united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

H4.SMR/1330-2

" Sixth Workshop on Non-Linear Dynamics and
Earthquake Prediction"

15 - 27 October 2001

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

I. Zaliapin, V. Keilis-Borok

International Institute of Earthquake Prediction Theory
and Mathematical Geophysics, Russian Academy of Sciences
Moscow, Russia
University of California at Los Angeles
Los Angeles, California, U.S.A.

M. Ghil
Department of Atmospheric Sciences and
Institute of Geophysics and Planetary Physics
University of California at Los Angeles
Los Angeles, California, U.S.A.

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

Ilya Zaliapin [*],

Vladimir Keilis-Borok[†], and

Michael Ghil[‡]

[*]International Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Science, Moscow, Russia, and Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California, E-mail: zal@ess.ucla.edu

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

[‡]Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California, 90095-1565. E-mail: mghil@igpp.ucla.edu, Phone: (310) 206-0651 Fax: (310) 206-5219, corresponding author.

Abstract

We consider a prominent and poorly understood feature of hierarchical nonlinear ("complex") systems: persistent recurrence of abrupt overall changes, called here "critical transitions." Unlike studying phase transitions in statistical physics, we consider large deviations from steady state that culminate in a critical transition, rather than the steady states themselves. This raises a general problem that comprises scenarios of development of critical transitions, their prediction and, possibly, control.

We consider prediction of critical transitions, using the model developed in the first part of this paper. The model merges the concept of *colliding cascades* and work of *Boolean delay equations.* It describes generation of critical transitions in hierarchical system by interaction between *direct cascades of loading* and *inverse cascades of failures.* Interaction is controlled by distinct delays between switching of elements from one state to another: loaded vs. unloaded and intact vs. failed.

The attention is focused on the earthquake prediction problem; accordingly, the model's heuristic constraints are taken from dynamics of seismicity. Many of our conclusions, though, may have wider applications: premonitory phenomena encountered in seismology seem to be common to hierarchical systems of different origin.

The model, its simple design notwithstanding, exhibits four major types of premonitory seismicity patterns (PSPs), which have been previously observed in earthquake prediction studies: (i) rise of earthquake clustering; (ii) rise of the earthquake flow's intensity; (iii) rise of the earthquakes' correlation range, and (iv)

certain transformation of Gutenberg-Richter relation (size distribution of earthquakes). We have found potentially important features of PSPs that are worth testing on observations. The individual and collective performance of PSPs in the modeled seismicity is evaluated by error diagrams. In real earthquake prediction such performance, particularly the collective one, would be regarded as quite good. Different prediction strategies are formulated and compared.

Keywords. Prediction of complex behavior; Cellular automata; Delay equations; Hierarchical modeling; Colliding cascades; Earthquake precursors.

1 Introduction

This is the second part of a study aimed at a prominent and yet insufficiently understood feature of hierarchical nonlinear, "complex" systems: persistent recurrence of abrupt overall changes, called here "critical transitions." The study's first part [1] described a hierarchical model for the development of critical transitions. The model's heuristic constraints were 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 that system. That model was applied to the study of multiple seismic regimes. In the study's Part II (the present paper), we focus on earthquake prediction problem.

We apply to the model previously developed earthquake prediction algorithms [2, 3], based on specific premonitory seismicity patterns (PSPs). PSPs are spatio-temporal

features of an earthquake sequence, which indicate the approach of a strong earthquake. The goal of the present study is twofold: (i) validation of the model, by showing that it fits the major heuristic constraints, PSPs in particular; and (ii) finding new premonitory phenomena that can be tested on observations; if successful, such a test would greatly increase the model's credibility.

This Introduction outlines the essentials of the model (Sects. 1.1 and 1.2), heuristic constraints on it (Sect. 1.3), PSPs used in this study, and the general scheme of prediction (Sect. 1.4). Sections 2 and 3 describe respectively the individual and the collective performance of PSPs in the modeled seismicity. A discussion of the results and their implications for real earthquake prediction follows in Sec. 4.

1.1 Colliding Cascades (CC) models

These recently introduced models [1, 4, 5, 6] synthesize three phenomena that play an important role in many complex systems: (i) The system has a hierarchical structure, with the smallest elements merging in turn to form larger and larger ones. (ii) The system is continuously loaded (or driven) by external sources. (iii) The elements of the system fail (break down) under the load, causing redistribution of the load and strength throughout the system. Eventually the failed elements heal, thereby ensuring the continuous operation of the system.

The load is applied at the top of the hierarchy and transferred downwards to the smallest elements, thereby forming a *direct cascade of loading.* Failures are initiated at

the lowest level of the hierarchy, and gradually propagate upwards, thereby forming an *inverse cascade of failures*, which is followed by healing. The interaction of direct and inverse cascades establishes the dynamics of the system: loading triggers the failures while failures redistribute and release the load.

1.2 Boolean delay equations (BDEs)

BDEs are a novel modeling language especially tailored for the mathematical formulation of conceptual models of systems that exhibit threshold behavior, multiple feedbacks and distinct time delays (see Fig. 2 in [1]). Originally inspired by theoretical biology, Ghil and associates [7, 8] recognized the potential of BDEs for modeling the multiple feedbacks between the components of the climate system [9]. They intended BDEs as a heuristic first step on the way to understand problems too complex to model using systems of partial differential equations at the present time. One hopes, of course, to be able to eventually write down and solve the exact equations that govern the most intricate phenomena. Still, in climate dynamics [10] as well as in solid-earth geophysics [11] and elsewhere in the natural sciences, much of the preliminary discourse is often conceptual.

We use the Boolean delay equation (BDE) framework to model the dynamics of colliding cascades. The phenomenon of colliding cascades comprises interaction of direct cascades of loading and inverse cascades of failures. Using the BDE framework, we replace the elementary interactions of elements in the system by their integral effect. This effect

is represented by the time delays between consecutive switching of an element's state, i. e. between unloaded vs. loaded and intact vs. failed states. In this way we bypass the necessity to reconstruct the detailed behavior of the system from complex and diverse interactions, for which our knowledge is very incomplete. This simplifies the systematic study of the model's dynamics without losing its essential features.

1.3 Heuristic constraints

In its applications to seismicity hierarchical structure of the model represents a faults network [12, 13], loading imitates the impact of tectonic forces, and failures imitate earthquakes. Heuristic constraints include the major regularities observed in the dynamics of real seismicity [14, 15, 16, 67]: (i) the seismic cycle; (ii) intermittency in the seismic regime; (iii) the size distribution of earthquakes, usually called the Gutenberg-Richter (G-R) relation; (iv) specific clustering of earthquakes in space and time; (v) long-range correlations in earthquake occurrence; and (vi) a variety of seismicity patterns premonitory to a strong earthquake [3, 11, 14, 67, 68].

The first part of this study [1] shows that the model reproduces the first four features, (i)-(iv). Here we will focus on the last one — the premonitory seismicity patterns (PSPs). The existence of long-range correlation will follow from our analysis of PSPs.

1.4 Premonitory seismicity patterns (PSPs) and prediction algorithms

Our point of departure is the methodology of earthquake prediction, realized in a family of intermediate-term prediction algorithms [2, 3, 11]. We briefly describe this methodology below.

1.4.1 Four types of premonitory seismicity patterns

Studies of observed and modeled seismicity have demonstrated that an earthquake of magnitude m is often preceded by PSPs formed within an area and magnitude range depending on m. Unambiguously defined and better validated are the intermediate-term patterns that have a characteristic duration of alarms years.

We consider here PSPs of the following four types [3, 18, 19, 20, 21]: (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 [22, 23, 24, 25, 26, 27]; patterns of other two types — first in models and then in observations [4, 28, 29, 30, 31, 32]. The intermediate-term patterns of the first two types have been validated by the statistically significant predictions of real earthquakes [34, 35, 36]. Other patterns are in different stages of testing.

1.4.2 General scheme of prediction

(i) A sequence of earthquakes in a given area is described by a catalog

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

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 main features of such a sequence are robustly captured by the functionals $F_k(t, s_k)$, $k = 1, 2, \ldots, K$, each depicting a certain PSP (Fig. 1). With a few exceptions, our functionals are defined in a sliding time window $(t - s, t)$; note that 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 in Section 2 below.

(ii) Emergence of a PSP is defined by the condition

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

Here $F_k(t)$ is one of the functionals that depict this PSP; the threshold C_k is usually chosen as a certain percentile of the probability density function (pdf) for functional F_k. Since a single PSP may be captured by several different functionals, when talking about emergence of a PSP we suppose that appropriate functional F_k that depict this PSP is chosen. When it is imporatnt to distinguish between different functionals that depict the same PSP we use phrases like "emergence of a functional", etc.

(iii) An alarm is triggered when a single pattern or a certain combination of the

patterns emerges. Different prediction algorithms use different patterns. An alarm is declared for a time interval τ_k. In the algorithms considered here the alarm is terminated after a major earthquake occurs or the time τ_k expires, whichever comes first. The possible outcomes of prediction are shown in Fig. 2. Obviously this scheme is open for the use of other data, not necessarily seismological ones (e.g. [41, 42, 43, 44, 45]).

1.4.3 The pattern recognition approach to prediction.

In the absence of an adequate theory, the prediction algorithms considered here have inevitably been found by a pattern recognition analysis of observed and/or modeled seismicity.

The relevance of pattern recognition is clear from Fig. 1: we have to find a set of functionals $F_k(t)$ that discriminate between the two possibilities: a strong earthquake will or will not occur in the area considered during a specified time interval. The methodology of pattern recognition of infrequent events has been developed by the school of I. M. Gelfand [13, 37, 38, 39] 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 [40].

A distinctive feature of this methodology is robustness. To properly deal with the high complexity of seismic phenomena, the functionals and the corresponding prediction algorithms have to be given robust definitions that makes them applicable to different conditions, such as region, magnitude range, and seismic regime; otherwise the test of a

prediction algorithm would be practically impossible. Such applicability is achieved at a price, namely the reduction of the predictions' accuracy in individual cases.

This is a particular instance of the usual trade-off between the number of parameters fitted and result's statistical significance. E. Fermi coined the phrase *"With four exponents I can fit an elephant"* to highlight the problem of using too large a set of parameters for a given data set. The pattern recognition approach leaves a considerable freedom in the retrospective development of an algorithm, in particular in the choice of adjustable parameters. A good prediction algorithm has to be insensitive to various free choices of its definition. The sensitivity analysis for such an algorithm comprises an exhaustive set of numerical experiments, that make up a major part of the efforts in the algorithm development. This set includes: (i) evaluating the stability of predictions to variations in the algorithm and its parameters, and (ii) validating the results by application of the algorithm to "out of sample" data. Finally, real-data advance prediction provides further confidence in the algorithm. The results of such analyses are summed up in the error diagrams that show the trade-off between different prediction errors [46, 47, 48, 49]. They show the actual prediction skill and provide the necessary basis for the optimization of disaster preparedness measures that may be undertaken in response to such a prediction [48, 49, 50].

Error diagrams for the PSPs considered in this study are given in Section 2.

1.4.4 Four paradigms

The studies described above led to the following findings that are important for fundamental understanding of the dynamics of seismicity, as well as for further prediction research [14, 51, 52]: (i) Long-range correlations occur in fault system dynamics and, accordingly, premonitory phenomena are formed in the large-size areas. (ii) There exist four types of premonitory phenomena (see Sect. 1.4.1). (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 for a wide class of nonlinear systems; others are specific to the geometry of the faults' network or to a certain physical mechanism that controlls the stress and strength field in the lithosphere.

Table 1 provides a more detailed comparison between the "holistic" approach to earthquake prediction being pursued here and the complimentary "reductionist" approach. The table also provides additional information on characteristic sizes of the fault networks and time scales of interest for long-, intermediate- and short-term prediction.

2 Individual performance of each seismicity pattern

In this section, we study the PSPs that appear in the synthetic seismicity produced by the model introduced in [1]. These are PSPs of the four types discussed above in Sect. 1.4.1, and we thus apply the predictions based on those PSPs, one at a time. The brief overview of the patterns is given in Table 2. Each PSP considered is depicted by a

set of functionals; as a result 23 different functionals are considered. We evaluate the performance of each functional by the corresponding error diagram (see Sect. 2.3 and Fig. 8) and juxtapose the emergence of different functionals before individual major earthquakes (see Sect. 2.9, Figs. 9 – 11)

2.1 Earthquake sequence

In this study we analyze the synthetic earthquake sequence shown in Fig. 3. It corresponds to the model's intermittent regime, denoted by **I** in [1]. The exact model parameter values that yield this sequence are given in the captions of the figure. As demonstrated in [1], this sequence exhibits major features of real seismicity, as summarized in Sect. 1.3. This sequence includes 43 earthquakes with $m = 7$, maximal possible in this version of the model. They are targeted for prediction in the subsequent analysis.

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 main shock are taken here as follows: $T(7) = 3000, T(6) = 1500, T(5) = 600, T(4) = 300, T(3) = 150$, and $T(2) = 70$.

