JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 99, NO. B4, PAGES 6769–6788, APRIL 10, 1994

Prediction of large events on a dynamical model of a fault

S. L. Pepke and J. M. Carlson[1]
Department of Physics, University of California, Santa Barbara

B. E. Shaw[2]
Institute for Theoretical Physics, University of California, Santa Barbara

Abstract. We present results for long-term and intermediate-term prediction algorithms applied to a simple mechanical model of a fault. The long-term techniques we consider include the slip-predictable and time-predictable methods and prediction based upon the distribution of repeat times between large events. Neither the slip-predictable nor time-predictable method works well on our model. In comparison, the time interval method is much more effective and is used here to establish a benchmark for predictability. We consider intermediate-term prediction techniques which employ pattern recognition to identify seismic precursors. These methods are found to be significantly more effective at predicting coming large events than methods based on recurrence intervals. The performances of four specific precursors are compared using a quality function Q, which is similar to functions used in linear cost-benefit analysis. When the quality function equally weights (1) the benefit of a successful prediction, (2) the cost of maintaining alerts, and (3) the cost of false alarms, we find that Q is optimized in algorithms based on the most conventional precursors when alarms occupy 10-20% of the mean recurrence interval and approximately 90% of the events are successfully predicted. The measure Q is further used to explore optimization questions such as variation in the space, time, and magnitude windows used in the pattern recognition algorithms. Finally, we study the intrinsic uncertainties associated with seismicity catalogs of restricted lengths. In particular, we test the hypothesis that many shorter catalogs are as effective as one long catalog in determining algorithm parameters, and we find that the hypothesis is valid for the model when the catalogs are of the order of the mean recurrence interval.

1. Introduction

Prediction of the occurrence of large earthquakes within a narrow space-time window on a fault has proven to be a difficult problem for several reasons: 1) timescales over which reliable and detailed seismological records are available are often small compared to recurrence times within a fault zone, 2) complexity of fault geometry and dynamics leads to great variability in premonitory phenomena, 3) initiating mechanisms for large events are not completely understood (which inhibits the determination of the relative importance of various precursors), and 4) knowledge of the strain distribution and yield points along faults is insufficient to indicate the locations of future epicenters. The above make it difficult not only to predict well but also to determine how inherently predictable the system is and to find optimal forecasting methods.

With relatively little certain knowledge concerning Earth's fault system, one must be concerned with optimizing prediction using the data at hand as well as objectively evaluating the quality of the predictions that are made. In seismology, these goals cannot be met using seismicity catalogs alone, because they represent only a brief record of the system relative to the timescale of the seismic cycle. One possible path forward is through the use of artificial catalogs to compare methods of forecasting, since one may numerically generate a wealth of statistics for them and also have the ability to vary system parameters. These features allow for algorithm development and the study of optimization procedures to an extent which is not possible for real catalogs. Volumes of literature have been published documenting possible precursory phenomena and event distributions, yet little is written on how best to use such information. It is with respect to this question of optimal use of available information that synthetic catalogs may prove most useful.

[1] Also at Institute for Theoretical Physics, University of California, Santa Barbara.

[2] Now at Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York.

Paper number 93JB03125.
0148-0227/94/93JB03125$05.00

In this paper we present results for long-term and intermediate-term prediction algorithms applied to catalogs generated from a dynamical model of a fault. Our objective is not to prove that a particular model will quantitatively duplicate the complex seismicity patterns observed on real faults. Instead, this study will address issues related to algorithm optimization and the intrinsic limitations of algorithms given the sparsity of data for Earth. We consider long-term prediction techniques, such as the time-predictable and slip-predictable algorithms, which are based solely on characteristics of the most recent large event. In a similar spirit, we also make predictions based on the distribution of time intervals between large events, which can be determined to arbitrary accuracy on a model such as that which will be considered here. However, the principal results of this paper involve intermediate-term prediction techniques analogous to those which have been developed by *Keilis–Borok and Rotwain* [1990] and *Keilis-Borok and Kossobokov* [1990] and have recently been the object of much attention (see, for example, *Healey et al.* [1992]). The aim of these algorithms is to provide an objective means for assessing the probabilities of large earthquakes based on a collection of precursor functions, the values of which are determined by regional small and medium size events. The precursor functions include overall activity, rate of change of activity, and clustering of events, and they are evaluated in coarse grained space-time windows. Simple pattern recognition techniques are used to establish threshold levels for signaling an alert.

On the model we consider the simplest versions of the algorithms using only single precursors and find that they do perform better than long-term prediction techniques. However, for the most standard precursors, such as the level of seismic activity, the algorithms still require the alarms to be on for a significant fraction of the mean repeat time (of the order of 10–20%). This still leaves open the possibility that algorithms which utilize a combination of precursors will be more effective. In addition, we will introduce a new precursor function, which is related to the degree to which seismicity extends throughout the region, which significantly outperforms the others measures, and leads to reliable predictions on a timescale of the order of 3–5% of the mean repeat time.

For the purposes of prediction studies, it is important to distinguish the two main types of artificial catalogs, each of which provides a means to a different end. First, there are purely statistical catalogs, which are constructed to satisfy certain statistical constraints observed in real catalogs, such as consistency with the Gutenberg–Richter law, Omori's law, and/or spatio–temporal clustering. However, because these catalogs are otherwise random, as a test for predictability they are most useful as lower bounds on the effectiveness of an algorithm [*Dieterich*, 1992; *Minster and Williams*, 1993]. A good algorithm should detect some inherent correlation which has not been put in by hand and thus, if such correlations exist on Earth, the algorithm should do better on real catalogs than it does on any statistically generated catalog.

In contrast, one can consider artificial catalogs generated by dynamical models as we do here. We will use the model-generated catalogs as a means to the end of algorithm optimization, letting the physical mechanisms guide us in determining which properties of the catalogs are important for prediction. Here, unlike the purely statistical catalogs, no features are a priori built in. Our goal is to identify features which are generic to a class of physical models and discover how those features function in prediction algorithms which may be easily adapted to different fault systems worldwide.

The model which we consider is a one-dimensional homogeneous model for a fault which has recently been studied in a variety of contexts including the statistical analysis of intrinsic scaling laws [*Carlson et al.*, 1991] and applications to dynamical fracture [*Langer and Tang*, 1991; *Langer*, 1992]. The model is a good candidate for studies of seismic phenomena not only because the physical basis for the generated seismicity is sound (the dynamics depend upon a stick-slip, velocity-weakening friction law) but also because its fundamental statistical features are reminiscent of behavior which is observed for real earthquakes. For example, the magnitude versus frequency distribution is similar to what is observed for a single fault or narrow fault zone [*Carlson and Langer*, 1989a, b], and the model generates moment spectra similar to those inferred from seismographic observations [*Shaw*, 1993a]. Our model has the advantage of requiring no fine-tuning of parameters. In particular, there is a minimal number of input parameters, all of which are physically meaningful, yet with respect to which the qualitative catalog features are robust over a wide range of values. Another important feature of the model is its simplicity, which allows us to numerically generate data equivalent to millions of years of seismic activity with perfect detection of events, so that we are able to evaluate quantitatively the success of algorithms in a manner that is impossible for real catalogs.

The model studied here is a particularly good candidate for studies of predictability. It is deterministically chaotic and hence technically unpredictable at long enough times. Furthermore, the model exhibits a sharp distinction between small and large events. The smaller, more numerous events tend to cluster in the neighborhood of an epicenter of a future large event [*Shaw et al.*, 1992]. This local increase in activity is a generic precursor in the model and is the primary statistical basis for predictability at shorter timescales. While the analogous behavior is much less systematic for Earth [*Kanamori*, 1981], a similar rise in regional activity has been observed on some occasions prior to large events and is one of the signals used for prediction in the algorithms such as CN and M8 which were introduced and studied by *Keilis–Borok and Rotwain*

[1990] and *Keilis-Borok and Kossobokov* [1990]. However, unrealistically, the model we consider does not exhibit aftershocks, which play an important role in the CN and M8 algorithms. Thus, instead of testing these algorithms explicitly, here we will be testing the more general pattern recognition techniques on which they are based. Methods for generating aftershocks within the model are being studied, however, and will be considered in the context of prediction in future work. Another significant shortcoming of the model catalog is the absence of a source of noisy background events which are clearly present in real catalogs.

Because the behavior of the model is relatively systematic compared to Earth, we expect to be able to predict events on the model better than on the real system. However, that is to the point: to establish predictability limits for these nonlinear dynamical models and to optimize on a "clean" system. We hope that the things we learn by doing so will shed light on the roles of various phenomena occurring during the preparatory period for a large event and allow us to distinguish correctly between meaningful causal premonitory characteristics and misleading happenstance trends.

The organization of the paper is as follows. In section 2 relevant characteristics of the uniform Burridge–Knopoff (UBK) model are reviewed. This includes descriptions of pertinent length and time scales which play a role in the model's predictability. In section 3 we introduce the quality function Q, which is the means by which we evaluate the success of the algorithms. In section 4 we present results for long-term prediction, including the slip-predictable model, the time-predictable model, and prediction based upon recurrence intervals. Section 5 contains our results for prediction based upon intermediate-term precursors. The forecasting method is outlined and applied to individual activity-based precursor functions on the model. Section 6 discusses the robustness of our results with respect to variation of the algorithm and catalog parameters, including variations in catalog length. Section 7 gives a summary and addresses outstanding problems.

2. Relevant Model Characteristics

In the finite difference approximation, the model considered here is one of a class first applied to earthquake dynamics by *Burridge and Knopoff* [1967]. It was reintroduced in its simplest form and analyzed in a modern context by *Carlson and Langer* [1989a, b]. The one-dimensional uniform Burridge–Knopoff (UBK) model represents the motion of one side of a lateral fault which is driven by a slow shear deformation relative to the other side of the fault and which is subject to a velocity–weakening slip-stick friction law at the interface. The system consists of N blocks. Each block is coupled to its nearest neighbors with coil springs representing the linear elastic response of the system to compressional deformations. A leaf spring attaches each block to a fixed upper surface and represents the linear elastic response of the system to shear deformations. The blocks are constrained to move on the surface of the sliding plate. The system is loaded slowly by moving the lower plate at a velocity ν until one of the blocks exceeds the frictional threshold. That block then begins to slide, dissipating energy as determined by the friction law. The initially negative slope of the friction law leads to the essential dynamical instability which maintains the complexity of the series of events over arbitrarily long time periods. An event is considered over when all of the blocks have come to rest.

In the continuum limit the partial differential equation describing the motion is

$$\frac{\partial^2 U}{\partial t^2} = \frac{\partial^2 U}{\partial s^2} - U - \phi[\frac{\partial U}{\partial t}] + \nu t \qquad (1)$$

where $U(s,t)$ is the displacement measured with respect to the fixed upper plate as a function of position s and time t. The dimensionless time t, has been scaled by the characteristic slip time for a (homogeneous) large event. Displacements U have been scaled by the corresponding characteristic slip distance. In equation (1), lengths s are measured in units of a stiffness length which is given by the distance a sound wave travels (of the order of 10 km in the Earth) in the characteristic slip time $\Delta t = 1$. The finite difference version of equation (1), which we solve numerically in this paper, gives the seismological model of *Burridge and Knopoff* [1967]. We refer to it as the uniform Burridge-Knopoff model to emphasize the fact that all of the material parameters are uniform along the model fault, that is, there are no built-in inhomogeneities. The small length scale given by the equilibrium block spacing $1/\ell$ plays an important role in the dynamics of the UBK model (see, for example, *Langer and Tang* [1991]). Recent measurements of microearthquakes suggest that there may be a small-scale cutoff in Earth as well. In particular, the measurements of changes in the distribution of sizes of very small events [*Malin et al.*, 1989; *Aki*, 1987] and measurements which suggest that the smallest earthquakes may have a nearly constant rupture area (see, for example, *Bakun et al.* [1976] and *Archuletta et al.* [1982]) lead to a small length cutoff ranging from meters to a few hundred meters. Thus in units of the characteristic stiffness length (10 km) realistic values of ℓ are of the order of 10^2 to 10^4, i.e., a large number. Here we will take $\ell = 10$ for numerical convenience. *Carlson et al.* [1991] considered the scaling as a function of ℓ. The parameter ν in (1) is the dimensionless pulling speed, given by the ratio of the risetime to the loading period between large events. Realistic values of ν are thus 10^{-8} or less. Here we will take ν to be small enough to preserve the separation of time scales between individual events and the loading mechanism. For technical reasons when we do this it is convenient to introduce the small parameter σ into the velocity-weakening friction law which we will take to be

$$\phi(z) = \begin{cases} (-\infty, 1), & z = 0; \\ (1-\sigma)/\{1 + [2\alpha z/(1-\sigma)]\}, & z > 0\,. \end{cases} \quad (2)$$

Here σ replaces ν in setting the scale for the displacement of the smallest events, and we can set $\nu = 0$ while the blocks slip. The friction parameter α is the ratio of the characteristic slipping speed to the speed at which the friction is reduced by half the difference between the threshold value and the value it attains at high speeds. It is difficult, if not impossible, to determine realistic values of α from laboratory measurements. For a wide range of α the behavior of the UBK model is not particularly sensitive to the exact value. Throughout this paper we will take $\alpha = 3$ which is in this range. As pointed out by *Vasconcelos et al.* [1992] and by *Carlson et al.* [1991], for α small enough the behavior does change substantially, though we believe that this regime is less appropriate to seismic studies. The results which we present in this paper are relevant in the regime where ℓ is large, σ is small, and α is large enough to generate the generic behavior.

