STOCHASTIC SYNTHESIS OF EARTHQUAKE CATALOGS

Y. Y. Kagan and L. Knopoff

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

Abstract. A model of earthquake occurrence is proposed that is based on results of statistical studies of earthquake catalogs. We assume that each earthquake generates additional shocks with a probability that depends on time as $t^{-1.07}$. This assumption together with one regarding the independence of branching events on adjacent branches of the event 'tree,' is sufficient to permit the generation of complete catalogs of earthquakes that have the same time-magnitude statistical properties as real earthquake catalogs. If $\theta$ is about 0.5, the process generates sequences that have statistical properties similar to those for shallow earthquakes: many well-known relations are reproduced including the magnitude-frequency law, Omori's law of the rate of aftershock and foreshock occurrence, the duration of a recorded seismic event versus its magnitude, the self-similarity or lack of scale of rate of earthquake occurrence in different magnitude ranges, etc. A value of $\theta$ closer to 0.8 or 0.9 seems to simulate the statistical properties of the occurrence of intermediate and deep shocks. A formula for seismic risk prediction is proposed, and the implications of the model for risk evaluation are outlined. The possibilities of the determination of long-term risk from real or synthetic catalogs that have the property of self-similarity are discussed.

1. Introduction

Most theoretical solutions to problems in fracture dynamics yield relatively simple source-time functions: the velocity of slip functions generally rise monotonically and then decrease monotonically until all motion eventually ceases. On the other hand, there is some circumstantial evidence that real earthquake source-time functions are not simple functions of time; the velocity of slip functions may be rather more serrated, which would indicate that if the motion due to a single event is simple, the source must be a superposition of a number of rupture events. In this paper, we make this postulate and derive source-time functions based on it.

We have noted elsewhere [Kagan and Knopoff, 1978, 1980b] that Omori's law of the rate of occurrence of aftershocks seems to be a universal phenomenon. This law, which is that the rate of occurrence of aftershocks varies as $1/t$, has been verified for catalogs of the largest earthquakes on a worldwide scale as well as for catalogs on smaller regional scales, at least when the available catalogs are reasonably well documented. Extended back to the origin time of an earthquake, the rate of occurrence literally approaches infinity. We shall make the assumption that this rate of occurrence relation holds even at such a short time scale that the interval between events is less than the duration of the relative faulting motions of an individual event.

In this paper we make several simplifying assumptions. First, we shall assume that the source is a point in space. Second, we assume that the individual shocks in a complex source event are discrete in such a way that each component event of the complex sequence contributes a waveform to the motion having the same time-dependence as all the others. Third, we assume that the source-time function is a simple step function of displacement; this may be wrong in detail, but it is sufficient for the purposes of simulating real catalogs, as will be seen below.

By virtue of the time scale independence of power law functions we may imagine that the source-time function of a single earthquake, composed of a superposition of overlapping individual events, is identical to that of the history of deformation of the earth due to repeated earthquakes over years or even centuries under a suitable change of time scale. The best statistical information we have today relates to earthquake occurrences over intervals of years. In this paper we assume that we can extend the known statistical properties of earthquake catalogs to the time scale of earthquake durations at the source. We use a model which employs a minimum number of parameters. We will show that this model provides an adequate source-time function that simulates well the statistics of individual earthquake events and simulates the relationships between earthquakes in real catalogs as well. We simulate individual earthquakes by a Monte Carlo procedure, obtain synthetic catalogs of earthquakes, investigate statistical properties of these catalogs, and compare the results with the corresponding properties of real earthquake catalogs.

2. Stochastic Self-Similarity of the Earthquake Process

Before embarking on a description of the properties of earthquake source-time motions synthesized according to models incorporating some of the ideas sketched above, we review some of our recent results relating to the stochastic interaction of earthquakes. The most important feature which emerges as a result of these investigations [Kagan and Knopoff, 1977, 1978, 1980a, b], is the stochastic self-similarity of the seismic process, i.e., the absence of any particular scale connected with time-distance-magnitude patterns of earthquake occurrence. This self-similarity manifests itself in the appearance of power law distributions of all the features of the earthquake process investigated thus far: the power law distribution of seismic energies or moments, Omori's power law for the rate of occurrence of aftershocks and foreshocks...
of shallow earthquakes, the power law distribution of the energies of foreshocks and aftershocks, and the inverse power law dependence of the spatial moment on the separation between two foci. Some generalizations of the ideas of stochastic self-similarity are presented by Mandelbrot [1977].

