Premonitory spreading of seismicity over the faults’ network in southern California: Precursor Accord

I. Zaliapin and V. Keilis-Borok
International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, Russia

G. Axen
Department of Earth and Space Sciences, University of California, Los Angeles, California, USA

Received 31 October 2000; revised 16 March 2002; accepted 21 March 2002; published 10 October 2002.

[1] We establish a connection between geometry of the faults’ network in a region and seismicity preceding the strong earthquakes in the intermediate-term timescale of years. Previous studies of observed and computer-simulated seismicity demonstrated that strong earthquakes are preceded in that timescale by the rise of seismic activity in a lower magnitude range. Here, we explore a complementary phenomenon: spreading of activity over the fault network. This phenomenon is depicted by the seismicity pattern Accord, defined as a simultaneous rise of seismicity in a sufficiently large number of fault zones (hence its name). Pattern Accord has been found recently in synthetic seismicity generated by the colliding cascades (CC) model. The present study demonstrates this pattern in observed seismicity of southern California. We applied to observations a prediction algorithm based on the pattern Accord. Scaled to the target magnitude 7.5, the pattern Accord emerges within a few years before each of the three largest earthquakes in southern California (Kern County, 1952; Landers, 1992; and Hector Mine, 1999) and at no other time. Scaled to the target magnitude interval from 6.5 to 7.4, the pattern Accord precedes six out of nine earthquakes. The duration of alarms in both cases is about 30% of the time considered. We repeated prediction with different combinations of adjustable numerical parameters of the algorithm, summing up the success and error scores on an error diagram. This numerical experiment shows that the prediction results are stable under moderate variations of adjustable parameters. The final test would be, as always, an advance prediction. The pattern Accord has a simple physical explanation connected with equilibrium of the crustal blocks’ system under the impact of tectonic driving forces. At the same time, Accord is a manifestation of a broader phenomenon: premonitory increase of earthquakes’ correlation range. The latter phenomenon was found for the first time in the CC model, although it was hypothesized by one of the authors much earlier.

INDEX TERMS: 0910 Exploration Geophysics: Data processing; 3220 Mathematical Geophysics: Nonlinear dynamics; 7223 Seismology: Seismic hazard assessment and prediction; KEYWORDS: earthquake prediction, long-range earthquake correlations, nonlinear dynamics


1. Introduction

1.1. Long-Range Correlations in Seismicity

[2] One of the prominent features of seismicity is the correlation in the earthquakes’ occurrence at long distances, greatly exceeding the source dimensions. That correlation is expressed in many ways: simultaneous change of seismic activity within large regions [Mogi, 1968]; migration of earthquakes along fault zones [Mogi, 1968; Vilkovich and Schnirman, 1982; Ma et al., 1990]; alternate rise of seismicity in distant areas [Press and Allen, 1995] and even in distant tectonic plates [Romanowicz, 1993]; and seismicity patterns, premonitory to strong earthquakes (see section 1.2 below). Global correlations have been found also between seismicity and other geophysical phenomena, such as Chandler wobble, variations of magnetic field, and velocity of Earth’s rotation [Press and Allen, 1995; Romanowicz, 1993; Press and Briggs, 1975]. To explain these
phenomena, several mechanisms (not mutually exclusive) have been suggested. They may be divided into two groups.

1. Some authors attribute long-range correlations to a specific large-scale process, controlling stress and strength in the lithosphere. Among such processes are microrotation of tectonic plates [Press and Allen, 1995]; interaction of crustal blocks within the fault zones [Gabrielson and Keilis-Borok, 1983; Rundkvist and Rotwain, 1994; Soloviev et al., 1999]; microfluctuations in the direction of mantle currents [Ismail-Zadeh et al., 1999]; migration of fluids in fault systems [Barenblatt et al., 1983; Yamashita, 1998, 1999]; hydrodynamic waves in the upper mantle, triggering the strong earthquakes [Romanowicz, 1993]; and perturbations of the ductile layer beneath the seismically active zone [Aki, 1996].

2. In another approach, long-range correlations are not attributed to any specific mechanical process. The lithosphere is regarded as a complex system with complexity contributed by a multitude of instability mechanisms [Allegre et al., 1982; Keilis-Borok, 1990; Turcotte, 1997; Newman et al., 1994; Bak and Tang, 1989; Keilis-Borok, 1992; Mandelbrot, 1983; Rundle et al., 2000b; Ben-Zion, 2001]. The long-range correlation between earthquakes is considered then as a general feature of chaotic systems in a near-critical state [Turcotte et al., 2000; Smalley et al., 1985; Sornette and Sammis, 1995; Bowman et al., 1998; Rundle et al., 1997, 2000a; Zoeller et al., 2001].

1.2. Premonitory Seismicity Patterns

1. We discuss here the intermediate-term patterns preceding a strong earthquake with characteristic lead time of years. Studies of observed and computer-simulated seismicity demonstrated that an earthquake of magnitude \( M \) can often be preceded by certain spatiotemporal patterns of seismicity in the magnitude range \((M-\delta M, M)\), with \( \delta M \) varying from 2 to 4 for different patterns [Keilis-Borok and Shebalin, 1999]. Relatively better validated among them are the patterns reflecting premonitory rise of seismic activity and earthquake clustering. They are used, separately or jointly, in the earthquake prediction algorithms, which have been successfully tested by advance prediction; statistical significance of predictions is above 95% [Molchan et al., 1990; Keilis-Borok and Rotwain, 1990; Kossobokov et al., 1999a; Vorobieva, 1999].

