Likelihood analysis of earthquake catalogues

Y. Y. Kagan

Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024, USA

SUMMARY

We apply several classes of stochastic multidimensional models to statistical analysis of earthquake catalogues using likelihood methods. We investigate the importance of including different earthquake parameters in the model: epicentral coordinates, hypocentral depth, time limits for interearthquake interaction, and especially spatial distribution of earthquakes as well as spatial aftershock patterns. Results of this study combined with other investigations, suggest that most distributions controlling earthquake interaction have a fractal or scale-invariant form. Developed models are used for statistical analysis of several earthquake catalogues to evaluate parameters of earthquake occurrence. These parameters are shown to be similar for shallow earthquakes of different magnitude ranges and seismogenic regions, confirming self-similarity of the earthquake process. Whereas intermediate earthquakes seem to emulate the pattern of shallow earthquake occurrence, albeit at a much smaller aftershock rate, deep earthquakes differ significantly in their properties. Predictability of standard shallow earthquake catalogues has been analysed; we present evidence that for the best available catalogues the predictability is close to 10 bits per earthquake. Several synthetic earthquake catalogues have been created and processed through the likelihood inversion scheme. The results from likelihood analysis of these catalogues confirm our approach.

Key words: earthquake catalogues, likelihood function, predictability of earthquakes, scale-invariance.

1 INTRODUCTION

In this paper we apply the likelihood analysis to infer statistical properties of earthquake catalogues. To this end we formulate an appropriate stochastic model of earthquake occurrence, parametrize the model, and then define a numerical algorithm for finding maximum likelihood estimate (MLE) of the parameters. The reason we use the likelihood method instead of many other available statistical procedures is that the MLE has certain optimal properties. These estimates are used here to formulate a global model of earthquake occurrence and, finally, to extrapolate catalogues for evaluating a future earthquake hazard.

Statistical analysis of earthquake catalogues has two major aspects. One is the practical or engineering question of how to use available earthquake information to calculate a probable occurrence within certain well-defined time-space limits. The second issue is finding a satisfactory theoretical description of the earthquake source. When such a description, well-founded in continuum mechanics, becomes available, we will be able to assess predictive seismic risk fully; these assessments must include possible errors in such predictions. Such errors may be due to the inadequacy of available data and/or the inherent stochasticity of the problem.

The best available earthquake catalogues show that earthquakes can be characterized by their origin time, location, and focal mechanism, and hence are of a strongly multidimensional nature with clear imprint of randomness. Thus to describe patterns of seismicity, we should model them as multidimensional stochastic processes. Since this model provides a straightforward, formal method of representing the process, we can perform temporal extrapolations, i.e., predict earthquakes. The prediction is therefore implemented as a straightforward application of the theory of stochastic processes to earthquake data. Both the effectiveness and accuracy of the prediction can be evaluated in an objective, quantitative manner. Indeed, preliminary results of these investigations have been used by Kagan & Knopoff (1987b) to develop and test a statistical earthquake prediction technique.

In recent years several attempts have been made to use likelihood methods for interpreting earthquake catalogues (Hawkes & Adamopoulos 1973; Ozaki 1979; Vere-Jones &
The distribution of the clusters in space is more complicated statistically independent although individual earthquakes in the cluster are dependent events. The sequences are multidimensional stochastic model of earthquake rupture process, in which clusters or sequences of earthquakes are and will be discussed below.

In constructing a stochastic model of the earthquake process, we need to strike a balance between the model's mathematical tractability and the degree of approximation in the process description. In terms of the stochastic model, mathematical tractability involves certain assumptions of statistical independence. Unfortunately, at the present time there is no other method for model-building other than a 'trial and error' procedure. Thus, to formulate a model of an earthquake occurrence, we need some results from statistical studies of earthquake catalogues as well as physical insight into the earthquake generation. We have addressed both problems in our work on stochastic seismology in the last few years. We have carried out statistical analyses of the time-space-magnitude relations in several earthquake catalogues (see Kagan & Knopoff 1980b, and references therein), and have analysed the geometry of earthquake faults (Kagan 1982, and references therein).

Based on these studies we constructed a kinematic multidimensional stochastic model of earthquake rupture (Kagan & Knopoff 1981; Kagan 1982).

The assumptions we make in constructing our model may be summarized as follows.

(1) Seismicity can be approximated as a Poisson cluster process, in which clusters or sequences of earthquakes are statistically independent although individual earthquakes in the cluster are dependent events. The sequences are assumed to be a Poisson time series with a constant rate. The distribution of the clusters in space is more complicated and will be discussed below.

(2) The major assumption regarding the interrelationships between events within a cluster is that the propagation of an earthquake rupture is closely approximated by a stochastic space–time critical branching process. Under this assumption, only one earthquake influences the occurrence of a subsequent event, i.e., there is a sole trigger for any given dependent event. The time–space distribution of interrelated earthquake sources within a sequence is controlled by simple relations justified by analysing the available statistical data on seismicity.

In large measure these are ad hoc assumptions; because of this, our model must be characterized as phenomenological. These assumptions have been chosen in part because the statistical and stochastic models that can be developed from them have a logical simplicity as well as a mathematical and numerical tractability.

In this paper we review different models of earthquake occurrence with special attention to modelling of spatial patterns of earthquake hypocentres. These models are used to obtain likelihood functions for earthquake catalogues and numerically invert these catalogues for values of model parameters. We study (1) the amount of information available in these catalogues, and (2) how information depends on the model and on the type of earthquake parameters used in the MLE optimization. Due to the problem's extreme complexity, our results are far from final; it will require a major effort in theory and data collection to fully understand a stochastic earthquake interaction.

2 LIKELIHOOD FUNCTION
COMPUTATION FOR EARTHQUAKE CATALOGUES

2.1 Likelihood function

The use of the maximum likelihood procedure has been described earlier (Kagan & Knopoff 1980b); here we comment on new features introduced in our study. We also use this opportunity to review and justify the assumptions that we have made in our previous likelihood analysis of earthquake catalogues.

The previous likelihood model (Kagan & Knopoff 1980b) was based on the hypothesis that foreshock and aftershock distribution depends on the magnitude difference between them and the main event. Our research, however, convinced us that there is no principal physical difference between shocks in an earthquake sequence (see more in Kagan & Knopoff 1981). Thus, we assume that all earthquakes, disregarding their position in a foreshock–mainshock–aftershock sequence, have the same seismic moment (magnitude) distribution. It means in effect, that the difference in size distribution of fore-, main-, and aftershocks is due not to physical processes during an earthquake rupture, but to a subdivision of the sequence during an interpretation of seismograms and a subsequent analysis of catalogues. As a result, in analysing an earthquake occurrence pattern statistically, we can separate the seismic moment distribution from all other parameters of the seismicity. In the new model we may lose some information present in the difference of the b-values, but we gain significantly in simplicity and ease of prediction.