We have found [Kagan and Knopoff, 1978, 1980b] that self-similarity holds over a magnitude range extending from \( M = 1.5 \) to the greatest earthquakes. Because of the absence of dimensional scaling parameters we make the assumption that the rate of occurrence of foreshocks of small earthquakes and small aftershocks of small earthquakes is similar to that for large events with only the time and distance scales changed appropriately. In this paper, we only use the assumption with regard to the time scale.

The time-magnitude rate of occurrence of statistically dependent earthquakes is shown schematically in Figure 1. The upper figure shows the time rate of occurrence for a given magnitude difference between the events. It is symmetric for the case in which the magnitude difference between a foreshock and its succeeding larger event and a larger event and its succeeding smaller aftershock is virtually zero. The asymmetry of the time behavior of the seismic process, that is, that there are few if any foreshocks and an abundance of aftershocks, is explained by their different magnitude distributions (see lower part of Figure 1). The number of aftershocks increases as \( \Delta M \) increases, whereas the number of foreshocks decreases with increasing \( \Delta M \).

As we have indicated, the large rate of occurrence of dependent events near the origin time of a main earthquake means that we assume that we can model each earthquake as a merger of many shocks [Kagan and Knopoff, 1978]. This multishock aspect of earthquake occurrence is illustrated schematically in Figure 2, where the single event in the left-hand portion of the diagram is replaced by several shocks whose origin times occur according to the statistical laws illustrated in Figure 1. There are some indications in real earthquakes for patterns of multishock occurrence [Trifunac and Hudson, 1973].

We assume that there is nesting of shocks: referring to Figure 2, perhaps each of the smaller steps in the right-hand part of the diagram should be replaced by an even smaller set of replicas of the whole process occurring on a shorter time scale. If our hypothesis is correct, then a real record shows a somewhat smoothed picture of the superposition of traces from many shocks having a dense distribution in time.

3. Stochastic Model of Earthquake Process

We describe a stochastic model of an earthquake process whose properties depend on only two significant intrinsic parameters. The model is designed to generate a simulated seismicity that is statistically equivalent to the time-magnitude properties of real earthquakes. Although there is no intrinsic size of earthquakes in our model, for convenience a unit earthquake with seismic moment \( m_0 \) is introduced on an ad hoc basis. Independent of the previous history of an event, it can act as a generator of other events of the same size according to a random process. We bypass the computational difficulties associated with the infinity in the occurrence rate of aftershocks as the time interval becomes shorter and shorter by the following device. In our model the probability that one event gives 'birth' to another in the time interval \( dt \) is

\[
\phi(t)dt = 0 \quad t < t_0
\]

\[
\phi(t)dt = (1-\kappa)8t_0^{\theta}e^{{-(1+\theta)}t}dt \quad t \geq t_0 \quad (1)
\]

where \( t_0 \) can be imagined to correspond to the rupture time of an elementary earthquake; during this time the parent earthquake cannot produce other shocks. The coefficient \( \theta \) specifies the 'fading memory' [Truesdell and Noll, 1965, p. 101] of the earthquake, and the coefficient \( \kappa \) defines the criticality of the process, which we discuss below. Secondary shocks are assumed to occur independently of each other; these have seismic moments equal to those of all the other shocks, including the 'main' shock. In their turn, secondary shocks produce new dependent shocks according to the same law (1); the process cascades indefinitely. Since all elementary shocks are equal in this model, we do not have to apply the
weighting factors of the lower part of Figure 1. Instead we shall show in section 7 that the dependence of the rate of occurrence of fore- shocks and aftershocks on magnitude difference arises as a consequence of the model we have assumed. Since the model is self-similar or scale-independent, the size of each shock m can be taken to be as small as necessary if t o is diminished correspondingly (see below).

The resulting process is known in the mathematical literature as an age-dependent, continuous state, branching process [Jirina, 1958; Harris, 1963].

We have normalized the integral of the memory function (1) over all time to be 1 - k. If k = 0, then the 'aftershocks' are said to be generated by a critical branching process; this means that each elementary event gives birth to one dependent or secondary event on the average. If k > 0 the process is subcritical [Harris, 1963]. By virtue of the independence of all offspring the number of such offspring of any event will be a Poisson variable with mean (1 - k). The probability that an event will have only one offspring is that while most Monte Carlo simulations of the process may give a finite number of events, the total expected number of events in such a process is infinite. Thus the simulations of the process are highly unstable and can lead to runaway cascading of seismicity. On the other hand, a subcritical process always has a finite total number of events. We call the first shock a 'first generation' event, the events produced by this shock as second generation events, etc. in agreement with the literature on branching processes [Harris, 1963]. We will have recourse to the index n of nth generation events.

