JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 102, NO. B8, PAGES 17,785–17,795, AUGUST 10, 1997

Techniques and parameters to analyze seismicity patterns associated with large earthquakes

Mariana Eneva

Department of Physics, University of Toronto, Toronto, Ontario, Canada

Yehuda Ben-Zion

Department of Earth Sciences, University of Southern California, Los Angeles

Abstract. A pattern recognition algorithm is developed to provide potential improvements over existing earthquake prediction practices. The parameters employed in the analysis include degree of spatial nonrandomness in two distance ranges, spatial correlation dimension, spatial repetitiveness of earthquakes with a similar size, average depth of events, time interval for the occurrence of a constant number of events, and ratio of the numbers of events in two magnitude intervals. The parameter temporal variations are compared quantitatively with the time series of large events using a technique of association in time. The significance of the association frequencies is evaluated by comparison with chance associations estimated from corresponding simulated random time series. The developed techniques differ from existing approaches in the following aspects. The parameters here emphasize the spatial distribution of earthquakes. Possible correlations among the parameters are evaluated to determine the final set of parameters to be monitored. Threshold values for the assumed anomalies are chosen with consideration of properties of the available earthquake catalogs, such as the number of large events to be retrospectively predicted. Equal weight is given to both locally high and locally low parameter values. Care is taken to distinguish between anomalies preceding large events and those following previous events. It is shown that the relationship between precursory local extrema and precursory trends is nonunique, with precursory local extrema of the same type frequently associated with opposite observable precursory trends. The application of the seismicity parameters and pattern recognition techniques is demonstrated using synthetic earthquake catalogs generated by models of segmented fault systems in a three-dimensional elastic solid [*Ben-Zion*, 1996].

Introduction

Earthquake prediction is one of the ultimate goals of seismology. Despite works emphasizing the importance of a rigorous quantitative treatment in the search for precursory patterns [e.g., *Matthews and Reasenberg*, 1988; *Habermann*, 1988; *Wyss and Habermann*, 1988; *Rhoades and Evison*, 1989] and the effect of artificial changes in earthquake catalogs [*Habermann*, 1987, 1988], qualitative observations of anomalies preceding isolated large events are still widely reported. The parameters most commonly employed in reports of precursory changes of seismicity are the b value of the frequency-size statistics of earthquakes and seismicity rates. In a number of cases, the presumed variations have been used to draw conclusions about the physical processes associated with the nucleation of larger earthquakes, without taking into account the uncertainties characterizing the observations. In addition, most reports on earthquake precursors involve observations of a single parameter. A newer generation of studies makes use of multiple parameters and pattern recognition techniques which employ relatively well-defined quantitative elements in the search for precursors. Every pattern recognition scheme includes a learning stage using past activity to uncover significant precursory traits and an application stage for future activity. The most elaborate early techniques of this type have been applied in the former Soviet Union for seismic zoning [e.g., *Gvishiani et al.*, 1988, and references therein]. In addition to seismicity parameters, geological and geophysical parameters were also included, describing the topography, fault configuration, crustal structure, and gravitational field in the vicinity of past large earthquakes. The best known contemporary algorithms that emerged from this approach are CN (named after an application to the California-Nevada region) and M8 (named after its first applications to $M \geq 8$ earthquakes) [e.g., *Keilis-Borok and Rotwain*, 1990; *Keilis-Borok and Kossobokov*, 1990]. The M8 algorithm is used currently in forward earthquake prediction tests [*Healy et al.*, 1992]. It incorporates only seismicity parameters and is intended for the purpose of intermediate-term prediction of the time, location, and magnitude of impending large events.

Pattern recognition techniques in current use do not account properly for measurement errors characterizing the earthquake catalogs [*Habermann and Creamer*, 1994]. However, if a reliable data set is used, these techniques can identify statistically significant precursors with much more confidence than any sporadic observations. Other general shortcomings of the existing procedures motivating some of our developments are as follows:

1. The M8 algorithm identifies a certain percentage of the highest values in temporal variations of seven seismicity parameters calculated in overlapping time windows. Using these high values, so-called "times of increased probabilities" (TIPs) for the

Paper number 97JB00994.
0148-0227/97/97JB-00994$09.00

occurrence of large events are identified. The number of TIPs is equal to or smaller than the number of high values considered, depending on whether the high values are separated or grouped in time. This approach offers an unambiguous way for fast, near-real-time applications. However, the following problems can be pointed out: (1) The smallest values may be equally, or more, important than the highest ones. (2) The highest values can occur on prolonged flat portions of the variation curves, which defies the intuitive notion of anomaly. (3) Using time windows of the same length and same arbitrary percentile (e.g., 10%) in different catalogs disregards important aspects of the variability and dynamics of seismicity. Thus the same number of highest values is identified for data covering time periods of the same length in different catalogs, regardless of the fact that the numbers of large events to be predicted retrospectively may vary greatly among catalogs. (4) Determining a certain percentage of the largest values is essentially useless if for any reason a long-term trend, unrelated to the occurrence of large events, is present in the time series, as all TIPs would cluster at the highest level.

2. Concerning parameter choices in M8, all seismicity parameters in the original algorithm are based on activity (counts of events, i.e., seismicity rates). Thus, by default, the parameters must be highly correlated. Interparameter correlations and their effect on the declaration of TIPs have not been clarified. Another set of parameters, based on the so-called active zone size, has been recently introduced in the application of M8 to synthetic data [*Pepke et al.*, 1994] and applied subsequently to real earthquake data [*Kossobokov and Carlson*, 1995]. Still, correlations among the new parameters have not been discussed, even though the correlation with the older, activity-based parameters has been addressed.

3. In the absence of an appropriate physical basis, all quantitative approaches use empirically determined threshold (alarm) values to identify significant anomalies. The threshold values vary in M8 from case to case, but these variations are only determined by the rather rigid requirement for a specific percentage (e.g., 10%) of the highest signals to be considered. In other applications, the variations of significance are left to be determined loosely by whatever level worked in particular cases. For example, 40% to 90% changes in seismicity rates between two time periods have been reported to mark the onset of seismic quiescence in different areas [*Wyss and Habermann*, 1988]. More generally, problems similar to the one described in point 3 of paragraph 1 above characterize these empirical choices. Regardless of the way threshold levels are chosen, it is usually further required to calculate "success rates," "failures to predict," and "false alarms" in order to quantify the portions of events that were predicted, missed, or never occurred [e.g., *Wyss and Habermann*, 1988; *Rhoades and Evison*, 1989]. Other related quantities are the "validity" and "reliability" of precursors [*Matthews and Reasenberg*, 1988]. However, since all of these depend on the threshold levels, they may be less informative than usually assumed.

4. Some authors have reported "low" or "high" precursory values of certain parameters, while others have referred to "increasing" or "decreasing" trends. The relationship among these two approaches has not been clarified. In particular, conflicting trends have been indicated for the same precursors. It has not been discussed to what extent such differences reflect the complexity and variability of the seismic process or whether they are of statistical nature. Indeed, a trend can appear to reverse itself even for the same data set when different windows are used in the calculations. Variations also arise from the intrinsic nonlinearity of the process. This leads to a nonunique relationship between observable types of precursory local extrema and trends, as is shown below. Thus it is important to clarify the implications of using one or the other indicator for the purpose of earthquake prediction.

5. Typically, a distinction is not made between precursory effects that precede a given large event and aftereffects that follow the previous large event. This issue becomes particularly troublesome if the frequency of events to be "predicted" increases, because of either natural variations in activity among regions or the use of lower magnitude thresholds. For example, in a reexamination of the so-called Varotsos, Alexopoulos, and Nomicos (VAN) precursors [e.g., *Varotsos and Lazaridou*, 1991], *Mulargia and Gasperini* [1992] suggested that such signals are more readily found to follow, rather than precede large events (for recent debate on the subject, see articles in the special section "Debate on "VAN,"" in *Geophysical Research Letters*, *23*(11), 1996).

6. Estimates of the probability that large earthquakes and their presumed precursors are associated by chance are also frequently missing from published literature. As an exception, *Wyss and Allmann* [1996] argued that the VAN predictions can be attributed to pure chance. See other papers in the special section "Debate on "VAN"" for related statistical arguments.