2. A prominent feature of these patterns is that they are nonlocal. The patterns preceding an earthquake of magnitude \( M \) with linear source dimension \( L(M) \) are observed within the region of linear dimension \( 5L(M) \sim 10LM \) [Keilis-Borok and Rotwain, 1990; Keilis-Borok et al., 1980; Kossobokov et al., 1999a, 1999b; Vorobieva, 1999; Bowman et al., 1998].

3. An early (and probably the first) evaluation of the size of that region, still valid, has been given by Keilis-Borok and Malinovskaya [1964]. That study introduced a premonitory pattern Sigma: the measure \( \Sigma(t) \) of seismic activity is the sum of the areas of earthquake sources (fault breaks) in the medium magnitude range. The areas are coarsely estimated from the magnitudes. Essential for our study is the large size of the regions where pattern Sigma was observed. This size grows with magnitude, approximately by a power law, from \( 10^5 \text{ km}^2 \) for \( M_0 = 5 \), to \( 10^6 \text{ km}^2 \) for \( M_0 = 8 \) (Figure 1). Similar estimations for other precursors are summed up in Table 1. It shows linear dimension

![Figure 1. Size Q (in km$^2$) of the regions where pattern Sigma was observed prior to the earthquakes of magnitude M. After Keilis-Borok and Malinovskaya [1964].](image)

<table>
<thead>
<tr>
<th>Size (d(M))</th>
<th>Measure/Algorithm</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>up to 10L</td>
<td>total area of fault breaks</td>
<td>Keilis-Borok and Malinovskaya [1964]</td>
</tr>
<tr>
<td>(\sim)10L</td>
<td>distant aftershocks</td>
<td>Prozorov [1975] and Prozorov et al. [1990]</td>
</tr>
<tr>
<td>3L–5L</td>
<td>Algorithm CN</td>
<td>Keilis-Borok and Rotwain [1990]</td>
</tr>
<tr>
<td>5L–10L</td>
<td>Algorithm M$\delta$</td>
<td>Kossobokov et al. [1999a]</td>
</tr>
<tr>
<td>L–3L</td>
<td>Algorithms M$\delta$ and M$\delta \text{S}$</td>
<td>Kossobokov et al. [1990]</td>
</tr>
<tr>
<td>(\sim)5L</td>
<td>number of earthquakes</td>
<td>Knopoff et al. [1996]</td>
</tr>
<tr>
<td>(\sim)5L</td>
<td>Benioff strain release</td>
<td>Bufe and Varnes [1993] and Bowman et al. [1998]</td>
</tr>
<tr>
<td>(\sim)100L$^a$</td>
<td>seismic activity</td>
<td>Romanowicz [1993]</td>
</tr>
<tr>
<td>50L–100L$^b$</td>
<td>near-simultaneous pairs of earthquakes</td>
<td>Press and Allen [1995]</td>
</tr>
<tr>
<td>(\sim)3L</td>
<td>correlation range via Single Link Cluster analysis</td>
<td>Shebalin et al. [2000]</td>
</tr>
<tr>
<td>(\sim)5L</td>
<td></td>
<td>Zöller et al. [2001]</td>
</tr>
</tbody>
</table>

$^a$ L is the linear size of rupture during an approaching earthquake.

$^b$ The works of Romanowicz [1993] and Press and Allen [1995] are related to the timescale of tens of years.
d(M) of a region where premonitory patterns were observed. On the one hand, the dimension d(M) may be reduced down to 3L−L [Kossobokov et al., 1990]. On the other hand, an even larger dimension, up to 100L, was observed in the longer timescale, tens of years. According to Press and Allen [1995, p. 6428] the Parkfield (California) earthquake, with M about 6, L ≤ 10 km, “is not likely to occur until activity picks up in the Great Basin or the Gulf of California,” hundreds of kilometers away.

2. Pattern Accord

[6] Previous studies impose no limitations on territorial distribution of premonitory activity. The earthquakes forming a premonitory pattern may be either spread over the whole region or concentrated in a smaller area not necessarily close to the epicenter of an incipient strong earthquake [Keilis-Borok and Malinovskaya, 1964; Keilis-Borok et al., 1980; Caputo et al., 1983; Keilis-Borok et al., 1990; Keilis-Borok and Kossobokov, 1990; Keilis-Borok and Rotwain, 1990; Bufe and Varnes, 1993; Knopoff et al., 1996; Shaw et al., 1997; Bowman et al., 1998; Kossobokov et al., 1999a, 1999b]. The difference between these situations is schematically illustrated in Figure 2. Pattern Accord is designed to reflect premonitory spreading of precursory activity over the faults network.

2.1. Definitions

[7] Pattern Accord consists, qualitatively, of simultaneous rise of seismic activity in several branches of the fault network. This pattern has been found recently in synthetic seismicity generated by the colliding cascades (CC) model [Gabrielov et al., 2000a, 2000b]. It is defined as follows.

[8] Consider a network of major faults $F = \{F_i, i = 1,2,\ldots,n\}$ in a region $R$. The latter is divided into subregions $R_i$, $R = \{R_i, i = 1,2,\ldots,n\}$ such that a subregion $R_i$ corresponds to the fault (or, more generally, to subnetwork) $F_i, i = 1,\ldots,n$.

