

# Statistical Short-Term Earthquake Prediction

Y. Y. KAGAN AND L. KNOPOFF

**A statistical procedure, derived from a theoretical model of fracture growth, is used to identify a foreshock sequence while it is in progress. As a predictor, the procedure reduces the average uncertainty in the rate of occurrence for a future strong earthquake by a factor of more than 1000 when compared with the Poisson rate of occurrence. About one-third of all main shocks with local magnitude greater than or equal to 4.0 in central California can be predicted in this way, starting from a 7-year database that has a lower magnitude cutoff of 1.5. The time scale of such predictions is of the order of a few hours to a few days for foreshocks in the magnitude range from 2.0 to 5.0.**

AS FAR AS WE KNOW, THE ONLY practical approach to short-term earthquake forecasting that makes use of seismological data is one that uses foreshock activity to identify a larger, ensuing event. Typically, most investigators identify foreshocks long after the major event that follows them. After most large earthquakes have occurred, it is usually easy to identify precursory foreshocks. Although it is difficult to identify a seismological precursor while it is in progress, Jones (1) found there is a strong probability that smaller shocks are frequently followed by stronger events within a short time interval. In this report we apply a well-defined stochastic model for the probability that one earthquake will be followed by another of any size to the problem of the prediction of the likelihood of occurrence of a stronger shock. The model, which has only three adjustable parameters, has been derived from an independent, albeit significantly simplified, model of quasi-static fracture growth. A number of models of fracture give insights into the origins of certain features of seismicity, but, to our knowledge, this is the first time that the consequences of a theoretical model of fracture have been directly coupled to seismic data to derive a quantitative, nonempirical forecasting procedure.

We define a "prediction" to be a formal rule whereby the available space-time-seismic moment manifold of earthquake occurrence is significantly contracted and for which the probability of occurrence of an earthquake is anticipated to be significantly increased. We make no specification of the size of the second event of the pair here, except to require it to be stronger than the first.

The statistical reliability of most forecasts of the size, date, and place of a future individual earthquake is difficult to measure quantitatively. Strictly speaking, a forecasted earthquake that occurs just outside the pre-specified time, space, and magnitude limits should be considered a failure, whereas, intuitively, the prediction was almost cor-

rect. Suppose that instead of the occurrence of one predicted strong earthquake, a cluster of slightly smaller, closely related earthquakes occurs that releases about as much energy. Has the prediction failed because the peak accelerations that were expected did not occur, or has it been successful because the total energy released is nearly equal to the predicted value? To respond to these questions we have developed a quantitative measure of the effectiveness or reliability of such predictions. As far as we know, no such measure has been proposed thus far, at least in the multidimensional case of interest to us.

At the present time the reliability of proposed prediction techniques is low. The occurrence of one false alarm (prediction of an earthquake that did not occur) does not disprove the validity of the arguments used in a prediction. If such predictions were formulated as a formal rule and applied to many earthquakes, it might be possible that the rule would actually "predict," that is, perform better than a Poisson random guess. Conversely, a single "successful prediction" does not validate a predictive procedure: an earthquake may occur just by chance, and we cannot tell whether the prediction was successful on its own merits or succeeded by coincidence.

We are concerned with the statistical prediction of strong earthquakes on time scales that are short (on the order of a few hours to a few days) when compared with the recurrence times of the strongest earthquakes, or with the prediction of aftershocks on time scales that are short when compared with aftershock inter-event times. Thus we avoid problems with the instability and variability of long-term earthquake sequences (2, 3). Over short time scales, the dominant feature of statistical earthquake occurrence is a strong clustering of events in time and space (2-4).

Quantitative prediction requires that we estimate future occurrence rates on the basis of probability at all points of the space-time-seismic moment manifold for any possible earthquake sequence. To do this we define a continuous function that is derived

from a stochastic model and is parametrically fitted to the available history of seismicity. We then extrapolate the fitted function to perform the prediction; the effectiveness of the prediction and its accuracy can be evaluated quantitatively. The seismic histories are contained in one of the catalogs of earthquake occurrence (5). These catalogs list earthquakes and give their origin times, their locations in three dimensions, and occasionally focal mechanisms in the form of either fault-plane or seismic moment tensor solutions or both.

