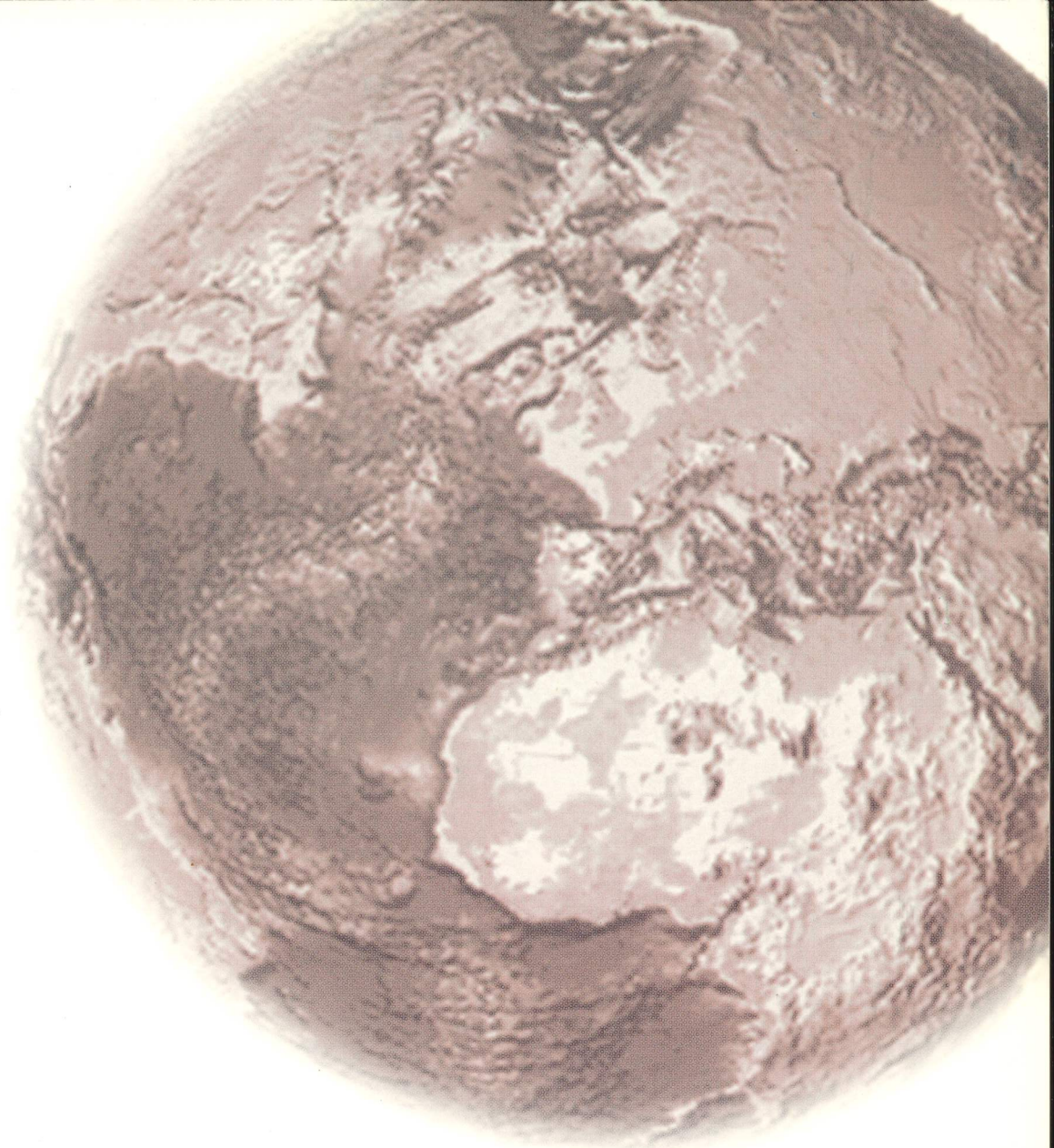




2002

Computer algorithms for testing earthquake forecasting hypotheses

R. Console and A.M. Lombardi

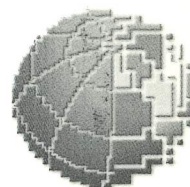
n. 13

Istituto Nazionale di Geofisica e Vulcanologia

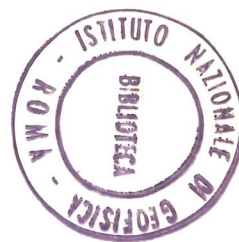
Via di Vigna Murata 605 - 00143 Roma

tel. 06518601 • fax 065041181

www.ingv.it







Computer algorithms for testing earthquake forecasting hypotheses

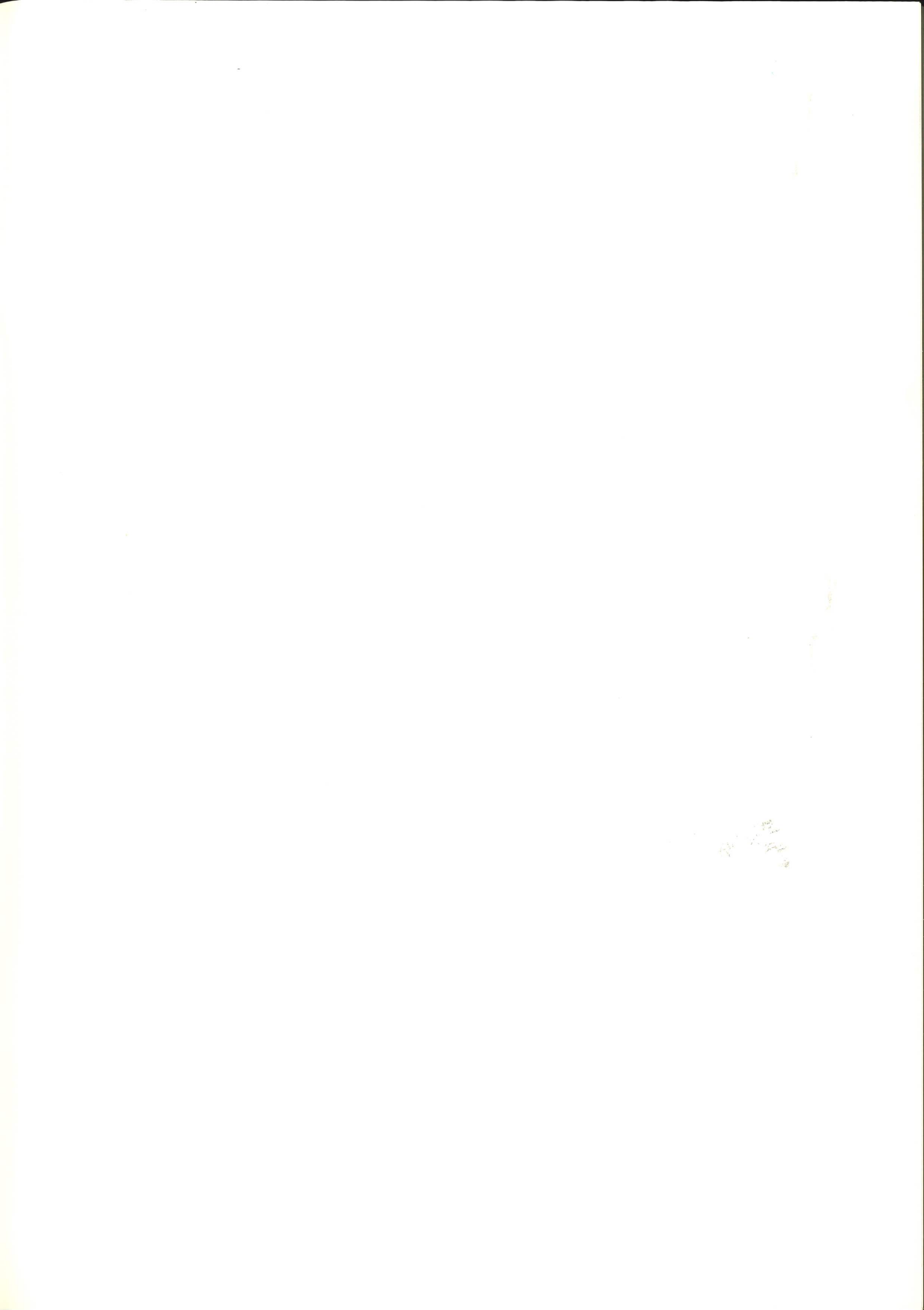
R. Console and A.M. Lombardi

Istituto Nazionale di Geofisica e Vulcanologia

Rome, December 2001

Contents

Abstract.....	5
Riassunto.....	5
Introduction.....	6
1. Statistical background.....	6
1.1 The Bayes theorem	6
1.2 Likelihood function.....	9
1.3 Alternative formulations	10
1.4 Likelihood ratio	11
2. Application to seismic catalogs	13
2.1 Extension to the continuum	13
2.2 The case of time independence	14
2.3 Time-independent model	16
2.4 Induced seismicity	17
2.5 The epidemic model.....	19
3. Computer programs	21
3.1 JMA2REC, JUN2REC, PDE2REC	21
3.2 REC2JMA, REC2JUN, REC2PDE	21
3.3 REMAFT	21
3.4 SMOOTH.....	21
3.5 GRID	22
3.6 LIKELY	22
3.7 TIMEHIST	23
3.8 DENSMAP	25
3.9 SEISYNT	28
3.10 CLUSYNT.....	28
References.....	30



Abstract

Earthquakes are regarded as the realization of a point process modeled by a generalized Poisson distribution. No hypothesis is inferred on the physical model of such a process. The occurrence of a seismic event within an infinitesimal volume of the space of its characterizing parameters is supposed to be completely random, while the behaviour in a reasonably large volume of the same space may exhibit some average non-random features.