As previously mentioned, the UBK model is homogeneous in its material properties. We observe complex behavior as a consequence of a dynamical instability associated with the friction law. Beginning with a small heterogeneity in the initial condition, we evolve the system through several loading cycles until it reaches a statistically steady state, at which point the statistical properties are independent of the initial conditions. In Figure 1 we plot a small fraction (both in space and time) of the catalog which will be used in this paper, which begins after the initial transient period. For each event, a line segment is drawn through all of the blocks that slip, and a cross marks the position of the epicenter for each large event. While there are clear correlations at shorter timescales, because of the underlying homogeneity of the UBK model the long time average of the locations of epicenters of large events is independent of position; that is, an event could happen anywhere with equal probability.

The behavior of this system is found to resemble that of an earthquake fault in several important respects. Defining the seismic moment M to be the total slip during an event:

$$M = \int_{\text{event}} \delta U(s) ds \quad (3)$$

and the magnitude $\mu = \ln M$, for wide ranges of the above parameter values, for small to medium size events the UBK model generates frequency-magnitude distributions described by the Gutenberg–Richter relation $D(\mu) = Ae^{-b\mu}$. Here $D(\mu)d\mu$ is the frequency of occurrence of events in the magnitude interval $[\mu, \mu + d\mu]$ per unit length per unit time (see Figure 2). As long as α is sufficiently large ($\alpha \geq 2.5$), we obtain $b = 1$ robustly in this region. In contrast, the largest events follow a different distribution indicating an overfrequency of these

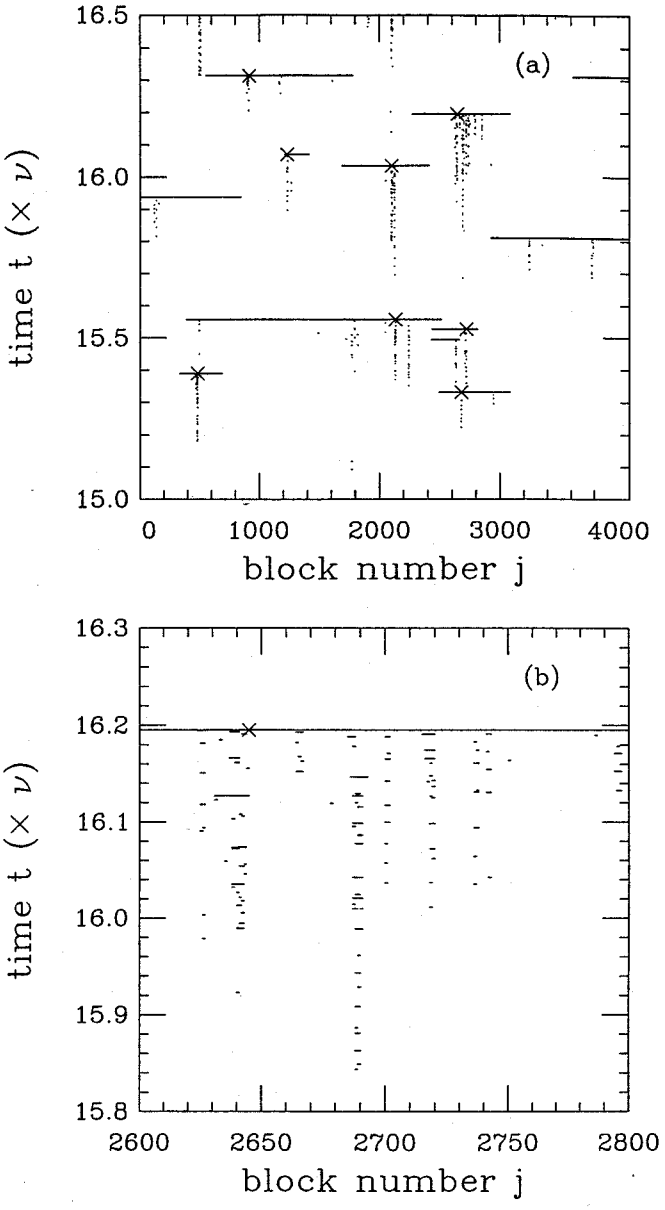

Figure 1. Sample catalog for the uniform Burridge-Knopoff (UBK) model. Time is measured relative to the inverse loading speed. A line segment is drawn through blocks which slip during an event, and a cross marks the position of the epicenter of each large event. Individual events have spatially irregular slip (not shown). Figure 1a represents only a fraction of the full catalog, and Figure 1b is an expansion of the precursory activity for one the large events in Figure 1a. Unless explicitly stated otherwise, the numerical results will be for a system of size $N = 8192$, with $\ell = 10$, $\sigma = 0.01$ and $\alpha = 3$. The full catalog consists of 114,000 events, 1848 of which are large events and corresponds to a total displacement of $t\nu = \delta U = 122$.

events relative to the frequency extrapolated from the distribution of small events. The large events are responsible for nearly all of the moment release in the UBK model. The overfrequency implies that there will be a characteristic repeat time between large events, as discussed by *Carlson* [1991], which can be exploited for

the purposes of long term prediction. The change in behavior between the small and large events is characterized by the length $\tilde{\xi} = 2/\alpha \ln(4\ell^2/\sigma)$. *Carlson and Langer* [1989b] showed that events which are triggered in regions of size less than $\tilde{\xi}$ tend to remain localized, that is, the slip pulses which are generated will decay rapidly when they encounter regions which are far from threshold. In contrast, when the initial triggering zone is larger than $\tilde{\xi}$, the slip pulses will tend to propagate much farther. The length scale $\tilde{\xi}$ coincides well with the upper bound on the clustering of small events such as those illustrated in Figure 1, and the associated crossover magnitude $\tilde{\mu} = \ln(2/\alpha)$ coincides with the minimum in the magnitude versus frequency distribution (Fig. 2). The smaller events smooth the spatial configuration on length scales less than $\tilde{\xi}$, preparing a triggering region for a large roughening event. This leads to spatio-temporal clustering of small-scale activity along the fault prior to a large event, as reported by *Shaw et al.* [1992]. While this systematic increase in activity is a key feature leading to the relative success of the intermediate-term prediction algorithms on the UBK model, this observation alone is not sufficient to determine whether ultimately the UBK model will be more or less predictable than Earth.

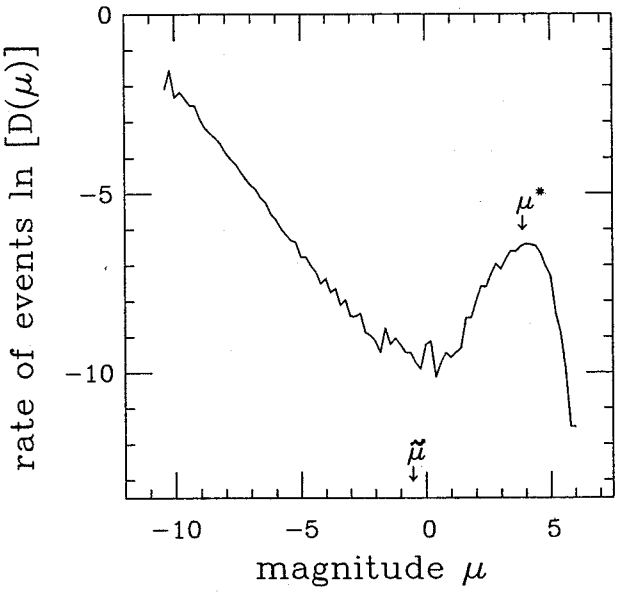

Figure 2. Log frequency $\ln[D(\mu)]$ versus magnitude μ for the uniform Burridge-Knopoff model. $D(\mu)d\mu$ is the number of events per unit length per unit time (measured in units of the inverse loading speed) in the magnitude range $[\mu, \mu + d\mu]$. The small events satisfy the Gutenberg–Richter law with $b = 1$, while the large events occur at an excess rate. The crossover between small and large events is denoted by $\tilde{\mu} = \ln\widetilde{M}$, where $\widetilde{M} \approx 2/\alpha$. The peak in the large events distribution corresponds roughly to a rupture length $\xi^* \approx 10\tilde{\xi}$, well below the system size. The magnitude μ^* associated with such events is indicated. In the corresponding integrated distribution the peak is replaced by a flat shoulder. We plot the differential distribution because it clearly illustrates the crossover $\tilde{\mu}$.

Throughout this work we distinguish between "small" premonitory events and the "large" events we wish to forecast. This distinction can be made precisely in the UBK model: small events are taken to be those for which $\mu < \tilde{\mu}$ while large events have $\mu \geq \tilde{\mu}$. While the existence of a sharp feature is useful to obtain quantitative results, the relatively small number of events of size near $\tilde{\mu}$ compared to events of lesser or greater magnitude implies that our results will not depend strongly on the exact criterion for the crossover that is used. This division is certainly much less precise in real data. It is interesting to note that when data from an individual fault or narrow fault zone are considered, as in the UBK model, an overfrequency of large events is observed [*Wesnousky et al.*, 1983; *Schwartz and Coppersmith*, 1984; *Davison and Scholz*, 1985]. In these cases typical magnitudes of the large events and geodetic measurements of the plate rates can be used to estimate the recurrence time interval and to make long-term predictions. In comparison, data accumulated over a broad region typically do not show an overfrequency of large events. This is a consequence of the fact that faults of different sizes contribute to regional seismicity. In this case, results by *Pacheco et al.* [1992] suggest a bend, or change in b value reflecting an underfrequency of large events, in the magnitude versus frequency distribution of California seismicity at roughly magnitude 6, which coincides with estimates of the magnitude of events which just span the full depth of the seismogenic zone. These types of regional seismicity catalogs are typically used in intermediate-term algorithms such as CN and M8.

Simplifications which are inherent in the UBK model used here are its low dimensionality, single fault dynamics, and lack of aftershocks. Systems of interacting faults and different fault geometries may ultimately prove interesting but would be most effectively represented in the context of a fully two-dimensional elastic medium. Studies of a homogeneous two-dimensional model are currently under way. In addition, *Shaw* [1993b] has proposed an aftershock mechanism which will lead to additional precursory phenomena. Ultimately, we plan to study other models in the context of prediction as well.

3. Evaluation of Forecasting Algorithms

In a spatially extended dynamical system, a prediction typically consists of a projection, based on the current status of the system, of the time and location of a coming event. The forecasts which are made in seismology are necessarily less precise and involve relatively long time scales. Rather than predicting how far in the future an event is likely to take place, one forecasts the likelihood of an event occurring between now and the end of some alarm period. In algorithms CN and

M8 one signals an alert or *time of increased probability* (TIP) to forecast a large event in a specified region in space and time. The TIP begins when the algorithm detects that the system is in a state of readiness, and, in order to minimize the cost associated with signaling an alarm, one must wait as long as possible prior to the event before turning on the alarm.

In this context it is worth noting that there is a technical distinction between algorithms CN and M8 and the intermediate-term prediction algorithms which we will consider here. In CN and M8 the alarm duration is initially fixed to be a specific time interval, typically 5 years. After this period has passed, the status of the region is reevaluated, and a decision is made regarding whether or not to extend the alert. In contrast, after each event we reevaluate the status of the alarms, and the average alarm duration is determined by the threshold for signaling an alert. In a model such as the UBK model, where the precursor functions tend to increase monotonically in the neighborhood of the epicenter of a future large event, alarms will typically (but not always) stay on until a large event has occurred.

With this in mind we now discuss methods for evaluating the long-term and intermediate-term prediction algorithms which will be considered in subsequent sections. While the choice of a specific function for such evaluation is to a large degree based on public policy rather than mathematics or physics, the decision is essential to algorithm optimization.

The success curve, which is defined to be the fraction of events predicted versus the fraction of time an alert is on, provides one meaningful way to evaluate an algorithm. Here each point on the success curve corresponds to a different value of the threshold for signaling an alert. Such curves are useful to compare different methods of forecasting and thus will be considered extensively in the remainder of the paper (see, for example, Figures 6 and 8). However, in order to set the threshold that will ultimately be used, one must define a function on such a curve which will select a particular value of the threshold over the others. One function which is commonly considered in the context of algorithms CN and M8 is called the success ratio which is defined at each point on the success curve to be

$$S = \frac{\text{fraction of large events predicted}}{\text{fraction of time the TIP is on}}. \quad (4)$$

Thus S specifies the gain relative to purely random prediction, for which $S = 1$.