Most of the premonitory patterns considered have been defined on the sequence of main shocks. However the number of aftershocks is retained for each main shock. As

in analysing observations, the few immediate foreshocks are not differentiated from the main shocks.

The synthetic sequence of main shocks is represented by a catalog D similar to C of Eq. 1:

$$D = \{(t_l, m_l, h_l, B_l) : \ l = 1, 2, \ldots, L, \ t_l \leq t_{l+1}\}. \tag{3}$$

Here, B_l is the number of aftershocks of the $l - th$ main shock within a time interval $(t_l, t_l + \delta)$; where δ is usually very small. We use $\delta = 300$ time units.

Prediction is targeted at the "major" earthquakes that have magnitude $m = 7$, the maximal possible in the sequence considered. Premonitory patterns are formed by the earthquakes with smaller magnitudes, from $m = 6$ down to $m = 1$.

2.2 Scale invariance.

The energy distribution of earthquakes, commonly known as G-R relation, is a fundamental feature of dynamics of seismicity [16, 87, 88, 89, 90, 91]. Statistical-physics-type models of seismicity often use the simplest power-law form of this relation,

$$\log_{10} N(m) = a - bm. \tag{4}$$

Here $N(m)$ is the average annual number of earthquakes with magnitude m within a certain territory. Since m is the logarithmic measure of energy, Eq. (4) is the power law.

This power law naturally evokes far-reaching analogies with scale invariance and self-organized criticality. However, Eq. (4) is merely a zeroth-order approximation of seismicity in a large-scale fault systems, but not on a local scale, e.g. on a single active

fault, commensurate with the size of earthquake sources. Equation (4) is essentially a multiscale description of seismicity, since the larger m the larger the area, where the value of $N(m)$ has to be estimated. The power law of Eq. (4) is valid only after substantial averaging over time and territory; even so, it is valid only in a limited magnitude range. For example, B. Gutenberg and C. Richter themselves emphasized the downward bend of $N(m)$ at large m for the whole global seismicity. As demonstrated by Molchan and Podgaetskaya [89], the global distribution of earthquake sizes is better fitted by Kolmogoroff's log-normal distribution [92]. A good heuristic constraint for modeling seismicity is rather the deviation from the power law (4) changing in time and from area to area. For instance, a more adequate mesoscale approximation of the G-R relation is a broken line in log-log coordinates with constant values of b within certain magnitude intervals [33].

G-R relation for our synthetic sequence is shown in Fig. 4 for orientation in the subsequent analysis. Power law of Eq. (4) fits well the model data. The b-value is less for main shocks than for aftershocks that is in accordance with observations on real earthquakes. Premonitory transformations of GR relation are discussed in Sect. 2.8 below.

2.3 Error diagrams

Definitions. Error diagrams are a key element in evaluating a prediction algorithm. They have been introduced into earthquake prediction studies by G. Molchan [46, 47, 48, 49]. The definition of an error diagram is the following: Consider prediction by a single

functional. In the present context, a precursor is a functional on the seismic sequence D of Eq. (3). We continously monitor seismicity, declare alarms when that functional exceeds a threshold value as described in Sect. 1.4.2, and count the prediction outcomes (Fig. 2).

Prediction relates to a specified 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 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 step in time the alarm is declared with some probability τ and not declared with probability $1 - \tau$. Given a particular prediction that depends on a set of parameters, different points on error diagram correspond to different values of these parameters.

Error diagrams thus sum up the score of prediction's successes and errors. This score depends on an algorithm's adjustable parameters. For example, raising the threshold C_k in Eq. (2) will reduce the number of alarms A but may increase the number of failures to predict N_F. Raising τ_k, on the other hand, will increase the duration of alarms D but may reduce the number of failures to predict N_F, etc. A prediction algorithm is useful if: (i) prediction score is higher than the random one that corresponds to the diagonal $n + \tau = 1$ of the square $\{(n, \tau) : \ 0 \leq n \leq 1, 0 \leq \tau \leq 1\}$; and (ii) this score is insensitive to variation of parameters.

2.4 Rise of seismic activity: the functionals Σ and N

Definitions. A premonitory rise of seismic activity is depicted by the following functionals [2, 14, 61, 62]: $N_m(t,s)$ is the number of main shocks of magnitude m within the time interval $(t-s,t)$ and

$$\Sigma_m(t,s) = \sum_{m'=1}^{m} N(m')S(m'). \tag{5}$$

Here $N(m')$ is the number of main shocks with magnitude m', while the weight $S(m')$ is proportional to the source area of an earthquake with magnitude m'. In analysis of observations it is coarsely estimated from the earthquake magnitude: $S(m) \sim 10^{bm}$. Here we use the same expression with $b = \log_{10}(3)$, suggested by the ternary hierarchical structure of our system.

Performance. We computed each of these functionals for the sequence shown in Fig. 3 with $s = 3000$ and for m from 1 to 6; note that according to (5) the patterns N_1 and Σ_1 coincide. Altogether we consider 11 functionals that depict rise of seimic activity. Error diagrams for prediction with these functionals are given in Figs. 8a–d below. "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. This is done for all the error diagrams shown in Fig. 8. For advance prediction of real earthquakes the performance shown in Fig. 8 would be quite satisfactory.

2.5 Rise of seismic activity: accelerated Benioff strain release

Definition. Aside from the functionals N_m and Σ_m discussed in Sect. 2.4, the cumulative Benioff stress release also captures rise of seismic activity [21, 26, 67, 68, 69, 70, 71, 72]. It is dfined as

$$\epsilon(t) = \sum_k E_k^{1/2}, \; t_0 \leq t_k \leq t. \tag{6}$$

Here E_k is the energy and t_k is the moment of the k-th earthquake within an analyzed sequence. In the analysis of observations energy is estimated from the magnitude. The summation is taken over all the earthquakes in an area under consideration, without elimination of aftershocks. Analysing the model we define the function $\epsilon(t)$ as

$$\epsilon(t) = \sum_k 10^{bm_k}, \; t_0 \leq t_k \leq t. \tag{7}$$

The function $\epsilon(t)$ determined from observations is approximated by a power law [67, 68]

$$\epsilon(t) = A - B(t_f - t)^\alpha, \tag{8}$$

where t_f is the time of a strong earthquake, targeted for prediction. This power law is often allowed to be modulated by log-periodic variations:

$$\epsilon(t) = A - B(t_f - t)^\alpha \left\{1 + C\cos\left[w\ln(t_f - t)\right]\right\}. \tag{9}$$

Since t_f is a priori unknown, a number of different time series analysis methods can be used to optimize its determination from the observations [73].

Composite seismicity. Figure 5 shows that the modeled seismicity indeed exhibits on average the premonitory power law rise described by Eq. (8). Aside from the power-law

behavior of the function ϵ captured well by Figs. 5a,b, Fig. 5c hints strongly at its periodic modulations amplifying as the major earthuqake approaches.

In this study, we do not explore this pattern for individual major earthquakes.

2.6 Rise of clustering: "bursts of aftershocks"

Definitions. Premonitory clustering is captured by "pattern B" or "burst of aftershocks" [53]. It consists of a main shock with a large number of aftershocks. One of the functionals depicting this premonitory phenomenon is $B_m(t_l, \delta)$. Here m and t_l are the magnitude and occurrence time of the $l-th$ main shock, while B_m is the number of its aftershocks within the time interval δ. The PSP "bursts of aftershocks" depicted by the functional B_m was the first to have its statistical significance carefully established [35]. In real earthquake prediction B_m is counted within a short time interval of $\delta = 2$ days, while the whole aftershock sequence may be much longer — a year or more. The updated application of this pattern in southern California was recently described by Rotwain and Liu [59].

We consider also a generalized functional, which employs the function Σ_m of Eq. (5) to weigh aftershocks according to their magnitude:

$$B_m^{\Sigma}(t_l, \delta) = \sum_{m'=1}^{m} 10^{bm'} \cdot N_l(m'). \tag{10}$$

Here $N_l(m')$ is the number of those aftershocks of the $l-th$ main shock that have magnitude m'; clearly, $m' \leq m$ and $\sum_{m'=1}^{m} N_l(m') = B_m$. The exponent b is taken from the G-R relation for the aftershocks. As in the definition of B_m, the summation is taken

over the time interval δ after the main shock. V. Kossobokov used the functional B_m^Σ in [60] for analysing observed seismicity.

The weighting introduced in (10) equalizes the contribution of the aftershocks with different magnitudes. Without the weighting, $b = 0$, and thus B^Σ and B coincide. We computed both B_m^Σ and B_m with $\delta = 300$ for $m = 6$, 5, and 4, thus considering 6 different functionals that depict the PSP "bursts of aftershocks".

Composite seismicity. Emergence of this pattern in the composite sequence of Fig. 3 is summed up in Fig. 6. The figure shows a composition of time intervals $(t_f - 10^4,\ t_f)$ before each major earthquake, where t_f is its occurrence time. The following features in Fig. 6 are noteworthy:

- High values of B_6 and B_6^Σ — that is bursts of aftershocks triggered by main shocks with $m = 6$ — abruptly emerge at about 700 time units prior to a major earthquake.

  Burst of aftershocks triggered by main shocks with $m = 5$ have a larger lead time, of 3,000 to 2,000 time units, and are more widely spread. This happens for an obvious reason: due to similarity of our hierarchical model at different scales they are premonitory not only to earthquakes of $m = 7$ but also to smaller ones, $m = 6$.

- B_m^Σ yields a better performance than B_m: it pays off to weight the aftershocks by magnitude.

- There is an apparent drop of B_m^Σ and B_m close to a major event. This effect was recently observed for real earthquakes in Southern California [59].

Performance. Performance of the functionals B_m^{Σ}, $m = 4, 5, 6$ in prediction of individual earthquakes is characterized by the error diagram in Fig. 8, panels e and f. The large rate of false alarms for $m = 5$ and 4 has the same explanation, as their wider spread in the composite sequence (Fig. 6): the first pattern may be premonitory to earthquakes with $m = 6$ and second — with $m = 6$ and 5, while we concentrate only on predition of earthquakes with $m = 7$, the strongest ones. Emergence of these patterns before individual major earthquakes is shown in Figs. 9 – 11 below, along with the false alarms.

2.7 Rise of correlation range: functionals Accord, Π, and ROC

Long-range correlations in 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 premonitory seismicity patterns are formed [2, 39, 61]; the correlation of the strongest earthquakes worldwide between themselves, and with perturbations of the Chandler wobble and Earth rotation [38, 86]; migration of seismicity along active faults to distances up to 10^4 km [84, 94]; and near-simultaneous occurrence of earthquakes at large distances [75, 76, 77, 81, 82, 83] This phenomenon has been also established by modeling seismic dynamics [78, 79, 80].

The physical mechanisms that are responsible for long-range correlations are summarized in [14]. Here we do not consider the large correlation range *per se*, but concentrate on its premonitory rise. This has been found to occur in observed seismicity, on the time

scales of months [30], years [31, 32], and decades [39], as well as in colliding cascades models [1, 4, 5]. No statistical significance for this phenomenon is yet established in observations. Further study in the present model is intended to provide pointers to the analysis of real seismicity.

Definitions. We consider several functionals that register premonitory rise of correlation.

The functional A, *"Accord"*. This functional measures simultaneous rise of seismic activity in the major branches of the system. Our ternary model is naturally divided into three major branches, descending 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. (5) with $s = 3000$. The functional $A_6(t)$ is defined as the number of branches for which $\Sigma_6(t, s)$ simultaneously exceed a common threshold C_A. By definition, the functional $A_6(t)$ may assume integer values from 0 to 3; we took $C_A = 2$.

In the analysis of observations the "major branches" are the fault zones comprising the region [12, 31].

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 $C_A = 3$.

The functional Π. This functional measures the rise of activity in a sufficiently large part of the system. As for the functional *Accord*, we measure activity of each major branch by the function $\Sigma_6(t, s)$ with $s = 3000$. The functional $\Pi_6(t)$ is the sum of the

activities of the two most active branches.

The functional *ROC ("range of correlation").* This functional captures the rise of the distance between nearly simultaneous earthquakes. The functional $R_m(t, s)$ 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 [74], i. e., the minimal number of edges that connect two elements; according to the definition of ultrametric distance, siblings are at a distance null from each other. 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 P. Shebalin et al. [30] 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 A. Prozorov [75, 76, 77] for California and some other regions.

A premonitory rise of the functionals *Accord* and *ROC* has been found first in the colliding cascades model [4, 6] and then in observations. So far this has been found only in one region for each functional, in S. California for *Accord* [31], and the Lesser Antilles for *ROC* [30]. The premonitory rise of Π has been found first in this study. An alternative approach for detecting the rise of correlation range was suggested by Zoeller et. al. [32].

Performance. Error diagrams for prediction with the patterns *Accord*, *ROC* and Π are given in Fig. 8, panels g-j. Performance of *ROC* is weaker, comparing to the other patterns. Possibly, being a short-term precursor, it works better as a second approximation to intermediate-term one; this was observed by Shebalin et al. [30] in Lesser

Antilles.

