm volume and the average rate at which target events occur over the whole target volume $= N_S /(N_A V_A))/(N_E /V_T))$, has been introduced. In principle, a method of prediction can be considered significant if it achieves a probability gain greater than one.

A prediction consisting of just the statement that an earthquake of large magnitude will definitely occur within a specific alarm volume may be considered too crude. More frequently, the observation of some precursors allows only an estimation that the probability of occurrence of an earthquake has increased by a certain factor with respect to the past in a given area. So, the *conditional probability* of occurrence is a parameter that should hopefully supplement the usual time-location-magnitude parameters of the

prediction. In this respect, in parallel to what is done for weather forecasting, the language in use refers to synoptic earthquake forecasting, rather than to earthquake prediction.

Suppose that the target volume is divided into non-overlapping sub-volumes, or cells, that fill it completely. The probability p_i of occurrence of at least one target event for each sub-volume $i = 1,\ldots, P$ should be estimated. After a suitable period of observation, a number N of target events will have really occurred in some of the sub-volumes denoted by $j = 1,\ldots, N$. The log-likelihood of observing that particular realisation of the earthquake process under the hypothesis defining the probabilities p_i is (see Kagan and Jackson, 1995)

$$\log L = \sum_{i=1}^{P} \left[c_i \log(p_i) + (1 - c_i) \log(1 - p_i) \right] \quad (2.1)$$

where $c_i = 0$ denotes the cells in which no events have been observed, and $c_i = 1$ those in which at least one event has occurred. The first term in the square brackets derives from the probability of occurrence of a target event in the cell i, the second derives from the probability of non-occurrence of such event in the same cell. A model will be considered more valid than another, if it attributes higher probabilities to the cells where the earthquakes occur, and if it implies a smaller total number of expected earthquakes. The situation is represented in fig. 1. One can imagine that the integer i represents the time steps in years. According to eq. (2.1), the likelihood of the observed realisation is positively affected by the occurrence of earthquakes in sub-volumes (years) where a probability larger than 0.5 had been estimated ($i = 6$, 16 and 19), and by the non-occurrence of earthquakes in sub-volumes where a probability smaller than 0.5 had been estimated ($i = 2, 3, 4$, etc.). On the contrary, it is negatively affected by the occurrence of earthquakes in sub-volumes where a probability smaller than 0.5 had been estimated ($i = 12$), and by the non-occurrence of earthquakes in sub-volumes where a probability larger than 0.5 had been estimated ($i = 1$ and 11).
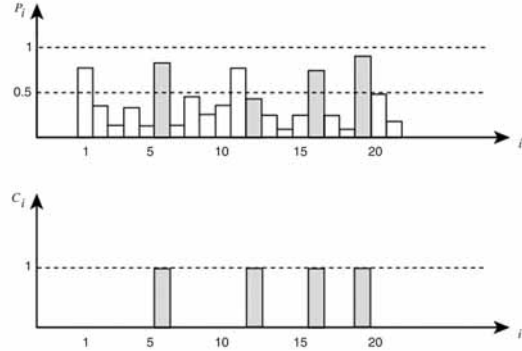
Fig. 1. Probability of occurrence for seismic events in non overlapping sub-volumes $i = 1,\ldots, N$, estimated by a forecasting model (top), and realisation of the process observed in the same sub-volumes (bottom) (after Console, 2001).

If all the probabilities p_i are small with respect to the unit, eq. (2.1) can also be written as

$$\log L \approx \sum_{i=1}^{P} \left[c_i \log p_i - p_i \right] = \sum_{j=1}^{N} \log p_j - N'. \quad (2.2)$$

The statistical procedures based on the division of the target volume into discrete regions can be extended to the formulation of the hypothesis of occurrence in terms of continuous variables, by the definition of the *occurrence rate density* (see Console and Murru, 2001)

$$\lambda(x, y, t, m) = \lim_{\Delta x, \Delta y, \Delta t, \Delta m \to 0} \frac{P_{\Delta x, \Delta y, \Delta t, \Delta m}(x, y, t, m)}{\Delta x \Delta y \Delta t \Delta m} \quad (2.3)$$

where Δx, Δy, Δt and Δm are increments of longitude, latitude, time and magnitude and $P_{\Delta x, \Delta y, \Delta t, \Delta m}(x, y, t, m)$ is the probability of occurrence of an event in the volume $\{x, x + \Delta x; y, y + \Delta y; t, t + \Delta t; m, m + \Delta m\}$. Equation (2.2) becomes

$$\ln L = \sum_{j=1}^{N} \ln \left[\lambda(x_j, y_j, t_j, m_j) V_0 \right] - \\ - \int_X \int_Y \int_T \int_M \lambda(x, y, t, m) dx dy dt dm \quad (2.4)$$

where $V_0 = \int_X \int_Y \int_T \int_M dxdydtdm$ is a dimensional coefficient.

