Journal of Statistical Physics, Vol. 111, Nos. 3/4, May 2003 (© 2003)

A Boolean Delay Equation Model of Colliding Cascades. Part II: Prediction of Critical Transitions

Ilya Zaliapin,[1] Vladimir Keilis-Borok,[2] and Michael Ghil[3]

Received October 23, 2001; accepted October 11, 2002

We consider here prediction of abrupt overall changes ("critical transitions") in the behavior of hierarchical complex systems, using the model developed in the first part of this study. The model merges the physical concept of *colliding cascades* with the mathematical framework of *Boolean delay equations*. It describes critical transitions that are due to the interaction between *direct cascades of loading* and *inverse cascades of failures* in a hierarchical system. This interaction is controlled by distinct delays between switching of elements from one state to another: loaded vs. unloaded and intact vs. failed.We focus on the earthquake prediction problem; accordingly, the model's heuristic constraints are taken from the dynamics of seismicity. The model exhibits four major types of premonitory seismicity patterns (PSPs), which have been previously identified in seismic observations: (i) rise of earthquake clustering; (ii) rise of the earthquakes' intensity; (iii) rise of the earthquake correlation range; and (iv) certain changes in the size distribution of earthquakes (Gutenberg–Richter relation). The model exhibits new features of individual PSPs and their collective behavior, to be tested in turn on observations. There are indications that the premonitory phenomena considered are not seismicity-specific, but may be common to hierarchical systems of a more general nature.

KEY WORDS: Cellular automata; colliding cascades; delay equations; earthquake precursors; hierarchical modeling; prediction of complex behavior.

[1] International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, Russia, and Institute of Geophysics and Planetary Physics, University of California, Los Angeles 90095-1567, USA; e-mail: zal@ess.ucla.edu

[2] International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, Russia, and Institute of Geophysics and Planetary Physics and Department of Earth and Space Sciences, University of California, Los Angeles, USA; e-mail: vkb@ess.ucla.edu

[3] Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, USA, and Département Tere-Atmosphère-Océan and Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, Paris, France; e-mail: ghil@atmos.ucla.edu

0022-4715/03/0500-0839/0

1. INTRODUCTION

This is the second part of a study aimed at a prominent feature of hierarchical nonlinear complex systems: persistent recurrence of abrupt overall changes, called here "critical transitions." The study's first part[1] introduced a model for the development of critical transitions. The model is based on the physical concept of colliding cascades, in which direct cascades of loading interact with inverse cascades of failures to produce critical transitions.[2, 3] The mathematical framework of Boolean delay equations[4, 5] allowed us to replace complex elementary interactions by their robust integral effect, which takes the form of delayed switches in the state of the system.

The model's heuristic constraints are taken from the dynamics of seismicity. The seismically active crust of the Earth is regarded as a complex, hierarchical, dissipative system and strong earthquakes as critical transitions in the system. In ref. 1 the model was applied to the study of multiple seismic regimes. Here, we consider the problem of predicting critical transitions.

The point of departure is provided by the studies in algorithmic prediction of large earthquakes, based on analyzing seismicity in a lower magnitude range.[6–8] We apply the previously developed prediction algorithms to the model's seismicity. The good performance of these algorithms in predicting synthetic seismicity confirms that the model behavior captures essential features of observed seismicity used in developing the algorithms. Besides, we explore new premonitory phenomena that can be tested on observations. The major elements of our analysis are briefly described below.

1.1. General Prediction Scheme

(i) Prediction of large earthquakes is based on the sequence of smaller earthquakes in a given area, called a catalog

$$\{(t_e, m_e, h_e): e = 1, 2,\ldots; t_e \leqslant t_{e+1}\}, \tag{1}$$

here t_e is the starting time of the rupture; m_e is the magnitude—a logarithmic measure of the energy released by the earthquake; and h_e is the position vector of the hypocenter. The latter represents a point approximation of the area where the rupture started. The synthetic counterpart of such a catalog for our model is defined in Eq. (5) of ref. 1.

(ii) Premonitory seismicity patterns (PSPs) are captured by functionals $F_k(t, s_k)$, $k = 1, 2,\ldots$, defined on the catalog (1). With a few exceptions, the functionals are defined in a sliding time window $(t - s_k, t)$. The

value of a functional is attributed to the end t of the window. The specific functionals F_k used in this study are defined below.

Emergence of a PSP is defined by the condition

$$F_k(t) \geqslant C_k. \tag{2}$$

Here $F_k(t)$ is one of the functionals that depict the PSP; the threshold C_k is usually chosen as a certain percentile of the functional's distribution.

(iii) An *alarm* is declared for time interval of length τ_k when a single pattern or a certain combination of patterns emerges. The alarm is terminated after a major earthquake occurs or the time τ_k expires, whichever comes first. The possible outcomes of such prediction are shown in Fig. 1. Obviously this scheme lends itself to the use of other data, not necessarily seismological ones (e.g., refs. 9 and 10).

1.2. Four Types of Premonitory Seismicity Patterns (PSPs)

Studies of observed and modeled seismicity have demonstrated that an earthquake of magnitude m is often preceded by specific PSPs formed within an area and magnitude range depending on m.

We consider here PSPs of the following four types (see Table I):[6, 7, 11–15] (i) rise of seismic activity; (ii) rise of earthquake clustering; (iii) rise of earthquake correlation range; and (iv) certain changes in the earthquakes' size distribution (G-R relation). Patterns of the first two types have been found in observations first and then in models;[16–22] patterns of the other two types—first in models and then in observations.[2, 23–28] The intermediate-term patterns of the first two types have been validated by statistically significant predictions of real earthquakes.[29–31] Other patterns are in different stages of testing.

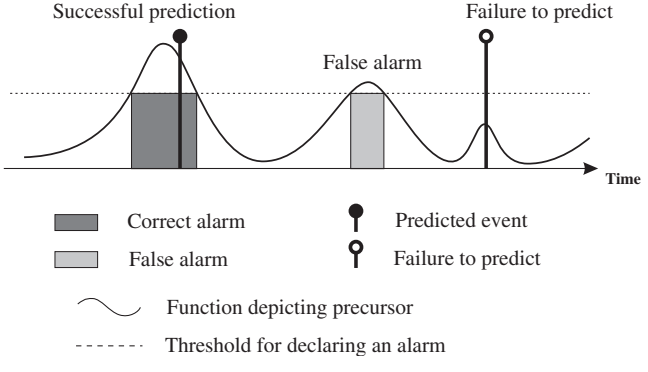

Fig. 1. Possible outcomes of prediction.

Table I. Premonitory Seismicity Patterns (PSPs) Considered in this Study

Notation	Description	Type	Definition in the text	References
N_m	Number of events	Intensity	Section 3.1.1	40, 63, 61
Σ_m	Weighted number of events; coarsely estimates the area of faultbreaks	Intensity	Eq. (3a)	39, 63, 61
B_m	Weighted number of immediate aftershocks	Clustering	Eq. (7)	45
R_m	Near-simultaneous occurrence of distant events	Correlation range	Section 3.3	26
A_m	Simultaneus activation of distinct branches of a system	Correlation range	Section 3.3	27
Π_m	Total activity of most active branches of a system	Correlation range	Section 3.3	This study
W	Ratio of N_m for different m	Transformation of GR relation	Eq. (8)	63

In lieu of an adequate theory, the PSPs considered here have been found by a pattern recognition analysis of observed and/or modeled seismicity. This analysis is based on the methodology developed by the school of I. M. Gelfand[32–34] for the study of rare phenomena of highly complex origin, a situation where classical statistical methods are inapplicable. It is in a way akin to the exploratory data analysis developed by J. Tukey.[35]

1.3. Error Diagrams

E. Fermi coined the phrase *"With four exponents I can fit an elephant"* to highlight the problems generated by using too large a number of parameters to fit a given data set. The pattern recognition approach allows considerable freedom in the retrospective development of a prediction algorithm. To avoid the problems highlighted by Fermi, such an algorithm has to be insensitive to the variations in its adjustable parameters, choice of premonitory phenomena, definition of alarms, etc. The sensitivity tests comprise an exhaustive set of numerical experiments, which represent a major part in the algorithm's development.