We have introduced two quantities that are related to the prediction problem (6). The first is the predictive ratio  $\Lambda(t, \mathbf{x}, M)/\Lambda_0(\mathbf{x}, M)$ , where  $\Lambda(t, \mathbf{x}, M)\Delta t$  is the probability that an earthquake occurs at time  $t$  during a small time interval  $\Delta t$ , at location (and possibly focal parameters)  $\mathbf{x}$ , and with scalar seismic moment  $M$ . The conditional probability  $\Lambda\Delta t$  is computed for some model to be tested (2) for a given history of seismicity. The quantity  $\Lambda_0(\mathbf{x}, M)\Delta t$  is the same probability according to the Poisson hypothesis. The predictive ratio (or its maximum value) has been adopted by some investigators to characterize the effectiveness of a forecast (7, 8).

The second quantity is the information content ( $I$ ), which is the base 2 logarithm of the integral of the predictive ratio over the space-time-seismic moment manifold (4, 6, 9). This quantity can be written as the sum of two terms,

$$I = -\log_2 e \int_0^T \int_{\mathbf{X}} \int_{M_c}^{\infty} [\Lambda(t, \mathbf{x}, M | \boldsymbol{\chi}) - \Lambda_0(\mathbf{x}, M | \boldsymbol{\chi}_0)] dt d\mathbf{x} dM + \int_0^T \int_{\mathbf{X}} \int_{M_c}^{\infty} \log_2 \frac{\Lambda(t, \mathbf{x}, M | \boldsymbol{\chi})}{\Lambda_0(\mathbf{x}, M | \boldsymbol{\chi}_0)} dN(t, \mathbf{x}, M) = (\ell - \ell_0)/\ln 2 \quad (1)$$

where the  $\boldsymbol{\chi}$  values are the parameters of the model;  $T$  and  $\mathbf{X}$  are, respectively, the time and space spans of the catalog;  $M_c$  is the seismic moment cutoff of the catalog;  $dN(t, \mathbf{x}, M)$  are multidimensional delta functions corresponding to the earthquakes in the catalog;  $\ell$  is the logarithm of the likelihood for the particular model of earthquake occurrence; and  $\ell_0$  is the logarithm of the likelihood for the Poisson model. Because the Poisson process has maximum entropy for a given rate of occurrence, it is an ideal reference model for measuring the information content of any competing model. Frequently a single catalog entry spans a large

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

number of subevents that have occurred between the beginning of rupture and the next event in the catalog. Thus the total time span in Eq. 1 includes a “dead time” after each earthquake in the catalog. We systematize these dead times against inconsistencies in identifying aftershocks by taking the dead time to be the coda time computed according to Eq. 4. These dead times can be considered to be one version of a weighting function that suppresses very short time influences. Other rules for incorporating weighting functions are possible.

The values of the parameters  $\chi$  in Eq. 1 are chosen to maximize the logarithm of the likelihood function  $\ell - \ell_0$ . These parameters are selected in view of certain assumptions regarding the stresses at the edge of an earthquake fracture (10). Depending on the details of the particular model and the assumptions, the number of parameters in our models is between three and seven.

We assume that the clusters of foreshocks and aftershocks have a Poisson distribution, although the individual events within the cluster do not. The conditional density or hazard function  $\Lambda(t, \mathbf{x}, M)$  is (11–13)

$$\Lambda(t, \mathbf{x}, M) = \lambda \cdot \phi(\mathbf{x}, M) + \sum_i \psi_{M_i}(t - t_i, \mathbf{x} - \mathbf{x}_i, M) \quad (2)$$

where  $\lambda$  is the rate per unit time of Poisson occurrence of a cluster, the first shock of which has seismic moment greater than or equal to  $M$  in the volume ( $\mathbf{X}$ );  $\phi(\mathbf{x}, M)$  is their space–seismic moment distribution;  $\psi_{M_i}(t - t_i, \mathbf{x} - \mathbf{x}_i, M)$  is the conditional distribution of later events occurring at time  $t$  and coordinates  $\mathbf{x}$ , if earlier earthquakes have occurred at times  $t_i$  and coordinates  $\mathbf{x}_i$ . For the purposes of the illustration below, we shall assume that an “alarm” is declared when the hazard function exceeds a certain “threshold” rate.