The present paper attempts to reduce the problems and inconsistencies mentioned above and to provide better defined concepts and techniques. We develop several analysis steps to improve on current techniques for identifying anomalies, to distinguish between aftereffects and precursors, to evaluate quantitatively the association between anomalies and large events, and to address problems with interpreting precursory trends. The examples here and the analysis in a follow-up paper [*Eneva and Ben-Zion*, 1997] employ synthetic data generated by theoretical fault models [*Ben-Zion*, 1996]. However, the issues listed above are addressed in general ways applicable to real data as well. Results from a partial application to real data from mining-induced seismicity are given by *Eneva* [1997] and *Eneva and Villeneuve* [1997]. While there are no conceptual limitations for the application of our techniques to real data, significant technical limitations, such as poor quality, scarcity, and ambiguity of observed earthquake catalogs, are common.

The models of *Ben-Zion* [1996] simulate seismicity along four cases of segmented fault systems in elastic continua. The techniques introduced in the present work are illustrated using synthetic data from two model realizations; results of application to all the models are reported by *Eneva and Ben-Zion* [1997]. The use of synthetic data has certain advantages relevant to the study of the involved physical processes, as well as to the statistical reliability of the results. This is because the synthetic catalogs, unlike real data, are devoid of man-made errors in the determinations of magnitudes and hypocentral locations, and the synthetic data can be generated for long sequences containing large numbers of events. To evaluate the effects that changes in detection, magnitude shifts, and hypocentral errors encountered in reality would have on our analysis tools, such changes and errors can be incorporated into the synthetic catalogs, using techniques like those applied by *Habermann* [1987], *Eneva et al.* [1994], *Eneva and Pavlis* [1991], and *Kagan and Knopoff* [1991]. However, such an analysis is outside the scope of the present work.

The seismicity parameters used in our work cover the earthquake distribution in space, time, and size. Particular emphasis is placed on quantifying various aspects of the spatial distribution

of hypocenters. The techniques developed include a number of new concepts associated with the search for precursory anomalies and their quantitative relation to the occurrence of large events.

Parameters

The seismicity parameters described here are intended to extract a large amount of quantitative information from data typically included in earthquake catalogs, such as event locations, origin times, and sizes. While these parameters are not necessarily unique, they can be used efficiently for a comprehensive description of the temporal, spatial, and size distribution of earthquakes. The temporal variations of the parameters are examined by calculating their values in time windows covered by overlapping groups of a constant number of events, with specific attention to the temporal changes of the spatial distribution of earthquakes. This emphasis was prompted by some works featuring crustal seismicity [*Eneva and Pavlis*, 1991; *Eneva et al.*, 1992; *Eneva*, 1994a] and mining-induced activity [*Eneva and Young*, 1993; *Eneva*, 1994b, 1996, 1997]. Incorporating seismicity parameters based on interevent distances, these works suggest that the spatial distribution can be at least as informative as the more commonly used seismicity rates and *b* values. In the works mentioned above, increasing spatial clustering of hypocenters was found to be associated with some large crustal events and rockbursts in mines. Unlike the temporal distribution, only in the last two decades did the studies of the spatial distribution of seismicity evolve from simple visual inspections of hypocentral maps and cross sections to a quantitative treatment. In addition to the parameters used in the above references, techniques like the single-link cluster analysis [*Frohlich and Davis*, 1990], investigations based on the so-called Morishita's index [*Ouchi and Uekawa*, 1986], studies using the spatial correlation dimension [e.g., *Hirata*, 1989], and active zone size [*Pepke et al.*, 1994; *Kossobokov and Carlson*, 1995] were all intended to examine the spatial distribution of events in a quantitative way. A number of precursory changes in spatial patterns were reported by these authors. Quantifying the spatial patterns of seismicity is also relevant to works attempting to assess the degree of heterogeneity in the spatial distribution of faults (e.g., the so-called "participation ratio" of *Sornette et al.* [1993]).

In the present work, we suggest the use of seven parameters of which five are spatial. The specific choice of parameters is based on generally expected features of real seismicity. However, a certain autonomy is preserved in the learning stage of the pattern recognition by neither assuming that too much is already known nor using any specific algorithm already in existence. The parameters employed in our work are discussed below.

Degree of Spatial Nonrandomness at Short (*NS*) and Long (*NL*) Distances

These two parameters have been used to study seismic activity in California [*Eneva and Pavlis*, 1991], central Asia [*Eneva et al.*, 1992], and eastern Canada [*Eneva*, 1994a], as well as mining-induced activity [*Eneva and Young*, 1993; *Eneva*, 1997; *Eneva and Villeneuve*, 1997]. They are based on distances between points, employed in the same way as in the two-point correlation function used by *Kagan and Knopoff* [1980] but with the addition of a direct comparison to random simulations. The degree of spatial nonrandomness measures the discrepancy between the frequency distribution of interevent distances and a reference distribution obtained from points simulated using a corresponding uniform random distribution. Because of the use of event pairs, this type of analysis has been referred to as "pair analysis." In the random simulations, the generated points fall with equal probability everywhere in the study volume. Thus the degree of nonrandomness is equivalent to the degree of nonuniformity or the degree of heterogeneity in the spatial distribution of seismic activity. The difference between the actual and the random distributions of interevent distances results in a residual distribution. A typical residual distribution is shown in Figure 1. The degree of nonrandomness for any N points in a volume can be measured over any distance range $[a, b]$ as

$$c_{[a,b]} = \pm(100\%)\sqrt{\left|\sum_{i=i_a}^{i=i_b} \frac{\Delta p_i}{P}\right|}, \qquad (1)$$

where i_a and i_b mark the discrete distance intervals in which a and b fall, Δp_i is the residual number of event pairs in interval i with $i_a \le i \le i_b$, $P = N(N-1)/2$ is the total number of nonrepeating pairs for N events, and the sign depends on whether the anomaly over $[i_a, i_b]$ is represented by excess or deficiency of pairs. The degree of spatial nonrandomness is measured in percent of the equivalent portion of events involved in the respective anomalies. An excess over small distances can be associated with the more intuitive perception of spatial clustering. As Figure 1 demonstrates, excess of pairs can be observed for both small and large distances. Any of the terms "nonrandomness," "nonuniformity," "heterogeneity," or "clustering"/"deficiency" over short or large distances can be applied to describe these anomalies. In this type of analysis, however, excess and deficiency of pairs appear only as relative terms, and deficiency of pairs reflects the relatively higher degrees of clustering in neighboring distance ranges rather than true deficiency of event pairs. The terms "clustering" and "deficiency" therefore may be misleading. Thus the excess of event pairs at relatively small and large distances is referred to as "short-" and "long-distance nonrandomness," respectively. Obviously, the specific distances considered as

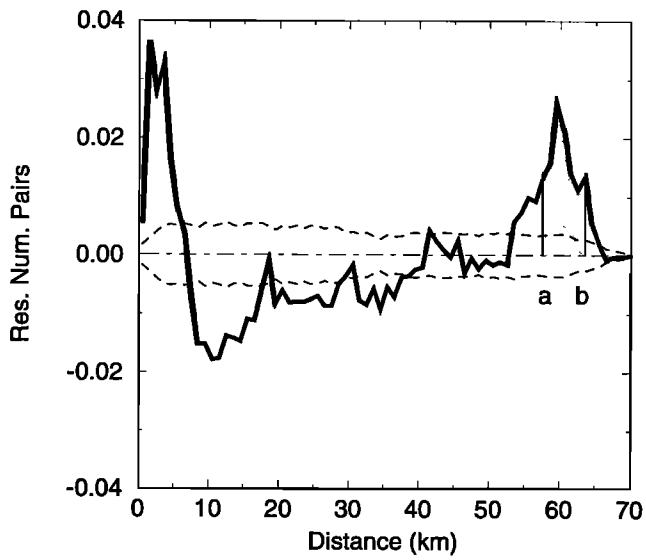

Figure 1. Schematic representation of a residual distribution from pair analysis of interevent distances. Shaded area may be used to measure the degree of spatial nonrandomness over the distance range $[a, b]$. Dashed lines show tolerance limits covering 90% of reference random distributions with 95% confidence [e.g., *Eneva and Pavlis*, 1991].

"short" or "long" depend on the size of the study system. Here and in previous works [e.g., *Eneva and Pavlis,* 1991], distances larger than about two thirds of the maximum linear size of the study area are referred to as "long," while distances smaller than one third of the minimum linear size of the system could be referred to as "short." Examples of distance ranges [*a*, *b*] used for *NS* in various applications of (1) above are [0, 5] km used by *Eneva and Ben-Zion* [1997], [0, 30] km used by *Eneva* [1994a], [0, 2] to [0, 13] km used by *Eneva et al.* [1992] and *Eneva and Pavlis* [1991], and [0, 75] m in the study of mining-induced activity by *Eneva* [1997] and *Eneva and Villeneuve* [1997]. Examples of distance ranges for *NL* are [60, 65] km used by *Eneva and Ben-Zion* [1997], [15, 30] km used by *Eneva and Pavlis* [1991], and [135, 210] m used by *Eneva and Villeneuve* [1997]. In all cases, these ranges include the distances over which significant anomalies, in terms of excess of event pairs by comparison with random distribution, are observed.