In its discrete approximation with nonzero values of m o and t o the model can also be regarded as a special case of the self-exciting point process proposed by Hawkes [Hawkes and Adamopoulois, 1973, and references therein]. The new features introduced in our model are the self-similarity of the occurrence of earthquakes (equation 1) and the continuous state character of the process. These features, taken together with the critical behavior of the branching process, will enable us to explain in a consistent manner both the magnitude-size and time distributions of earthquakes from a set of simple assumptions. The importance of criticality was brought to our attention by Vere-Jones' [1976] earthquake branching model, which can be considered as an immediate forerunner of our model. The new ingredients introduced in our model are the time dependence of the rate of occurrence of aftershocks, and the additive character of the seismic moment of a series of small events that overlap in time. Vere-Jones has attributed the additive property to the fracture lengths of the individual shocks; in his model, time was not considered as a dimension of the earthquake process.

4. Synthetic Seismograms

The model lends itself to Monte Carlo simulations by virtue of the digitization of the process through the use of the finite lower bound values of m and t. To facilitate the discussion, we identify log 10 m with the threshold magnitude of homogeneity of an earthquake cata-

log. We take the characteristic time t o to be proportional to m 1/2 in consideration of a suggested relationship between magnitudes and earthquake rupture times [Kagan and Knopoff, 1978, Figure 3]. For example, if we set log 10 m o = 5 for an elementary shock and t o = 1.0 s, our simulated result can be compared with the NOAA catalog [Kagan and Knopoff, 1980a]; on the other hand, if we set log 10 m o = 1, and t o = 0.01 s, we match the results to the USGS catalog [Kagan and Knopoff, 1978].

Examples of such simulations are shown in the upper parts of Figures 3a and 3b; these two functions are similar in appearance, since they have been produced by the same random number sequence.

The graphs of cumulative seismic moment clearly show patterns of clustering of different orders (cf. Mandelbrot, 1977, p. 100). The source-time function appears to consist of several steps, each of which starts with a sharp onset and usually becomes more gradual after some time. This mode of behavior is to be expected in our model because the cumulative number of events in a critical branching process should increase at the beginning of the sequence [Harris, 1963]. The time intervals between any 'parent' event and its aftershock are then calculated using (1). Finally, the resulting event sequence is smoothed in time. Each subevent contributes a step function of moment m to the source-time functions such as those shown in Figures 3a and 3b; these two functions are similar in appearance, since they have been produced by the same random number sequence.

The source-time function in the upper part of the plot can be regarded as a sequence of events occurring over several decades, over several minutes, or over several hours, as in the figure. The synthesized source-time functions are not directly observable quantities. To compare these results with real seismograms, we must convolve the time derivative of these functions with the Green's function of ground displacement corresponding to an elementary dislocation. The resulting process is no longer self-similar because of the introduction of two characteristic times: the first is connected with the travel times of body waves and instrumental response, while the second is associated with the dispersion of surface waves and scattering processes that are manifested in the coda.

We have convolved our complex source histories with a selected theoretical seismogram; for this purpose we have used the time scale for the source function appropriate to the NOAA catalog, i.e., log t o = 5, t o = 1 s. The theoretical seismograms for an elemental step-function, displayed in the middle picture of Figures 3a and
Fig. 3. Simulated source-time functions and seismograms for shallow (a) and intermediate (b) earthquake sources. The upper trace in each plot is a synthetic source-time function. The middle plot is a theoretical seismogram described in the text, and the lower trace is a convolution of the derivative of source-time function with the theoretical seismogram.

3b, were calculated by a program from Liao et al. [1978]. The epicentral distance in both cases is close to 7000 km; the focal depths are 0 and 140 km, respectively; the component displayed is SH. The cross section is a continental structure for the shallow source and more-or-less typical oceanic structure for the source at intermediate depth. The first case (Figure 3a) corresponds to a strike slip dislocation on a vertical fault plane; the second case is that for dip slip motion, also on a vertical fault plane. The seismograms are the result of passing the ground motion through a filter equivalent to a standard 15–100 s World–Wide Standard Seismographic Network (WWSSN) seismograph. The convolution in our case is computationally simple because the kernel is convolved with a comb of delta functions.

The traces in the lower parts of Figures 3a and 3b are the results of the above computation and show the complex form not atypical of recordings of strong earthquakes and differ only in one major regard. We have not introduced codas; scattering by random inhomogeneities is not taken into account in the theoretical seismograms.

In principle, these seismograms, or others generated by similar methods, could be tried out on a practitioner of the compilation of earthquake catalogs from long-period teleseismic recordings to see which discrete events, plus
associated magnitudes, will be identified. This is a formidable task and will not be undertaken at this time.