For each subregion $R_i$ we define the measure $\Sigma_i(t)$ of seismic activity as follows:

$$\Sigma_i(t) = \sum_{k:M_i \leq m_k \leq M_s} \left(10^{H(t-t_k,s)} - 10^{H(t-t_k,s)}\right),$$

where $H(t) = 10^{\lambda}$, $0 \leq t < s$, and $\lambda$ is a numerical parameter.

With finite positive $\lambda$ we introduce an attenuating “memory,” giving a larger weight to the more recent earthquakes. Similar memory was introduced in the “seismic flux,” a smoothed description of seismicity, suggested by Khokhlov and Kossobokov [1994].

The pattern Accord is depicted by the function

$$A(t) = \sum_{i} \left(\Sigma_i(t) \geq \Sigma_0, i = 1,\ldots,n\right).$$

In this definition, $M_1$, $M_2$, $B$, $s$, and $\lambda$ are numerical parameters.

The general scheme of prediction is taken from the past studies [Keilis-Borok, 1990; Keilis-Borok and Shebalin, 1999]:

1. We consider the earthquake sequence with aftershocks eliminated.

2. The target of prediction is the “strong” earthquakes, defined by the condition $M \geq M_0$; $M_0$ is a parameter of the
algorithm. The target magnitude may be truncated from above as well: $M_q > M \geq M_0$. Note that the threshold $\Sigma_0$ in equation (3) is taken as a specific fraction of contribution to $\Sigma(t)$ from an earthquake of magnitude $M_0$: $\Sigma_0 = 10^{R M_0 / R}$, $R$ being a parameter of the algorithm.

3. Applications for Southern California

[13] Here we explore the pattern Accord in the observed seismicity. We apply the prediction algorithm retrospectively, in the same way as it would be applied in actual (advance) prediction: at any moment $t$ the algorithm uses only information on seismicity prior to $t$ and not on subsequent seismicity. First, we try to predict the three strongest ($M > 7.5$) southern California earthquakes. Then we turn to the earthquakes with magnitudes from 6.5 to 7.4. Table 2 provides information about the earthquakes analyzed.

### Table 2. Earthquakes Analyzed

<table>
<thead>
<tr>
<th>Date</th>
<th>Magnitude</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 Dec. 1934</td>
<td>6.5</td>
<td>-115.5</td>
<td>32.3</td>
<td>Laguna Salada</td>
</tr>
<tr>
<td>31 Dec. 1934</td>
<td>7.0</td>
<td>-114.8</td>
<td>32.0</td>
<td>Colorado River delta</td>
</tr>
<tr>
<td>19 May 1940</td>
<td>6.7</td>
<td>-115.5</td>
<td>32.7</td>
<td>El Centro</td>
</tr>
<tr>
<td>21 Oct. 1942</td>
<td>6.5</td>
<td>-116.0</td>
<td>33.0</td>
<td>Borrego Valley</td>
</tr>
<tr>
<td>4 Dec. 1948</td>
<td>6.5</td>
<td>-116.4</td>
<td>33.9</td>
<td>Desert Hot Springs</td>
</tr>
<tr>
<td>21 July 1952</td>
<td>7.7</td>
<td>-119.0</td>
<td>35.0</td>
<td>Kern County</td>
</tr>
<tr>
<td>9 Feb. 1956</td>
<td>6.8</td>
<td>-119.5</td>
<td>31.8</td>
<td>San Miguel</td>
</tr>
<tr>
<td>15 Oct. 1979</td>
<td>7.0</td>
<td>-115.3</td>
<td>32.6</td>
<td>Imperial Valley</td>
</tr>
<tr>
<td>2 May 1983</td>
<td>6.7</td>
<td>-120.3</td>
<td>36.2</td>
<td>Coalinga</td>
</tr>
<tr>
<td>24 Nov. 1987</td>
<td>6.5</td>
<td>-118.5</td>
<td>33.1</td>
<td>Elmore Ranch</td>
</tr>
<tr>
<td>24 Nov. 1987</td>
<td>6.7</td>
<td>-118.5</td>
<td>33.0</td>
<td>Superstition Hills</td>
</tr>
<tr>
<td>28 June 1992</td>
<td>7.6</td>
<td>-116.4</td>
<td>34.2</td>
<td>Landers</td>
</tr>
<tr>
<td>17 Jan. 1994</td>
<td>6.8</td>
<td>-118.5</td>
<td>34.2</td>
<td>Northridge</td>
</tr>
<tr>
<td>16 Oct. 1999</td>
<td>7.6</td>
<td>-116.2</td>
<td>34.4</td>
<td>Hector Mine</td>
</tr>
</tbody>
</table>

3.1. Division of the Region

[14] By the definition developed in the CC model we have to divide our system into major subregions. [15] We use the traditional criteria of division based upon fault orientation, subjective grouping, and sense of slip, along with tectonic setting. In general this subdivision was not a difficult task because southern California geology is naturally divided. Northern regions are separated from southern ones by the Transverse Ranges, where east-west striking left-lateral faults dominate. The relatively stable Sierra Nevada-Great Valley block and the Peninsular Ranges separate the eastern areas, which have been extended and thinned, from western areas that are mostly underlain by diverse oceanic rock types.