In this paper M denotes the scalar seismic moment, and m denotes the magnitude of an earthquake. To simplify rather cumbersome expressions in this section, we normalize all seismic moments of earthquakes by dividing them by the moment (M_c) corresponding to the threshold of a catalogue; thus M thereafter means M/M_c. We assume that the scalar seismic moment is distributed according to the gamma distribution (see more in Kagan 1991, and references therein):

\[
\phi(M) = C^{-1}M^{-1-\beta} \exp\left(-\frac{M}{M_c}\right) \quad \text{for } M_c \leq M < \infty. \tag{1}
\]

Here \(M_c\) is the parameter that controls the distribution in the upper range of \(M\) and \(C\) is the normalizing coefficient: \(C = M_c^{-\beta} \gamma(-\beta, M_c/M_c)\), and \(\gamma(x, y)\) is the incomplete gamma function.

We construct an earthquake intensity function, \(\Lambda(t, x, M)\) which is the probability that an earthquake will occur at time \(t\), at location \(x\), with seismic moment \(M\), given the
Likelihood analysis of earthquake catalogues

2.2. Time–seismic moment model

We first consider the likelihood function for a catalogue of earthquakes in which only two arguments of $\Lambda$ in (2), i.e., time and seismic moment, are known. Because the distribution density $\phi(m)$ is common to both right-hand terms in (2) (see equations 3 and 4), the log-likelihood is separated into two terms

$$l_m = l_0 + l_m,$$

where $l_m$ is given by equation (6) in Kagan (1991), thus we need to study only $l_0$ here.

We take for the time distribution density

$$\psi_t(\tau) = \frac{\theta \tau \tau_{M_0}^{\theta - 1}}{\tau^{\theta - 1}} (T_M^{\theta - 1} - T_m^{\theta - 1}), \quad T_M \geq \tau \geq T_m,$$

where $\theta$ is an ‘earthquake memory’ factor. The parameter $t_m$ is a coda duration time of an earthquake with the seismic moment $M$. We assume that earthquakes which occur during the coda are not included in a catalogue. We assume for the coda function

$$t_m = t, M^{1/3},$$

where $t$, is the coda duration time of an earthquake with the reference seismic moment $M$, the dependence of the duration time on the seismic moment is deduced from the results by Lee, Bennett & Meagher (1972) and equation (2) in Kagan (1991). We have set $t = 3.45 \times 10^{-3}$ days for $M = 15.0$, which approximately corresponds to the local magnitude 4.0. The maximum time limit $T_M$ is calculated similarly:

$$T_M = T_T M^{1/3},$$

where $T_T$ is the upper limit of interaction memory for an earthquake with the seismic moment $M$. Previously (Kagan & Knopoff 1980b), in order to save computation time, we set $T_T = 34.5$ days. In most present calculations, to avoid edge effects, we set $T_T > T$, where $T$ is the total time length of a catalogue (or effectively we take $T_T = \infty$, or impose no time limit on interearthquake interaction). Then (7) becomes

$$\psi_t(\tau) = \theta \tau_{M_0}^{\theta - 1} \tau^{\theta - 1} - \theta, \quad \tau \geq t_m.$$

The non-normalized function $\psi_m(M)$ which corresponds to the number of dependent shocks generated by the main earthquake with seismic moment $M$, is

$$\psi_m(M) = \mu M^{1/3}.$$

In our new likelihood model the seismic moment of an ‘offspring’ earthquake does not depend on the seismic moment of the ‘parent’ event, so it might be even larger than the latter. In the standard interpretation of an earthquake sequence, the first earthquake will be considered as a foreshock of the second. The parameters $\mu$ and $\delta$ in our simulation model (Kagan & Knopoff 1981) have value 1.0. In that work we considered the forward simulation case in which we know exactly the temporal internal structure of the source. In the case of earthquake catalogues, we do not know this structure, and we are interested in the number of dependent events which occur after the coda of an earthquake has died out. At that time
most elementary sources which comprise the earthquake focus are 'healed' and no longer active. Thus, we need to estimate parameters $\mu$ and $\delta$ from catalogues of earthquakes.

### 2.3 Poisson model and information content of a catalogue

We compare the branching process of an earthquake occurrence (2) with the null-hypothesis of a completely random occurrence of earthquakes, or a Poisson temporal process of events. For this process, the log-likelihood is

$$ l_i^o = -\alpha T + \sum_{i=1}^{N} \ln (\alpha T) + C(t, M) $$

where $\alpha = \alpha_c$ is the temporal rate of occurrence of events with $M > M_c$ and the additive constant $C(t, M)$ depends on units of time and seismic moment measurement. The term $\alpha T$ is the total number of events in a catalogue. Since the last two terms in (12) are the same as in other likelihood functions, a modified Poisson log-likelihood is more convenient for our later use

$$ l_0 = -\alpha T + \sum_{i=1}^{N} \ln (\alpha T). $$

If we do not take into account the spatial distribution densities in (3) and (4), we obtain for the process described by (2)

$$ l_i = T_i + T_2 + T_3 + C(t, M) = -\lambda T - \mu \sum_{i=1}^{N} M_i^o \Psi(t_i, T_i) $$

$$ + \sum_{i=1}^{N} \ln \left[ \mu + \mu M_i^o \sum_{j=1}^{N} \psi_i(t_j) \right] + C(t, M) $$

$$ = -\lambda T - \mu \sum_{i=1}^{N} M_i^o \Psi(t_i, T_i) $$

$$ + \sum_{i=1}^{N} \ln [\lambda T + \mu M_i^o T S \sum_{j=1}^{N} \psi_i(t_j)] + C(t, M) - N \ln T $$

$$ = \hat{l}_i + C(t, M) - N \ln T, $$

where $T_1$, $T_2$ and $T_3$ are the first three terms of the log-likelihood, and

$$ \Psi(t_i, T_i) = \int_{t_i}^{T_i} \psi_i(t) \, dt, $$

where, in turn, $t_i$ corresponds to $t_M$ in (8) for the $i$th earthquake, and $T_i$ is equal either to $t_M$ in (9) or to $(T - t_i)$, whichever is smaller. Thus, in this model we exclude from the log-likelihood calculations the parts of time intervals outside the time limits of a catalogue. The term $\lambda T$ ($T_i$) corresponds to the total number of independent events (or sequences of earthquakes), and $T_2$ is an estimate of the total number of dependent events. It is easy to see that

$$ l_0 = \hat{l}_i, \quad \text{if } \mu = 0 \text{ and } \lambda = \alpha. $$

We can use (14) alone to estimate the values of the model parameters, but we are also interested in the difference

$$ l_i - l_0 = \hat{l}_i - l_0. $$

This difference is equal to the total information content of a catalogue.