The conditional distribution of the  $j$ th shock, which depends on the occurrence of the  $i$ th independent shock ( $j > i$ ) with seismic moment  $M_i$ , is

$$\psi_{M_i}(\tau_{ij}, \rho, \zeta, M_j) = \nu(M_i) \cdot \psi_\tau(\tau_{ij}) \cdot \psi_\rho(\rho) \cdot \psi_\zeta(\zeta) \cdot \phi_M(M_j) \quad (3)$$

where  $\tau_{ij} = t_j - t_i$  ( $\tau > 0$ );  $\rho$  is the horizontal distance between the  $i$ th and  $j$ th epicenters;  $\zeta = z_j - z_i$  is the vertical distance between the hypocenters;  $\nu(M_i)$  is the total number of dependent shocks “generated” by a shock with seismic moment  $M_i$ ;  $\psi_*(^*_{ij})$  are conditional density distribution functions (the asterisk indicates the three possible arguments of the functions); and  $\phi_M(M_j)$  is the unconditional distribution density function of the seismic moment (14).

We have estimated the values of the pa-

rameters of the model of Eq. 2 and the information content for the part (1971 through 1977) of the California Network (CALNET) (U.S. Geological Survey) catalog for central California (15) that is in final form (16); more recent catalogs have not been winnowed to eliminate spurious events. The magnitude threshold for the catalog has been taken to be 1.5 (4, 15). There are 7360 events in the catalog.

The occurrence of one earthquake raises the probability level immediately; it then decays rapidly at a rate that depends on the scalar seismic moment  $M$  of the earthquake and the time that has elapsed since the earthquake’s occurrence  $\tau$  (2, 3):

$$\psi_M(\tau) = \nu(M) \cdot \psi_\tau(\tau) = \frac{\mu}{2} \left[ \frac{M}{M_c} \right]^{2/3} \left[ \frac{t_M}{\tau^3} \right]^{1/2} \text{ for } \tau \geq t_M \quad (4)$$

where the total number of dependent events is  $\mu(M/M_c)^{2/3}$  and  $t_M$  is the time span of the coda waves of an earthquake with seismic moment  $M$ . The values of  $t_M$  and  $M_c$  depend on the properties of the seismographic network. For the CALNET catalog we take  $t_M$  to be  $3.46 \times 10^{-3}$  day for an earthquake with scalar seismic moment  $\log_{10} M = 22.4$  [local magnitude ( $m_L$ ) about 4.0] (17) and  $\log_{10} M_c = 18.6$  ( $m_L$  about 1.5). The coda duration  $t_M$  is proportional to the cube root of the seismic moment (3). We set  $\mu = 0.075$ . Numerical tests show that the results of the predictions are not influenced significantly by changes in the value of  $\mu$ . A change of  $\mu$  is simply equivalent to a change of the alarm threshold rate for large values of the alarm threshold rate; the value of  $\mu$  influences the prediction efficiency only for values of the threshold rate close to the Poisson rate.

Introduction of the epicentral coordinates into the fitting process (Eq. 3) increases the information content by a factor of from 5 to 10. The use of depth data in Eq. 3 increases the information content by only about 15%; this effect is small because of the high redundancy in the narrow depth range of the hypocenters in the CALNET catalog. To illustrate the use of Eq. 3, we neglect all of the conditional distributions other than those for origin times and one-dimensional locations of a projection of the epicenters on the trace of the San Andreas fault, which we take to be approximately N37°W in this region. In Table 1, the values of  $I$  and  $\ell$  are calculated for subcatalogs consisting of 1-year intervals and five equal 73-km portions of the San Andreas fault zone, with the first segment being the northernmost.

Although these quantities fluctuate strongly over the temporal and spatial subdivisions of the catalog, closer inspection

shows that the value of  $I$  depends strongly on the occurrence of large clusters of earthquakes; most often, but not necessarily, these clusters are the aftershocks of some strong event. An indicator of the presence of a large aftershock sequence is the appearance of a large value of the maximum magnitude  $m_{\max}$  in the catalog.

We estimate the probability of occurrence of future earthquakes from Eq. 2. The adjustable parameters  $\chi$  are given by a maximum likelihood procedure. Some of the parameters are obtained from other physical or geological considerations (3, 10, 14); both the unconditional and conditional predicted distributions of earthquake sizes (Eq. 3) are assumed to obey the Gutenberg-Richter law or its modification for the distribution of the seismic moment tensor (14). As a direct consequence of the latter assumption, we make no attempt here to “predict” the seismic moment or any other measure of the size of a future earthquake. Only the time–multidimensional space probability or seismic activity has been extrapolated from the available data.

