The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific
and Cultural Organization

International Atomic
Energy Agency

SMR.1676 - 40

8th Workshop on Non-Linear Dynamics and Earthquake Prediction

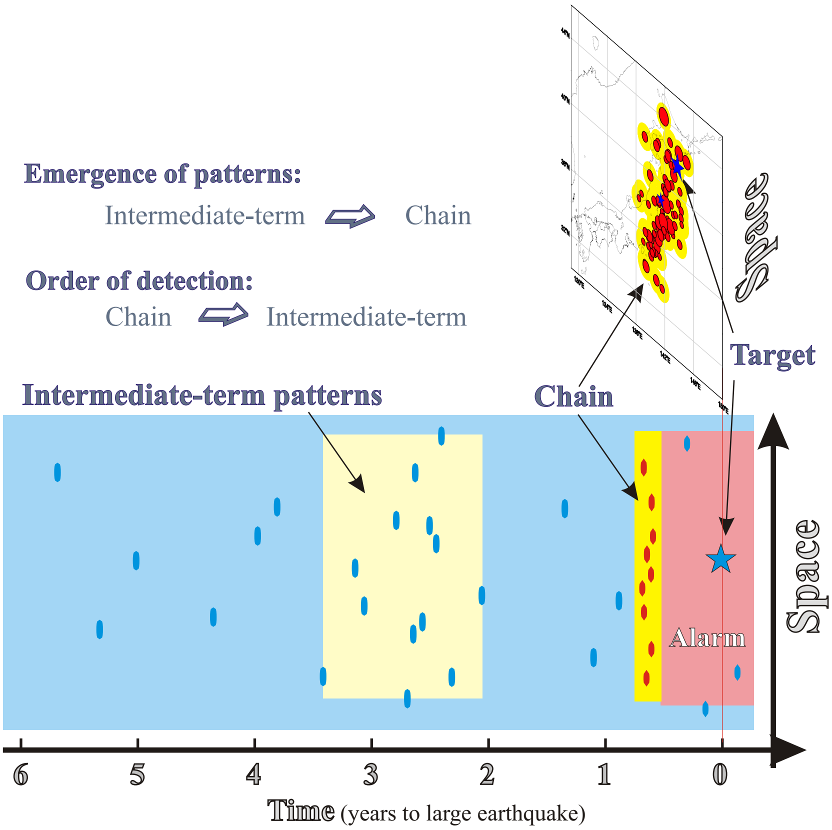
3 - 15 October, 2005

Test in Month-in-Advance Earthquake Prediction by RTP Algorithm. Estimating False Alarm and Random Success Probabilities.

Peter Shebalin
Russian Academy of Sciences
International Inst. of Earthquake Prediction Theory and
Matematical Geophysics
117556 Moscow
Russia
&
Institut de Physique du Globe,
Paris,
France

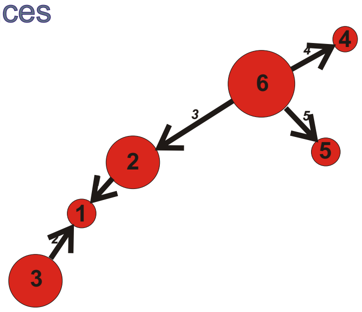
Strada Costiera 11, 34014 Trieste, Italy - Tel. +39 040 2240 111; Fax +39 040 224 163 - sci_info@ictp.it, www.ictp.it

Reverse tracing of precursors (*RTP*)



Pattern "Chain of epicenters"

Chain is a sequence of earthquakes that follow each other closely in time and space, quickly extending over large distances



The chain is formed by union of all pairs of "neighbors" with

$\Delta r_{ij} <= r_0\, 10^{\,c\,\{\min(M_i, M_j)-2.5\}}$ and

$\Delta t_{ij} <= \tau_0$

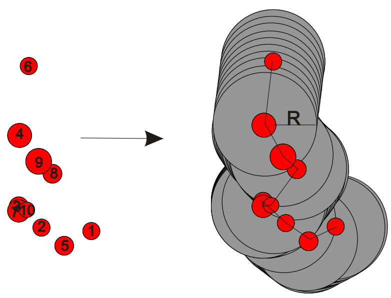
Each element of the chain has at least one neighbor inside the chane and has no neighbors outside

Pattern ***Chain*** depicts the rise of *earthquakes correlation range*. Accordingly, only large chains are considered (at least k_0 elements and length >= l_0).

Only main shocks with $M>=M_{min}$ are considered.

List of parameters: $M_{min},\ \tau_0,\ r_0,\ c,\ k_0,\ l_0$.

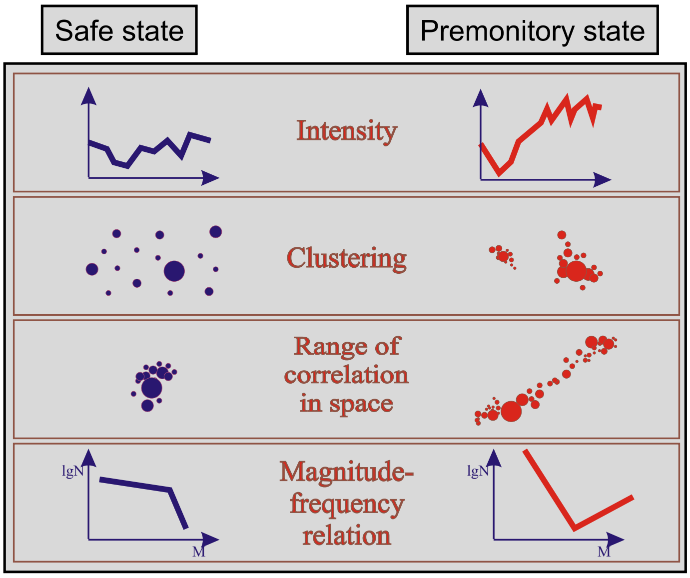
R-vicinity of the chain



Time order is not taken into account

Intermediate-term patterns

considered in the R-vicinity of the chain



- Benioff strain release
- Total area of faultbreaks
- Number of earthquakes

[KB & Malinovskaya 1964; Bufe & Varnes, 1993; Bowman et al., 1998; Knopoff et al., 1996; Jaume & Sykes, 1999]

- Burst of immediate aftershocks
- Swarms of mainshocks

[KB, Knopoff, Rotwain, 1980; Molchan et al, 1990]

- Accord: simultaneous activation of distant areas
- ROC: simultaneous occurrence of distant intermediate events

[Kossobokov, Carlson, 1995; Gabrielov et al., 2000; Shebalin et al., 2000; Zaliapin et al., 2002]

- Upward bend of GR law

[Narkunskaya, Schnirman, 1990; Gabrielov et al., 2000; Zaliapin et al., 2003]

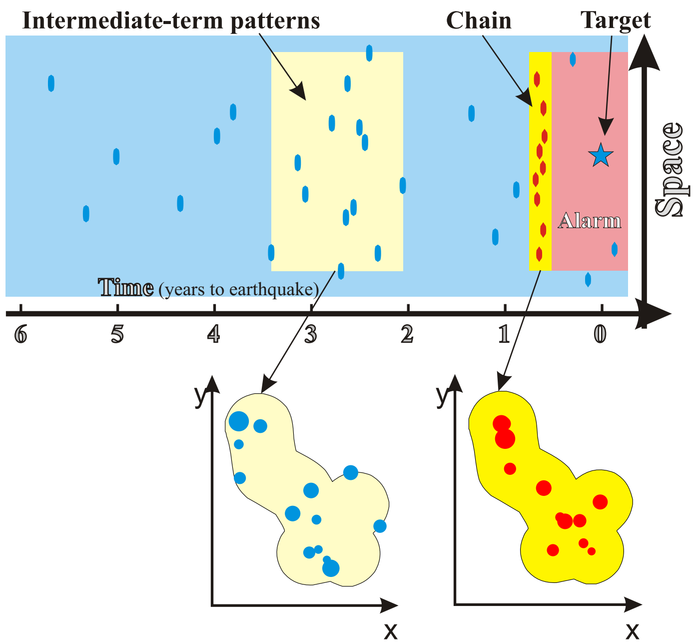
Intermediate-term premonitory seismicity patterns (lead time years) are studied in the *R*-vicinities of the ***Chains*** (lead time months). ***Pattern recognition*** is applied to select precursory chains. Available case histories are used for learning.

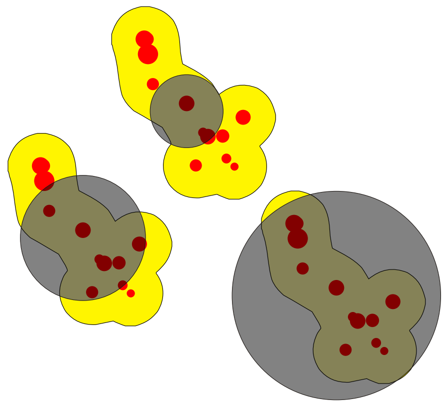
What are the advantages of *RTP*?

***Hypothesis 1*:**
shorter-term pattern (chain) geometrically correlates with intermediate-term patterns.

RTP gives reduction in dimensionality of parameter space where premonitory patterns are looked for.

If circles, what size?

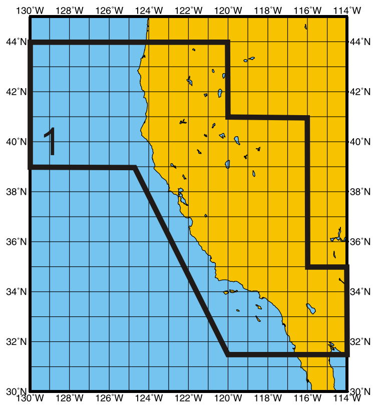


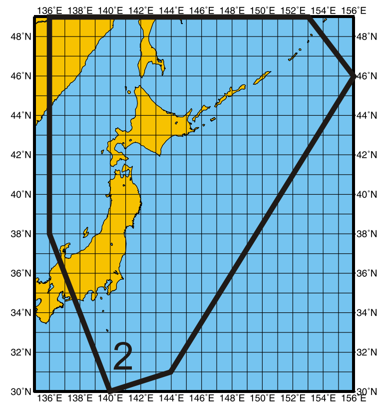


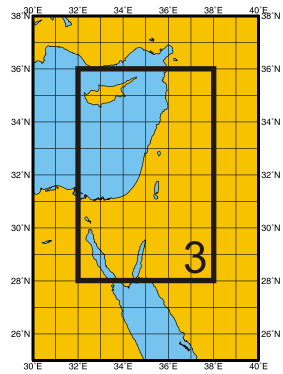
***Hypothesis 2*:**
Chains are self-adapted to size and form of the area where precursors emerged

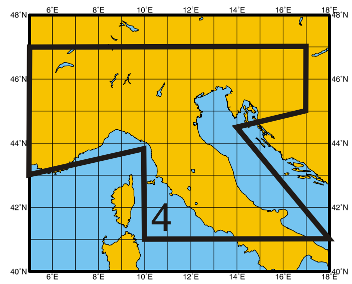
Experiment in month-in-advance earthquake prediction by RTP algorithm

http://www.igpp.ucla.edu/prediction/rtp/









1. California and Western Nevada
2. Honshu-Hokkaido-Southern Kurils
3. Eastern Meditteranean
4. Central Apennines, Po valley with mountain setting

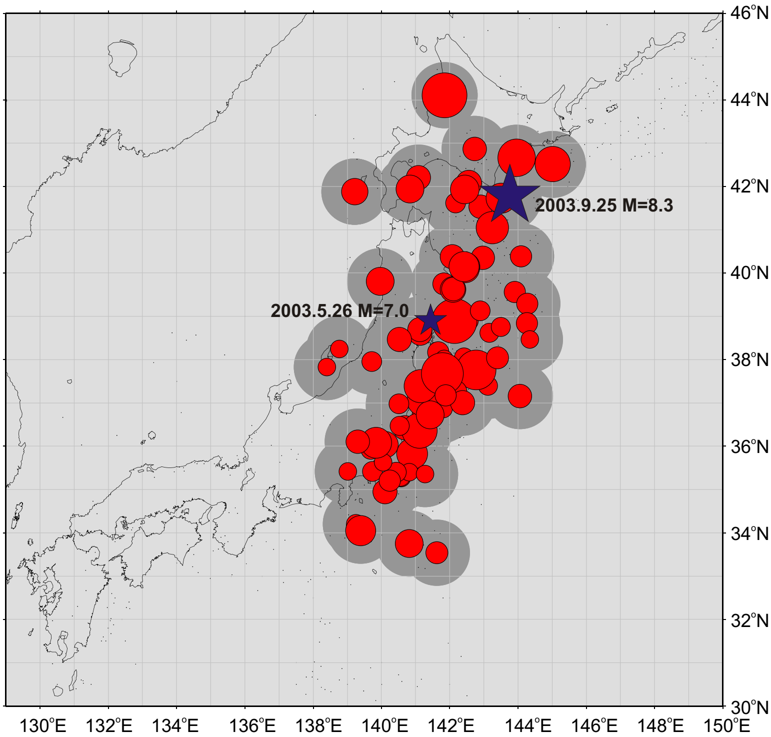
Regions covered by prediction

Summary of the test in month-in-advance earthquake prediction by RTP algorithm

#	*Region/target earthquakes*	*Period of alarm*	*Prediction was put on record on*	*Target earthquake*	*Prediction outcome*	*Probability of a random success*	*Probability of a random success with R_1=2.5R*
1	Japan $M_{JMA}\geq7.0$	Mar 27, 2003- - Jan27, 2004	July 1, 2003	**Sep 25, 2003, M_w=8.3** ***within the alarm***	Correct	0.25	0.34
2	California $M_{ANSS}\geq6.4$	May 5, 2003 - -Feb 27, 2004	June 24, 2003	**Dec 22, 2004, M=6.5** ***within the alarm***	Correct	0.05	0.07
3	Southern California $M_{ANSS}\geq6.4$	Oct 29, 2003- - Sep 05, 2004	May 12, 2004		False alarm	0.08	0.10
4	Honsu, Japan $M_w\geq7.2$	Feb 8, 2004 - - Nov 8, 2004	June 1, 2004	**Sep 5, 2004, M_w=7.4** ***outside the region; 127 km outside alarm***	Near miss (correct with R_1=2.5R)	0.07	0.11
5	Northern Dinarides $M_w\geq5.5$	Feb 29, 2004 - - Nov 29, 2004	May 12, 2004	**Jul 12, 2004, M_w=5.2, M_L=5.7** ***within the alarm***	Near miss (correct for $M_L\geq5.5$)	0.07 estimate was made for $M_L\geq5.5$	0.08 for $M_L\geq5.5$
6	Southern California $M_{ANSS}\geq6.4$	Nov 14, 2004 - - Aug 14, 2005	Nov 16, 2004		False alarm	0.05	0.07
7	Oregon off coast $M_{ANSS}\geq6.4$	Nov 16, 2004 - - Aug 16, 2005	Jan 29, 2005	**Jun 15, 2005, M_w=7.2** ***60 km outside alarm***	Near miss (correct with R_1=2.5R)	0.01	0.03
9	Honsu, Japan $M_w\geq7.2$	June 14, 2005 - - Mar 14, 2006	Oct 1, 2005	**Aug 16, 2005, M_w=7.2** ***within the area of alarm***	Due to the technical delay of data, the alarm was determined after the earthquake	0.05	0.14

An alarm is turned on if the estimated probability that alarm is false is <50%

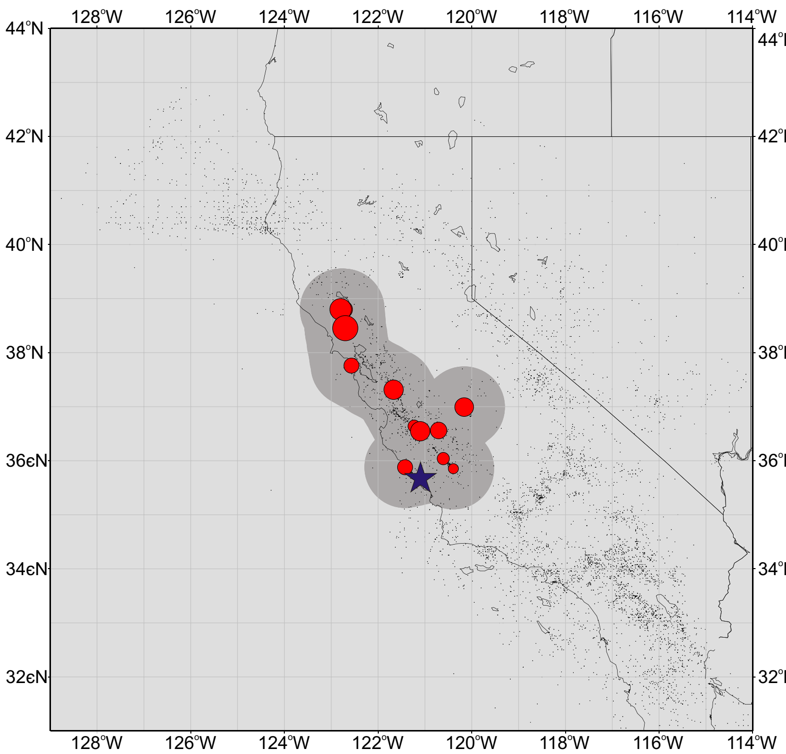
Advance prediction of Tokachi-oki earthquake, Japan, Sept. 25, 2003, M = 8.3



Dots show earthquakes, forming precursory chain. Stars - target earthquakes.