Using S as a measure of the quality of predictions has some inherent complications, as pointed out by *Molchan* [1991] in the context of some time interval distributions. For example, when the alarm time decreases more rapidly than the number of events predicted, S diverges as the alarm time and the fraction of events predicted approach zero. We often observe this behavior in S when intermediate-term algorithms are applied to the UBK model. Algorithm optimization becomes an ill-defined problem in this case.

To solve this problem, we introduce a quality function Q which will be used to optimize the intermediate-term algorithms. The specific choice of a Q function is somewhat arbitrary. Desired features of the Q function are that (1) it should be as simple as possible, with meaningful parameters, so that policy decisions are easily mapped onto parameter settings, (2) the definition should be flexible enough that alternative considerations can easily be incorporated, (3) it should provide a clear measure of when an algorithm is performing well with respect to some simple base measure such as random prediction or doing nothing, (4) it should avoid the pathological behavior as alarm time goes to zero which we observe in S, and (5) the definition should be robust to small changes in algorithm parameters so that determining a maximum value of Q is not an overly delicate procedure.

To satisfy the above criteria, we define Q in its most general form to be

$$Q = \sum_{i=1}^{N} A_i p_i, \quad (5)$$

where the A_i are constant coefficients. For the purposes of this paper we define Q such that the p_i are the probabilities of a set of outcomes with $0 \leq p_i \leq 1$. The probability based function Q in equation (5) is a standard linear measure. Except near the extreme values of p_i its performance is similar to an analogous product-based function of the form $\Pi {p_i}^{\alpha_i}$. In our analysis we choose to use a sum rather than a product to preserve the symmetry between using the success or failure rate of a given outcome, i.e., $p_i \rightarrow (1-p_i)$ and $A_i \rightarrow -A_i$, is preserved in Q (adding or subtracting a constant from Q is irrelevant). Alternatively, one might consider a rate based function $Q = Q_R$ similar to (5), but where the p_i are rates measured relative to the overall rate of large events. From a public policy point of view, the rate-based function Q_R is potentially the most useful since Q_R may be interpreted as a cost-benefit function, and the coefficients can be set according to the relative costs of the set of possible outcomes. Linear cost-benefit functions have also been used in algorithms applied to real catalogs [*Molchan and Kagan*, 1992]. For our purposes the behaviors of the probability- and rate-based functions are entirely similar.

There are three outcomes which we will include in Q: we define p_1 to be the fraction of large events successfully predicted, p_2 to be the fraction of the total observation time for which alarms are declared, and p_3 to be the fraction of the total number of alarms that are issued which turn out to be false. The penalty for false alarms is necessary to avoid the pathological behavior referred to above in which the solution to the optimization problem is ill-defined and leads to "flickering" alarms which are alternately turned off and on at very short intervals in space and time.

Finally, we consider the coefficients A_i. Since multiplying Q by a constant does not change our ultimate

conclusions, we can normalize Q so that $A_1 = 1$. Thus

$$Q = p_1 - |A_2|p_2 - |A_3|p_3 \quad (6)$$

where the coefficient of p_1 is positive, since it is a benefit, and the coefficients of p_2 and p_3 are negative, since they are costs. Taking the optimal values for each of the outcomes (i.e., successful prediction of all of the events $p_1 = 1$, with negligible costs $p_2 = p_3 = 0$) we obtain the upper bound $Q \leq 1$. Alternatively, if one simply does nothing then $p_1 = p_2 = p_3 = 0$, and we obtain $Q = 0$. We thus use Q to say something interesting about an algorithm: given the costs which have been specified, are there outcomes of the algorithm having $Q > 0$; that is, are there strategies that are better than doing nothing?

The coefficient $|A_2|$ measures the cost of maintaining an alarm. Clearly, as $|A_2| \rightarrow 0$ the best strategy is to leave the alarm on all the time, in which case $Q = 1 - |A_2|$. Similarly, as $|A_2|$ increases, eventually the best strategy will be to do nothing, in which case $Q = 0$. Here we will take $|A_2| = 1$ so that only algorithms which do not maintain a constant state of alert will have a chance of doing better than doing nothing at all.

The parameter $|A_3|$ sets a tolerance on the number of false alarms that may occur. In some cases one might be willing to tolerate quite a few false alarms, as in, for example, the case of short term prediction, where the relative cost of issuing a false alarm is much less than the losses that might be spared by successfully predicting a large earthquake. In comparison, for intermediate-term prediction one might have a lesser tolerance for false alarms because of the expense associated with maintaining a state of readiness over extended time periods. In the sections that follow we consider some of the implications of different choices of this coefficient on, for example, the optimization procedure. In most cases for simplicity we will take $|A_3| = 1$, and while we expect that this value is somewhat larger than the value which would be used in practice, we find that even this choice is not unreasonable, as it leads to a false alarm rate in the intermediate-term algorithms which is comparable to those obtained in algorithms CN and M8.

4. Long-Term Prediction: Results Based Upon Recurrence Intervals

Long-term prediction methods are used to estimate earthquake hazards on a timescale of the order of tens of years. The simplest such schemes make use of only the magnitude or time of occurrence of the last large event and are referred to as the time-predictable and slip-predictable models for hazard assessment [*Shimazaki and Nakata*, 1980]. Various applications of both time-predictable [*Scholz*, 1985; *Bakun and McEvilly*, 1984] and slip-predictable [*Kiremidjian and Anagnos*, 1984] models to real faults have been made. Historical records and geological information are used to construct plots of accumulated slip as a function of time for a particular fault or fault segment. The analogous plot is constructed for the UBK model in Figure 3, where the results correspond to the accumulated slip for one representative patch along the fault over a time interval which is long compared to available catalogs for real earthquakes, but short compared to the catalogs which we consider later for the UBK model. Essentially all of the slip is associated with large events, which is a feature that is common to both the UBK model and real faults. In the slip-predictable scenario, the magnitude of the coming large event is correlated with the time since the last large event. Here the basic assumption is that were an event to occur today it would relieve all of the accumulated strain. When this is valid the upper corners of the staircase in Figure 3 would fall on a line. In the time-predictable case, the time interval preceding the coming large event is correlated with the magnitude of the last large event. In this case, the assumption is that there is some roughly constant threshold which the local stresses must achieve before a large event will be triggered, and the system must reaccumulate a slip deficit comparable to that which was relieved in the last large event before the next event will be triggered. If this were valid, the lower corners of the staircase in Figure 3 would fall on a line. The best linear least squares fits to both models are shown in Figure 3. The time-predictable model works somewhat better than the slip-predictable model on this short UBK model catalog, though neither method works well.

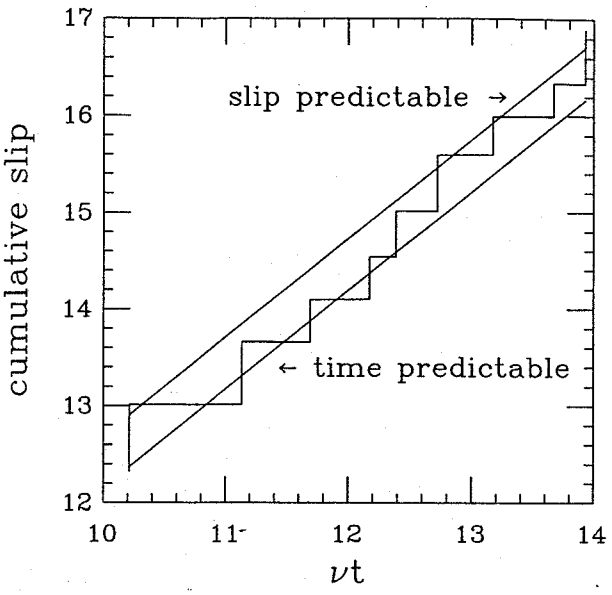

Figure 3. Cumulative slip as a function of time for a representative patch along the fault. While the small events are by far the most frequent, all of the visible displacement is associated with large events. The lines illustrate the best least squares fits to the upper and lower corners of the staircase and thus represent fits to the slip-predictable and time-predictable models, respectively. Clearly, the sequence of events is not periodic, and although neither the slip-predictable model nor the time-predictable model works well, the time-predictable model is somewhat better.

To confirm this more generally, in Figure 4 we test the

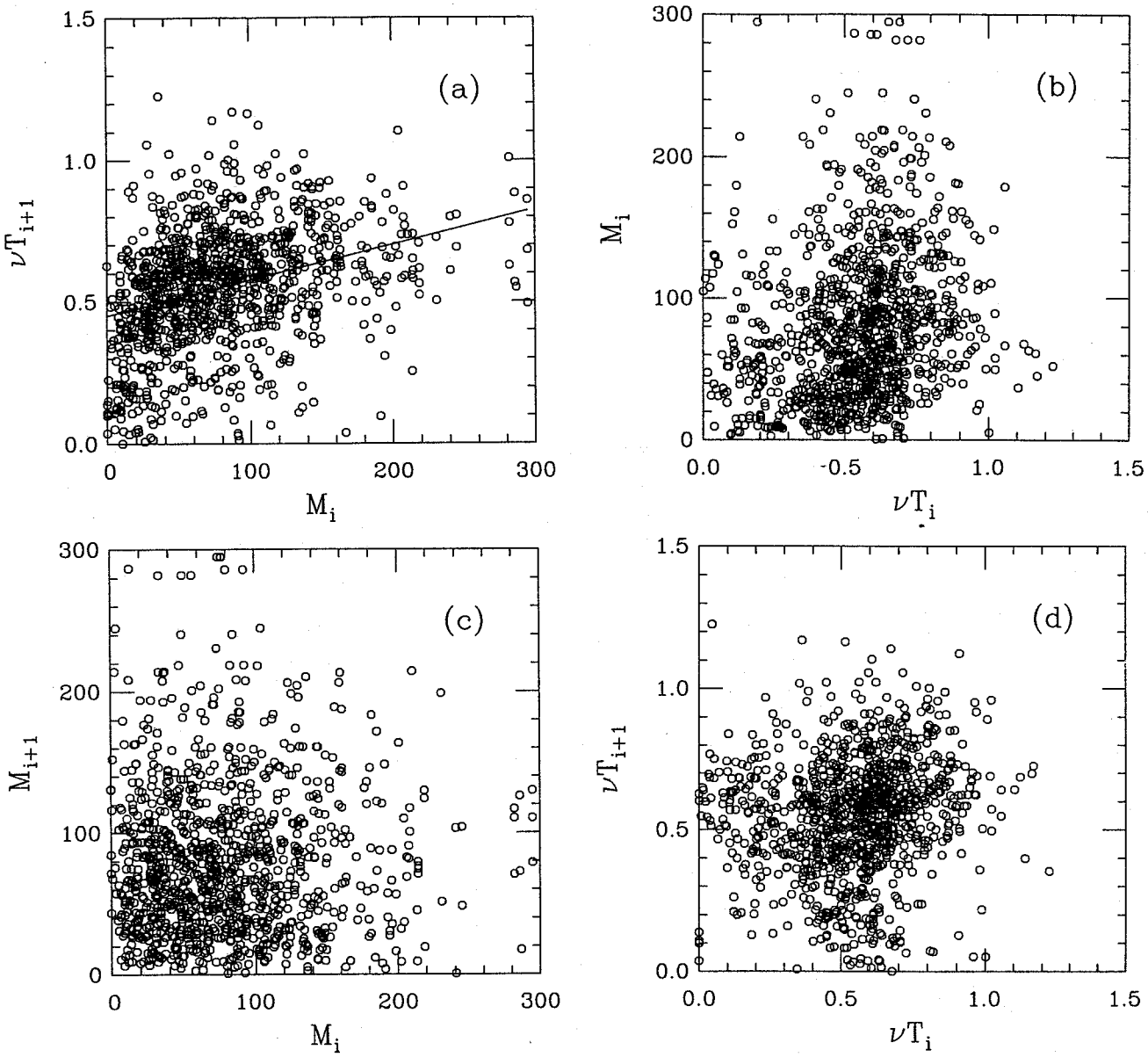

Figure 4. Tests for correlations between seismic moments and time intervals for subsequent large events. Here T_i and T_{i+1} are the time intervals preceding large subsequent events of moment M_i and M_{i+1}, respectively. Pairs of events in which the epicenters lie within $\tilde{\xi}$ of one another are considered. Figure 4a illustrates a test of the time-predictable model which would lead to collapse of the data for T_{i+1} versus M_i onto a line. Figure 4b illustrates a test of the slip-predictable model which would lead to a collapse of the data for M_i versus T_i onto a line. In Figures 4c and 4d we test for correlations between M_i and M_{i+1} and T_i and T_{i+1}, respectively. While none of the graphs show significant correlations, compared to the others, Figure 4a minimizes the least squares deviation to the best linear fit, which is shown. However, our primary conclusion from these data is that while short catalogs may sometimes suggest some correlation such as those tested here, none of these prediction schemes is reliable even for the uniform Burridge-Knopoff model.

time-predictable (Figure 4a) and slip-predictable (Figure 4b) models on a much longer artificial catalog. We also test for correlations between the moments of subsequent large events (Figure 4c) and subsequent time intervals (Figure 4d). A strict adherence to either the time-predictable or slip-predictable model would lead to a concentration of points along a straight line in Figure 4a or 4b, respectively. The errors associated with a linear least squares fit to the data indicate that the strongest correlation is observed in Figure 4a which implies that the time-predictable model works best. In fact, none of the others shows any significant correlation at all. In view of the threshold dynamics which govern the UBK model, it is not surprising that the time-predictable model better approximates the behavior we observe. What is more surprising is that even in

this case the correlation is very weak, as indicated by the broad scatter in the data. Significantly, however, for short enough catalogs one may find much better correlation with one of the above long-term prediction models than actually exists when a sufficient amount of data has been taken into account. For real faults, the case is indeterminate due to the few number of data points available for any individual region, as well as problems associated with accurate slip calculation during an event (see, for example, *Thatcher* [1984]). The most reliable data seem to favor time-predictability, however, and the US Geological Survey (USGS) [*WGCEP*, 1988, 1990] incorporates this model into its long-term seismic hazard estimates.