Figure 1 shows an example of the hazard function, or the prediction, as a function of time. We take the spatial distribution to be Gaussian

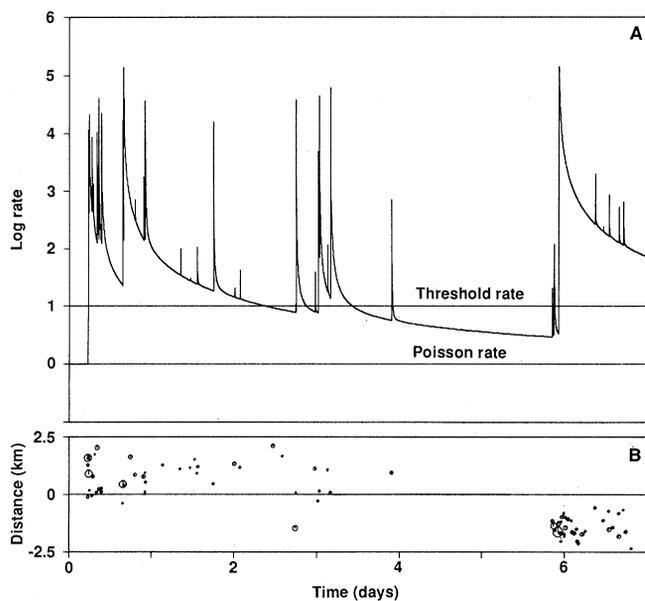
$$\psi_\xi(\xi) = (\sigma\sqrt{2\pi})^{-1} \cdot \exp[-\xi^2/(2\sigma^2)] \quad (5)$$

where  $\xi$  is the distance between shocks along the line. The value of  $\sigma$  is estimated to be about 0.5 km for an earthquake with  $\log_{10} M = 22.4$  ( $m_L = 4.0$ );  $\sigma$  is proportional to the cube root of the scalar seismic moment.

The hazard function displayed in Fig. 1A

**Table 1.** Likelihood function for the CALNET catalog of earthquakes and its time and space subdivisions. Abbreviations:  $n$ , number of events;  $m_{\max}$ , maximum magnitude;  $\ell$ , logarithm of the likelihood;  $I/n$ , information content ratio.

Time or space interval	$n$	$m_{\max}$	$\ell$	$I/n$
<i>Time intervals, all space</i>				
1971	784	3.9	773.0	1.42
1972	1511	5.0	2760.2	2.64
1973	1013	4.6	912.7	1.30
1974	946	5.2	876.6	1.34
1975	1004	4.9	583.8	0.84
1976	1061	4.3	791.3	1.08
1977	1041	4.3	1239.4	1.72
<i>Space intervals, all time</i>				
1	96	3.2	45.6	0.69
2	736	4.3	1370.6	2.69
3	1750	5.2	1210.2	1.00
4	3719	5.0	4211.9	1.63
5	1059	4.9	401.9	0.55
<i>All space-time intervals</i>				
7360	5.2	8080.2	1.58	



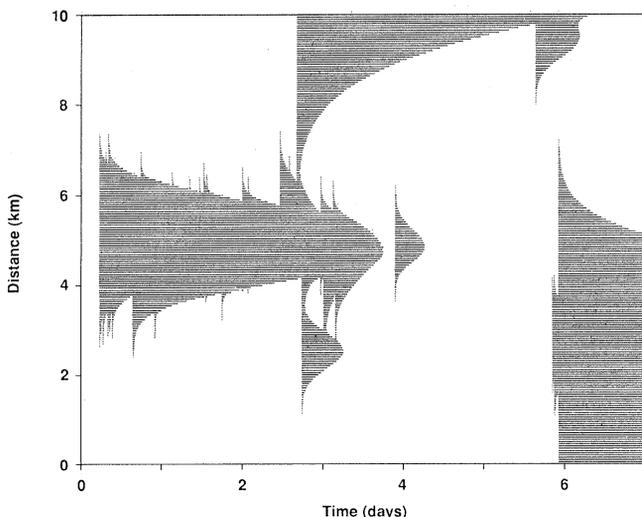
**Fig. 1.** (A) Conditional rate of occurrence of earthquakes or hazard function for Bear Valley, California;  $t = 0$  is 22 February 1972, 0000 Greenwich Mean Time. The threshold rate is ten times the Poisson rate. (B) A time-distance plot of earthquake occurrence; the sizes of the symbols are roughly proportional to the magnitudes of the events. The center line, marked at 0, in (B) corresponds to the reference point of the diagram in (A).

spans the interval from 22 to 29 February 1972 for one point on the San Andreas fault in Bear Valley, California (18), a time of particularly vigorous activity. Most of the time-location space is significantly less densely studded with alarms. The number of independent events in the full catalog is estimated to be 5355. The reference Poisson rate is 0.0058 earthquake/(day times kilometers), which we assume to be a constant for the entire fault and the entire period of the catalog. For purposes of illustration, we chose the threshold rate for an alarm to be ten times the Poisson rate.