[16] Let us clarify some essential details of this division. Some individual faults form the geological boundaries between blocks. However, we do not wish to divide earthquakes associated with a given fault into two separate subregions, so faults must lie within subregions, not along their borders. The Garlock fault is an example of this situation; it separates the tectonically and morphologically distinct Mojave Desert from the western Great Basin and southern Sierra Nevada. We outline the Garlock fault with its own narrow subregion. Similarly, the San Andreas fault subnetwork is generally continuous through southern California, transacting the entire southern California fault network, but is too large to be put into its own subregion as was done for the Garlock fault. Thus, we split the San Andreas among subregions that were otherwise defined. Some of the geologically most obvious subnetworks are too large. Examples of this are the faults west of the central San Andreas and in the northeastern part of southern California. These two subnetworks were divided into southern and northern parts; the northern parts are not considered further.

[17] Finally, given these considerations, we defined seven subregions shown in Figure 3a. The faults shown there are taken from Jennings [1977, 1994]. [18] The southwestern subnetwork includes mainly northwest striking dextral faults in the offshore continental borderland. The southeastern subnetwork also includes northwest striking dextral faults that cut through the Peninsular Ranges or lie within the strongly extended Salton Trough-northern Gulf of California region, where smaller sinistral faults and normal faults are present. The northwestern subnetwork includes mostly northwest striking dextral faults north of the Western Transverse Ranges, which separate it from the southwestern subnetwork. The Western Transverse Ranges are dominated by east striking sinistral and thrust faults. Between the southeastern and Garlock subnetworks is a complicated subnetwork that includes the Eastern Transverse Ranges and Mojave Desert. The Eastern Transverse Ranges are also dominated by east striking sinistral and thrust faults. The Mojave Desert is largely underforming in its westernmost part, but includes a broad array of northwest striking dextral faults in its central part and a small array of east striking sinistral faults in its northeastern corner. The seventh subnetwork includes north to northwest striking dextral and normal faults in the western Great Basin.

[19] What determines the number of subregions in our division? The basic factor is how the fault network is organized into a hierarchy. The nature of premonitory pattern Accord, as it was originally defined during the analysis of the CC model, demands to consider first only subregions of the highest rank. In case of observed fault network this corresponds to considering first only the major fault zones. The quantitative approach to such a division was suggested by Alekseevskaya et al. [1977]. Here we followed qualitative criteria, described above, which appear to be sufficient for the subregions at the scale we considered.

3.2. Pattern Accord and the Three Largest Earthquakes

[20] Here, we juxtapose the pattern Accord and the three largest earthquakes in southern California during the time considered: Kern County, $M = 7.7$, 21 July 1952; Landers, $M = 7.6$, 28 June 1992; and Hector Mine, $M = 7.6$, 16 October 1999. We explore whether the pattern Accord could...
predict them. With our choice of magnitude this would correspond to $M_0 = 7.5$.

### 3.2.1. Data Source

[21] We analyze the data from the U.S. Geological Survey-National Earthquake Information Center (USGS-NEIC) earthquake catalog for the time period 1 January 1928 to 30 January 2000. The data for the period from 1 January 1928 to 31 December 1989 are from the USGS-NEIC Global Hypocenter Data-Base (1990, available at http://neic.usgs.gov/products_and_services.html) with duplicates removed as described by Shebalin [1992]. The data for the period from 1 January 1990 to 30 January 2000 are from the NEIC/Preliminary Determination of Epicenters reports (monthly reports at http://gldfs.cr.usgs.gov/pde; weekly reports and QED at http://gldfs.cr.usgs.gov/weekly). The resulting catalog is preprocessed by the same rules as in other prediction algorithms of that kind [Keilis-Borok and Shebalin, 1999]: (1) The catalog contains magnitudes in different scales. For each earthquake we use the largest value; for the region and magnitude range considered this is practically equivalent to using Ms scale. (2) The aftershocks are eliminated from the catalog. We choose here an intentionally crude definition of aftershocks that has been used for other premonitory patterns in this region [Molchan et al., 1990; Keilis-Borok and Rotwain, 1990]. More precise definitions of aftershocks are described by Molchan and Dmitrieva [1991], Keilis-Borok et al. [1980], Gardner and Knopoff [1974], and Knopoff et al. [1982]. Distribution of main shocks among the subregions is illustrated in Figure 3b.

### 3.2.2. An Example of Prediction

[22] To apply the prediction algorithm, we need to choose the values of its numerical parameters, defined in section 2. Let us consider first the values given in the first row of Table 3. Function $A(t)$ computed with these parameters is shown in Figure 4. We see that $A(t)$ exceeds 5 only in the vicinity of the earthquakes targeted for prediction. This suggest the threshold $C_A = 5$; the corresponding value of $q$ is 80%. With this threshold each strong earthquake is preceded by an alarm, and there are no false alarms. The total duration of alarms is 29 years or 41% of the time considered. Figure 4 may sound not very impressive. Note, however, that we consider here the three strongest earthquakes of the century. More than twofold reduction of the time might be useful in different ways: (1) as initial approximation for more accurate prediction, (2) for instal-

---

**Figure 3.** The fault network of southern California. (a) Faults and subregions (fault zones). (b) Subregions and seismicity. The epicenters are shown for the period 1928–2000. See sections 3.1. and 3.2.