5. Simulation of Synthetic Catalogs of Earthquakes

We bypass the problem of convolution by constructing a simplified version of the seismogram for the specific purposes of this study, which are catalog simulation and analysis. For the purposes of event identification a simple envelope drawn around the highly oscillatory kernel seismogram probably suffices. Here much of the problem of identification lies in the effective description of coda waves, since the persistent coda of strong earthquakes may mask the largest amplitude arrivals due to later, smaller earthquakes. Evernden [1976] has suggested an exponential form for this envelope. If we take this point of view, then we model a complex source as yielding a seismogram that is a superposition of exponentials of different initial heights and different occurrence times; the use of the exponential envelope function greatly reduces computation cost of simulations. The envelope function is the same as an equivalent source-time function, which is a convolution of the stepwise source-time function considered above with a decaying exponential. By using the equivalent source-time function we ignore the effects of change in the length of the coda with epicentral distance.

An example of an equivalent source-time function constructed as a sum of exponentials is shown in Figure 4, where the shallow source-time function of Figure 3a is convolved with an exponential decay function with a time constant $\beta = 400$ s; this value of $\beta$ corresponds roughly to that suggested for the envelope of the coda of distant shallow earthquakes [Evernden, 1976]. We can match the duration magnitude curve for individual earthquakes in the USGS catalog [Lee et al., 1972] in the magnitude range $0 < M_L < 1$ by using an exponential function with decay time near 3.0 s. In the frequency and distance range represented by the USGS network in California we are evidently using a part of the seismograms other than the surface wave coda to generate a kernel seismogram, and consequently, a smaller value of $\beta$ is called for. The values of the ratio $\beta/t_o$ for the two cases are roughly the same. As we will see below, the dimensionless quantity $\beta/t_o$ defines the values of the parameters used in the statistical analysis of earthquake catalogs. The fact that this ratio is approximately the same for both catalogs is consistent with the similarity of the values of the parameters that describe the aftershock sequences [Kagan and Knopoff, 1980b]. The similarity of the values of the parameters found in the statistical analysis of these earthquake catalogs seems to imply that the properties of these seismographic networks and the properties of the scattering of seismic waves in worldwide versus local seismic wave propagation are scaled in such a way as to ensure the similarity of $\beta/t_o$.

We define individual earthquakes from the equivalent source-time function by the following formal rule. We introduce a cutoff threshold, which is the horizontal line in Figure 4; for computational purposes we have taken this threshold $M_o$ to be close to $\log_{10}M_o - 0.5$. Each time the trace reaches this threshold, an earthquake is presumed either to have started or stopped and thus defines the duration of an earthquake. The number of elementary events in this time interval is taken to be proportional to the seismic moment $M_o$ of the earthquake.

We define a quantity $M_B$ that is the logarithm of the peak of the source-time moment function convolved with decaying exponential for each burst (Figure 4). This quantity, which we call the $B$ magnitude of an earthquake, evidently depends strongly on the value of $\beta$: if $\beta$ were infinite, as in the upper curve of Figure 4, $M_B$ will be equal to the logarithm of the total seismic moment of a simulated earthquake sequence. We have introduced this quantity by obvious analogy with conventional methods of measuring magnitudes. Our measure does not take instrumental response into account and hence is not directly comparable with conventional magnitudes. Nevertheless, the prospect of using this measure of the size of an earthquake is too tempting to bypass; below we shall show some similarities between $B$ magnitudes and conventional magnitudes. In this paper, most of our examples are for the case $\beta = 400$ s.

The replacement of the theoretical seismogram by the exponential envelope function has certain drawbacks. At instrumental periods less than the interval between a pair of subevents the waves in the coda add incoherently, and the sum will be proportional to the square root of the number of the subevents; hence our equivalent source-time function will no longer have a simple additive property. The impact of our failure to take this effect into account on the statistics of synthetic earthquakes is twofold: First, the incoherent addition of codas will tend to reduce the overall decay time of a burst of events; if
the decay rate for the equivalent source-time function is shorter than the one we have used, some small earthquakes formerly reported as a part of a larger multiple shock event, will be reported as independent events; this increases the b value. Second, the moment-duration relation will be skewed to shorter durations for a given moment; while the duration may be changed, the magnitude, however, will not be much affected, since the maximum of an equivalent source-time function most probably corresponds to the densest concentration of events in a cluster: the addition of these codas is more likely to be coherent because of the short time interval between these events.