The spikes are the functions  $t^{-3/2}$ . The peaks of the functions for the stronger events are not as large as those for the smaller events, since the conditional rate function is plotted only from the end of the coda, and larger earthquakes have larger coda times. However, the influence of stronger shocks can be recognized on the display

by longer decay times. Figure 1B shows those earthquakes in this time period whose epicentral projections are close to the "prediction" point. The reduced influence of more remote earthquakes can be discerned.

The display is characterized by two significant alarm bursts above threshold. The strongest earthquake ( $m_L = 3.5$ ) in the first burst occurred about 0.5 day after the alarm had been declared at about 0.2 day. The largest event ( $m_L = 4.6$ ) occurred shortly after the second alarm at about 6 days. Two of the three small earthquakes that immediately preceded the largest event triggered separate alarms, but in each case the shocks were so small that the probability function fell rapidly below threshold. Considered rigorously, all of the shocks above the threshold in Fig. 1 were forecast successfully except those that "turned on" alarms; however, most of the shocks must be considered aftershocks, for which we cannot take much



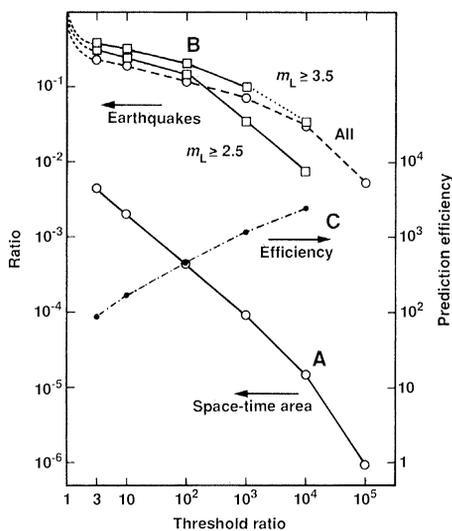
**Fig. 2.** Earthquake hazard map for a 10-km segment of the San Andreas fault in Bear Valley, California. The time scale is the same as in Fig. 1. The reference point for Fig. 1 is at coordinate 4 km. The shaded areas indicate predicted rates that exceed the threshold rate. The threshold rate is ten times the Poisson rate.

credit. A first "predicted" earthquake starts an alarm and "predicts" a second, even stronger earthquake; if no second event occurs that is stronger during the prolongation of the burst, we have generated a false alarm. Thus, every burst must be a generator of one false alarm, even if it contains one or more successful predictions. Similarly, the declaring of an alarm is by definition a failure-to-predict, and every burst is identified with one such failure.

Some of these relations become clearer if we consider the space-time hazard function. Figure 2 shows the locations and times that are above the threshold for a 10-km segment of the San Andreas fault in the Bear Valley area for the same time interval as in Fig. 1. The disconnected alarms of Fig. 1 near  $t = 6$  days are in reality connected; the disjoint nature of the alarm intervals in Fig. 1 arose because of our failure to display spatial variations. Figure 1 represents a section through Fig. 2 at a coordinate that cuts across the decays for the first two events. The complexity of the interconnections is associated with the complexity of the geometry of earthquake occurrence, which has been shown to have a fractal character (2, 19). The complexity of the blotchy pattern will probably increase as the number of dimensions of the prediction space increases.

The ratio of the total size of all alarm zones to the total space-time size of the catalog is shown in Fig. 3 as a function of the threshold level. As may have been expected, the space-time areas of the alarm zones decrease as the prediction threshold is raised. Also displayed is the ratio of the number of events that fall into the alarm zones to the total number of earthquakes in a given catalog. The latter ratios have been calculated (i) for main earthquakes with  $m_L = 3.5$  or greater that have been preceded by foreshocks, that is, by smaller earthquakes that triggered the alarm; and (ii) for all events with  $m_L \geq 1.5$ ; in this case, we count aftershocks as well as main shocks as "successes."