Molchan[36] introduced a particular type of error diagrams to evaluate earthquake prediction algorithms. The definition of such an error diagram is the following: Consider the outcomes of prediction during a time interval of length T. During that time interval, N strong events occurred and N_F of

them were not predicted. The number of declared alarms is A, with A_F of them being false alarms. The total duration of alarms was D. The error diagram shows the trade-off between the relative duration of alarms $\tau = D/T$, the rate of failures to predict $n = N_F/N$, and the rate of false alarms $f = A_F/A$. In the (n, τ)-plane, the straight line $n+\tau = 1$ corresponds to a random binomial prediction—at each epoch an alarm is declared with probability τ and is not declared with probability $1-\tau$.

Different points in the error diagram correspond to different variants of a prediction algorithm. In particular, the variants may differ by the value of the threshold for declaring an alarm (see Fig. 1): the higher the threshold, the shorter D. On the other hand, this may lead to a higher N_F. An algorithm is useful if: (i) the prediction score is higher than the random one, i.e., the corresponding point in the error diagram lies below the line $n+\tau = 1$; and (ii) this score is fairly insensitive to variations of the algorithm.

Error diagrams also provide the basis for the optimization of disaster preparedness measures to be undertaken in response to prediction.[36, 37] Error diagrams for the PSPs considered in this study are given in Section 3.

1.4. Four Paradigms

The previous studies led to the following findings that are important for a fundamental understanding of the dynamics of seismicity, as well as for further prediction research;[7, 8, 38] (i) Long-range correlations occur in fault system dynamics and, accordingly, premonitory phenomena arise over large areas. (ii) There exist four types of premonitory phenomena (see Section 1.2). (iii) These phenomena exhibit at least partial similarity worldwide. (iv) Premonitory phenomena are dual in nature: some of them, like the PSPs considered here, are common to a wide class of nonlinear systems; others are specific to the geometry of a fault network or to a certain physical mechanism that controls the stress and strength field in the lithosphere.

The remainder of this paper is organized as follows: In Section 2 we discuss the synthetic earthquake sequence that we study. The PSPs used for prediction are defined in Section 3, and their individual performance is discussed. Section 4 describes the collective performance of PSPs in the modeled seismicity. A discussion of the results and their implications for earthquake prediction follows in Section 5.

2. EARTHQUAKE SEQUENCE

In this study we analyze the synthetic earthquake sequence shown in Fig. 2. It represents the model's intermittent regime—the most interesting

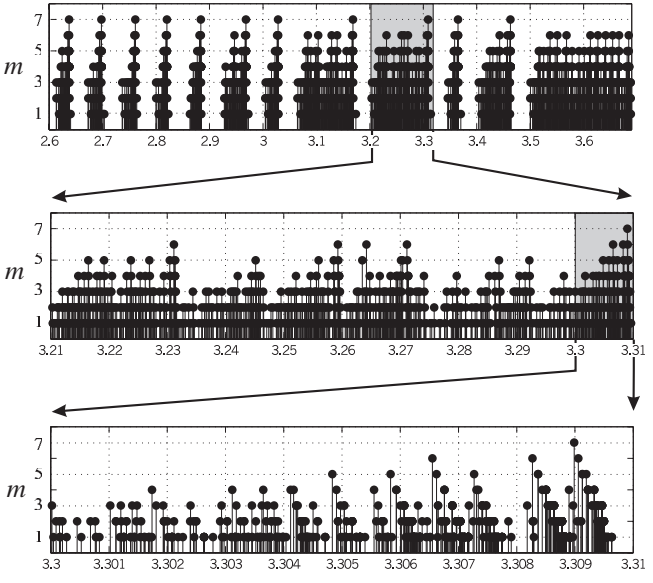

Fig. 2. Synthetic earthquake sequence, consecutively zoomed. The sequence is generated by the colliding-cascade model of Part I[(1)] for the parameter values $L = 7$, $\lambda = 0.2 \cdot 10^{-4}$, $\Delta_L = 5 \cdot 10^3$, $\Delta_D = 5 \cdot 10^2$, $\Delta_H = 10^3$, $c = 2/3$, $p = 3$, $k = 3$. Shadowing marks the zoomed intervals. The model shows a rich variety of behavior at different time scales. Note that the entire sequence (not shown) is $5 \cdot 10^6$ time units long and that the difference in time scales betweeen the top and bottom panels is by a factor of 10^3.

and realistic one (see ref. 1 for details). The parameter values corresponding to this sequence are given in the caption to the figure; its duration is $5 \cdot 10^6$ time units.

The sequence includes 43 earthquakes with $m = 7$, the largest magnitude possible in this version of the model; they are the targets of prediction in the subsequent analysis. Premonitory patterns for the magnitude-7 earthquakes are formed by the earthquakes with smaller magnitudes, from $m = 6$ down to $m = 1$.

In observations, one distinguishes between foreshocks, main shocks, and aftershocks. Most of the premonitory patterns considered hare have been defined on the sequence of main shocks. However, the number of aftershocks is retained for each main shock. As in analyzing observations, the few immediate foreshocks are not differentiated from the main shocks.

An aftershock is defined in the model as a descendant of a main shock of magnitude m which occurred within $T(m)$ time units from the main shock. A descendant is a child, child of a child, etc., of a main shock within

the system considered. In accordance with this definition, events of magnitude $m = 1$, the smallest possible in the model, have no aftershocks. The values of $T(m)$ for different magnitudes m of a main shock are taken here as follows: $T(7) = 3000$, $T(6) = 1500$, $T(5) = 600$, $T(4) = 300$, $T(3) = 150$, and $T(2) = 70$.

3. PSPS AND THEIR INDIVIDUAL PERFORMANCE

3.1. Increase of Seismic Intensity

A premonitory rise of seismic flow intensity is captured by two types of functionals. They are described in the following two subsections.

3.1.1. Functionals Σ and N

Definitions. The following two functionals[39, 40] depict an increase of activity: the number $N_m(t, s)$ of main shocks of magnitude m within the time interval $(t-s, t)$ and the weighted number of main shocks

$$\Sigma_m(t, s) = \sum_{m'=1}^{m} S(m')\, N_{m'}(t, s). \tag{3a}$$

In analyses of observed seismicity, the weight $S(m)$ in (3a) is proportional to the source area of an earthquake of magnitude m. Here we use

$$S(m) = 10^{bm}, \qquad b = \log_{10}(3), \tag{3b}$$

as suggested by the ternary hierarchical structure of our system.

Performance. We computed each of these functionals for the sequence shown in Fig. 2 with $s = 3000$ and m varying from 1 to 6; note that, according to (3a), the functionals N_1 and Σ_1 coincide. Altogether we consider 11 functionals that depict increase of seimic activity. Error diagrams for prediction with these functionals are given in Figs. 3a–d. "Quiet" intervals, with seismicity below $m = 4$, are eliminated from consideration, to avoid artificial improvement of statistics: neither strong earthquakes nor alarms may emerge in our model during such intervals, which occupy 77% of the total duration of our earthquake sequence. The same is done for all the error diagrams in Fig. 3. For advance prediction of real earthquakes the overall performance shown in Figs. 3a–d would be quite satisfactory.