2.8 Transformation of G-R relation: "Upward Bend" pattern

Definition. According to [29, 93, 94, 95, 96] a major earthquake with magnitude m_0 is preceded by an "upward bend", 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 7 clearly demonstrates such a change in the composite seismicity. Panels (a) and (c) show the G-R relation for the main shocks and all events, respectively. The change is particularly strong when considering only the main shocks. To depict this transformation we use the functional

$$U(t) = \frac{\sum_{m'=4}^{6} N_{m'}(t,s)}{\sum_{m'=1}^{3} N_{m'}(t,s)}. \tag{11}$$

This functional was calculated with $s = 2000$ for the composite seismicity and is shown in Figs. 7b and d. One can see a rapid growth of the $U(t)$ that starts a few thousand time units before a major earthquake. This provide a clear evidence of the presence of the upward-bend pattern in our model. However, the error diagram (panels k and l of Fig. 8) shows too many false alarms produced by this pattern in predicting individual earthquakes. This is discussed in a few more detail in Sect. 4.

2.9 Track record of single patterns

Alarms. Alarms raised by each functional are juxtaposed in Fig. 9. Each box corresponds to a major 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 major earthquake. The horizontal axis shows the time before that earthquake. Each line corresponds to a premonitory pattern depicted by a specific functional in a certain magnitude range; both functional and magnitude range are indicated at the left. Shaded areas are the alarms rised by this pattern.

We juxtapose in Fig. 9 the predictions by all the 23 functionals considered in this study.

Prediction strategies. The trade-off between correct predictions and errors is controlled by a prediction algorithm's adjustable parameters. Three panels in Fig. 9 correspond to different choices of the trade-off that implement the three "prediction strategies". Each strategy minimizes one of the following characteristics of prediction: the rate of failures to predict n (top panel); the sum of errors $n + \tau + f$ (middle panel); and the rate of false alarms f (lower panel). One can easily see that the lead time of precursors decreases in that order.

Figure 10 illustrates the continuous performance of each pattern, including long intervals where major earthquakes did not occur, so that most of the alarms are false. Earthquakes of $m = 6$ are indicated in the figure by shadowed vertical lines to emphasize their association with the false alarms.

Success score. The alarms preceding each of the 43 major earthquakes are juxtaposed in Fig. 11. It presents the complete summary of correct predictions and failures to predict, for each of the 23 functionals considered.

Next, we explore collective performance of these patterns.

3 Collective performance of seismicity patterns

Minimax strategy. Figures 8–11 clearly show that our prediction does have many errors even in a simple model. Is it possible to improve the prediction for this model, and with 23 functionals considered? To answer this question, we explore the following "minimax" strategy: (i) tune the prediction by individual functional at minimizing the rate of false alarms, at the cost of large rate of failures to predict (Figs. 9a, 10a), and (ii) reduce the failures to predict by combining different predictions.

Error diagrams. We applied the minimax strategy as follows: Six functionals, defined for $m = 6$, are used for prediction: $\Sigma_6, N_6, B_6, R_6, A_6$, and Π_6. Individual alarms are declared as described in 1.4.2. 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 predictions, corresponding to different values of M. The corresponding error diagrams are shown in Figs. 12a,b.

In the same way we consider the five functionals defined for $m = 5$: 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. 12c,d.

Figures 9–11 provide a clear evidence that emergences of different functionals and patterns are strongly correlated, while Fig. 12 shows, that prediction indeed improved by combining single functionals. This confirms that while the onsets of different patterns are strongly correlated, the remaining discrepancy is still sufficient to improve prediction by combining them. Note, that studies of PSPs in observed seismicity led to the similar

conclusion: *"Emergence of observed patterns is strongly correlated, suggesting that they all reflect a single underlying phenomenon"* [101]. Moreover, joint application of different patterns is reasonably successful in many prediction algorithms [102, 103, 104, 105, 106]. The "minimax" strategy shown in Fig. 12 was not yet used in these algorithms, and seems very promising, even though such a spectacular success could hardly be repeated with observations.

4 Discussion

1. *Relevance.* The relevance of the model to reality is well supported by the fact, that it fits the major heuristic constraints, including a wide range of premonitory seismicity patterns.

2. *Premonitory patterns.* We have found in the model potentially important features of premonitory patterns, to be tested in observations:

- *"Magnitude ladder".* Patterns defined for lower magnitudes tend to emerge earlier. For instance the functional B_5 rises prior to the rise of the functional B_6, as shown in Fig. 6. This is also true for other functionals as can be seen from Figs. 9 and 10. The individual performance of pattrens that are defined for smaller magnitudes is weaker, but they may be used in consecutive approximations to provide early warnings. *Further problem* is to organize these patterns into scenario of development of a major earthquake.

- *False alarms* are clustered in time. Most of them are associated with earthquakes of magnitude 6. *Further problem* is to discriminate such alarms, thus increasing the accuracy of predictions.

- *Collective performance* is much better than individual one, despite of strong correlation of the patterns. *Further problem* is to find optimal combination of patterns. This might be done by analysis of observations, taking advantage of existing prediction algorithms.

- *Lead time* of the patterns is noticeable different before different earthquakes. That is also the case in prediction of real earthquakes when the lead time varies from months to years. *Further problem* is to evaluate the lead time in advance.

3. *Yet unexplored possibilities.*

- *Other premonitory seismicity patterns.* We have focused here on the relatively better validated PSPs and did not explore several potentially important others:

  (i) Other forms of premonitory clustering — swarms of the main shocks of about the same magnitude [54, 55, 56] and narrow clusters of the foreshocks [57, 58].

  (ii) More sophisticated diagnostics of transformations of G-R relation, like those suggested in [29, 89, 98, 99, 100].

  (iii) Premonitory quiescence in different time scales [64, 65, 66].

  (iv) Log-periodic variations of seismic activity.

- *Alternative strategies.* The strategy of prediction opposite to the one considered in Sect. 3 deserves attention: to tune the patterns at minimizing the rate of failures to predict at the cost of the large rate of false alarms; then reduce the false alarms by combining different patterns.

- *Prediction within other seismic regimes.* As demonstrated in this study's first part [1] our model reproduces three types of seismic regime that exhibit (i) periodic, (ii) intermittent, and (iii) low-level seismic activity. Here we considered prediction only in the intermediate regime. The periodic regime is hardly realistic and prediction there is trivial. The low-seismicity regime is indeed realistic. However, prediction in that regime is quite different problem: this regime has neither clear critical phenomena nor scale invariance (see Fig. 4 of [1]).

4. *Key problems.* Our results open a new access to the following two problems, pivotal in the quest for better prediction methodology:

- *Control parameters.* Emergence of the patterns is strongly correlated in time. This emphasizes the problem to express them through the limited number of control parameters, lowering dimensionality of the system.

- *A simpler model?* While introduction of colliding cascades and the BD approach happened to be very helpful, some of the results could be possibly obtained with a simpler model that would underly the rich variety of the previously explored ones.

5. *Other applications.* While the model considered was tested against observed dy-

namics of seismicity, nothing in the model is particularly earthquake-specific. It seems worth to explore application of the model to other geological and geotechnical disasters, involving hierarchical structure, and the loading $\rightarrow$ failures $\rightarrow$ healing sequense.

Acknowledgements

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 to NSF grant ATM-00-82131 (I. Z. and M. G.).

Table 1. Two complementary approaches to earthquake prediction.

"REDUCTIONISM" **(from details to the whole)**	**"HOLISM"** **(from the whole to details)**
Premonitory phenomena preceding an earthquake with linear source dimension L are formed	
near the incipient source	**in a network of faults**
	of the linear size / **in time scale**
	$10^2 L$ / tens of years
	$10L$ / years
	L / years to months
	possibly, fractions of L, i.e. in a vicinity of the hypocenter / month to days
Premonitory phenomena	
are specific to mechanisms controlling the strength, e.g. friction, rock-fluids interaction, stress corrosion, buckling, etc.	**are divided into:** • "universal" ones common for many chaotic systems • those depending on the geometry of the fault network • mechanism-specific ones
Premonitory phenomena in different regions and energy ranges	
are different	**are to a considerable extent similar**

Table 2. Premonitory seismicity patterns (PSPs) considered in this study.

Notation	Description	Type	Definition in the text	Use in analysis of observations	References
N_m	Number of events.	Intensity	Sect. 2.4	YES	[62, 102, 103]
Σ_m	Weighted number of events. Coarsely estimates the area of faultbreaks.	Intensity	Eq. (5)	YES	[61, 102, 103]
B_m	Number of immediate aftershocks.	Clustering	Sect. 2.6	YES	[53, 35, 102, 103, 59]
B_m^{Σ}	Weighted number of immediate aftershocks.	Clustering	Eq. (10)	YES	[60]
R_m	Near-simultaneous occurrence of distant events.	Range of correlation	Sect. 2.7	Retrospectively	[30]
A_m	Simultaneus activation of distinct branches of a system.	Range of correlation	Sect. 2.7	Retrospectively	[31]
Π_m	Total activity of most active branches of a system.	Range of correlation	Sect. 2.7	NO	This study
U	Ratio of N_m for different m.	Transformation of GR relation.	Eq. (11)	YES	[102]

References

[1] Zaliapin, I., Keilis-Borok, V. and Ghil, M., 2001. A Boolean delay equation model of colliding cascades. Part I: Multiple seismicity regimes, J. Stat. Phys., submitted.

[2] Keilis-Borok, V. I. (Ed.), 1990. Intermediate-term earthquake prediction: Models, phenomenology, worldwide tests. Phys. Earth Planet. Inter. 61, Special Issue, 1,2, 144.

[3] Keilis-Borok, V. I. and Shebalin, P. N. (Eds.), 1999. Dynamics of Lithosphere and Earthquake Prediction. PEPI 111, Special Issue, III, 179-330.

[4] Gabrielov, A., Keilis-Borok, V., Zaliapin, I., and Newman, W.I., 2000. Critical transitions in colliding cascades, Phys. Rev. E, 62, 237-249.

[5] Gabrielov, A.M., Keilis-Borok, V.I., Zaliapin, I.V., and Newman, W.I., 2000. Colliding cascades model for earthquake prediction. JGI, 143, 427-437.

[6] Gabrielov, A., Zaliapin, I., and Keilis-Borok, V., 2001. Premonitory seismicity patterns in a cellular automaton with colliding cascades. Manuscript.

[7] Dee, D., and Ghil, M., 1984. Boolean difference equations, I: Formulation and dynamic behavior. SIAM J. Appl. Math. 44, 111-126.

[8] Ghil, M., and Mullhaupt, A. P., 1985. Boolean delay equations. II: Periodic and aperiodic solutions. J. Stat. Phys. 41, 125-173.

[9] Ghil, M., Mullhaupt A., and Pestiaux, P., 1987. Deep water formation and Quaternary glaciations, Climate Dyn., 2, 1-10.

[10] Ghil, M., and Childress, S. 1987. Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory and Climate Dynamics, Springer-Verlag, New York/Berlin/London/Paris/ Tokyo.

[11] Keilis-Borok, V. I., 1990. The lithosphere of the Earth as a non-linear system with implications for earthquake prediction. Review of Geophysics, 28, No 1, pp.19-34. Translated into Chinese

[12] Alekseevskaya, M., Gabrielov, A., Gelfand, I., Gvishiani, A., and Rantsman, E., 1977. Formal morphostructural zoning of mountain territories. J. Geophys., 43, 227-233

[13] Gelfand, I.M., Guberman, Sh. A., Keilis-Borok, V. I., Knopoff, L., Press, F., Ranzman, E. Ya., Rotwain, I. M., and Sadovsky, A. M., 1976. Pattern recognition applied to earthquake epicenters in California. Phys. Earth and Planet. Inter., 11: 227-283.

[14] Keilis-Borok, V. I., 1996. Intermediate-term earthquake prediction. Proc. Natl. Acad. Sci. USA, Vol.93, pp.3748-3755.

[15] Scholz, C. H., 1990. The Mechanics of Earthquakes and Faulting. Cambridge University Press, Cambridge.

[16] Turcotte, D. L., 1997. Fractals and Chaos in Geology and Geophysics, 2nd ed. Cambridge University Press, Cambridge.

[17] Turcotte, D. L., Newman, W. I., and Gabrielov, A., 2000. 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, 83-96.

[18] Keilis-Borok, V. I., 1994. Symptoms of instability in a system of earthquake-prone faults, Physica D, 77, 193-199.

[19] Rundle, J., Turcotte, D. and Klein, W.(Eds.), 2000. Geocomplexity and the Physics of Earthquakes. American Geophysical Union, Washington, DC.

[20] Allegre, C. J., Le Mouel, J. L., and Provost, A., 1982. Scaling rules in rock fracture and possible implications for earthquake prediction. Nature, 297, 47-49.