### 2.4 Uniformly homogeneous spatial distribution model

Definition of the log-likelihood function for a spatial field is more complex, mostly because it is difficult to formulate a proper comparison model for spatial distribution of events. The simplest 'naive' solution would be to assume that all epicentres are distributed uniformly randomly over area $S$ spanned by a catalogue. Then we have $\psi_i(s) = S^{-1}$ for $\psi_i(s)$ in (3). Because the density $\psi_i(s)$ is normalized, the log-likelihood could then be written similarly to that of (14) with only $T_i$ and additive terms different from that of above:

$$ \sum_{i=1}^{N} \ln \left[ \lambda T / S + \mu M_i^o T \sum_{j=1}^{N} \psi_i(t_j) \psi_j(t_i) \right] + C(t, M) - N \ln T $$

$$ = \sum_{i=1}^{N} \ln \left[ \lambda T + \mu M_i^o T S \sum_{j=1}^{N} \psi_i(t_j) \psi_j(t_i) \right] $$

$$ + C(t, M) - N \ln (TS), $$

where we use the same notation as in (4) and (14). In the right part of the equation we multiplied the argument under logarithm in $T_i$ by $S$. This does not change the difference $l_{iS} - l_{0S}$ if we do the same operation with the Poisson comparison field:

$$ l_{iS} - l_{0S} = l_{iS} - l_0, $$

so we need to calculate here only $l_{iS}$ (the first three terms of the log-likelihood). Although the Poisson comparison (null-hypothesis) process is different in (12) and (18), the procedure for estimating the difference $l_i - l_0$ could be made the same, as soon as we normalize the arguments of the logarithm in $T_i$ in such a way that the first subterm of the argument is equal to $\lambda T$. The same procedure is applied in this section to log-likelihood calculations of other models below.

After considering seismicity maps, it is clear that polygons with widely different total areas can confine the same, or almost the same, number of earthquakes. For the CALNET (USGS) catalogue, for example, we can easily expand the area toward the Pacific Ocean, without including a significant number of new epicentres. As we discussed earlier (Kagan & Knopoff 1980a), this effect is due to the fractal (Mandelbrot 1983) nature of the earthquake hypocentre distribution over the space. If we compare the two terms under the logarithm in (17), we can see that whereas the first term is constant (equal to the total number of main or independent events), the second term depends on the value of $S$, i.e., it could be made (by enlarging the area spanned by a catalogue without including new earthquakes in the list) as large as possible. This means that the log-likelihood difference ($l_{iS} - l_{0S}$) is not bound by any value. Thus the 'naive' solution does not work.

### 2.5 Requirements for models of spatially inhomogeneous distribution of epicentres

How can the inhomogeneous Poisson field of earthquake epicentres on the surface of the Earth be defined? This field should satisfy the following requirements.
(a) Most important, the method should have a predictive value, i.e., it should not only describe a past seismicity pattern, but also anticipate such features as ‘filling up’ seismic gaps, earthquakes occurring on subsidiary faults, etc.

(b) It should not depend strongly on slight perturbation of the boundaries of the region.

(c) The method should give an asymptotically correct answer for all source distribution geometries we know, such as a uniform distribution of sources over a plane, or a planar figure; a line distribution of sources; a Cauchy flightstopover (Mandelbrot 1983) distribution, etc.

(d) The method should be simple, formal, algorithmic, hence accessible to computer simulation and estimation.

(e) Preferred orientations of earthquake faults or anisotropy of fault distribution should be taken into account.

(f) The method should work for geologically complex regions with many intersecting, branching, en echelon faults.

We discuss two possible solutions for the above problem.

(1) Projection of all epicentres on an appropriately chosen line.

(2) Use of a two-point epicentral moment to normalize the function $\phi_s(\rho)$ in (4).

2.6 Linear fault model

The advantage of the first method is simplicity. For the seismicity on the San Andreas fault (the CALNET catalogue), we take the projection line with the azimuth $-37^\circ$. If we define the $x$ axis along the fault trace, and the $y$ axis as an orthogonal to $x$, then the function $\psi_s(\rho)$ could be decomposed into the product of two functions

$$\psi_s(\rho) = \psi_s(x)\psi_s(y).$$

Similarly for the function $\phi_s(x)$ in (3) we obtain

$$\phi_s(x) = \phi_s(y) = \psi_s(y).$$

It means that the difference $(l - \rho)$ does not depend on the $y$ distribution density, so we need to specify only function $\psi_s(x)$

$$\psi_s(z) = (\sigma_\rho \sqrt{2\pi})^{-1} \exp(-z^2/(2\sigma_\rho^2)).$$

where $z = x_y - x_x$. For the distribution along the $z$ axis we specify similarly

$$\psi_s(z) = (\sigma_\rho \sqrt{2\pi})^{-1} \exp(-\xi^2/(2\sigma_\rho^2)).$$

where $\xi = x_z - x_z$. The other method for the Poisson field representation we use here goes back to the idea of the ‘random catalogue’ by Kellis-Borok, Podgaetskaya & Prozorov (1972), or to the ‘data based simulation test’ (see Kendall & Young 1984, p. 642). In this method we define the distribution density function $\phi_s(\rho)$ in (3) as a normalized second-order moment of epicentres:

$$\bar{\phi_s}(\rho) = \frac{N_2(\rho)}{N(N - 1)}.$$ (27)

where $N_2(\rho)$ is the number of pairs of epicentres separated by the distance in the interval $(\rho, \rho + \Delta \rho)$ (see more in Kagan & Knopoff 1980a).

For the horizontal distribution density $\psi_s(\rho)$ we take two variants: the distribution of epicentral errors can be approximated by a Rayleigh distribution,

$$\psi_s(\rho) = \frac{\rho}{\sigma_\rho} \exp(-\rho^2/(2\sigma_\rho^2)).$$

where $\sigma_\rho$ is an appropriate standard deviation. To emulate a scatter of aftershocks around mainshocks, we take a one-sided Gaussian function,

$$\psi_s(\rho) = 2(\sigma_\rho \sqrt{2\pi})^{-1} \exp(-\rho^2/(2\sigma_\rho^2)).$$

The standard deviation depends on the standard errors of hypocentre determination and on the seismic moment of the main event (cf. equation 22)

$$\sigma_\rho = (\varepsilon_\rho \sigma_x^2) \{M_s/M_e\}^{3/2};$$

here $\varepsilon_\rho$ is the standard error in epicentre determination and
3 RESULTS FOR EARTHQUAKE CATALOGUES

3.1 Earthquake catalogues

We apply the likelihood technique to several available earthquake catalogues. The catalogue where we can test most of the formulae of the preceding section is the CALNET catalogue (see Marks & Lester 1980 and references therein). The accuracy of the hypocentre determination for this catalogue is very high and reasonably uniform in time and space. Other catalogues are used mostly to obtain appropriate estimates of seismicity parameters. These catalogues are listed in Table 1. The HARVARD
3.2 Comparing new and old models

We start our analysis by comparing the previous statistical model as discussed in Kagan & Knopoff (1980b) with the new models formulated in the preceding section. We compare the least differing models first, so in the new model we keep the form of the interaction function as described by equation (33) (see rows 1 and 2 in Table 2). The depth distribution of hypocentres has not been taken into account (except for a subdivision of a catalogue into shallow, intermediate, and deep subcatalogues) in our earlier (Kagan & Knopoff 1980b) calculations. To shorten the discussion, in this section we will usually replace \( l \) by \( l_0 \).