For a threshold ratio of  $10^3$ , for example, the total size of the dangerous zones is  $9 \times 10^{-5}$  of the total time-distance area; 7.1% of all earthquakes with  $m_L \geq 1.5$ , 6.6% of 301 main shocks with  $m_L \geq 3.0$ , 10% of 58 main shocks with  $m_L \geq 3.5$ , and 14% of 21 main shocks with  $m_L \geq 4.0$  occur in these zones. For low levels of the threshold ratio, the maximum percentage of successes, that is, the number of earthquakes with  $m_L \geq 4.0$  that are preceded by foreshocks, is about 33%. If we take the threshold to be equal to the Poisson rate and assume that shocks occur uniformly over the entire fault, almost all of the earthquakes



**Fig. 3.** Prediction efficiency for a 7-year, 364-km interval of the San Andreas fault in central California. Threshold rates are in units of the Poisson rate. Curve A, ratio of the total space-time size of all dangerous zones to the total space-time size of the catalog. Curves B, ratio of the number of events falling into dangerous zones to the total number of earthquakes. The dashed line represents main shocks and aftershocks for all events with  $m_L \geq 1.5$ . Curve C, prediction efficiency for main shocks with  $m_L \geq 3.5$ .

will be predicted, since almost all of the space-time span of the catalog will be an alarm zone; the value of this "prediction" is nil.

Similarly, the occurrence of a single very large earthquake raises the probability function to such a high level that an alarm, presumably for aftershocks, is declared that may endure for a number of years. Under the assumption that the triggering earthquake is the only one holding the alarm open, for a limited range of local magnitudes the number of days that an alarm will be open is  $0.0154 \times 10^{5m_L/6} \times r_i^{-2/3}$ , where  $r_i$  is the relative threshold rate, for distance  $\xi = 0$  (Eq. 5). Most, if not all, of the strong aftershocks of any earthquake can be predicted by our procedure. Even though predictions of aftershocks do not have the same visibility as the prediction of main shocks, strong aftershocks can be dangerous, and they represent an area of significant engineering and public concern.

The ratio of the percentage of earthquakes predicted to the percentage size of the dangerous zones gives an estimate of the improvement in the prediction over the Poisson assumption. We call this ratio the efficiency of the prediction. For earthquakes with  $m_L \geq 3.5$  the efficiency is about 1100 for a threshold ratio of  $10^3$  (Fig. 3). We prefer to use the efficiency instead of the false alarm rate as a measure of success or failure of the prediction; our representation always predicts false alarms. The efficiency

of prediction of future main shocks increases with increase in the magnitude of the main shock.

The drop in the success rate with increasing threshold ratio is small up to a ratio of about  $10^2$ ; hence, up to this threshold, the space-time size of the alarm zones can be reduced strongly without major reduction in the success rate. Beyond this threshold the success rate falls off rapidly, at least for this earthquake catalog. The difference between the success rate and 100% is the rate of failures-to-predict; as the threshold level increases, both the efficiency of the prediction and the rate of failures-to-predict increase (Fig. 3).

Criteria for alarm onset and call-off may have to be modified according to the user's needs. Such modifications could result in different failure-to-predict rates than those reported here. The choice of the appropriate weighting function (Eq. 1) and the threshold level are discretionary parameters for the users of the prediction technique.

Figure 2 displays the space-time zones of increased probability of occurrence of earthquake epicenters. To predict the damage caused by larger earthquakes, we will have to take into account the size of the rupture zone, as well as other engineering features such as propagation effects, soil conditions, and attenuation; these aspects of the problem are not considered here.

We estimate the maximum effectiveness of our procedures by calculating the specific information content per event,  $I/n$ , in a model that simulates the complete process. The output of a stochastic model of earthquake occurrence that simulates the Poisson cluster process well (2, 3, 19), with parameters appropriate to the CALNET catalog, yields from 10 to 15 bits of information per earthquake, if the synthetic catalog is processed in the same way as above. We believe the major reason for the difference is that although our synthetic catalogs include large earthquakes, there are no large earthquakes in the period 1971 through 1977 for the CALNET catalog; the largest earthquake ( $m_L = 5.2$ ) ruptured only about 3% of the total length of the fault. Support for this interpretation is found by noting that during the prominent Bear Valley sequence of earthquakes of 1972 (18) the value of  $I/n$  was significantly larger than during all of the other years of the catalog (Table 1). Comparison with other earthquake catalogs shows that an increase in the length of the catalog does not necessarily reflect an increase in the bit rate per earthquake; the latter quantity depends on the quality of the seismographic network and on the presence of large earthquakes in the catalog (4).