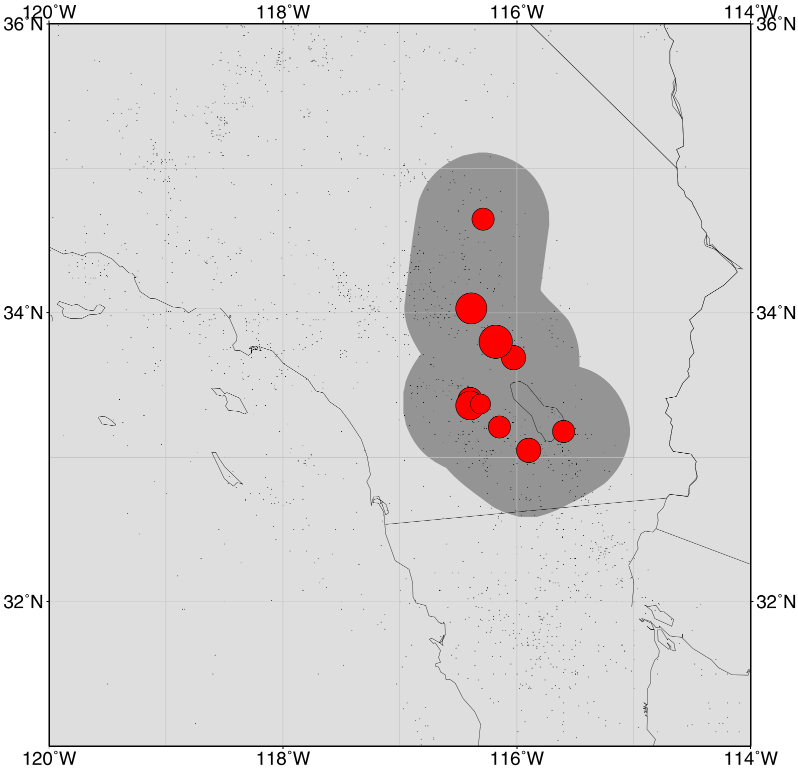
Case history, 2003
March 27: *Precursory chain* of earthquakes was formed. It indicates that an earthquake with magnitude 7 or more will occur in gray area within 9 months.
May 26: *Earthquake with magnitude 7.0* occurred in gray area; precursor was not reported in advance.
July 2: *Precursor reported* at IUGG (Sapporo, Japan).
Sept. 25: *Tokachi-oki earthquake* in gray area.

Advance prediction of San Simeon earthquake in central California, M=6.5



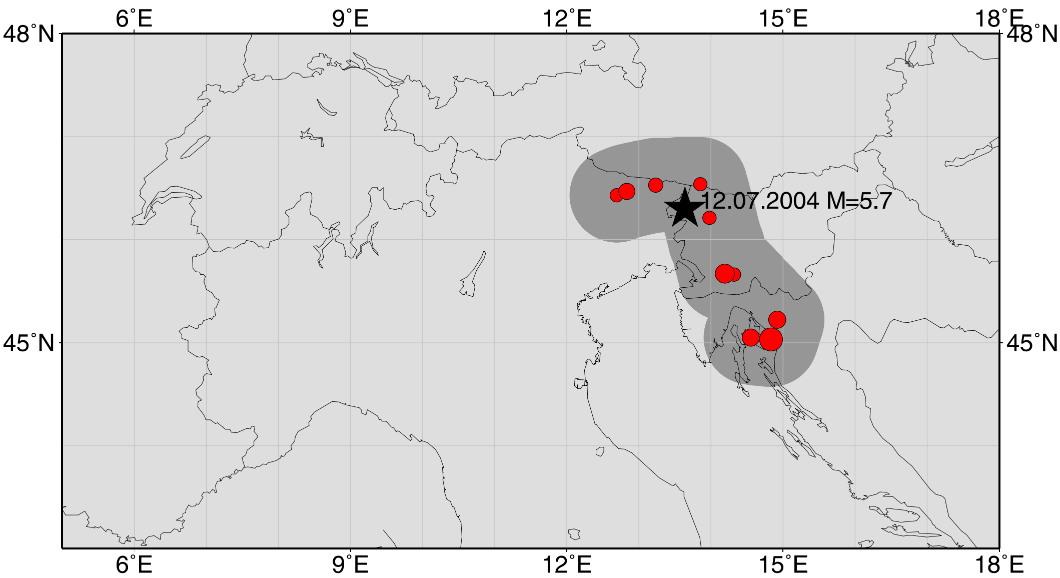
Case history, 2003
May 5: *Precursory chain* of earthquakes was formed. It indicates that an earthquake with magnitude 6.4 or more will occur in gray area within 9 months.
June 21: *Prediction was distributed* among relevant scientists and administrators.
Dec. 22: *San Simeon earthquake* (star).

False alarm in southern California



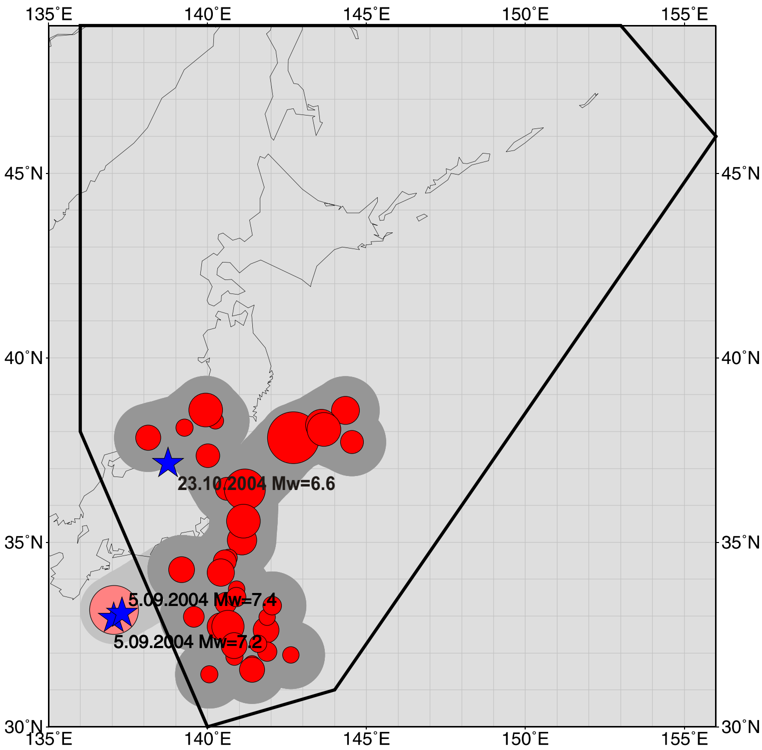
Case history
***November 29, 2003*:** *Precursory chain* of earthquakes was formed and kept growing until December 5, 2003. It indicates that an earthquake with magnitude 6.4 or more will occur in gray area by September 5, 2004.
December 24, 2003: *Prediction was distributed* among relevant scientists and administrators.
September 6, 2004: *Alarm terminates*. No such earthquake occurres.

Bovec earthquake, Slovenia, M_L=5.7



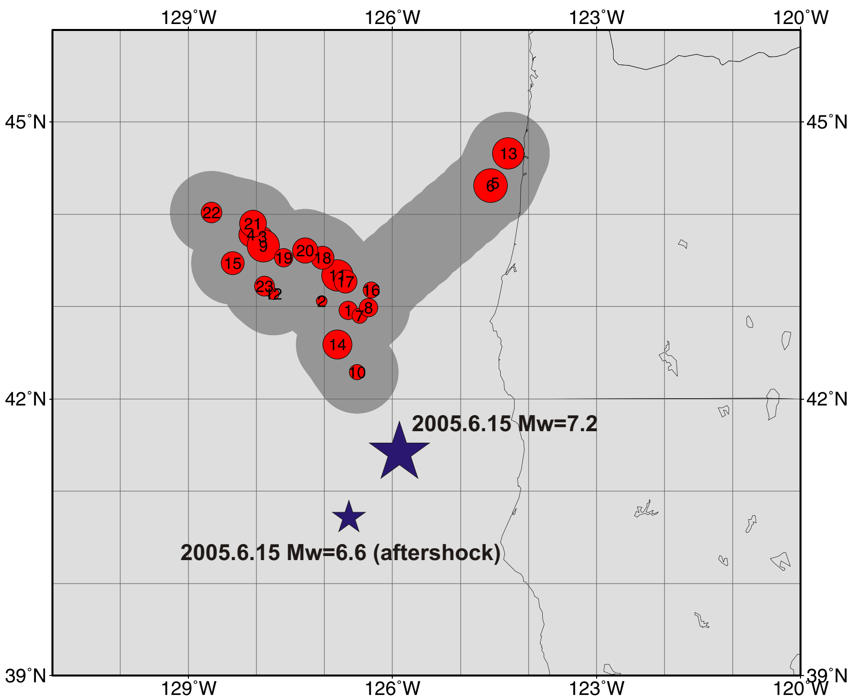
Case history, 2004
***February, 29*:** *Precursory chain* of earthquakes was formed. It indicates that an earthquake with magnitude $M_w \geq 5.5$ or more will occur in gray area by November 29, 2004.
***May, 12*:** *Prediction was distributed* among relevant scientists and administrators.
July, 12: Bovec earthquake, M_L=5.7(M_w=5.3) has occurred in the area of alarm.

Two earthquakes south to Honsu, M=7.2 and M=7.4. Near miss



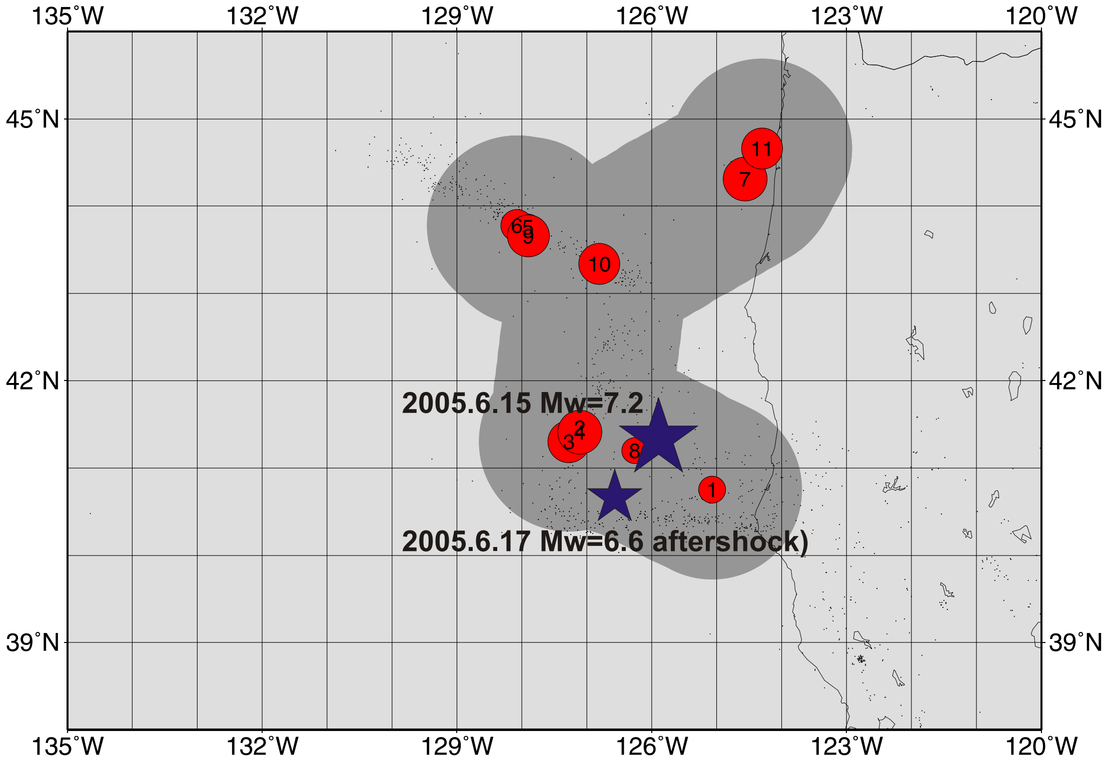
Case history, 2004
February, 8: *Precursory chain* of earthquakes was formed. It indicates that an earthquake with $M_w \geq 7.2$ will occur in gray area by November 8, 2004.
June, 1: *Prediction was distributed* among relevant scientists and administrators.
September, 1: Two earthquakes, M_w=7.2 and M_w=7.4 have occurred ***near*** the area of alarm. Successful prediction is not scored.

Gorda plate earthquakes, M=7.2 and M=6.6. Near miss

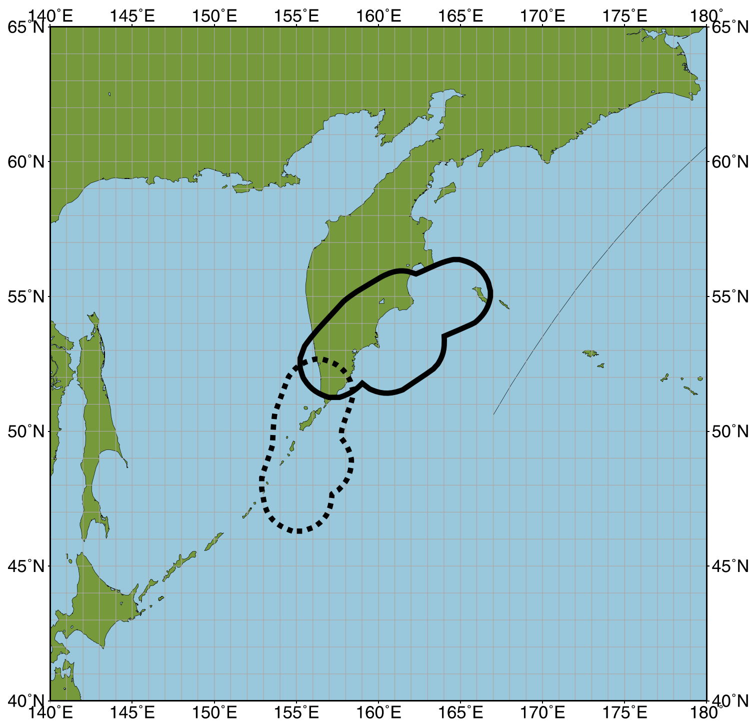


Case history
November, 16, 2004: *Precursory chain* of earthquakes was formed. It indicates that an earthquake with M≥6.4 will occur in gray area by August 16, 2005.
January, 29, 2005: *Prediction was distributed* among relevant scientists and administrators.
June, 15 and 17, 2005: Two earthquakes, M=7.2 and M=6.6 have occurred ***near*** the area of alarm. Successful prediction is not scored.

RTP applied to $M_w \geq 7.2$ in Northern Pacific
(Japan-Kurils-Kamchatka-Aleutians-Alaska-Canda-Oregon-California-N. Mexico).
Alarm until 19 August, 2005



RTP applied to $M_w \geq 7.2$ in Northern Pacific
(Japan-Kurils-Kamchatka-Aleutians-Alaska-Canda-Oregon-California-N. Mexico).



Current alarms. Solid line shows the area where an earthquake with magnitude $M_w \geq 7.2$ is predicted to occur until December 13, 2005 with probability 70%. Dashed line shows the area where an earthquake with magnitude $M_w \geq 7.2$ is predicted to occur until December 18, 2005 with probability 50%.

Functions for intermediate-term patterns in time and event scales

Time scale	Event scale
Rise of activity: *1. Activity*	
$F_U(t,s)=\dfrac{N_{t-s\le t_k<t}}{s}$	$\widetilde{F}_U(t_j,N)=\dfrac{N}{t_j-t_{j-N+1}}$
2. *Sigma*	
$F_\Sigma(t,s)=\log_{10}\sum\limits_{t-s\le t_k<t}10^{M_k-M^*}$	$\widetilde{F}_\Sigma(t_j,N)=\log_{10}\sum\limits_{k=j-N+1}^{j}10^{M_k-M^*}$
3. *Rise of magnitudes*	
$F_M(t,s)=\dfrac{\sum\limits_{t-s\le t_k<t}M_k}{N_{t-\frac{s}{2}\le t_k<t}}-\dfrac{\sum\limits_{t-2s\le t_k<t-\frac{s}{2}}M_k}{N_{t-s\le t_k<t-\frac{s}{2}}}$	$\widetilde{F}_M(t_j,N)=\dfrac{2}{[N/2]}\left(\sum\limits_{k=j-N+1}^{j-N+[N/2]}M_k-\sum\limits_{k=j-[N/2]+1}^{j}M_k\right)$
4. Acceleration	
$F_C(t,s)=\dfrac{\sum\limits_{t-\frac{s}{2}\le t_k<t}\dfrac{1}{t_k-t_{k-1}}}{N_{t-\frac{s}{2}\le t_k<t}}-\dfrac{\sum\limits_{t-s\le t_k<t-\frac{s}{2}}\dfrac{1}{t_k-t_{k-1}}}{N_{t-s\le t_k<t-\frac{s}{2}}}$	$\widetilde{F}_C(t_j,N)=\dfrac{1}{[N/2]}\left(\sum\limits_{k=j-N+1}^{j-N+[N/2]}\dfrac{1}{t_k-t_{k-1}}-\sum\limits_{k=j-[N/2]+1}^{j}\dfrac{1}{t_k-t_{k-1}}\right)$
Rise of clustering: *5. Swarm*	
$F_W(t,s)=1-\dfrac{A_r(t)}{\pi r^2 N_{t-s\le t_k<t}}$	$\widetilde{F}_W(t_j,N)=1-\dfrac{A_r(t_j)}{\pi r^2 N}$
6. B-micro	
$F_{B_\mu}(t,s)=\log_{10}\sum\limits_{t-s\le t_k<t}\sum\limits_{l}10^{M_{kl}-M^*}$	$\widetilde{F}_{B_\mu}(t_j,N)=\log_{10}\sum\limits_{k=j-N+1}^{j}\sum\limits_{l}10^{M_{kl}-M^*}$
Rise of earthquakes correlation range: *7. Accord*	
$F_A(t,N)=\widetilde{F}_A(t_j,N)=\dfrac{A_r(t)}{\pi r^2}$	
Transformation of Gutenberg-Richter relation: *8. Gamma*	
$F_G(t,N)=\widetilde{F}_G(t_j,N)=\dfrac{1}{N_{M_k\ge M_{1/2}}}\sum\limits_{M_k\ge M_{1/2}}(M_k-M^*)$	

Parameters of the functions representing intermediate-term precursors

Time scale

Set No.	R, km	s, months	n*
1	50	24	20
2	50	60	20
3	100	24	20
4	100	60	20

Event scale

Set No.	R, km	N	n*
1	50	20	10
2	50	50	10
3	100	20	20
4	100	50	20

For each set two values of T is used: 6 months and 24 months, in total this forms 8 sets for 8 functions.