**Table 3.** Numerical Parameters of the Algorithma

<table>
<thead>
<tr>
<th>$M_0$</th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$\lambda$, years</th>
<th>$\Delta$, years</th>
<th>$q$,%</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>5.0</td>
<td>7.4</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>7.5</td>
<td>5.0</td>
<td>6.4</td>
<td>2</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*See sections 3.2.2., 3.3.1, and 3.5 and Figures 4, 6, and 10.*
lation of additional observations, and (3) for low-key enhancement of earthquake preparedness measures. Moreover, as is shown in section 3.2.3, this time can be reduced to 25% without loss of prediction stability.

3.2.3. Stability of Prediction

[23] For a single set of parameters the above scores, per se, are of little relevance, because parameters of the algorithm are retrospectively data fitted. We have to explore how stable such prediction is to variation of parameters. For that purpose we repeat prediction using different combinations of parameters’ values and compare the results. The pivotal tool for such an analysis is the error diagram introduced in earthquake prediction studies by Molchan [1997]. Its definition follows.

[24] Consider prediction made during a time interval $T$. Within that time $N$ strong events occurred and $N_f$ of them were not predicted; altogether $A$ alarms were declared and $A_f$ of them were false; the total duration of alarms was $D$ time units. The error diagram brings together the three major scores, characterizing the quality of prediction: the relative duration of alarms $\tau = D/T$; the rate of failures to predict $n = N_f/N$; and the rate of false alarms $f = A_f/A$. A prediction algorithm, with adjustable parameters fixed, corresponds to a point in the $(\tau, n, f)$-space. We plot an error diagram in two projections: $(\tau, n)$ and $(f, n)$. The random prediction, when we randomly pick up $p\%$ of the time to be covered by alarms, corresponds to the diagonal $n + \tau = 1$.

[25] First, each adjustable parameter was varied independently of the others, within a broad range indicated in Table 4, test a. The results are summed up on the error diagram in Figure 5. Dots correspond to different sets of parameters’ values. Note that parameters of our algorithm might be not independent. For example, reducing the “memory” $l$, we have to reduce also the threshold $C_0$ in order to produce the same prediction scores. The following relation between parameters $l$ and $R$ is obtained empirically:

$$R = \lambda 0.55 10^{1.5}.$$  

(4)

The stars in Figure 5 correspond to predictions with $R$ calculated by this relation while the other parameters are varied independently; specific values of parameters are given in Table 5; 36 variants are considered altogether. The obvious improvement of the results stresses the necessity to explore the connections among the adjustable parameters.

[26] Could the parameters of our algorithm be interdependent? Equation (4) above illustrates that this is indeed possible. Could our results be artificially improved by other correlations that are overlooked? Quite to the contrary: as clearly illustrated by Figure 5, the results are greatly improved when the correlation of parameters is known. Not knowing the interdependence of parameters, one can only increase the number of unacceptable versions of an algorithm (like those lying close to the diagonal of the error diagram in Figure 5).

Table 4. Test of Stability: Independent Variation of Parameters

<table>
<thead>
<tr>
<th></th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$s$, years</th>
<th>$B$</th>
<th>$\lambda$, years</th>
<th>$\Delta$,</th>
<th>$q$, %</th>
<th>$R$</th>
<th>Sections and Figures in Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test A</td>
<td>From</td>
<td>4.5</td>
<td>7.0</td>
<td>5</td>
<td>0.9</td>
<td>10</td>
<td>2</td>
<td>80</td>
<td>15</td>
</tr>
<tr>
<td>To</td>
<td>5.1</td>
<td>7.4</td>
<td>15</td>
<td>0.9</td>
<td>$\infty$</td>
<td>3</td>
<td>90</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Test B</td>
<td>From</td>
<td>4.5</td>
<td>5.5</td>
<td>2</td>
<td>0.9</td>
<td>5</td>
<td>0.5</td>
<td>70</td>
<td>5</td>
</tr>
<tr>
<td>To</td>
<td>4.5</td>
<td>6.1</td>
<td>4</td>
<td>0.9</td>
<td>10</td>
<td>0.5</td>
<td>70</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>To</td>
<td>5.1</td>
<td>6.4</td>
<td>10</td>
<td>0.9</td>
<td>$\infty$</td>
<td>2</td>
<td>90</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>To</td>
<td>5.5</td>
<td>7.0</td>
<td>12</td>
<td>0.9</td>
<td>$\infty$</td>
<td>2</td>
<td>90</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Test C</td>
<td>From</td>
<td>4.0</td>
<td>7.4</td>
<td>10</td>
<td>0.9</td>
<td>5</td>
<td>5</td>
<td>80</td>
<td>10</td>
</tr>
<tr>
<td>To</td>
<td>5.5</td>
<td>7.4</td>
<td>10</td>
<td>0.9</td>
<td>$\infty$</td>
<td>6</td>
<td>90</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Test D</td>
<td>From</td>
<td>4.5</td>
<td>6.4</td>
<td>2</td>
<td>0.8</td>
<td>5</td>
<td>2</td>
<td>90</td>
<td>10</td>
</tr>
<tr>
<td>To</td>
<td>5.2</td>
<td>7.0</td>
<td>7</td>
<td>1</td>
<td>$\infty$</td>
<td>5</td>
<td>95</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Precursor Accord and the three strongest earthquakes in southern California. Precursor is depicted by the function $A(t)$ (see equation (3)). Vertical lines show the moments of strong earthquakes. Horizontal line shows the threshold $C_A = 5$ for declaring an alarm. Periods of alarm are highlighted.
3.3. Pattern Accord and Earthquakes With Magnitude 6.5–7.4