In our new model the interaction between earthquakes of the same magnitude (magnitudes are rounded off to the closest decimal point in most catalogues) has been taken into account, whereas in the old one this interaction was neglected. This omission decreases the value of \( l \) by 3 per cent, a difference observed in this case too (Table 2). Although the second model has half the number of degrees of freedom, the quality of the fit is essentially the same in both.

In the above calculations the value of \( T_r \) in (9) has been set at \( 3.45 \times \sqrt{10} \) days for an earthquake with local magnitude 1.5. Next we studied the effect of enlarging this variable, so that \( T_r > T \) (total time interval of the catalogue), i.e., in effect excluding this parameter from the model. The value of log-likelihood for this model (row 3 in Table 2) has been found to be 6940, thus the extension of the upper time limit does not contribute significantly to the \( l \) value. To account for location errors in the CALNET catalogue, the formula for the distance boundaries is

\[
R_i = \sqrt{x^2 + (\rho_1 \times 3^{m-4.0})^2}^{1/2}
\]

where \( r_i = 0.35, 0.70, 1.1, 1.6 \text{~km}, \rho_1 = 0.50, 1.25, 2.5, 5.0 \text{~km}, \) and \( m \) is a magnitude of the main shock (see also Kagan & Knopoff 1980b). These distance intervals are used in (32) and (33).

We determined two parameters of (11) for model (33). An approximate estimate of the number of active sources might provide the theory of branching processes (Athreya &
Ney 1972, p. 20): in the beginning of a critical process the number of 'particles' (events) should increase linearly with the number of generations \( n \). It means that the number of active members of the population should be proportional to \( n \), whereas the total number of particles produced by the branching is proportional to \( n^2 \). Thus, the number of 'active' events is proportional to the square root of the total number of events generated so far.

The search for the maximum of the likelihood function yielded a value \( \delta \) close to 0.50, which is predicted by the above theoretical arguments. A similar expression has been used by Reasenberg (1985, equation 9) to approximate the number of aftershocks as a function of the magnitude of a main earthquake. If we recalculate Reasenberg's formula using a seismic moment, the value of the exponent is, depending on a conversion coefficient, between 0.4 and 0.6. Ogata (1988, table 3) obtained the value of the exponent which is equivalent to \( \delta = 0.47 \) using likelihood methods.

The equality of the \( \delta \) coefficients to 0.5 does not mean that the rate of aftershock (or foreshock) occurrence is proportional to \( M^\beta \) (see equation 11). In our parametrization, the rate is influenced by the value of \( t_M \) (see equations 7 to 10). Thus, if we take the value of the coefficient \( \theta = 0.5 \), we obtain the dependence of the conditional aftershock rate with the seismic moment \( M \):

\[
\psi(M)(t) \propto M^{2\beta t - 3/2}, \quad \text{for } t \geq t_M.
\]  

(36)

The value of \( \mu \) was found to be about 0.1, which means that towards the end of the earthquake coda, most (0.9) dependent shocks have already occurred (essentially they comprise the source function of the first or the main earthquake).

3.3 Estimating the number of dependent shocks

Using these values of parameters, we roughly estimate the number of dependent shocks in the CALNET catalogue. To avoid cumbersome equations we assume that \( M_s \gg M_p \). The number of events \( v \) generated by any earthquake is calculated as

\[
v = C^{-1} \int_{M_p} \phi_M(M) \psi_M(M) \, dM.
\]  

(37)

The normalizing coefficient \( C \) is

\[
C = M_p^{-\beta} \exp\left(-M_s/M_p\right)/\beta.
\]

The condition for a branching process to be finite is \( \nu < 1 \). After replacing \( \phi_M(M) \) by (1), and \( \psi_M(M) \) by (11), we obtain

\[
v = \frac{\mu \beta}{\nu} \frac{M_p^\beta - M_s^\beta}{\nu - \beta - 1}.
\]  

(38)

If we introduce the above mentioned values of parameters \( \beta, \delta \) and \( \mu \) into this equation, we obtain \( \nu = 0.6 \). The average total number of dependent events in a sequence \( (D) \), and the variance \( (V) \) of this number are equal respectively (Good 1949)

\[
D = \frac{\nu}{1 - \nu}, \quad V = \frac{\nu}{(1 - \nu)^2}.
\]  

(39)

Thus, we obtain \( D = 1.5, V = 9.4 \) for the above model. This means that for each main shock there are 1.5 \( \pm \) 3.1 dependent shocks in the catalogue. The relative standard error \( \nu \) (the coefficient of variability) is

\[
\nu = \frac{D}{\sqrt{V}} = \frac{\nu}{\sqrt{1 - \nu}}.
\]  

(40)

The minimum value of \( \nu \) is 2 (for \( \nu = 0.5 \)). Once again we can see the great variability of the earthquake process. Unfortunately, as we have found out, the estimates obtained through (39) are very unstable for different models of earthquake occurrence.

<table>
<thead>
<tr>
<th>No</th>
<th>Catalog Depth</th>
<th>Model’s parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>name</td>
<td>limits</td>
</tr>
<tr>
<td>1</td>
<td>CALNET</td>
<td>0-15</td>
</tr>
<tr>
<td>2</td>
<td>CALNET</td>
<td>0-15</td>
</tr>
<tr>
<td>3</td>
<td>PDE</td>
<td>0-70</td>
</tr>
<tr>
<td>4</td>
<td>PDE</td>
<td>0-35</td>
</tr>
<tr>
<td>5</td>
<td>PDE</td>
<td>36-70</td>
</tr>
<tr>
<td>6</td>
<td>PDE</td>
<td>71-280</td>
</tr>
<tr>
<td>7</td>
<td>PDE</td>
<td>71-140</td>
</tr>
<tr>
<td>8</td>
<td>PDE</td>
<td>141-280</td>
</tr>
<tr>
<td>9</td>
<td>PDE</td>
<td>281-700</td>
</tr>
<tr>
<td>10</td>
<td>HARVARD</td>
<td>0-70</td>
</tr>
<tr>
<td>11</td>
<td>DUDA</td>
<td>0-70</td>
</tr>
<tr>
<td>12</td>
<td>DUDA</td>
<td>0-70</td>
</tr>
</tbody>
</table>

variable has been constrained during optimisation:

* - from below;
# - from above.
3.4 θ estimation

We analyse the likelihood function, following in general the same sequence of models as in the previous section. While calculating (14) we encounter poor convergence of the likelihood optimization with regard to θ. The value of θ is found to be between 0.0 and 0.5 for shallow and intermediate earthquakes, but it is close to 1.0 for deep earthquakes (Tables 2 and 3). It means that fore-/aftershocks of deep earthquakes are on the average more concentrated around the main event than dependent events of shallow earthquakes are.