Time and space window for aftershock determination

Magnitude of the main shock	Time window, days	Space window, km
$M < 2.5$	6	20
$2.5 \le M < 3.0$	11	23
$3.0 \le M < 3.5$	22	26
$3.5 \le M < 4.0$	42	30
$4.0 \le M < 4.5$	83	35
$4.5 \le M < 5.0$	155	40
$5.0 \le M < 5.5$	290	47
$5.5 \le M < 6.0$	615	54
$6.0 \le M < 6.5$	790	61
$6.5 \le M$	915	70

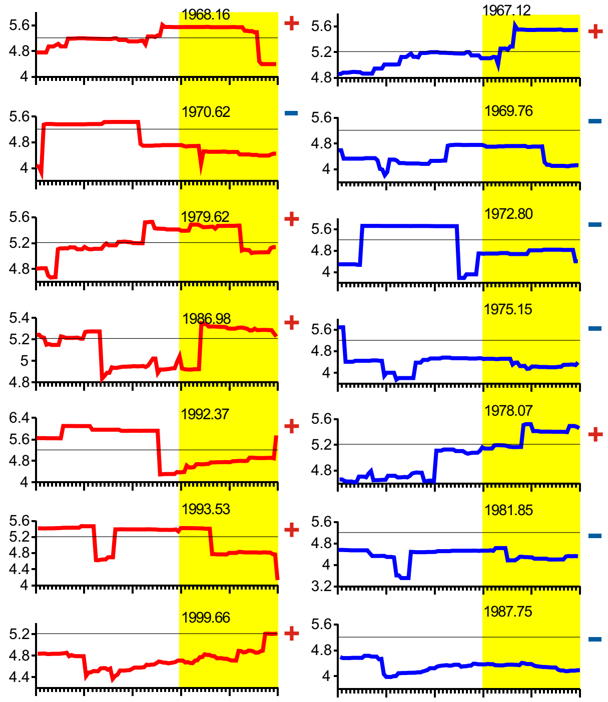
Magnitude of an aftershock is less or equal to the magnitude of main shock. If an event occurs within time-space window of an aftershock of some main shock, but outside the time-space window of that main shock, then the event is not considered as an aftershock, unless it is formally the direct aftershock of another main shock.

Pattern recognition: algorithm "Hamming distance"

Formally, the Hamming distance between two Boolean vectors of the same length is defined as the number of their non-coincident symbols. Here, the Hamming distance gives the number of emergent intermediate-term premonitory patterns.

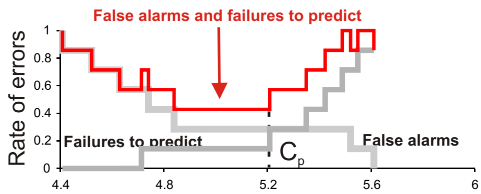
Emergence of a pattern at the moment *t* is captured by the condition $F_P(t) \ge C_P$. Each threshold C_P is determined automatically at the learning stage. It minimises the sum $n + f$; here *n* is the rate of failures to predict and *f* is the rate of false alarms in prediction with a single pattern *P*.

Rise of clustering: “b-micro”



$$b_\mu = \log_{10}\left(\sum_{k=j-N+1}^{j} \sum_l 10^{m_{kl}}\right) - m^*$$

m_{kl} - magnitudes of the aftershocks during first two days after the k-th main shock, l = 1, 2,….;

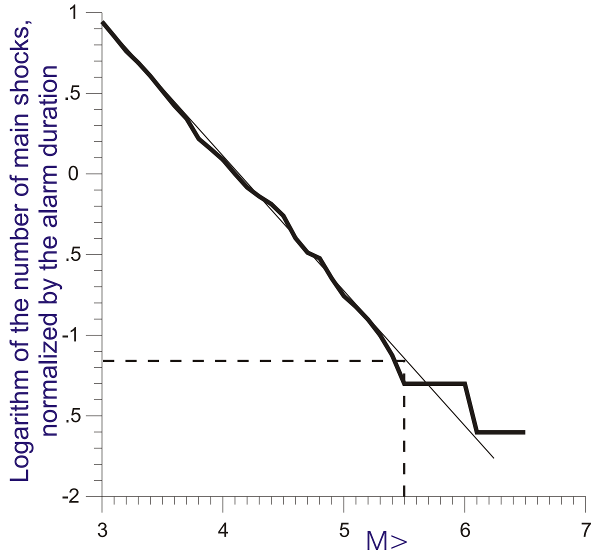


Definition of the earthquake prediction

[C. Allen et al., 1976]

- ·Time interval
- ·Spatial area
- ·Magnitude range
- ·Estimation of the probability of a random occurence of a target earthquake in time-space of alarm
- ·Estimation of the degree of reliability (probability that the alarm will happen to be false)
- ·Prediction should be documented to make possible the verification

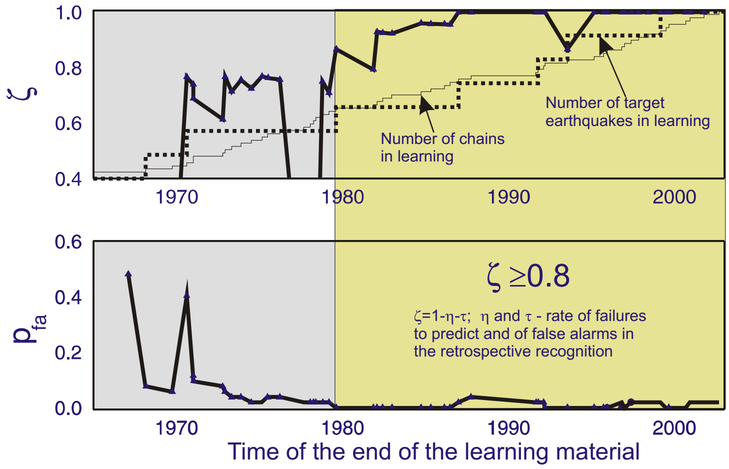
Probability that a target earthquake will occur in the time-space of alarm at random



This probability is estimated from the available statistics of large earthquakes in the area of alarm using cumulative magnitude-frequency graph

Probability p_{fa} that the alarm will happen to be false

Test “seismic hisrtory”

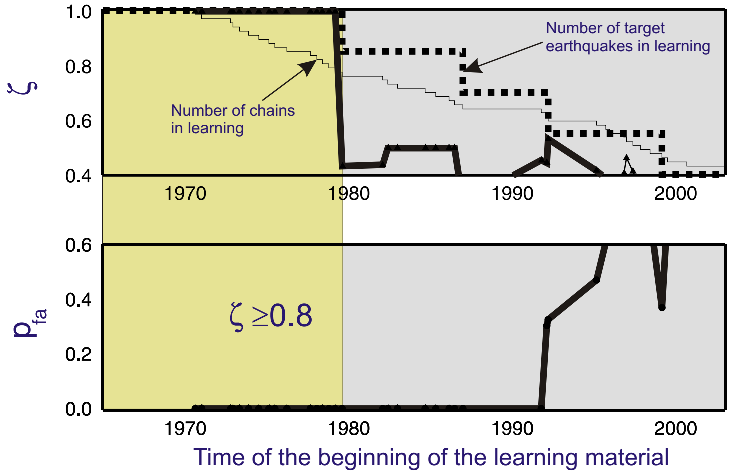


Probability p_{fa} is estimated as the ratio of the number of false alarms with pattern recognition using intermediate-term premonitory patterns equal or better than that of the tested chain to the total number of such chains.

The number of chains in the learning set is increased one by one. Individual estimates of p_{fa} are then used in the global distribution.

“Seismic history” is repeated in two groups of tests:
24 tests with variation of parameters of the chains and of the aftershocks and 32 tests with change of pattern recognition elements.

Test “backward hisrtory”

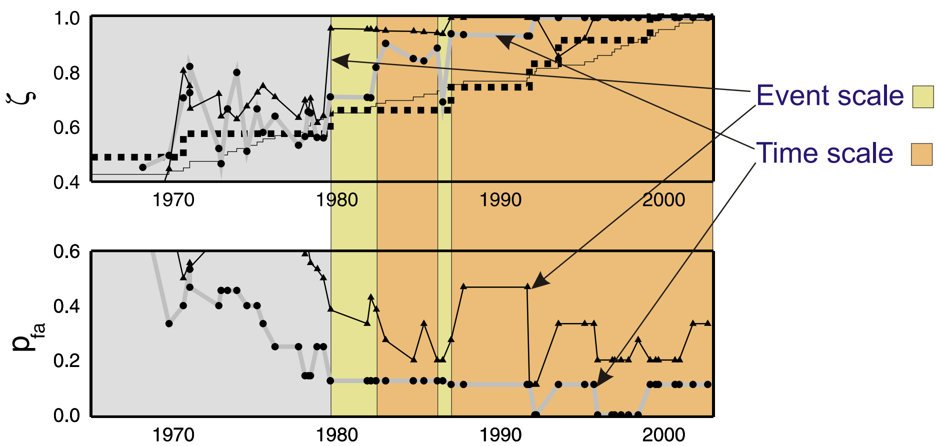


The number of chains in the learning set is increased one by one starting the latest chains (possible current alarms excluded)..

“Backward history” is repeated in both groups of tests; individual estimates of p_{fa} (at condition $\zeta \geq 0.8$) are also used in the global distributions.

Functions in *time* scale and in *event* scale

Test “seismic hisrtory”



Individual estimates of p_{fa} (at condition $\zeta \geq 0.8$) with functions both in time and event scales are used in the global distributions.

First group of tests: variations at the Step 1(Chains)

I. Modification of the catalogue used to detect chains (parameters τ_0 and k_0 are readjusted)

1. Aftershocks are not eliminated at all.
2. Spatial windows to eliminate aftershocks are decreased by factor 2/3.
3. Spatial windows to eliminate aftershocks are decreased by factor 1/2.
4. Spatial windows to eliminate aftershocks are increased by factor 3/2.
5. Temporal windows to eliminate aftershocks are decreased by factor 2/3.
6. Temporal windows to eliminate aftershocks are decreased by factor 1/2.
7. Temporal windows to eliminate aftershocks are increased by factor 3/2.
8. Both spatial and temporal windows to eliminate aftershocks are decreased by factor 2/3.
9. Both spatial and temporal windows to eliminate aftershocks are decreased by factor 1/2.
10. Both spatial and temporal windows to eliminate aftershocks are increased by factor 3/2.

II. Variation of the parameters of the chains

11. c=0.2. Parameters r_0 and k_0 are readjusted.
12. c=0.3. Parameters r_0 and k_0 are readjusted.
13. c=0.4. Parameters r_0 and k_0 are readjusted.
14. c=0.5. Parameters r_0 and k_0 are readjusted.
15. τ_0 is decrease by 10%. Parameter k_0 is readjusted.
16. τ_0 is increase by 10%. Parameter k_0 is readjusted.
17. r_0 is decrease by 10%. Parameter k_0 is readjusted.
18. r_0 is increase by 10%. Parameter k_0 is readjusted.
19. k_0 is decrease by 10%.
20. k_0 is increase by 10%.
21. l_0 is decrease by 10%.
22. l_0 is increase by 10%.
23. M_{min} is decrease by 0.1. Parameter k_0 is readjusted.
24. M_{min} is increase by 0.1. Parameter k_0 is readjusted.

2-nd group of tests: variations at the Step 2 (pattern recognition)

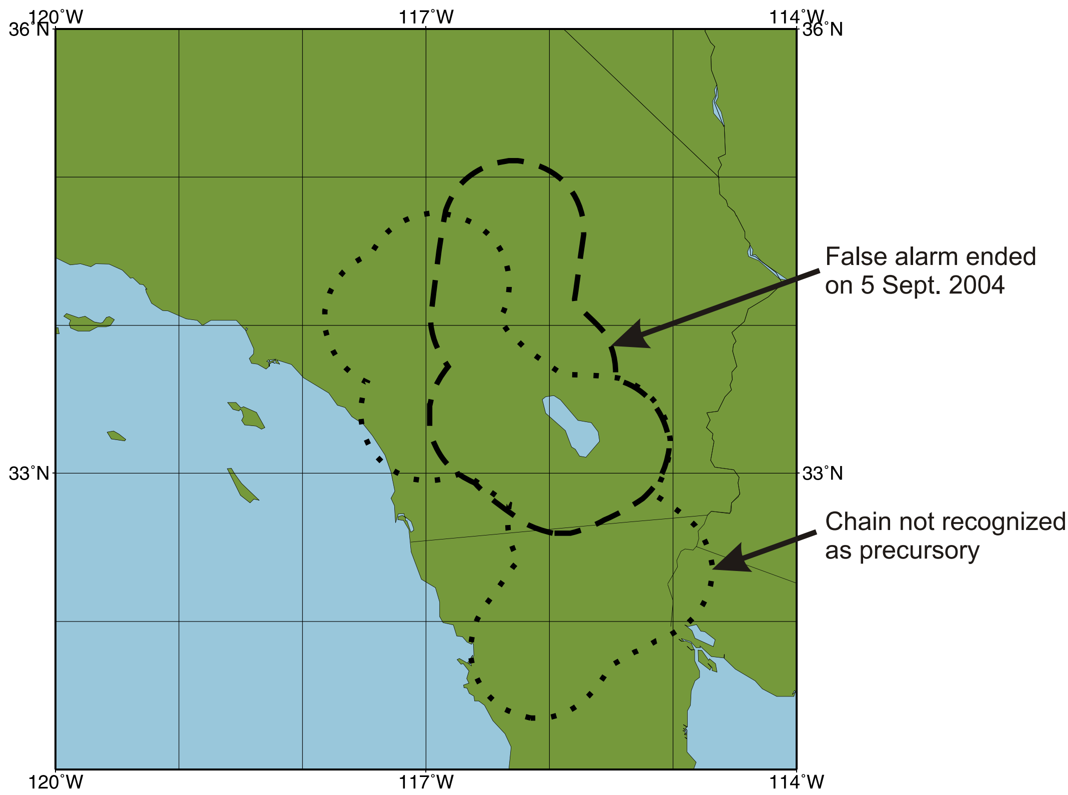
I. Variants of the catalogue of main shocks are used to calculate functions

1. Main variant: historic experiment and backward experiment in time and event scales.
2. Spatial windows to eliminate aftershocks are decreased by factor 2/3.
3. Spatial windows to eliminate aftershocks are decreased by factor 1/2.
4. Spatial windows to eliminate aftershocks are increased by factor 3/2.
5. Temporal windows to eliminate aftershocks are decreased by factor 2/3.
6. Temporal windows to eliminate aftershocks are decreased by factor 1/2.
7. Temporal windows to eliminate aftershocks are increased by factor 3/2.
8. Both spatial and temporal windows to eliminate aftershocks are decreased by factor 2/3.
9. Both spatial and temporal windows to eliminate aftershocks are decreased by factor 1/2.
10. Both spatial and temporal windows to eliminate aftershocks are increased by factor 3/2.

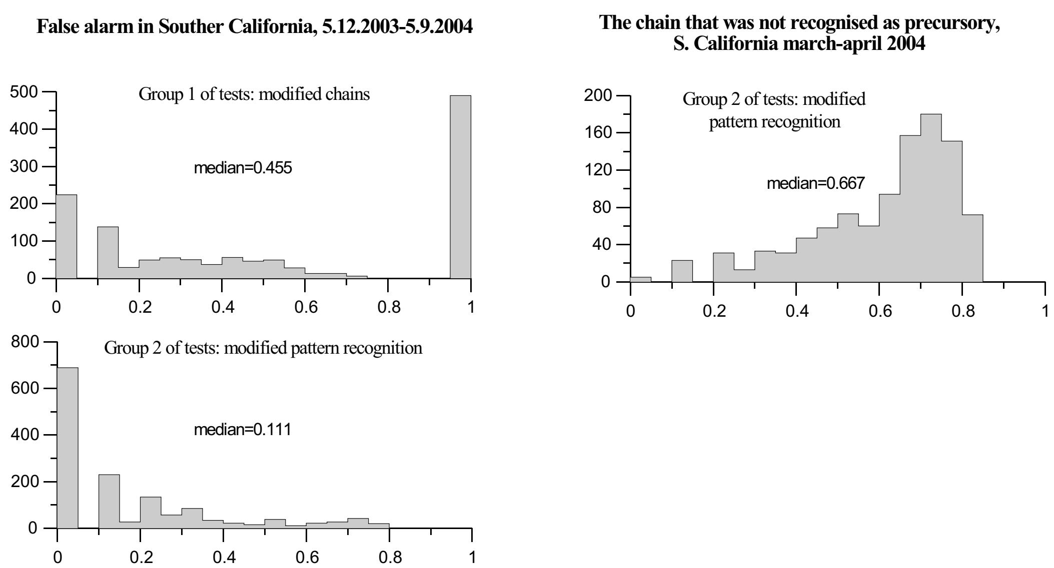
II. Variants of pattern recognition rule

11. Only cases ($n_P + f_P < 0.8$) in the learning are taken into account.
12. Only cases ($n_P + f_P < 0.7$) in the learning are taken into account.
13. Only cases ($n_P + f_P < 0.6$) in the learning are taken into account.
14. 8 cases with best ($n_P + f_P$) in the learning are taken into account.
15. 16 cases with best ($n_P + f_P$) in the learning are taken into account.
16. 32 cases with best ($n_P + f_P$) in the learning are taken into account.
17. "Hard C_P" in the learning false alarms are given more weight to determine C_P.
18. "Soft C_P" in the learning failures to predict are given more weight to determine C_P.
19. 32 cases: R=50 km (100 km excluded).
20. 32 cases: R=100 km (50 km excluded).
21. 32 cases: s=24 months (60 months excluded) in time scale; N=20 (50 excluded) in event scale.
22. 32 cases: s=60 months (24 months excluded) in time scale; N=50 (20 excluded) in event scale.
23. 32 cases: T=6 months (24 months excluded).
24. 32 cases: T=24 months (6 months excluded).

25-32. 56 cases: one function of 8 excluded.



Distribution of the individual estimates p_{fa}



Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics of the Earth and Planetary Interiors 145 (2004) 75–85
PHYSICS OF THE EARTH AND PLANETARY INTERIORS
www.elsevier.com/locate/pepi

Reverse tracing of short-term earthquake precursors

V. Keilis-Borok [a,b,*], P. Shebalin [a], A. Gabrielov [c], D. Turcotte [d]

[a] *International Institute for Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy Science, Warshavskoe sh. 79, Korp. 2, Moscow 113556, Russia*
[b] *Department of Earth and Space Sciences and Institute of Geophysics and Planetary Physics, UCLA, Los Angeles, CA 90095-1567, USA*
[c] *Departments of Mathematics and Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN 47907-1395, USA*
[d] *Department of Geology, University of California, Davis, CA 95616, USA*

Received 15 December 2003; accepted 23 February 2004

Abstract

We introduce a new approach to short-term earthquake prediction named "*Reverse Tracing of Precursors*" (RTP), since it considers precursors in reverse order of their appearance. First, we detect the "candidates" for the short-term precursors; in our case, these are newly introduced chains of earthquakes reflecting the rise of an earthquake correlation range. Then we consider each chain, one by one, checking whether it was preceded by an intermediate-term precursor in its vicinity. If *yes*, we regard this chain as a precursor; in prediction it would start a short-term alarm. The chain indicates the narrow area of possibly complex shape, where an intermediate-term precursor should be looked for. This makes possible to detect precursors undetectable by the direct analysis.