[21] Sammis, C. G., Sornette, D., and Saleur, H., 1996. Complexity and earthquake forecasting. In: J. Rundle, D. L. Turcotte, and W. Klein (Eds.), SFI Studies in the Science of Complexity, v. XXV, Addison-Welsey.

[22] Yamashita T. and Knopoff, L., 1992. Model for intermediate-term precursory clustering of earthquakes. J.Geophys. Res., 97, 19873-19879.

[23] Gabrielov, A., and Newman, W. I. 1994. Seismicity modeling and earthquake prediction: A review. In W.I.Newman, A.Gabrielov, and D.L.Turcotte (eds), Nonlinear Dynamics and Predictability of Geophysical Phenomena. Am. Geophys. Un., Int. Un. of Geodesy and Geophys., 7-13 (Geophysical Monograph 83, IUGG Vol. 18).

[24] Gorshkov, A.I., Keilis-Borok, V. I., Rotwain, I.M., Soloviev, A.A., and Vorobieva, I.A., 1997. On dynamics of seismicity simulated by the models of blocks-and-faults systems, Annali di Geofisica, XL, 5, 1217-1232.

[25] Huang, Y., Saleur, H., Sammis, C., and Sornette, D., 1998. Precursors, aftershocks, criticality and self-organized criticality. Europhys. Lett. 41, 43-48.

[26] Newman, W. I., Turcotte, D. L. and Gabrielov, A., 1995. Log-periodic behavior of a hierarchical failure model with applications to precursory seismic activation. Phys. Rev. E, 52, 4827-4835.

[27] Pepke, G. F., Carlson, J. R., and Shaw, B. E., 1994. Prediction of large events on a dynamical model of fault. J. Geiphys. Res., 99, 6769-6788.

[28] Narkunskaya, G. S. and Shnirman, M. G., 1990. Hierarchical model of defect development and seismicity. Phys. Earth. Planet. Inter., 61, 29-35.

[29] Narkunskaya, G. S. and Shnirman, M. G., 1994. An algorithm of earthquake prediction. Computational Seismology and Geodynamics (AGU, Washington, D.C.), 1, 20-24.

[30] Shebalin, P., Zaliapin, I., and Keilis-Borok, V., 2000. Premonitory rise of the earthquakes' correlation range: Lesser Antilles. Phys. Earth Planet. Int. 122 (3-4), 241-249.

[31] Zaliapin, I., Keilis-Borok, V. I., and Axen, G., 2000. Premonitory spreading of seismicity over the fault network in Southern California: Precursor Accord. Submitted to J. Geophys. Res.

[32] Zoeller, G., Hainzl, S., and Kurths, J., 2001. Observation of growing correlation length as an indicator for critical point behaviour prior to large earthquakes. J. Geophys. Res., 106, 2167-2176.

[33] Knopoff, L., 2000. The magnitude distribution of declustered earthquakes in Southern California. Proc. Natl. Acad. Sci., 97, 11880-11884.

[34] Kossobokov, V. G., Romashkova, L. L., Keilis-Borok, V. I., and Healy, J. H., 1999. 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.

[35] Molchan, G. M., Dmitrieva, O. E., Rotwain, I. M., and Dewey, J., 1990. Statistical analysis of the results of earthquake prediction, based on burst of aftershocks. Phys. Earth Planet. Int., 61, 128-139.

[36] Vorobieva, I. A., 1999. Prediction of a subsequent large earthquake. Phys. Earth and Planet. Inter., 111, 3-4: 197-206.

[37] Bongard, M. M., Vaintsveig, M. I. Guberman, Sh. A., Izvekova, M. L., and Smirnov, M. S., 1966. The use of self-learning programs in the detection of oil containing layers. Geol. Geofiz., 6: 96-105.

[38] Press, F., and Briggs, P., 1975. Chandler wobble, earthquakes, rotation, and geomagnetic changes. Nature (London), 256, 270-273.

[39] Press, A., and Allen, C., 1995. Pattern of seismic release in the southern California region. J. Geophys. Res., 100, 6421-6430.

[40] Tukey, J. W., 1977. Exploratory Data Analysis. Addison-Wesley Series in Behavioral Science: Quantitative Methods, Reading, Mass. Addison-Wesley.

[41] Lichtman, A. J., and Keilis-Borok, V. I. 1989. Aggregate-level analysis and prediction of midterm senatorial elections in the United States, 1974-1986. Proc. Natl. Acad. Sci. USA, 86: 10176-10180.

[42] Keilis-Borok, V. I., and Lichtman, A. J., 1993. The self-organization of American society in presidential and senatorial elections. In Yu. A. Kravtsov (Ed.), Limits of Predictability, Springer-Verlag, Berlin-Heidelberg, 223-237.

[43] Lichtman, A. J., 1996. The Keys to the White House. Lanham (ed.): Madison Books.

[44] Keilis-Borok, V., Stock, J. H., Soloviev, A., and Mikhalev, P., 2000. Pre-recession pattern of six economic indicators in the USA. Journal of Forecasting, 19, 1: 65-80.

[45] Keilis-Borok, V. I., Soloviev, A. A., Allegre, C. B., Sobolevskii, A. N., and Intriligator, M. D., 2000. Patterns of macroeconomic indicators preceding the fast acceleration of unemployment in France, manuscript.

[46] Molchan, G. M., 1990. Strategies in strong earthquake prediction, Phys. Earth Planet. Int., 61, 84-98.

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

[48] Molchan, G. M. 1994. Models for optimization of earthquake prediction. Computational Seismology and Geodynamics, AGU, Washington, DC, 2, 1-10.

[49] Molchan, G. M., 1997. Earthquake prediction as a decision-making problem. Pure Appl. Geophys. 149: 233-247.

[50] Kantorovich, L. V., and Keilis-Borok, V. I., 1991. 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: 586-593.

[51] Keilis-Borok, V. I., 1996. Non-seismological fields in earthquake prediction research. In: Sir James Lighthill (Ed.) A Critical Review of VAN, pp. 357-372, World Scientific, Singapore-New Jersey-London-Hong Kong.

[52] Newman, W. I., Gabrielov, A., and Turcotte, D. L. (Eds.), 1994. Nonlinear Dynamics and Predictability of Geophysical Phenomena, Geophys. Monographs Ser. AGU, Washington, DC, 83.

[53] Keilis-Borok, V. I., Knopoff, L. and Rotwain, I. M., 1980. Bursts of aftershocks, long-term precursors of strong earthquakes. Nature (London), 283, 258-263.

[54] Caputo, M., Gasperini, P., Keilis-Borok, V. I., Marcelli, L., and Rotwain, I. M., 1977. Earthquakes' swarms as forerunners of strong earthquakes in Italy. Annali di Geofisica (Roma), vol.XXX, No.3,4, pp.269-283.

[55] Gabrielov, A. M., Caputo, M., Keilis-Borok, V. I., Console, R., and Sidorenko, T. V., 1983. Long-term premonitory seismicity patterns in Italy. Geophys. J.R. Astr. Soc. 75, pp.71-75.

[56] Keilis-Borok, V. I., Lamoreaux, R., Johnson, C., and Minster, B., 1982. Swarms of main shocks in Southern California. In: Earthquake Prediction Research, Rikitake, T. (Ed.), Elsevier, Amsterdam.

[57] Molchan G., and Dmitrieva O., 1990. Dynamics of the magnitude-frequency relation for foreshocks. Phys. Earth Planet. Int., 61, pp. 99-112.

[58] Molchan G. M., Kronrod T. L., and Nekrasova A. K., 1999. Immediate foreshocks: time variation of the b-value. Phys. Earth Planet. Int., 111, 229-240.

[59] Rotwain, I., and Liu, C., 2000. The Precursor Bursts of Aftershocks in Southern California. Manuscript.

[60] Kossobokov V. G., Rastogi, B. K., and Gaur, V. K., 1989. On self-similarity of premonitory patterns in the regions of natural and induced seismicity, P. Indian AS-Earth, 98 (4), 309-318.

[61] Keilis-Borok, V. I. and Malinovskaya, L. N., 1964. One regularity in the occurrence of strong earthquakes. J. of Geophys. Res., 69, 3019-3024.

[62] Knopoff, L., Levshina, T., Keilis-Borok, V. I., and Mattoni, C., 1996. Increased long-range intermediate-magnitude earthquake activity prior to strong earthquakes in California. J. Geophys. Res., 101, B3: 5779-5796.

[63] Gabrielov, A., Newman, W. I., and Turcotte, D. L., 1999. An exactly soluble hierarchical clustering model: inverse cascades, self-similarity and scaling, Phys. Rev. E, 60, 5293-5300.

[64] Ogata, Y., 1992. Detection of precursory relative quiescence before great earthquakes through a statistical model. J. Geophys. Res. 97, 19845-19871.

[65] Schreider, S. Yu., 1990. Formal definition of premonitory seismic quiescence. Phys. Earth Planet. Inter. 61, 113-127.

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

[67] Varnes, D. J., 1989. Predicting earthquakes by analyzing acceleration precursory seismic activity. Pure Appl. Geophys., 130, 661-686.

[68] Bufe, C. G. and Varnes, D. J., 1993. Predictive modeling of the seismic cyrcle of the greater San Francisco bay region. J. Geophys. Res. 98, 9871-9883.

[69] Bufe, C. G., Nishenko, S. P. and Varnes, D. J., 1994. Seismicity trends and potential for large earthquakes in the Alaska-Aleutian region. Pure Appl. Geophys., 142, 83-99.

[70] Johansen, A., Sornette, D., Wakita, H., Tsunogai, U., Newman, W. I., and Saleur, H., 1996. Discrete scaling in earthquake precursory phenomena: evidence in the Kobe earthquake, Japan. J. Phys. I France, 6: 1391-1402.

[71] Sornette, D., and Sammis, C. G., 1995. Complex critical exponents from renormalization group theory of earthquakes: Implications for earthquake predictions. J. Phys. I France, 5: 607-619.

[72] Bowman, D. D., Ouillon, G., Sammis, C. G., Somette, A., and Sornette, D., 1998. An observational test of the critical earthquake concept, J. Geophys. Res., 103, 24 359-24 372.

[73] Yiou, P., Sornette, D., and Ghil, M. 2000. Data-adaptive wavelets and multi-scale SSA, Physica D,124,254-290.

[74] Rammal, R., Toulous, G., and Virasoro, M. A., 1986. Ultrametricity for physicists, Rev. Mod. Phys., 58, 765-788.

[75] Prozorov, A. G., 1975. Variations of seismic activity related to locations of strong earthquakes. Vychislitel'naya Seismologiya, 8, 71-72.

[76] Prozorov, A. G., 1975. Changes of seismic activity connected to large earthquakes. Interpretation of data in seismology and neotectonics. Computational Seismology, 8, Moscow, Nauka, 71-82.

[77] Prozorov, A. G., 1994. A new test for the statistical significance of distant interaction of large earthquakes. Seismisity and related processes in the environment. Global Changes of Environment and Climate. Federal Research Program of Russia, 1., Russian Ac. Sci., 69-73.

[78] Ferguson, C. D., Klein, W., and Rundle, J. B., 1998. Long range earthquake fault models, Computers in Physics, 12, 34-40.

[79] Preston, E., de la Martins, J. S., Rundle, J. B., Anghel, M., and Klein, W., 2000. Models of earthquake faults with long-range stress transfer, Computing in Science and Engineering, 2, 34-41.

[80] Soloviev, A. and Vorobieva, I., 1999. 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, 18 pp.

[81] Rice., J., and Gu, J., 1983. Earthquake after effects and triggered seismic phenomena. Pure Appl. Geophys., 121: 187-219.

[82] Gomberg, J., Blanpied, M. L., and Beeler, N. M., 1997. Transient triggering of near and distant earthquakes. BSSA 87, 2: 294-309.

[83] Hill, D. P., Reasenberg, P. A., Michael, A., Arabasz, W. J., Beroza, G., Brumbaugh, D., Brune, J. N., Castro, R., Davis, S., dePolo, D., Ellsworth, W. L., Gomberg, J., Harmsen, S., House, L., Jackson, S. M., Johnston, M., Jones, L., Keller, R., Malone, S., Munguia, L., Nava, S., Pechmann, J.C., Sanford, A., Simpson, R.W., Smith, R. S., Stark, M., Stickney, M., Vidal, A., Walter, S., Wong, V., and Zollweg, J., 1993. Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. Science, 260: 1617-1623.

[84] Vilkovich, E. V. and Shnirman, M. G., 1982. Epicenter migration waves: examples and models. Comput. Seism., 14, 27-37.

[85] Prozorov, A. G. and Schreider, S. Yu., 1990. Real time test of the long-range aftershock algorithm as a tool for mid-term earthquake prediction in Southern California. PAGEOPH 133, 329-347.

[86] Romanowicz, B., 1993. Spatiotemporal patterns in the energy-release of great earthquakes, Science, 260, 1923-1926.

[87] Aki, K., 1996. Scale dependence in earthquake phenomena and its relevance to earthquake prediction. Proc. Natl. Acad.Sci. USA, 93, 3748-3755.