Ogata (1988) obtained the value θ = 0 as the result of the maximum likelihood analysis of Japanese earthquake sequences. This zero value of θ corresponds to the standard form of Omori’s law of aftershock frequency decay: Φ(t) = t^-1. However, the zero value of θ is unacceptable for the following two reasons.

(a) The integral of the distribution (10) diverges when time tends to infinity. It means that each earthquake has an infinite number of aftershocks. Therefore, the distribution has to be truncated (or have its form changed) for large time intervals. In effect, this means introducing an additional parameter in the model.

(b) The values of θ other than 1/2 generate incorrect numbers of dependent shocks in the stochastic model of shallow earthquake occurrence (Kagan & Knopoff 1981). In Kagan & Knopoff (1987a) we show that θ equals 1/2 as a consequence of natural assumptions on the stochastic behaviour of stress loads at the edges of an earthquake fault.

3.5 Testing spatial models of earthquake occurrence

The information content of the catalogue for the time–seismic moment model (14) is very low, reflecting loss of much information concerning the spatial dimension of an earthquake process. We can compare the results obtained by Vere-Jones (1978) and by Vere-Jones & Ozaki (1982) for the historic Japanese catalogue (South Kanto region). If we calculate the information per one earthquake for Vere-Jones’ (1978) analysis, the result is from 0.1 to 0.6 bit/eq (bits per earthquake), which is close to our values above. The information rate l bit/eq corresponds to a reduction of uncertainty in the seismicity level by a factor of 2.0^l. Similar comparison can be made with the results obtained by Ogata & Katsura (1986, pp. 300–302) for the earthquake catalogue of the Canberra (Australia) area. Their result is 0.7 bit/eq; the number of independent events is 0.84 of the total number of earthquakes in the catalogue (see Table 2).

As we mentioned in the Introduction, the above investigations do not take into account the magnitude of earthquakes. Ogata (1988, table 2) compared two sets of computations, with and without magnitude information. The addition of magnitudes to likelihood procedures increases information content of the catalogue by 0.1–0.2 bit/eq. When comparing these results with ours (see rows 4 and 5 in Table 2), it should be noted that the value of the likelihood function in the former cases strongly depends on the ratio of the region size to the size of the largest earthquake cluster in a catalogue. If this ratio is large, the l value is small, because the regional seismicity consists of largely independent groups of events.

We then tried model (17): a uniformly homogeneous Poisson distribution over the whole surface area, spanned by the CALNET catalogue. A conditional spatial distribution of dependent shocks (4) is set here to be the same as in model (31). The value of I/N (Table 2, row 6) is the highest of all the models, but as we discussed in Section 2.4 the values of the likelihood function are not constrained in this model by any reasonable limit; we should consider these values as artifacts.

Next we studied the influence of spatial moment (27) normalization on the value of the log-likelihood. In Table 2 we tabulate four cases with a different normalization. The first two (rows 7 and 8) are taken from the subcatalogue covering only 1977. In the first of these items the two-point spatial moment used for the normalization is that of the whole catalogue 1971–1977; in the latter case we have used the spatial moment of 1977 data only. As we have explained in Section 2.7 (see discussion around equation 34), this effect is expected. The determination of the spatial moment in a catalogue of small time duration tends to increase the value of the Poisson rate for small distance intervals; thus the value of l is lowered.

The above features are represented more dramatically in the next two cases (rows 9 and 10 in Table 2), where we tabulate the results of the optimization for the southern 1/5 part (S5) of the CALNET catalogue. Again the first row of values is for the normalization taken from the whole catalogue, and the latter results are for the normalization based on the S5 data only. Reasons for such a large difference in the two optimizations are as follows: the southern part of the area spanned by the catalogue is more densely 'populated' by epicentres than the whole area on the average; the values of the two-point spatial moment are about 2 to 4 times larger for S5 area than for the whole catalogue itself. We will return to this topic later in discussing Table 4.

In rows 11 to 14 of Table 2 calculations are made using equation (31) with horizontal scatter modelled by (29). In the first two rows we compare the inversions without (row 11) and with (row 12) use of depth data in the catalogue. Including depth increases the information content of the catalogue by about 16 per cent. Comparing rows 12 and 13 again shows that the influence of the time limits T on the information content is insignificant, whereas the θ value (rows 13 and 14) has a strong influence. Nevertheless, we see that the values of geometrical parameters of the earthquake occurrence (c, ε, and ε'), as well as the value of δ, are not influenced significantly by the change of θ and T.

The last four rows of Table 2 are for the ‘linear’ model of Section 2.6. We show again that the influence of the time limits (9) is not significant, at least for the value of parameter θ = 0.5. In the last row of Table 2 we added data from the California catalogue for years 1900–1971 to the CALNET catalogue. In this new catalogue we have also taken into account how past seismicity influenced the earthquakes in the CALNET catalogue. As we see from Table 2 those influences are not very large, the value of log-likelihood increasing only slightly.

We note that the value of ε is smaller than that of ε'.
which has been calculated for the model (31). The value $s_\alpha$ is similar to that obtained for equation (25). One possible explanation is that (29) is a better approximation for distribution of dependent shocks around a mainshock than for distribution of horizontal errors.

In Table 3 we tabulate values of seismicity parameters for several earthquake catalogues obtained by using likelihood function (31) with the Rayleigh distribution (28) for the horizontal scatter. In this case $e_\alpha$ for the CALNET catalogue is similar to that obtained using (25) in Table 2, but for $s_\alpha$ the values in Table 3 are lower than those in Table 2. The explanation is similar to that in the previous paragraph: the Rayleigh distribution is a reasonably good approximation for epicentral errors, but it might be inappropriate for an aftershock distribution. Values of other parameters for the CALNET catalogue are almost identical in Table 3 and in the appropriate entries of Table 2, demonstrating robustness of the inversion procedure.

For the PDE catalogue we produce several maximum likelihood inversions for earthquakes in different depth ranges. For shallow earthquakes seismicity seems to be distributed more or less uniformly through the upper 70 km of the lithosphere: if we subdivide it into two layers (rows 4 and 5), the sum of log-likelihoods is smaller than the log-likelihood for the whole 70 km. It may mean that some earthquake interactions cross from the upper part of the lithosphere into the lower part. It is quite possible, however, that since the depths of shallow hypocentres are often inaccurate, this interaction is an artifact caused by location errors.