RTP can best be described on an example of its application; we describe retrospective prediction of two prominent Californian earthquakes—Landers (1992), $M = 7.6$, and Hector Mine (1999), $M = 7.3$, and suggest a hypothetical prediction algorithm. This paper descripes the RTP methodology, which has potentially important applications to many other data and to prediction of other critical phenomena besides earthquakes. In particular, it might vindicate some short-term precursors, previously rejected as giving too many false alarms.

Validation of the algorithm per se requires its application in different regions with a substantial number of strong earthquakes. First (and positive) retrospective results are obtained for 21 more strong earthquakes in California ($M \geq 6.4$), Japan ($M \geq 7.0$) and the Eastern Mediterranean ($M \geq 6.5$); these results are described elsewhere. The final validation requires, as always, prediction in advance for which this study sets up a base. We have the first case of a precursory chain reported in advance of a subsequent strong earthquake (Tokachi-oki, Japan, 25 September 2003, $M = 8.1$).

Possible mechanisms underlying RTP are outlined.

Keywords: Short-term earthquake prediction; Long-range correlation; Earthquake chains; Precursors; Instability

1. Introduction

1.1. Generation of strong earthquakes—a non-localized process

Seismicity is commonly recognized as a part of the geodynamics (Aki, 2003; Bird, 1998; Keilis-Borok,

* Corresponding author.
E-mail addresses: vkb@ess.ucla.edu (V. Keilis-Borok), shebalin@mitp.ru (P. Shebalin).

0031-9201/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.pepi.2004.02.010

1990; King et al., 2002; Press, 1965; Rundquist and Soloviev, 1999; Scholz, 1990); in seismically active areas the earthquakes accommodate a considerable fraction of tectonic development of the lithosphere. That development goes on in multiple time-, space-, and energy-scales and preparation of strong earthquakes is not an exception. Accordingly, while the target of earthquake prediction—a strong earthquake—is a localized event, the process of its generation is not localized. Strictly speaking, its time scales range from geological to seconds in time, and spatial scales—from global to microscopic (Turcotte, 1997; Keilis-Borok, 1990; Gabrielov et al., 1999); however, in prediction research a truncated scaling is usually considered: from tens of years to days, and from hundreds of kilometers to kilometer.

This multiplicity of scales is reflected in the general concept of the seismically active lithosphere as a hierarchical dissipative non-linear system, persistently self-organizing from time to time into the critical phenomena—the strong earthquakes (Blanter and Shnirman, 1997; Bowman et al., 1998;Gabrielov et al., 1994, 2000; Jaume and Sykes, 1999; Keilis-Borok, 1990; Rundle et al., 2000; Sornette, 2000; Turcotte, 1997; Zaliapin et al., 2002a). Among manifestations of that selforganization are premonitory seismicity patterns—the spatio-temporal patterns of seismicity emerging as a strong earthquake approaches (Aki, 2003; Buffe and Varnes, 1993; Caputo et al., 1983; Gabrielov and Newman, 1994; Jin et al., 2003; Keilis-Borok, 1990, 1996, 2000; Keilis-Borok et al., 1990a,b, 1964, 1999, 2002; Knopoff et al., 1996; Kossobokov et al., 1995, 2003; Ma et al., 1990; Mogi, 1985; Newman et al., 1995; Novikova et al., 2002; Press, 1965; Press and Allen, 1995; Romanowicz, 1993; Rotwain and Novikova, 1999; Shebalin et al., 2000; Turcotte, 1997; Zaliapin et al., 2002a,b, 2003b; Zöller et al., 2001). A multitude of such patterns have been reported in rather different scales. Systematically tested are the intermediate-term patterns (with characteristic lead time of years). Here, we suggest a method to detect the short-term patterns, which have the lead time of months.

1.2. Reverse Tracing of Precursors (RTP)

We consider the short-term patterns in conjunction with intermediate-term ones. This is done by RTP analysis, in which these patterns are detected in the reverse order of their appearance: short-term patterns are analyzed first, although they emerge later. Our findings can best be described on a specific example of data analysis.

1.3. Region and data

We describe detection of short-term patterns before two prominent Californian earthquakes—Landers (1992), $M = 7.6$, and Hector Mine (1999), $M = 7.3$. These are the largest Californian earthquakes since 1965—the period, when the earthquake catalog is sufficiently complete for our analysis. Territory considered is shown in Fig. 1. The earthquake catalog is taken from (ANSS/CNSS and NEIC).

2. Chains

Our point of departure is provided by the short-term patterns *Roc* and *Accord* capturing a premonitory increase in earthquake correlation range. They were found first in models (Gabrielov et al., 2000) and then in observations (Keilis-Borok et al., 2002; Shebalin et al., 2000; Novikova et al., 2002). Other patterns capturing that phenomenon are suggested in Zöller et al. (2001) and Zaliapin et al. (2002b). Here, we introduce the pattern *chain* which is a generalization of *Roc* and *Accord*. Qualitatively, a chain is a rapidly extended sequence of small earthquakes that follow each other closely in time and space.

2.1. Definitions

2.1.1. Earthquake catalog

As in most premonitory patterns of that family (Keilis-Borok, 1996; Kossobokov and Shebalin, 2003) aftershocks are eliminated from the catalog; however, an integral measure of aftershocks sequence b is retained for each remaining earthquake (main shocks and foreshocks). We use a common representation of the earthquake catalog $\{t_j, \varphi_j, \lambda_j, h_j, m_j, b_j\}$, $j = 1, 2, \ldots$. Here, t_j is the time of an earthquake, $t_j \geq t_{j-1}$; φ_j and λ_j, latitude and longitude of its epicenter; h_j, focal depth; and m_j, magnitude. We consider the earthquakes with magnitude $m = m_{\min}$. Focal depth is not used in this study.

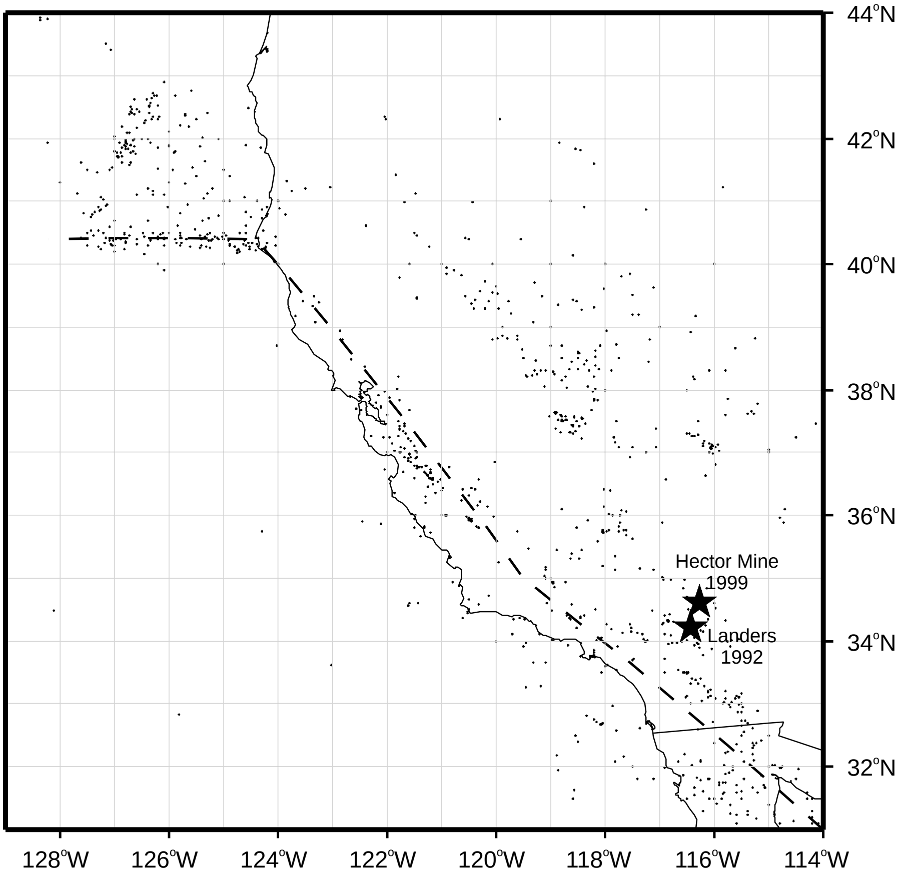


Fig. 1. Territory considered. Stars mark large earthquakes, targeted for prediction. Dots show background seismicity for the time considered (1965–2003): epicenters of earthquakes with magnitude $m \geq 3$ with aftershocks eliminated. Dashed line is used for time–distance projection of epicenters (Fig. 3 below).

2.1.2. Chain

Let us call two earthquakes "*neighbors*" if: (i) their epicenters lie within a distance r; and (ii) the time interval between them does not exceed a threshold τ_0. A chain is the sequence of earthquakes connected by the following rule: *each earthquake has at least one neighbor in that sequence*; *and does not have neighbors outside the sequence*. The threshold r is normalized by the average distance between the earthquakes with lowest magnitude m in a pair considered. We use a coarse normalization $r = r_0 10^{cm}$, c being a numerical parameter.

Let k be the number of earthquakes thus connected and l—the maximal distance between their epicenters. We look for precursors only among the chains with $k \geq k_0$ and $l \geq l_0$. These thresholds ensure that our chains are exceptional phenomena.

2.1.3. Chain's vicinity

To compare location of a chain with locations of strong earthquakes we consider its R-vicinity formally defined as the union of circles of the radius R centered at the epicenters of the chains forming the chain. To smooth the borders of that area we add the dense sequence of circles along the lines connecting each epicenter in the chain with the two closest ones. The envelope of all the circles is the border of R-vicinity of the chain; it is similar to the "Wiener sausage", introduced by N. Wiener in the theory of probability.

Table 1
Parameters for detecting the chains

$m_{\min}$	r_0 (km)	c	τ_0 (days)	k_0	l_0 (km)	R (km)
3.3	50	0.35	20	8	350	75

Notations are given in the text, Section 2.1.

2.2. Data analysis

We detected the chains defined above using numerical parameters listed in Table 1. Aftershocks have been identified by a coarse windowing, as described in (Keilis-Borok et al., 2002). The remaining catalog contains 3940 earthquakes. We have found among them nine chains, altogether containing 116 earthquakes: this shows that our chains are indeed exceptional phenomena. Maps of the chains are shown in Fig. 2; shaded areas are their vicinities, defined above. Vital characteristics of each chain are given in Table 2. Fig. 3 juxtaposes the chains and strong earthquakes on the time–distance plane; distance is counted along the dashed line shown in Figs. 1 and 2.

As we see in Fig. 2 (two panels in the bottom row) and Fig. 3, only the two last chains (#8 and #9) might be regarded as the local short-time precursors to the Landers and Hector Mine earthquakes: short-term—because they emerge with the short-term lead times (respectively, 1.7 and 4.6 months); and local—because the target earthquakes occur in their vicinities. However, the other seven chains, if used as precursors, would give false alarms. To reduce their number we introduce the RTP analysis.

3. Precursory chains

3.1. Hypothesis

We hypothesize that *a precursory chain* (*as opposed to a chain giving a false alarm*) *is preceded by the local intermediate-term precursors formed in the chain's R-vicinity*. This vicinity is not known, until the chain is formed, and its shape might be rather complicated (see Fig. 2). To overcome that impasse we introduce the two-step RTP analysis schematically illustrated in Fig. 4.

(i) *Search for the chains and determination of their R vicinities* (Section 2). Each chain is regarded as a "candidate" for a short-term precursor.

(ii) *Search for the local intermediate-term patterns in the R-vicinities of each chain*. They are looked for within T years before the chain; T is an adjustable numerical parameter. If (and only if) such patterns are detected, we regard this chain as a short-term precursor; in prediction it would start a short-term alarm.

To complete that description we have to specify intermediate-term patterns used at the second step.

3.2. Definitions

We use the *pattern* Σ which reflects premonitory rise of seismic activity. This pattern, introduced in Keilis-Borok and Malinovskaya (1964), is successfully used in different prediction algorithms, alone or in combination with other patterns (Keilis-Borok, 1990, 1996, 2000; Keilis-Borok et al., 1999, 2002; Kossobokov et al., 1995, 2003; Rotwain and Novikova, 1999). It is defined as a premonitory increase of the total area of the earthquake sources. Emergence of this pattern is captured by the function $\Sigma(t)$ defined in a sliding time-window (Keilis-Borok and Malinovskaya, 1964):

$$\sum(t/s,\, B) = \sum 10^{Bm_i},\; m_i \ge m_{\min}; \quad t - s < t_i \le t$$

Summation is taken over all main shocks within the time window $(t-s,\ t)$ in the R-vicinity of the chain. We take $B \sim 1$, so that the sum is coarsely proportional to the total area of the fault breaks in the earthquakes' sources (Keilis-Borok, 2002); with $B = 0$ this sum is the number of earthquakes, with $B = 3/2$ it is proportional to their total energy. The emergence of pattern Σ is identified by condition $\Sigma(t) \ge \Sigma_0$; this threshold depends on the magnitude of target earthquakes. In previous applications cited above pattern Σ was used as non-local one. We renormalize its numerical parameters to make it local.

3.3. Data analysis

We detected precursory chains and determined their R-vicinities (Section 2). In each vicinity we computed the function $\Sigma(t)$ within time interval $T = 5$ years and summation interval $s = 6$ months. We identified as precursory three chains preceded by largest peaks of

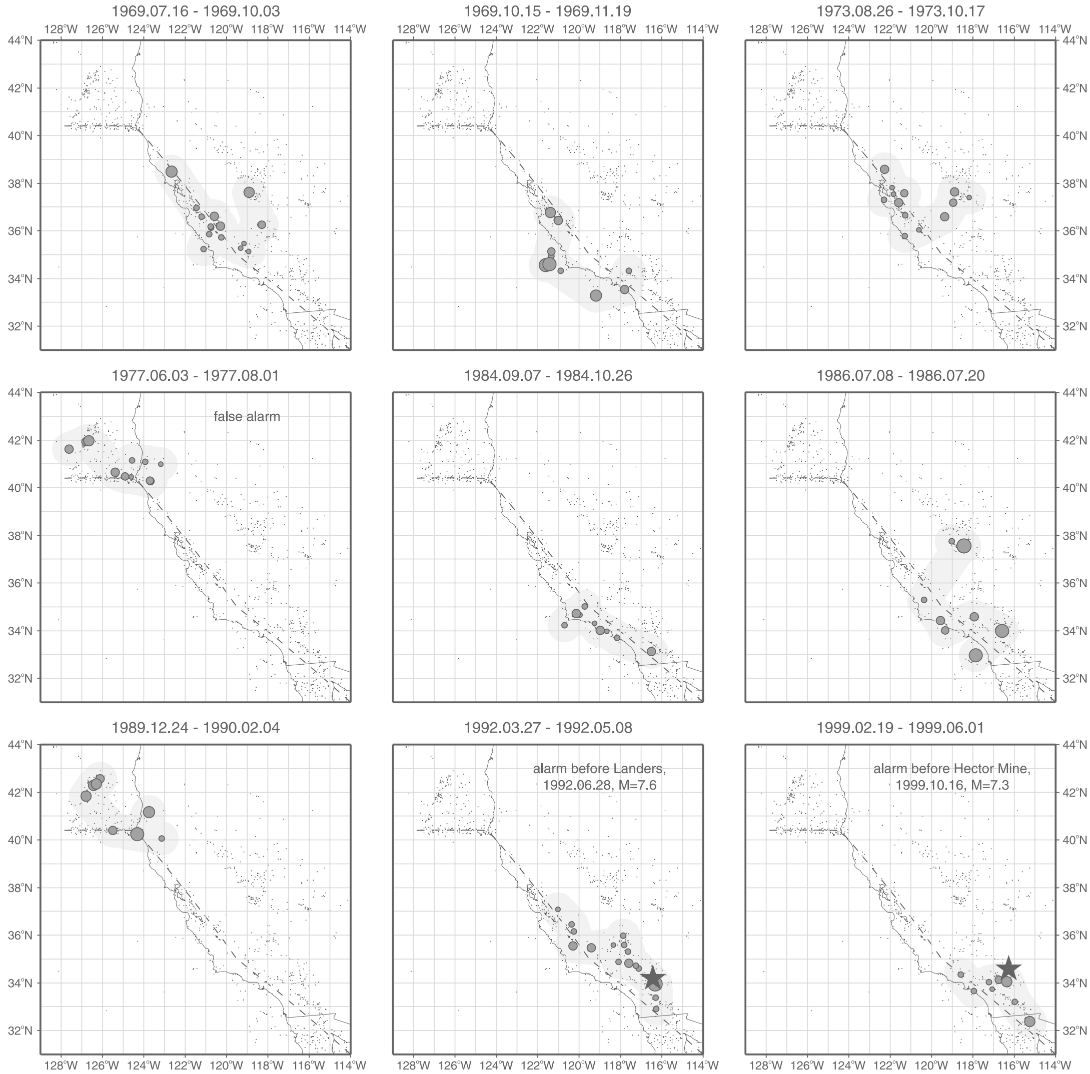


Fig. 2. Maps of the chains. Detected chains are shown in separate boxes. Circles show epicenters of earthquakes in a chain; their size is proportional to magnitude. The shadowed areas show *R*-vicinities of the chains. Dates of the beginning and the end of a chain are given at the top of each box. Three chains (1977, 1992, and 1999) shown in bold are identified as precursory ones. The first chain gives a false alarm; two other chains are followed within few months by target earthquakes, Landers and Hector Mine. Other notations are the same as in Fig. 1.