Another method of long-term prediction which has been studied extensively in connection with real earthquakes is the use of probability distributions of recurrence times for large earthquakes on individual faults or fault segments. The difficulty in this method lies in determining the correct distribution, given the sparsity of data for large events on a given fault. By combining data from many different faults, *Nishenko and Buland* [1987] obtained a reasonably good fit to a lognormal distribution. Others [*McNally and Minster*, 1981] have argued that a Weibull distribution is most appropriate. While it is unlikely that the distribution will ever be known exactly, a better understanding of the constraints would be useful because the hazard assessments often rely on features which are several standard deviations away from the mean repeat time. In fact, *Davis et al.* [1989] showed that, given the sparsity of data, and the uncertainty in the recurrence time interval distribution, large deviations can substantially alter the estimates for future earthquake hazards.

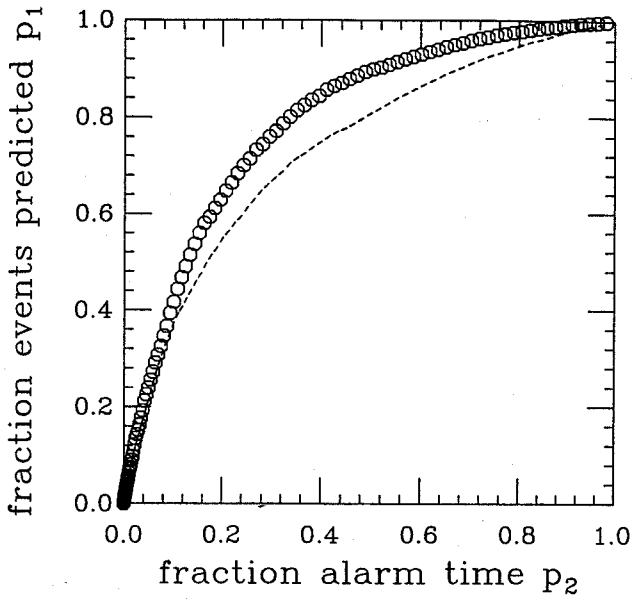

Figure 6. Success curve for time interval prediction. The octagons correspond to local predictions, while the dashed line corresponds to predictions made in spatially coarse grained windows of size $3\tilde{\xi}$ (we show the second curve for comparison to the intermediate-term algorithms). Local predictions are optimal for long-term forecasts, and predicting 80% of the large events requires that alarms be declared roughly 35% of the time.

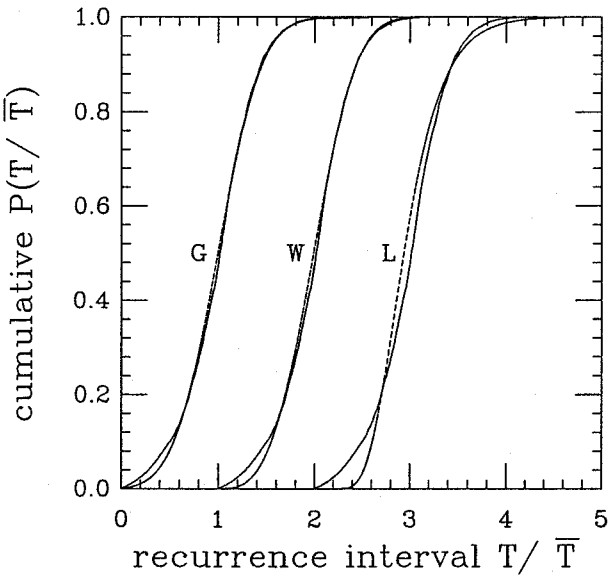

Figure 5. Local recurrence intervals for large events on the uniform Burridge-Knopoff model, scaled by the mean $\overline{T} = 0.558/\nu$. The figure represents the cumulative distribution along with the best fits to Gaussian (G), Weibull (W), and lognormal (L) distributions (shifted for convenience). Comparison of the root-mean-square deviation for each of the fits indicates that the Gaussian (omitting unphysical negative intervals) provides the best fit, followed closely by the Weibull, and, finally, by the lognormal.

In contrast, the corresponding distribution can be determined to arbitrary accuracy for a model such as the UBK model. This was done in the case of a short fault (in which large events spanned the entire system) by *Carlson* [1991]. The cumulative time interval distribution is illustrated in Figure 5 for the long fault catalog that we consider here. The best fit to Gaussian (restricted to positive time intervals), Weibull, and lognormal distributions are shown for comparison. Both the Gaussian and Weibull fit the distribution reasonably well, however, the Gaussian does slightly better. This is primarily due to the fact that the Gaussian somewhat better approximates the nonnegligible weight at very short times in Figure 5 which arises from temporal correlations between large events in neighboring (and, in fact, slightly overlapping) regions. By comparison the lognormal provides a substantially worse fit. For the UBK model the standard deviation for each fit is of the order of the mean repeat time, with $\sigma/\overline{T} \approx 0.36$ for the model. This lies between the ratio of $\sigma/\overline{T} \sim 0.75$ obtained by *Ward* [1992] for a different model and the intrinsic spread of $\sigma/\overline{T} \sim 0.21$ proposed by *Nishenko and Buland* [1987] based on sparse data from real faults. Our results are comparable to the value of $\sigma/\overline{T} \sim 0.4$ which was ultimately used in the *WGCEP* [1988, 1990]

forecasts, though there the amount in excess of 0.21 was ascribed to limited observations rather than intrinsic spread as in the case of the UBK model.

In order to quantitatively compare our results for time interval prediction with the corresponding results for the intermediate-term prediction algorithms discussed in the next sections, it is useful to evaluate this scheme in terms of the success curve discussed in section 3. In the most conventional method of time interval based prediction, given some presumed distribution and known time since the last large event, the probability of a large event in, say, the next 30 years is estimated [*WGCEP*, 1988, 1990]. In contrast, here we specify a threshold time t_0 since the last large event in a region and turn on an alert, or TIP, in that region once a time t_0 has passed. The TIP is turned off once the large event has occurred. For each value of t_0 we then calculate the fraction of earthquakes predicted and the fraction of time the TIP was on. In Figure 6 the resultant success curve is shown. The data points correspond to predictions which are made independently for each position in space and may be calculated directly from the distribution of time intervals. Because these results are clearly an improvement over random prediction, we expect this method ultimately will play a role in the optimal prediction algorithm for the UBK model.

5. Intermediate-Term Prediction: The Pattern Recognition Algorithms

Intermediate-term prediction algorithms are used to make earthquake hazard assessments on the timescale of 1 to 5 years. Because forecasts are made on relatively short timescales compared to long-term prediction, more detailed information about the local state of the system must somehow be deduced. Regional small and medium size events provide one possible probe. If a fault or fault segment is near the threshold for slipping then one might expect that the small scale seismicity would also reflect the fact that the system is close to an instability. This sort of behavior occurs more generally in a wide variety of complex systems. For example, in laboratory fracture experiments, the rate of microcrack production accelerates prior to material failure [*Mogi*, 1962]. As stated previously, in the UBK model we observe an increase in the rate of small to medium size events prior to a large event. While a local increase in seismicity has been observed prior to some large earthquakes, and is in part the basis of some intermediate-term forecasts of, for example, the Loma Prieta earthquake [*Keilis-Borok et al.*, 1990c], this behavior is much less systematic in Earth. In fact, in some cases a local decrease in small scale seismicity, or quiescence, is observed prior to a large event [*Wyss*, 1985], and in others no change in the local rate of seismicity is observed [*Kanamori*, 1981]. Herein lies the difficulty of intermediate-term prediction. The complexity of the earth yields many different patterns so that it is difficult, perhaps impossible, to look at one specific measure to make forecasts worldwide.

For that reason, *Keilis–Borok and Rotwain* [1990] and *Keilis-Borok and Kossobokov* [1990] have developed pattern recognition algorithms, such as CN and M8, which can be applied to interpret objectively the seismicity patterns in earthquake catalogs. Earlier model studies were reported by *Gabrielov et al.* [1990]. In these algorithms as many as 9 different possible precursory phenomena are considered, and each precursor casts a vote as to whether or not an alarm should be turned on. The hope is that if a large event is not preceded by a signal in a particular precursor, then it might be preceded by a signal in another. Algorithms CN and M8 use seven precursor functions and yield a success ratio (equation (4)) of roughly $S = 4$. In particular, roughly 80% of events are successfully predicted, when the TIPs are on 20% of the time. In comparison, the value of S for an individual precursor function is typically of the order of S =2 or 3, corresponding to successful predictions of only 40% to 60% when the alarm time is 20%. While these values exceed what one would obtain for purely random prediction $S = 1$, alarm times which are a smaller fraction of the seismic cycle would be more useful for intermediate-term predictions.

On one hand, algorithms such as CN and M8 are the best candidates for objective means to signal alerts, while on another, they remain somewhat speculative because a systematic evaluation using seismicity data is impaired due to the limited amount of available data. In fact, *Dieterich* (in *Healy et al.* [1992]) has suggested that instead of comparing the success of the algorithm to purely random prediction, one should compare to the success obtained when the TIPs are randomly set but with a bias determined by the local rate of large events, thus building in an element of long term prediction. We will see, however, for the model catalogs even in the case of single precursor functions the intermediate-term algorithms do perform well in comparison to long-term prediction methods previously discussed.

The algorithms we use require precursor functions $\{f_i(\Delta s, \Delta t)\}$, where $f_i(\Delta s, \Delta t)$ is the ith precursor function, evaluated in the spatial region Δs (these are large overlapping circles for Earth, and overlapping line segments in the one-dimensional UBK model which we will take to be a distance l apart) during a sliding time window Δt. Throughout this section we will take $\Delta s = 3\tilde{\xi}$ and $\Delta t/\overline{T} = 0.36$ (i.e. time windows which are 36% of the mean recurrence interval). In section 6 we will show that the performance of the pattern recognition algorithms employed here are relatively insensitive to these choices.

The statistics used to evaluate the success curve and quality function Q (see section 3) are compiled individually within each spatial region R, then combined to determine the cumulative result. For each R, the events which are counted in R are those events for which the

epicenter lies in R. A false alarm will be counted if the alarm is turned off before the region in question has contained the epicenter of a large event. Thus false alarms typically arise in two fashions. For precursor function values near threshold, occasionally an alarm may be turned off then reinstated prior to a large event. Alternatively, a region may experience the typical precursory upswelling of activity but then be preempted as the epicenter by an event which is triggered in a nearby region.

For algorithms CN and M8, precursor functions take a discrete set of values simply referred to as high, medium, or low. A high value (the threshold level "high" is determined from existing catalogs) casts a vote in the favor of issuing a TIP. In our case, a threshold divides the precursor function into high–"on"– values and low–"off"– values. The threshold value is set to optimize the Q of the forecast. The most effective precursor functions will minimize the overlap between the distribution of values taken near the time of a large event and the set of values taken over all time. Examples of both the distribution of values just prior to a large event and the background distribution are illustrated in Figure 7 for the activity precursor function f_1. Here activity is defined to be the total number of small to medium size earthquakes, independent of magnitude, within a space-time window (we omit single block events for convenience).