[88] Gutenberg B. and Richter, C. F., 1954. Seismicity of the Earth and Associated Phenomena. Princeton University Press, Princeton.

[89] Molchan G. M., and Podgaetskaya, V. M. 1973. Parameters of global seismicity. Comput. Seism. 6, 44.

[90] Rundle, J. B., 1993. Magnitude-frequency relations for earthquakes using a statistical mechanical approach. J. Geophys. Res. 98, 21943-21949.

[91] Turcotte, D. L., 1999. Seismicity and self-organized criticality, Phys. Earth Planet. Inter., 111, 275-294.

[92] Kolmogorov A. N., 1941. On the logarithmically normal law of distribution of the size of particles under pulverization. Dokl. Akad. Nauk SSSR, 31, 99-101.

[93] Haberman R. E., 1981. Precursory seismicity pattern: Stalking the mature seismic gap. In: Earthquake Prediction: An International Review. Maurice Ewing Series, American Geophysical Union, Washington, DC, 4, 29-42.

[94] Mogi, K., 1981. Seismicity in Western Japan and long-term forecasting. In: Earthquake Prediction: An International Review. Maurice Ewing Series, American Geophysical Union, Washington, DC, 4, 43-51.

[95] Smith, W., 1981. The b-value as an earthquake precursor. Nature, 289, 136-139.

[96] Prozorov, A., Sabina, F. J., 1984. Study of seismic properties in the Mexico region. Geophys. J.R. Astr. Soc., 76, 317-336.

[97] Rotwain, I., Keilis-Borok, V. and Botvina, L., 1997. Premonitory transformation of steel fracturing and seismicity, Phys. Earth Planet. Int., 101, 61-71.

[98] Molchan, G. M., Kronrod, T. L., and Panza, G. F., 1997. Multiscale seismicity model for seismic risk. Bull. Seismol. Soc. Amer., 87 (5).

[99] Okal, E. A. and Romanovicz, B. A., 1994. On the variation of b-values with earthquake size. Phys. Earth Planet. Inter. 87, 55-76.

[100] Shnirman, M. G. and Blanter, E. M., 1999. Mixed hierarchical model of seismicity: scaling and prediction, Phys. Earth planet. Inter., III, 295-304.

[101] Keilis-Borok, V. I., 1982. A worldwide test of three long-term premonitory seismicity patterns - a review. Tectonophysics, Vol.85, pp.47-60.

[102] Keilis-Borok, V. I., Rotwain, I. M., 1990. Diagnosis of time of increased probability of strong earthquakes in different regions of the world: algorithm CN. Phys. Earth Planet. Inter., 61, 1-2: 57-72.

[103] Keilis-Borok, V. I., and Kossobokov, V. G., 1990. Premonitory activation of earthquake flow: algorithm M8. Phys. Earth Planet. Inter., 61, 1-2: 73-83.

[104] Kossobokov, V. G., Keilis-Borok, V. I., and Smith, S. W., 1990. Localization of intermediate-term earthquake prediction, J. Geophys. Res., 95, 19763-19772.

[105] Rotwain, I. and Novikova, O., 1999. Performance of the earthquake prediction algorithm CN in 22 regions of the world, Phys. Earth Planet. Inter., 111, 207-214.

[106] Vorobieva, I. A. and Levshina, T. A., 1994. Prediction of a second large earthquake based on aftershock sequence Computational Seismology and Geodynamics, AGU, Washington, DC, vol. 2, pp. 27-36.

List of Figures

Figure 1. Scheme of prediction. See text for details.

Figure 2. Possible outcomes of prediction.

Figure 3. Synthetic earthquake sequence, consecutively zoomed; the sequence is generated by the colliding-cascades 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.

Figure 4. Magnitude-frequency relation (Gutenberg-Richter law). The solid dots represent the actual value of the number $N(m)$ of events with magnitude m, while the solid line is the least-square fit of a power law $\log_{10} N(m) = a - bm$. a) All events. b) Main shocks. c) Aftershocks.

Figure 5. Premonitory acceleration of cumulative Benioff stress release $\epsilon(t) = A - B(t_f - t)^\alpha$ vs. $t - t_f$: 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) $\epsilon(t)$ in linear scale; b) $\epsilon(t)$ in log-log scale; and c) residuals: the difference between the data and the fitting curve in a linear scale. Note the accelerating oscillations prior to a major earthquake in panel c).

Figure 6. Bursts of aftershocks for the composite synthetic sequence. Each panel shows a measure of aftershock activity vs. time to a major event $(t - t_f)$, t_f being an occurrence time of the latter. Two different measures are considered. Both are counted

at the very beginning of the whole aftershock sequence, within 300 time units after a main shock. Panels at the top show the functional B_m^{Σ} that is the weighted number of aftershocks [see Sect. 2.6, Eq. (10)]: a) $m = 6$, b) $m = 5$. Panels at the bottom show the the number of aftershocks B_m [see Sect. 2.6] c) $m = 6$, d) $m = 5$. Vertical arrows indicate the moment of each major event, all plotted at $t - t_f = 0$ on the abscissa. Significant growth of activity is observed at several hundred time units prior to major events. The measure B_m^{Σ} indicates premonitory rise of clustering better than the measure B_m. See text for details.

Figure 7. Premonitory transformation of the model's magnitude distribution. Panels on the left side show magnitude distribution in a composite seismicity for two distinct time intervals prior to a major event; t_f is its occurrence time. a) main shocks, c) all events. Circles: $-50001 < t - t_f < -2000$; squares: $-2001 < t - t_f < 0$. Figures show explicit premonitory upward bend at magnitudes $m \geq 5$; it is stronger for the main shocks. Panels on the right side show the functional $U(t)$ that depicts such a bend (cf. Eq. (11)). b) main shocks, d) all events. The premonitory rise of $U(t)$ is apparent in both panels. See details in the text.

Figure 8. Error diagrams for all premonitory patterns considered. Panels on the left side show the rate of the total duration τ of alarms vs. the rate n of failures to predict. Panels on the right side show the rate f of false alarms vs. the rate n of failures to predict. a), b) Weighted number of events, Σ_m, $2 \leq m \leq 6$ (Sect. 2.4); c), d) number of events, N_m, $m = 2, \ldots, 6$ (Sect. 2.4); e), f) bursts of aftershocks, B_m, $m = 4, \ldots, 6$

(Sect. 2.6); g), h) accord, A_m, $m = 5, 6$ and, Π (Sect. 2.7); i), j) correlation range, R_m, $m = 2, \ldots, 6$ (Sect. 2.7); and k), l) change in magnitude distribution, U (Sect. 2.8). See text for details.

Figure 9. Performance of single premonitory seismicity patterns (PSPs) prior to 10 major earthquakes. Figure juxtaposes the alarms generated by each of 23 functionals considered. Each box shows the emergence of a pattern prior to a major event; the events sequential index (out of total 43) is shown at the top of the box. The right edge of the box is the epoch of a major event. Each horizontal row represents the track record of a pattern, identified at the left side of the panel. Shaded areas show the time intervals when an alarm was declared by the pattern. The three panels correspond to different prediction strategies: a) minimization of failures to predict ($n \to \min$); b) minimization of sum of errors ($n + \tau + f \to \min$); and c) minimization of the number of false alarms ($f \to \min$).

Figure 10. Performance of the single premonitory seismicity patterns (PSPs) in a continuous time interval. The figure juxtaposes alarms generated by all 23 functionals considered. Each panel shows the alarms generated during the time interval $[3.05, 3.70] \cdot 10^6$ time units. The horizontal line represents the track record of a pattern identified at the left side of the panel. Shaded areas show the time intervals when an alarm was declared. Heavy solid vertical lines and vertical arrows depict the epoch of major events ($m = 7$). Heavy shaded vertical lines depict the epoch of main shocks with magnitude $m = 6$. The three panels correspond to different prediction strategies, as in Fig. 9.

Figure 11. Emergence of the single premonitory seismicity patterns (PSPs) prior to each major earthquake. Each panel shows in separate boxes emergence of patterns prior to each of 43 major events: shaded boxes represent predicted events while empty boxes represent failures to predict. The sequential index of each major event (out of total 43) is shown at the top of each box. Each horizontal line represents the track record of a specific pattern. Shaded boxes indicate succesful predictions. The two panels correspond to different prediction strategies: a) minimization of sum of errors ($n + \tau + f \to \min$); and b) minimization of the number of false alarms ($f \to \min$).

Figure 12. Collective performance of the patterns. Error diagrams: a), b) Six patterns 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, $M = 1, ..., 6$. 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 patterns are tuned to minimize the rate of false alarms. The straight line $n + \tau = const$ is shown in each panel for comparison. See text for details.