The critical branching process described by (1) with \( \kappa = 0 \) cannot be used for numerical simulations because of the instability of this case. For some simulations of the critical process, very large numbers of events are produced; because dependent shocks are distributed in time according to the power law of (1), these events are usually concentrated in a short time interval and correspond to the occurrence of an unacceptably strong earthquake. To avoid problems connected with instabilities, stabilization was introduced into the simulation by the use of a slightly subcritical process (section 3) with \( \kappa = 0.01 \). To make up for the deficiency due to normalization, we introduced into our simulation a number of independent occurring elementary events which can be interpreted, for example, as representing aftershocks of earthquakes occurring before the start of the catalog. The use of a subcritical instead of a critical process also means, in effect, that a bend in the log moment-frequency curve will have been introduced at the large moment end of the curve. The bend can be parameterized by the introduction of a maximum magnitude into the magnitude-frequency relation [Knopoff and Kagan, 1977]. In our case, the subcritical branching process yields a \( b \)-magnitude-frequency distribution function that is almost linear for small magnitudes and has an increasing slope with increasing magnitude [Vere-Jones, 1976] instead of an abrupt truncation at the value of the maximum magnitude.

In the simulations we used 300 independent shocks to 'restart' the sequence due to the termination by the subcritical nature of the process; the total number of elementary events generated was usually slightly less than 100,000. The length of the simulated catalog was \( 10^{10} \). We have only used events with magnitudes greater than \( M_{CO} = \log_{10} m_0 + 0.5 \); this effectively removes 'quantization noise' arising from the use of elementary shocks of finite size. The parameters we have used to describe the occurrence of earthquakes [Kagan and Knopoff, 1980b] are largely insensitive to variations of either \( M_{CO} \) or \( \kappa \), which, in the latter case, is equivalent to maximum magnitude, as noted.

An example of a simulated catalog is shown in Figure 5a; for comparison, Figure 5b shows a portion of the USGS catalog for 1976. Both catalogs appear to be similar, although USGS has seemingly more events which do not belong to an easily identifiable cluster of earthquakes. We may suppose that these shocks are connected statistically with earthquakes that occurred before the start of the section of the catalog that is displayed. In the simulated catalogs, as indicated above, these 'orphan' shocks were modeled as independent initial events. In the mathematical literature these events are called 'immigrants' [Harris, 1963].

6. First-Order Analysis of Synthetic Catalogs

To analyze the synthetic catalogs, as with real catalogs, we first draw the \( b \)-magnitude-frequency distribution, which is the first-order moment of the process. The \( b \)-magnitude-frequency plots for several values of \( \beta \) are shown in Figure 6, where we have plotted \( M_0 + \log_{10} N \) as the ordinate to display more clearly small differences of the \( b \) values from unity, if there are any [Kagan and Knopoff, 1980b, Figure 2]; in this plot a horizontal line corresponds to the value \( b = 1 \). For \( \beta = 400 \) s we get \( b \) close to 0.5, which is a value close to that for the distribution of seismic moments or energies in real catalogs.

The result \( b = 0.5 \) for \( \beta/t \) large is not unexpected for these synthetic sequences, since if, for example, \( \beta \) is close to infinity our 'earthquake' would include almost all of the events in a critical branching process. The dis-
The b values for synthetic catalogs with \( \beta = 0.5 \) and those with \( \beta = 0.8 \) and 0.9 (Table 1) are almost the same as for those for real catalogs of shallow and deep earthquakes, respectively [Kagan and Knopoff, 1980b]; the b values for deep shocks are slightly smaller of the two. If we associate a higher \( \beta \) value with deep shocks, we see that a larger fraction of a deep earthquake sequence occurs in the same time interval in comparison with shallow earthquake sequences (compare Figures 3a and 3b).

Thus the same reasoning that yields the dependence of b on \( \beta \) applies in this case as well: the value of b for deep shocks should be closer to 0.5 for our simulated catalogs.

Because of the use of an equivalent source-time function that focuses on the coda as the feature that distinguishes one earthquake from another we have effectively removed any possibility of discussing ordinary magnitudes which are strongly dependent on instrumental responses at periods much shorter than \( \beta = 400 \) s. If we wish to describe ordinary seismic magnitudes, we have to gain access to it through the coda. The dependence of the duration D of the earthquake record on the \( \beta \) magnitude is similar to that for shallow earthquakes [Kagan and Knopoff, 1978, Figure 3] 

\[
\log_{10} D = 0.5 M + \text{const} \quad (2)
\]

if the time rate of occurrence of dependent shocks varies as \( t^{-1.5} \), which is appropriate for shallow shocks. We offer the conjecture that \( \beta \) magnitudes are simply related to ordinary magnitudes. As \( \beta \) increases to 0.8 or 0.9, the value of the exponent in (2) decreases to 0.3; thus our model predicts that deep and intermediate earthquakes should have shorter wave trains than shallow earthquakes if our conjecture is correct.