Contrary to shallow events, for intermediate earthquakes, their subdivision increases the total value of the likelihood function (compare row 6 with rows 7 and 8). This means that most earthquake interaction is concentrated in the upper part of the intermediate layer, so if we take the layer’s boundaries as 71–280 km, we ‘dilute’ this interaction. The value of the log-likelihood for the 141–280 km layer is only slightly above the level which corresponds to a statistically significant non-Poisson process. This dependence of the log-likelihood on depth confirms previous results that intermediate earthquakes in the depth range of 141–280 km have the minimal number of aftershocks (Kagan & Knopoff 1980; Frolich 1987).

Deep earthquakes exhibit, on the other hand, a new increase of the likelihood function, indicating resumption of aftershock activity. This fact combined with the sudden change in the $\theta$ value (see Section 3.4) suggests that deep seismicity should be explained by a physical mechanism different from that for shallow and intermediate earthquakes. This conclusion is confirmed by inspecting deep dependent events in the catalogue. We calculate that the total number of dependent events for deep earthquakes is 6.3. However, more than 90 per cent of the total is due to six after-foreshock sequences in which dependent earthquakes occur closely in a wake of the causative event.

Shallow earthquakes in two other catalogues (HARVARD and DUDA) exhibit similar values of seismicity parameters. Lower values of the log-likelihood for the DUDA catalogue could be explained by its smaller magnitude range as well as larger location and, possibly, magnitude errors.

The absolute values of errors, both horizontal and vertical, for the CALNET catalogue correspond well to that of hypocentre inversions (Marks & Lester 1980, and references therein). We note that our determination of the standard errors involves relative uncertainties; see equations (22), (23), (24) and (30). The estimate of a focal zone size of an earthquake with local magnitude 4.0 ($s_\alpha$ in Table 2) agrees with other size estimates done by more traditional methods. Our goal in these investigations was to demonstrate general possibilities in studying geometrical parameters of the earthquake interaction using the likelihood approach. Therefore, the results tabulated in Tables 2 and 3 are to be considered preliminary. More thorough analysis of the size of aftershock zones and location errors should include investigating correlations between statistical estimates of these quantities, as well as application of prior available information.

### 3.6 Catalogue subdivision

Model (25) allows us to check easily enough the influence of a catalogue’s temporal and spatial limits on the value of $l$. In Table 4 we display the values of $l$ as well as the values of $\beta$ for the above-mentioned two subdivisions of the CALNET catalogue. The values of all other parameters of the model have been set equal to that of row 15 of Table 2. We see again a great variability of log-likelihood both in time and space. A closer consideration shows that the value of the information content depends strongly on the presence of large clusters of earthquakes, usually aftershocks of some strong event. The one possible indicator of a large aftershock sequence is the value of $m_\text{max}$ (see Table 4). Another indication of the presence of large earthquakes in a subcatalogue is the average seismic moment $\bar{M}$ which is

<table>
<thead>
<tr>
<th>Year, space intervals</th>
<th>N</th>
<th>$\lambda$</th>
<th>$b_\alpha$</th>
<th>$m_\text{max}$</th>
<th>$\bar{l}$</th>
<th>$1/N$</th>
<th>$\lambda/N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>784</td>
<td>612</td>
<td>0.57</td>
<td>3.9</td>
<td>29</td>
<td>773</td>
<td>1.42</td>
</tr>
<tr>
<td>1972</td>
<td>1511</td>
<td>875</td>
<td>0.56</td>
<td>5.0</td>
<td>172</td>
<td>2760</td>
<td>2.64</td>
</tr>
<tr>
<td>1973</td>
<td>1013</td>
<td>779</td>
<td>0.52</td>
<td>4.6</td>
<td>90</td>
<td>913</td>
<td>1.30</td>
</tr>
<tr>
<td>1974</td>
<td>946</td>
<td>711</td>
<td>0.58</td>
<td>5.2</td>
<td>246</td>
<td>877</td>
<td>2.74</td>
</tr>
<tr>
<td>1975</td>
<td>1004</td>
<td>804</td>
<td>0.62</td>
<td>4.9</td>
<td>127</td>
<td>584</td>
<td>1.54</td>
</tr>
<tr>
<td>1976</td>
<td>1061</td>
<td>860</td>
<td>0.65</td>
<td>4.3</td>
<td>25</td>
<td>791</td>
<td>1.08</td>
</tr>
<tr>
<td>1977</td>
<td>1041</td>
<td>761</td>
<td>0.61</td>
<td>4.3</td>
<td>44</td>
<td>1239</td>
<td>1.72</td>
</tr>
<tr>
<td>av.</td>
<td>1051</td>
<td>772</td>
<td>0.59</td>
<td></td>
<td>1.48</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>dev.</td>
<td>4223</td>
<td>490</td>
<td>$\pm0.04$</td>
<td></td>
<td>$\pm0.58$</td>
<td>$\pm0.08$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year, space intervals</th>
<th>N</th>
<th>$\lambda$</th>
<th>$b_\alpha$</th>
<th>$m_\text{max}$</th>
<th>$\bar{l}$</th>
<th>$1/N$</th>
<th>$\lambda/N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1971</td>
<td>784</td>
<td>612</td>
<td>0.57</td>
<td>3.9</td>
<td>29</td>
<td>773</td>
<td>1.42</td>
</tr>
<tr>
<td>1972</td>
<td>1511</td>
<td>875</td>
<td>0.56</td>
<td>5.0</td>
<td>172</td>
<td>2760</td>
<td>2.64</td>
</tr>
<tr>
<td>1973</td>
<td>1013</td>
<td>779</td>
<td>0.52</td>
<td>4.6</td>
<td>90</td>
<td>913</td>
<td>1.30</td>
</tr>
<tr>
<td>1974</td>
<td>946</td>
<td>711</td>
<td>0.58</td>
<td>5.2</td>
<td>246</td>
<td>877</td>
<td>2.74</td>
</tr>
<tr>
<td>1975</td>
<td>1004</td>
<td>804</td>
<td>0.62</td>
<td>4.9</td>
<td>127</td>
<td>584</td>
<td>1.54</td>
</tr>
<tr>
<td>1976</td>
<td>1061</td>
<td>860</td>
<td>0.65</td>
<td>4.3</td>
<td>25</td>
<td>791</td>
<td>1.08</td>
</tr>
<tr>
<td>1977</td>
<td>1041</td>
<td>761</td>
<td>0.61</td>
<td>4.3</td>
<td>44</td>
<td>1239</td>
<td>1.72</td>
</tr>
<tr>
<td>av.</td>
<td>1051</td>
<td>772</td>
<td>0.59</td>
<td></td>
<td>1.48</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>dev.</td>
<td>4223</td>
<td>490</td>
<td>$\pm0.04$</td>
<td></td>
<td>$\pm0.58$</td>
<td>$\pm0.08$</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.** Values of the likelihood function (l) for the CALNET catalogue of earthquakes and its time and space subdivisions. The values of parameters are taken from the full catalogue optimization (the last row in the table).
The result is an increase of information content per earthquake with the increased magnitude cut-off in the values of all model parameters, with the exception of CALNET catalogue. For the PDE catalogue the result of sequence, the value of during all other years in the catalogue (Table 5). Were we able to average the seismicity over the whole seismic cycle or several cycles (200 or 300 years), the value of the ratio would probably be significantly smaller. In 1972 during the Bear Valley earthquake sequence, the value of /N has been significantly lower than during all other years in the catalogue (Table 4).