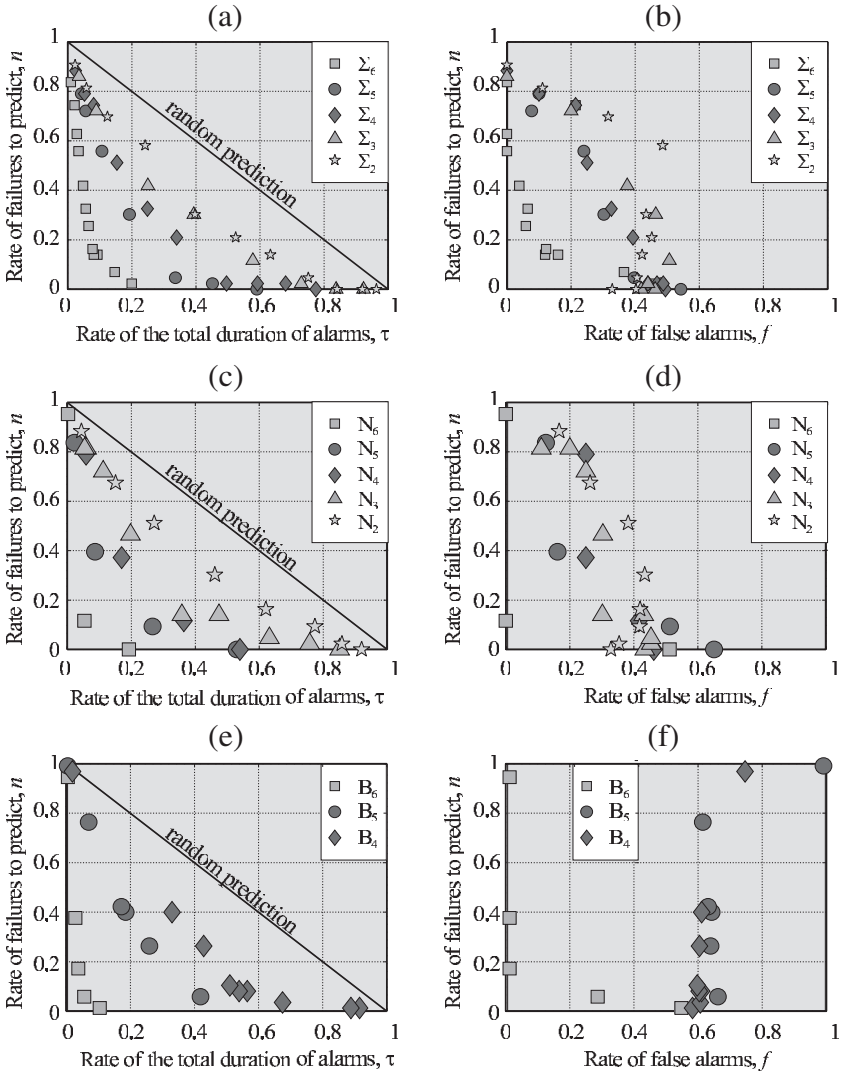

Fig. 3. Error diagrams for the premonitory patterns considered. The diagrams show the tradeoff between major characteristics of prediction: the relative duration of alarms τ, the rate of failures to predict n, and the rate of false alarms, f. Panels correspond to prediction with different single PSPs, with n vs. τ on the left and n vs. f on the right. (a), (b) Weighted number of events, Σ_m, $2 \leqslant m \leqslant 6$ (Section 3.1.1); (c), (d) number of events N_m, $m = 2,\ldots, 6$ (Section 3.1.1); (e), (f) bursts of aftershocks B_m, $m = 4,\ldots, 6$ (Section 3.2); (g), (h) Accord A_m, $m = 5, 6$, and Π (Section 3.3); (i), (j) correlation range R_m, $m = 2,\ldots, 6$ (Section 3.3); and (k), (l) change in magnitude distribution W (Section 3.4). For a given functional, the points in each diagram differ from each other only by the alarm threshold (not indicated); see text for details.

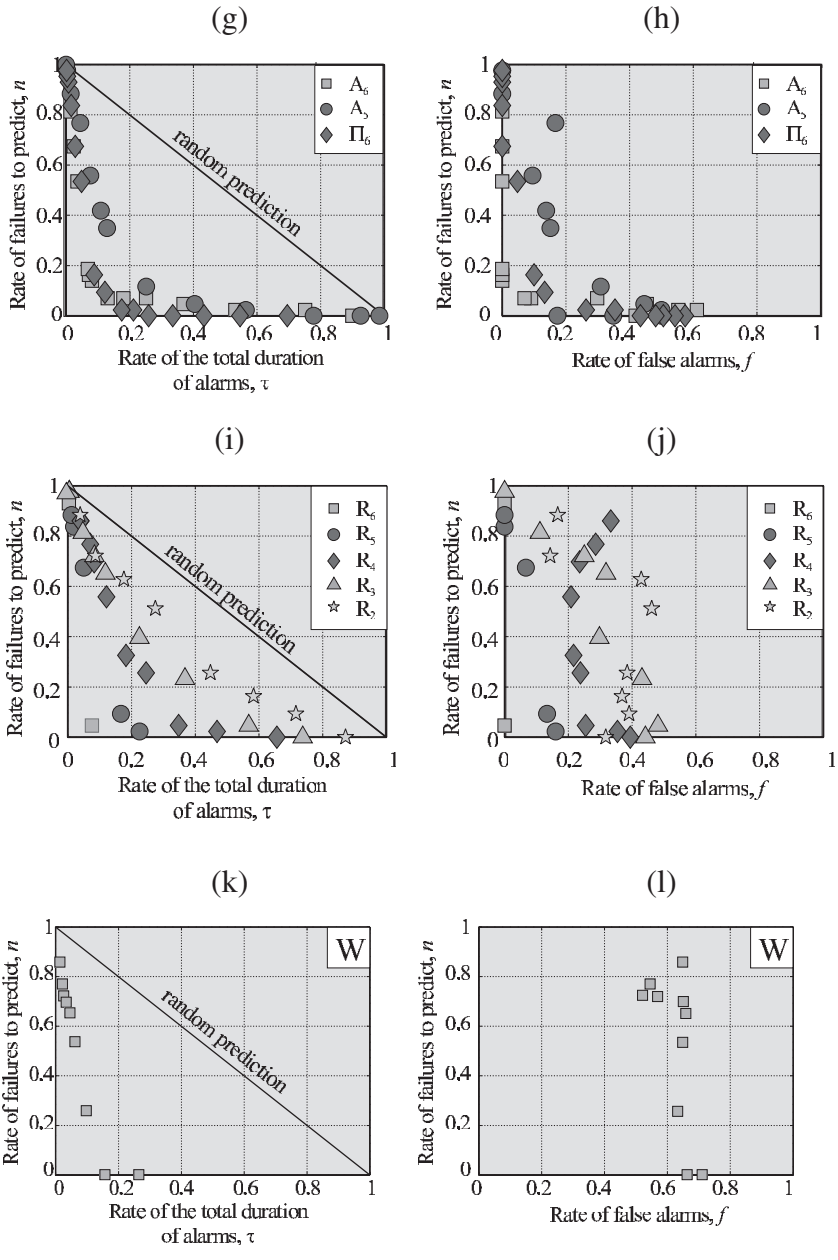

Fig. 3. (Continued.)

3.1.2. Accelerated Benioff Strain Release

Definition. The rise of seismic activity is also captured by the cumulative Benioff stress release.[8, 20, 41, 42] It is usually defined as

$$\epsilon(t) = \sum_e E_e^{1/2}, \qquad t_0 \leqslant t_e \leqslant t. \tag{4}$$

Here E_e is the energy and t_e is the occurrence time of the eth earthquake. The summation is taken over all the earthquakes in an area under consideration, including aftershocks. In the analysis of observations, the energy is estimated from the magnitude. The function $\epsilon(t)$ determined from observations is approximated by a power law[8, 41]

$$\epsilon(t) = A - B(t_f - t)^{\alpha}, \tag{5}$$

where t_f is the time of a large earthquake, targeted for prediction.