Another feature that bears on our proposal that \( \beta \) magnitudes and surface wave magnitudes are similar, arises from comparison of the seismic moment \( \beta \) magnitude curve for synthetic catalogs and the moment-magnitude curve for real catalogs (Figure 7). In the synthetic case we have used \( \log_{10} M_o = 3 \) or \( M_o = 5.5 \) to correspond to our estimates for the NOAA catalog. The reference curve for the case of real catalogs is that proposed by Aki [1972]. The similarity is strong.

**TABLE 1. Values of Parameters for Simulated Catalogs**

<table>
<thead>
<tr>
<th>Simulation Numbers</th>
<th>Simulation Numbers</th>
<th>Simulation Numbers</th>
<th>Simulation Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>( b_0 )</td>
<td>( \gamma_0 )</td>
<td>( a_0 )</td>
</tr>
<tr>
<td>0.5</td>
<td>0.622+0.067</td>
<td>1.154</td>
<td>22.12</td>
</tr>
<tr>
<td>0.5</td>
<td>0.634+0.057</td>
<td>1.157</td>
<td>30.91</td>
</tr>
<tr>
<td>0.5</td>
<td>0.683+0.075</td>
<td>1.70</td>
<td>18.93</td>
</tr>
<tr>
<td>0.5</td>
<td>0.642+0.065</td>
<td>1.159</td>
<td>23.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.589+0.089</td>
<td>1.145</td>
<td>12.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.602+0.0</td>
<td>1.149</td>
<td>11.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.634+0.057</td>
<td>1.157</td>
<td>22.12</td>
</tr>
<tr>
<td>0.5</td>
<td>0.683+0.075</td>
<td>1.70</td>
<td>18.93</td>
</tr>
<tr>
<td>0.5</td>
<td>0.642+0.065</td>
<td>1.159</td>
<td>23.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.589+0.089</td>
<td>1.145</td>
<td>12.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.602+0.0</td>
<td>1.149</td>
<td>11.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.634+0.057</td>
<td>1.157</td>
<td>22.12</td>
</tr>
<tr>
<td>0.5</td>
<td>0.683+0.075</td>
<td>1.70</td>
<td>18.93</td>
</tr>
<tr>
<td>0.5</td>
<td>0.642+0.065</td>
<td>1.159</td>
<td>23.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.589+0.089</td>
<td>1.145</td>
<td>12.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.602+0.0</td>
<td>1.149</td>
<td>11.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.634+0.057</td>
<td>1.157</td>
<td>22.12</td>
</tr>
<tr>
<td>0.5</td>
<td>0.683+0.075</td>
<td>1.70</td>
<td>18.93</td>
</tr>
<tr>
<td>0.5</td>
<td>0.642+0.065</td>
<td>1.159</td>
<td>23.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.589+0.089</td>
<td>1.145</td>
<td>12.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.602+0.0</td>
<td>1.149</td>
<td>11.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.634+0.057</td>
<td>1.157</td>
<td>22.12</td>
</tr>
<tr>
<td>0.5</td>
<td>0.683+0.075</td>
<td>1.70</td>
<td>18.93</td>
</tr>
<tr>
<td>0.5</td>
<td>0.642+0.065</td>
<td>1.159</td>
<td>23.89</td>
</tr>
<tr>
<td>0.5</td>
<td>0.589+0.089</td>
<td>1.145</td>
<td>12.24</td>
</tr>
<tr>
<td>0.5</td>
<td>0.602+0.0</td>
<td>1.149</td>
<td>11.89</td>
</tr>
</tbody>
</table>

**Fig. 6.** Cumulative \( \beta \)-magnitude-frequency relations for synthetic catalogs produced using different values of the time constant \( \beta \).**
that both foreshocks and aftershocks are a manifestation of essentially the same process, namely, the stochastic interaction of earthquakes. Simply stated, if the second earthquake of a pair of interrelated shocks happens to be larger than the first, we call the first member of the pair a foreshock. From the values of $v_+$ in Table 1 the number of cases for which the above possibility is encountered in simulated catalogs is in good agreement with the corresponding number for real catalogs.

Although all events were produced by the same stochastic mechanism, foreshocks and aftershocks seem to have different magnitude distributions as reflected in the values of $c_-$ and $c_+$, respectively: the number of dependent events in the magnitude interval $dM_B$ is proportional to $c_- dM_B$. The values of $c_+$ in real catalogs [Kagan and Knopoff, 1980b] are approximately similar to the values for synthetic catalogs, although the values of $c_-$ in the simulated catalogs are lower than their real counterparts.