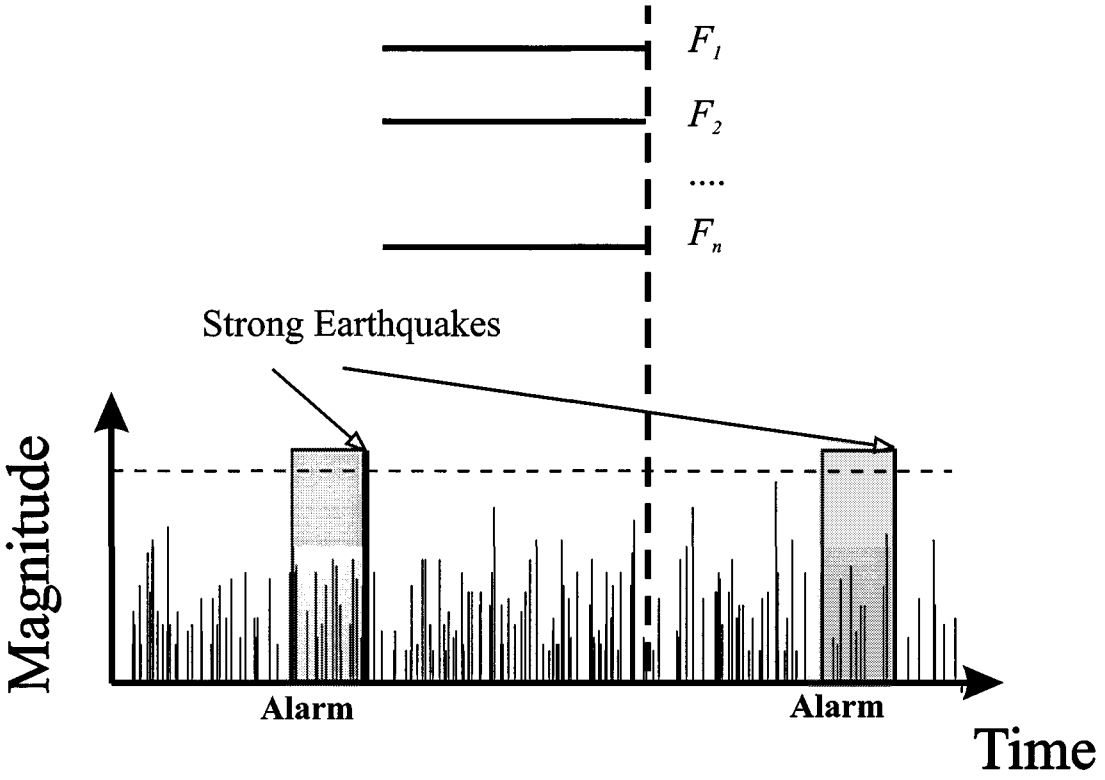

Figure 1.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

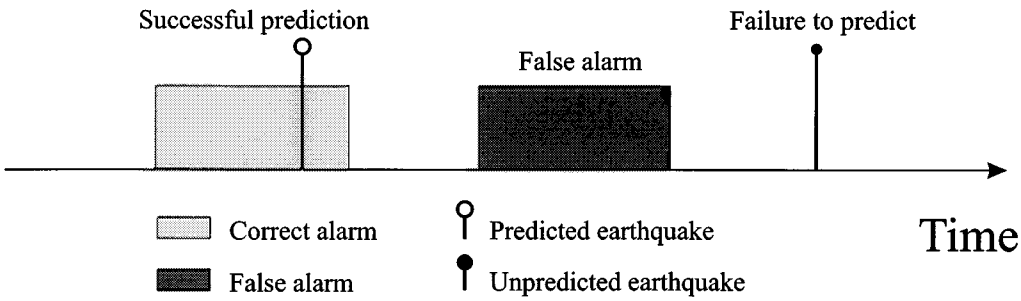

Figure 2.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

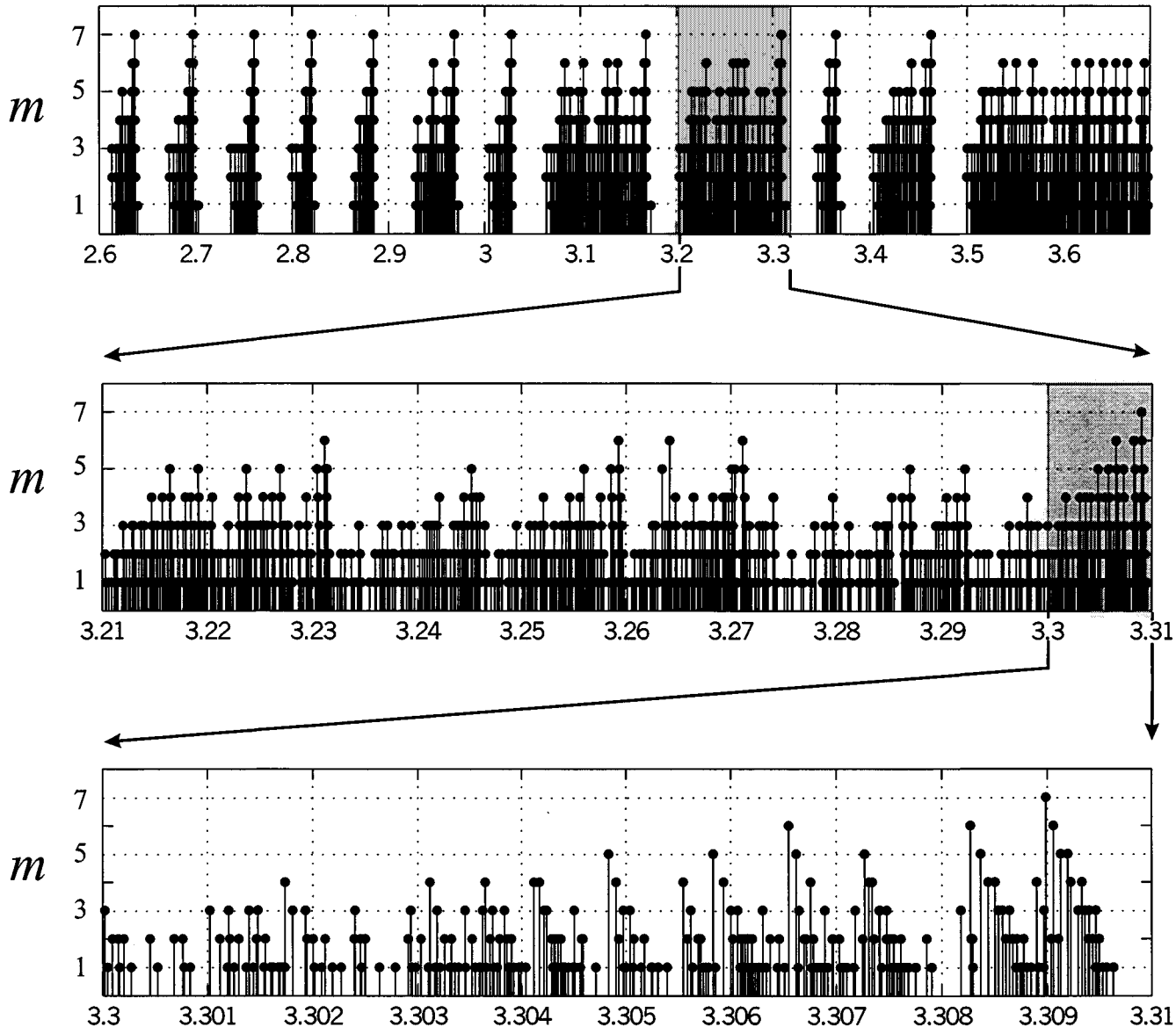

Figure 3.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

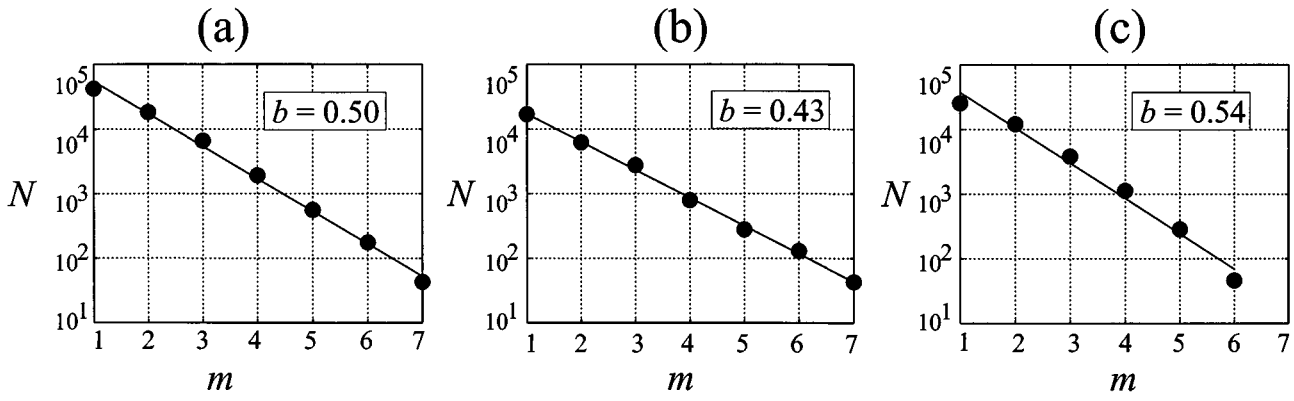

Figure 4.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