The description of seismicity as a point process in space, time and magnitude is suitable for the application of statistical tools for comparing the performance of different models. Here the Bayes theorem is introduced aiming to its application in computing the probability that a given model is true in light of the results of experimental observations. The use of the Bayes theorem requires the computation of the likelihood ratio for the two models to be compared. It follows that both the hypotheses are fully and quantitatively described.

In the application of the statistical algorithms to seismicity, the Gutenberg-Richter law is assumed to describe the magnitude distribution of all the earthquakes in a sample, with a constant b value. The occurrence rate density of earthquakes in space and time is modeled as the sum of two terms, one representing the independent, or spontaneous, activity and the other representing the activity induced by previous earthquakes. The first term depends only on space and is modeled by a continuous function of the geometrical coordinates, obtained by smoothing the discrete distribution of the past instrumental seismicity. The second term depends on both space and time, and it is factorized in two terms that depend on the space distance (according to an isotropic normal distribution) and on the time difference (according to the generalized Omori law), respectively, from the past earthquakes. Knowing the expected rate density, the likelihood of any realization of the process (actually represented by an earthquake catalog) can be computed straightforwardly.

This paper is concluded with a short outline of some computer tools written for the application of the above mentioned algorithms.

Riassunto

Il processo sismogenetico può essere considerato come la realizzazione di un processo di punto marcato, ossia come "un insieme aleatorio di punti", ognuno dei quali è individuato dall'istante di accadimento di ogni singolo evento, dalla magnitudo e dalle coordinate epicentrali. Avvalendosi dell'ausilio di strumenti statistici, è possibile confrontare diversi modelli stocastici e determinare quello che meglio sintetizza le caratteristiche del fenomeno studiato. In particolare, nel seguito, è illustrato come, mediante l'applicazione del Teorema del Bayes, sia possibile calcolare la probabilità che un dato modello sia migliore di un altro, alla luce delle informazioni ottenute dalle osservazioni.

Il modello utilizzato per descrivere la sismicità assume valida la legge di Gutenberg-Richter per tutti i terremoti con un b -value costante. L'occorrenza spazio-temporale degli eventi è, invece, attribuita alla somma di due processi uno rappresentante l'attività spontanea, l'altro responsabile dei terremoti indotti da eventi precedenti. Il tasso degli eventi spontanei è assunto costante nel tempo (*Processo di Poisson*); la loro distribuzione spaziale è, invece, ottenuta da uno smoothing della distribuzione degli eventi osservati nel passato nella zona studiata. Il tempo di occorrenza degli eventi indotti è invece individuato a partire dall'istante di accadimento del terremoto inducente, in accordo con la legge di Omori; è assunto, inoltre che la distanza tra due eventi, uno inducente e l'altro indotto, abbia una distribuzione normale.

Alla fine del seguente lavoro, sono riportati anche brevi informazioni su alcuni programmi usati per applicare la metodologia ed il modello illustrati.

Introduction

This report contains information on the concepts and on the practical use of the algorithms developed by one of the authors during his stage at the ERI of the University of Tokyo through September 22 - December 21, 1998.

Section 1 is devoted to a statistical background presented in elementary way. It introduces the concept of the Bayes theorem, followed by the concepts of likelihood and likelihood ratio, that are the tools for estimating and comparing the reliability of different hypotheses.

Section 2 deals with the application of the statistical tools to the seismicity of a region described by a seismic catalog, starting with the extension to the continuum of the concepts previously introduced for discrete regions, time intervals and magnitude intervals. In this way the occurrence rate of earthquakes is modeled as a continuous rate density and the likelihood of an observation is computed consequently. The model used in this paper contains both a time-independent term, function of the space coordinates, and a time dependent term, conditioned by the previous events in order to model the mutual interaction among earthquakes (short term earthquake clustering).

Section 3 contains a list and a short description of a set of computer programs, each of which performs one of the steps necessary for processing a seismic catalog. In this respect, most of the programs are linked together by their output and input, making the whole set a sort of dedicated software package.

The main intention of the report is more methodological than scientific. For this reason, the bibliography is intentionally scarce and limited to the papers where the concepts used here were earlier introduced or are given in a suitably elementary way.

1. Statistical background

In the context of this section an *event* is any circumstance, or set of circumstances, that can be described objectively, even if just in an ideal way. Along the report, a *seismic event* is the occurrence of an earthquake in a defined volume of the space of its characterizing parameters. In this respect, earthquakes can be considered from the view point of the point process theory (Vere-Jones, 1995). The parameters considered in this report are those usually reported in earthquake catalogs, like the geographical coordinates (x and y), the origin time t and the magnitude m . Depth and other parameters could be introduced in straightforward way.

1.1 The Bayes theorem

Event A is true when the circumstance which it is referred to happens in reality. The probability that event A is true is indicated with $P[A]$. By intersection of two events $A \cap B$, we intend an event whose description includes both the descriptions of A and B jointly (Fig.1).

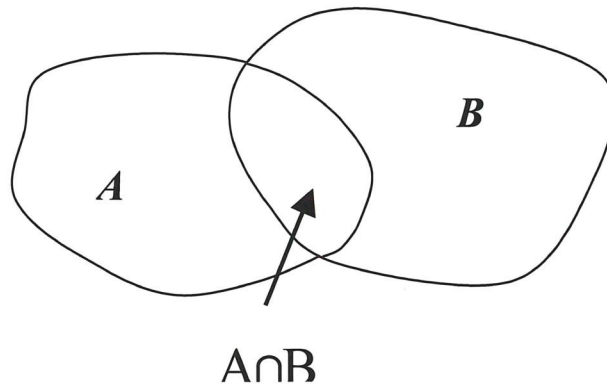


Fig. 1 – Intersection of two events.

The *conditional probability* that event A is true when also event B is true is given by:

$$P[A | B] = \frac{P[A \cap B]}{P[B]}. \quad (1.1.1)$$

If event B can be thought as subdivisible in various mutually exclusive sub-events B_i (if one is true, none of the others can be true), and if event A is *totally* contained in B (Fig. 2), we can write:

$$P[A] = \sum_{i=1}^N P[A \cap B_i] \quad (1.1.2)$$

and, using (1.1.1):

$$P[A] = \sum_{i=1}^N P[A | B_i] P[B_i]. \quad (1.1.3)$$

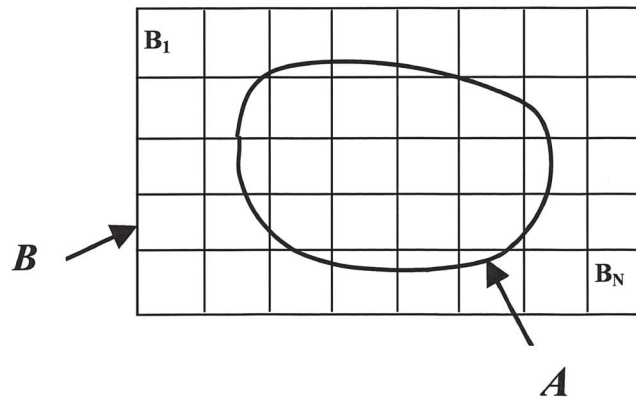


Fig. 2 – Intersection of two events. Event A is totally contained in B , and B is divided in various mutually exclusive sub-events B_i ($i=1, \dots, N$).

Let's now consider the probability that one of the sub-events of B is true, conditioned by event A , repeating what is expressed by (1.1.1):

$$P[B_j | A] = \frac{P[B_j \cap A]}{P[A]} = \frac{P[A \cap B_j]}{P[A]} \quad (1.1.4)$$

Using again (1.1.1) for rewriting the numerator of the last term of (1.1.4), and (1.1.3) for rewriting the denominator, we obtain:

$$P[B_j | A] = \frac{P[A | B_j]P[B_j]}{\sum_{i=1}^N P[A | B_i]P[B_i]} \quad (1.1.5)$$

Equation (1.1.5) expresses the Bayes theorem, that can be written also in words:

$$P[\text{hypothesis}_j | \text{observ.}] = \frac{P[\text{observ.} | \text{hypothesis}_j]P[\text{hypothesis}_j]}{\sum_{\text{all-hypothesis}} P[\text{observ.} | \text{hypothesis}_i]P[\text{hypothesis}_i]} \quad (1.1.6)$$

i.e. the probability that a certain hypothesis is true at the light of a certain observation is equal to the product of the probability that such observation is realized under that hypothesis by the probability that such hypothesis is true in absolute, divided by the sum carried out over all the possible hypotheses, of the probability that such observation is realized under each single hypothesis, multiplied by the absolute probability that such hypothesis is true. The probability that a certain observation is realized under a given hypothesis is called *likelihood* of that observation.

The Bayes theorem tells us how the *a-priori* probability of a given hypothesis is modified, after that hypothesis has been tested through an experiment with a certain outcome. The conditional probability (new probability in light of the observation) can be either larger or smaller than the *a-priori* probability, and its computation requires the knowledge of all the latter.