$\Sigma(t)$; they can be recognized with the threshold $\Sigma_0 = 10^{6.7}$. Table 2 shows these chains in bold. As we see, identification of the first chain, in 1977, is wrong; in prediction it would give a false alarm. Identification of two other chains, in 1992 and 1999, is correct; each is followed by a target earthquake within few months. The same chains would the selected with the tenfold smaller time interval, $T = 6$ months. The corresponding threshold is $\Sigma_0 = 10^{5.4}$; it is smaller since smaller number of earthquakes is included in summation.

3.4. Hypothetical prediction algorithm

It remains to define alarms triggered by a precursory chain. This is a final step in transition from a precursor to algorithmic prediction. We adapt the standard

Table 2
Characteristics of the chains

#	Start	End	Duration (days)	Lead time (months)	Distance from a strong earthquake (km)	Number of earthquakes, k	Maximal distance, l (km)	Largest magnitude	Area of the R-vicinity, $\times 10^3$ (km^2)
28.06.1992: Landers earthquake, $M = 7.6$									
1	16.07.1969	03.10.1969	80			17	499	5.3	150
2	15.10.1969	19.11.1969	35			12	485	5.6	113
3	26.08.1973	17.10.1973	53			13	381	4.5	150
4	**03.06.1977**	**01.08.1977**	**60**			11	377	**4.7**	**104**
5	07.09.1984	26.10.1984	49			9	408	4.6	90
6	08.07.1986	20.07.1986	12			10	543	5.9	122
7	24.12.1989	04.02.1990	41			8	373	5.7	101
8	**27.03.1992**	**08.05.1992**	**42**	**1.7**	**29**	17	635	**6.1**	**161**
16.10.1999: Hector Mine earthquake, $M = 7.4$									
9	**19.02.1999**	**01.06.1999**	**102**	**4.6**	**60**	11	380	**4.9**	**98**

Chains recognized as "precursory" by RTP analysis (Section 3) are shown in bold. Chain #4 would trigger in prediction a false alarm, Chains #8 and #9 would trigger correct alarms.

general scheme of prediction algorithms, widely used in intermediate-term earthquakes prediction and many other problems (Keilis-Borok, 2002; Kossobokov and Carlson, 1995, and references therein).

(i) Prediction is targeted at the main shocks with magnitude M or more; usually the magnitude intervals (M, $M + 1$) are considered separately.
(ii) When a precursory chain is detected, a short-term alarm is triggered. It predicts a target earthquake in R-vicinity of the chain, within time interval (t_e, $t_e + \tau$); here t_e is the moment when chain emerged, τ a numerical parameter (duration of alarm). Results of the data analysis suggest to take $\tau = 6$ months.

Possible outcomes of such prediction are illustrated in Fig. 5. Probabilistic component of prediction is represented by the total time–space covered by alarms and probabilities of false alarms and failures to predict (Molchan, 2003).

4. Discussion

4.1. Summary

This paper introduces RTP analysis in the evaluation of seismicity, culminated by a strong earthquake. Precursors with different lead times are considered in reverse order of their appearance. First, we detect the candidates for short-term precursors; in our case, those are the chains of small earthquakes capturing the rise of earthquake correlation range. A chain determines its narrow vicinity where we look for the local intermediate-term precursor(s), pattern Σ in our case. Its presence in turn indicates the precursory chains. We describe RTP on an example: detecting precursory chains months before two prominent California earthquakes, Landers (1992) and Hector Mine (1999), well isolated in time and space from other comparable earthquakes in that region.

4.2. Methodological advantage of RTP

The opposite (direct) analysis would start with tracing of the intermediate-term patterns hidden in the background seismicity. Almost all of them, known so far, are not local, pattern Σ included. They emerge in the areas whose linear size is up to 10 times larger than the source of the incipient target earthquake (Bowman et al., 1998; Keilis-Borok and Malinovskaya, 1964; Keilis-Borok and Soloviev, 2003); some patterns—even up to 100 times larger (Press and Allen, 1995; Romanowicz, 1993). We have found pattern Σ that became local after renormalization: it emerges in the same narrow area (R-vicinity of the chain), where epicenter of a target earthquake lies. As we see in Fig. 2, the shape of that area might be rather complex, and its size—diverse. To find this area by trying different shapes, sizes, and locations is

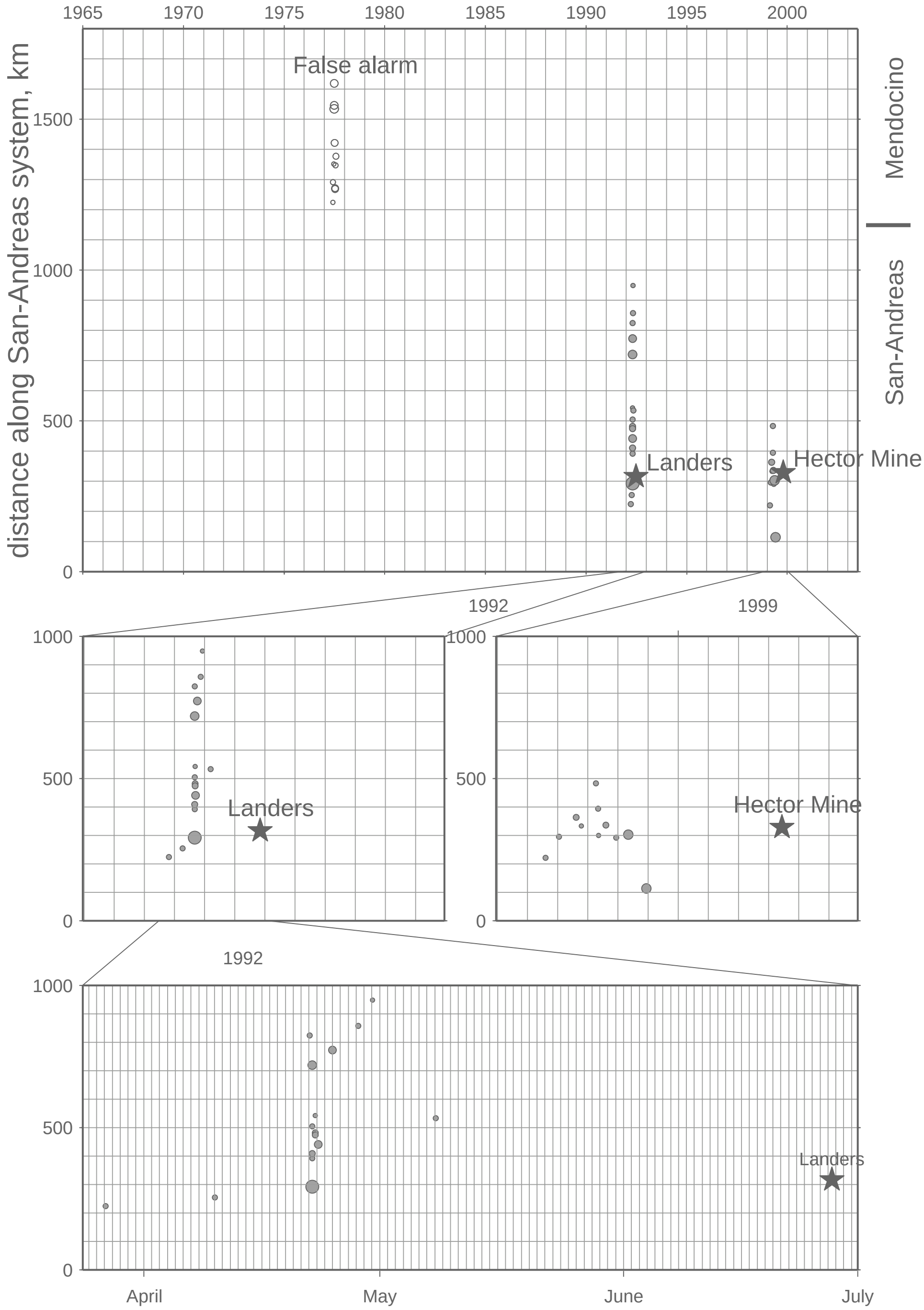


Fig. 3. Chains and strong earthquakes on the time–distance plain. Distance is counted along the dashed line shown in Fig. 1. Filled and open circles show the chains identified, respectively, as precursory and non-precursory. Other notations are the same as in Fig. 1.

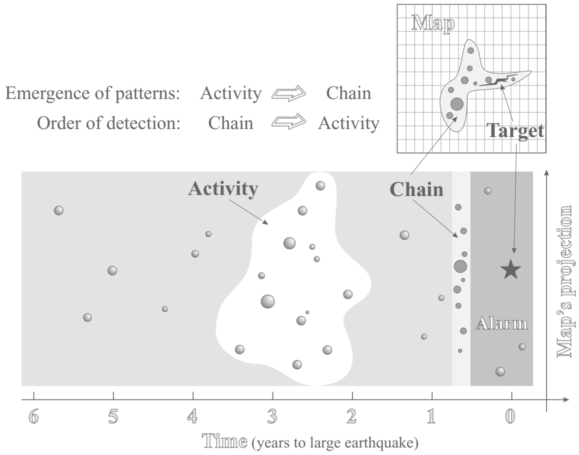


Fig. 4. Schematic illustration of the *Reverse Tracing of Precursors* (RTP). (Top) Map showing precursory chain and the source of the target earthquake (black). (Bottom) Scheme of analysis in time–space projection. Circles show epicenters forming the chain (dark gray) and preceding it (light gray). The "R-vicinity" of the chain is shown in light gray. Star is projection of the epicenter of the target earthquake. The gray rectangle before the chain shows the time–space where rise of activity (pattern Σ) is looked for. White area shows the time–space where this pattern was found; its presence indicates a precursory chain. The chain is detected first, although it emerges after the pattern Σ. Note how a narrow chain determines a much larger time interval where a pattern Σ is looked for. Dark gray area shows the time–space covered by an alarm: within 6 months after precursory chain a target earthquake is expected in its R-vicinity.

not realistic. Reverse analysis resolves this impasse, indicating a limited number of chains to consider.

4.3. Physical interpretation

RTP seems to be a promising general approach to prediction of critical phenomena in complex systems: it identifies a rare small-scale phenomenon that carries a memory of the larger scale history of the system. At the same time, this approach has a natural earth-specific explanation: it follows from the concept that strong earthquake is a result of a lasting large-scale process whose different stages involve different parts of the fault network. Earthquakes in the chain mark the part of the fault network that has started to move in unison months before a target earthquake. Pattern Σ indicates that this synchronization started much earlier, albeit expressed in a more subtle form. A similar step-by-step escalation of instability was observed in direct analysis: by algorithms M8&MSc (Kossobokov and Shebalin, 2003), and by some other algorithms (Aki, 2003; Shebalin et al., 2000; Keilis-Borok and Soloviev, 2003).

Both the chains and the peaks of Σ are sporadic short-lived phenomena not necessarily reflecting the steady trends of seismicity. This is typical for all premonitory patterns of that family (Keilis-Borok, 2002; Kossobokov and Shebalin, 2003). Probably, both patterns are the symptoms but not causes of a strong earthquake: they signal its approach but do not trigger it. Similarly sporadic are many observed precursors to other critical phenomena, e.g. economic recessions (Keilis-Borok et al., 2000).

4.4. Implications for earthquake prediction

- We have applied RTP analysis to target earthquakes of more diverse magnitudes in California and two other regions, Japan and E. Mediterranean, normalizing the parameters of the algorithm and considering all known (eight) major types of intermediate-term patterns (Keilis-Borok and Soloviev, 2003). We have first two earthquakes predicted in advance: Tokachi-oki earthquake in Northern Japan (M8.1, 25 September 2003) and San Simeon in Central California, M6.5, 22 December 2003). The results, highly encouraging, are described in Shebalin et al., in press.
- It seems natural to apply the RTP analysis to earthquake precursors, expressed in other fields. First positive results are obtained with precursors gauging interaction between the ductile and brittle layers of the crust (Aki, 2003; Jin et al., 2003; Zaliapin et al., 2003a). Other promising applications include electromagnetic fields (Uyeda and Park, 2002), fluid regime (Keilis-Borok, 1990; Ma et al., 1990), GPS, InSAR, etc.
- We detect intermediate-term patterns only after a chain has emerged so that its vicinity can be determined; this is too late to declare an intermediate-term alarm. Accordingly, our results concern only short-term prediction.
- "Pre-chain" precursors might emerge with a short lead time too.
- There are no reasons not to explore RTP analysis for prediction of different critical phenomena

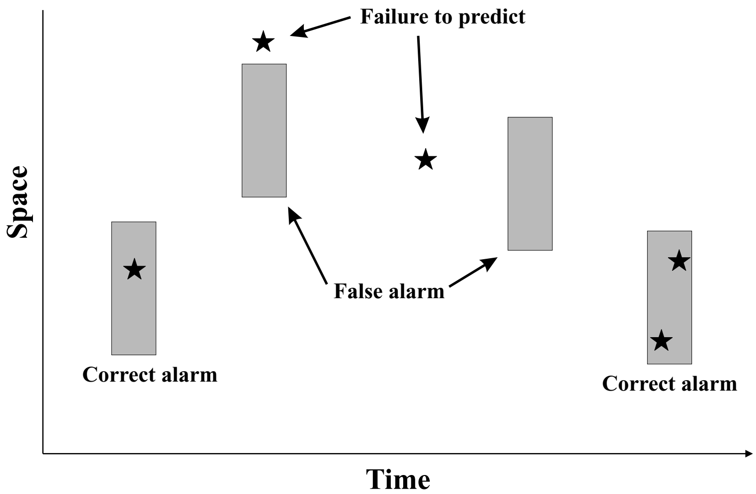


Fig. 5. Possible outcomes of prediction. Stars mark epicenters of strong earthquakes, targeted by prediction. A box to the right of the chain (dark gray) is the time–space covered by an alarm. A prediction is correct if a strong earthquake occurs within an alarm. Otherwise, this is a false alarm. Failure to predict is the case when a strong earthquake occurs outside of an alarm. Probabilistic component of prediction is represented by the rates of false alarms and failures to predict and the time–space covered by alarms (in % to total time–space considered).

in hierarchical non-linear systems: other geological disasters; geotechnical, and even socio-economic disasters. Qualitatively similar approach is routinely used in medicine, criminology, etc.

- However, accurate the short-term prediction would be it will not render unnecessary the predictions with a longer lead time. One can find in seismological literature a reappearing mistake: that only precise short-term (or even immediate) prediction is practically useful. Actually, protection from earthquakes requires a hierarchy of preparedness measures, from building codes, insurance, and issuing bonds, to reinforcement of high risk objects, to red alert. It takes different time, from decades, to years, to seconds to undertake different measures. Accordingly, earthquake preparedness requires all stages of prediction (Keilis-Borok, 2002; Molchan, 2003; Kantorovich and Keilis-Borok, 1991). Such is the case in preparedness to all disasters, war included.

4.5. Questions arising

- We considered only one short-term precursor—a chain of earthquakes—and one intermediate-term one—the pattern Σ. In subsequent applications (Shebalin et al., in press), all major types of intermediate-term seismicity patterns have been used with similar renormalization. The question arises which set of precursors provides the optimal prediction strategy, as defined for example in (Molchan, 2003; Zaliapin et al., 2003b).
- It is not yet clear how to make the scaling of RTP analysis self-adapting to the regional seismic regime, e. g. to parameters of the Gutenberg–Richter relation.
- Earthquake precursors emerge with the broader range of the lead times than considered here. They are divided, albeit fuzzily, into *long-term* (*tens of years*) ⇒ *intermediate-term* (*years*) ⇒ *short-term* (*months*) *and* ⇒ *immediate* (*days or less*). The question arises how to apply RTP analysis to the whole sequence or to its different parts.

Summing up, the RTP approach seems to open new possibilities in the quest for the short-term prediction. We hope that this study sets up a base for further development of this approach in the intertwined problems of earthquake prediction, fundamental understanding of dynamics of the lithosphere, and non-linear dynamics.

Acknowledgements

This study was supported by the 21st Century Collaborative Activity Award for Studying Complex Systems, granted by the James S. McDonnell Foundation. Exceedingly useful has been insightful criticism by K. Aki, G. Barenblatt, M. Ghil, R. Mehlman, A. Soloviev, J. Vidale, I. Zaliapin, and discussions at the Seventh Workshop on Non-Linear Dynamics and Earthquake Prediction (Trieste, 2003). We are grateful to D. Shatto and V. Ewing for patient and competent help in preparation of the manuscript.

References

Aki, K., 2003. Introduction to seismology for earthquake prediction. In: Proceeding of the Seventh Workshop on Non-Linear Dynamics and Earthquake Prediction, International Center for Theoretical Physics, Trieste (Preprint).

ANSS/CNSS Worldwide Earthquake Catalog. Produced by Advanced National Seismic System (ANSS) and hosted by the Northern California Data Center (NCEDC), 1965–2003, http://quake.geo.berkeley.edu/anss.

Bird, P., 1998. Testing hypotheses on plate-driving mechanisms with global lithosphere models including topography, thermal structure, and faults. J. Geophys. Res. 103 (B5), 10115–10129.

Blanter, E.M., Shnirman, M.G., 1997. Simple hierarchical systems: stability, self-organized criticality and catastrophic behavior. Phys. Rev. E 55 (6), 6397–6403.

Bowman, D.D., Ouillon, G., Sammis, C.G., Sornette, A., Sornette, D., 1998. An observational test of the critical earthquake concept. J. Geophys. Res. 103, 24359–24372.

Buffe, C.G., Varnes, D.J., 1993. Predictive modeling of the seismic cycle of the Great San Francisco Bay region. J. Geophys. Res. 98, 9871–9883.

Caputo, M., Console, R., Gabrielov, A.M., Keilis-Borok, V.I., Sidorenko, T.V., 1983. Long-term premonitory seismicity patterns in Italy. Geophys. J. R. Astron. Soc. 75, 71–75.

Gabrielov, A., Newman, W.I., 1994. Seismicity modeling and earthquake prediction: a review. In: Newman, W.I., Gabrielov, A., Turcotte, D.L. (Eds.), Non-linear Dynamics and Predictability of Geophysical Phenomena, vol. 18. American Geophysics University, International University of Geodesy and Geophysics, Geophysical Monograph 83, IUGG, pp. 7–13.

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

Gabrielov, A.M., Zaliapin, I.V., Newman, W.I., Keilis-Borok, V.I., 2000. Colliding cascades model for earthquake prediction. Geophys. J. Int. 143, 427–437.

Jaume, S.C., Sykes, L.R., 1999. 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.