Equation (2.4) preserves the same meaning of eq. (2.2), if we consider that the integral in (2.4) gives the total expected number of events according to the concerned hypothesis.

3. Long term recurrence

Testing earthquake forecasting hypotheses in the context of short-term precursors is a fairly easy task, thanks to the abundant data provided by a modern seismological network covering a seismic region. It has been shown, for instance, how the equations seen in the previous section can be applied to clustering analysis (Console and Murru, 2001). Conversely, the understanding of how large earthquakes recur through time, and thus of the seismic cycle, is a more difficult and critical issue. One of the main problems encountered in attempting recurrence models and particularly in verifying them, is the lack of data (*i.e.* dated earthquakes) due to the short time window within which we can observe how earthquakes recurred in the past. Even using the historical record, that may span time windows extremely variable between a few centuries to a few millennia, we have a low probability to catch more than one or two events on the same fault.

In fact, large earthquakes rupture the same portion of a fault too infrequently when compared to the time interval spanned by instrumental and even historical seismicity (fig. 2). Paleoseismology, that is the geological record of earthquakes of the past, can extend the information on earthquakes back in time up to several millennia representing a great opportunity to study how earthquakes recur through time and thus providing innovative contributions to seismic hazard assessment.

In this paper, we present a test carried out using data extracted by the «Database of earthquake recurrence data from paleoseismology» developed in the frame of the ILP project «Earthquake Recurrence Through Time» (Pantosti, 2000). This database includes information on the studied sites (fault, segmentation, location, kinematics, slip rates, etc.) and on the definition of paleoearthquakes: type of observation for event recognition, type of dating, age, size of movement, uncertainties, etc. This test was developed only on data from the Mediterranean area. In this area we were able to construct seismic histories with two or more inter-events (*i.e.* three or more paleoearthquakes) only for four faults: the Irpinia fault in Southern Italy, the Fucino fault in Central Italy, the El Asnam fault in Algeria, and the Skinos fault in Central Greece (fig. 3). The seismic history for each fault was
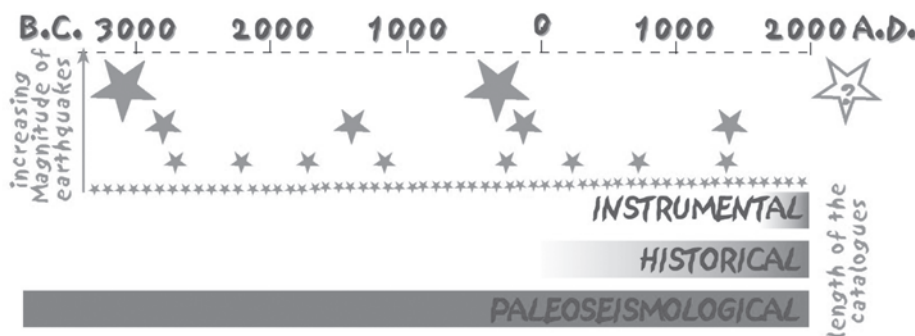
Fig. 2. Length of different types of seismic records compared with earthquake recurrence intervals on individual faults in the Mediterranean area. Note that, because of the relatively long recurrence intervals of large earthquakes typical of this area (1000 or more years), the historical and instrumental records may miss evidence of large earthquakes. In the best cases the historical record may contain evidence for one inter-event interval (*i.e.* two events on the same fault).

Fig. 3. Location of the four faults for which a seismic history was built on the basis of paleoseismological data extracted from the «Database of earthquake recurrence data from paleoseismology» of the ILP.