Because the ratio of  $I/n$  for the theoretical

result is larger than that derived from the observations, it might be possible that the uncertainty in earthquake occurrence can in principle be reduced by a factor of  $2^{10}$  when compared with the Poisson model, since each bit of information reduces the uncertainty by a factor of 2. These estimates are somewhat suspect because we do not know the numbers of degrees of freedom either in the stochastic model or in the result of processing by Eq. 1. We cannot assess the influence of inadequacies due to the use of the branching process simulation model in the stochastic model, nor can we assess the influences of misidentification, origin time and location errors, the selection of arbitrary thresholds, and the use of models such as the magnitude-frequency law, in the analysis of actual data. However, we know that the absence of large earthquakes in the test period gives a significantly low value of the specific information content. Although both foreshock-main shock and main shock-aftershock sequences are used in the computation of the information content, most of the information is supplied by the latter sequences, which are of less interest for purposes of prediction. However, the results of statistical analysis of earthquake catalogs (3, 4) and those of the modeling of the occurrence of earthquakes by a self-similar stochastic model (2, 19) indicate that both foreshocks and aftershocks are manifestations of essentially the same process, namely, the stochastic interaction of earthquakes. We interpret our model as an indication that we are able to derive the likelihood of the triggering of one earthquake by another but that we do not know whether the succeeding earthquake will be larger than the preceding one; the next level of refinement of this model will involve computation of the relative magnitudes of the events.

In view of these comments it is startling that the estimates of the effectiveness of predictions based on the information content are of the same order of magnitude as those obtained from the hazard function (Fig. 3). This agreement may be coincidental because the methods may have different numbers of degrees of freedom and constraints. The information content is suited for theoretical studies of earthquake occurrence, whereas the hazard function is more useful for earthquake forecasting. The information content for the CALNET catalog decreases rapidly if the dead time is increased beyond the coda time. After 1 hour, 10% of the predictive information is lost; 1 day later, the reduction is one-third; and 10 days later it is one-half (3). In addition, the failure-to-predict rate increases rapidly with increasing dead time. The information content for the best seismological data presently

available is not an upper limit. But it may be increased significantly through the introduction of seismic moment tensor information and estimates of stress from geological and geodetical investigations as well as from past earthquakes.

Our results indicate that only about one-third of the strong earthquakes are preceded by foreshocks that are separated in time from an independent shock by more than  $t_M$ . For other modern catalogs of earthquakes that are similar in quality to the central California CALNET catalog, the number of failures-to-predict is about the same, that is, of the order of two-thirds (1, 3, 4).

A modification of the above strategy for prediction is called for in the case of the occurrence of strong earthquakes. Strong earthquakes have the potential for serving as foreshocks of even stronger earthquakes, or they may be the main shock in the sequence, just as weaker earthquakes can serve both functions. However, the coda time  $t_M$  is about 15 minutes for an earthquake with  $m_L = 5$ , and this time increases by a factor of about  $\sqrt{10}$  for a unit increase in earthquake magnitude. If an earthquake with  $m_L = 6$  were to occur, with or without prior warning according to the scenario above, no alarm for a possibly even stronger earthquake would be sounded for about  $t_M = 50$  minutes, which might be an unacceptably long delay for issuing a warning. This difficulty is circumvented if we reduce the dead time for large earthquakes to a value less than the coda time.

This modification for strong earthquakes indicates that response strategies can also be developed with time delays of the order of seconds. As suggested by Heaton (20), it may be possible to predict some large earthquakes through the analysis of small starting phases of complex events that later blossom into large earthquakes. These small starting phases are genuine earthquakes whose signals overlap with those of their successors and raise the probability level for a short time, thereby triggering an alarm for the larger event. The number of these preshocks should increase as  $t_b^{-3/2}$ , where  $t_b$  is the time before the start of the main phase of a strong earthquake (2-4). In the present method, we are not restricted to dead times of the order of rupture times, but instead we are able to use longer delays of the order of a few minutes. With this procedure there should be far fewer failures-to-predict for very strong earthquakes. Automated response strategies could take advantage of these predictions in a well-developed technology.