The greatest difference between the second-order properties of simulated and real catalogs relate to the values of the branching rate $\mu$ for short-time aftershocks. The values of $\mu$ in aftershock time intervals 1 and 2 are significantly higher for the simulated catalogs than the corresponding estimates for real earthquakes [for definition of time intervals see Kagan and Knopoff 1978, 1980b]. This discrepancy may be connected with the incomplete identification of weak shocks in the wake of strong main events in real earthquake catalogs. In our maximum likelihood optimizations we have excluded earthquakes which occurred a short time after the main shocks from both the real and simulated catalogs [Kagan and Knopoff, 1978, 1980b]. The residual real catalogs may still be biased due to the possible failure to take into account some weak aftershocks which occur at the shortest time intervals (1 and 2).

The partial absence of short time aftershocks in real catalogs may also explain the higher values of the likelihood function $L - L_0$ [Kagan and Knopoff, 1978] for simulated catalogs (Table 1) and the lower values of the coefficient $c_-$ in the simulated compared to real catalogs. If so, then our earlier result that $c_- = 1/c_+$ [Kagan and Knopoff, 1978] may be coincidental. The entry $L/n$ in Table 1 shows the average amount of information in a catalog in units of bits per earthquake [Kagan and Knopoff, 1977]. We have noted earlier that the values of $L - L_0$ and $L/n$ that both foreshocks and aftershocks are a manifestation of essentially the same process, namely, the stochastic interaction of earthquakes. Simply stated, if the second earthquake of a pair of interrelated shocks happens to be larger than the first, we call the first member of the pair a foreshock. From the values of $v_+$ in Table 1 the number of cases for which the above possibility is encountered in simulated catalogs is in good agreement with the corresponding number for real catalogs.

Although all events were produced by the same stochastic mechanism, foreshocks and aftershocks seem to have different magnitude distributions as reflected in the values of $c_-$ and $c_+$, respectively: the number of dependent events in the magnitude interval $dM_B$ is proportional to $c_- dM_B$. The values of $c_+$ in real catalogs [Kagan and Knopoff, 1980b] are approximately similar to the values for synthetic catalogs, although the values of $c_-$ in the simulated catalogs are lower than their real counterparts.

The greatest difference between the second-order properties of simulated and real catalogs relate to the values of the branching rate $\mu$ for short-time aftershocks. The values of $\mu$ in aftershock time intervals 1 and 2 are significantly higher for the simulated catalogs than the corresponding estimates for real earthquakes [for definition of time intervals see Kagan and Knopoff 1978, 1980b]. This discrepancy may be connected with the incomplete identification of weak shocks in the wake of strong main events in real earthquake catalogs. In our maximum likelihood optimizations we have excluded earthquakes which occurred a short time after the main shocks from both the real and simulated catalogs [Kagan and Knopoff, 1978, 1980b]. The residual real catalogs may still be biased due to the possible failure to take into account some weak aftershocks which occur at the shortest time intervals (1 and 2).

The partial absence of short time aftershocks in real catalogs may also explain the higher values of the likelihood function $L - L_0$ [Kagan and Knopoff, 1978] for simulated catalogs (Table 1) and the lower values of the coefficient $c_-$ in the simulated compared to real catalogs. If so, then our earlier result that $c_- = 1/c_+$ [Kagan and Knopoff, 1978] may be coincidental. The entry $L/n$ in Table 1 shows the average amount of information in a catalog in units of bits per earthquake [Kagan and Knopoff, 1977]. We have noted earlier that the values of $L - L_0$ and $L/n$
are determined mostly by the short-time interval aftershocks [Kagan and Knopoff, 1977, 1978]; this is due to our use of a maximum likelihood procedure which is influenced most by the high-density regions of the population. The coefficients describing the time interaction between shocks $\nu^+$ are significantly different for different values of $\theta$ and hence for different depths, a result that is similar to our observations for real catalogs [Kagan and Knopoff, 1980b]. We have already commented above that both the synthetic and real catalogs have $b$ values that are essentially independent of depth.

The model has properties that are in excellent agreement with those of catalogs of shallow earthquakes; this agreement is all the more remarkable, since we have been able to generate almost all the known time-magnitude statistical features of shallow earthquake catalogs with only four parameters $\theta$, $m_0^{(2)}/t_0$, $\beta/t$ and $(\log 10 M_0 - M_0)$. Two additional parameters, $\kappa$ and $m_0^{(0)}$, play a relatively insignificant role as determinants of the statistical properties of the result, as long as $\kappa$ is small. We note that the magnitude-frequency relation is derived in our model from the inverse power law memory function of an earthquake process (1) and the assumption of the branching property of the process. Thus the two basic laws of statistical seismology, namely, Omori's law and the magnitude distribution, have been shown to be nonindependent; the second is a consequence of the first.

With regard to seismicity at greater depth, unfortunately, we have only two statistical results with which to compare: the magnitude-frequency relation and the upper bound for the rate of earthquake interaction [Kagan and Knopoff, 1980b]. It is possible, in principle, that the coincidence of predictions from the model with these two features is fortuitous.