In Figure 7 we see that f_1 is typically small. The large spike at $P(f_1) = 0$ reflects the extended quiescent period which is observed just after a large event in the UBK model. In contrast, the conditional probability of the value of f_1 at the time of a large event $P(f_1|\text{large event})$ has very little weight at $f_1 = 0$, because the neighborhood of the epicenter nearly always exhibits some activity just prior to a large event. In fact, within the specified space-time window and for the parameter values we have taken, the activity before a large event can reach values as high as a few hundred events. From Figure 7 we see that the average activity just before a large event is, for the model parameters used here, roughly $f_1 = 30$. At that point we observe a ratio of approximately 5 between the conditional probability and the background value, suggesting that activity should be a good precursor function.

In addition to (1) activity, we will consider the following precursor functions: (2) rate of change of activity f_2 defined to be the slope of a linear least squares fit to the activity as a function of time within a space-time window, (3) fluctuations in activity f_3 defined to be the root-mean-square deviation from the linear least squares fit used in function 2, and (4) active zone size f_4 defined to be the number of blocks that have slipped within the space-time window, independent of how many times they have slipped. Note that this is the only precursor which is set to zero after a large event occurs (zeroing the activity affects the results very little).

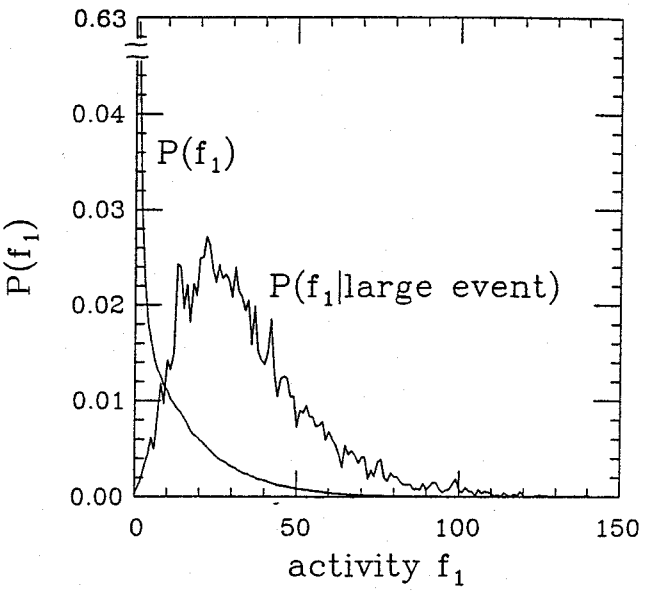

Figure 7. Distribution of values of activity $f_1(\Delta s, \Delta t)$ evaluated in space-time windows $\Delta s = 3\tilde{\xi}$ and $\Delta t/\overline{T} = 0.36$. The background distribution $P(f_1)$ and the distribution of values just prior to a large event $P(f_1|$ large event) are illustrated. The relatively small overlap suggests that activity should be a reasonably good intermediate-term precursor for the uniform Burridge-Knopoff model.

Active zone size is a measure of the extent to which seismicity in a given region is diffuse. It is not simply a size or moment weighted activity measure, nor is it a measure of clustering. Instead, it is more directly a measure of the broadening of small to moderate size events, which leads to the development of a nucleation region associated with a coming large event. To measure active zone size in the Earth we might use a box counting algorithm, in which the large spatial regions taken in algorithms CN and M8 would be subdivided into many smaller regions, and the number of these smaller regions containing seismicity at or above a certain level would define the regional active zone size. Of course, in the Earth this measure is somewhat more complex due to the variable complexity of fault networks in different regions. However, properly normalized to account for such differences, it may be worth considering because of the exceptional performance of this measure for the UBK model. Compared to activity (Figure 7) the signal to background ratio is somewhat greater for active zone size and is somewhat less for rate of change of activity and fluctuations in activity.

Next we implement the intermediate-term prediction algorithm individually for each of the above functions. For each precursor we vary the threshold level for signaling a TIP, and we monitor the resulting fraction of the large events ($\mu > \tilde{\mu}$) successfully predicted as well as the fraction of time TIPs were on. This yields the success curves illustrated in Figure 8. Note that for each of the precursor functions the algorithm leads to

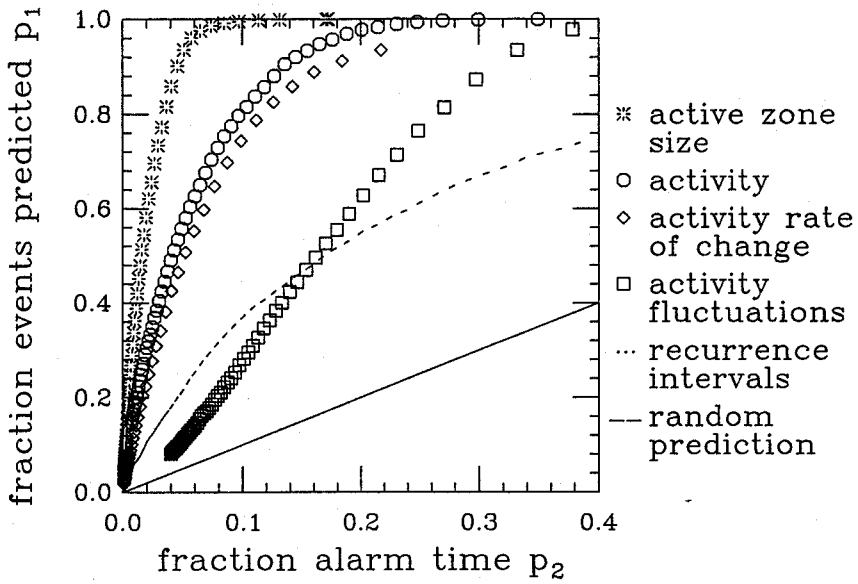

Figure 8. Success curves for intermediate-term precursor functions $f_i(\Delta s = 3\tilde{\xi}, \Delta t/\overline{T} = 0.36)$. All intermediate-term precursors outperform random prediction (solid line) and prediction based on recurrence intervals (dashed line). Active zone size clearly outperforms the other precursors, predicting nearly all of the large events when alarms occupy roughly 5% of the total time.

an enhancement over random prediction and a clear improvement over the results obtained in the last section using long-term techniques. The activity measure f_1, which is most easily interpreted seismologically, yields results for the UBK model which are somewhat better than those quoted above for algorithms CN and M8, and is a significant improvement over the results obtained using single precursors on real catalogs. In comparison to f_1, rate of change of activity f_2 performs somewhat worse, and fluctuations in activity f_3 performs significantly worse. This is not too surprising in light of the fact that in the UBK model the increase in activity is essentially monotonic. What is more surprising is the extent to which active zone size f_4 outperforms the others. In that case nearly all of the events are successfully predicted when the alarms are on only 5% of the time. The corresponding Q curves are illustrated in Figure 9. Note that the ordering of the precursors in terms of effectiveness is the same as for the success curves.

In the UBK model the effectiveness of f_4 as a pre-

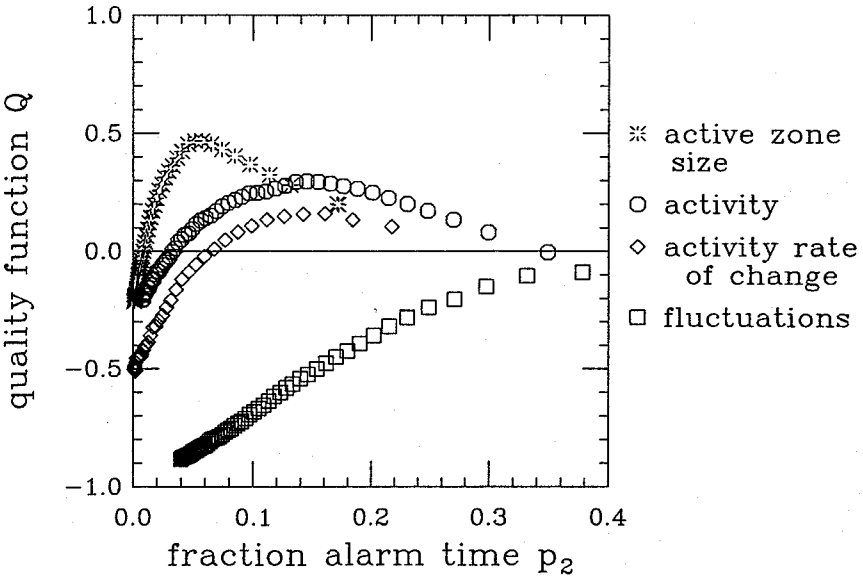

Figure 9. Quality function $Q = p_1 - p_2 - p_3$ for the intermediate-term precursor functions illustrated in Figure 8. Active zone size, activity, and activity rate of change all perform well compared to doing nothing, with their Q values rising above zero by the time alarms occupy 5% of the observation period. Fluctuations in activity almost always has $Q < 0$, indicating that it is a poor measure. Simulations were also performed for the cost-benefit function Q_R with $Q_R = p_1 - p_2 - p_3/3$ where the p_i are taken to be rates, and the results were found to be similar.

cursor can be traced to the fact that most of the energy dissipation occurs during large events. In comparison, small to moderate size events relieve relatively little stress. Thus when a small event has occurred it is a direct signal that the blocks involved are poised at the threshold of instability. For a given region therefore the active zone size is a more direct measure of the density of blocks that are close to threshold than activity, although these two quantities are clearly related. One may think of activity as being an indicator of temporal correlations in the system while active zone size is a measure of the development of spatial correlations. In systems for which behavior on one length scale is correlated with behavior on a larger scale, it would be interesting to compare and contrast the relative roles of spatial versus temporal correlations in determining the predictability of the system. To this end, we are currently investigating the relative performance of measures analogous to activity and active zone size in other models.

It remains a topic of current research to determine the extent to which our results might be improved by combining the different precursors. For example, appropriate combinations of the precursor functions may lead to threshold functions which result in an increase of the number of successful predictions or a reduction in the false alarm rate and the amount of time occupied by alarms. This is illustrated in Figure 10, where each point corresponds to the simultaneous measurement of the activity f_1 and the rate of change of activity f_2 over all time (Figure 10a) and just before a large event (Figure 10b). For the parameter values taken here, Q is optimized for activity f_1 alone when alarms are declared for activity at or above the threshold level $F_1 = 12$. Similarly, for rate of change of activity f_2 alone at $Q = Q_{max}$, alarms are declared at or above the threshold level $F_2 = 120$. At these values of F_1 and F_2 the set of events which are predicted using f_2 is a subset of those which are predicted using f_1. However, by comparing the background and conditional distributions, it is clear that in some cases while f_1 maintains a relatively high value, f_2 has dropped far below the value it takes at the time of a large event. Thus by choosing a threshold function such that alarms are declared, for example, only when the values of both activity and rate of change of activity exceed some specified individual thresholds one might find that the number of false alarms and the total alarm time could be reduced while the number of successfully predicted events might decrease very little. We are currently developing multidimensional optimization techniques with which we may address this problem in more detail, and we hope the results will be useful for further evaluation of the pattern recognition algorithms CN and M8.

6. Intermediate-Term Prediction: Variation of Parameters

In this section we consider the question of optimizing the pattern recognition algorithm studied in section 5. Restricted magnitude windows will also be considered. In particular, we maximize $Q(F_i, \Delta s, \Delta t, \Delta\mu)$, where F_i is the alarm threshold for the ith precursor function

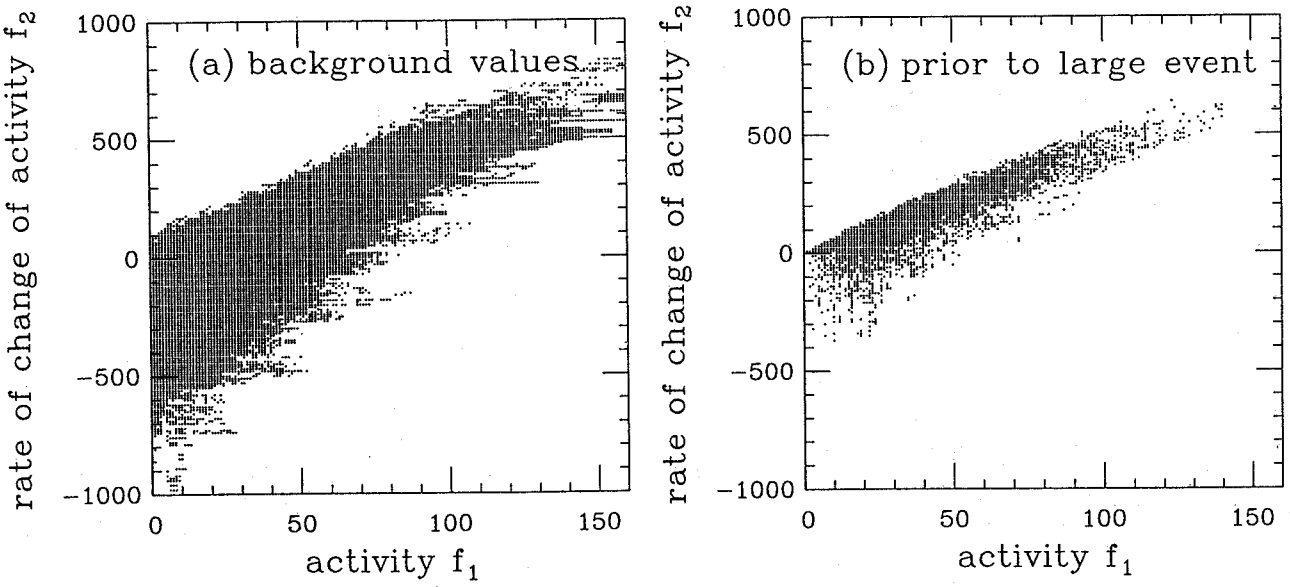

Figure 10. Test for independence of activity f_1 and rate of change of activity f_2. Each point in Figure 10a represents simultaneous values of f_1 and f_2 at unrestricted points in time while the values plotted in Figure 10b are just prior to a large event. These plots constitute projections of the joint probability distributions for f_1 and f_2 analogous to those illustrated for f_1 in Figure 8. While the values are clearly correlated, the nonnegligible spread in the joint distribution in Figure 10b suggests that our predictions may be enhanced through a judicious combination of the two functions.