Systematic changes in *NS* have been observed to precede and follow some rockbursts in mines and large earthquakes in their aftershock zones (see references above). These variations indicate mainly precursory increase, highest values during the aftershock sequences, and subsequent decrease at the end of the sequences and during the transition to background seismicity. In addition, there has been an observation [*Eneva*, 1994a] of *NS* taking the highest value coseismically to a large event in a neighboring zone. The parameter *NL* appears to exhibit systematic changes as well [*Eneva and Pavlis*, 1991]. In contrast to results from pair analysis of earthquake hypocenters, applications of the same technique to the interevent time intervals did not yield any significant variations [*Eneva*, 1994a].

Spatial Correlation Dimension (*CD*)

The use of correlation dimensions to estimate fractal dimensions was introduced by *Grassberger and Procaccia* [1983]. It has since been widely used in various fields of nonlinear dynamics. Similar to the pair analysis above, applications of correlation dimension analysis to the temporal distribution of events did not suggest any significant variations in time [e.g., *Eneva*, 1994a], while the spatial correlation dimension appeared to be informative. Applications of the correlation dimension analysis to the spatial distribution of earthquakes have been reported by *Hirata* [1989], *Radulian and Trifu* [1991], *Main* [1992], and *De Rubeis et al.* [1993] among others. *Coughlin and Kranz* [1991], *Trifu et al.* [1993], *Eneva and Young* [1993], and *Eneva and Villeneuve* [1997] have used the same method to study the spatial distribution of mining-induced seismicity.

Figure 2 shows a typical correlation integral using the interevent distances $r < R$:

$$C(R) = \frac{p(r < R)}{P} = \frac{2p(r < R)}{N(N-1)}, \qquad (2)$$

where $C(R)$ represents the number of pairs $p(r < R)$ normalized by the total number of pairs $P = N(N-1)/2$ for N events (same as in (1)). The number of pairs in (1) is counted in each distance interval, whereas the number of pairs in (2) is cumulative. For fractal sets, it has been shown that the correlation integral can be represented by a power law $C(R) \sim R^d$ for small reference distances R, where the power is the correlation dimension d. The slope of the straight segment in the log-log plot in Figure 2 can be used to estimate the exponent d; since d is estimated only over the distance range [*A*, *B*] where scaling is observed, it is more appropriately denoted by $d_{[A, B]}$. Examples of the distance range [*A*, *B*] are [5, 10] km used by *Eneva and Ben-Zion* [1997], [1, 10] km used by *Eneva* [1994a], and [30, 120] m used by *Eneva and Young* [1993].

The correlation dimension is inversely related to the degree of spatial clustering, as the slope of the straight segment decreases when the number of event pairs with relatively small interevent distances increases. Precursory changes, mainly of decreasing correlation dimensions (i.e., increasing spatial clustering), have been reported in the above references and other works. It is important to note that one is severely limited by the way the scaling distance range [*A*, *B*] can be chosen, while the range [*a*, *b*] in the pair analysis can be quite arbitrary and does not require the presence of scaling at all. This represents a significant advantage of the pair analysis over the method of correlation dimensions.

Various aspects concerning the choice of most appropriate scaling range, work with small data sets, and the effect of finite size and irregular geometry of study volumes are discussed for spatially scattered data by *Eneva* [1994b, 1996]. The ambiguity due to such effects is addressed through comparison of the estimated correlation dimensions with results from randomly generated data subjected to the same limitations as the real data. Both uniform and monofractal random simulations are used in these works to demonstrate that estimated correlation dimensions are generally biased downward and that apparent multifractality suggested by real data may be only an artifact. A similar approach has been used by *Ouillon et al.* [1996] and *Ouillon and Sornette* [1996], who correct for finite size and irregular geometry in the mutifractal analysis of fault patterns. This is achieved through comparison of estimates from real fault networks with estimates based on reference networks created by assigning the real faults at random to the initial fault domain.

In the present work the scaling range [*A*, *B*] is kept the same for all overlapping groups of events. Its choice is based on the best straight segment from a statistical point of view and the distances of "depopulation" and "saturation," as described in detail by *Eneva* [1996]. It was also shown in that work that small data sets can be effectively used to follow the temporal changes

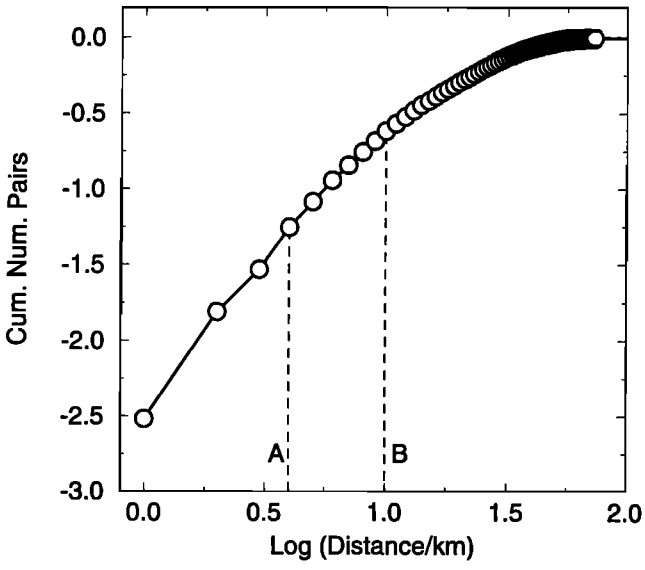

Figure 2. Schematic representation of a correlation integral using interevent distances. The slope of the straight segment over distance range [*A*, *B*] is used to estimate the spatial correlation dimension $CD = d_{[A, B]}$.

of the correlation dimension. This is because typical biases downward in the estimates of the spatial correlation dimensions are related mainly to geometry and not to the number of events, although the range of random fluctuations in the estimates does increase with decreasing number of points. Using a relatively narrow scaling range over which the correlation dimension is estimated represents another source of fluctuations in the estimates from group to group. The comparison with random simulations, however, can provide estimates of both the bias from the true fractal dimension and the range of random fluctuations. As an example, the width of the range covered by the *CD* values calculated for groups of 100 events each given by *Eneva and Ben-Zion* [1997] is 0.67 to 0.87 for the various fault models. In contrast, the range covered by the correlation dimensions evaluated from corresponding randomly simulated data is less than half of these values. Thus significant temporal variations in *CD* estimated from the seismicity along the fault models can be readily assumed, despite variations due to random fluctuations and instability of the estimates.

Degree of Spatial Repetitiveness (*SR*)

This parameter represents the number of event pairs for which

$$x \le x_0;\ |\Delta M| \le m_0,$$

where Δx and ΔM are the interevent distance and magnitude difference, respectively. The values of x_0 and m_0 are chosen to be small, so that this parameter reflects the occurrence of events of similar size at essentially the same site. Obviously, some of the information contained in *SR* is already included in two of the three parameters introduced above, *NS* and *CD*. Note, however, that both *NS* and *CD* involve much larger distance ranges and do not contain information on the size of earthquakes. The use of *SR* is motivated by recent reports on repeating small events in spatial clusters on the Parkfield segment of the San Andreas fault [*Nadeau et al.*, 1994, 1995], suggesting that decrease in clustering may precede large events. In addition, *SR* can be closely connected to another spatial parameter, active zone size, that has been incorporated recently in the M8 algorithm for earthquake prediction [*Pepke et al.*, 1994; *Pepke and Carlson*, 1994; *Kossobokov and Carlson*, 1995]. This parameter appears to perform better than the original M8 parameters based only on activity (counts of events). The relation between the active zone size and *SR* is generally inverse, as larger values of the degree of spatial repetitiveness imply a smaller degree of hypocentral diffusion and thus a smaller effective area covered by the earthquakes.

Average Depth (*AZ*)

Another useful parameter is the average depth of hypocenters in a given time window. *Nadeau et al.* [1995] suggested that increase in relatively deep seismicity may precede large events; this would translate into a precursory increase in *AZ* here.