reconstructed by integrating results from the different paleoseismological sites investigated along it, following the interpretations proposed by the scientists that performed the investigations (Pantosti *et al*., 1993; Meghraoui and Doumaz, 1996; Collier *et al*., 1998; Galadini and Galli, 1999). By using the age of the paleoearthquakes with their associated uncertainty (fig. 4), we tested the null hypothesis that the observed inter-event times come from a uniform random dis-tribution, generally known as the Poisson model. For this test we consider the parameter Cv, defined as the ratio between the standard deviation σ and the mean inter-event time Tr of the observed inter-event times. This parameter is equal to 1 for a uniform random distribution, when the number of observations is very large. $Cv < 1$ denotes a periodical occurrence of events and $Cv > 1$ characterises clustering of events. Cv is easily computed for any observed set of inter-events, if all the events are given their occurrence time without any uncertainty. To take into account, the uncertainties we used a Monte Carlo procedure, computing the average and the stand-ard deviation of Cv, for 1000 inter-event sets. These sets were randomly obtained by choosing the occurrence time of each event within the limits of uncertainty provided by the observations (fig. 5). The obtained values are

$Cv = 0.368 \pm 0.148$ for the Irpinia fault,
$Cv = 0.187 \pm 0.017$ for the Fucino fault,
$Cv = 0.806 \pm 0.102$ for the El Asnam fault,
$Cv = 0.656 \pm 0.214$ for the Skinos fault.

Even if these values, especially that of the Fucino fault, appear substantially smaller than 1, we proceeded with the test of the null hypothesis in a more rigorous and quantitative way. In fact, by means of synthetic tests, it was observed that when estimating Cv on a small number of inter-event times, it appears systematically biased in defect. In this test, the values of Cv obtained experimentally were compared with those obtained by a Monte Carlo procedure, building up to 1000 synthetic random sets of the same number of events in the same total time covered by the real
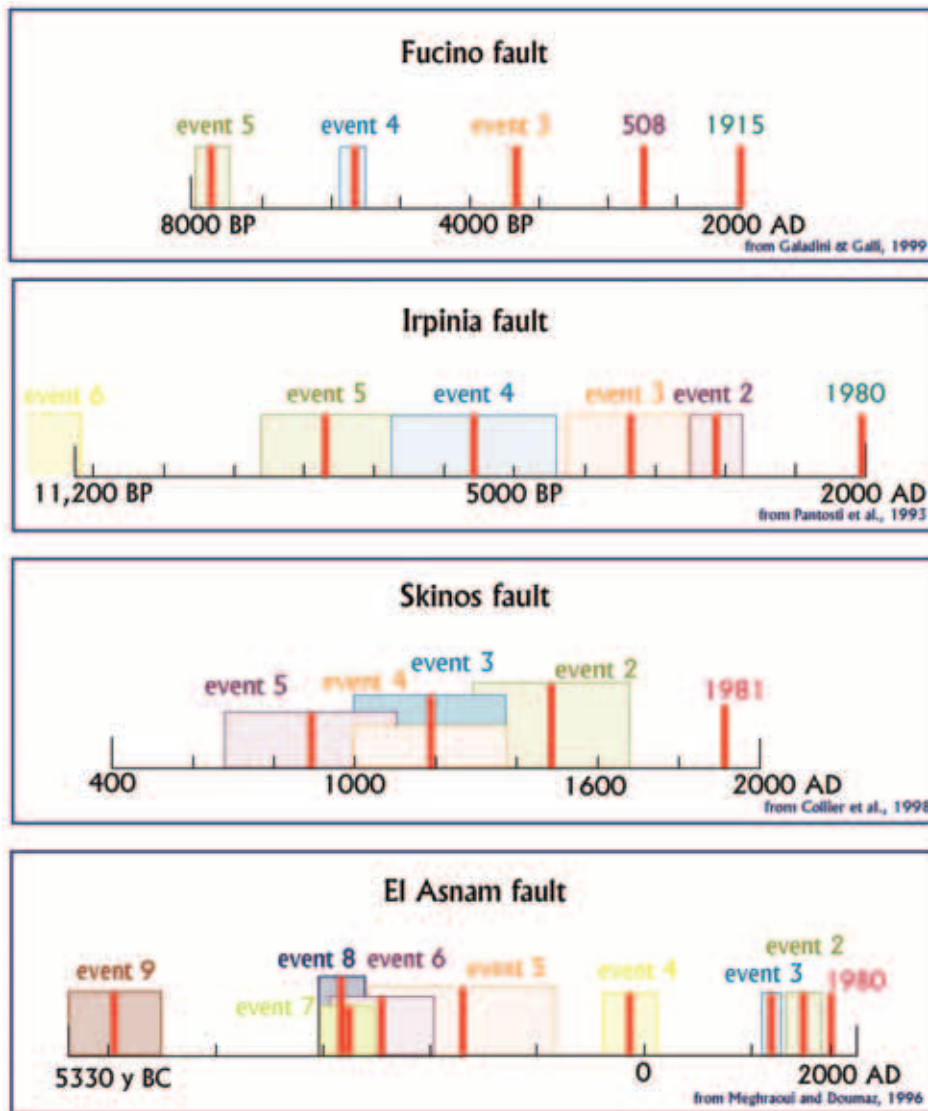
Fig. 4. Ages of paleoearthquakes and associated uncertainty for the four faults used in the test (data from Pantosti *et al.*, 1993; Meghraoui and Doumaz, 1996; Collier *et al*., 1998; Galadini and Galli, 1999). Note that large uncertainties are generally associated with paleoseismological data.