8. Implications for the Problem of Seismic Risk

Our model suggests a simple formula in regard to earthquake risk prediction: the average predicted seismicity at a time $t$ immediately after the end of a catalog can be estimated as

$$\lambda(t) = \text{const} \sum_{i=1}^{n} M_{oi} (t - t_i)^{-3/2}$$

(3)

where $M_{oi}$ is the seismic moment of the $i$th earthquake in the catalog and $t_i$ is the time of this earthquake; the normalization coefficient is given in (1).

To calculate the seismic risk at some future time $t'$, we must take into account not only the influence of earthquakes in the existing catalog (3) but also the contribution of shocks which would have occurred in the time interval $(t'-t)$. The latter part can be evaluated, for example, by Monte Carlo simulations of several possible extrapolations of the process from $t$ to $t'$ and then by averaging these simulations. The uncertainty in the estimate of $\lambda(t')$ can be found by the same procedure.

To illustrate formula (3), suppose that there are two earthquake faults such that the only available information is one earthquake on each fault. We suppose that one of these earthquakes had a magnitude close to $8.0$ ($\log M_0 = 28$) and occurred about $300$ years ago ($10^8$ a); and in the other case we take $M = 3.0$ ($\log M_0 = 22$) and assume that it occurred about $12$ days ago ($10^6$ a); both of these faults have equal probabilities of producing a new shock 'today.' If we assume for simplicity that no new earthquakes have occurred on these faults one year from 'now,' the probabilities that each would produce a new earthquake differ drastically: that due to the second is $185$ times less than that due to the first.

We consider the implications of the self-similarity of the seismic process for estimating long-term seismic risk parameters. If we calculate the recurrence time of an earthquake process having a conditional rate of occurrence varying as $t^{-\theta}$, we find that the average recurrence time is infinite for $1 > \theta > 0$ [Mandelbrot, 1977]. It can be argued that the power law distribution may not be valid for very long time intervals. We know that the spatial distribution of earthquakes [Kagan and Knopoff, 1980a] loses its self-similarity at distances greater than $2000$ or $3000$ km. At present we do not have techniques for estimating the corresponding time scale of disappearance of self-similarity, if there is any; it may even be as long as the time scale of mantle convection.

The infinite recurrence time of earthquakes means that the earthquake process is nonstationary. It is of interest to note that some other natural processes exhibit the same nonstationary behavior [Mandelbrot and Wallis, 1969]. Our model implies that almost all earthquakes are statistically and causally interdependent, a conclusion that contradicts attempts to divide the full catalog of earthquakes, either into sets of independent or main sequence events or into sets of dependent events (aftershocks and forerunners). If this picture applies even for the strongest earthquakes, and our results in the previous sections and elsewhere seem to confirm this, then all earthquakes occur in superclusters with very long time spans which may exceed the time spans of all the earthquake catalogs we have at our disposal.

A partial but qualitative confirmation of this possibility can be found in the historical catalogs of China [Lee et al., 1976; York et al., 1976] and the Middle East [Ambreses, 1971; Allen, 1975]; in these catalogs, periods of high seismic activity lasting several centuries are interspersed with periods of significantly lessened seismicity. If even longer fluctuations of activity also occur, we are simply unable to identify them in catalogs of only (sic) a few hundred or a few thousand years duration. These periods of seismic activity might correspond to the lifetime of an active fault; on a geological time scale, even major tectonic features have a limited lifetime, since they are replaced after long intervals by other tectonic features.

In summary, from all of these arguments we assert with high probability that the statistical moments of the seismic historical process do not at present provide us with an instrument to evaluate the seismic risk in the very distant future. In the catalogs presently available to us, seismicity may have a peak (or it can have a minimum as well) in recent centuries of seismic activity; our formulas for earthquake interaction (1) and (3) do not have a characteristic period;
hence they cannot be used for long-term forecasting. But such risk predictions may not be too important, since in practice we only need estimates of seismic risk for a maximum period of, let us say, the next 50 to 100 years, and these estimates can be obtained by an extension of the methods discussed in this paper.

Acknowledgements. This research was supported by grant PFR 77-24742 of the Directorate for Applied Science and Research Applications (ASRA) of the National Science Foundation. We are grateful to F. Schwab for providing us with theoretical seismograms and useful discussion. We also extend our thanks to D. Vere-Jones for valuable discussion and comments. We are grateful to two referees for their useful remarks. Publication No. 2044, Institute of Geophysics and Planetary Physics, University of California, Los Angeles.

References


(Received February 27, 1980; revised May 27, 1980; accepted June 11, 1980.)