Time Interval for the Occurrence of a Constant Number of Events (*TI*)

This parameter characterizes the temporal distribution of events and is reciprocal to the better known "seismicity rate" measuring the number of events per unit time. Thus increasing time interval indicates decreasing seismicity rate. Both precursory seismicity-rate decrease (seismic quiescence) and increase have been reported in the literature. Some authors have studied in detail the statistical grounds of the seismic quiescence hypothesis [*Matthews and Reasenberg*, 1988; *Reasenberg and Matthews*, 1988]; others have emphasized the need for distinguishing artificial from natural changes in seismicity rates [e.g., *Habermann*, 1987; *Wyss and Habermann*, 1988; *Eneva et al.*, 1994]. The synthetic catalogs used here are free of artificial changes and measurement errors of the types observed in real catalogs. It is possible to simulate such errors in the model catalogs and see how they affect the temporal changes of the study parameters; however, this line of study is not pursued here.

Ratio of Numbers of Events in Two Different Magnitude Ranges (*MR*)

This parameter, briefly referred to as "magnitude ratio," characterizes the size distribution of earthquakes. It represents the ratio of the number of $M \ge M_0$ events to the number of $M < M_0$ events:

$$r_{M_0} = \frac{N(M \ge M_0)}{N(M < M_0)},$$

where M_0 is chosen to be a certain representative magnitude determined by inspecting the frequency-size statistics of earthquakes. As an example, $M_0 = 4$ is used by *Eneva and Ben-Zion* [1997]. For a total N_T number of events, $N(M \ge M_0) + N(M < M_0) = N_T$. A parameter similar to *MR* has been previously used by *Radulian and Trifu* [1991] to identify a precursory anomaly prior to an *M*7.3 event. *MR* replaces here the better known *b* values of the frequency-size statistics; it is especially useful in cases where for various reasons *b* values cannot be properly defined. This can be due to strong deviations from a power law scaling, which characterize most synthetic and natural earthquake catalogs of individual fault systems [*Ben-Zion and Rice*, 1993, 1995; *Ben-Zion*, 1996; *Wesnousky*, 1994; *Stirling et al.*, 1996]. When both *b* values and magnitude ratios can be defined for a given data set, they are inversely related; lower *b* values indicate higher relative numbers of larger events and hence larger ratios.

Two types of precursory changes relevant to *MR* used here have been reported in the literature. The first type uses *b* values; low values [e.g., *Henderson et al.*, 1992] or increasing values, sometimes followed by a short-term decrease [*Smith*, 1986], have been reported to precede some large events. These are not necessarily contradictory reports, as they may reflect different portions of the same type of precursory anomaly. Isolated observations of such type, however, may have to be viewed with some scepticism, as they can be greatly affected by the same factors dividing the seismicity rate variations into real and artificial changes [*Habermann*, 1987; *Eneva et al.*, 1994]. On the other hand, some physical considerations have been put forward as to what kind of precursory changes are to be expected. *Main et al.* [1989] discuss a fracture mechanics model that suggests a sequence of precursory minimum, maximum, and a second minimum in the *b* values. However, owing to varying levels of seismic activity, not all of these features can necessarily be resolved in practice, as a sufficient number of events is needed for calculation of the *b* values. A second type of precursory changes relevant to *MR* used here is much less known than the *b* values. It concerns deviations from the Gutenberg-Richter law [*Okuda et al.*, 1992], which appear to be the smallest (i.e., the frequency-size statistics show best fits to straight lines) prior to large earthquakes.

Techniques

The analysis involves the following steps: (1) A moving window with a given number of events is used to divide the study data into overlapping groups. (2) Each test parameter is calculated for each of these groups. Thus the temporal changes of each parameter are represented by a parameter time series. (3) Possible correlations among parameters are evaluated. If a high correlation coefficient is found for a given pair of parameters, one of the parameters may be eliminated from consideration. (4) Each of the remaining parameter time series is used to construct two subseries; one consists of local maxima, and the other consists of local minima. (5) Each time series of local extrema is compared with the time series of large events using a quantitative technique of association in time. (6) The significance of association is evaluated by comparison with randomly simulated time series. (7) Significant precursors to large events are identified in terms of precursory local maxima or minima. Precursory trends associated with these are evaluated separately, as the relationship between types of extrema and trends is nonunique.

Time Series and Local Extrema

Comparing results obtained from data samples having different numbers of events and filling volumes of different size and shape can be very misleading. Effects of this type concerning the correlation dimension are discussed in detail by *Eneva* [1994b, 1996]. For this reason, parameter variations in time are examined here by constructing overlapping groups of a constant number of events covering time windows of variable size, rather than using time windows of the same length. The advantage of using number windows is that statistical variations resulting from the use of samples of different size are removed. Number windows have been previously used by some other authors [e.g., *Hirata*, 1989; *De Rubeis et al.*, 1993].

Each study parameter is evaluated for each sample group of events. The parameter group values calculated for a given catalog form a time series reflecting the temporal variations of that parameter. We denote the estimated value of any given parameter by y_j, $j = 1,..., n^{gr}$, where the subscript j marks the group for which the estimate has been made and n^{gr} denotes the total number of groups in a catalog. The parameter value y_j is assigned to a time t_j which is the end time of group j, that is, the occurrence time of the last event in the group. Since the time intervals within which a constant number of events occurs are of different length, the end times t_j are not separated by equal time intervals. Figure 3 shows an example of a parameter time series.

Once the parameter time series are calculated, it is necessary to evaluate the possible correlation for each pair of parameters, $\{y_j^k, y_j^m\}$, where k and m are parameter indexes with $k = 1,..., 7$ and $m = k+1,..., 7$ for the seven study parameters here. Most works in which several test parameters are used assume implicitly that the parameters are independent. Some of the seismicity parameters, however, can be expected to correlate by default. In particular, all spatial parameters are somewhat related in the present work. This is especially true for *NS* and *CD* reflecting the degree of spatial clustering in overlapping distance ranges. In addition, there have been reports for weak negative correlation between *CD* and *b* values [*Hirata*, 1989; *Main*, 1992; *Henderson et al.*, 1992, 1994], which might translate into a positive correlation between *CD* and *MR* here. We recall that *CD* is derived only from the hypocentral distribution, whereas *MR* (and the *b* value) reflects only the size distribution of events. *Main* [1992] shows that both positive and negative correlations between *CD* and *b* values can be expected depending on the type of damage sustained. The presence of negative correlation is believed to be indicative of predominance of crack coalescence [*Henderson et al.*, 1992], whereas a positive correlation may be due to a fluid flow in the fault zone [*Henderson et al.*, 1994]. Significant interparameter correlation (e.g., > 80 to 90%) between any two parameters indicates that one of them may be redundant. On the other hand, the presence of significant but relatively weak correlation in such complex sequences does not justify the removal of parameters.

After a decision is made which of the parameter time series are to be retained for further analysis, they are compared with the time series of large events. Obviously, there is some arbitrariness in deciding which events should be considered as "large." The earthquake catalogs analyzed here are generated by models of two-dimensional strike-slip faults in a three-dimensional elastic half-space [*Ben-Zion*, 1996]. Each fault model has a horizontal extent of 70 km and a depth extent of 17.5 km. As discussed by *Ben-Zion* [1996], large earthquakes are events which rupture most of the fault segment under consideration and hence relieve stress from that fault as a whole. In the context of the models, this holds for $M \geq 6$ events. It is also reasonable to assume that if detectable precursors are present, the largest events in the catalogs would likely dominate the process, and it may be impossible to distinguish between their precursors and precursors to events of smaller size. We denote the number of large events with $N^{large} = N(M \geq 6)$. Each event in the time series of large earthquakes is represented by its magnitude M_k and occurrence time t_k, $k = 1,..., N^{large}$.