paleoseismological data for each fault. In this way we obtained the probability that a value equal to or smaller than each of the observed Cv comes by chance from casual fluctuations of a uniform random distribution. As a result of the test, this probability is estimated approximately equal to 8.4% for the Irpinia fault, 0.5% for the Fucino fault, 49% for the El Asnam fault and 42% for the Skinos
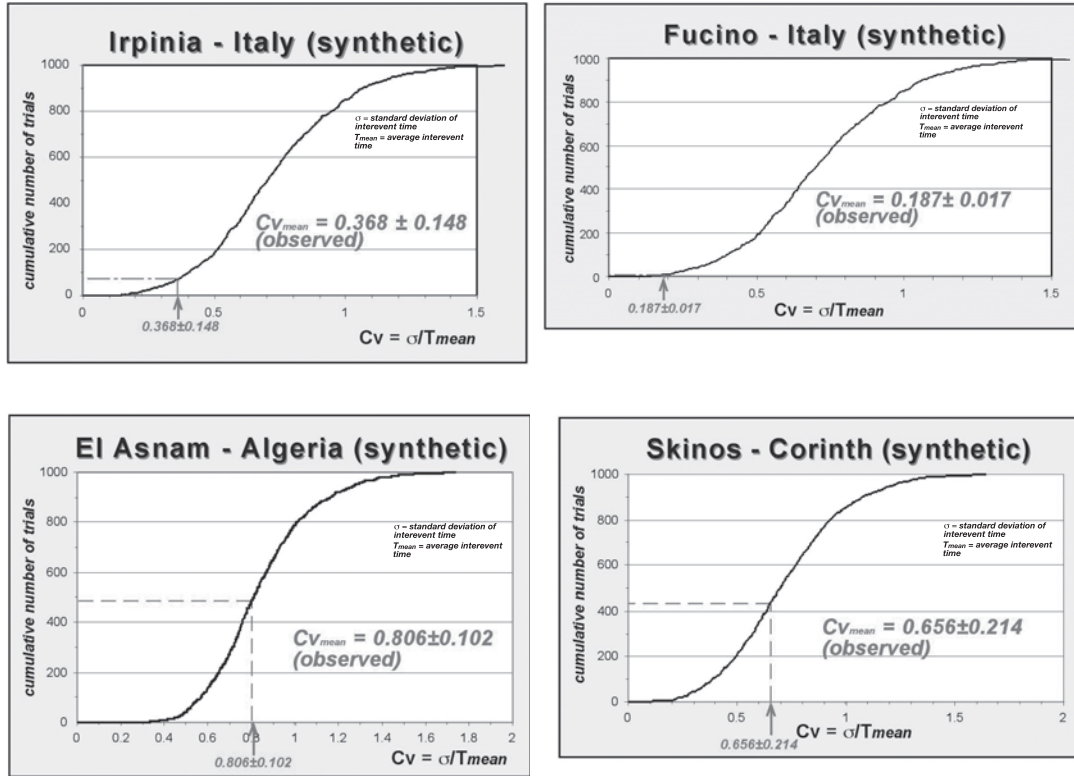
Fig. 5. Comparison between the synthetic random sets of recurrence and the observed *Cv* for each fault. This comparison allows us to derive the probability that the observed intervals are from a random distribution. The average of the coefficient of variation (*Cv*) and of the standard deviation of *Cv* for the fours faults are derived using a Monte Carlo procedure for 1000 inter-event sets randomly obtained by choosing the occurrence time of each event within the limits of uncertainty provided by the observations.

fault. So, the null Poisson hypothesis can be rejected with a confidence level of 99.5% for the Fucino fault, but it can be rejected only with a confidence level between 90% and 95% for the Irpinia fault, while it can not be rejected for the other two cases (fig. 5).

4. Economic impact of an earthquake prediction system

The choice of the best prediction strategy can be considered an issue for decision makers. Different forecasting strategies, even if they are statistically validated, can imply different practical consequences. Even if a decision in matter of disaster mitigation is always connected with social and political issues, some statistical analysis can provide elements to indicate a possible strategy.