$f_i(\Delta s, \Delta t)$, Δs and Δt are the space and time windows within which the function f_i is evaluated, and $\Delta\mu$ is the magnitude window from which events are taken. Finally, we consider stability of the maximization procedure as a function of catalog length.

Although the four-parameter optimization problem for $Q(F_i, \Delta s, \Delta t, \Delta\mu)$ can certainly be done in a brute force fashion, it seems sensible to maximize Q with respect to F_i, Δs, and Δt simultaneously in order to first optimize the space and time windows. Then the variation of $Q_{\max}$ with $\Delta\mu$ (once again allowing F_i to vary) may be studied separately. Examining such cross sections allows for the future possibility of including the magnitude dependence within the precursor function definitions, as is done in CN and M8.

Below we will restrict our attention to the activity precursor f_1, because it is the most easily interpreted seismologically.

6.1. Optimization With Respect to Space and Time Window Size

The pattern recognition algorithms CN and M8 are typically evaluated with time windows of the order of 5 years and spatial windows of the order of 1000 km (the specific size scales with the moment of the large event). Once a TIP is alerted within a particular space-time window, the system is sometimes reevaluated within the original window on a more finely coarse grained scale. While these secondary predictions are somewhat less reliable and do not work well when applied independently, on occasion they do serve to more closely pinpoint the target region for a coming large event.

One interesting and somewhat surprising feature of CN and M8 is the large spatial window sizes that are used for the primary predictions. Spatial boxes are typically set to be an order of magnitude larger than the target event. The notion that correlations might extend over anomalously broad regions is akin to the ideas of self-organized criticality introduced by *Bak et al.*, [1987], which postulate that a large class of driven dissipative systems may be attracted to dynamical states which display large correlations reminiscent of equilibrium critical points and that this behavior arises due to instabilities associated with threshold dynamics. Such long-range correlations are typically indicated by power law frequency spectra, like the Gutenberg-Richter law describing fault systems. However, because statistics associated with individual faults or narrow fault zones do not exhibit power laws over the whole frequency–magnitude spectrum, we do not anticipate that individual faults should be described within the context of a simple critical phenomena theory. It is also for this reason that we would not expect the UBK model to exhibit correlations on scales as large as those used in CN and M8.

In Figure 11 we evaluate $Q_{\max}(\Delta s, \Delta t, \Delta\mu)$ for the activity precursor function $f_1(\Delta s, \Delta t)$ as a function of the spatial Δs and temporal Δt window size. For each

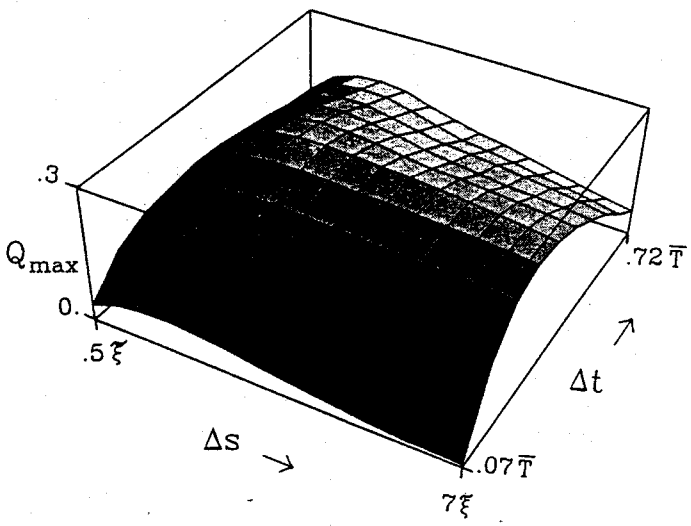

Figure 11. Optimization of space and time windows for the activity f_1. Here $Q_{\max}$ is plotted as a function of Δs and $\Delta t/\overline{T}$. At each space and time window, Q has already been optimized with respect to the threshold value F_1. The maximum with respect to Δs and Δt is broad indicating that the algorithm is robust with respect to variations in these parameters. Variations in $Q_{\max}$ are less than 10% for time windows from ranging from 0.2 to 0.5 of the mean recurrence interval and space windows ranging from $2\tilde{\xi}$ to $7\tilde{\xi}$.

data point, the optimization with respect to threshold level F_1 has already been performed, and we include all of the small and medium size events $\mu_{\min} \le \mu \le \tilde{\mu}$, so that $\Delta\mu = \tilde{\mu} - \mu_{\min}$ with $\mu_{\min}$ fixed at the minimum magnitude of a two-block event (the smallest event we have retained), in our evaluation of the activity. The fact that the surface contains a broad maximum, with no sharp features, is an indication that the algorithm is reasonably robust with respect to variations in the parameters. The fact that the maximum is roughly L-shaped indicates that as long as either the spatial windows or the time windows are taken to be near the optimal size, there is reduced sensitivity to variations in the other parameter. The optimal setting is $\Delta s \approx 3\tilde{\xi}$ and $\Delta t/\overline{T} \approx 0.36$. For these space-time windows, the optimal activity threshold value is $F_1 = 12$, which leads to a value of $p_1 = 0.92$ for the fraction of events successfully predicted, with alarms on $p_2 = 0.14$ of the time, and a $p_3 = 0.48$ rate of false alarms. The value of $Q_{\max}$ is thus $Q_{\max} = 0.30$.

The natural spatial scale for correlations in the UBK model is $\tilde{\xi}$, so it is not surprising that the optimal box size is of the order of this length. Figure 12a illustrates a slice of an optimization surface (Figure 11) along the spatial direction. It is clear that Q rises dramatically up to a length close to $\tilde{\xi}$ then becomes less sensitive beyond that length. In contrast, in fault models which exhibit correlations over lengths up to the system size one might expect the optimal spatial window setting to scale with the size of events one is attempting to predict. As discussed in more detail below, the coefficient $|A_3|$

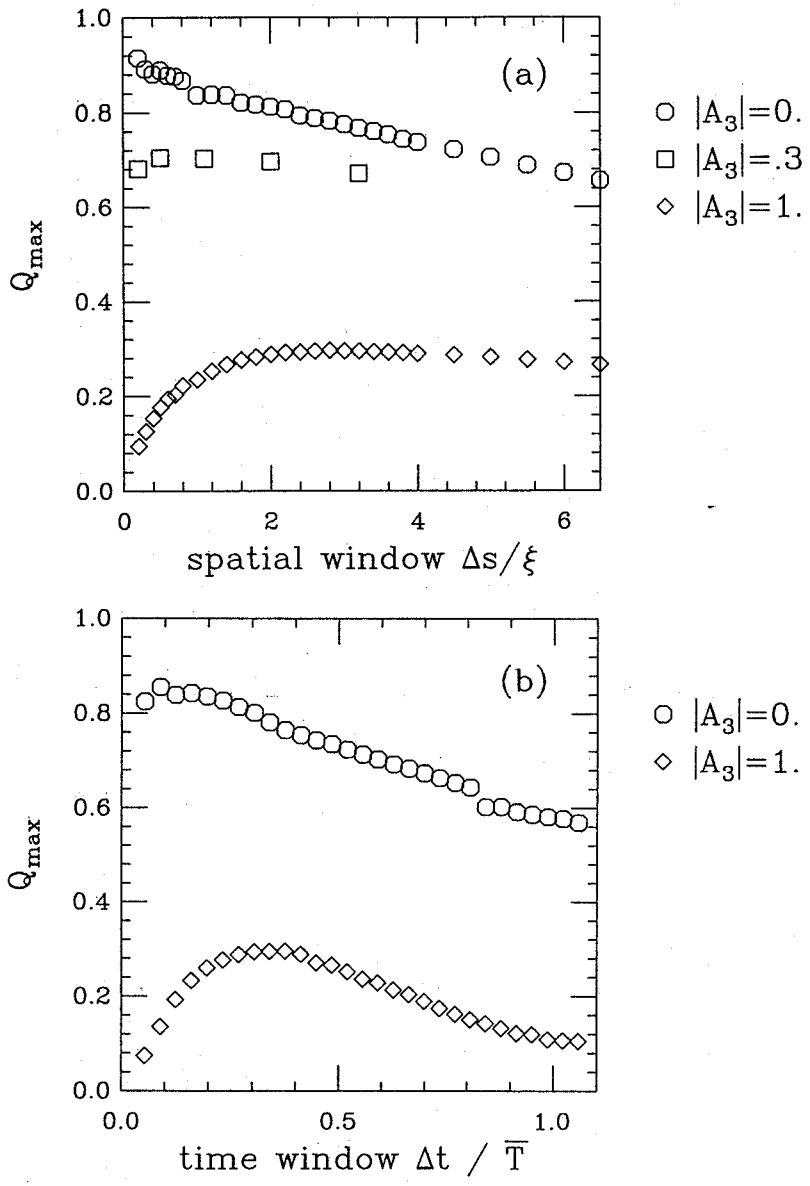

Figure 12. Cross sections for (a) space and (b) time window optimization. Here we illustrate the danger of an injudiciously chosen quality function Q. In each case the lower curve is simply a cross section of the results plotted in Figure 12 ($|A_3| = 1$), while the upper curve corresponds to the results obtained when the penalty for false alarms is not included ($|A_3| = 0$). In Figure 12a the intermediate value $|A_3| = 0.3$ is included to illustrate that the size of the optimal spatial window is sensitive to the choice of this coefficient. The fact that the upper curves ($|A_3| = 0$) are maximized for essentially infinitesimal space and time windows leads to an ill-defined optimization problem.

for the false alarm penalty in Q plays an important role in setting the tolerance for the minimum window size for which reasonable values of Q are obtained.

The optimal time window of $\Delta t/\overline{T} = 0.36$ is relatively large, corresponding to time windows which are 36% of the mean recurrence interval. One can see this in Figure 12b which illustrates a typical slice of the optimization surface (Figure 11) taken along the temporal direction. As observed previously by *Shaw et al.* [1992], there is a significant increase in the rate of change of the cumulative activity after on average 2/3 of the cycle time has passed. Because the activity increase is essentially monotonic, and because we have made the alarm time independent of the time window size, there is no cost for having large time windows. As a result the optimization over time windows is essentially picking out the activity "turn-on" time. Note that the optimal time window is relatively much longer than those employed in algorithms CN and M8 (typically 6 years).

The optimal spatial and temporal windows for a few other values of the UBK model parameters were also evaluated (specifically $l = 10, \alpha = 2$ and $l = 8, 16, \alpha = 3$) and found to be consistent with the above results when expressed in terms of the length $\tilde{\xi}$ (calculated using the relevant l and α values) and the mean repeat time.

Finally, we emphasize that the physically relevant spatial and temporal window sizes which are selected by the optimization were not obtained when we attempted to optimize the algorithm with respect to the success ratio (equation (4)) or a Q function which neglects false alarms ($A_3 = 0$ in equation (6)). For example, in the case of the success ratio S, as the space-time windows were decreased S was found to increase essentially without bound. The analog of this behavior for the Q which neglects false alarms is illustrated in Figures 12a and 12b. As previously mentioned, this occurs because as the activity threshold is increased, the alarm time goes to zero faster than the fraction of events successfully predicted, leading to an ill-defined optimization problem.