The purpose of the analysis is to compare quantitatively the time series of each seismicity parameter $\{y_j, j = 1,..., n^{gr}\}$ with the time series of large events $\{M_k, k = 1,..., N^{large}\}$. Initially, it is preferable not to expect any particular relationship between event magnitude and amplitude or duration of precursory anomalies. Thus, instead of using the series of large events, we use the series of their occurrence times $\{t_k, k = 1,..., N^{large}\}$. The parameter time series are not used directly either; because the

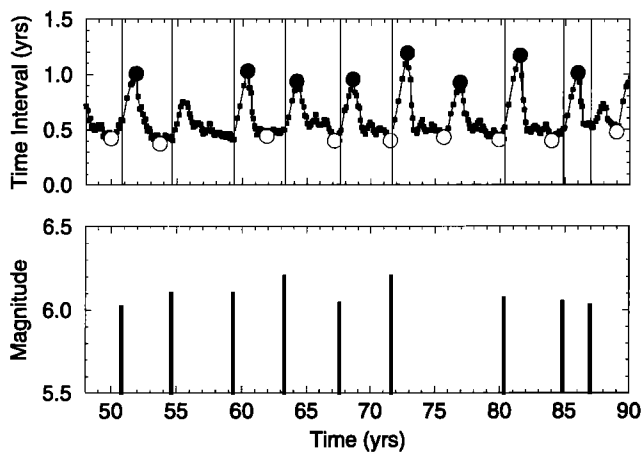

Figure 3. Example based on the *TI* time series for the asperity model of *Ben-Zion* [1996]. (top) Portion of the parameter time series. Squares denote individual entries of the time series, plotted at the end times of groups having 100 events each and moving forward in time by 20 events. Local maxima and minima are shown with solid and open circles, respectively. Vertical lines mark the occurrence times of $M \geq 6$ events. (bottom) Magnitudes and times of $M \geq 6$ events.

parameter values are not calculated at equal time intervals and the number of groups for which they are estimated is much larger than N^{large}, it is not possible to use regular cross-correlation techniques to compare them with the large events. To overcome this obstacle, not all group values in the parameter series are used but only the most pronounced ones. Each parameter series is reduced to two subseries; one consisting of the occurrence times of "local maxima" $\{t_{mx};\ mx = 1,..., n^{max}\}$ and the other of the occurrence times of "local minima" $\{t_{mn};\ mn = 1,..., n^{min}\}$, where n^{max} and n^{min} are the total numbers of local maxima and minima in a given series, respectively. Unless distinction between maxima and minima is required, these two subseries will be further denoted by $\{t_m;\ m = 1,..., n^{extr}\}$, where m is either mx or mn and n^{extr} is either n^{max} or n^{min}. The local extrema are determined using the following filters:

1. A group parameter value y_j at time t_j is considered tentatively to be a local maximum if $y_{j-i} < y_j > y_{j+i}$ for $i = 1,..., L$, where L is a preselected number of group values preceding and following y_j. It is not required that these values increase monotonically before the occurrence of y_j and decrease monotonically afterward. The tentative local minima are determined similarly, with reversed signs of the inequalities above.

2. It is further required that local maxima and minima alternate in time. When two or more tentative extrema of the same type occur consecutively, only the largest or the smallest is retained.

3. A tentative local extremum is retained only if it differs from the previous local extremum of opposite type by a certain preselected threshold amount s or more. Retrospectively, the extrema both prior to and following the extremum under consideration can be compared. In a near-real-time application, however, only earlier local extrema are known. Using only earlier extrema values to decide on the retention of a tentative extremum is not to be confused with the procedure described in step 1 above for the initial identification of tentative extrema when both earlier and later group values are used around any group value.

4. Step 2 is repeated, as step 3 might have left some nonalternating extrema.

The filters above assure that the final maxima and minima occur alternately. Note that the final extrema satisfy $y_{mx} - y_{mn} \geq s$ for any $t_{mn} < t_{mx} < t_{mn+1}$ or $t_{mx} < t_{mn} < t_{mx+1}$ and that the values of y_m are used only to determine when (t_m) the local extrema occur. The subseries $\{t_m;\ m = 1,..., n^{extr}\}$ are the ones that we compare with the series of occurrence times of large events $\{t_k;\ k = 1,..., N^{large}\}$ through a specific procedure of association in time. Unlike some current techniques, this approach gives equal weight to both sets of local maximum and minimum values, without preference to maxima.

Once the values for s and L are chosen, the procedure outlined above can be used even when there is a long-term trend in the temporal variations, as the local extrema are identified on the basis of relative rather than absolute amplitudes. The same is not true for existing methods using preset percentages of the highest parameter values, as these would fall only within the high end of the trend. The principle of selecting distinct rather than absolute maximum values may not make a significant difference for seismically active zones with no long-term seismicity trends. However, as discussed by, for example, *Li and Rice* [1987] and *Ben-Zion et al.* [1993], a realistic description of seismic cycles should include transients associated with viscous flow underneath the brittle seismogenic zone. Practices ignoring the possible existence of long-term trends are also questionable in conditions such as in mines where the level of induced seismic activity can vary significantly when blasting and shutdown periods alternate.

For any given parameter time series, we choose values for s and L such that the number of local extrema n^{extr} matches approximately $N^{large} = N(M \geq 6)$, and then check if any significant association exists between this series and the time series of large events. Thus we attempt to identify association between approximately equal numbers of large events and anomalies. Compared with the M8 algorithm, this procedure can be thought of as using a variable ($\approx 100\% \times N^{large}/n^{gr}$), rather than a constant, percentile depending on the number of events to be retrospectively predicted. In addition, only the highest/lowest of a cluster of high/low values is considered thus assuring that the local maxima/minima are well separated in time. As said before, equal weight is given here to both maximum and minimum values. Our approach assures that the numbers of default occurrences of "false alarms" and "failures to predict" stemming from differences between the number of events to be predicted and the number of identified anomalies are minimized. *Reasenberg and Matthews* [1988] defined the validity and reliability of a precursor as $V = n(S)/n(A)$ and $R = n(S)/n(E)$, respectively, where $n(A)$ is the number of anomalies, $n(E)$ is the number of earthquakes to be predicted, and $n(S)$ is the number of successful predictions. Using this terminology, the difference between the validity and reliability is minimized here ($V \approx R$) through the requirement that $n(A) \approx n(E)$.

In the present approach, the values for L and s are chosen through an exploratory procedure applied to each parameter time series for any given earthquake catalog. The values of L and s are varied until the requirement for approximate equality between the numbers of anomalies and events to be predicted is fulfilled. Once this is done, one can decide whether the value of s is reasonable or is unlikely to have practical meaning. If a larger value of s appears to be required, leading to a smaller retained number of local extrema, one has two choices: either (1) to accept that a certain portion of the large events is inherently unpredictable by statistical means or (2) to consider a higher magnitude cutoff for the large events in order to make their number comparable with the number of local extrema. The time series from the synthetic catalogs studied here do not exhibit significant long-term trends. Thus a comparison with the average parameter value for any given time series is a valid way to estimate s, even though the procedure used is not limited to such comparisons.

The above approach does not eliminate the arbitrariness of the initial empirical choices but shifts its weight to a different, perhaps more manageable problem. A higher level of formalization that also retains a reasonable flexibility does not appear feasible at this time. Admittedly, it may appear practically inconvenient and somewhat counter intuitive to apply varying values of thresholds s for different parameters and different catalogs. It can be argued, however, that the widespread practice of arbitrary choice of alarm levels, together with the unrelated choice of the magnitude threshold for events considered to be "large," seems to be more questionable. The traditional ways of determining false alarms and missed targets may be less useful than sometimes assumed.

An example of retained local extrema is shown in Figure 3. It is observed that local minima tend to precede the large events in this case. An apparent local maximum-minimum pair is missed around 55 and 60 model years, as the chosen value of s is not exceeded, and the result is a missed event. On the other hand, a

local minimum around 75 model years is not succeeded by a large event and represents a false alarm (in fact, an event of M = 5.9 has occurred following this minimum, but the magnitude cutoff in this study is M = 6.0).

Association in Time

Because regular methods of cross correlation cannot be applied to compare time series with entries at different times, we use an alternative technique to compare the time series $\{t_k;\ k = 1,..., N^{large}\}$ and $\{t_m;\ m = 1,..., n^{extr}\}$. The approach was inspired by an application reported by *Mulargia* [1992] for studying the association in time between 11 volcanic eruptions of Mount Etna and 12 local earthquakes.

Our procedure distinguishes between pre-event and post-event extrema. If independent information on the duration of the precursory and aftereffect periods exists, it could be readily incorporated in our algorithm. However, such information is usually absent, and we use here the following arbitrary criteria. Suppose that a given local extremum at time t_m occurs between the times t_{k-1} and t_k of two consecutive large events. This local extremum is assumed to be a potential precursor of the event at time t_k if t_m is closer to t_k than it is to t_{k-1}, and a potential aftereffect of the event at time t_{k-1} otherwise. If several precursory local extrema of the same type are identified prior to the large event occurring at time t_k, only the "last" one is considered. Similarly, if several aftereffects are obtained following the large event at time t_{k-1}, only the "first" one is considered.