In Table 5 we first repeated similar calculations; only this time the magnitude limit of the CALNET subcatalogues has been varied. Similarly to the calculations in Table 4, the values of all model parameters, with the exception of \( \beta \) and \( \lambda \), have been set equal to the values of row 15 in Table 2. The result is an increase of information content per earthquake with the increased magnitude cut-off in the CALNET catalogue. For the PDE catalogue the result of the magnitude cut-off change is opposite to that for the CALNET catalogue (Table 5). In general, we should expect decreased information content with a higher value of \( m_c \), since the catalogue should become more Poissonian. However, earthquakes in the CALNET catalogue are far from the maximum magnitude level. It is possible that for small earthquakes the interaction between events intensifies with their magnitude. Therefore, the information content of subcatalogues increases with a higher level of \( m_c \).

### 3.7 Time dependence of the information content

We study how the information content depends on the time lapsed since the end of an earthquake coda. We use two methods to investigate this dependence (Table 6). In the first method we introduce time delay between the end of the earthquake coda and starting time when the interaction between events is taken into account. This time delay is taken to be the same for all earthquakes in the catalogue. We use the parameter values obtained for the standard (zero time delay) optimization in all computations with non-zero time delays.

In Table 6 we see that the information content for a time delay of 1 min is actually higher than that of zero delay. This may be explained by incomplete reporting of weak aftershocks in the wake of a stronger event (Kagan & Knopoff 1980b). The information content decreases rather rapidly with increasing time delays: after 1 hr, one tenth of the predictive information is lost; 1 day later, the reduction is one third; and 10 days later it is one-half.

In the second set of calculations reported in Table 6, we change the value of \( t_c \) of (8), so that more earthquakes are removed from these catalogues. (We exclude every event which is closer than \( t_m \) to another earthquake from the catalogue.) These new catalogues are optimized independently of each other. The results are similar to the one reported above: the information content decreases drastically as soon as we have removed dependent events from the catalogue.

### 4 SIMULATION OF EARTHQUAKE CATALOGUES

The major goals of earthquake sequence and catalogue simulation have been to ascertain whether our inversion procedures described in Section 3 are unbiased and to

---

**Table 5.** Values of the likelihood function \( \lambda \) for catalogues for various magnitude cut-offs.

<table>
<thead>
<tr>
<th>Magnitude cutoff</th>
<th>CALNET catalog</th>
<th>PDE catalog</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>( \lambda )</td>
<td>( \lambda )</td>
</tr>
<tr>
<td>1.5</td>
<td>7360</td>
<td>5368</td>
</tr>
<tr>
<td>2.0</td>
<td>3029</td>
<td>2198</td>
</tr>
<tr>
<td>2.5</td>
<td>929</td>
<td>673</td>
</tr>
<tr>
<td>3.0</td>
<td>301</td>
<td>221</td>
</tr>
<tr>
<td>3.5</td>
<td>58</td>
<td>37</td>
</tr>
</tbody>
</table>

**Table 6.** Values of the information content \( I/N \) per earthquake for subdivisions of the CALNET catalogue.

<table>
<thead>
<tr>
<th>Time delays</th>
<th>0</th>
<th>1</th>
<th>10</th>
<th>1</th>
<th>10</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_c )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t_c )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t_c )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t_c )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( t_c )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
estimate the possible amount of information one can extract from earthquake catalogues like the CALNET. As we mentioned in Section 3, estimates of the information content are not stable with regard to temporal and spatial subdivisions of the catalogue. We explained earlier that we need the earthquake history of one or several seismic cycles in order to properly estimate this quantity: this is clearly impossible with the present data. One possibility is to determine information content from simulated sequences of earthquakes. We described previously simulation methods used to create realistic sequences of events (Kagan & Knopoff 1981; Kagan 1982). By creating these sequences and synthetic catalogues of earthquakes, using the values of parameters extracted from the real catalogues, we can then estimate the values of information content.

We have taken the value of parameter \( \phi_0 \) which controls a degree of spatial branching of earthquake faults (Kagan 1982), to be equal to \( 5 \times 10^{-6} \). This value seems to correspond to the four-point distribution of earthquake foci in the CALNET catalogue (Kagan 1982). The value of the criticality coefficient \( \kappa \), which controls the maximum size of simulated faults can be estimated from the value of \( M_s \) in (1):

\[
\kappa = 2.0 \sqrt{M_s / M_c}.
\]

The formula can be easily deduced from Vere-Jones' (1976, p. 721) considerations. In most of our calculations we take the value \( \kappa = 10^{-5} \) that roughly corresponds to the value inferred from Table 2 of Kagan (1991) (with \( M_c = 1.5 \)). For comparison purposes some other values of \( \kappa \) have been also used. Other parameters of the simulation follow: \( \theta = 0.5 \), \( t_c = 0.01 \) s, \( M_c = 2 \times 10^{11} \) N m, and \( x_c = 50 \) m, where \( t_c \) and \( x_c \) are characteristic rupture time and size of an earthquake with the seismic moment \( M_c \).

Synthetic catalogues are subjected to the same log-likelihood optimization as the catalogues of natural earthquakes, discussed in Section 3. As a rule, synthetic catalogues contained between a few hundred and a few thousand events, so the optimization procedure has not been as stable as in the case of the CALNET or PDE catalogues. Since for simulated catalogues both parameters of the simulation and of the inversion are to be investigated, the study is more difficult and cannot be considered complete and exhaustive. The major results of the synthetic catalogue analysis follow.

(1) The value of information content has been much larger than even the maximum obtained in Table 4 for possible subdivisions of the CALNET catalogue. In Table 7 several values of information content \( (I/N) \) for synthetic catalogues with different values of \( \kappa \) (see equation 41) are shown. For \( \kappa = 10^{-5} \) the value of \( I/N \) is about 15 bit/eq. If we decrease the value of \( \kappa \) to 0.5, the value of \( I/N \) is decreased to about 2 bit/eq. The other influencing factor is the average seismic moment \( \bar{M} \) which, as we learned during the simulation trials, can vary up to several orders of magnitude even for the same value of \( \kappa \). The value of the average seismic moment for the CALNET catalogue is 110. Even after the correction for \( \bar{M} \), the values of the information content seem to be too high for synthetic catalogues (see also discussion in Kagan & Knopoff 1987b).