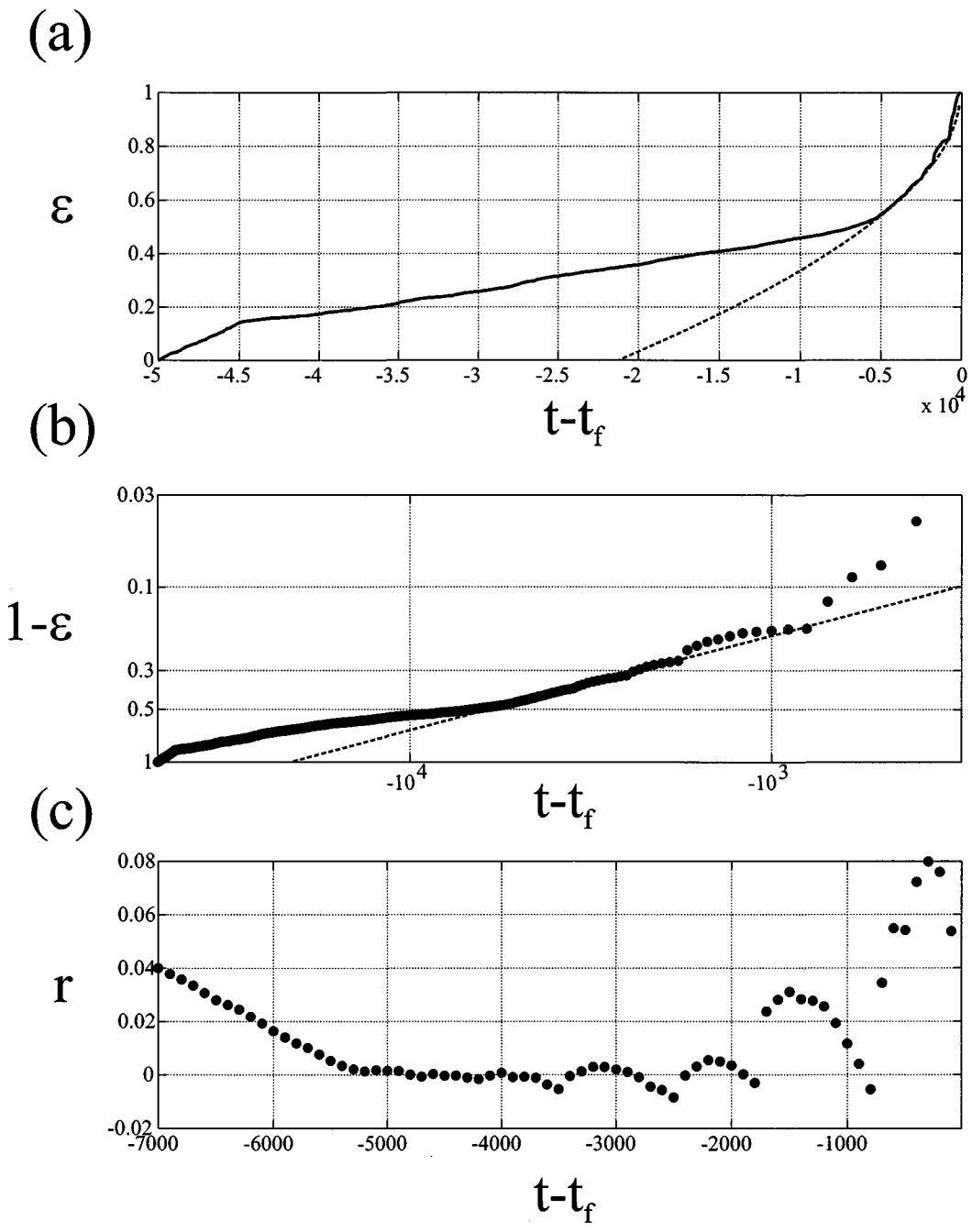

Figure 5.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

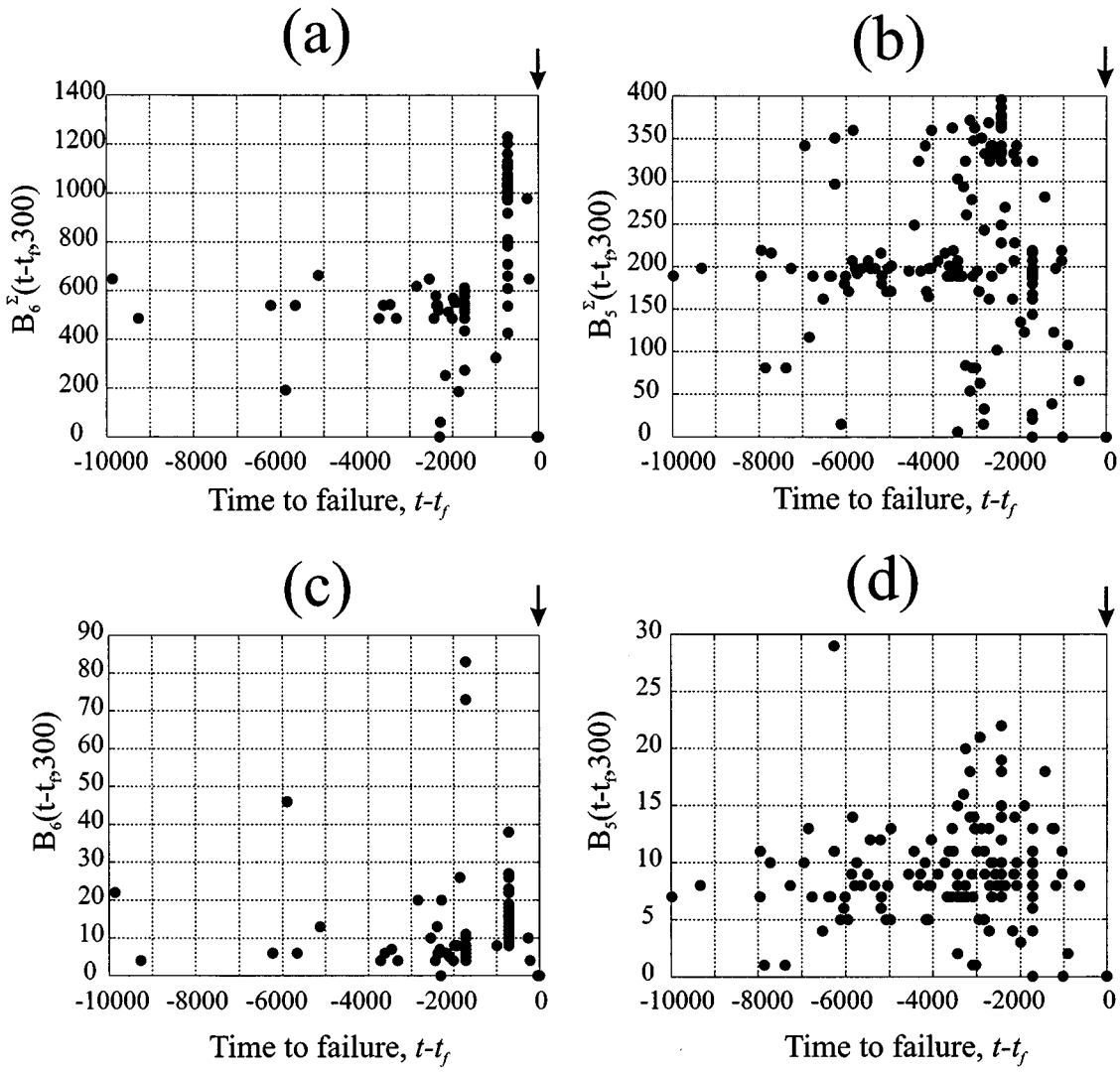

Figure 6.
CCM-BDE. PartII. 2001
I. Zaliapin, V. Keilis-Borok, M. Ghil

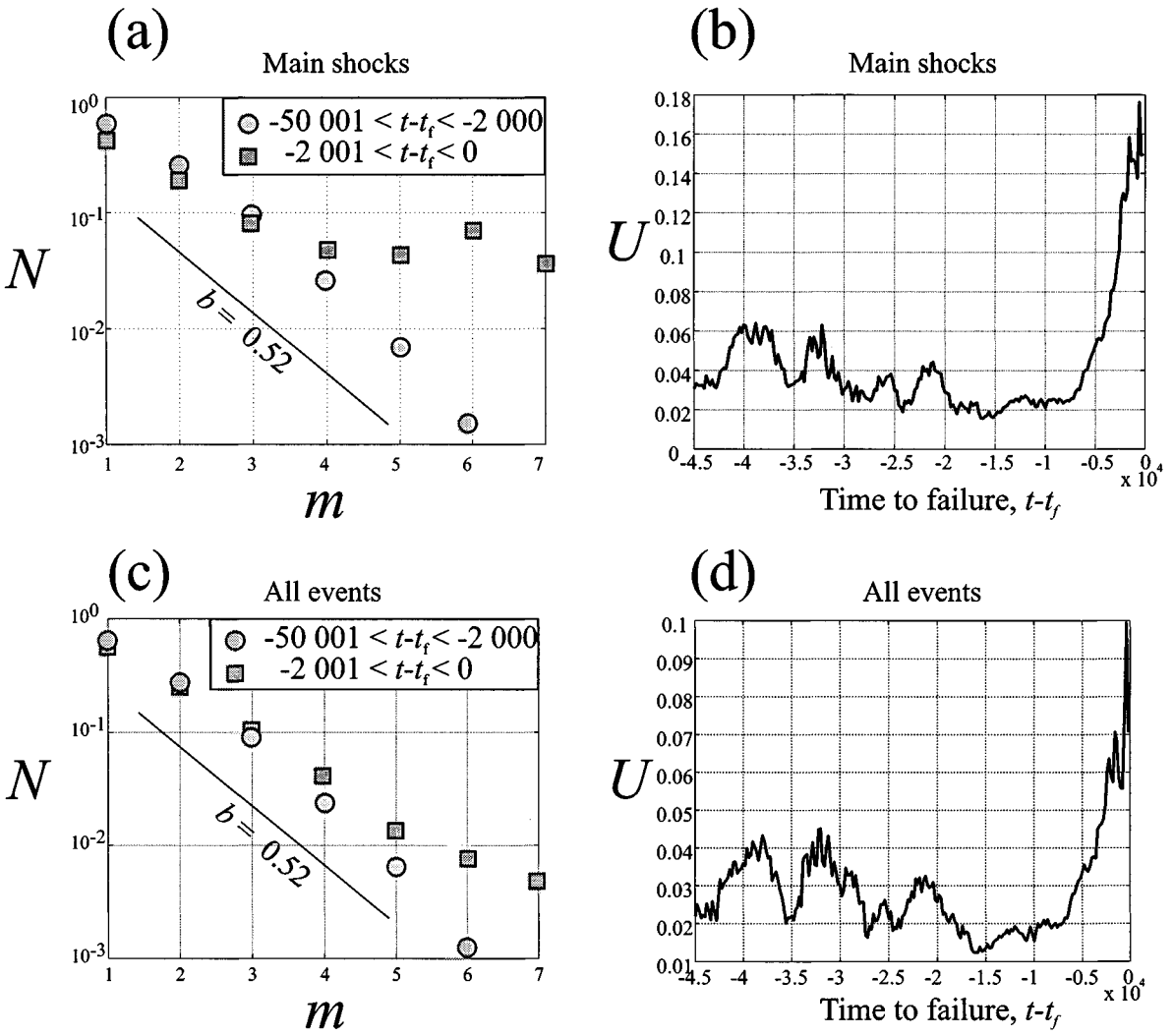

Figure 7.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

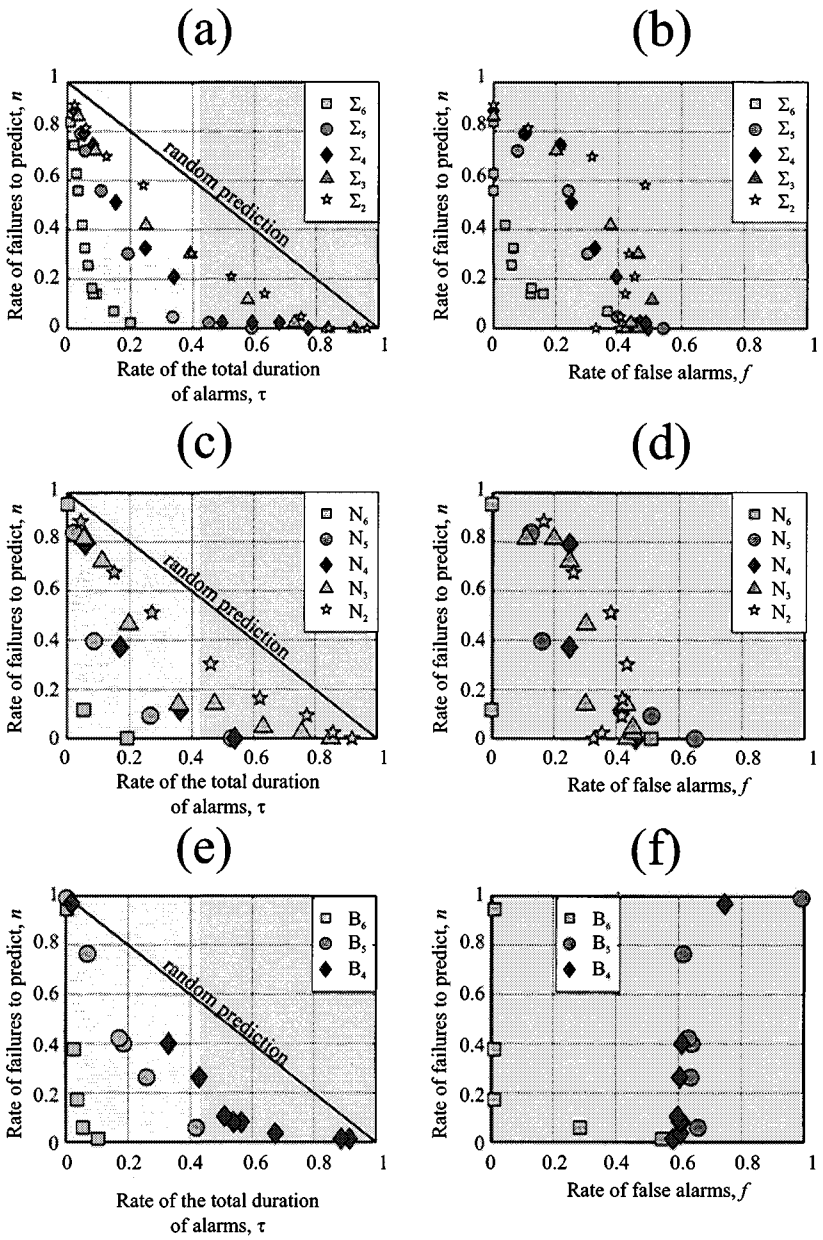

Figure 8 (to be continued.)