A substantial parameter for the evaluation of different strategies is represented by the cost that each of them would have for the community (Vere-Jones, 1995). Let C_{tot} be the amount of money paid by the community to the earthquakes, including both the cost of prevention measures and the cost of damages

$$C_{tot} = \alpha \, l(A) + C_p \, N_p(A) + C_u \, N_u(A) \quad (4.1)$$

where

A represents the level of protection that has to be implemented,

α is the cost of maintaining an alarm per unit time,

$l(A)$ is the total length of alarms,

C_p is the cost for recovering the damages after a predicted event,

C_u is the cost for recovering the damages after an unpredicted event,

$N_p(A)$ is the number of predicted events,

$N_u(A)$ is the number of unpredicted events.

C_{tot} depends, obviously on the threshold of the parameters of the prediction model, to be exceeded before an alarm is declared. In this simple formulation, the cost of implementing and operating the observation system is considered negligible.

It seems reasonable to consider C_p and C_u to depend on the intensity of the earthquakes, and C_u larger than C_p for any class of intensity.

If $N_e = N_p(A) + N_u(A)$ is the total number of earthquakes, eq. (4.1) can be written as

$$C_{tot} = C_u N_e + \alpha \, l(A) - (C_u - C_p) N_p(A). \quad (4.2)$$

In eq. (4.2), the first term of the right member is the total cost that the community should pay for the earthquakes if no prediction system was operated. The second term is the additional cost to be supported for maintaining the alarms, and the third term is the cost saved by the community because of the successful predictions. The best prediction strategy is the one that minimises C_{tot}.

To clarify the meaning of eqs. (4.1) and (4.2) with an example, let us assume that, in arbitrary units, the cost α of maintaining an alarm is 50/year, the cost C_p for recovering the damages of a predicted event is 200 and the cost C_u for recovering the damages caused by an unpredicted event is 300. We imagine to apply the warning system to a seismic area characterised, year by year, by the theoretical probabilities p_i and the real occurrence c_i shown in fig. 1, with $i = 1,\dots,$ 21. If we let the threshold of probability adopted for declaring an alarm decrease from 1 to 0, the number of alarms changes from 0 to 21. From the values of fig. 1 it is possible to determine, in

each case, the number of predicted and non predicted events, and consequently to compute the total cost by means of eq. (4.1) or (4.2). Figure 6 shows the results of these computations by three lines, representing respectively the cost of alarms ($\alpha \, l(A)$), the cost of earthquakes ($C_p \, N_p(A) + C_u \, N_u(A)$) and the total cost ($C_{tot}$), *versus* the number of alarms issued.

As the duration of each alarm is assumed to be constant, the trend of the cost of alarms is linear. The plot of the cost of earthquakes, obviously not increasing, has some flat interval, in correspondence of false alarms. From the plot of the total cost it is evident that the optimal strategy is achieved by the issue of only two alarms, and therefore with a probability threshold significantly larger than 0.5. Moreover, a large
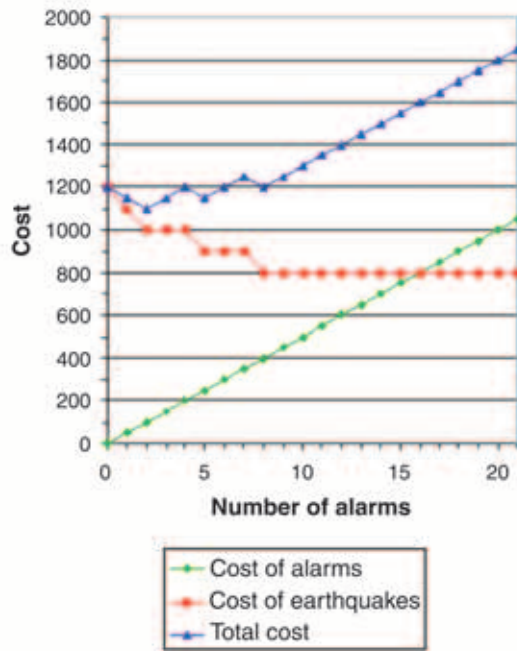
Fig. 6. Cost of alarms, cost of earthquakes, and total cost computed for the realization of the earthquake process shown in fig. 1, *versus* the number of alarms issued, following the criterion of issuing only alarms for the periods of time characterised by a theoretical probability taken in reverse order.

number of alarms gets the total cost even larger than that corresponding to the implementation of no alarm system at all.