Figure 4 shows an example where the precursors and aftereffects of large events are plotted versus respective time lags to the occurrences of the large events. The portion between model years 48 and 90 reflects the time series previously shown in Figure 3. The association frequency for a time lag u is obtained by shifting a given time series of local extrema $\{t_m;\ m = 1, ..., n^{extr}\}$ at u time units with respect to the time series of large events $\{t_k;\ k = 1, ..., N^{large}\}$ and counting all joint occurrences. The association frequency at time lag u can be written as

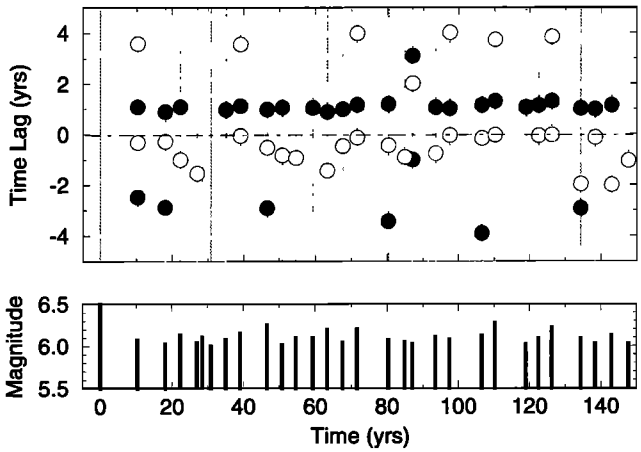

Figure 4. An example of association between large events and local extrema in the *TI* parameter for the asperity model of *Ben-Zion* [1996]. (top) Time lags at which local maxima (solid circles) and minima (open circles) appear with respect to the $M \geq 6$ events. Precursors and aftereffects are shown at negative and positive time lags, respectively. Vertical lines mark the occurrence times of the large events. (bottom) Magnitudes and times of $M \geq 6$ events.

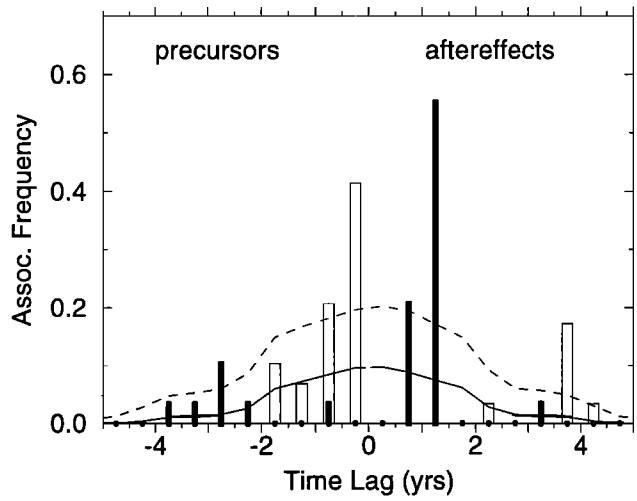

Figure 5. An example of association plot showing the *TI* parameter for the asperity model of *Ben-Zion* [1996]. Local maxima (solid bars) and minima (open bars) show a significant statistical association with the $M \geq 6$ events when exceeding the association frequencies obtained for corresponding random simulations. The average distribution of association frequencies +2 standard deviations from an ensemble of 1000 random simulations are shown with solid and dashed lines, respectively.

$$f(u) = \frac{1}{N^{large}} \sum_{k=1}^{N^{large}} \sum_{m=1}^{n^{extr}} H(h - |t_k + u - t_m|),$$

where $H(\,)$ is the Heaviside unit step function which is nonzero if $t_m - t_k$ is within $u \pm h$ units. In applying the above, the time lag u is varied with a step Δu, and h is a step of counting. When $h = \Delta u/2$, the counts are done for nonoverlapping adjacent lag intervals $[u-h, u+h)$. These intervals are semiopen and do not cross zero; this ensures that precursors and aftereffects are not mixed together. Negative time lags are associated with precursors and positive time lags are related to aftereffects. The significance of association is evaluated by comparing the observed association frequencies with a corresponding average association from a large ensemble of randomly generated time series, which can be considered as an estimate of the probability of chance association.

Figure 5 shows an example of observed association frequencies between large events and local extrema for various time lags, together with corresponding randomly simulated frequencies for the same model as in Figures 3 and 4. It is clearly seen that there is a significant association between the occurrence times of large events and the times of local minima in the specific parameter examined (*TI*) for negative time lags. More specifically, 61% of the large events are preceded by local minima within 1 model year, and 79% are preceded by local minima within 2 model years. Similarly, a significant association is observed with local maxima at positive time lags. In both cases, the association frequencies are much larger than the ones expected from random association.

Note that once reliable information is extracted as to what kind of local extrema significantly precede and/or follow large events, various other issues can be addressed. For example, one can check if there is any correlation between magnitude of large events and the particular values of precursory and aftereffect extrema, between magnitude and time lags, etc. Thus it is possible to check whether larger events are preceded by larger

anomalies within larger time intervals in these synthetic catalogs, an assumption frequently made regarding real seismic activity.

Local Extrema and Trends

Precursory trends (increasing or decreasing values) for various parameters have been reported in the seismological literature. However, there are problems with the practical identification of any trend in a given time series. One reason for uncertainty associated with identified trends is purely statistical. The specific choice of window size in combination with the resolution of data can obscure or even reverse the apparent precursory trends. Nonlinearities in the underlying physical process can also contribute to the existence of opposing trends, while consistent trends are expected from linear-based intuition. A distinction should be made between the "observable" ("identifiable") trend, largely dependent on the specific statistical treatment used and the resolution of data, and the "expected" trend related to the physics of the seismic process. We can only derive information from the observable trends, regardless of how representative they are of the expected ones. Thus the example in Figure 5 indicates that the large events tend to occur on the background of increasing values leading from the precursory local minima to aftereffect local maxima. However, Figure 6 shows two examples, in which a precursory extremum of the same type (minimum) is observed prior to two large events, but the observed trends are different. In the case on the left, the local minimum occurs shortly before the large event, and in a hypothetical near-real-time observation a negative precursory trend leading to this minimum is most likely to be reported. In the case on the right, the local minimum occurs early enough before the subsequent large event to allow a number of additional observations to be made after the minimum has been reached. In this case, a generally increasing trend of the signal most closely preceding the respective large event is likely to be reported. Thus, even though the physical process is such that precursory extrema of the same type characterize the given parameter time series, the interplay of the variability of the process (i.e., the variations in the time lags at which the precursory extrema appear) and the observational limitations are likely to lead to identifications of opposite precursory trends.

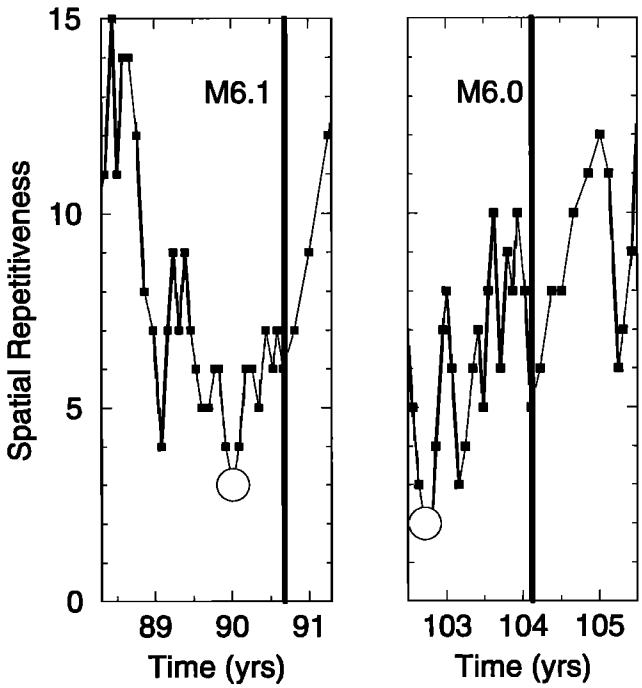

Figure 6. An example illustrating the nonunique relationship between precursory trends and type of precursory local extrema in the *SR* parameter for the uniform model of *Ben-Zion* [1996]. Squares mark group parameter values plotted at end times of groups of 100 events each, moving forward in time by 20 events. Vertical bars indicate occurrence times of $M \geq 6$ events. Both frames show precursory local minima, but the trends most likely to be reported are decreasing trend in the left frame and increasing trend in the right one.