Jin, A., Aki, K., Liu, Z., Keilis-Borok, V.I., 2003. Brittle–ductile interaction hypothesis for earthquake loading. Abstract, Hagiwara Symposium, XXIII General Assembly of IUGG, 2003, Sapporo, Japan.

Kantorovich, L.V., Keilis-Borok, V.I., 1991. Earthquake prediction and decision making: social, economic, legislative and civil defense domains. In: Proceedings of International Conference on Earthquake Prediction: State-of-the-Art, 15–18 October, Strasbourg, France, pp. 586–593.

Keilis-Borok, V.I., 1990. The lithosphere of the Earth as a non-linear system with implications for earthquake prediction. Rev. Geophys. 28 (1), 19–34.

Keilis-Borok, V.I., 1996. Intermediate-term earthquake prediction. Proc. Nat. Acad. Sci. U.S.A. 93, 3748–3755.

Keilis-Borok, V.I., 2002. Earthquake prediction: state-of-the-art and emerging possibilities. Annu. Rev. Earth Planet. Sci. 30, 1–33.

Keilis-Borok, V.I., Kossobokov, V.G., 1990a. Premonitory activation of earthquake flow: algorithm M8. Phys. Earth Planet. Inter. 61 (1-2), 73–83.

Keilis-Borok, V.I., Malinovskaya, L.N., 1964. One regularity in the occurrence of strong earthquakes. J. Geophys. Res. 69 (14), 3019–3024.

Keilis-Borok, V.I., Rotwain, I.M., 1990b. 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.

Keilis-Borok, V.I., Shebalin, P.N. (Eds.), 1999. Dynamics of Lithosphere and Earthquake Prediction. Phys. Earth Planet. Inter. (Special Issue), 111, 179–330.

Keilis-Borok, V.I., Shebalin, P.N., Zaliapin, I.V., 2002. Premonitory patterns of seismicity months before a large earthquake: five case histories in Southern California. Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16562–16567.

Keilis-Borok, V.I., Soloviev, A.A. (Eds.), 2003. Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer–Verlag, Heidelberg, 335 pp.

Keilis-Borok, V., Stock, J.H., Soloviev, A., Mikhalev, P., 2000. Pre-recession pattern of six economic indicators in the USA. J. Forecast. 19, 65–80.

King, S.D., Lowman, J.P., Gable, C.W., 2002. Episodic tectonic plate reorganizations driven by mantle convection. Earth Planet. Sci. Lett. 203, 83–91.

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

Kossobokov, V.G., Carlson, J.M., 1995. 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.

Kossobokov, V., Shebalin, P., 2003. Earthquake prediction. In: Keilis-Borok, V.I., Soloviev, A.A. (Eds.), Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer–Verlag, Heidelberg, pp. 141–207.

Ma, Z., Fu, Z., Zhang, Y., Wang, C., Zhang, G., Liu, D., 1990. Earthquake Prediction: Nine Major Earthquakes in China. Springer, New York.

Mogi, K., 1985. Earthquake Prediction. Academic Press, Tokyo.

Molchan, G.M., 2003. Earthquake prediction strategies: a theoretical analysis. In: Keilis-Borok, V.I., Soloviev, A.A. (Eds.), Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Springer–Verlag, Heidelberg, pp. 209–237.

NEIC, 2004. NEIC Worldwide Earthquake Catalog. Produced by United States Geological Survey National Earthquake Information Center, USGS, NEIC, http://neic.usgs.gov/neis/epic/epic.html.

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

Novikova, O., Shebalin, P., Keilis-Borok, V.I., 2002. A second approximation to intermediate-term earthquake prediction: two cases histories for southeastern Mediterranean. Problems of Theoretical Seismology and Seismicity. Comput. Seismol. 33, 186–206.

Press, F. (Eds.), 1965. Earthquake Prediction: A Proposal for a Ten Year Program of Research. Ad Hoc Panel on Earthquake Prediction, White House Office of Science and Technology, Washington, DC, 134 pp.

Press, F., Allen, C., 1995. Pattern of seismic release in the Southern California region. J. Geophys. Res. 100, 6421–6430.

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

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

Rundle, J., Turcotte, D., Klein, W. (Eds.), 2000. Geocomplexity and the Physics of Earthquakes. American Geophysical Union, Washington, DC, 284 pp.

Rundquist, D.V., Soloviev, A.A., 1999. Numerical modeling of block structure dynamics: an arc subduction zone. Phys. Earth Planet. Inter. 111 (3–4), 241–252.

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

Shebalin, P., Keilis-Borok, V., Zaliapin, I., Uyeda, S., Nagao, T., Tsybin, N., in press. Short-term advance prediction of the large Hokkaido earthquake, 25 September 2003, magnitude 8.1: a case history. Earth Planets Space, in press.

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

Sornette D., 2000. Critical Phenomena in Natural Sciences. Chaos, Fractals, Self-organization and Disorder: Concepts & Tools. Springer Ser. Synerg., Springer–Verlag, Heidelberg, 432 pp.

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

Uyeda, S., Park, S. (Eds.), 2002. In: Proceedings of the International Symposium on The Recent Aspects of Electromagnetic Variations Related with Earthquakes (Special Issue), 20 and 21 December 1999, J. Geodyn., 33, 4–5.

Zaliapin, I., Keilis-Borok, V.I., Axen, G., 2002a. Premonitory spreading of seismicity over the faults' network in southern California: precursor accord. J. Geophys. Res. 107 (B10), 2221.

Zaliapin, I., Liu, Z., Zöller, G., Keilis-Borok, V.I., Turcotte, D., 2002b. On increase of earthquake correlation length prior to large earthquakes in California. Comput. Seismol. 33, 141–161.

Zaliapin, I., Jin, A., Liu, Z., Aki, K., Keilis-Borok, V., 2003a. Temporal (un)correlations between coda Q-1 and seismicity—multiscale trend analysis. Pure. Appl. Geophys., in press.

Zaliapin, I., Keilis-Borok, V.I., Ghil, M., 2003b. A Boolean delay equation model of colliding cascades. Part II: Prediction of critical transitions. J. Stat. Phys. 111 (3–4), 839–861.

Zöller, G., Hainzl, S., Kurths, J., 2001. Observation of growing correlation length as an indicator for critical point behavior prior to large earthquakes. J. Geophys. Res. 106, 2167–2176.

Short-term earthquake prediction by reverse analysis of lithosphere dynamics

P. Shebalin [1,4], V. Keilis-Borok[1, 2, 3], A. Gabrielov[5], I. Zaliapin[1, 2] , D. Turcotte[6]

[1] *International Institute for Earthquake Prediction Theory and Mathematical Geophysics, Russian Ac. Sci., Warshavskoe sh., 79, korp. 2, Moscow, 113556, Russia*

[2] *Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California, 90095-1567, USA*

[3] *Department of Earth and Space Sciences, University of California, Los Angeles, California, 90095-1567, USA*

[4] *Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252, Paris Cedex 05, France*

[5] *Departments of Mathematics and Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907-1395, USA*

[6] *Department of Geology, University of California, Davis, CA 95616, USA*

Abstract

Short-term earthquake prediction, months in advance, is an elusive goal of earth sciences, of great importance for fundamental science and for disaster preparedness. Here, we describe a methodology for short-term prediction named RTP (Reverse Tracing of Precursors). Using this methodology the San Simeon earthquake in Central California (magnitude 6.5, Dec. 22, 2003) and the Tokachi-oki earthquake in Northern Japan (magnitude 8.1, Sept. 25, 2003) were predicted six and seven months in advance, respectively. The physical basis of RTP can be summed up as follows: An earthquake is generated by two interacting processes in a fault network: an accumulation of energy that the earthquake will release and a decrease in stability triggering this release. Energy is carried by the stress field, instability is carried by the difference between the stress and strength fields. Both processes can be detected and characterized by “precursory” patterns of seismicity or other relevant fields. Here, we consider an ensemble of premonitory seismicity patterns. RTP methodology is able to reconstruct these patterns by tracing their sequence backwards in time. The principles of RTP are not specific to earthquakes and may be applicable to critical transitions in a wide class of hierarchical non-linear systems.

1. Introduction

There is increasing evidence that variations in regional seismicity occur prior to intermediate and large earthquakes (Knopoff et al., 1996; Bowman et al., 1998). Using pattern recognition techniques a series of algorithms have been developed which provide intermediate-term and long-term predictions with lead times of years to decades (Keilis-Borok and Shebalin, 1999; Keilis-Borok, 2002; Keilis-Borok and Soloviev, 2003; Rundle et al., 2003). In this paper an algorithm is introduced that has successfully made short-term earthquake predictions, i.e. months in advance (Shebalin et al., 2004) . This algorithm is now tested by advance prediction in several seismically active regions and its performance is yet to be validated. However, the first successes along with the novelty of the methodology used in this algorithm prompt description of its essential underlying ideas and approaches.

It should be emphasized that there are two quite different approaches to earthquake prediction. The first is to make continuous predictions; in terms of earthquakes, this requires the specification of earthquake risk at all spatial points at each time instant. This approach is very useful when predicting a large number of small to intermediate earthquake since one can directly compare the observed and predicted seismic rates (probabilities, intensities, *etc.*) using the log-likelihood paradigm (Daley and Vere-Jones, 2004) or the least-square discrepancy (Whittle, 1963). The second approach that is used here is binary: an earthquake it forecast (predicted) for a specified area and time window, called *alarm region* or *alarm*. This approach is better justified when predicting extremely rare large events, so the direct comparison of the predicted continuous rate with a couple of observed earthquakes is rather problematic (Molchan, 2003). Our goal is to narrow down the area and time duration of alarms, within which a target earthquake is expected. Prediction is targeted at the large and therefore rare earthquakes; in a typical alarm area they occur on average once in 10-20 years. Our prediction should capture the target within an interval 20-30 times smaller, since a short-term alarm lasts months. Thus our alarms should be equally rare and each correct alarm would typically capture only one target; in very rare cases more than one.

It should be noted that an early theoretical discussion of the necessity of a discrete binary approach instead of a continuous one in predicting rare point events appear in 70s (Lindgren, 1975,1985; De Mare, 1980); and the difference between continuous and binary predictions has been widely recognized in weather forecasting (Jolliffe and Stephenson, 2003). An example of binary forecast is a tornado warning issued for a specified area and time window. Tornado warnings are analogous to the earthquake alarms considered in this paper. Possible outcomes of such predictions are illustrated in Fig. 1. In this scheme, we have two types of errors: failures to predict (target earthquake outside alarm region) and false alarms (no target earthquakes within an alarm); we also consider the total spatio-temporal coverage of alarms as a characteristic of a prediction algorithm. Probability of errors of different types is estimated using a sequence of predictions and is visually represented in the error diagrams (Sect. 3 below; Molchan, 2003). An analogous approach in weather forecasting is the *relative operating characteristic* diagram (Jolliffe and Stephenson, 2003).

2. Reverse Tracing of Precursors (RTP)

We will now outline the RTP (reverse tracing of precursors) approach to short-term earthquake forecasting. Details are given in the Appendix. Three aspects are important:

(i) Precursory chains that reflect the premonitory increase of the earthquakes' correlation range; qualitatively speaking, these chains are the dense, long, and rapidly formed sequences of small and medium sized earthquakes. Their definition generalizes premonitory seismicity patterns ROC and ACCORD. Heuristically, the pattern ROC ensures the ongoing increase of earthquake correlation range, expressed via the pair-wise correlation function; while ACCORD reflects simultaneous activation of several major parts of the regional fault network. They represent complimentary approaches to detecting the earthquake correlation. Formal definitions of these patterns as well as their performance in synthetic and observed seismicity can be found in (Gabrielov et al., 2000; Shebalin et al., 2000; Zaliapin et al., 2002, 2003; Keilis-Borok et al., 2002). An alternative approach to measuring the earthquake correlation was introduced in (Zöller and Hainzl, 2001; Zöller et al., 2001).

(ii) Intermediate-term patterns, originally found in the modeled and observed seismicity (Prozorov and Schreider, 1990; Keilis-Borok and Shebalin, 1999; Keilis-Borok, 2002; Keilis-Borok and Soloviev, 2003; Gabrielov et al., 2000; Zaliapin et al., 2003). They reflect four major types of

premonitory phenomena: rise of seismic activity, rise of earthquakes' clustering, rise of earthquakes correlation range, and a transformation of the magnitude-frequency (Gutenberg – Richter) relation towards an increasing share of relatively large magnitudes.

(iii) Pattern recognition of infrequent events is used to define the precursory combination of the patterns. Specifically, we used the Hamming algorithm which in our case is analogous to voting (Keilis-Borok and Soloviev, 2003); this algorithm is a standard tool in making a decision considering several "opinions". Formally, the Hamming distance between two Boolean vectors of the same length is defined as the number of their non-coincident symbols. Here, the Hamming distance gives the number of emergent intermediate-term premonitory patterns (see details below).

RTP analysis consists of the following stages. First, we detect chains - the "candidates" for the short-term precursors. We have found that precursory chains emerge within months before most of the target earthquakes. However, up to 90% of the chains are not followed so closely by strong earthquakes and in prediction they would cause false alarms. To eliminate false alarms, we next determine which intermediate-term precursors have occurred in the vicinity of each candidate within few years preceding it. Finally, we apply pattern recognition: knowing for each candidate what intermediate-term patterns have preceded it, we recognize which chains are precursory and which are false alarms. Specifically, in order to decide whether a chain is premonitory or not we use a set of M individual intermediate-term premonitory patterns. Some of them give premonitory signal (emerge) while other do not. The current state of the patterns is represented by a Mx1 Boolean vector indicating which pattern emerge (1) and which is not (0). A zero vector would indicate that none of the patterns emerge and the chain is most probably not precursory, while a vector consisting of all ones that all the patterns emerge and the chain is most probably precursory. Hamming distance, defined as the number of ones in our Boolean vector, shows how far the vector is from a zero one; in other words, how many patterns "voted" for making the chain precursory. If sufficient number of votes is accumulated (the threshold is established during the learning) the chain is considered precursory. This brings us to the prediction proper. The emergence of each precursory chain starts an alarm: a target earthquake is expected during τ months after the chain was formed and in its formally defined vicinity.

Thus, the precursory chain indicates the narrow area of a possibly complex shape (the chain vicinity) where intermediate-term precursors should be looked for. Their presence in turn validates the chain, as a short-term precursor. A chain is considered first although it emerges later – hence our analysis is called *reverse*.

3. Performance

We have tested our algorithm by advance prediction in Southern and Central California using the earthquake catalog ANSS/CNSS starting from January 1965. First, the data for 1965 – 1994 have been used for "learning", i.e. self-adaptation of some of the parameters (see Appendix **A3**). Then, the resulting rule was tested on independent data (i.e. the data not used for learning) for the period from January 1995 to May 2003. In June 2003 we have launched advance prediction. The only target earthquake that happened in the prediction region during the advance phase of the experiment (the San Simeon earthquake, December 22, 2003, M=6.5) was successfully predicted. The alarm capturing this earthquake started on May 5, 2003 – the date when the precursory chain triggering the alarm was completed. This alarm was reported on June 21 of the same year (Aki et al., 2003)

The algorithm has also been applied to the territories of Japan; Central Apennines, Alps, Northern Dinarides and Po valley; and Eastern Mediterranean (Figs. 2-5) with magnitude of target earthquakes $M \geq 7$, $M \geq 5.5$ and $M \geq 6.5$ respectively. In Japan, the learning was performed during 1975-2003, and advance prediction started on July 1, 2003 (Shebalin et al., 2003). In Central Apennines, Alps, Northern Dinarides and Po valley the learning was performed during 1970-1990, and advance prediction started on May 12, 2004. In Eastern Mediterranean the learning was performed during 1983-2003, and advance prediction started on May 12, 2004. The intermediate-term patterns (see **Appendix A3**) showed amazing self-adjustment: they were applicable within all three regions, and to all chains within each region, with the same values of their four numerical parameters.

The Tokachi-Oki, Japan, earthquake, 25 September 2003, M=8.1, has been also predicted in advance: the alarm started on 27 March, 2003 and was reported on 2 July 2003 (Shebalin et al., 2004).

During the time period covered by our advance prediction experiment, two target earthquakes have occurred; both of them have been predicted. Three false alarms were issued; one alarm is current. Figures 3-5 and Table 1 summarize the results of the experiment. It is worth noticing that a large earthquake (M_L = 5.7, M_W = 5.3) occurred within the alarm issued in Northern Dinarides; and that two target earthquakes (M_W= 7.4, M_W= 7.2) occurred near one of the alarms issued in Japan, outside the formal prediction region. A retrospective prediction with extended region led to successful prediction of those two earthquakes.

3. Prediction quality

As we mentioned in introduction, the problem of evaluating a binary prediction requires special tools. The main difference from evaluating a continuous prediction is that we can no longer use a single measure of discrepancy between prediction and observations (one faces the same situation in classical hypothesis testing where errors of two types are introduced). We use three interdependent measures of prediction quality, defined in **Appendix A5**: fraction of unpredicted earthquakes, n; fraction of false alarms, f; and the space-time τ covered by all alarms together, normalized by the whole space-time considered. The space is measured not in km^2 but in long-term average of seismicity. We used the average number of mainshocks with $m \geq 4$. The optimal tradeoff between different characteristics depends on a loss function $L(n,f,\tau)$ for preparedness measures.

The *error diagram* juxtaposes the prediction errors; each particular prediction corresponds to a single point in (n, τ, f) space. The error diagram is used to evaluate the predictive power of our prediction algorithm and its stability. For illustration, the error diagram for our prediction experiment in California during 1964-2005 is shown in Fig. 6; it shows the relative alarm coverage τ (10%) vs. the number of failures to predict (0); the number of false alarms (5) is indicated in parentheses. A more detailed discussion of error diagram approach is given in appendix **A6.**

Discussion

1. *A possible physical mechanism* underlying the RTP methodology is based on models of dynamical systems (Gabrielov et al., 2000; Zaliapin et al., 2003) and geodynamics (Rundquist and Soloviev, 1999). Precursory chains outline the areas where instability is accumulated months before a target earthquake. This instability reveals itself through an increase of the earthquake correlation

range. Intermediate-term premonitory seismicity patterns considered reflect the accumulation of energy and instability necessary and sufficient to trigger an earthquake, in the area outlined by a precursory chain, but years before the chain. In more general terms, RTP identifies a small-scale perturbation that carries a memory of the larger scale history of a complex system (in our case, the fault network). Increases of the correlation range are a known symptom of critical transitions in statistical physics and of bifurcation in nonlinear dynamics (Kadanoff, 2000). Typically for premonitory patterns of this kind precursors considered are sporadic short-lived phenomena not necessarily reflecting the steady trends of seismicity. This suggests that both patterns are symptoms but not the causes of a target earthquake: they signal its approach but do not trigger it. Such sporadic precursors to critical phenomena have been found also in socio-economic complex systems (Keilis-Borok et al., 2000).