Let's suppose, now, that we have to choose only between two hypotheses, i.e. that we are sure that either of them is true. We can rewrite (1.1.6) for each of the two hypotheses and compute their ratio, obtaining:

$$\frac{P[\text{hypothesis}_1 | \text{observ.}]}{P[\text{hypothesis}_0 | \text{observ.}]} = \frac{P[\text{observ.} | \text{hypothesis}_1]P[\text{hypothesis}_1]}{P[\text{observ.} | \text{hypothesis}_0]P[\text{hypothesis}_0]} \quad (1.1.7)$$

In light of a given observation, the ratio of the probabilities that one or the other of the hypotheses is true, is given by the ratio between the products of the probabilities (likelihoods) of making such observation under each hypothesis, by the *a-priori* probability of the respective hypothesis. A slightly different way of writing (1.1.7) is:

$$\frac{P[\text{hypothesis}_1 | \text{observ.}]}{P[\text{hypothesis}_0 | \text{observ.}]} = \left(\frac{P[\text{observ.} | \text{hypothesis}_1]}{P[\text{observ.} | \text{hypothesis}_0]} \right) \left(\frac{P[\text{hypothesis}_1]}{P[\text{hypothesis}_0]} \right) \quad (1.1.8)$$

that shows how the ratio of the absolute probabilities, which should be known, is modified by the ratio of the likelihoods obtained from the observation, to obtain the ratio of the conditional

probabilities. The ratio of the likelihoods, called by Rhoades & Evison [1997] *performance factor* of the new model, tells how much we trust more (or less) hypothesis 1 with respect to hypothesis 0 after the outcome of the experimental observation. It is function of time: when an event occurs where the expected rate density under hypothesis 1 is larger than that under hypothesis 0, it increases suddenly. It decreases in the contrary case. During the time intervals with no events, it increases if the expected rate density is smaller under hypothesis 1 than under hypothesis 0, and decreases in the contrary case. The ratio of the absolute probabilities is a quantity that represents our *a-priori* confidence in hypothesis 1 with respect to hypothesis 0. When hypothesis 0 is taken as reference, and hypothesis 1 is a new one to be tested, usually this ratio is given the value of 20 (confidence level equal to 95%).

1.2 Likelihood function

If the occurrence of an event is described as a stationary random process, and λ is the expected number of events in the realization of a test that can be carried out as many times as we can imagine, the probability that in a single realization the number of events is n is given by the Poisson distribution:

$$p(n | \lambda) = \frac{e^{-\lambda} \lambda^n}{n!}. \quad (1.2.1)$$

From (1.2.1), the probability of *no event* is:

$$p(0 | \lambda) = e^{-\lambda} \quad (1.2.2)$$

and, then, the probability of *at least one event* is:

$$p(n \geq 1 | \lambda) = 1 - e^{-\lambda}. \quad (1.2.3)$$

For the combination of events, we may use the rules of the theory of probability. If λ_1 and λ_2 are the expected number of events in two *independent* tests, the probability of *no event in any of the two tests* is:

$$p(0 | \lambda_1; 0 | \lambda_2) = p(0 | \lambda_1) \cdot p(0 | \lambda_2) = e^{-(\lambda_1 + \lambda_2)}; \quad (1.2.4)$$

so, the probability of *at least one event in either of the two tests* is:

$$1 - p(0 | \lambda_1; 0 | \lambda_2) = 1 - e^{-(\lambda_1 + \lambda_2)}. \quad (1.2.5)$$

In the same way, the probability of *at least one event in both tests* is:

$$p(n \geq 1 | \lambda_1; n \geq 1 | \lambda_2) = p(n \geq 1 | \lambda_1) \cdot p(n \geq 1 | \lambda_2) = (1 - e^{-\lambda_1}) \cdot (1 - e^{-\lambda_2}). \quad (1.2.6)$$

Generalizing (1.2.4) and (1.2.6), the probability of *no event in any of a series of P' tests*, each of which with expected number of events $\lambda_i (i = 1, \dots, P')$ is given by:

$$p(0 | \lambda_1; \dots; 0 | \lambda_{P'}) = \prod_{i=1}^{P'} p(0 | \lambda_i) = e^{-\sum_{i=1}^{P'} \lambda_i} \quad (1.2.7)$$

and the probability of *at least one event in all of a series of P'' tests*, each of which with expected number of events λ_j ($j = 1, \dots, P''$) is given by:

$$p(n \geq 1 | \lambda_1; \dots; n \geq 1 | \lambda_{P''}) = \prod_{j=1}^{P''} p(n \geq 1 | \lambda_j) = \prod_{j=1}^{P''} (1 - e^{-\lambda_j}). \quad (1.2.8)$$

We have come to the possibility of computing the probability that a series of P tests comes out with a pre-determined realization of successes and failures, in the sense that for each test it is known the expected number of events and it is requested that such test comes out with at least one event (success) or no event (failure). This probability is given by the product of (1.2.7) and (1.2.8), with $P = P' + P''$:

$$\begin{aligned} L &= p(0 | \lambda_i, i = 1, \dots, P'; n \geq 1 | \lambda_j, j = 1, \dots, P'') \\ &= \prod_{i=1}^{P'} p(0 | \lambda_i) \cdot \prod_{j=1}^{P''} p(n \geq 1 | \lambda_j) = e^{-\sum_{i=1}^{P'} \lambda_i} \prod_{j=1}^{P''} (1 - e^{-\lambda_j}). \end{aligned} \quad (1.2.9)$$

The probability given by (1.2.9) is called *likelihood* of the realization of the set of tests (that can be thought as a unique test as a whole). It depends, obviously on the result of the experiment, but it is function of the parameters by which the probabilities are determined (*a-priori* model of occurrence). From that comes the term *likelihood function*.

1.3 Alternative formulations

The likelihood function can be written in a more synthetic way by introducing the factor c_i ($i = 1, \dots, N$), the value of which is 1 for each successful test and 0 for each contrary case. In this way (1.2.9) is written:

$$L = p(0 | \lambda_i, i = 1, \dots, P'; n \geq 1 | \lambda_j, j = 1, \dots, P'') = \prod_{i=1}^P [p(0 | \lambda_i)^{(1-c_i)} \cdot p(n \geq 1 | \lambda_i)^{c_i}] \quad (1.3.1)$$

and, taking into account that $p(0 | \lambda_i) = 1 - p(n \geq 1 | \lambda_i)$,

$$L = \prod_{i=1}^P [p(n \geq 1 | \lambda_i)^{c_i} \cdot (1 - p(n \geq 1 | \lambda_i))^{(1-c_i)}]. \quad (1.3.2)$$

Finally, calling simply $p_i = p(n \geq 1, \lambda_i)$ the probability of success of the i th test, we have:

$$L = p(0 | \lambda_i, i = 1, \dots, P'; n \geq 1 | \lambda_j, j = 1, \dots, P'') = \prod_{i=1}^P [p_i^{(c_i)} \cdot (1 - p_i)^{(1-c_i)}]. \quad (1.3.3)$$

The likelihood functions may assume very small values, especially when the number of single tests is high and the probabilities connected with them are small. For this reason, it is practical using the logarithm of this function (*log-likelihood*) generally represented with $\log L$; so, passing to logarithms, (1.3.3) becomes (Kagan & Jackson, 1995; Jackson, 1996):

$$\log L = \sum_{i=1}^P [c_i \log(p_i) + (1 - c_i) \log(1 - p_i)] = \sum_{i=1}^P \left[c_i \log \frac{p_i}{1 - p_i} + \log(1 - p_i) \right] \quad (1.3.4)$$

It should be noted that only the first term of the latter expression under the sum depends on the actual realization of successes, while the second one is a constant that depends on the probabilities provided by the considered model. Kagan & Jackson [1995] consider this second term a sort of cost of betting. It would be the final cost, if no test came up with a success. The first term, that depends on the result of each test, is positive when the success is obtained for a test to which we assigned a probability p_i larger than 0.5, and negative in the contrary case. The circumstance of an event occurred in a test to which we had assigned zero probability, leads to a value of L equal to 0 and that of $\log L$ to $-\infty$: it must be avoided in the practice.

If all the probabilities p_i are small the following approximations hold:

$$p_i \cong \lambda_i \quad (1.3.5)$$

(probability of at least one event)

$$p(0 | \lambda_i) \cong 1 - \lambda_i \cong 1 - p_i \quad (1.3.6)$$

(probability of no event).

Moreover it results:

$$\log(1 - p_i) = \log(e^{-\lambda_i}) = -\lambda_i \quad (1.3.7)$$

(log. of probability of no event)

Under the approximations given by (1.3.5) and (1.3.7), formula (1.3.4) becomes:

$$\log L = \sum_{i=1}^P [c_i \log(p_i) - p_i] = \sum_{j=1}^N \log \lambda_j - N' \quad (1.3.8)$$

where N is the total number of occurred events and N' is the expected total number of events. In (1.3.8) we eliminated all the zero terms and kept only the sum over the events that occurred in reality.

1.4 Likelihood ratio

We have seen that the likelihood function depends on the hypotheses made for computing the α -

priori probability of event occurrence. In general, we are not interested in the value itself of this function (it depends on the way the extension and the number of the regions are defined), but we are interested mainly in the ratio of two of these functions computed under two different hypotheses or models, but with the same subdivision in regions (sub-volumes) and with the same catalog containing the observations. As seen in section 1.1, we shall attribute more credibility to the hypothesis that provides the larger value of likelihood. In order to define the best set of parameters characterizing a certain model of occurrence, we look for the set of parameters that provides the maximum value of L (or the minimum absolute value of $\log L$, which is negative). This procedure of maximization is called the *learning phase* for the hypothesis. Rhoades & Evison [1989] notice that a test for validating a new hypothesis should be carried out on data completely independent of those that are used in the learning phase for setting up the hypothesis.

The comparison between two L functions is in general carried taking one of them as reference (*null* hypothesis), defined in the simplest possible way. An example, in the subject of earthquake forecasting, can be that of taking as reference hypothesis a distribution that depends only on space and magnitude, but not on time, for which a Poisson random distribution is assumed.