It may be assumed that one must make at least K observations between a precursory extremum and the subsequent large event in order to identify the precursory trend as the one following the extremum (i.e., positive trends following minima and negative trends following maxima). If the number of observations, however, is below K, the precursory trend most likely to be identified is the one leading to the precursory extremum (i.e., positive trends leading to maxima and negative trends leading to minima). The value of K depends on the specific circumstances of observation. It is reasonable to choose $K \geq L$, that is, the required minimum number of observations to be identical to or larger than the number of neighboring points used to identify the local extrema as discussed before. This is implemented by *Eneva and Ben-Zion* [1997].

In light of the above and in view of the nonlinearity of the earthquake process, there is no compelling reason for expecting to observe the same precursory trend in a given seismicity parameter prior to different large events. It is possible that effects of the type discussed above are at least partly responsible for contradictory reports of opposing precursory trends from various seismically active areas or even prior to different large events in the same area. The same considerations can be also applied to the periods following large events. For this reason, it may be better to concentrate on identifying the precursors in terms of type of local extrema rather than in terms of trends.

Discussion and Conclusions

The seismicity parameters in this study were chosen with the intention of extracting a large amount of information from the origin times, locations, and magnitudes included in typical earthquake catalogs. A specific emphasis was given to parameters describing quantitatively the spatial distribution of events. This distribution has been traditionally ignored but found to be at least as informative [*Eneva and Pavlis*, 1991; *Eneva et al.*, 1992] as the better known parameters describing the temporal and size distribution of events.

The techniques we have developed for identifying precursory patterns are intended to address some problems in existing methods which search only for high precursory values, choose significance thresholds for anomaly identification independently of the specifics of the existing seismicity (e.g., number of large events to be "predicted" retrospectively), use single parameters in isolated observations, employ parameters whose intercorrelations and their effect on the predictions are not examined, mix patterns preceding and following large events, lack rigorous techniques to evaluate the significance of association between patterns and large events, and report precursory trends without considering sample size and data resolution.

Elements of the techniques discussed in this work can be used at the learning stage of any pattern recognition scheme aiming at

earthquake prediction. These techniques are not parameter-specific and additional parameters can be incorporated into the algorithm, such as event source characteristics and nonseismicity indicators (e.g., water level and electrical signals). A comprehensive application of the parameters and techniques described here to the synthetic earthquake catalogs of *Ben-Zion* [1996] is discussed by *Eneva and Ben-Zion* [1997]. Another application of these techniques and some of the proposed seismicity parameters was used to evaluate the potential for rockburst forecast based on patterns in the mining-induced microseismicity [*Eneva*, 1997; *Eneva and Villeneuve*, 1997].

Although a number of problems relevant to earthquake prediction were addressed in the present work, some issues are left for future work. The most important of those are as follows:

1. The significant number of large events and the absence of measurement errors in the studied synthetic catalogs make the identification of precursory patterns here easier than when working with real catalogs. It is possible, however, to use the same synthetic data sets for the purpose of studying quantitatively the effect of measurement errors by incorporating various such errors into the catalogs. In real catalogs, one would have to make some additional efforts to assure the reliability of the final results.

2. Only the learning procedure of the pattern recognition was applied here. The identification of local extrema as they unfold in near-real-time observations and the decision on alarms (or alerts) for the purpose of forward prediction may present a somewhat different challenge than the one encountered with retrospective identification.

3. In the work done so far, we have not attempted to optimize the retrospective prediction by jointly minimizing the total time when alarms are declared and maximizing the predictability of large events [e.g., *Molchan and Kagan*, 1991; *Pepke et al.*, 1994]. This will be needed, however, if these techniques and parameters are to be used in forward prediction.

4. The significance of association was evaluated only retrospectively. The "true" statistical significance cannot be tested on the same sets used for learning; it has to be tested on different, independent sets [*Mulargia*, 1992; *Rhoades and Evison*, 1989].

5. When working with real earthquake catalogs, it will be necessary to evaluate rigorously the stability of the applied procedure with respect to the choice of location and size of the areas to be monitored for precursory changes. As *Minster and Williams* [1992] demonstrated, depending on this choice, results from pattern recognition studies can vary greatly. Related issues are not addressed here, since the size of the modeled fault segments is comparable with the source size associated with the largest events ($M \geq 6$ in the present work) and the entire synthetic catalogs are used.

6. In the present work, although several parameters are used to monitor seismicity, their performance is evaluated one by one; that is, no attempt has been made yet to devise a combined parameter containing all the information provided by the individual parameters examined. Such an approach is common in the related literature [see *Pepke et al.*, 1994; *Pepke and Carlson*, 1994]. In the M8 algorithm, parameters are considered separately as here, although six out of seven are required to reach simultaneously high values for a TIP to be declared [*Keilis-Borok and Kossobokov*, 1990; *Kossobokov and Carlson*, 1995]. In the future, however, it may be useful for the learning stage of the pattern recognition to construct one or a few combined parameters. For this purpose, it may be possible to adopt a method incorporating statistical orthogonal decomposition of empirical functions, similar to the one employed by *Sobolev and Koltzov* [1988] in their analysis of precursors to failure in large-scale laboratory models. We note that the idea of combined parameters where various characteristics are appropriately weighed is conceptually very different from a search for a single "magic" parameter, an approach still used in most prediction studies.

Acknowledgments. We thank Jim Rice and Ted Habermann for fruitful discussions. Reviews of the original manuscript by John McCloskey, Didier Sornette, Chi-Yu King, and Nadya Williams helped to improve the paper. M. Eneva was partially supported by Inco, Ltd. (Ontario, Canada). Y. Ben-Zion was with the Department of Earth and Planetary Sciences at Harvard University during this work and was supported by grants from the Southern California Earthquake Center (subcontracts 569928 and 621911 from USC, based on NSF support).

References

Ben-Zion, Y., Stress, slip and earthquakes in models of complex single-fault systems incorporating brittle and creep deformations, *J. Geophys. Res.*, *101*, 5677-5706, 1996.

Ben-Zion, Y., and J. R. Rice, Earthquake failure sequences along a cellular fault zone in a three-dimensional elastic solid containing asperity and nonasperity regions, *J. Geophys. Res.*, *98*, 14,109-14,131, 1993.

Ben-Zion, Y., and J. R. Rice, Slip patterns and earthquake populations along different classes of faults in elastic solids, *J. Geophys. Res.*, *100*, 12,959-12,983, 1995.

Ben-Zion, Y., J. R. Rice, and R. Dmowska, Interaction of the San Andreas fault creeping segment with adjacent great rupture zones and earthquake recurrence at Parkfield, *J. Geophys. Res.*, *98*, 2135-2144, 1993.

Coughlin, J. and R. Kranz, New approaches to studying rock burst-associated seismicity in mines, in *Proceedings of the 32nd U.S. Symposium, Rock Mechanics as a Multidisciplinary Science*, edited by J. –C. Roegiers, J.-C. pp.491-500, A. A. Balkema, Rotterdam, 1991.

De Rubeis, V., P. Dimitriu, E. Papadimitriu, and P. Tosi, Recurrent patterns in the spatial behaviour of Italian seismicity revealed by the fractal approach, *Geophys. Res. Lett.*, *20*, 1911-1914, 1993.

Debate on "VAN," *Geophys. Res. Lett.*, *23*, 1291-1452, 1996.

Eneva, M., Investigation of space-time distribution of earthquakes in the Charlevoix seismic zone, Quebec, *Geol. Surv. of Can. Open File Rep. 2983*, 140 pp., 1994a.

Eneva, M., Monofractal or multifractal: a case study of spatial distribution of mining induced seismic activity, *Nonlinear Processes Geophys.*, *1*, 182-190, 1994b.

Eneva, M., Effect of limited data sets in evaluating the scaling properties of spatially distributed data: An example from mining induced seismic activity, *Geophys. J. Int.*, *124*, 773-786, 1996.

Eneva, M., In search for a relationship between induced seismicity and larger events in mines, *Tectonophysics*, in press, 1997.

Eneva, M., and Y. Ben-Zion, Application of pattern recognition techniques to earthquake catalogs generated by models of segmented fault systems in three-dimensional elastic solids, *J. Geophys.Res.*, in press, 1997.

Eneva, M., and G. L. Pavlis, Spatial distribution of aftershocks and background seismicity in central California, *Pure Appl. Geophys.*, *137*, 35-61, 1991.

Eneva, M., and T. Villeneuve, Retrospective pattern recognition applied to mining induced seismicity, in *Rockbursts and Seismicity in Mines*, edited by S. J. Gibowicz and S. Lasocki, A. A. Balkema, Rotterdam, in press, 1997.