2. It seems promising to apply RTP analysis to the detection of earthquake precursors in the other relevant and available data such as electromagnetic fields (Uyeda and Park, 2002), fluid regime (Ma et al., 1990), InSAR and GPS (Simons et al., 2002). The first positive result has been obtained with precursors gauging interaction between the ductile and brittle layers of the Earth crust; this opens a highly promising link of geodynamics and nonlinear dynamics approaches to prediction (Jin et al., 2004).

3. *The methodological advantage* of RTP over a direct analysis is in the drastic reduction in dimensionality of the parameter space where premonitory patterns are looked for. We have found here the patterns formed in narrow areas different from case to case, whose shape might be complicated, and with diverse size. To find these areas by a trial-and-error procedure would require trying different shapes, sizes, chains and locations, which is hardly realistic. Reverse analysis resolves this impasse, determining from the start a limited number of the areas to consider. Thus, RTP analysis provides a common methodological approach to the prediction of avalanches in a wide class of the complex systems, formed separately or jointly by nature and society.

4. The only decisive test of any prediction theory is an experiment in advance prediction. Such an experiment for the methodology described above was launched in June 2003 and is currently maintained by University of California Los Angeles (USA), Russian Academy of Sciences, and Institut de Physique du Globe de Paris (France). The complete results, including the San Simeon prediction, will be published elsewhere. The goal of this paper is to present the essential underlying concepts and report its first successes to a broad range of multidisciplinary experts, attracting their attention to the possibility of exploring premonitory patterns in diverse physical fields using the RTP methodology.

APPENDICES

A1. Earthquake catalogs

The data used in analysis are provided by the routinely compiled earthquake catalogs, which present at the moment the most accurate and complete information about the dynamics of seismicity. The earthquake catalog is taken from ANSS/CNSS and NEIC. We use a common representation of the earthquake catalog $\{t_j, \varphi_j, \lambda_j, M_j, b_j\}$, j = 1, 2, … Here t_j is the time of an earthquake, $t_j \geq t_{j-1}$; φ_j and λ_j – latitude and longitude of its epicenter; and M_j magnitude. We consider the earthquakes with magnitude $M \geq M_{min}$. As in most premonitory patterns of that family [Keilis-Borok 1996, 2002] aftershocks are eliminated from the catalog; however, an integral measure of aftershocks activity b_j

is retained for each remaining earthquake (main shocks and foreshocks); b_j is the number of aftershocks occurring immediately after an earthquake (e.g. within two days).

A2. Chains

A chain captures a rise of earthquakes' correlation range in its vicinity. Let us call two earthquakes "neighbors" if their epicenters are closer than r and their times are closer than τ_0. A chain is a sequence of earthquakes where each earthquake has at least one neighbor belonging to that sequence and, therefore, no neighbors outside the sequence. he average density of epicenters decreases with increasing magnitudes. Accordingly, r is normalized as $r = r_0 10^{c(\underline{m}-2.5)}$, where $\underline{m}$ is the smallest magnitude in the pair. The R-vicinity of a chain is outlined by the smoothed envelope of the circles of a radius R drawn around each epicenter in the chain. We consider only the chains with two sufficiently large characteristics: number of earthquakes $k \geq k_0$, maximal distance between epicenters $l \geq l_0$. Two parameters of the chains are common for all the regions: r_0=50 km, c = 0.35. Other parameters are common for all chains within a region, but differ between regions as follows: Southern California, τ_0=20 days, k_0=6, $M_{\min}$=2.9, l_0=175 km; Central California, τ_0 =30 days, k_0=10, $M_{\min}$=2.9, l_0=250 km; in Japan, τ_0 =20 days, k_0=10, $M_{\min}$=3.6, l_0=350 km, γ_0=0.4; Eastern Mediterranean, τ_0 =40 days, k_0=6, $M_{\min}$=3.0, l_0=200 km.

A3. Intermediate-term patterns

We look for intermediate-term patterns in the R-vicinity of each chain within T years preceding it. To detect a pattern P we compute a function $F_P(t_j)$ defined in the "event window" (Keilis-Borok and Soloviev, 2003) i.e. on the sequence of N consecutive earthquakes with indexes j-N+1, j-N+2,..., j. In R –vicinity of each chain we normalize seismicity by the lower magnitude cutoff M^*. The latter is derived from magnitude-frequency relation, by the condition $n(M^*) = n^*$; here $n(M^*)$ is the annual number of earthquakes with magnitude $M \geq M^*$.

Four functions represent **a rise of activity**. Namely

"Activity" $$F_U(t_j) = \frac{N}{t_j - t_{j-N+1}} \qquad \text{(A1)}$$

is inversely proportional to the time it took to accumulate the most recent N earthquakes;

"Sigma" $$F_\Sigma(t_j) = \sum_{k=j-N+1}^{j} 10^{M_k - M^*} \qquad \text{(A2)}$$

is shown to be a crude measure of total area of fault-breaks during the most recent N earthquakes (Keilis-Borok, 2002);

"Rise of magnitudes", $$F_M(t_j) = \frac{2}{[N/2]}\left(\sum_{k=j-N+1}^{j-N+[N/2]} M_k - \sum_{k=j-[N/2]+1}^{j} M_k\right) \qquad \text{(A3)}$$

is the difference between the average magnitude of the last [N/2] earthquakes and that of the first [N/2] earthquakes within a series of N;

"Acceleration" $$F_C(t_j) = \frac{1}{[N/2]}\left(\sum_{k=j-N+1}^{j-N+[N/2]} \frac{1}{t_k - t_{k-1}} - \sum_{k=j-[N/2]+1}^{j} \frac{1}{t_k - t_{k-1}}\right) \qquad \text{(A4)}$$

is connected to the function “Activity”. “Acceleration” increases if intercurrence time between earthquakes decreases with time.

Here $[x]$ denotes integer part of x.

Two functions depict **a rise of clustering**:

“Swarm” $F_W(t_j) = 1 - \frac{A_r(t_j)}{\pi\, r^2 N}$ (A5)

with A_r being the area of the union of the circles of radius r centered at N epicenters in the sequence reflects clustering of mainshocks; while

“b-micro” $F_{b_\mu}(t_j) = \sum_{k=j-N+1}^{j} \sum_{l} 10^{M_{kl} - M^*}$ (A6)

reflects clustering of aftershocks; here M_{kl}, l = 1, 2,… are the magnitudes of the aftershocks of the k-th main shock within the first 2 days after the main shock

The **rise of earthquakes correlation range** is depicted by function

“Accord” $F_A(t_j) = \frac{A_r(t_j)}{\pi\, r^2}$, (A7)

which increases if earthquakes are widely distributed in space and their r-neighbourhoods are barely overlapping. Finally, **the transformation of Gutenberg-Richter relation** is reflected by function

“Gamma” $F_\gamma(t_j) = \frac{1}{N_{M_k \ge M_{1/2}}} \sum_{M_k \ge M_{1/2}} (M_k - M^*)$, (A8)

which increases if the magnitude distribution is shifted to the larger magnitudes (e.g., if the GR slope is decreasing). Here $M_{1/2}$ is the median of magnitudes of N earthquakes in our sequence.

Altogether the eight functions are determined by five parameters. In each region we used the same eight combinations of these parameters: n^*=10 and R=50 km or n^*=20 and R=100 km, N=10 or 50, T=6 or 24 months, r=50 km. Emergence of a pattern at the moment t is captured by the condition $F_P(t) \ge C_P$. Each threshold C_P is determined automatically at the learning stage. It minimises the sum $n + f$; here n is the rate of failures to predict and f is the rate of false alarms in prediction with a single pattern P.

A4. Prediction

Final stage is recognition of precursory chain and issuing an alarm: A chain is recognised as precursory if it was preceded by C or more intermediate-term patterns out of the ensemble considered. The threshold C controls the trade-off between the rates of false alarms and failures to predict. Emergence of precursory chain triggers an alarm in its R-vicinity for the Δ months; statistics of past alarms suggests Δ = 9 months. A precursory chain may keep growing accumulating subsequent earthquakes. In that case the alarm is extended. If a target earthquake occurs in the R-vicinity of a chain, then the chain no longer grows, but the alarm (if it has been diagnosed for that chain) is not called off. After a target earthquake all other chains containing its epicenter within the *R*-vicinity are disregarded during the period Δ.

A5. Quality of prediction

Suppose that the prediction was performed during the time interval of length T (yr) within the region Ω with the area S (km^2); N large earthquakes occurred within this period; A alarms were declared and A_f of them were false; all the alarms together covered the spatio-temporal volume V_A (yr × km^2); N_f target earthquakes were unpredicted. Prediction is described by the following dimensionless errors: the fraction of unpredicted earthquakes, $n = N_f / N$; the relative alarm coverage, $\tau = V_A/(T \times S)$; the fraction of false alarms, $f = A_f / A$.
When calculating the alarm coverage, it might be advantageous to take into account the observed inhomogeneities of the earthquake spatial distribution. In our prediction experiment, the relative alarm coverage for an alarm that spans the time T_A and space S_A is calculated as

$$\tau_A = \frac{T_A}{T} \times \frac{\int_{S_A} dN_4(r)}{\int_S dN_4(r)} = \frac{T_A}{T} \times \frac{\#\{\text{EQ with } m \geq 4 \text{ within } S_A\}}{\#\{\text{EQ with } m \geq 4 \text{ within } S\}}. \qquad \text{(A9)}$$

Here by $N_4(r)$ we denote the 2D point process of earthquakes with magnitude $m \geq 4$. The total alarm coverage is the sum of that for all individual alarms.

A6. Significance level: Random Binomial prediction

To evaluate significance of a prediction one typically evaluates the chances of getting the same or better result (same or smaller values of errors) when there is no dependence between alarms and the occurrence of target earthquakes. An extremely simple but easily tractable model of prediction which produces alarms independent of the target earthquakes is *random binomial prediction* (Molchan, 2003): One divides the space-time considered for prediction into M small equal bins and declares alarm in each of them with fixed probability p. Indeed, this approach is highly unrealistic. Nevertheless, considered as a null (random) prediction model, it provides a good coarse estimation of the algorithm predictive power. Significance with respect to a random binomial prediction may serve as a necessary, but not sufficient, condition for validating an algorithm.

It is readily checked that expected values of alarm coverage τ and fraction f of failures to predict in the binomial prediction are given by:

$$E(\tau) = p, \qquad E(f) = 1 - p$$

so the point corresponding to this prediction is on the diagonal $f = (1 - \tau)$ in the 2D (τ, f)-section of the error diagram. The probability to predict exactly $N - N_f$ out of N target earthquakes, assuming that no more than one target earthquake may occur within a single bin, is given by Binomial distribution

$$\Pr\{\text{predict } N - N_f \text{ out of } N\} = \binom{N}{N_f} p^{N-N_f}(1-p)^{N_f}. \qquad \text{(A10)}$$

The probability to predict $N - N_f$ out of N target earthquakes issuing alarm within k bins out of M is given by Hypergeometric distribution

$$\Pr\{\text{predict } N-N_\mathrm{f} \text{ out of } N \text{ declaring alarm in } k \text{ bins out of } M\} = \frac{\binom{k}{N-N_\mathrm{f}}\binom{M-k}{N_\mathrm{f}}}{\binom{M}{N}}. \quad \text{(A11)}$$

The number of false alarms can also be obtained, but because of the simplistic binomial rules, the number of binomial alarms (and false alarms) will be significantly larger than that in any realistic prediction (where alarm is typically declared for considerable spatio-temporal area, not for a small bin). Thus here we do not make any inference about false alarms using the binomial prediction model.

Using the above probabilities (A10, A11) one can construct different significance measures for a given prediction with errors (τ^*, n^*). One approach is to use the 2D (τ, n) distribution under the binomial model using (A10) with $p=\tau^*$, and evaluate probability of obtaining a prediction of the same or better quality, say

$$\Pr\{(\tau,n): \tau+n \le \tau^*+n^*\} \quad \text{or} \quad \Pr\{(\tau,n): \tau \le \tau^* \;\&\; n \le n^*\}.$$

Another approach is to use (A11) to find the probability to predict the same or larger number of earthquakes with the same total duration of alarms. The difference between using (A10) and (A11) is that in the first case we assume fixed *probability* of declaring an alarm, while in the second – fixed *duration* of alarm. Indeed, in generic cases both approaches give very similar evaluation of prediction performance.

To illustrate the above approach, Fig. 6 shows the error diagram for the results of our prediction experiment in California. Shaded ball represents the errors of our prediction experiment during 1964-2005. The probability for a random binomial prediction with given value of τ to fall within the shaded area (*i.e.*, to predict more than $N(1-n)$ target earthquakes with given τ) is less or equal than 0.001 (0.1%). The point that corresponds to our experiment is well within this area, thus indicating very high predictive power. It should be emphasized that the results presented in this figure combine the information from the learning period, independent data, and advance prediction (we have too few alarms and target earthquakes during the advance phase to use them alone). Thus, this analysis is not equivalent to evaluating the real predictive power of the algorithm, where only advance results must be used. Nevertheless, the grey shadowed area that corresponds to the binomial model gives a good orientation for the expected significance of the results.

A7. Significance level: Empirical estimation

An alternative approach to testing significance of a prediction algorithm involves empirical estimations of occurrence rate for target earthquakes. Thus, the approach is unavoidably approximate due to the small number of target earthquakes; yet it is much more realistic comparing to the random binomial prediction.

Specifically, we assume that target earthquakes form a Poisson process $N(t,r)$ stationary in time but non-homogeneous in space. The expected number of earthquake within the interval of length t and spatial region R is given by

$$E(N(t,R)) = t \times \mu(R) \qquad \text{(A12)}$$

where $\mu(R)$ is some non-negative measure over the space. In practice, a first-order approximation to this measure can be obtained by considering the number N_R of target earthquakes within the region R per unit of time using observations over S years:

$$\mu(R) \approx N_R / S.$$

With our assumptions, the probability of having exactly k target earthquakes within the region R during time interval of length t is given by Poisson distribution

$$\Pr\{k \text{ target earthquakes within } R\} = e^{-\mu(R)t} \frac{\left(\mu(R)t\right)^k}{k!} \qquad \text{(A13)}$$

and the probability p to have at least one target earthquake is

$$p := \Pr\{\text{at least one target earthquakes within } R\} = 1 - e^{-\mu(R)t}. \qquad \text{(A14)}$$

When the rate of target earthquakes is small (which is indeed the case in our experiment), we can approximate p as

$$p := \Pr\{\text{at least one target earthquakes within } R\} \approx \mu(R)t \approx \frac{N_R\, t}{S}$$

Our final goal is to calculate the probability of predicting $N - N_{\mathrm{f}}$ target earthquakes out of N by a set of alarms $A_i = (t_i, R_i)$ that were declared for regions R_i and time intervals t_i. We denote by p_i the probability to have at least one target earthquake within A_i.

The probability for a given target earthquake to be predicted is calculated as the probability that it will be predicted by at least one of the alarms:

$$\Pr\{\text{given target EQ is predicted}\} = \sum_i \Pr\{\text{given target EQ is within } A_i\}$$
$$= \sum_i \Pr\{\text{given target EQ is within } R_i \text{ during } t_i\}$$
$$= \sum_i \Pr\{\text{given target EQ is within } R_i\} \times \frac{t_i}{T}$$
$$= \sum_i q_i \times \frac{t_i}{T}$$

Here we used the fact that alarms are not overlapping (by definition); factorization property (A12) of the target event process; and the fact that conditional distribution of the occurrence time of an event from Poisson process is uniform, given that this event occurred within the given time interval.

The probability for a given target earthquake to happen within the spatial region R_i can be estimated as

$$q_i = \Pr\{\text{given target EQ happened within } R_i\} = \frac{n_i}{N_\Omega},$$

where n_i is the number of target earthquakes within R_i during some period S and N_Ω is the total number of target earthquakes within the region Ω considered for prediction during the same time. Finally

$$Q := \Pr\{\text{given target EQ is predicted}\} \approx \sum_i \frac{n_i}{N_\Omega} \times \frac{t_i}{T} \approx \sum_i p_i \frac{S}{N_\Omega T},$$

and the distribution of the number of predicted target earthquakes out of N is given by the Binomial formula:

$$\Pr\{N - N_\mathrm{f} \text{ out of } N \text{ target EQs are predicted}\} = \binom{N}{N_\mathrm{f}} Q^{N-N_\mathrm{f}} (1-Q)^{N_\mathrm{f}}$$

$$\approx \binom{N}{N_\mathrm{f}} \left(\sum_i \frac{n_i t_i}{N_\Omega T} \right)^{N-N_\mathrm{f}} \left(1 - \sum_i \frac{n_i t_i}{N_\Omega T} \right)^{N_\mathrm{f}}$$

$$\approx \binom{N}{N_\mathrm{f}} \left(\sum_i p_i \frac{S}{N_\Omega T} \right)^{N-N_\mathrm{f}} \left(1 - \sum_i p_i \frac{S}{N_\Omega T} \right)^{N_\mathrm{f}}$$

We apply the above approach to California. Specifically, we consider the region Ω shown in Fig. 2 during the period 1965-2004 (S = 40 years); there were N_Ω = 10 target earthquakes. The advance prediction was performed within the same region during July 2003 – June 2005 (T = 2 years), and resulted in three alarms; N = 1 target earthquake occurred during this period. The probabilities p_i of having at least one target earthquake within each of the alarms are 5%, 8%, and 5% (see Table 1, and Eq. (A14)). The probability to predict the only target event by chance is estimated as 36%. Notice that this is the conditional probability given the actual number of target earthquakes and alarms. If one does not want to be conditioned by the number of actual target earthquakes, then we need to modify our results using (A13). In the case of California, where we had only one target earthquake, this will give:

$$\Pr\{\text{predict 1 target with our three alarms}\} =$$
$$\Pr\{\text{there is exactly one target}\} \times \Pr\{\text{it was predicted}\}.$$

The first probability is estimated using (A13):

$$\Pr\{\text{there is exactly one target}\} = \mu(R) T e^{-\mu(R)T}$$
$$\approx \frac{N_\Omega}{S} T e^{-\frac{N_\Omega}{S}T} = \frac{10}{40} \times 2 \times e^{-\frac{10}{40}\times 2} \approx 0.3 \quad .$$

Thus, the probability to have only one target event and predict it by chance is approximately 0.36×0.3 = 0.12 (or 12%).