Naming by p_i and p_{oi} the probabilities of occurrence of at least one event in the i th region under the "new" and null hypothesis respectively, and with L and L_0 their likelihoods, we infer, from (1.3.4):

$$\log R = \log \left(\frac{L}{L_0} \right) = \log L - \log L_0 = \sum_{i=1}^P \left[c_i \log \frac{p_i(1-p_{oi})}{p_{oi}(1-p_i)} + \log \frac{1-p_i}{1-p_{oi}} \right] \quad (1.4.1)$$

The reader could notice a correction with respect to equation [4] in Jackson (1996). Under the approximations given by (1.3.5) and (1.3.7), we have:

$$\sum_{i=1}^P c_i \log \frac{p_i(1-p_{oi})}{p_{oi}(1-p_i)} = \sum_{i=1}^P c_i \log \frac{p_i}{p_{oi}} = \sum_{i=1}^P c_i \log \frac{\lambda_i}{\lambda_{oi}} = \sum_{j=1}^N \log \frac{\lambda_j}{\lambda_{oj}}. \quad (1.4.2)$$

As in (1.3.8), we eliminated all the zero terms and kept only the sum over the events really occurred, obtaining the sum of the ratio of the occurrence rate for all the events.

At the same way, for the second term under the sum in (1.4.1):

$$\sum_{i=1}^P \log \frac{1-p_i}{1-p_{oi}} = \sum_{i=1}^P (-\lambda_i) - \sum_{i=1}^P (-\lambda_{oi}) = N'_0 - N' \quad (1.4.3)$$

where N'_0 and N' are respectively the total expected number of events under the null hypothesis and the "new" hypothesis. So, combining (1.4.1), (1.4.2) and (1.4.3) we have:

$$\log R = \log L - \log L_0 = \sum_{j=1}^N \log \frac{\lambda_j}{\lambda_{oj}} + N'_0 - N' \quad (1.4.4)$$

It should be noted that $N'_0 - N'$ are *not* numbers of events counted from the observation, but computed by the sum of the occurrence rates estimated over all the considered volume. This could not be trivial in practice, because it requires the full and detailed knowledge of the hypotheses and of their implication on the expected occurrence rates. In general, this computation doesn't give an integer number as result. The information drawn from the observation is, however, contained in the sum of the first term, computed for all the events really occurred.

2. Application to seismic catalogs

Kagan & Jackson [1995] apply the concept of likelihood to the validation of a hypothesis of forecast of seismic events, in the frame of the point process theory. Given N seismic events in the overall volume of the parameters (space, time, magnitude) spanned by a catalog, let's imagine such volume subdivided in *regions* extended in space, time and magnitude. The regions are small enough, so as in each of them the event occurrence is modeled by a stationary random process. The important aspect is that these regions, or sub-volumes, are disjoint, i.e. that if an event occurs in one of them, it can not also occur in another. Each event reported in the catalog is a success for the particular region ($c_i = 1$). The value of L for the realization of the experiment described by the set of c_i provided by the catalog, can be computed on the base of the particular set of p_i consequent to a model of occurrence. In Kagan & Jackson [1995] the regions are effectively geographical regions, the time extent is that spanned by the whole catalog and the magnitude extent is for all magnitudes exceeding the magnitude threshold. Then, in this case a success is the occurrence of at least one seismic event exceeding the magnitude threshold, in the total time of observation, in each specific geographical region.

2.1 Extension to the continuum

The considerations based on the subdivision of the space-time-magnitude volume in regions have the drawback that the occurrence probabilities can differ considerably from one region to the bordering one. So, the success or the failure of a hypothesis can strongly depend on small errors in the location and magnitude estimation. The statistical procedures should take into account the uncertainty in the measurements and the imperfect specificity in the definition of the hypothesis. One way of facing this problem is that of considering smaller regions, introducing a smoothing of the probability variations from a region to another. By a matter of fact, this leads to formulating the hypothesis of occurrence in terms of continuous variables, as function of the distance (in space and time) from earlier seismic events or of the localization of preceding geophysical anomalies. Let's consider a refinement of the concept of "region" which includes regions, or sub-volumes, of infinitesimal space-time-magnitude dimensions, for which we can use the approximations used in (1.4.2), (1.4.3) and (1.4.4), and apply it to (1.2.9):

$$\log L = \log e^{-\sum_{i=1}^{P'} \lambda_i} + \log \prod_{j=1}^{P''} (1 - e^{-\lambda_j}) = -\sum_{i=1}^{P'} \lambda_i + \sum_{j=1}^{P''} \log(1 - e^{-\lambda_j}) = \sum_{j=1}^{P''} \log \lambda_j - \sum_{i=1}^{P'} \lambda_i, \quad (2.1.1)$$

where, again, the first sum is taken over all the successes, and the second over all the failures. The two sums taken together include all the possible regions, i.e. the total volume of the samples. Here, the assumption that in each region the event occurrence is modeled as a stationary random process is implicit. The problem that the probability of occurrence in each single region is very low is

overcome by the possibility of ideally imagine the repetition of the test in the same conditions, as many times as we want.

The strict application of (2.1.1) to the case limit of $\lambda_i \rightarrow 0$ would lead to the value of $-\infty$ for the first term, because the number of addends is equal to the number of successes, while their value tends to zero. At this point, we introduce the definition of *occurrence rate density*:

$$\lambda(x, y, t, m) = \lim_{\Delta x, \Delta y, \Delta t, \Delta m \rightarrow 0} \frac{P_{\Delta x, \Delta y, \Delta t, \Delta m}(x, y, t, m)}{\Delta x \Delta y \Delta t \Delta m} \quad (2.1.2)$$

where $\Delta x, \Delta y, \Delta t$ and Δm are increments of longitude, latitude, time and magnitude, $P_{\Delta x, \Delta y, \Delta t, \Delta m}(x, y, t, m)$ is the probability of occurrence of an event in such volume. In many papers by Ogata, Vere-Jones, Rhoades & Evison, etc., the occurrence rate density is called *intensity function*. We prefer to follow the terminology given by Jackson [1996] who calls it *conditional rate density* in order to avoid confusion with the concept of seismic intensity.

If we substitute the occurrence rate density $\lambda(x, y, t, m)$ in the place of the expected number of events λ_i in (2.1.1), we obtain a function which doesn't diverge in the limit case. The effect of the division by $\Delta x \Delta y \Delta t \Delta m$ disappears when we divide two likelihoods by each other, regardless of the infinitesimal size of the volume $\Delta x \Delta y \Delta t \Delta m$, imaging that the sub-division in smaller volumes is carried out in identical way. So, from now on, we shall deal with the occurrence rate density $\lambda(x, y, t, m)$, using the same symbol we have used for the occurrence rate, but recalling that it maintains a finite value and it has dimensions of the inverse of the parameters used for the definition of the sub-volumes. The model we are considering here, where the rate density changes from point to point, while the occurrence rate in an infinitesimal volume surrounding the point is still modeled as a stationary random process, is called *generalized Poisson process*.

Equation (2.1.1) now becomes:

$$\log L = \sum_{j=1}^N \log[\lambda(x_j, y_j, t_j, m_j) V_{tot}] - \int \int \int \int_{XY, 0, M}^T \lambda(x, y, t, m) dx dy dt dm \quad (2.1.3)$$

where $V_{tot} = \int \int \int \int_{XY, 0, M}^T dx dy dt dm$ is a dimensional coefficient. Once again here we recall that the integral in (2.1.3) gives the total *expected number* of events according to the present hypothesis.

2.2 The case of time independence

The case of time independence is mainly relevant to the null hypothesis (Poisson model). By this hypothesis, we assume that the point process doesn't have memory and the probability of occurrence of future events is constant irrespectively of the past activity. It is generally assumed the validity of the Gutenberg-Richter law:

$$\mu(x, y, m) = e^{\alpha(x, y) - \beta(x, y)m}, \quad (2.2.1)$$

where $\mu(x, y, m)$ is the space density of earthquakes of magnitude equal to or larger than m .

If $\beta(x, y)$ is assumed independent of the space coordinates ($\beta = b \ln(10) = 2.30259b$) (while α is allowed to change from point to point), we obtain the decomposition

$$\mu(x, y, m) = \mu_o(x, y) \cdot e^{-\beta(m-m_o)}, \quad (2.2.2)$$

where m_o is the minimum magnitude of completeness and $\mu_o(x, y) = \mu(x, y, m_o)$ is the space density of earthquakes of magnitude equal to or larger than m_o .

By derivation of (2.2.2) with respect to m , in reverse order, the density magnitude distribution of earthquakes is obtained:

$$\lambda_o(x, y, m) = \mu_o(x, y) \beta e^{-\beta(m-m_o)}. \quad (2.2.3)$$

Given a catalog of events $\{(x_j, y_j, t_j, m_j), j = 1, \dots, N\}$, with $0 \leq t_j \leq T$ and $m_j \geq m_o$, (2.1.3) is written:

$$\begin{aligned} \log L &= \sum_{j=1}^N \log[\mu_o(x_j, y_j) \beta e^{-\beta(m_j-m_o)}] - \int \int \int_{XY} \int_{m_o}^{T+\infty} \mu_o(x, y) \beta e^{-\beta(m-m_o)} dx dy dt dm \\ &= \sum_{j=1}^N \log[\mu_o(x_j, y_j) \beta e^{-\beta(m_j-m_o)}] - T \int \int_{XY} \mu_o(x, y) dx dy. \end{aligned} \quad (2.2.4)$$

The second term is, obviously, the total expected number of earthquakes of magnitude equal to or larger than m_o in the considered space and time extent.