The incorporation of a penalty for false alarms leads to large space-time windows and greater reliability of the results. The optimal window size is sensitive to the choice of the coefficient $|A_3|$. For example, we find that the onset of the rapid decrease in $Q_{\max}$ as the spatial window size is reduced (Figure 12a) occurs for windows which scale approximately linearly with $|A_3|$; we estimate $\Delta s \approx 2|A_3|\tilde{\xi}$. Thus the optimization procedure may lead to smaller windows as the penalty for false alarms is decreased. In fact, in Figure 12a with the choice $|A_3| = 0.3$, the optimal window size is less than $\tilde{\xi}$, which results in preemptive false alarms being issued within the nucleating region (of roughly size $\tilde{\xi}$) of the coming large event, simply because the windows were not taken to be large enough (i.e. the active window did not receive credit for predicting the event since it was too small to contain the epicenter). For this reason $|A_3| \approx 1$ is a natural choice for the probability based Q.

6.2. Magnitude Windows

Next we consider the effect of restricting the range of magnitudes which is used to evaluate f_1. Because they are more frequent, small events have more reliable statistics. However, they also tend to swamp the medium sized events in measures such as total activity. By considering only events larger than specified cutoff magnitude, we determine whether there is a statistically significant increase in the moderate size events apart from that predicted by an extrapolation of the increased rate of smaller events which occurs as precursory phenomena. An increased rate of medium size events prior to a large earthquake has been observed in certain instances in the Earth (see, for example, *Sykes and Jaume* [1990]).

In Figure 13 we illustrate our results for $Q_{\max}$ as a function of the lower magnitude cutoff $\mu_{\min}$. In each case, we use the spatial and temporal windows sizes $\Delta s = 3\tilde{\xi}$ and $\Delta t/\overline{T} = 0.18$. In addition, the upper cutoff $\mu_{\max} = \tilde{\mu}$ is fixed. Variations in $\mu_{\max}$ should not significantly alter the results, because the total activity is dominated by the events at the lower cutoff. From Figure 13 it is clear that for a wide range of smaller values of μ_{min}, $Q_{\max}$ is quite insensitive to the lower magnitude cutoff. In fact, in previous sections we have used this feature to ignore the numerous one-block events in our catalog. In contrast, for large values of $\mu_{\min}$ there is eventually a sharp (linear) decline in $Q_{\max}$, which is dominated by a decrease in the number of large events predicted. It is clear that eventually a decline must be observed once the rate of small to moderate events that will be counted becomes comparable to the overall rate of large events (see Figure 2). This corresponds to $\mu = -7.2$ for the parameters we have chosen. The onset of this behavior occurs for a slightly smaller, but comparable, magnitude $\mu_{\min} = -7.5$. This coincides with a threshold of unity for signaling a TIP. There appears to be a slight maximum in $Q_{\max}$ just prior to the decline. This suggests that in the UBK model there may be some additional precursory feature associated with medium size events and that it may be useful to consider the moderate size events as a precursor separate from the small events as is done in algorithms CN and M8. This feature is more pronounced for time windows which are smaller than the optimal windows for the activity precursor (the figure shows our results for $\Delta t/\overline{T} = 0.18$ rather than the window $\Delta t/\overline{T} = 0.36$ for optimization with respect to activity for this reason) which reflects the fact that the largest precursory event also typically occurs relatively close in time to the main event.

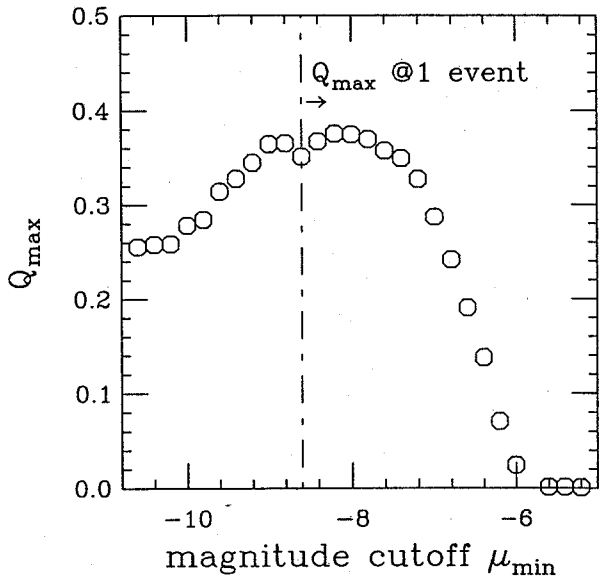

Figure 13. The effects of restricted magnitude windows. Here we plot $Q_{\max}$ for the activity f_1 as a function of lower magnitude cutoff $\mu_{\min}$. The steady rise in Q up to $\mu_{\min} \approx -7.5$ is due to the decreasing number of false alarms. The curve plateaus near where Q is optimized for a threshold value of one event and thereafter is governed by the probability of observing a single event of sufficient magnitude to trigger an alarm before the large event is initiated. The steep decline occurs roughly when the integrated probability of observing an event above the cutoff per unit length per unit time is less than the probability of observing a large event.

6.3. Variation of Catalog Length

Finally, we consider the effects of varying the length of the catalog. Because real catalogs are short compared to the timescale of the seismic cycle (30 years of reliable data on small to moderate size events while the cycle time is of the order of hundreds of years), restrictions on catalog length may significantly constrain the ability of any seismicity based algorithms to predict. Statistics for many large event cycles are needed to gain reliable information about the conditional probability distributions of precursor function values which may indicate an imminent rupture.

For the M8 algorithm, *Keilis-Borok and Kossobokov* [1990] set the parameters using catalogs from different regions which included over 100 earthquakes of magnitude greater than or equal to 8.0. Thus one relevant issue is the ergodicity of the seismological record. The assumption is that averaging over many different earthquake fault realizations in space is equivalent to averaging over a single fault for a much longer time. For the UBK model this reduces to a problem of comparing algorithm selection based upon many short-time catalogs with that obtained at longer times. Since the UBK model is deterministically chaotic, it is not too surprising that we observe ergodic behavior. The interesting question is what minimum catalog length is needed for this ergodic hypothesis to be useful? If the minimum length were much longer than a single cycle, one could not hope to gain reliable predictability from any number of catalogs containing less than one cycle's worth of data.

In order to examine this problem for our system, we consider the distribution $P_T(Q, F_1)$ of values of Q as a function of the threshold F_1 for an ensemble of n catalogs as the length T of each of the catalogs is varied. For each catalog length, we choose F_1 in order to maximize the average Q, defined to be $\overline{Q}_{\max}$, and then for that value of F_1 we also compute the fluctuations: $\sigma^2(Q) = (1/n)\sum_{i=1}^{n}(Q_i^2 - \overline{Q}_{\max}^2)$. In order for optimization of the algorithm over many short catalogs to be a useful procedure one must have available catalogs which are sufficiently long that $\overline{Q}_{\max}$ attains an acceptable value and the width $\sigma(Q)$ of the distribution is relatively small. Here, for convenience, we will set the spatial Δs window to coincide with the optimal value determined earlier in this section while we take $\Delta t/\overline{T} = .18$, which allows for greater range in the catalog lengths which can be compared. We compute the distribution $P_T(Q, F_1)$ by taking a single extremely long catalog and breaking it down into groups of sequential catalogs of shorter length ranging from the time window size to several large event cycles.

Figure 14 illustrates our results for the maximized mean value $\overline{Q}_{\max}$, as well as the standard deviation $\sigma(Q)$ as a function of catalog length T normalized by the mean large event cycle time $\overline{T}$. The behavior of $\overline{Q}_{\max}(T/\overline{T})$ indicates that predictability on the UBK model is poor overall until each short catalog on average encompasses one large event cycle. At this point $\sigma(Q)$ appears to decrease much less rapidly as well. Only minimal improvements are obtained beyond a single large event cycle time. Note that because the fault is sufficiently long that large events do not encompass the entire system, when the catalogs we use extend for

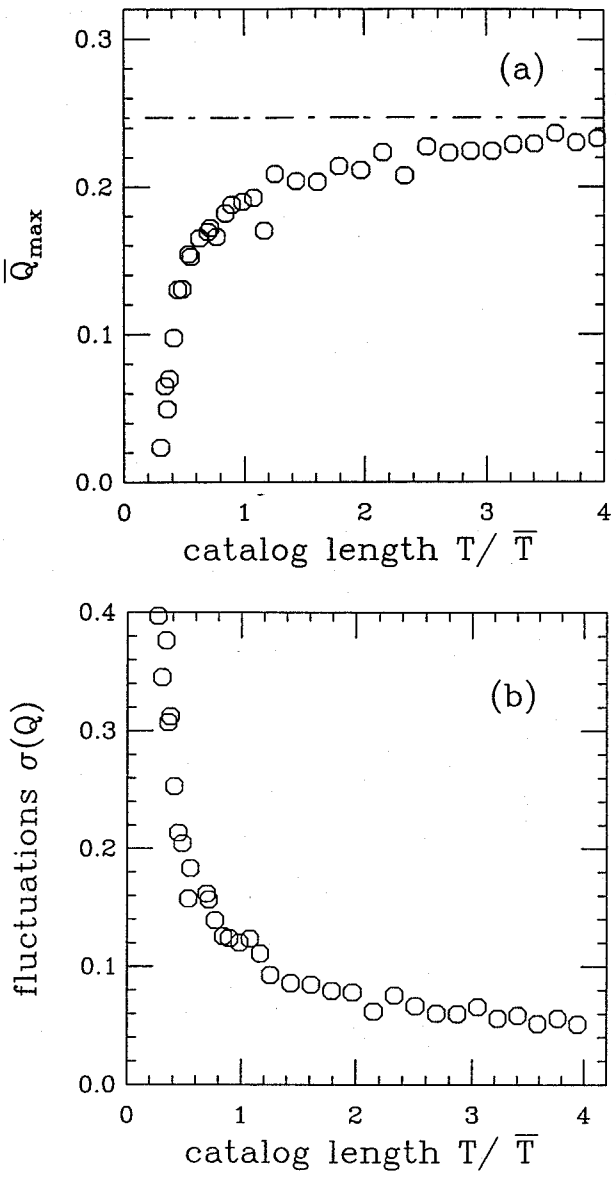

Figure 14. The effects of restricted catalog length. The space and time windows are fixed at $\Delta s = 3\tilde{\xi}$ and $\Delta t/\overline{T} = 0.18$. For each catalog length 100 catalogs were combined and optimized with respect to the threshold F_1, so that F_1 was set by the maximum average $\overline{Q}_{\max}$. In Figure 14a we plot this average value as a function of catalog length. In Figure 14b we plot the width of the distribution of Q values (for the independent catalogs with the threshold fixed at F_1) as a function of catalog length. The mean converges to its optimal value and the width approaches zero once the catalog length becomes roughly of the order of the mean recurrence interval.

one cycle time they will typically contain several large events. We expect that the amplitudes of $\overline{Q}_{\max}$ and $\sigma(Q)$ will thus depend on features such as the system length N, though variations in N should not effect the characteristic time for optimization, i.e. a single cycle time.

In order for a longer catalog on an individual fault to provide information which is not available from combining many shorter time records, the algorithm must make use of correlations which develop on longer timescales. The fact that the fluctuations in Q for the activity precursor die down on the timescale of a single cycle implies that the activity measure does not contain appreciable information which is relevant on timescales longer than the timescale over which activity builds prior to a large event. We suspect it is the same for the other precursors considered in section 5. If such relatively short time correlations are dominant for seismicity on real faults, then it is not too surprising that the pattern recognition algorithms such as M8 may lead to some measurable enhancement over random prediction and long-term techniques. On the other hand, a full cycle's worth of data still corresponds to a much longer catalog than is available for most regions of the Earth. Thus an improvement in the stability of algorithms may still occur as the catalogs lengthen, especially in regions with long repeat times.

In addition to ergodicity, the M8 algorithm assumes self-similarity between faults in different regions so that adjusting for the overall seismicity in an area should be sufficient to allow transfer of an algorithm from one fault to another. The question of self-similarity relates most closely to the variation of model parameters. While we have not yet addressed this issue in great detail, as stated previously it appears that optimization is primarily sensitive to the UBK model parameters ℓ and α through their determination of the length scale $\tilde{\xi}$ and the shape of the main events peak (for α and ℓ sufficiently large). These can be adjusted for by variation of the time and space windows utilized for prediction, and hence we expect will they present no great obstacle to combining UBK model catalogs.

7. Conclusions

We have shown how seismicity catalogs generated from a simple deterministic model of an earthquake fault can provide insights into the problem of prediction. The ample statistics available allow for a thorough study of both long and intermediate-term methods. Below we summarize the major results presented in this paper.

7.1. The Quality Function Q

One methodological question we have addressed is the evaluation of the performance of a prediction algorithm. For this purpose we have employed the function Q which is a linear combination of probabilities relevant to the prediction problem (reflecting the fractions of large events predicted, time occupied by alarms, and false alarms) and was used here to examine optimization questions in some detail. This function is complementary to the success curve and measures the extent to which an algorithm is able to fulfill all of the prediction goals ($Q = 1$) relative to the option of doing nothing at all ($Q = 0$).