5. Conclusions

Statistical analyses are an essential tool for the evaluation of models of earthquake forecasting. Their use is conditioned by a rigorous and objective definition of the precursors introduced in the models that have to be compared. In this paper we have dealt mainly with non-random time-space distribution of seismicity, considered as a phenomenon that can allow forecasting the evolution of subsequent seismic activity.

We have shown how the evaluation of an earthquake precursor can be carried out comparing the occurrence rate of earthquakes exceeding a given threshold magnitude in the alarm space-time volumes with the average occurrence rate in a seismic area. This estimate can be refined by the addition of the concept of likelihood, computed by the conditional probability of occurrence of earthquakes under a certain hypothesis. A further refinement is also the extension of the concept of likelihood to continuous distributions of the probability density.

The evaluation of a forecasting model must be made on a data set independent of the one used in the best fit for the parameters of the model. While this procedure has been demonstrated as feasible for sets of instrumental data collected in a seismic area, it may be hardly implemented for strong earthquakes. This is because the recurrence of strong earthquakes on individual faults implies a very long period of observation to collect statistically significant data sets. Paleoseismological data sets, spanning a time interval much longer than that possible for historical data, can provide invaluable information for this purpose. We carried out an analysis on the recurrence of strong earthquakes by selecting from the ILP «Database of earthquake recurrence data from paleoseismology» four cases in the Mediterranean area. In this way, it has been possible to show that the regularity by which large earthquakes have occurred on the Fucino fault is a statistically significant feature

(the null Poisson hypothesis can be rejected with a confidence level of 99.5%). However, the confidence level is only between 90% and 95% for the Irpinia fault, while it is much lower in the other two cases studied: the El Asnam and Skinos faults.

Even if a forecasting model is proved to be reliable, its practical application for social and economic purposes can be questionable. By means of elementary formulas and a simple example, it has been demonstrated that the incorrect use of a prediction strategy could lead to costs that exceed the benefits produced by the practical implementation of a prediction system.

REFERENCES

COLLIER, R., D. PANTOSTI, G. D'ADDEZIO, P.M. DE MARTINI, E. MASANA and D. SAKELLARIOU (1998): Paleoseismicity of the 1981 Corinth earthquake fault: seismic contribution to the extentional strain in Central Greece and implications for seismic hazard, *J. Geophys. Res.*, **103**, 30,001-30,019.

CONSOLE, R. (2001): Testing earthquake forecast hypotheses, *Tectonophysics*, **338**, 261-268.

CONSOLE, R. and M. MURRU (2001): A simple and testable model for earthquake clustering, *J. Geophys. Res.*, **106**, B5, 8699-8711.

GALADINI, F. and P. GALLI (1999): The Holocene paleo-earthquakes on the 1915 Avezzano earthquake faults (Central Italy): implications for active tectonics in Central Apennines, *Tectonophysics*, **308**, 142-170.

KAGAN, Y.Y. and D.D. JACKSON (1995): New seismic gap hypothesis: five years later, *J. Geophys. Res*, **100**, 3943-3959.

MEGHRAOUI, M. and F. DOUMAZ (1996): Earthquake induced flooding and paleoseismicity of the El Asnam, Algeria, fault-related fold, *J. Geophys. Res.*, **101**, 17,617-17,644.

PANTOSTI, D. (2000): Earthquake recurrence through time, in *Proceeding of the Hokudan International Symposium and School on Active Faulting,* Awaji Island, Hyogo, Japan, 17th-26th January 2000, 363-365.

PANTOSTI, D., D.P. SCHWARTZ and G.VALENSISE (1993): Paleoseismicity along the 1980 surface rupture of the Irpinia fault: implications for earthquake recurrence in Southern Apennines, Italy, *J. Geophys. Res*, **98**, 6561-6577.

RHOADES, D.A. and F.F. EVISON (1979): Long-range earthquake forecasting based on single predictor, *Geophys. J. R. Astron. Soc.*, **59**, 43-56.

RHOADES, D.A. and F.F. EVISON, (1989): On the reliability of precursors, *Phys. Earth. Planet. Int.*, **58**, 137-140.

VERE-JONES, D. (1995): Forecasting earthquakes and earthquake risk, *Int. J. Forecasting*, **11**, 503-538.

WYSS, M. (Editor) (1991): Evaluation of proposed earthquake precursors, *Am. Geophys. Un.*, pp. 94.