Let's suppose there is a region in which $\mu_o(x, y)$ is constant; in practice let's consider the events of a catalog regardless of the space coordinates and time of the events, and of which only the magnitude is relevant. We can obtain as an exercise the formula used for estimating the b-value of the catalog by the maximum likelihood method.

Let be

$$\mu(m) = \mu_o \cdot e^{-\beta(m-m_o)} \quad (2.2.5)$$

and

$$\lambda(m) = \mu_o \cdot \beta \cdot e^{-\beta(m-m_o)}. \quad (2.2.6)$$

By substitution into (2.2.4) we have

$$\log L = \sum_{j=1}^N [\log(\mu_o \beta) - \beta(m_j - m_o)] - T \mu_o. \quad (2.2.7)$$

In order to obtain the value of β corresponding to the maximum of $\log L$, we derive with respect to β :

$$\frac{\partial \log L}{\partial \beta} = \sum_{j=1}^N \left[\frac{1}{\beta} - (m_j - m_o) \right] = \frac{N}{\beta} - \sum_{j=1}^N m_j + N m_o \quad (2.2.8)$$

Equating to zero the derivative, we have:

$$\beta = \frac{1}{\frac{\sum_{j=1}^N m_j}{N} - m_o} = \frac{1}{\bar{m} - m_o} \quad (2.2.9)$$

from which

$$b = \frac{\log_{10} e}{\bar{m} - m_o} \quad (2.2.10)$$

In ideal way, given the magnitude distribution (2.2.5) and (2.2.6) as continuous, the same result can be obtained by computing the average magnitude of the events exceeding the magnitude threshold m_o :

$$\bar{m} = \frac{1}{\mu_o} \int_{m_o}^{+\infty} m \lambda(m) dm = \beta \int_{m_o}^{+\infty} m \cdot e^{-\beta(m-m_o)} dm = \frac{1}{\beta} + m_o \quad (2.2.11)$$

If $\beta = 2.3$ ($b = 1$) the average magnitude is about 0.43 units larger than the minimum.

2.3 Time-independent model

The computation of (2.2.4) for hypothesis test purposes requires the knowledge of the space density distribution $\mu_o(x, y)$ as a continuous, smooth function of the geographical coordinates (x, y) . In the present work, this is based entirely on instrumental catalogs of past seismicity, ignoring additional information such as intensity observations of historical earthquakes or surface faulting from prehistoric earthquakes. Each point source of the catalog is replaced with a sort of density function $f_i(x, y)$ (*kernel* function) which is greatest at the epicenter but tails off over some region surrounding it. The $\mu_o(x, y)$ distribution is obtained by summing all $f_i(x, y)$ for the earthquakes in the catalog and scaling appropriately.

Here, the method presented by Frankel [1995] is adopted. First, the region to be analyzed is subdivided in square cells of appropriate size. The number of earthquakes N_k in each cell k represents the maximum likelihood estimate of the number of earthquakes above m_o for that cell. The grid of N_k values is then smoothed spatially by multiplying by a Gaussian function with correlation distance c . For each cell k , the smoothed value $\tilde{N}_k$ is obtained from:

$$\tilde{N}_k = \frac{\sum_l N_l e^{-\Delta_{kl}^2/c^2}}{\sum_l e^{-\Delta_{kl}^2/c^2}} \quad (2.3.1)$$

where Δ_{kl} is the distance between the centers of the k th and l th cells. In equation (2.3.1), $\tilde{N}_k$ is normalized to preserve the total number of events. In order to implement equation (2.3.1) one has to choose c : the larger c is, the smoother $\tilde{N}_k$ will be, but also the larger will be the contamination between different areas. In this work, the optimal choice of c is made by trial and error, adjusting c until the maximum likelihood is obtained from a half of the catalog and a seismicity grid derived from the other half. The values of $\tilde{N}_k$ obtained so far still change abruptly from one cell to another. In order to obtain a continuous function, the single value of $\mu_o(x, y)$ is computed by interpolation between the four cells whose centers surround the point (x, y) . The result is given as number of events per cell size, per total time spanned by the catalog. In order to have it in units of number of events per unit time and per unit area, it must be divided by the total time of the input catalog and the area of a cell.

2.4 Induced seismicity

As commonly shown by the experience, the occurrence rate of seismic events in a particular region, increases soon after a strong earthquake, and then it goes down again, according to the so called *modified Omori law* (Ogata, 1983) expressed, for events exceeding the threshold magnitude m_o by:

$$\lambda(t, m \geq m_o) = K(t + c)^{-p}, \quad (2.4.1)$$

with K , c and p constants that are characteristic of any particular case. The expected total number of events in the period of time $[0, T]$ is obtained by integrating (2.4.1) over the time as:

$$N(T, m \geq m_o) = \begin{cases} K [\log(T + c) - \log c], & \text{if } p = 1 \\ K [(T + c)^{1-p} - c^{1-p}] / (1 - p) & \text{if } p \neq 1, \end{cases} \quad (2.4.2)$$

that allows the computation of K , from the observation of seismic events of magnitude equal to or larger than m_o over a given time window.

Let's assume that the realization of a point process described by (2.4.1) is observed as a sequence of seismic events the occurrence times of which are $\{t_i; i = 1, \dots, N\}$ in the time interval $[S, T]$. In this case, ignoring the space dependence of the rate density, (2.1.3) can be written as:

$$\log L = \sum_{i=1}^N \log \lambda(t_i) - \int_S^T \lambda(t) dt \quad (2.4.3)$$

By substitution of (2.4.1) in (2.4.3), it becomes (Ogata, 1983):

$$\log L(K, c, p) = N \log K - p \sum_{i=1}^N \log(t_i + c) - KA(c, p), \quad (2.4.4)$$

where

$$A(c, p) = \begin{cases} \log(T + c) - \log(S + c), & \text{if } p = 1 \\ [(T + c)^{1-p} - (S + c)^{1-p}] / (1 - p) & \text{if } p \neq 1, \end{cases} \quad (2.4.5)$$

The estimate of the parameters (K, c, p) which best fit the observations is carried out by the method of the maximum likelihood seeking for the maximum of (2.4.4).

The earthquakes that induce further seismicity can be a single *mainshock* (in the usual subjective meaning), as in the case of a typical *aftershock* sequence, or more than one mainshock, as in the case of a multiple aftershock sequence, or even all the previous events, as in the *epidemic model* (Ogata, 1988). The resultant expected rate density of seismic events, taking into account the contribution of the previous inducing earthquakes, will be given by:

$$\lambda(x, y, t, m) = f_r \lambda_o(x, y, m) + \sum_{i=1}^N H(t - t_i) \cdot \lambda_i(x, y, t, m), \quad (2.4.6)$$

where f_r is a factor called *failure rate*, $H(t)$ is the step function, and $\lambda_i(x, y, t, m)$ is the single contribution, the time dependence of which is expressed by (2.4.1). The first and the second term to the right of the sign of equality in (2.4.6) represent the *independent* and the *induced* seismicity respectively. The rate density corresponding to any earthquake is, in general, constituted by the superposition of both of these two components.

Following the Reasenberg & Jones [1989] hypothesis, the contribution to the occurrence rate density of any previous earthquake (x_i, y_i, t_i, m_i) is decomposable (for $t > t_i$) into three terms, respectively representing the time, magnitude and space distribution, as:

$$\lambda_i(x, y, t, m) = K(t - t_i + c)^{-p} \beta e^{-\beta(m - m_i)} f(x - x_i, y - y_i), \quad (2.4.7)$$

where $f(x, y)$ is the spatial distribution of the induced seismicity with respect to the coordinates of the inducing event.

The decomposition given in (2.4.7) implies several important hypotheses. First, we assume that β, K, c and p are independent of the geographical coordinates (x, y) and (x_i, y_i) . This is, of course, an initial approximation, that can be refined if there is enough experience for formulating a more complex situation.

Then, modelling the spatial distribution $f(x, y)$ independently of the magnitude of the inducing earthquake is another strong approximation. Refer, e.g. to Ogata [1998] for different forms of the spatial term.

Finally, it must be noted that putting the magnitude m_i of the inducing events in the same term of the magnitude m of the induced events and with the same coefficient β is a strongly subjective hypothesis, that implies the self-similarity of the process over all the ranges of magnitude. With this assumption, the probability of an earthquake of magnitude 3.0 after an earthquake of magnitude 4.0 is the same as that of an earthquake of magnitude 5.0 after an earthquake of magnitude 6.0, and so on. While it is not proved this assumption is true, we use it because of its simplicity, in the context of the parsimony of hypotheses that is a rule all through this work.

We model the spatial distribution of the induced seismicity by a function $f(x - x_i, y - y_i)$ that has circular symmetry around the point of coordinates (x_i, y_i) and is normalized to one, so as to preserve the total expected number of induced events of magnitude equal to or larger than m_o , related to K through (2.4.2). Defining the distance from the point (x_i, y_i) as $r = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ the first condition is written as:

$$f(r) = ce^{(-r^2/2\sigma^2)}, \quad (2.4.8)$$

where σ is a suitable constant which defines the "width" of the distribution and c is a normalization constant. The normalization is obtained imposing that the integral over the whole plane is equal to 1:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x - x_i, y - y_i) dx dy = 2\pi \int_0^{+\infty} f(r) dr = 1 \quad (2.4.9)$$

This, by substitution of (2.4.8), gives

$$c = \frac{1}{2\pi\sigma^2}. \quad (2.4.10)$$

Then, for the space distribution, we'll assume:

$$f(r) = \frac{1}{2\pi\sigma^2} \cdot e^{(-r^2/2\sigma^2)}. \quad (2.4.11)$$

2.5 The epidemic model

We shall now build the full occurrence rate density for the model called by Ogata [1988] as epidemic model. In this model *all* the events are likely to have influence on the occurrence rate of the following ones and *all* the events are likely to be influenced by the previous ones. The extreme simplicity of this model allows to avoid the use of arbitrary and subjective definitions as "foreshocks", "mainshocks", "aftershocks", independent and induced events, etc., that would introduce further free parameters in the model: all the earthquakes are considered identical from this view point. More sophistication in the model can be introduced at any time.