The differences between our proposed forecasting technique and methods that use

empirically derived probabilities of fore-shock-main shock occurrence may be summarized as follows. (i) Since our model is based on a formulation derived from a multidimensional stochastic process, it is not necessary to use arbitrary windows to analyze seismicity, nor is it necessary to delete aftershocks from a catalog to make the catalog amenable to statistical analysis. Therefore, our forecasts are not dependent on a post-factum classification of earthquakes into fore-, main, and aftershocks, a subdivision that may not be possible in real time. (ii) Since the parameters of our seismicity model are obtained through a maximum likelihood procedure, the model is optimal in a quantitative way. The choices of the parameters can be justified on the basis of a well-defined theoretical model of earthquake occurrence (2, 10, 19). Furthermore, the model itself is consistent with all of the other aspects of statistical seismicity that have been well documented, and it has not been derived for the sole purpose of developing the foreshock-main shock relations. The model has only three parameters that are adjusted to the properties of the local seismicity: the rate of occurrence of independent earthquakes and the coefficients that specify the occurrence of dependent earthquakes,  $\mu$  and  $\sigma$ . In one sense the exponent 2/3 in Eq. 4 is also an adjustable parameter, but since this can be derived on formal grounds (3), we have considered it as fixed. (iii) The likelihood function we have derived allows us to estimate the effectiveness of the proposed prediction scheme in terms of information content of an earth-

quake catalog. The procedures outlined here can be adapted to predicting schemes other than the one we have used, as soon as the quality and quantity of the data describing these precursors reach the stage where they can be processed by similar multidimensional statistical techniques.

#### REFERENCES AND NOTES

1. L. M. Jones, *Bull. Seismol. Soc. Am.* **75**, 1669 (1985).
2. Y. Y. Kagan and L. Knopoff, *J. Geophys. Res.* **86**, 2853 (1981).
3. ———, in preparation.
4. ———, *Geophys. J. R. Astron. Soc.* **55**, 67 (1978).
5. W. Rinehart, H. Meyers, C. A. von Hake, *Summary of Earthquake Data Base* (National Geophysical Data Center, Boulder, CO, 1985).
6. Y. Kagan and L. Knopoff, *Phys. Earth Planet. Inter.* **14**, 97 (1977).
7. D. Vere-Jones, *J. Phys. Earth* **26**, 129 (1978).
8. K. Aki, in *Earthquake Prediction, An International Review*, D. W. Simpson and P. G. Richards, Eds. (Maurice Ewing Series, no. 4, American Geophysical Union, Washington, DC, 1981), p. 566.
9. F. Papangelou, *Z. Wahrscheinlichkeitstheorie verw. Geb.* **44**, 191 (1978).
10. Y. Y. Kagan and L. Knopoff, *Geophys. J. R. Astron. Soc.* **88**, 723 (1987).
11. I. Rubin, *IEEE Trans. Inf. Theory* **IT-18**, 547 (1972).
12. T. Ozaki, *Ann. Inst. Stat. Math.* **B31**, 145 (1979).
13. D. Vere-Jones and T. Ozaki, *ibid.* **B34**, 189 (1982).
14. Y. Y. Kagan and L. Knopoff, *Proc. 8th Int. Conf. Earth Eng.* **1**, 295 (1984).
15. S. M. Marks and F. W. Lester, *U.S. Geol. Surv. Open File Rep.* **80-1264** (1980), and references therein.
16. See (4), p. 69.
17. W. H. K. Lee et al., *U.S. Geol. Surv. Open File Rep.* (1972).
18. W. L. Ellsworth, *Bull. Seismol. Soc. Am.* **65**, 483 (1975).
19. Y. Y. Kagan, *Geophys. J. R. Astron. Soc.* **71**, 659 (1982).
20. T. H. Heaton, *Science* **228**, 987 (1985).
21. Supported in part by grant CEE-84-07553 from the National Science Foundation. We acknowledge useful discussion with G. M. Molchan. Publication 2962, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024.

4 August 1986; accepted 1 April 1987

## Homozygosity Mapping: A Way to Map Human Recessive Traits with the DNA of Inbred Children

ERIC S. LANDER AND DAVID BOTSTEIN

An efficient strategy for mapping human genes that cause recessive traits has been devised that uses mapped restriction fragment length polymorphisms (RFLPs) and the DNA of affected children from consanguineous marriages. The method involves detection of the disease locus by virtue of the fact that the adjacent region will preferentially be homozygous by descent in such inbred children. A single affected child of a first-cousin marriage is shown to contain the same total information about linkage as a nuclear family with three affected children. Calculations show that it should be practical to map a recessive disease gene by studying DNA from fewer than a dozen unrelated, affected inbred children, given a complete RFLP linkage map. The method should make it possible to map many recessive diseases for which it is impractical or impossible to collect adequate numbers of families with multiple affected offspring.

IN HIS CLASSIC STUDY OF INBORN ERRORS of metabolism, Garrod (1) noted that an unusually high proportion of patients with alkaptonuria were progeny of consanguineous marriages. Almost immedi-

E. S. Lander, Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142; Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; and Harvard University, Cambridge, MA 02138.  
D. Botstein, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.