[27] Here we explore in a similar way the emergence of the pattern Accord before earthquakes with magnitudes from 6.5 to 7.4. Eleven such earthquakes occurred during the time period considered. Two of them, Colorado River Delta of 31 December 1934, $M = 7.0$, and Superstition Hills of the 24 November 1987, $M = 6.7$, are eliminated from consideration, because they follow the preceding strong earthquakes within a too-short time interval, 29 and 12 hours respectively. Such intervals are incompatible with characteristic resolution time of our analysis, which is of order of months.

3.3.1. An Example of Prediction

[28] We consider for a start the values of adjustable parameters given in the second row of the Table 3. All the parameters are readjusted to the new target magnitude in about the same way as was done for other premonitory patterns [Keilis-Borok et al., 1980]. Figure 6 shows the function $A(t)$ and alarms determined with these parameters. Six out of nine strong earthquakes are “predicted” with no false alarms; three earthquakes are missed. The total duration of alarms is 20 years or 28% of the time considered.

[29] The question naturally arises: Why is this value less than that obtained above in section 3.2.2 for the large and presumably easier to predict earthquakes? In fact, 28% correspond to the case when three earthquakes out of nine were not predicted, while for the strongest earthquakes we have no failures to predict. As shown in section 3.3.2 below, the duration of alarms increases up to 62% if parameters of the algorithm are readjusted to predict all nine earthquakes.

3.3.2. Stability of Prediction

[30] We varied adjustable parameters of the algorithm within the limits indicated in Table 4, test b (lines 1 and 4). Acceptable prediction score ($\tau < 0.4; n < 0.4; f < 0.5$) is obtained within more narrow limits indicated in lines 2 and 3 of Table 4, test b. The corresponding error diagram is shown in Figure 7. Again, we see that the quality of prediction remains acceptable not only for some special combinations of parameters but for a reasonably wide domain in parameter space. Thus the prediction is stable. In particular, parameters $s$, $\lambda$, and $R$ can be varied much more than one could dare to hope. It is encouraging also that combinations leading to good scores are not scattered randomly over parameter space but form a connected cluster. The overall quality of prediction is slightly lower than that for the three largest earthquakes (see section 3.2).

3.4. Variation of Basic Definitions

[31] In sections 3.2 and 3.3 above we varied adjustable parameters of the algorithm to check its stability. However, we also have freedom of choice in determination of sub-regions and in the choice of the function that depicts our pattern. In this section we introduce an alternative regionalization based entirely on the territorial distribution of seismicity and explore to what extent the seismic regionalization is different from the geological one and how this difference affects the result. We consider also an alternative measure of correlation of seismic activity. The whole previous analysis, including the stability tests, is repeated here with these two alternative choices.

3.4.1. Alternative Subregions

[32] Consider the case of zoning, based solely on territorial distribution of seismicity. Figure 8 shows sub-

---

Table 5. Test of Stability: Parameter $R$ is Determined by $\lambda^a$

<table>
<thead>
<tr>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$s$, years</th>
<th>$B$</th>
<th>$\lambda$, years</th>
<th>$\Delta$, years</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5</td>
<td>7.0</td>
<td>10</td>
<td>0.9</td>
<td>10</td>
<td>3</td>
<td>80</td>
</tr>
<tr>
<td>5.0</td>
<td>7.4</td>
<td>12</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>7.1</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$See section 3.2.3 and Figure 5.

---

Figure 5. Error diagram for retrospective prediction of the three strongest earthquakes in southern California, 1928–2000. Dots correspond to independent variation of parameters; stars to empirical relation between $R$ and $\lambda$. See section 3.2.3.
regions obtained this way; note that there are only five of
them, compared to seven in Figure 3. The seismological
regionalization is less detailed because the clouds of epi-
centers associated with different faults merge or even
overlap.

3.4.2. Alternative Measure of Correlation

It is defined as follows:

\[
A^* (t) = \left( r_1^2 + r_2^2 + \cdots + r_n^2 \right)^{-1},
\]

\[
i = 1, \cdots , n; \sum_i (t) = \Sigma_i (t) + \Sigma_2 (t) + \cdots + \Sigma_n (t).
\]

Function \( A^* (t) \) shows how uniform the distribution of
seismicity is among the subregions. It is easy to see that
\( A^* (t) \) takes the maximal value of \( n \) when seismic activity
\( \Sigma_i (t) \) is the same in all subregions and decreases to the
minimal value of unity when activity is totally concentrated
within a single subregion.

3.4.3. Stability of Prediction

With two alternative sets of boundaries (Figures 3 and 8) and two measures of correlation (\( A \) and \( A^* \)) we have
four versions of the prediction algorithm. One of them,
which uses geological zoning and the measure \( A(t) \), was
considered above.

Figure 9 shows the error diagrams for prediction by
the other three versions. The ranges of parameters’ variation
are shown in Table 4, tests c and d. Table 4, test c, is related
to parameters scaled for prediction of the three largest
earthquakes, while Table 4, test d, is related to parameters
scaled for prediction of earthquakes with magnitudes greater
than 6.5.