By substitution of (2.4.11) into (2.4.7), and of the latter into (2.4.6), together with (2.2.3), we have the complete expression of the rate density:

$$\lambda(x, y, t, m) = f_r \mu_o(x, y) \beta e^{-\beta(m-m_o)} + \frac{\beta K}{2\pi\sigma^2} \sum_{i=1}^N \frac{H(t-t_i)}{(t-t_i+c)^p} e^{-\beta(m-m_i)} e^{(-r^2/2\sigma^2)}, \quad (2.5.1)$$

with $r = \sqrt{(x - x_i)^2 + (y - y_i)^2}$. This expression can be substituted into (2.1.3) for computing the log-likelihood of the observation of a set of seismic events $\{(x_j, y_j, t_j, m_j), j = 1, \dots, N\}$ of magnitude equal to or larger than m_o in the time span $[0, T]$.

The first term of (2.1.3) is written explicitly:

$$\begin{aligned} \sum_{j=1}^N \log[\lambda(x_j, y_j, t_j, m_j) V_{tot}] &= \sum_{j=1}^N \log \left[f_r \mu_o(x_j, y_j, m_j) + \sum_{i=1}^{j-1} \lambda_i(x_j, y_j, t_j, m_j) \right] + \sum_{j=1}^N \log V_{tot} = \\ &= \sum_{j=1}^N \log \left[f_r \mu_o(x_j, y_j) \beta e^{-\beta(m_j - m_o)} + \frac{\beta K}{2\pi\sigma^2} \sum_{i=1}^{j-1} \frac{e^{-\beta(m_j - m_i)} e^{(-r_{ij}/2\sigma^2)}}{(t_j - t_i + c)^p} \right] + \sum_{j=1}^N \log V_{tot}. \end{aligned} \quad (2.5.2)$$

where $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$.

The second term becomes:

$$\begin{aligned} \int \int \int_{XY} \int_0^{T+\infty} \int_{m_o}^{\infty} \lambda(x, y, t, m) dx dy dt dm &= T \int \int_{XY} f_r \mu_o(x, y) dx dy + \frac{\beta K}{2\pi\sigma^2} \int \int \int_{XY} \int_0^{T+\infty} \sum_{\substack{i=1 \\ t_i \leq t}}^N \frac{e^{-\beta(m - m_i)} e^{(-r^2/2\sigma^2)}}{(t - t_i + c)^p}. \end{aligned} \quad (2.5.3)$$

The first part of the second member of (2.5.3) represents the expected total number of "spontaneous" events, and the second part represents the expected total number of "induced" events. Notice that in order to compute this quantities it is not necessary to define which particular event is "spontaneous" and which is "induced". The integration over the time in the second part of the second member of (2.5.3) can be transformed changing the variable of integration t into $t - t_i$ and the limits of integration of any term of the sum from t_i to T into 0 to $T - t_i$ (while the integration over the space is considered equal to 1), so obtaining:

$$\begin{aligned} \int \int \int_{XY} \int_0^{T+\infty} \int_{m_o}^{\infty} \lambda(x, y, t, m) dx dy dt dm &= T \int \int_{XY} f_r \mu_o(x, y) dx dy + K \sum_{i=1}^N \int_0^{T-t_i} \frac{e^{-\beta(m_o - m_i)}}{(t + c)^p} dt = \\ &= T \int \int_{XY} f_r \mu_o(x, y) dx dy + K \sum_{i=1}^N e^{-\beta(m_o - m_i)} \frac{[(T - t_i + c)^{1-p} - c^{1-p}]}{1-p}, \quad (p \neq 1) \end{aligned} \quad (2.5.4)$$

Looking at (2.5.1), it is easily recognizable that the magnitude distribution of both "spontaneous" and "induced" events follows the *Gutenberg-Richter law* (2.2.1) with the same coefficient β . Though it has been frequently claimed that this hypothesis is contradicted by the so called "*Bath's Law*" (i.e. the difference in magnitudes of the "main shock" and the largest aftershock in a sequence

is roughly 1.2), it was shown by Vere-Jones [1969] that the contradiction appears to be largely, if not entirely, due to an accidental bias in the selection of data.

3. Computer programs

The computer programs described in this section are written in FORTRAN. The users can easily compile and link them with practically any FORTRAN compiler, eventually modifying the dimensions of the arrays. Most of the programs require input files and produce output files, the names of which are entered interactively from the keyboard. Several free parameters of the algorithms are also requested interactively from the keyboard. The input and output formats are studied in order to use the output files of some programs as input for some others. Some output files are suitable for using them in input to commercial software packages as "EXCEL", "MATLAB", etc. for obtaining graphic displays.

3.1 JMA2REC, JUN2REC, PDE2REC

These programs convert a file containing a catalog of earthquakes respectively from JMA, JUNE and PDE format, into another file where:

- time is expressed in days from the first day of the starting year entered from the keyboard, and
- coordinates are converted from geographical into rectangular (in *kms*) computed with respect to the origin point, whose geographical coordinates are entered from the keyboard.

The events are selectable by the minimum magnitude and the maximum depth, entered from the keyboard.

The output format of these programs is that used for the catalogs in input to all the following programs.

3.2 REC2JMA, REC2JUN, REC2PDE

These programs perform the inverse conversion of the previous three ones. Also here it is necessary to enter the starting year and the geographical coordinates of the origin point of the rectangular coordinates from the keyboard. In order to recover the same figures of the original catalogs, it is necessary that these parameters are identical to those used in one of the previous programs for the direct conversion.

3.3 REMAFT

Selects the data contained in the input file, within given time and magnitude limits, removing all the aftershocks, and copies them on the output file. The criterion for removing aftershocks is based on the space-time volume of influence from the previous events as in Kagan [1996]. Larger limits, if wanted, can be entered from the keyboard.

3.4 SMOOTH

Creates a grid of points representing the seismicity distribution of a given area, by smoothing the epicenters of the input catalog. The value f at each node represents the number of events exceeding the minimum magnitude contained in a square cell of size dx . The values are normalized so as to preserve the total number of the original catalog.

It also computes the maximum likelihood "b-value" of the input catalog.

The limits of the rectangular area $xmin$, $xmax$, $ymin$, $ymax$ (km), the grid size (km), the minimum magnitude $magmin$, and the "c" correlation distance (km) of the smoothing algorithm are entered from the keyboard.

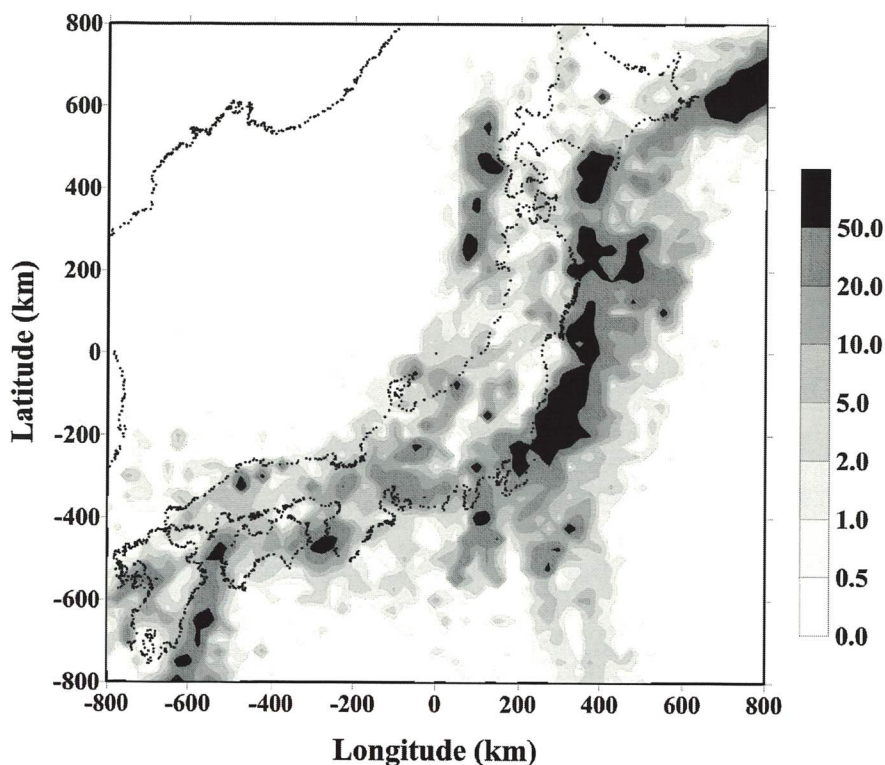


Fig. 3 – Smoothed seismicity of the Japan region, obtained from the JMA catalog 1970-1994 by program "SMOOTH". Different tones of gray represent the number of events in an area of a cell size, for the total time duration of the catalog. The total number of events of the original catalog is preserved. Saturation of values larger than 50 events/day is represented by a black tone.

The header of the output file contains: the left-bottom coordinates, the grid size (km), the number of grid points in x and y directions, the minimum magnitude $magmin$, the "b-value" b , the time extent (days), and the total number of events.

Fig. 3 shows an example of the result of smoothing through SMOOTH for the seismicity of Japan, obtained with a correlation distance of 20 kms.