For the model we define the function $\epsilon(t)$ as

$$\epsilon(t) = \sum_{m'=1}^{m} S(m')\, N^*_{m'}(t, s). \tag{6}$$

Here $N^*_m(t, s)$ is the number of all earthquakes, including aftershocks, within the interval $(t-s, t)$; $S(m)$ was defined in (3b).

Composite Seismicity. The function $\epsilon(t)$ averaged over the 43 seismic cycles is shown in Figs. 4a, b. The figures confirm that the modeled seismicity does exhibit, on average, the premonitory power-law rise of Eq. (5) as a large earthquake approaches. Away from a large earthquake the function $\epsilon(t)$ is linear. The average power-law rise is accompanied by roughly periodic oscillations, shown in Fig. 4c. The amplitude of these oscillations grows as a large earthquake approaches (compare refs. 42 and 43).

In the present study, we do not explore the predictive skill of this functional for individual major earthquakes.

3.2. Increase of Clustering: "Bursts of Aftershocks"

Definition. Premonitory clustering prior to large earthquakes is captured by "pattern B" or "bursts of aftershocks."[44, 45] It consists of a

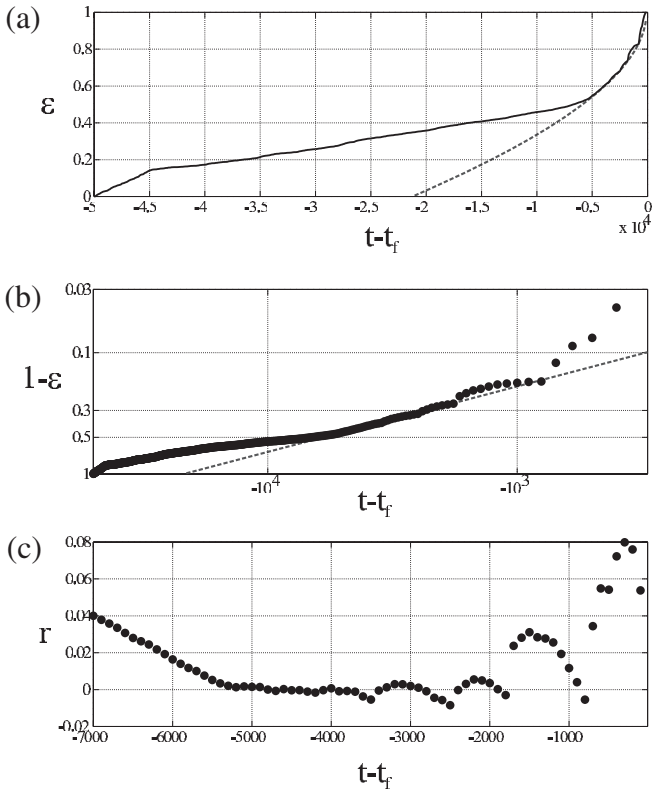

Fig. 4. Premonitory acceleration of cumulative Benioff stress release $\epsilon(t)$ vs. time to a major event $t-t_f$. Heavy solid (panel (a)) or filled circles (panels (b) and (c))—$\epsilon(t)$; light dashed line—least square fit of a power law $\epsilon(t) = A - B(t_f - t)^{\alpha}$. The values of $\epsilon(t)$ are normalized so that $A = 1$; the fitting curve in panels (a) and (b) corresponds to $B = 2.36$, $\alpha = 0.55$. (a) $\epsilon(t)$ on a linear scale; (b) $\epsilon(t)$ on a log-log scale; and (c) residuals, i.e., the difference between the data and the fitting curve. Note the accelerating oscillations prior to a major earthquake in panel (c).

main shock of medium magnitude with a large number of aftershocks. One of the functionals that depicts this premonitory phenomenon is

$$B_m(t_l, \delta) = \sum_{m'=1}^{m} S(m')\, M_l(m'). \tag{7}$$

Here $M_l(m')$ is the number of those aftershocks of the lth main shock that have magnitude m', while $S(m)$ was defined in (3b), with the exponent b taken now from the GR relation for the aftershocks (see ref. 1). The summation is taken over the time interval δ after the main shock.

The PSP "bursts of aftershocks" depicted by the functional (7) with $b = 0$ is the first for which statistical significance has been rigorously established in observed seismicity.[30] In practice, aftershocks are counted within a short time interval of $\delta = 2$ days, while the whole aftershock sequence may be much longer, lasting up to a year or more. We have computed B_m with $\delta = 300$ for $m = 6$, 5, and 4.

Performance. The performance of the functionals B_m, $m = 4, 5, 6$, in predicting individual earthquakes is characterized by the error diagrams in Figs. 3e, f. The large rate of false alarms for $m = 5$ and 4 has a natural explanation: we concentrate only on the prediction of earthquakes with $m = 7$, while B_5 may be premonitory to earthquakes with $m = 6$ and B_4 to earthquakes with $m = 6$ and 5 as well.

3.3. Increase of Correlation Range

Long-Range Correlations in Observed Seismicity. The occurrence of major earthquakes is correlated at large distances that exceed considerably the size of the rupture in the earthquakes' sources. Different manifestations of this phenomenon include the large size of the areas where PSPs are formed;[7, 8, 34] the correlation of the strongest earthquakes worldwide between themselves, and with perturbations of the Chandler wobble and Earth rotation;[33, 46] migration of seismicity along active faults to distances up to 10^4 km;[47, 48] and near-simultaneous occurrence of earthquakes at large distances.[49–52] This phenomenon has also been reproduced in models of seismic dynamics.[53–55]

The physical mechanisms responsible for long-range correlations are summarized in ref. 7. Here we do not consider the large correlation range *per se*, but concentrate on its premonitory rise. This rise has been found in observed seismicity on the time scales of months,[26] years,[27, 28] and decades,[34] as well as in colliding-cascade models.[1–3]

Definitions. We consider several patterns that reflect premonitory increase of correlation range.

Pattern Accord[2, 3] reflects simultaneous rise of seismic activity in the distinct major branches of the system. Our ternary model is naturally divided into three major branches that descend from the second highest level, corresponding to $m = 6$. We measure seismic activity of a branch by the functional $\Sigma_6(t, s)$ defined by Eq. (3a) with $s = 3000$. The functional $A_6(t)$, which depicts the pattern *Accord*, is defined as the number of branches for which Σ_6 simultaneously exceeds a common threshold Σ_0.

By definition, the functional $A_6(t)$ may assume integer values from 0 to 3; we took the alarm threshold C_A for $A_6(t)$ to be $C_A = 2$. We considered also a functional $A_5(t)$ defined, similarly to $A_6(t)$, for the nine branches that descend from the third-highest level of the hierarchy, corresponding to $m = 5$; in this case we took $C_A = 3$. In the analysis of observations the "major branches" are the fault zones comprising the region.[27]

Pattern Π reflects the rise of activity in a sufficiently large part of the system. It is measured by the sum $\Pi_6(t)$ of the activities of the two (out of three) most active branches. As for the functional *Accord*, the activity of each major branch is measured by the function $\Sigma_6(t, 3000)$.

Pattern ROC ("Range of Correlation")[2, 3] captures the rise of the distances between nearly simultaneous earthquakes. The functional $R_m(t, s)$, which quantifies this pattern, is defined as the number of pairs of events with magnitude m which occur within time s from each other at the maximal possible distance. In the model we use the ultrametric distance along the tree,[56] i.e., the minimal number of edges that connect two elements. We considered this functional with $s = 5000$ for magnitudes m from 2 to 6. The functional that generalizes $R_m(t, s)$ to real observations has been explored by Shebalin *et al.*[26] in the quest for short-term prediction, since it could have a smaller lead time than other functionals considered here. On longer time scales a similar phenomenon was reported by Prozorov[49] for California and some other regions.