(2) One possible explanation of why synthetic catalogues yield different results from the real catalogues in likelihood inversion is the presence of errors in real catalogues. These errors involve misidentification of events, origin time, location errors, etc. We simulated location errors in our synthetic catalogue to observe their influence on the inversion procedure. In Table 8 we collected the results for two sets of synthetic catalogues in which horizontal and vertical Gaussian location errors have been introduced (columns 'Simulation parameters'). We see that the inversion procedure recovers with a reasonable success the values of the standard deviations \( \sigma_x \) and \( \sigma_z \) for the location errors which have been introduced into catalogues. The value of the information content decreases by about 10–20 per cent due to the location errors. The value of \( s_x \) obtained for the case of no errors (the upper rows for both

### Table 7. Values of the information content \((I/N)\) per earthquake for synthetic catalogues of earthquakes.

<table>
<thead>
<tr>
<th>( \kappa ) value</th>
<th>( 10^{-5} )</th>
<th>( 10^{-4} )</th>
<th>( 10^{-3} )</th>
<th>( 5 \times 10^{-3} )</th>
<th>( 5 \times 10^{-2} )</th>
<th>( 5 \times 10^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I/N )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.2</td>
<td>14.3</td>
<td>15.7</td>
<td>12.0</td>
<td>6.0</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>( M )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35084</td>
<td>839</td>
<td>1208</td>
<td>197</td>
<td>24</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

### Table 8. Values of parameters for various synthetic catalogues.

<table>
<thead>
<tr>
<th>Simulation parameters</th>
<th>Inversion parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa )</td>
<td>( \epsilon_x )</td>
</tr>
<tr>
<td>( km )</td>
<td>( km )</td>
</tr>
<tr>
<td>1 ( 10^{-5} )</td>
<td>-</td>
</tr>
<tr>
<td>1 ( 10^{-5} )</td>
<td>0</td>
</tr>
<tr>
<td>1 ( 10^{-5} )</td>
<td>.3</td>
</tr>
<tr>
<td>1 ( 10^{-5} )</td>
<td>.6</td>
</tr>
<tr>
<td>2 ( 5 \times 10^{-3} )</td>
<td>-</td>
</tr>
<tr>
<td>2 ( 5 \times 10^{-3} )</td>
<td>0</td>
</tr>
<tr>
<td>2 ( 5 \times 10^{-3} )</td>
<td>.6</td>
</tr>
</tbody>
</table>

* * - means that the variable has been constrained during optimization.
catalogues) is reasonably close to that of Table 2 for the CALNET catalogue. When we introduced the errors, the optimization procedure would not always converge to any positive value of \( s_e \), so we had to constrain this variable (see Table 8).

(3) As we mentioned in Section 3, the parameters \( \mu \) and \( \delta \) have the value of 1.0 during simulations. In the simulation case we consider infinitesimal subevents which multiply and develop according to a critical branching process, so the above values cannot be transferred without adjustment to the case of earthquake catalogues, simulated or real (see Sections 2.2, 3.2 and 3.3). We can see from Table 8, that for synthetic catalogues we obtain values of parameters \( \mu \) and \( \delta \) reasonably close to those obtained for the CALNET catalogue. This serves as additional proof that our simulation and inversion procedures are correct and unbiased.

(4) The value of the parameter \( \theta \) which controls the 'memory' of the earthquake process (see equations 7 and 10) corresponds roughly (see Table 8) to the value used in the simulation (\( \theta = 0.5 \)). The variations of this parameter caused in Table 8 by the introduction of the location errors are relatively large. We noticed also that, in general, the values of this parameter obtained through the inversion are unstable with regard to slight perturbation of the model conditions and parameters. This may explain why in our likelihood search for the value of \( \theta \) we often have to constrain this parameter (see Section 3). The \( \theta \) estimates might have large standard deviations because we use relatively inefficient statistical procedures to evaluate parameters of stable distribution (7). Zolotarev (1980) discusses more appropriate statistical methods to estimate stable distribution parameters.

5 CONCLUSIONS

(1) A multidimensional model of earthquake occurrence based on the theory of stochastic branching processes has been developed and tested to obtain the parameters of earthquake occurrence from earthquake catalogues and to simulate the earthquake process.

(2) The stochastic model we have developed is fully quantitative; it does not require any expert parameter adjustment, arbitrary smoothing, aftershock removal, or drawing of boundaries to seismogenic regions. The model has a few adjustable parameters: five are needed to simulate earthquake sequences; three more are necessary to identify the existence of an earthquake in a manner consistent with modern seismographic network procedures.

(3) Most distributions controlling earthquake interaction have a power-law form, i.e., they are fractal or scale-invariant. These distributions include the Omori's law for time dependence of additional earthquake ruptures, and dependence of the number of aftershocks on the size of the mainshock, which are analysed in this study. Our other investigations suggest that the scale-invariance underlies the seismic moment-frequency relation (Kagan 1991) and the spatial distribution of hypocentres (Kagan & Knopoff 1980a).

(4) We investigate the importance of including different earthquake parameters into the model: hypocentral depth, time limits for interearthquake interaction, and especially the modelling of spatial patterns for mainshocks and aftershocks. We find that whereas the first two parameters do not influence strongly the likelihood function, appropriate models for the spatial distribution of earthquakes are vital for the likelihood calculation.

(5) The number of dependent events first decreases steadily as the depth increases; for a depth range 141–280 km the number of aftershocks reaches the minimum. However, for deep events the number of dependent events increases again. Combined with a sudden change in the \( \theta \) value, this suggests that temporal properties of deep seismicity could be explained by a physical mechanism unlike those for shallow and intermediate earthquakes. Unfortunately, our results do not allow to differentiate shallow and intermediate earthquakes with regard to their temporal properties.

(6) The results obtained here may be easily used for extrapolating past seismicity in the future, using standard methods of stochastic processes. As a result of such prediction the average uncertainty of future earthquake occurrence can be reduced, using available earthquake catalogues, by about 10 bits of information per earthquake, i.e., by a factor up to 1000 (2^{10}).

ACKNOWLEDGMENTS

This research was supported by Grants of the National Science Foundation CEE-82-14203, CEE-84-07553 and EAR 88-04883. I am grateful to L. Knopoff, and D. D. Jackson of UCLA for their very valuable comments. I thank Ritsuko Matsu'ura of the Tokyo University for her help in compiling some of the computer programs used in this work. I would also like to thank W. Smythe, R. Mehlman, and J. Piotrowski of JPL for giving me access to VAX computers during their idle time. I thank an anonymous referee for useful remarks which helped to clarify the paper. Publication No. 3451, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024.

REFERENCES


Kagan, Y. Y. & Knopoff, L., 1980a. Spatial distribution of