3.5 GRID

Converts a file with a list of entries representing the smoothed seismicity of an area (such as in output from SMOOTH) into another file containing the same data as a table suitable for representation of the data in graphics way by means of a computer tool like EXCEL or MATLAB.

3.6 LIKELY

Given a grid of points representing the average (static) seismicity of a given area, computes the likelihood of the input list (catalog) of events, based on a model of induced seismicity the limit case of which is the Poisson random occurrence. Both the likelihoods for the Poisson (null) hypothesis and for the epidemic model, as well as the performance factor, are presented in output on the screen.

The following parameters of the occurrence model are entered from the keyboard:

- the minimum magnitude $mag0$ of the events used in the computation of the expected number of events,
- (optionally) the failure rate fr , equal to the ratio between the number of spontaneous events and the total number of events under the Poisson hypothesis.
- the standard deviation σ of the space distribution of induced events,
- the parameters K , c , p of the generalized Omori formula,
- the initial time of the computation of the likelihood, $tstart$,
- the final time at which the likelihood is computed, $tfin$.

The header of the input file with the grid points contains:

- the left-bottom coordinates, the grid size (km),
- the number of grid points in x and y directions, the minimum magnitude $magmin$,
- the "b-value" b ,
- the time extent (days),
- the total number of events of the input seismicity.

Only the events of the input list occurring inside the rectangle on which the geographical distribution of the seismicity is represented, are used for the estimation of the likelihood.

Program LIKELY can be used for two different purposes:

- for the search of the parameters that best fit a given set of observations (learning phase); in this case the failure rate fr should be left free and the other four parameters of the model (σ , K , c , p) should be changed from one test to another, according to the values of the likelihood that are obtained, until the maximum is found; it is suggested to change only one parameter at time;
- for testing the validity of a model against a new set of observations (testing phase); in this case the failure rate fr should be fixed together the other four parameters. It is possible to test any time section of the input catalog.

3.7 TIMEHIST

Given a grid of points representing the Poisson model of the seismicity of a given area, and a catalog of seismic events in the same area, computes the time history of the occurrence rate based on a model of induced seismicity according to a circular normal space distribution, to the Omori time distribution, and to the GR frequency-magnitude distribution.

The occurrence rate computed in this program is the expected number of events between the present time and the present time plus the time step dt , by integration of the occurrence rate density. The

occurrence rate can be referred to a single point (rate density normalized to the size of a grid cell) or to the whole area.

For comparison, also the real number of events occurred in the same area and time windows is counted and written in output.

The following parameters are entered from the keyboard:

- the minimum magnitude $mag0$ used in this computation,
- the initial and the final time of the time history,
- the time step of the time history,
- the failure rate fr , equal to the ratio between the number of spontaneous events and the total number of events under the Poisson hypothesis.
- the standard deviation σ of the space distribution,
- the parameters K , c , p of the Omori law.

The header of the input file contains:

- the left-bottom coordinates,
- the grid size (km),
- the number of grid points in x and y directions,
- the minimum magnitude $magmin$,
- the " b -value" b ,
- the time extent (days),
- the total number of events of the input grid of reference seismicity.

Fig. 4 and Fig. 5 respectively show a plot of the time history of the expected occurrence rate for the whole Japanese territory, and the rate density for a single point.

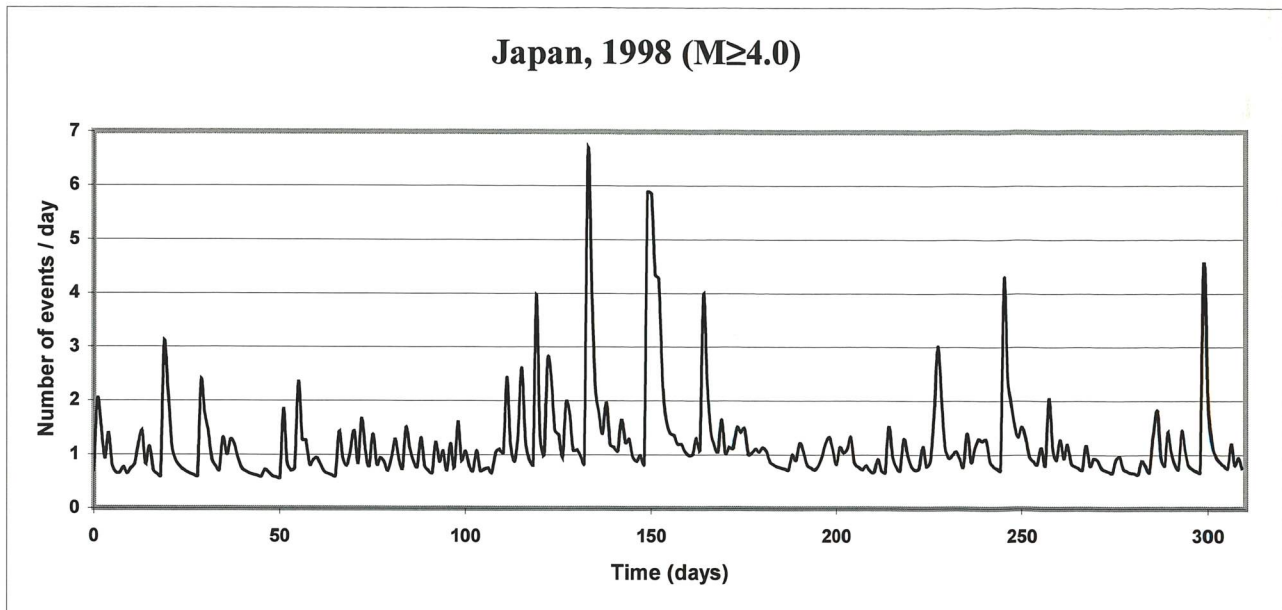


Fig. 4 – Time history of the expected rate of events of magnitude equal to or larger than 4.0 in Japan for 1998, obtained plotting the output table from program “TIMEHIST”.

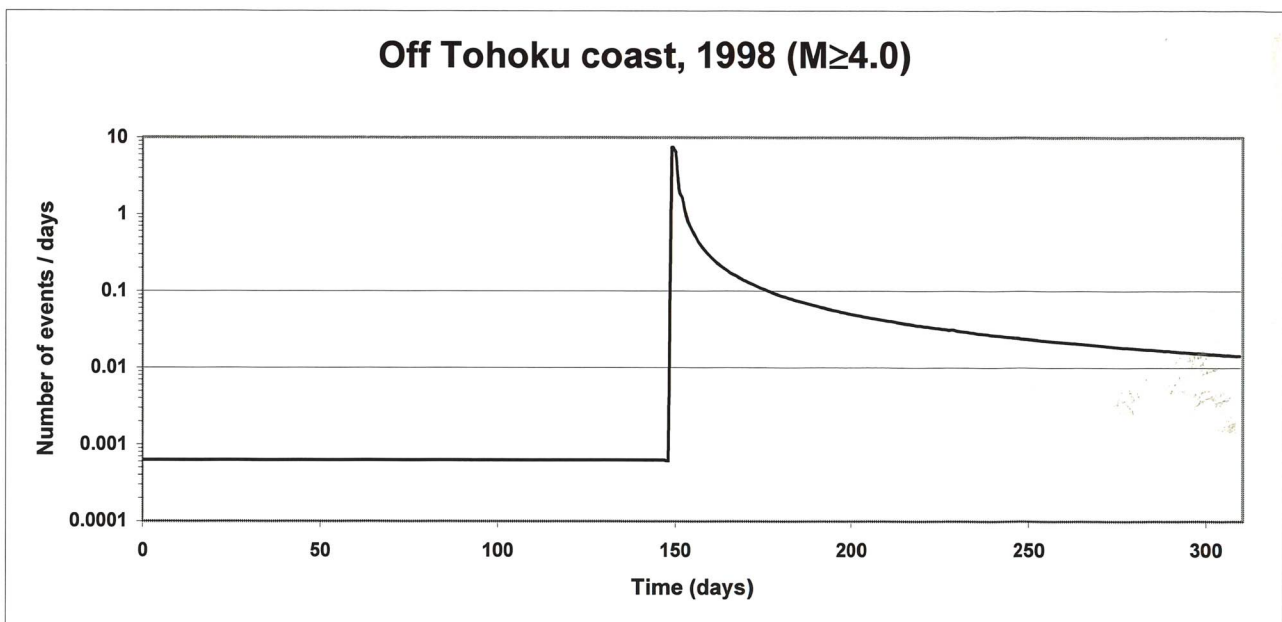


Fig. 5 – Time history of the expected rate density on the epicenter of the 30 May 1998, at 18:18 GMT, magnitude 6.2 earthquake, occurred offshore of the Tohoku coast, obtained as explained in Fig. 4.

3.8 DENSMAP

Given a grid of points representing the average (static) seismicity of a given area, and an input catalog of observed events, creates another grid of points representing the density of occurrence rate

at the final time, based on a model of induced seismicity, the limit case of which is the Poisson random occurrence.

The following parameters of the occurrence model are entered from the keyboard:

- the minimum magnitude $mag0$ of the events used in the computation of the expected number of events,
- the failure rate fr , equal to the ratio between the number of spontaneous events and the total number of events under the Poisson hypothesis.
- the standard deviation σ of the space distribution of induced events,
- the parameters K , c , p of the generalized Omori formula,
- the initial time at which the computation is started, $tstart$
- the final time at which the rate density is computed, $tfin$.