Premonitory emergence of the patterns *Accord* and *ROC* has been found first in the colliding-cascade model[2, 3] and then in observations.[26, 27] The premonitory rise of Π is found for the first time in this study. An alternative approach for detecting the rise of correlation range was suggested by Zöller *et al.*[28]

Error diagrams for prediction with the patterns *Accord*, *ROC* and Π are given in Figs. 3g–3j.

3.4. Change in the Size Distribution

According to refs. 24, 48, 57–59 a large earthquake with magnitude m_0 is preceded by an "upward bend" of earthquake size distribution, i.e., a decrease of the b-value for the G-R relation in the adjacent magnitude interval $(m_0 - c_2, m_0 - c_1)$, $c_1 < c_2$. Figure 5a clearly demonstrates such a change in the composite seismicity.

Definition. We measure this transformation by the functional

$$W(t) = \frac{\sum_{m'=4}^{6} N_{m'}(t, s)}{\sum_{m'=1}^{3} N_{m'}(t, s)}. \tag{8}$$

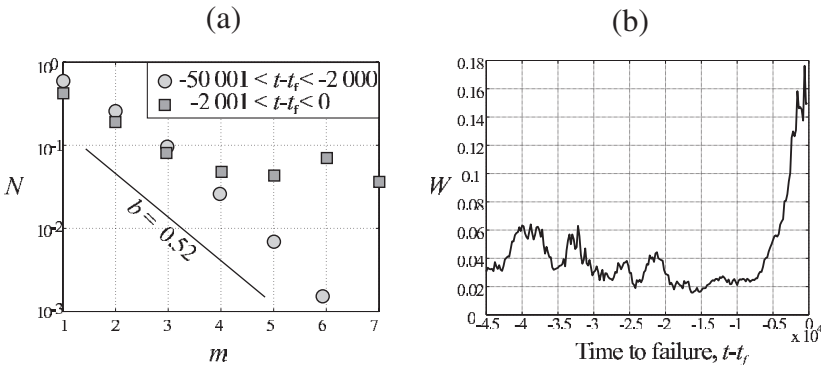

Fig. 5. Premonitory change in the earthquake size distribution. (a) Magnitude distribution of composite seismicity for two distinct time intervals prior to a large event that occurred at t_f. Circles: $-50001 < t - t_f < -2000$; squares: $-2001 < t - t_f < 0$. (b) The functional $W(t)$ that depicts the upward bend, cf. Eq. (8).

This functional was calculated with $s = 2000$ for the composite seismicity obtained by superimposing the 43 seismic cycles, each of which culminates in a maximal-magnitude earthquake (Fig. 5b). One can see a rapid growth of $W(t)$ that starts a few thousand time units before a major earthquake; the upward-bend pattern is thus clearly present in our model. However, the error diagram (panels (k) and (l) of Fig. 3) shows too many false alarms produced by this functional in predicting individual earthquakes. This illustrates a well-known difficulty in prediction research, where good statistical performance of a precursor does not necessary imply reliable prediction of individual events.

3.5. Error Diagrams for Individual Patterns

Figure 3 shows error diagrams for all the PSPs defined above. In previous sections we have defined four types of PSPs. Each pattern can be depicted by one or several functionals in different magnitude ranges; 23 functionals have been considered altogether. Different points in the error diagrams for a specific functional correspond to different thresholds for declaring an alarm; in the case of the functionals $A_m(t)$, different points correspond to different thresholds Σ_0.

The predictive skill of the functionals is better when they are defined at higher magnitudes; this is expressed by the fact that the corresponding points lie closer to the origin in the error diagram. Despite the large differences among the PSPs considered, the functionals defined for the same magnitude range give a similar performance in terms of the (n, τ) graph in

the error diagram (left panels in Fig. 3); the same observation was reported for observed seismicity in ref. 44.

As was mentioned above, the value of the threshold for declaring an alarm plays a major role in the performance of a given PSP: the higher the threshold, the shorter the relative duration τ of the alarms and the higher the rate n of failures to predict. But what happens to the fraction f of false alarms? As one can see from Fig. 3, the PSPs are divided into two groups according to the dependence of f on the alarm threshold. For the patterns "bursts of aftershocks," ROC, and "upward bend" these quantities are threshold-independent. At the same time, for the patterns Σ, N, Accord and Π, the rate of false alarms does depend on the alarm threshold: the higher the threshold, the lower the rate of false alarms. It is remarkable that both threshold-dependent and -independent functionals can be defined for the same PSP type.

4. COLLECTIVE PERFORMANCE OF PSPS

4.1. Track Record of Single PSPs

The alarms declared by each of the 23 functionals considered are juxtaposed in Fig. 6. Each box corresponds to a large earthquake; its sequential number is indicated at the top of the box. Our earthquake sequence includes 43 major earthquakes; for brevity we show only 10 of them. The right edge of a box is the time of a large earthquake. The horizontal axis shows the time before that earthquake. Each horizontal line corresponds to a premonitory pattern depicted by a specific functional in a certain magnitude range; both the functional and magnitude range are indicated at the left. Shaded areas are the alarms declared by the functional considered.

The trade-off between correct predictions and errors is controlled by a prediction algorithm's adjustable parameters and one tries to tune these parameters to optimize the algorithm. The three panels in Fig. 6 correspond to different choices of this trade-off. Each strategy minimizes one of the following characteristics of prediction: the rate of failures to predict n (panel (a)); the sum of errors $n+\tau+f$ (panel (b)); and the rate of false alarms f (panel (c)). One can easily see that the lead time of precursors decreases in that order. All the panels in Fig. 6 demonstrate the "magnitude ladder" phenomenon: patterns defined for lower magnitudes tend to emerge earlier.

4.2. Minimax Prediction Strategies

Figures 3 and 6 clearly show that prediction by individual PSPs produces many errors. Is it possible to improve the prediction for this model,