The following results are noteworthy: (1) All four
versions give almost the same errors’ scatter; this confirms
the robustness of the algorithm. (2) The points on the error
diagram are well separated from the diagonal, that is, from
random prediction. (3) The rate of false alarms is small.
These results are highly encouraging for further exploration
of the pattern Accord.

3.5. What Does Work in the Pattern Accord?

Here we compare the pattern Accord with the activity
of the whole southern California measured by the
function \( \Sigma(t) \) described above (section 1.2.2). The functions

Figure 6. Precursor Accord and earthquakes with magnitudes from 6.5 to 7.4 in southern California.
“Predicted” earthquakes are marked with solid triangles; “unpredicted” are marked with empty triangles.
Other notations are the same as in Figure 4.

Figure 7. Error diagram for retrospective prediction of the strong (6.5 \( \leq M < 7.5 \)) earthquakes in
southern California, 1928–2000. Parameters of the algorithm are varied independently. Dots correspond
to the broad range of the variation; stars correspond to the narrowed range. See section 3.3.2.
\( \Sigma(t) \) and \( A(t) \) are juxtaposed in Figure 10. Both functions are calculated with the same values of numerical parameters, given in Table 3. Figure 10a and the first line of Table 3 correspond to prediction of the three strongest earthquakes \( (M \geq 7.5) \), while Figure 10b and the second line of Table 3 correspond to prediction of earthquakes with \( 6.5 \leq M \leq 7.4 \). Alarms declared by the pattern Accord are shown at the top panel of each figure. Figure 10a demonstrates that both functions \( \Sigma(t) \) and \( A(t) \) have two explicit peaks over the considered time interval: one of these peaks precedes Kern County earthquake, and another precedes Landers and Hector Mine. Figure 10b demonstrates a more diverse behavior. It shows that the pattern Accord is not equivalent to the total activity measured by \( \Sigma(t) \). There are periods where \( A(t) \) is abnormally high while \( \Sigma(t) \) remains within its usual boundaries (e.g., 1940–1944), and vice versa (e.g., 1981–1983). At the same time, there are periods where both the functions are abnormally high (e.g., 1994–1995) or low (e.g., 1959–1961) simultaneously. Similar results are obtained with different sets of numerical parameters.

Figure 8. Alternative zoning of southern California based solely on seismicity. The epicenters are shown for the period 1928–2000.
fore what does work in the pattern Accord is not only the general rise of activity but its spreading over the faults’ network as well.

[39] As one can see in Figure 10b, this version of the pattern Accord fails to predict three earthquakes: Colorado River delta of 31 December 1934, Imperial Valley of 15 October 1979, and Coalinga of 2 May 1983. The natural question for future study is: What is the reason for these errors (particularly the two recent ones that are relatively better studied)?

3.6. Alarm From a Decision Maker’s Point of View

[40] A prediction algorithm has errors; that is certainly the case for currently known algorithms and possibly is inevitable [Holland, 1995; Gell-Mann, 1994; Turcotte, 1997; Rundle et al., 2000b]. There is, however, a freedom to
choose a trade-off between different kinds of errors as illustrated in the error diagrams of Figures 5, 7, and 9. What version of a prediction algorithm is “the best”? Let us consider that question from the point of view of a decision maker who has to choose the safety measures undertaken in response to a prediction. The question arises because earthquake preparedness requires a wide variety of safety measures; see their overview in Kantorovich and Keilis-Borok [1991]. The choice of safety measures depends on two factors: (1) expected damage and (2) reliability of the prediction: probabilities of false alarm and failure to predict and duration of the alarm.

[41] Different measures are justified in different cases. Some critical objects (e.g., high dams) have to be protected even at high cost, and expensive safety measures are justified even if the probability of false alarm is high. In other cases
the safety measures are justified only if the probability of false alarm is small. The choice of disaster preparedness measures belongs to a decision maker responsible for disaster management; Molchan [1997] suggested a theoretical framework for optimization of a response to prediction. Accordingly, contrary to the established tradition, we propose here not to select a single version of the algorithm, that is, a “best” set of adjustable parameters, but provide a decision maker with the whole available information. Then prediction at a given time moment is presented as the error diagram marking the versions yielding alarms. An example of such a presentation is shown in Figure 11. It shows performance of several versions of the prediction algorithm from Figure 7. Comparing Figure 11 to Figure 7, we retained only the versions that give minimal duration of alarm for the same rate of failures to predict. Panels refer to different time moments: panel a refers to a moment one month before the Imperial Valley earthquake of 15 October 1979, \( M = 7.0 \). (b) Alarms within a month before the Elmore Ranch, \( M = 6.5 \), and Superstition Hills earthquakes, \( M = 6.7 \), of 24 November 1987.

Figure 11. The end-user-oriented presentation of prediction. Error diagrams are the same as in Figure 7. Two versions of prediction that give minimal \( \tau \) are considered for each value of \( n \). Stars indicate versions of prediction algorithm, which give an alarm at the moment considered. Other versions are indicated by triangles. (a) Alarms within a month before the Imperial Valley earthquake of 15 October 1979, \( M = 7.0 \). (b) Alarms within a month before the Elmore Ranch, \( M = 6.5 \), and Superstition Hills earthquakes, \( M = 6.7 \), of 24 November 1987.
ber 1979; panel b refers to one month before the Elmore Ranch and Superstition Hills earthquakes of 24 November 1987.