The header of the input file with the grid points contains:

- the left-bottom coordinates,
- the grid size (km),
- the number of grid points in x and y directions,
- the minimum magnitude $magmin$,
- the "b-value" b ,
- the time extent (days),
- the total number of events of the input grid of reference seismicity.

Only the events of the input list occurring inside the rectangle on which the geographical distribution of the seismicity is represented, are used for the estimation of the likelihood.

Fig. 6 shows a snap-shot of the rate density for an area of the Japanese territory, and Fig. 7 shows the rate density 24 hour later, after the occurrence of an earthquake of magnitude 6.2.

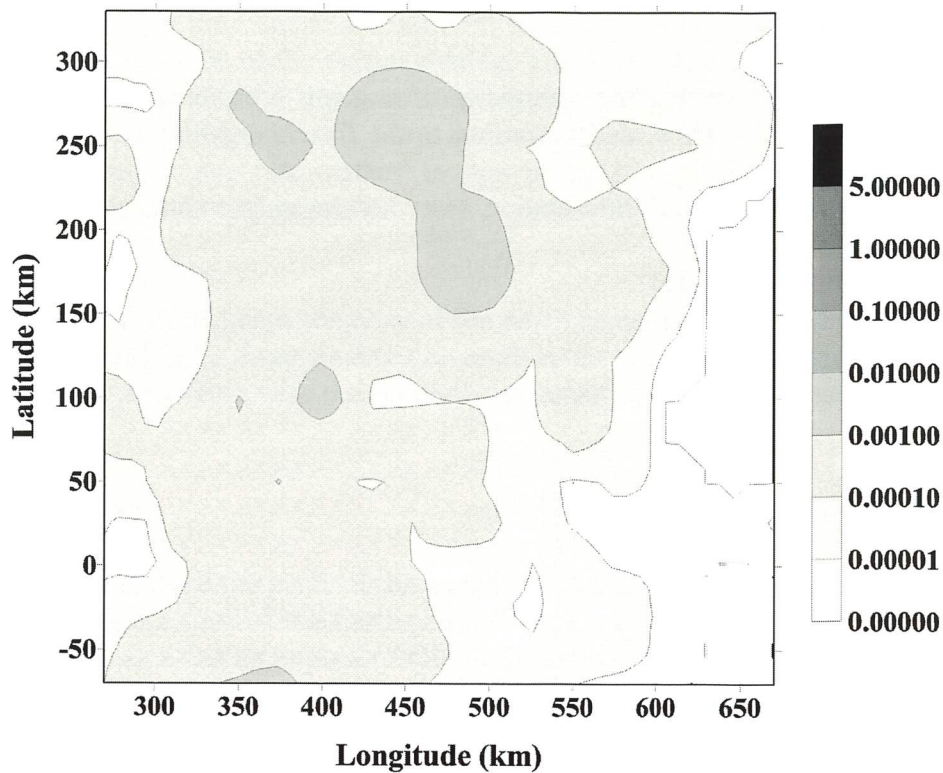


Fig. 6 – Geographical distribution of the expected rate density of events of magnitude equal to or larger than 4.0 for the region of the 30 May 1998, magnitude 6.2 earthquake of Fig. 5. The map has been obtained plotting the output table from program “DENSMAP, computed for the time May 30, 1998 at 00:00 GMT.

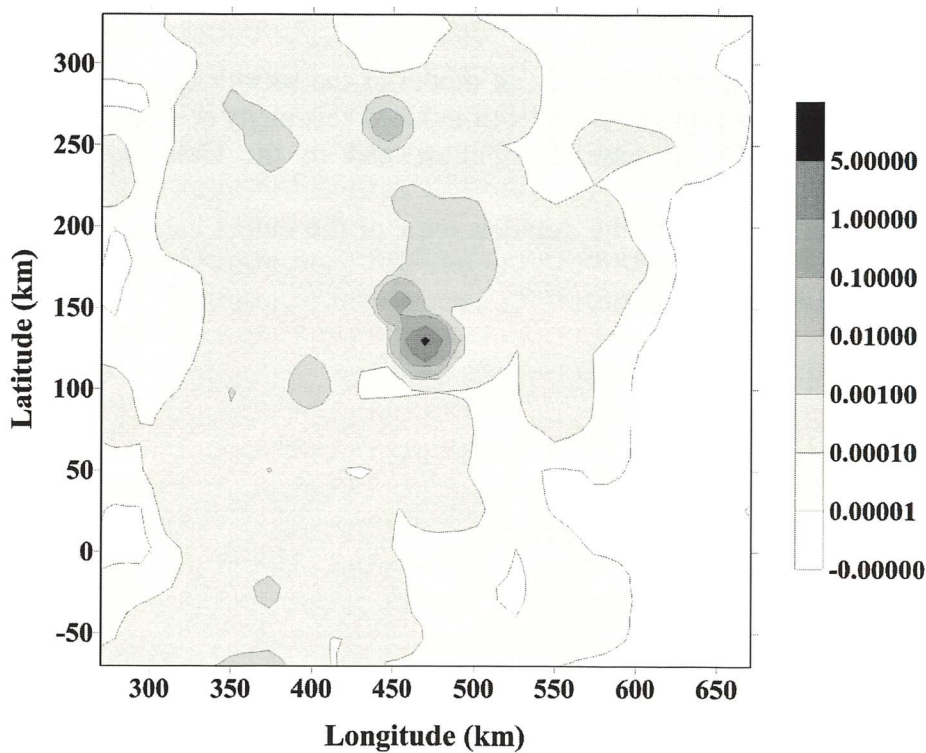


Fig. 7 – As in Fig. 6, after the magnitude 6.2 earthquake of 30 May 1998, at 18:18 GMT, computed for the time May 31, 1998 at 00:00 GMT.

3.9 SEISYNT

Given a grid of points representing the seismicity of a given area, creates a catalog of random events uniformly distributed in time, and according to the Gutenberg-Richter frequency-magnitude distribution.

The minimum magnitude *mag0* and the duration *tmax* of the output catalog are entered from the keyboard.

The header of the input file contains:

- the left-bottom coordinates,
- the grid size (*km*),
- the number of grid points in *x* and *y* directions,
- the minimum magnitude *magmin*,
- the "*b*-value" *b*,
- the time extent (days),
- the total number of events of the input grid of reference seismicity.

The events are drawn randomly and then their origin times are sorted at the conclusion of the program.

3.10 CLUSYNT

Given a grid of points representing the Poisson model of the seismicity of a given area, creates a catalog of random events generated by a model of induced seismicity according to a circular normal space distribution, to the Omori time distribution, and to the Gutenberg-Richter frequency-magnitude distribution.

The minimum magnitude *mag0* and the duration *tmax* of the output catalog are entered from the keyboard.

The header of the input file contains:

- the left-bottom coordinates,
- the grid size (*km*),
- the number of grid points in *x* and *y* directions,
- the minimum magnitude *magmin*,
- the "*b*-value" *b*,
- the time extent (days),
- the total number of events of the input grid of reference seismicity.

The events are created in time steps: a minimum time step is computed depending on the total number of events and duration of the output catalog. The time interval is adjusted as the inverse of the current expected total rate. If the time interval is smaller than the minimum, one event is generated in a random moment of this time interval and in a spacial position with probability proportional to the occurrence rate density in that point. If the time interval is larger than the minimum, the event is, or is not generated, with a probability equal to the inverse of the rate per time interval.

At any time, a special routine keeps memory of the past events which are still significantly likely to induce new events, and only these events are used in the computation of the rate density. The selection is made by comparison with the average rate density of the input seismicity grid.

Acknowledgments

The work that has constituted the basis of this paper has been granted by the University of Tokyo. It has been supported by the staff of the Earthquake Research Institute, and by Prof. K. Shimazaki in particular. Dr. D. Rhoades had an important role in the preparation of the methodology. Gratitude is also due to Prof. F. Evison, Prof. R. Geller, Prof. Y. Ogata and Prof. Vere-Jones for their suggestions and encouragement.

References

- Frankel, A., 1995. Mapping seismic hazard in the Central and Eastern United States. *Seism. Res. Lett.*, 66, 4, 8-21.
- Jackson, D.D., 1996. Hypothesis testing and earthquake prediction. *Proc. Natl. Acad. Sci. USA*, 93, 3772-3775.
- Kagan, Y.Y., 1996. VAN earthquake prediction-an attempt at statistical evaluation. *Geoph. Res. Lett.*, 23, 1315-1318.
- Kagan, Y.Y. & Jackson, D.D., 1995. New seismic gap hypothesis, Five years later. *J. Geoph. Research*, 100, 3943-3959.
- Ogata, Y., 1983. Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure. *J. Phys. Earth*, 31, 115-124.
- Ogata, Y., 1988. Statistical models for earthquake occurrence and residual analysis for point processes. *J. Am. Statist. Ass.*, 83, 401, 9-27.
- Ogata, Y., 1998. Space-time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.*, 50, 2, 379-402.
- Reasenber, P.A., and Jones, L.M., 1989. Earthquake hazard after a mainshock in California, *Science*, 243, 1173-1176.
- Rhoades, D.A. & Evison, F.F., 1989. On the reliability of precursors. *Ph. Earth Plan. Int.*, 58, 137-140.
- Rhoades, D.A. & Evison, F.F., 1997. The precursory earthquake swarm in New Zealand: hypothesis tests II. *New Zealand Journal of Geology and Geophysics*, 40, 537-547.
- Vere-Jones, D., 1969. A note on the statistical interpretation of Bath's Law. *Bull. Seism. Soc. Am.*, 59, 4, 1535-1541.
- Vere-Jones, D., 1995. Forecasting earthquakes and earthquake risk. *Int. J. of Forecasting*, 11, 503-538.