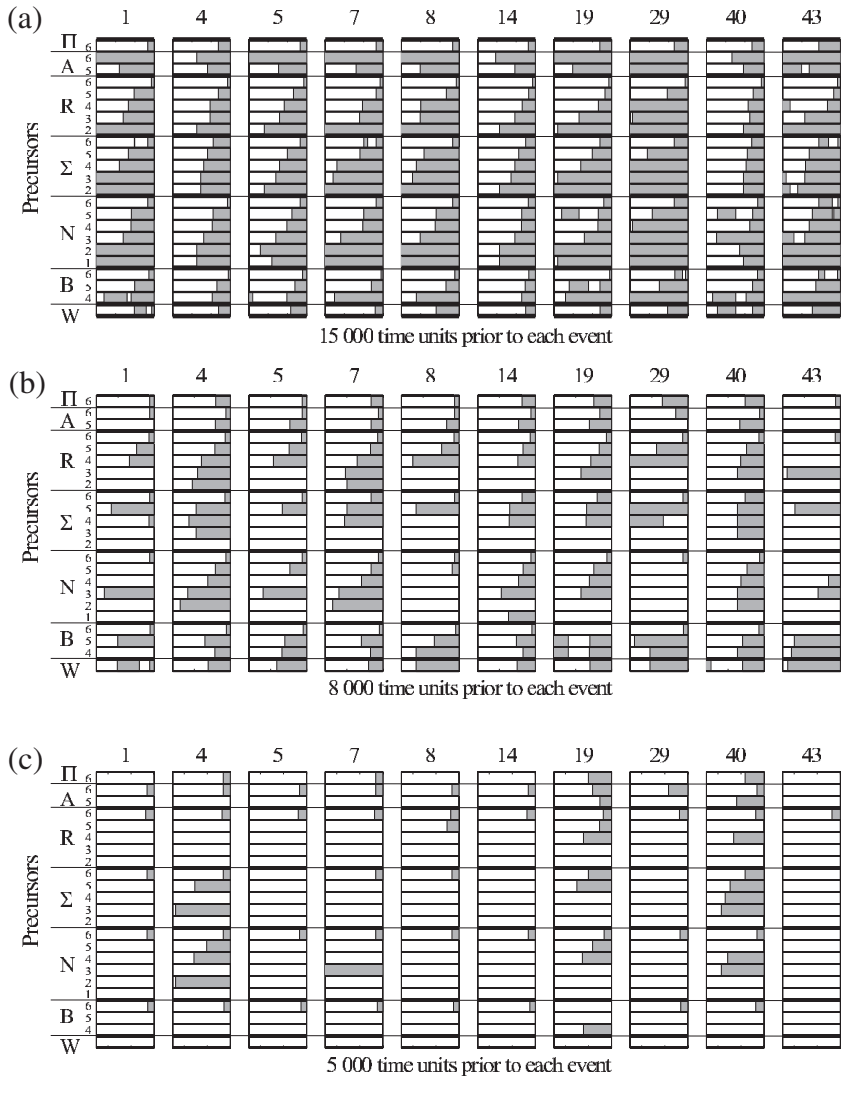

Fig. 6. Performance of single premonitory seismicity patterns (PSPs) prior to 10 major earthquakes. The figure juxtaposes the alarms generated by each of the 23 functionals considered. Each box shows the emergence of a pattern prior to a large earthquake; the sequential index of each event (out of a total of 43) is shown at the top of the box. The right edge of the box is the occurrence time of a large earthquake. Each horizontal row represents the track record of a particular functional, identified at the left side of the panel. Shaded areas show the time intervals when an alarm was declared by that functional. The three panels correspond to different prediction strategies: (a) minimization of failures to predict ($n \to \min$); (b) minimization of the sum of errors ($n+\tau+f \to \min$); and (c) minimization of the number of false alarms ($f \to \min$).

by a judicious combination of the functionals considered? To answer this question, we explore the following "minimax" strategy: (i) tune the prediction by an individual functional to minimize the rate of false alarms, at the cost of a large rate of failures to predict (Fig. 6c); and (ii) reduce the failures to predict by combining individual predictions.

We applied this minimax strategy as follows: Six functionals, all defined for $m = 6$, are used for prediction; they are Σ_6, N_6, B_6, R_6, A_6, and Π_6. Individual alarms are declared as described in Section 1.1. A collective alarm is declared when M out of the six possible alarms are triggered at the same time, $M = 1,\ldots, 6$. In that way we obtain six versions of our prediction, one for each value of M. The corresponding error diagrams are shown in Figs. 7a, b. Clearly, the number of failures to predict increases with M, while the durations of collective alarms decrease in that order.

In the same way we consider the five functionals defined for $m = 5$; they are Σ_5, N_5, B_5, R_5, and A_5. The collective alarm is declared when two individual alarms arise simultaneously. The corresponding error diagram is shown in Figs. 7c, d. Note that, in both cases, the minimax strategy allowed us to eliminate false alarms entirely.

Figure 7 shows that prediction can be substantially improved by combining single functionals: points corresponding to the minimax strategy lie closer to the origin than individual PSPs. The quantity $n+\tau$ is often used as a measure of prediction performance.[36] We show the line $n+\tau = const.$ in the error diagrams for comparison. The higher the angle between this line and the straight-line segment that connects the performance of an individual PSP with that of the minimax strategy, the higher the improvement in predictive skill.

Our analyses confirm that, while the onsets of different patterns are strongly correlated, the remaining discrepancy between the predictive performance of various functionals can still be exploited in order to improve overall prediction by combining them. Studies of PSPs in observed seismicity led to a similar conclusion.[44] Moreover, application of several patterns jointly has been reasonably successful in many prediction algorithms.[61–65] The more rigorous "minimax" strategy of combining PSPs, as tested here, has not yet been used in practical algorithms, and seems very promising.

5. DISCUSSION

Using a Boolean delay equation (BDE) model of colliding cascades we have reproduced in synthetic seismicity the four basic types of premonitory seismicity patterns (PSPs), which had been previously found in observations. The model allows one to explore the individual and collective performance of these patterns in different magnitude ranges, the correlation between patterns, and a prediction strategy based on a set of PSPs.

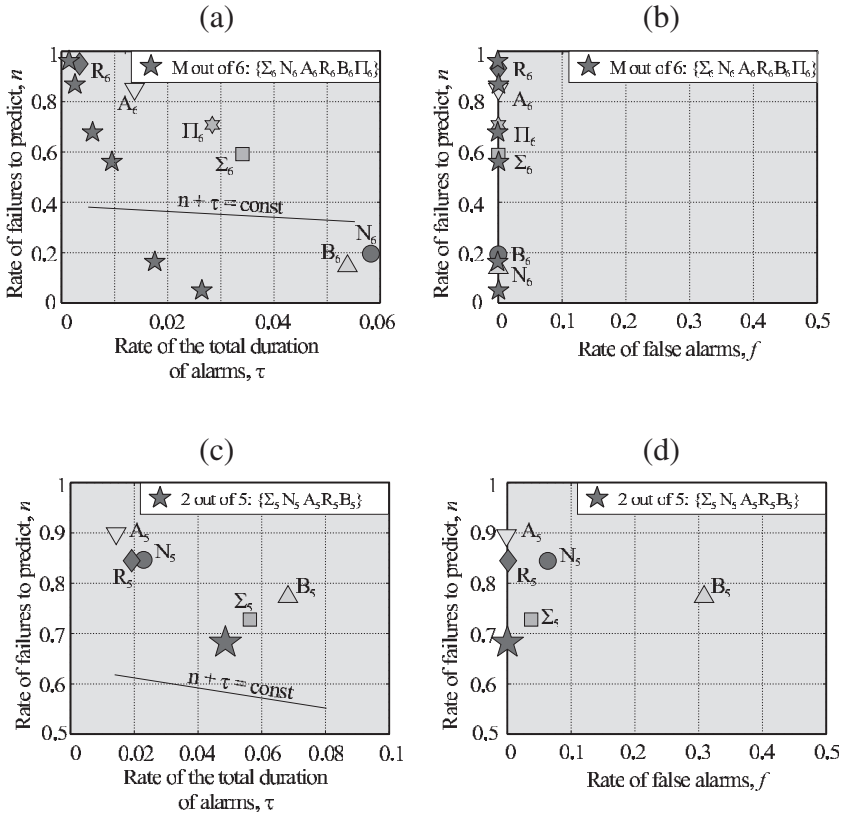

Fig. 7. Minimax strategy for combining individual PSPs. Error diagrams: (a), (b) Six functionals defined for $m = 6$ are considered: Σ_6, N_6, B_6, A_6, R_6, and Π_6; an alarm is triggered when at least M patterns emerge, $M = 1,..., 6$. As M increases, the total duration τ of alarms decreases and the ratio n of failures to predict increases. (c), (d) Five patterns defined for $m = 5$ are considered: Σ_5, N_5, B_5, A_5, and R_5; an alarm is triggered when at least two patterns emerge. In each case the single functionals are tuned to minimize the ratio of false alarms. The line $n+\tau = const$. is shown for comparison. See text for details.

The relevance of the model to the dynamics of seismicity is well supported by the fact that it fits the major heuristic constraints on earthquake sequences. It is remarkable that these constraints include heuristically defined PSPs.