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

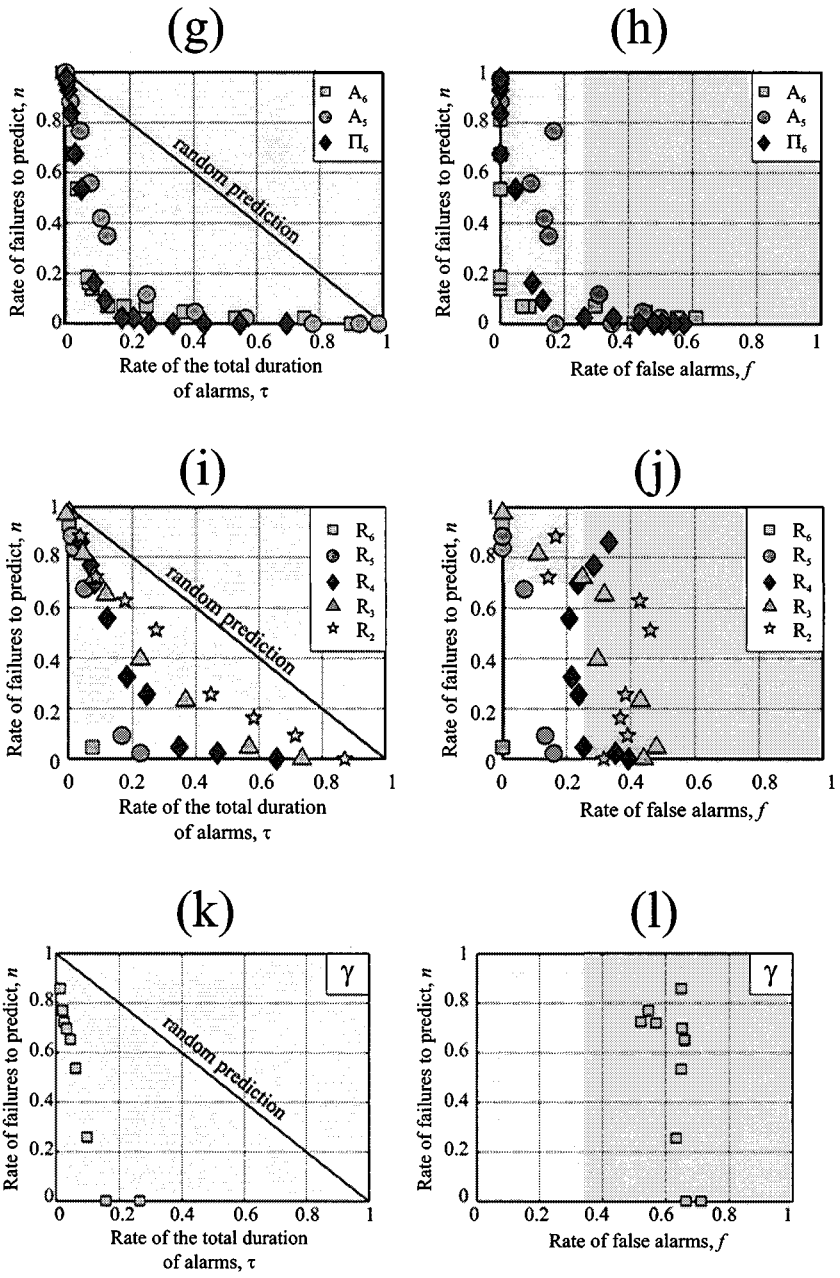

Figure 8 (cont.)
CCM-BDE. PartII. 2001
I. Zaliapin, V. Keilis-Borok, M. Ghil

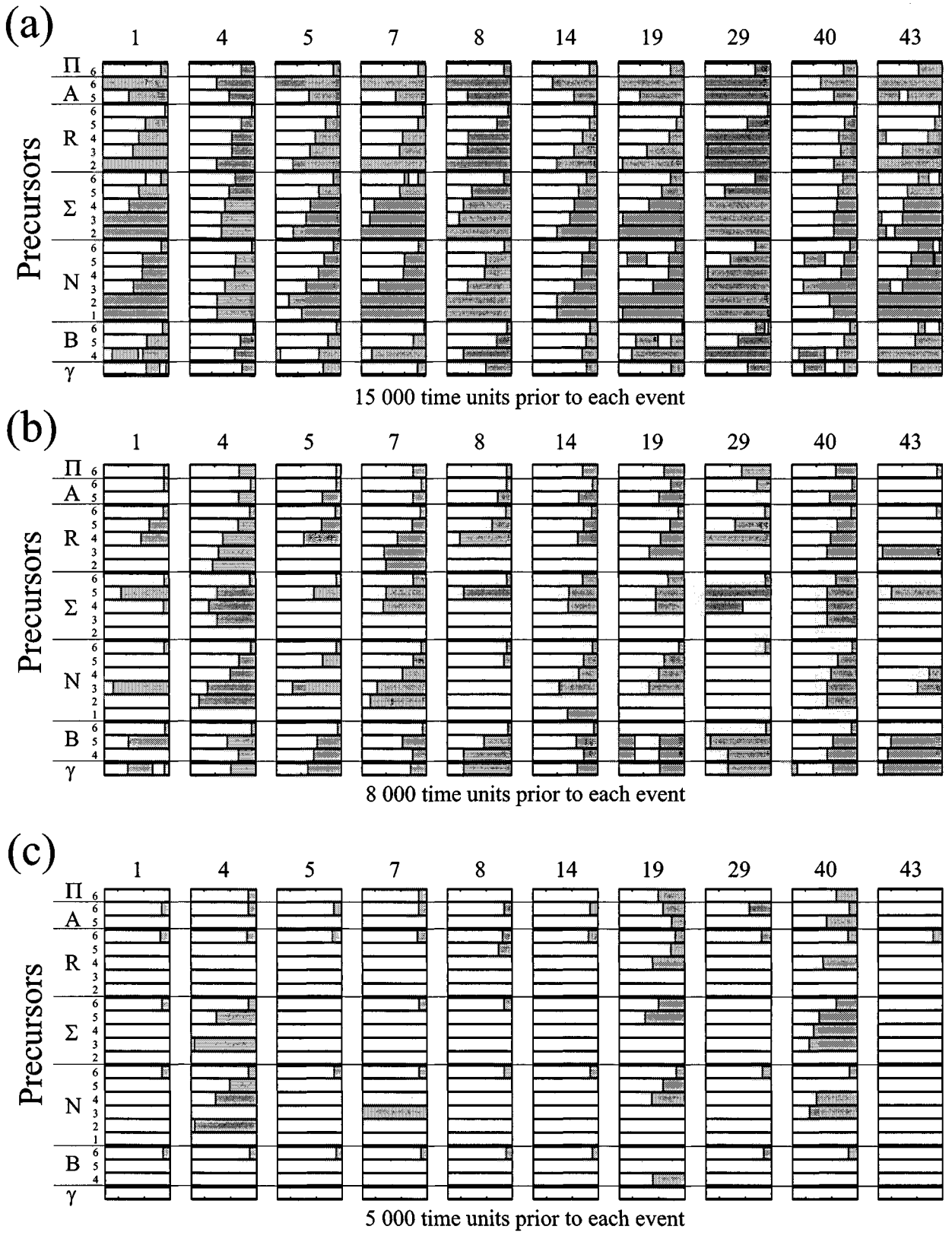

Figure 9.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

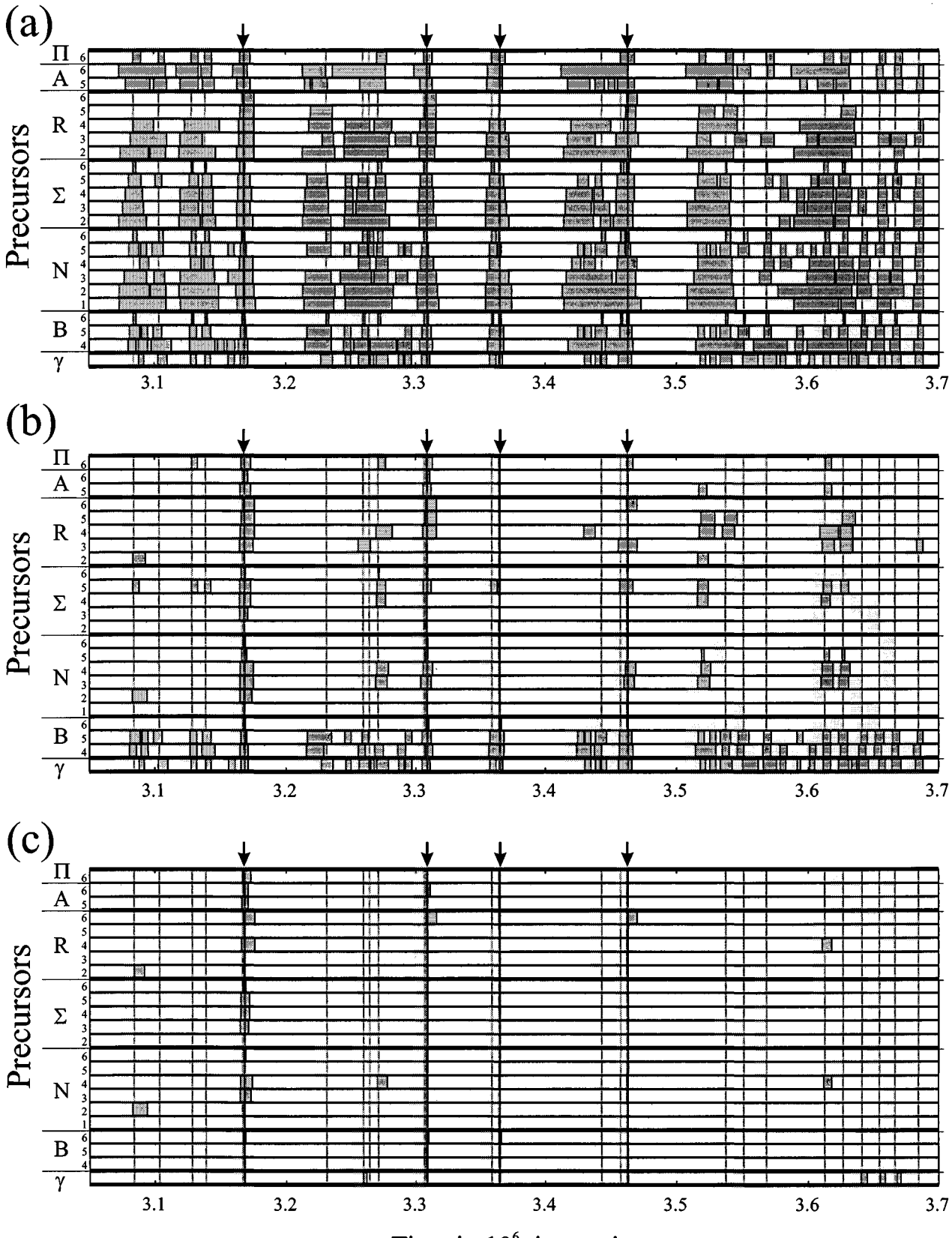

Figure 10.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

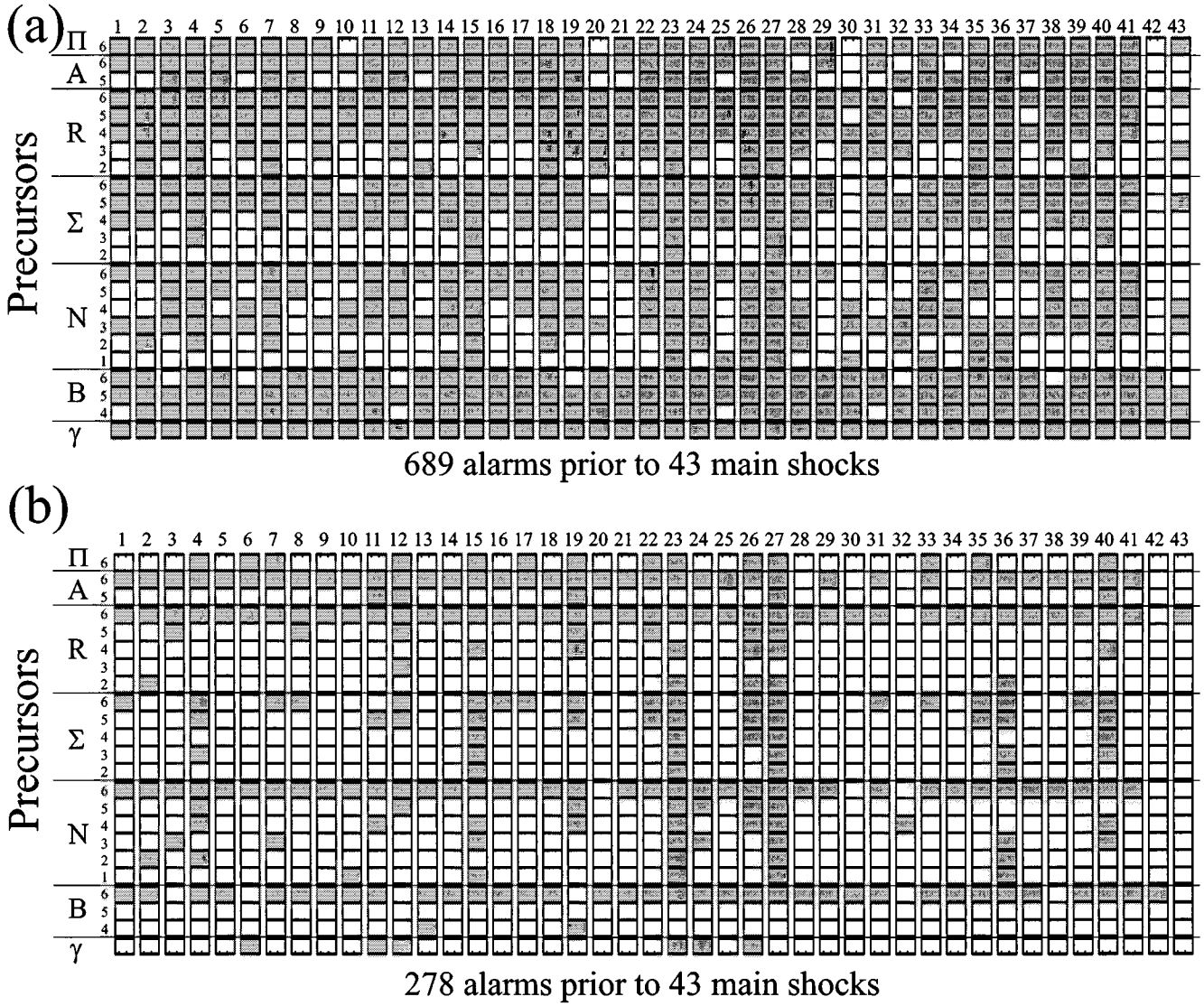

Figure 11.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil

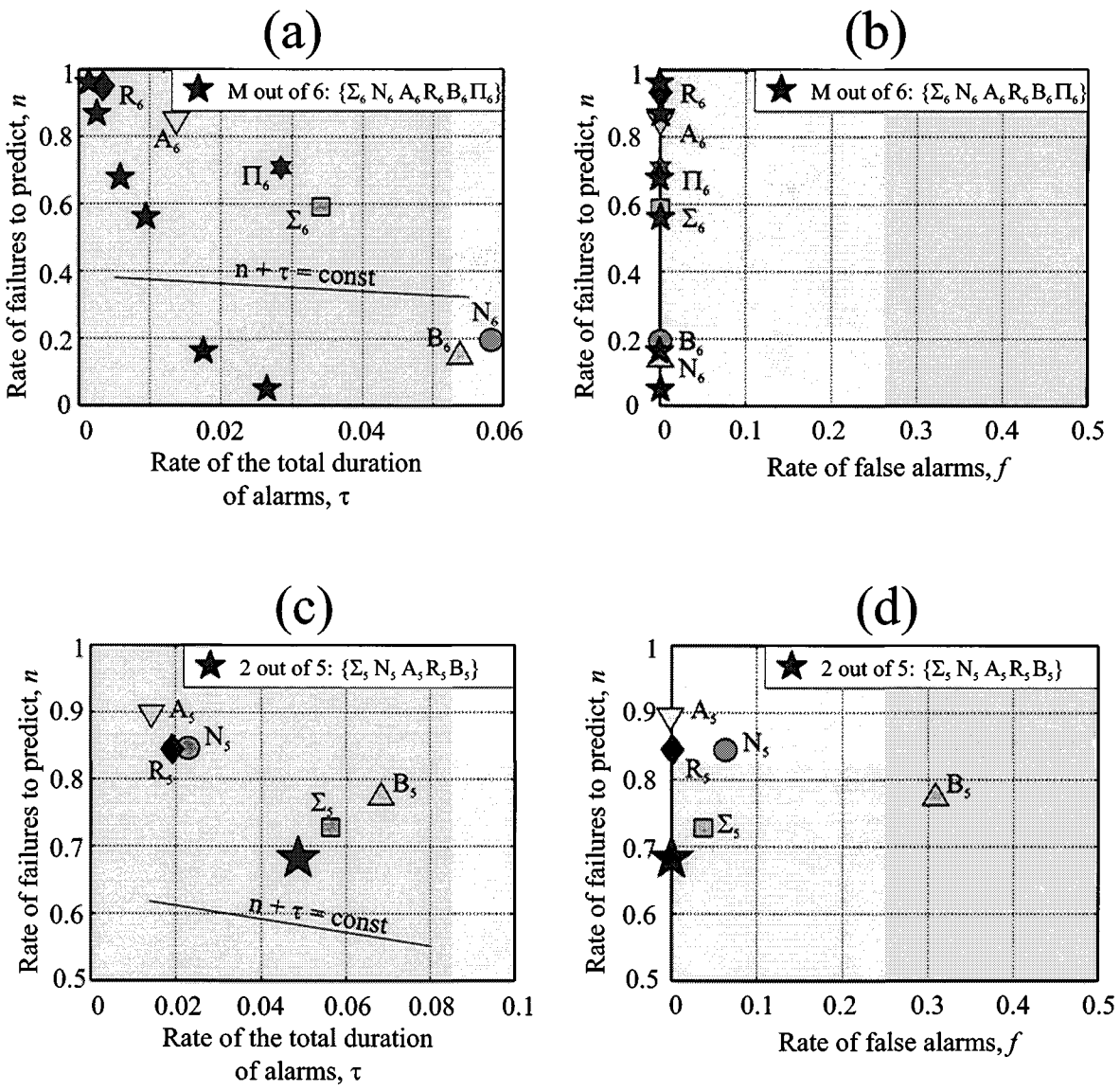

Figure 12.

CCM-BDE. PartII. 2001

I. Zaliapin, V. Keilis-Borok, M. Ghil