[42] A new element on these diagrams is the following: we show which version of prediction algorithm produces an alarm at the given moment (stars) and which versions do not (triangles). In these particular cases predictions are very stable, being generated by a majority of the versions considered. Implications for the decision maker are beyond the scope of the present paper.

4. Discussion

[43] 1. The definition of the algorithm Accord, developed first for the abstract model of colliding cascades [Gabrielov et al., 2000a], is adapted here to analysis of observations. The retrospective application of the algorithm to seismicity of southern California is highly encouraging. Predictions yield acceptable success-failure scores in a reasonably wide range of adjustable parameters and choice of subregions. Thus this paper establishes a hypothesis that may be tested on independent data. In support of our hypothesis is the fact that it was first found in a model [Gabrielov et al., 2000a], then in a shorter timescale in real seismicity [Shebalin et al., 2000], and finally in an intermediate timescale considered here. The final test would be the advance prediction.

[44] 2. It is encouraging also that the pattern Accord is directly connected with geometry of fault networks. Two mechanical explanations of that pattern may be hypothetically suggested: (1) strength drop (weakening) at the fault’s segments broken by the earthquakes and (2) stress redistribution after the earthquakes. If either of them occurs in several subregions, it is likely to affect the equilibrium of a whole system more than the equivalent change in a single subregion.

[45] 3. There are further as-yet unexplored possibilities to improve the prediction algorithm described here. (1) The size of the fault network considered probably has to be scaled by the size \( L(M) \) of the source of the earthquakes which we are trying to predict. (2) One may consider instead of \( \Sigma \) the pattern Active Zone Size (AZS). Its difference from \( \Sigma \) is that overlapping parts of several fault breaks are counted only once. This pattern was suggested by Kossobokov and Carlson [1995], who found that AZS yields better predictions than \( \Sigma \). (3) It would be most interesting to consider instead of fault subregions the “nodes”: fractured structures, formed around faults’ intersections and junctions [Gabrielov et al., 1996, and references therein]. As was demonstrated by Gelfand et al. [1976] and Gorskov et al. [2001, and references therein], strong earthquakes are nucleated mainly in such nodes. Moreover, the tendency of the fault system to fracturing and stress accumulation is concentrated at fault intersections [Gabrielov et al., 1996; Minster and Jordan, 1984].

[46] 4. The pattern Accord is a manifestation of a broader phenomenon: premonitory increase of earthquake correlation range. This phenomenon has been formally defined and statistically analyzed for the first time during the analysis of the CC model, although it was previously hypothesized by Keilis-Borok [1994, 1996]. Its qualitative explanation by analogy with phase transitions in statistical physics was suggested by Turcotte et al. [2000] and Bowman et al. [1998]. Zoeller et al. [2001] have reported the growth of correlation range prior to the strong earthquakes for observed seismicity of southern California. The CC model suggests also a similar phenomenon in the shorter timescale of months; its first test on observations [Shebalin et al., 2000] is encouraging too.

[47] 5. Although long-range correlations are well known in seismology, they are sometimes regarded as counterintuitive in earthquake prediction studies. This is probably due to the fact that they cannot be explained by the linear elastic model of stress redistribution after a rupture. In such models redistribution is confined to the distances comparable with the rupture dimension (the “Saint Venant principle”). However, many other mechanisms, such as mentioned in section 1.1, do explain long-range correlations. It is interesting to note in this context that Charles Richter, who was generally skeptical about feasibility of earthquake prediction, made an exception to the pattern Sigma specifically because it was based on long-range correlations. [Richter, 1964, p. 3025] wrote: “It is important that [the authors] confirm the necessity of considering a very extensive region including the center of the approaching event. It is very rarely true that the major event is preceded by increasing activity in its immediate vicinity”.

[48] 6. More evidence of long-range correlations comes from statistical modeling. It was shown recently by Liu et al. [1999] that the Stress Release model, introduced by Vere-Jones [1978], does fit seismicity of North China much better if long-range correlations are introduced.

[49] 7. Contrary to the established tradition, it seems better not to choose for prediction the “best” version of an algorithm, but to use, in parallel, different acceptable versions. Then one may present a prediction by the error diagram marking the versions yielding the current alarm (Figure 11). Such representation is important for a decision maker, who has to choose preparedness measures to be undertaken in response to prediction.

Acknowledgments. We are grateful to M. Harrison, P. Bird, A. Gabrielov, and W. Newman for inspiring discussions and constructive criticism. This study was partially supported by a subcontract with Cornell University, Geological Sciences, under grant EAR-9804859 from the National Sciences Foundation. This study was also partially supported by grants from the International Science and Technology Center (Project 1293-99), USGS/SCC, INTA 0748, and James S. McDonnell Foundation.

References


G. Axen, Department of Earth and Space Sciences, University of California, Los Angeles, 595 Circle Drive East, Los Angeles, CA 90095-1567, USA. (gaxen@ess.ucla.edu)

V. Keilis-Borok and I. Zaliapin, Institute of Geophysics and Planetary Physics, University of California, Los Angeles, 3845 Slichter Hall, Los Angeles, CA 90095-1567, USA. (vkb@ess.ucla.edu; zal@ess.ucla.edu)