We have found in the model several potentially important features of premonitory patterns, to be tested in observations.

1. *"Magnitude ladder."* The patterns defined for lower magnitudes emerge earlier. The individual performance of lower-magnitude functionals

is weaker, but they are important within the general scenario for the development of a major earthquake.

2. *Minimax strategy*. We have developed a prediction strategy using a minimax approach of combining individual PSPs: each pattern is tuned to minimize the rate of false alarms, at the cost of a large rate of failures to predict; then the failures to predict are reduced by the condition that at least few of them have emerged ("OR" operator).

An opposite minimax strategy can also be considered: each pattern is tuned to minimize the rate of failures to predict, at the cost of a large rate of false alarms; then the number of false alarms is reduced by requiring most of the patterns to have emerged ("AND" operator).

Such minimax strategies might be quite useful in practice, since combinations of several patterns are commonly involved in earthquake prediction algorithms.

The emergence of distinct patterns is strongly correlated in time, in the case of both correct and false alarms; moreover, the performance of different patterns defined in the same magnitude range is similar. This implies the existence of a common phenomenon that underlies the emergence of different PSPs. Empirical evidence for the existence of such a phenomenon is described in ref. 44. Furthermore, it seems that not only the premonitory phenomena but the model itself might be simplified even further.

While our model was tested against the observed dynamics of seismicity, the model transcends earthquake-specific processes and phenomena. The broader aspects it encompasses include regime switching, clustering (demonstrated in ref. 1), long-range correlations, and deviations from scale invariance. It seems worthwhile to explore its application to other hierarchical dissipative systems, which involve the loading → failure → healing sequence. Likely candidates are geotechnical constructions, volcanic regions, and socio-economic systems.

ACKNOWLEDGMENTS

This study was supported by the James S. McDonnell Foundation's 21st Century Collaborative Activity Award for Studying Complex Systems (I.Z. and V.K.-B.) and by an international supplement (I.Z.) to NSF Grant ATM-00-82131 (M.G.).

REFERENCES

1. I. Zaliapin, V. Keilis-Borok, and M. Ghil, A Boolean delay equation model of colliding cascades. Part I: Multiple seismic regimes, *J. Stat. Phys.* **111**:815–837 (2003).

2. A. Gabrielov, V. Keilis-Borok, I. Zaliapin, and W. I. Newman, Critical transitions in colliding cascades, *Phys. Rev. E* **62**:237–249 (2000).
3. A. M. Gabrielov, I. V. Zaliapin, W. I. Newman, and V. I. Keilis-Borok, Colliding cascades model for earthquake prediction, *Geophys. J. Int.* **143**:427–437 (2000).
4. D. Dee and M. Ghil, Boolean difference equations, I: Formulation and dynamic behavior, *SIAM J. Appl. Math.* **44**:111–126 (1984).
5. M. Ghil and A. P. Mullhaupt, Boolean delay equations. II: Periodic and aperiodic solutions, *J. Stat. Phys.* **41**:125–173 (1985).
6. V. I. Keilis-Borok and P. N. Shebalin, eds., Dynamics of lithosphere and earthquake prediction, *Phys. Earth Planet. Inter.* **111**, Special Issue, III:179–330 (1999).
7. V. Keilis-Borok, Earthquake prediction: State-of-the-art and emerging possibilities, *Annu. Rev. Earth Planet. Sci.* **30**:1–33 (2002).
8. S. C. Jaume and L. R. Sykes, Evolving towards a critical point: A review of accelerating seismic moment/energy release prior to large and great earthquakes, *Pure Appl. Geophys.* **155**:279–306 (1999).
9. A. J. Lichtman, *The Keys to the White House*, Lanham, ed. (Madison Books, 1996).
10. V. Keilis-Borok, J. H. Stock, A. Soloviev, and P. Mikhalev, Pre-recession pattern of six economic indicators in the USA, *J. Forecasting* **19**:65–80 (2000).
11. D. L. Turcotte, W. I. Newman, and A. Gabrielov, A statistical physics approach to earthquakes, in *Geocomplexity and the Physics of Earthquakes*, J. B. Rundle, D. L. Turcotte, and W. Klein., eds. (American Geophysical Union, Washington, DC, 2000), pp. 83–96.
12. V. I. Keilis-Borok, Symptoms of instability in a system of earthquake-prone faults, *Physica D* **77**:193–199 (1994).
13. J. Rundle, D. Turcotte, and W. Klein, eds., *Geocomplexity and the Physics of Earthquakes* (American Geophysical Union, Washington, DC, 2000).
14. C. J. Allegre, J. L. Le Mouel, and A. Provost, Scaling rules in rock fracture and possible implications for earthquake prediction, *Nature* **297**:47–49 (1982).
15. C. G. Sammis, D. Sornette, and H. Saleur, Complexity and earthquake forecasting, in *SFI Studies in the Science of Complexity*, v. XXV, J. Rundle, D. L. Turcotte, and W. Klein, eds. (Addison–Wesley, 1996).
16. T. Yamashita and L. Knopoff, Model for intermediate-term precursory clustering of earthquakes, *J. Geophys. Res.* **97**:19873–19879 (1992).
17. A. Gabrielov and W. I. Newman, Seismicity modeling and earthquake prediction: A review, in *Nonlinear Dynamics and Predictability of Geophysical Phenomena*, W. I. Newman, A. Gabrielov, and D. L. Turcotte, eds., Geophysical Monograph, Vol. 83; IUGG, Vol. 18 (1994).
18. A. I. Gorshkov, V. I. Keilis-Borok, I. M. Rotwain, A. A. Soloviev, and I. A. Vorobieva, On dynamics of seismicity simulated by the models of blocks-and-faults systems, *Annali di Geofisica* **XL 5**:1217–1232 (1997).
19. Y. Huang, H. Saleur, C. Sammis, and D. Sornette, Precursors, aftershocks, criticality and self-organized criticality, *Europhys. Lett.* **41**:43–48 (1998).
20. W. I. Newman, D. L. Turcotte, and A. Gabrielov, Log-periodic behavior of a hierarchical failure model with applications to precursory seismic activation, *Phys. Rev. E* **52**: 4827–4835 (1995).
21. G. F. Pepke, J. R. Carlson, and B. E. Shaw, Prediction of large events on a dynamical model of fault, *J. Geophys. Res.* **99**:6769–6788 (1994).
22. M. G. Shnirman and E. M. Blanter, Mixed hierarchical model of seismicity: Scaling and prediction, *Phys. Earth Planet. Inter.* **III**:295–304 (1999).
23. G. S. Narkunskaya and M. G. Shnirman, Hierarchical model of defect development and seismicity, *Phys. Earth. Planet. Inter.* **61**:29–35 (1990).