REFERENCES

K. Aki, V. Keilis-Borok, A. Gabrielov, A. Jin, Z. Liu, P. Shebalin, I. Zaliapin. On the current state of the lithosphere in Central California. Letter with San Simeon prediction sent to a group of leading experts on June 21, 2003. Available at http://www.math.purdue.edu/~agabriel/Workshop/SanSimeon_let.pdf

ANSS/CNSS Worldwide Earthquake Catalog. (http://quake.geo.berkeley.edu/cnss) (Produced by Advanced National Seismic System (ANSS) and hosted by the Northern California Data Center (NCEDC), 1965-2003).

Bowman, D. D., Ouillon, G., Sammis, C. G., Sornette, A., and Sornette, D., 1998. An observational test of the critical earthquake concept. J. Geophys. Res. 103, 24359-24372.

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

Jin, A., Aki, K., Liu, Z., and Keilis-Borok, V., 2004. Seismological evidence for the brittle-ductile interaction hypothesis on earthquake loading. Earth Planets Space, 56, 823-830.

Jolliffe, I. T., and Stephenson, D. B., 2003. Forecast Verification, Wiley, Chichester.

Kadanoff, L. P., 2000. Statistical Physics: Statics, Dynamics, and Renormalization, World Scientific Publishing, Singapore.

Keilis-Borok, V., 2002. Earthquake prediction: State-of-the-art and emerging possibilities, Ann. Rev. Earth Planet. Sci., 30, 1-33.

Keilis-Borok, V. I., and Shebalin, P. N. (eds), 1999. Dynamics of the lithosphere and earthquake prediction. Phys. Earth Planet. Inter. (special issue) 111, 179-330.

Keilis-Borok, V., Shebalin, P., Gabrielov, A., and Turcotte, D., 2004. Reverse tracing of short-term earthquake precursors. Phys. Earth Planet. Inter., 145, 75-85.

Keilis-Borok, V. I., Shebalin, P. N., and Zaliapin, I. V., 2002. Premonitory patterns of seismicity months before a large earthquake: Five case histories in Southern California. Proc. Natl. Ac. Sci. 99, 16562-16567.

Keilis-Borok, V. I., and Soloviev, A. A. (eds), 2003. Nonlinear Dynamics of the Lithosphere and Earthquake Prediction, Springer, Heidelberg.

Keilis-Borok, V., Stock, J. H., Soloviev, A., and Mikhalev, P., 2000. Pre-recession pattern of six economic indicators in the USA. J. Forecast., 19, 65-80.

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

Lindgren, G., 1975. Prediction from a random time point. Ann. Probability, 3, 412-423.

Lindgren, G., 1985. Optimal prediction of level crossings in Gaussian processes and sequences. Ann. Probab. 13, 804-824.

Ma, Z. et al., 1990. Earthquake Prediction: Nine Major Earthquakes in China, Springer, New York.

De Mare, J., 1980. Optimal prediction of catastrophes with applications to Gaussian processes. Ann. Probability, 8, 841-850.

Molchan, G. M., 2003. Earthquake prediction strategies: A theoretical analysis, in Nonlinear Dynamics of the Lithosphere and Earthquake Prediction, eds Keilis-Borok, V. I., and Soloviev, A. A., pp. 209-237, Springer, Heidelberg.

Prozorov, A. G., & Schreider, S.Yu, 1990. 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.

Rundle, J. B., Turcotte, D. L., Shcherbakov, R., Klein, W. and Sammis, C., 2003. Statistical physics approach to understanding the multiscale dynamics of earthquake fault systems. Rev. Geophys. 41, 1019.

Rundquist, D. V., and Soloviev, A. A., 1999. Numerical modeling of block structure dynamics: an arc subduction zone. Phys. Earth Planet. Inter. 111, 241-252.

Shebalin, P., Keilis-Borok, V., Zaliapin, I., Uyeda, S., Nagao, T., and Tsybin, N., 2004. Advance short-term prediction of the large Tokachi-oki earthquake, September 25, 2003, M=8.1 A case history, Earth Planets Space, 56, 715-724.

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

Simons, M. Fialko, Y., and Rivera, L., 2002. Coseismic deformation from the 1999 Mw 7.1 Hector Mine, California, earthquake as inferred from InSAR and GPS observations. Bull. Seismol. Soc. Am. 92, 1390-1402.

Uyeda, S., and Park, S. (eds), 2002. Proceedings of the International Symposium on The Recent Aspects of Electromagnetic Variations Related with Earthquakes, 20 and 21 December 1999. J. Geodynamics (special issue) 33, 4-5.

Daley D. J. and Vere-Jones D., 2004. Scoring probability forecasts for point processes: The entropy score and information gain. J. Appl. Prob., 41A: 297-312.

Whittle, P., 1963. Prediction and Regulation by Linear Least-Squares Methods, English University Press, London.

Zaliapin, I., Keilis-Borok, V. I., and Axen, G., 2002. Premonitory spreading of seismicity over the faults' network in southern California: Precursor Accord. J. Geophys. Res. B 107, 2221.

Zaliapin, I., Keilis-Borok, V.I., and Ghil, M., 2003. A Boolean delay equation model of colliding cascades, Part II: Prediction of critical transitions. J. Stat. Phys. 111, 839-861.

Zöller G., Hainzl S., 2001. Detecting premonitory seismicity patterns based on critical point dynamics. Natural Hazards and Earth System Sciences, 1, 93-98.

Zöller G., Hainzl S., Kurths J. 2001. Observation of growing correlation length as an indicator for critical point behavior prior to large earthquakes. J. Gephys. Res., 106. 2167-2176.

Acknowledgements. We are grateful to K. Aki, M. Ghil, A. Jin, L. Knopoff, J. McWilliams, and A. Soloviev for tough constructive criticism; to A. Kashina for inspired editing; and to D. Shatto for help in preparing the manuscript. This study was supported by the 21st Century Collaborative Activity Award for Studying Complex Systems from the James S. McDonnell Foundation, the National Science Foundation under grant ATM 0327558, and ISTC project 1538. The earthquake catalogs are compiled by Advanced National Seismic System (ANSS), Geophysical Institute of Israel (GII), and Japanese Meteorological Agency (JMA). The JMA catalog was received through the Japan Meteorological Business Support Center.

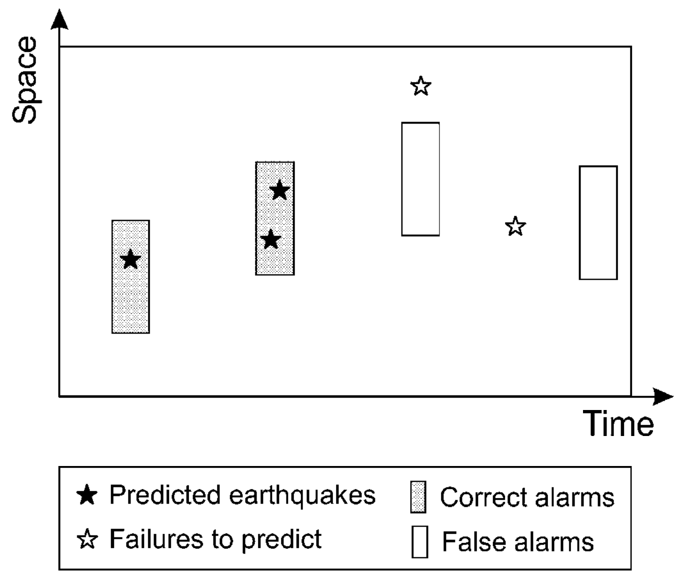


Figure 1 Possible outcomes of prediction. For simplicity the territory where the prediction is made is represented by a 1*D* 'Space' axis. Rectangles – space-time areas covered by correct (gray) and false (white) alarms respectively.

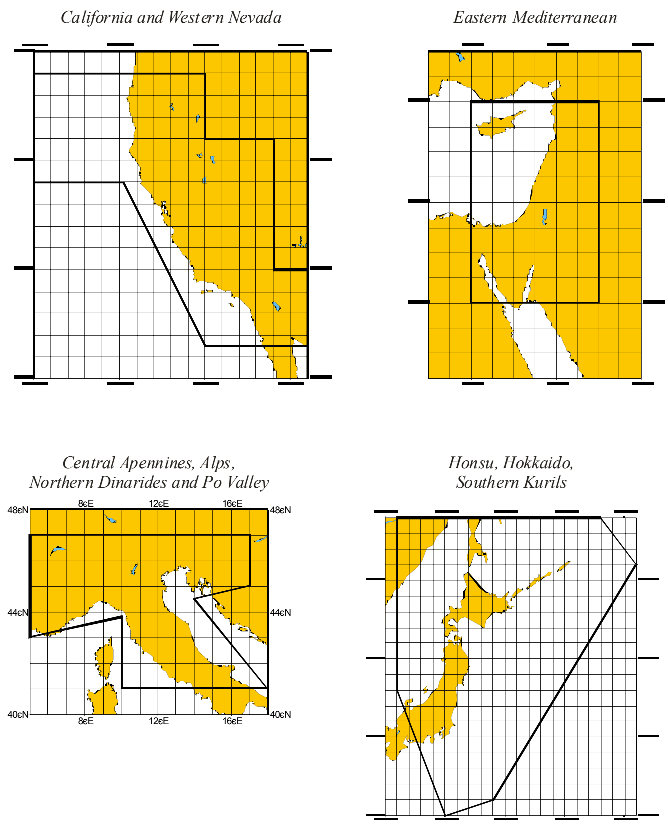


Figure 2 Regions where the proposed algorithm was tested by advance prediction. See text for details.

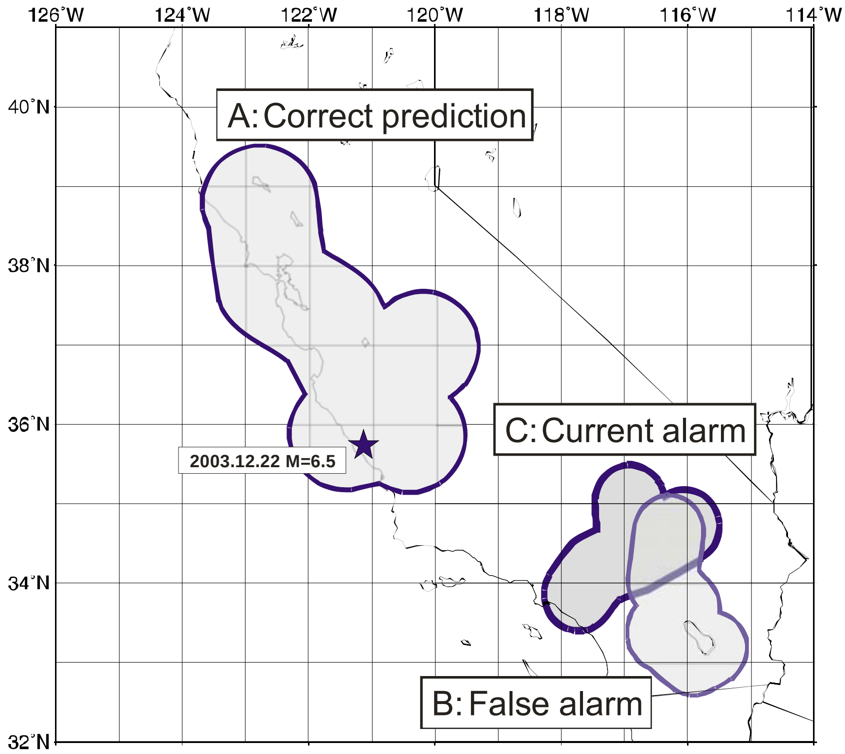


Figure 3 Results of advance prediction in California. The advance prediction started in July 2003. Three alarms were issued: one correct (marked A), one false (B), one current (C).

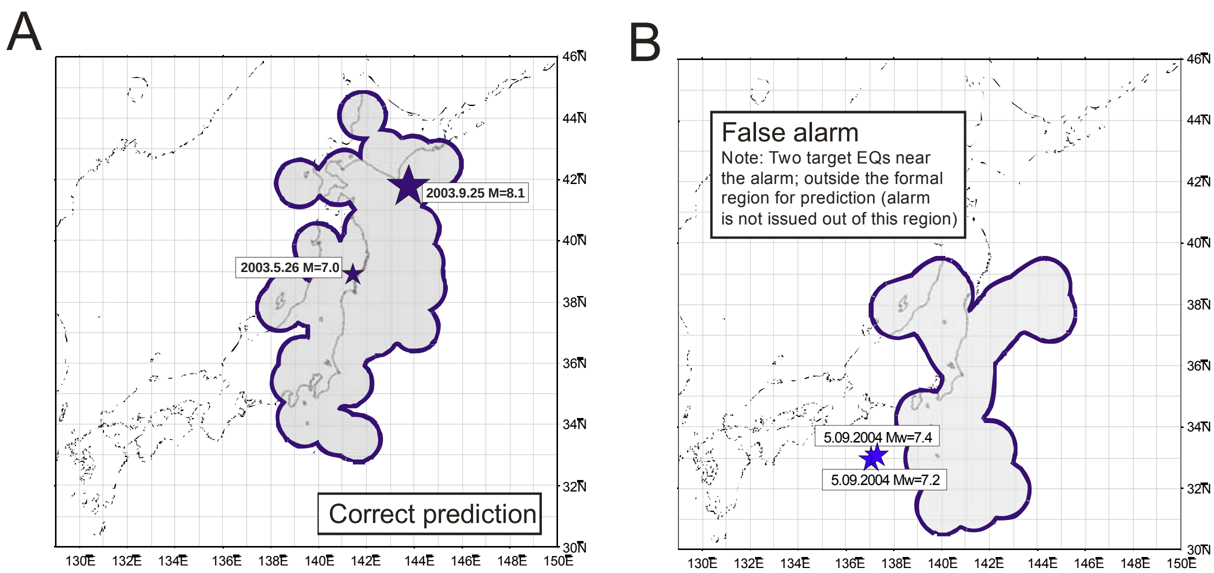


Figure 4 Results of advance prediction in Japan. The advance prediction started in July 2003. Two alarms were issued: one correct (panel a), and one false (panel b). We notice that two target earthquakes occurred outside the formal prediction region near the boundaries of our false alarm; if one extended the prediction region to include these two target earthquakes, they would be successfully predicted with the current values of algorithm parameters.

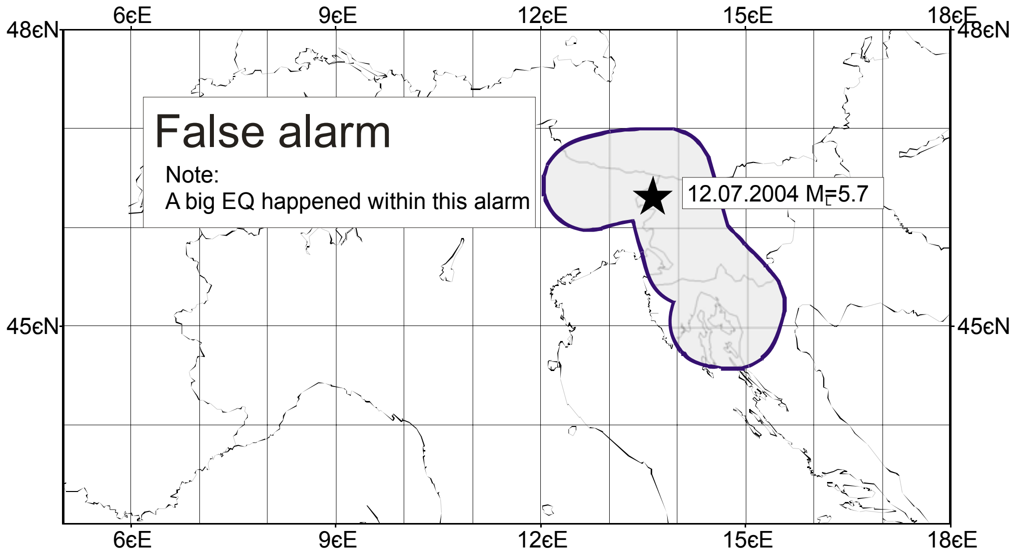


Figure 5 Results of advance prediction in Northern Dinarides**.** The advance prediction started in May 2004. One false alarm was issued. We notice that the only big earthquake (M_L=5.7) in the considered region since May 2004 happened within our false alarm.

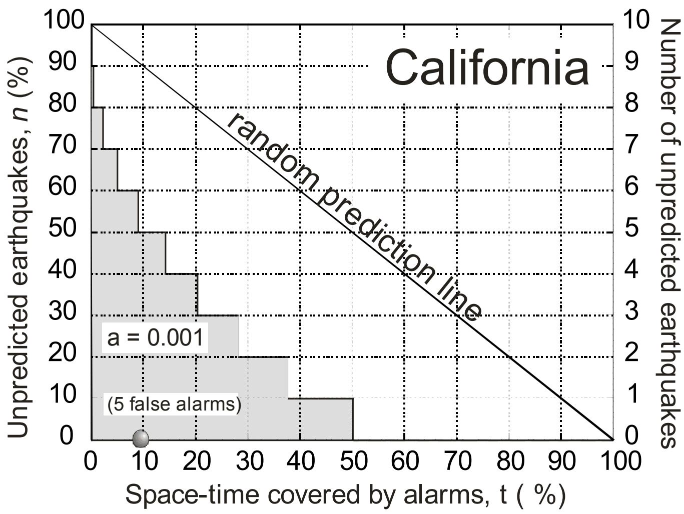


Figure 6 Significance of prediction in California: an illustration (see Appendix **A6** for details). Shaded ball shows performance of the prediction algorithm during the time interval considered. A perfect prediction would lie in the origin. Random binomial predictions (alarm is declared for each elementary spatio-temporal unit with a fixed probability τ) asymptotically occupy the diagonal, but might deviate from it with finite number of target earthquakes. Random predictions with fixed τ fall in the grey area with probability α = 0.001. Note that the shape of the grey area depends on the number of the target earthquakes that actually happened within the prediction region.