7.2. Long-Term Prediction

Our results for predictability based upon the time-predictable and slip-predictable models show that neither describes the UBK fault behavior adequately, although the time-predictable case does appear slightly better. It is interesting to note that for either model it is possible to find good correlation over just a few subsequent large events. At any point, however, such a short-time "pattern" could be broken, with the next event differing dramatically from the proposed model. Hence these sorts of models are certainly not sufficient when used alone for prediction on the UBK model fault. This is significant for time-predictable models applied to real faults where positive correlation is often based on knowledge of only two or three events in a given fault region.

Prediction based upon recurrence intervals of large events was examined and performs reasonably well on the timescales that are relevant for long-term assessments, although these methods are intrinsically limited in effectiveness by the breadth of the distribution of recurrence times. At all times this strategy is better than doing nothing because there are no false alarms in the way we have chosen to implement it for the UBK model. Clearly, this is not the case for the Earth, where, for instance, the more complicated geometry allows the accumulated strain to be relieved along neighboring interacting faults which may not have previously hosted such a large event.

7.3. Intermediate-Term Prediction

intermediate-term precursor functions were studied using a pattern recognition technique similar to those introduced and studied by *Keilis-Borok and Rotwain* [1990] and *Keilis-Borok and Kossobokov* [1990]. Through comparison of the success curves we find that for all precursor functions the pattern recognition algorithms perform significantly better than prediction based on recurrence intervals. When the results for different intermediate-term precursors are compared with one another using the Q functions, it is found that $Q > 0$ when active zone size, activity, and activity rate of change are each employed as individual precursors. Not surprisingly, individual precursors are much more effective at predicting events on the UBK model than in the Earth, and, in particular, the active zone size precursor, which in the earth is analogous to the extent to which seismicity is broadly distributed in space, performs remarkably well. For this function, as well as activity

and rate of change of activity, the results we obtain are comparable and sometimes better, both in terms of the success curve and the value of $Q_{\max}$, than those obtained using the M8 algorithm which employs multiple precursors on real seismicity data.

On a more pessimistic note, using the most standard measure of seismicity, that is, the overall activity (with or without restricted magnitude windows), the intermediate term prediction algorithms studied here do not lead to predictions which are as precise as we had hoped. In particular, relative to the cycle time, the alarm time is still much too large: of the order of 10-20% rather than the 1-5% which is desired. Although we are in the process of examining functions of multiple precursors, it is not obvious that these changes will improve the results substantially. While we suspect that because of its simplicity and the rise in activity which clearly precedes almost every large event, the UBK model will ultimately be more amenable to prediction than will the Earth, this need not a priori be so. Because of the breadth of the distribution of activity values prior to a large event, we suspect some other feature will be necessary in order to make predictions with alarm times which are significantly less. Thus it may ultimately be that the heterogeneity of the earth will lead to phenomena which probe a wider degree of timescales than those which are available in the UBK model and ultimately then also to more precise predictions.

7.4. Algorithm Optimization

In the UBK model, optimization of the Q function for activity with respect to space and time windows leads to selection of lengths which scale with $\tilde{\xi}$ and time windows of the order of 1/3 the mean recurrence interval, which are the length and time scales over which correlations have been previously observed to grow prior to a large event [*Carlson and Langer*, 1989b; *Shaw et al.*, 1992]. These considerations may prove less simple when the algorithms are applied to fault models for which there is no break in the scaling behavior. In such models (see, for example, *Chen et al.* [1991]) typically the dynamics have no inherent length or time scale so the behavior of Q as a function of box size may be very different.

The optimization issue perhaps most relevant to applications of the pattern recognition algorithms to the Earth, is that of stability of prediction for independent data (catalogs not used during the learning procedure). In the model, we find that for activity alone, the algorithm is as adequately optimized on many relatively short catalogs (in which the length is approximately equal to the mean repeat time) as it is on one longer catalog (containing over 1800 large events). If in the Earth correlations in seismicity-based precursors also do not extend in time beyond a single cycle, this result suggests that predictions based on algorithms such as CN and M8 may improve over time as the available catalogs become longer. While, based on the results of the model, the timescale over which improvements might be expected is certainly appreciable (of the order of the seismic cycle), our results do suggest that one need not wait forever to obtain substantial gains.

Acknowledgments. In the course of this work we have profited greatly from discussions with V. Keilis-Borok, A. Gabrielov, D. Turcotte, J. Dieterich, G. Swindle, and especially J. S. Langer. The work of J. M. C. was supported by a grant from the Alfred P. Sloan Foundation, a fellowship from the David and Lucile Packard Foundation, and NSF grant DMR-9212396. The work of S. L. P was supported by an INCOR grant from the CNLS at Los Alamos National Laboratories. The work of B. E. S. was supported by the SCEC grant USC-572726, and USGS grant 1434-93-G-2284. The work of J. M. C and B. E. S. was also supported by the National Science Foundation under Grant PHY89-04035.

References

Aki, K.J., Magnitude-frequency relation for small earthquakes: a clue to the origin of f_{max} of large earthquakes, *J. Geophys. Res., 92*, 1349-1355, 1987.

Archuletta, R.J., E. Cranswick, C. Mueller, and P. Spudich, Source parameters of the 1980 Mammoth Lakes, California earthquake sequence, *J. Geophys. Res., 87*, 4595-4607, 1982.

Bak, P., C. Tang, and K. Wiesenfeld, Self-organized criticality: An explanation of 1/f noise, *Phys. Rev. Lett., 59*, 381-384, 1987.

Bakun W.H. and T.V. McEvilly, Recurrence models and Parkfield, California, earthquakes, *J. Geophys. Res., 89*, 3051-3058, 1984.

Bakun, W.H., C.G. Buffe, and R.M. Stewart, Body wave spectra of central California earthquakes, *Bull. Seismol. Soc. Am., 66*, 363-84, 1976.

Burridge, R., and L. Knopoff, Model and theoretical seismicity, *Bull. Seismol. Soc. Am., 57*, 3411-3471, 1967.

Carlson, J.M., Time intervals between characteristic earthquakes and correlations with smaller events: An analysis based on a mechanical model of a fault, *J. Geophys. Res., 96*, 4255-4267, 1991.

Carlson, J.M., J.S. Langer, B.E. Shaw, and C. Tang, Intrinsic properties of a Burridge-Knopoff model of an earthquake fault, *Phys. Rev. A, 44*, 884-897, 1991.

Carlson, J.M., and J.S. Langer, Properties of earthquakes generated by fault dynamics, *Phys. Rev. Lett., 62*, 2632-2635, 1989a.

Carlson, J.M., and J.S. Langer, Mechanical model of an earthquake fault, *Phys. Rev. A, 40*, 6470, 1989b.

Chen, K., P. Bak, and S. Obukov, Self-organized criticality in a crack-propagation model of earthquakes, *Phys. Rev. A, 43*, 625-630, 1991.

Davis, P.M., D.D. Jackson, and Y.Y. Kagan, The longer it has been since the last earthquake, the longer the expected time till the next?, *Bull. Seismol. Soc. Am., 79*, 1439-1456, 1989.

Davison, F.C., Jr., and C.H. Scholz, Frequency-moment distribution of earthquakes in the Aleutian Arc: A test of the characteristic earthquake model , *Bull. Seismol. Soc. Am., 75*, 1349-1362, 1985.

Dieterich, J.H., An alternate null hypothesis, *U.S. Geol. Surv. Open File Rep., 88-398*, 1992.

Gabrielov, A.M., T.A. Levshina, and I.M. Rotwain, Block

model of earthquake sequence, *Phys. Earth Planet. Inter., 61*, 18-28, 1990.

Healy, J.H., V.G. Kossobokov, and J.W. Dewey, A test to evaluate the earthquake prediction algorithm M8, *U. S. Geol. Surv. Open File Rep., 92-401*, 1992.

Kanamori, H., The nature of seismicity patterns before large earthquakes, in *Earthquake Prediction-An International Review*, Maurice Ewing Ser., vol. 4, edited by D.W. Simpson and P.G. Richards, pp. 1-19, AGU, Washington, D.C., 1981.

Keilis-Borok, V.I., and I.M. Rotwain, Diagnosis of time of increased probability of strong earthquakes in different regions of the world: Algorithm CN, *Phys. Earth Planet. Inter., 61*, 57-72, 1990.

Keilis-Borok, V.I., and V.G. Kossobokov, Premonitory activation of earthquake flow: Algorithm M8, *Phys. Earth Planet. Inter., 61*, 73-83, 1990.

Keilis-Borok, V.I., L. Knopoff, V.G. Kossobokov, and I. Rotwain, Intermediate-term prediction in advance of the Loma Prieta earthquake, *Geophys. Res. Lett., 17*, 1461-1464, 1990.

Kiremidjian, A.S., and T. Anagnos, Stochastic slip-predictable model for earthquake occurrences, *Bull. Seismol. Soc. Am., 74*, 739-755, 1984.

Langer, J.S., Models of crack propagation, *Phys. Rev. A, 46*, 3123-3131, 1992.

Langer, J.S., and C. Tang, Rupture propagation in a model of an earthquake fault, *Phys. Rev. Lett., 67*, 1043-1046, 1991.

Malin, P.E., S.N. Blakeslee, M.G. Alvarez, and A.J. Martin, Microearthquake imaging of the Parkfield asperity, *Science, 2 44*, 557-559, 1989.

McNalley, K.C., and J.B. Minster, Nonuniform seismic slip rates along the Middle American Trench, *J. Geophys. Res., 86*, 4949-4959, 1981.

Minster, B., and N. Williams, 1993 (private communication).

Mogi, K., Study of elastic shocks caused by the fracture of heterogeneous material and its relation to earthquake phenomena, *Bull. Earthquake Res. Inst., Univ. Tokyo, 40*, 125-173, 1962.

Molchan, G.M., Structure of optimal strategies in earthquake prediction, *Tectonophysics, 193*, 267-276, 1991.

Molchan, G.M., and Y.Y. Kagan, Earthquake prediction and its optimization, *J. Geophys. Res., 97*, 4823-4838, 1992.

Nishenko, S.P., and R. Buland, A generic recurrence interval distribution for earthquake forecasting, *Bull. Seismol. Soc. Am., 77*, 1382-1399, 1987.

Pacheco, J.F., C.H. Scholz, and L.R. Sykes, Changes in frequency-size relationship from small to large earthquakes, *Nature, 235*, 71-73, 1992.

Scholz, C.H., Earthquake prediction and seismic hazard, *Earthquake Predict. Res., 3*, 11-23, 1985.

Schwartz, D.P., and K.J. Coppersmith, Fault behavior and characteristic earthquakes: Examples from the Wasatch and San Andreas fault zones, *J. Geophys. Res., 89*, 5681-5698, 1984.

Shaw, B.E., Moment spectra in a simple model of an earthquake fault, *Geophys. Res. Lett., 20*, 643-646, 1993a.

Shaw, B.E., Generalized Omori Law for aftershocks and foreshocks from a simple dynamics, *Geophys. Res. Lett., 20*, 907-910, 1993b.

Shaw, B.E., J.M. Carlson, and J.S. Langer, Patterns of seismic activity preceding large earthquakes, *J. Geophys. Res., 97*, 479-488, 1992.

Shimazaki, K., and T. Nakata, Time predictable recurrence model for large earthquakes, *Geophys. Res. Lett., 7*, 279-282, 1980.

Sykes, L.R., and S.C. Jaume, Seismic activity on neighboring faults as a long-term precursor to large earthquakes in the San Francisco Bay area, *Nature, 348*, 595-599, 1990.

Thatcher, W., The earthquake deformation cycle, recurrence, and the time- predictable model, *J. Geophys. Res., 89*, 5674-5680, 1984.

Vasconcelos, G.L., M.S. Vieira, and S.R. Nagel, Phase transitions in a spring-block model of earthquakes, *Phys. A, 191*, 69-74, 1992.

Ward, S.N., An application of synthetic seismicity in earthquake statistics: The Middle America Trench, *J. Geophys. Res., 97*, 6675-6682, 1992.

Wesnousky, S., C.H. Scholz, K. Shimazaki, and T. Matsuda, Earthquake frequency distribution and the mechanics of faulting, *J. Geophys. Res., 88*, 9331-9340, 1983.

Working Group on California Earthquake Prediction (WGCEP), Probabilities of large earthquakes occurring in California on the San Andreas fault, *U.S. Geol. Surv. Open File Rep., 88-398*, 1988.

Working Group on California Earthquake Prediction, Probabilities of large earthquakes in the San Francisco Bay Region, California, *U.S. Geol. Surv. Circ., 1053*, 1990.

Wyss, M., Precursors to large earthquakes, *Earthquake Predict. Res., 3*, 519-543, 1985.

S. L. Pepke and J. M. Carlson, Department of Physics, University of California, Santa Barbara, CA 93106.

B. E. Shaw, Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964.

(Received May 10, 1993; revised September 10, 1993; accepted October 29, 1993.)