24. G. S. Narkunskaya and M. G. Shnirman, An algorithm of earthquake prediction, in *Computational Seismology and Geodynamics*, Vol. 1 (AGU, Washington, D.C., 1994), pp. 20–24.
25. V. G. Kossobokov and J. M. Carlson, Active zone size vs. activity: A study of different seismicity patterns in the context of the prediction algorithm M8, *J. Geophys. Res.* **100**: 6431–6441 (1995).
26. P. Shebalin, I. Zaliapin, and V. Keilis-Borok, Premonitory rise of the earthquakes' correlation range: Lesser Antilles, *Phys. Earth Planet. Int.* **122**:241–249 (2000).
27. I. Zaliapin, V. I. Keilis-Borok, and G. Axen, Premonitory spreading of seismicity over the fault network in Southern California: Precursor Accord, *J. Geophys. Res.* **107**:2221 (2002).
28. G. Zoeller, S. Hainzl, and J. Kurths, Observation of growing correlation length as an indicator for critical point behaviour prior to large earthquakes, *J. Geophys. Res.* **106**: 2167–2176 (2001).
29. V. G. Kossobokov, L. L. Romashkova, V. I. Keilis-Borok, and J. H. Healy, Testing earthquake prediction algorithms: Statistically significant advance prediction of the largest earthquakes in the Circum-Pacific, 1992–1997, *Phys. Earth Planet. Int.* **111**:187–196 (1999).
30. G. M. Molchan, O. E. Dmitrieva, I. M. Rotwain, and J. Dewey, Statistical analysis of the results of earthquake prediction, based on burst of aftershocks, *Phys. Earth Planet. Int.* **61**:128–139 (1990).
31. I. A. Vorobieva, Prediction of a subsequent large earthquake, *Phys. Earth Planet. Inter.* **111**:197–206 (1999).
32. I. M. Gelfand, Sh. A. Guberman, V. I. Keilis-Borok, L. Knopoff, F. Press, E. Ya. Ranzman, I. M. Rotwain, and A. M. Sadovsky, Pattern recognition applied to earthquake epicenters in California, *Phys. Earth Planet. Inter.* **11**:227–283 (1976).
33. F. Press and P. Briggs, Chandler wobble, earthquakes, rotation, and geomagnetic changes, *Nature (London)* **256**:270–273 (1975).
34. A. Press and C. Allen, Pattern of seismic release in the Southern California region, *J. Geophys. Res.* **100**:6421–6430 (1995).
35. J. W. Tukey, *Exploratory Data Analysis* (Addison–Wesley, Reading, Mass., 1977).
36. G. M. Molchan, Earthquake prediction as a decision-making problem, *Pure Appl. Geophys.* **149**:233–247 (1997).
37. L. V. Kantorovich and V. I. Keilis-Borok, Earthquake prediction and decision-making: Social, economic and civil protection aspects, in *International Conference on Earthquake Prediction: State-of-the-Art, Strasbourg, France* (Scientific-Technical Contributions, CSEM-EMSC, 1991), pp. 586–593.
38. W. I. Newman, A. Gabrielov, and D. L. Turcotte, eds., Nonlinear dynamics and predictability of geophysical phenomena, in *Geophys. Monographs Ser.* (AGU, Washington, DC, 1994), p. 83.
39. V. I. Keilis-Borok and L. N. Malinovskaya, One regularity in the occurrence of strong earthquakes, *J. Geophys. Res.* **69**:3019–3024 (1964).
40. L. Knopoff, T. Levshina, V. I. Keilis-Borok, and C. Mattoni, Increased long-range intermediate-magnitude earthquake activity prior to strong earthquakes in California, *J. Geophys. Res.* **101**:5779–5796 (1996).
41. C. G. Bufe and D. J. Varnes, Predictive modeling of the seismic cycle of the Greater San Francisco Bay region, *J. Geophys. Res.* **98**:9871–9883 (1993).
42. D. D. Bowman, G. Ouillon, C. G. Sammis, A. Sornette, and D. Sornette, An observational test of the critical earthquake concept, *J. Geophys. Res.* **103**:24359–24372 (1998).
43. P. Yiou, D. Sornette, and M. Ghil, Data-adaptive wavelets and multi-scale SSA, *Physica D* **142**:254–290 (2000).

44. V. I. Keilis-Borok, L. Knopoff, and I. M. Rotwain, Bursts of aftershocks, long-term precursors of strong earthquakes, *Nature (London)* **283**:258–263 (1980).
45. V. G. Kossobokov, B. K. Rastogi, and V. K. Gaur, On self-similarity of premonitory patterns in the regions of natural and induced seismicity, *P. Indian AS-Earth* **98**:309–318 (1989).
46. B. Romanowicz, Spatiotemporal patterns in the energy-release of great earthquakes, *Science* **260**:1923–1926 (1993).
47. E. V. Vilkovich and M. G. Shnirman, Epicenter migration waves: Examples and models, *Comput. Seism.* **14**:27–37 (1982).
48. K. Mogi, Seismicity in Western Japan and long-term forecasting, in *Earthquake Prediction: An International Review*, Maurice Ewing Series, Vol. 4 (American Geophysical Union, Washington, DC, 1981), pp. 43–51.
49. A. G. Prozorov and S. Yu. Schreider, Real time test of the long-range aftershock algorithm as a tool for mid-term earthquake prediction in Southern California, *Pure Appl. Geophys.* **133**:329–347 (1990).
50. J. Rice and J. Gu, Earthquake after effects and triggered seismic phenomena, *Pure Appl. Geophys.* **121**:187–219 (1983).
51. J. Gomberg, M. L. Blanpied, and N. M. Beeler, Transient triggering of near and distant earthquakes, *BSSA 87* **2**:294–309 (1997).
52. D. P. Hill, P. A. Reasenberg, A. Michael, W. J. Arabasz, G. Beroza, D. Brumbaugh, J. N. Brune, R. Castro, S. Davis, D. dePolo, W. L. Ellsworth, J. Gomberg, S. Harmsen, L. House, S. M. Jackson, M. Johnston, L. Jones, R. Keller, S. Malone, L. Munguia, S. Nava, J. C. Pechmann, A. Sanford, R. W. Simpson, R. S. Smith, M. Stark, M. Stickney, A. Vidal, S. Walter, V. Wong, and J. Zollweg, Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake, *Science* **260**:1617–1623 (1993).
53. C. D. Ferguson, W. Klein, and J. B. Rundle, Long range earthquake fault models, *Computers in Physics* **12**:34–40 (1998).
54. E. Preston, J. S. de la Martins, J. B. Rundle, M. Anghel, and W. Klein, Models of earthquake faults with long-range stress transfer, *Computing in Science and Engineering* **2**:34–41 (2000).
55. A. Soloviev and I. Vorobieva, Long-range interaction between synthetic earthquakes in the model of block structure dynamics, *Fifth Workshop on Non-Linear Dynamics and Earthquake Prediction, 4–22 October 1999, Trieste*, ICTP, H4.SMR/1150-4 (1999), 18 pp.
56. R. Rammal, G. Toulous, and M. A. Virasoro, Ultrametricity for physicists, *Rev. Mod. Phys.* **58**:765–788 (1986).
57. R. E. Haberman, Precursory seismicity pattern: Stalking the mature seismic gap, in *Earthquake Prediction: An International Review*, Maurice Ewing Series, Vol. 4 (American Geophysical Union, Washington, DC, 1981), pp. 29–42.
58. W. Smith, The *b*-value as an earthquake precursor, *Nature* **289**:136–139 (1981).
59. A. Prozorov and F. J. Sabina, Study of seismic properties in the Mexico region, *Geophys. J.R. Astr. Soc.* **76**:317–336 (1984).
60. V. I. Keilis-Borok, A worldwide test of three long-term premonitory seismicity patterns: A review, *Tectonophysics* **85**:47–60 (1982).
61. V. I. Keilis-Borok and V. G. Kossobokov, Premonitory activation of earthquake flow: Algorithm M8, *Phys. Earth Planet. Inter.* **61**:73–83 (1990).
62. V. G. Kossobokov, V. I. Keilis-Borok, and S. W. Smith, Localization of intermediate-term earthquake prediction, *J. Geophys. Res.* **95**:19763–19772 (1990).
63. V. I. Keilis-Borok 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).

64. I. Rotwain and O. Novikova, Performance of the earthquake prediction algorithm CN in 22 regions of the world, *Phys. Earth Planet. Inter.* **111**:207–214 (1999).
65. I. A. Vorobieva and T. A. Levshina, Prediction of a second large earthquake based on aftershock sequence, in *Computational Seismology and Geodynamics*, Vol. 2 (American Geophys. Union, Washington, DC, 1994), pp. 27–36.