Eneva, M., and R. P. Young, Evaluation of spatial patterns in the distribution of seismic activity in mines: A case study of Creighton Mine, northern Ontario (Canada), in *Rockbursts and Seismicity in Mines*, edited by R. P. Young, pp. 175-180, A. A. Balkema, Rotterdam, 1993.

Eneva, M., M. W. Hamburger, and G. A. Popandopulo, Spatial distribution of earthquakes in aftershock zones of the Garm region, Soviet Central Asia, *Geophys. J. Int.*, *109*, 38-53, 1992.

Eneva, M., R. E. Habermann, and M. W. Hamburger, Artificial and natural changes in the rates of seismic activity: A case study of the Garm Region, Tadjikistan (CIS), *Geophys. J. Int.*, *116*, 157-172, 1994.

Frohlich, C., and S. D. Davis, Single-link cluster analysis as a method to evaluate spatial and temporal properties of earthquake catalogues, *Geophys. J. Int.*, *100*, 19-32, 1990.

Grassberger, P., and I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.*, *50*, 346-349, 1983.

Gvishiani, A. D., A. I. Gorshkov, E. Y. Rantzman, A. Sisternas, and A. A. Solovev, *Forecasting Earthquake Sites in Regions of Moderate Seismicity* (in Russian), 176 pp., Nauka, Moscow, 1988.

Habermann, R. E., Man-made changes of seismicity rate, *Bull. Seismol. Soc. Am.*, *77*, 141-157, 1987.

Habermann, R. E., Precursory seismic quiescence: Past, present, and future, *Pure Appl. Geophys.*, *126*, 279-318, 1988.

Habermann, R. E., and F. Creamer, Catalog errors and the M8 earthquake prediction algorithm, *Bull. Seismol. Soc. Am.*, *84*, 1551-1559, 1994.

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.

Henderson, J., I. Main, P. Meredith, and P. Sammonds, The evolution of seismicity at Parkfield: Observation, experiment and a fracture-mechanical interpretation, *J. Struct. Geol.*, *14*, 905-913, 1992.

Henderson, J., I. G. Main, R. G. Pearce, and M. Takeya, Seismicity in north-eastern Brazil: Fractal clustering and the evolution of the b-value, *Geophys. J. Int.*, *116*, 217-226, 1994.

Hirata, T., A correlation between the *b* value and the fractal dimension of earthquakes, *J. Geophys. Res.*, *94*, 7507-7514, 1989.

Kagan, Y. Y., and L. Knopoff, Spatial distribution of earthquakes: The two-point correlation function, *Geophys. J. R. Astron. Soc.*, *62*, 303-320, 1980.

Kagan, Y. Y., and L. Knopoff, Likelihood analysis of earthquake catalogues, *Geophys. J. Int.*, *106*, 135-148, 1991.

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., 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.

Kossobokov, V. G., and J. M. Carlson, Active zone size versus activity: A study of different seismicity patterns in the context of the prediction algorithm M8, *J. Geophys. Res.*, *100*, 6431-6441, 1995.

Li, V. C., and J. R. Rice, Crustal deformation in great California earthquake cycles, *J. Geophys. Res.*, *92*, 11,533-11,551, 1987.

Main, I. G., Damage mechanics with long-range interactions: Correlation between the seismic *b*-value and the fractal two-point correlation dimension, *Geophys. J. Int.*, *111*, 531-541, 1992.

Main, I., P. G. Meredith, and C. Jones, A reinterpretation of the precursory seismic *b*-value anomaly from fracture mechanics, *Geophys. J. Int.*, *96*, 131-138, 1989.

Matthews, M. V., and P. A. Reasenberg, Statistical methods for investigating quiescence and other temporal seismicity patterns, *Pure Appl. Geophys.*, *126*, 357-372, 1988.

Minster, J.-B., and N. P. Williams, The "M8" intermediate-term earthquake prediction algorithm: An independent assessment (abstract), *Eos Trans. AGU*, *73*(43), Fall Meet. Suppl., 366, 1992.

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

Mulargia, F., Time association between series of geophysical events, *Phys. Earth Planet. Inter.*, *71*, 147-153, 1992.

Mulargia, F., and P. Gasperini, Evaluating the statistical validity beyond chance of "VAN" earthquake precursors, *Geophys. J. Int.*, *111*, 32-44, 1992.

Nadeau, R., M. Antolik, P. A. Johnson, W. Foxall, and T. V. McEvilly, Seismological studies at Parkfield III: Microearthquake clusters in the study of fault-zone dynamics, *Bull. Seismol. Soc. Am.*, *84*, 247-263, 1994.

Nadeau, R. M., W. Foxall, and T. V. McEvilly, Clustering and periodic recurrence of microearthquakes on the San Andreas fault at Parkfield, California, *Science*, *267*, 503-507, 1995.

Okuda, S., T. Ouchi, and T. Terashima, Deviation of magnitude frequency distribution of earthquakes from the Gutenberg-Richter law: Detection of precursory anomalies prior to large earthquakes, *Phys. Earth Planet. Inter.*, *73*, 229-238, 1992.

Ouchi, T., and T. Uekawa, Statistical analysis of the spatial distribution of earthquakes - Variation of the spatial distribution of earthquakes before and after large earthquakes, *Phys. Earth Planet. Inter.*, *44*, 211-225, 1986.

Ouillon, G., and D. Sornette, Unbiased multifractal analysis: Application to fault patterns, *Geophys. Res. Lett.*, *23*, 3409-3412, 1996.

Ouillon, G., C. Castaing, and D. Sornette, Hierarchical geometry of faulting, *J. Geophys. Res.*, *101*, 5477-5487, 1996.

Pepke, S. L., and J. M. Carlson, Predictability of self-organizing systems, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, *50*, 236-242, 1994.

Pepke, S. L., J. M. Carlson, and B. E. Shaw, Prediction of large events on a dynamical model of a fault, *J. Geophys. Res.*, *99*, 6769-6788, 1994.

Radulian, M., and C.-I. Trifu, Would it have been possible to predict the 30 August 1986 Vrancea earthquake?, *Bull. Seismol. Soc. Am.*, *81*, 2498-2503, 1991.

Reasenberg, P. A., and M. V. Matthews, Precursory seismic quiescence: A preliminary assessment of the hypothesis, *Pure Appl. Geophys.*, *126*, 373-406, 1988.

Rhoades, D. A., and F. F. Evison, On the reliability of precursors, *Phys. Earth Planet. Inter.*, *58*, 137-140, 1989.

Smith, W. D., Evidence for precursory changes in the frequency-magnitude b-value, *Geophys. J. R. Astron. Soc.*, *86*, 815-838, 1986.

Sobolev, G. A. and A. V. Koltzov, *Large-scale Modeling of Earthquake Preparation and Precursors* (in Russian), 205 pp., Nauka, Moscow, 1988.

Sornette, A., P. Davy, and D. Sornette, Fault growth in brittle-ductile experiments and the mechanics of continental collisions, *J. Geophys. Res.*, *98*, 12,111-12,139, 1993.

Stirling, M. W., S. G. Wesnousky, and K. Shimazaki, Fault trace complexity, cumulative slip, and the shape of the magnitude-frequency distribution for strike-slip faults: A global survey, *Geophys. J. Int.*, *124*, 833-868, 1996.

Trifu, C.-I., T. I. Urbancic, and R. P. Young, Non-similar frequency-magnitude for $M < 1$ seismicity, *Geophys. Res. Lett.*, *20*, 427-430, 1993.

Varotsos, P., and M. Lazaridou, Latest aspects of earthquake prediction in Greece based on seismic electric signals, *Tectonophysics*, *188*, 321-347, 1991.

Wesnousky, S. G., The Gutenberg-Richter or characteristic earthquake distribution, which is it?, *Bull. Seismol. Soc. Am.*, *84*, 1940-1959, 1994.

Wyss, M., and A. Allmann, Probability of chance correlations of earthquakes with predictions in areas of heterogeneous seismicity rate: The VAN case, *Geophys. Res. Lett.*, *23*, 1307-1310, 1996.

Wyss, M., and R. E. Habermann, Precursory seismic quiescence, *Pure Appl. Geophys.*, *126*, 319-332, 1988.

M. Eneva, Department of Physics, University of Toronto, 60 St. George Street, Toronto, Ontario, Canada M5S 1A7. (e-mail: eneva@geophy.physics.utoronto.ca)

Y. Ben-Zion, Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089-0740. (e-mail: benzion@jade.usc.edu)

(Received December 10, 1996; revised March 15, 1997;
accepted April 2, 1997.)