united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
**abdus salam**
international centre for theoretical physics

**H4.SMR/1519-50**

**"Seventh Workshop on Non-Linear Dynamics and Earthquake Prediction"**

**29 September - 11 October 2003**

# Real-Time Earthquake Prediction in Two Approximations

*V. Kossobokov*

**International Institute of Earthquake Prediction Theory**
**and Mathematical Geophysics, Russian Academy of Sciences**
**Moscow, Russia**
**&**
**Institute de Physique du Globe de Paris**
**France**

# Real-time Earthquake Prediction in Two Approximations


V. Kossobokov[1,2]

International Institute of Earthquake Prediction Theory and Mathematical Geophysics,
Russian Academy of Sciences,
79-2 Warshavskoye Shosse, Moscow 113556, Russian Federation

Institute de Physique du Globe de Paris,
4 Place Jussieu, 75252 Paris, Cedex 05, France

*E-mail:* volodya@mitp.ru *or* volodya@ipgp.jussieu.fr

# Outline

- Earthquake prediction definition
- Intermediate-term middle-range earthquake prediction algorithm M8
- How to reduce earthquake prediction uncertainty from middle-range to narrow? Algorithm MSc
- Global Test of M8-MSc

# What is earthquake prediction?

**The United States National Research Council, Panel on Earthquake Prediction of the Committee on Seismology suggested the following definition (1976, p.7):**

*"An earthquake prediction must specify the expected magnitude range, the geographical area within which it will occur, and the time interval within which it will happen with sufficient precision so that the ultimate success or failure of the prediction can readily be judged. Only by careful recording and analysis of failures as well as successes can the eventual success of the total effort be evaluated and future directions charted. Moreover, scientists should also assign a confidence level to each prediction."*

# Stages of earthquake prediction

- Term-less prediction of earthquake-prone areas
- Prediction of time and location of an earthquake of certain magnitude

| Temporal, ***in years*** | | Spatial, ***in source zone size L*** | |
|---|---|---|---|
| Long-term | 10 | Long-range | up to 100 |
| Intermediate-term | 1 | Middle-range | 5-10 |
| Short-term | 0.01-0.1 | Narrow | 2-3 |
| Immediate | 0.001 | Exact | 1 |

- Moreover, the Gutenberg-Richter law suggests limiting magnitude range of prediction to about one unit. Otherwise, the statistics would be essentially related to dominating smallest earthquakes.

# How earthquake prediction methods work?

“Predicting earthquakes is as easy as one-two-three.

**Routine seismological data bases, e.g. US GS/NEIC**

- Step 1: Deploy your precursor detection instruments at the site of the coming earthquake.

**Reproducible intermediate-term algorithms, e.g. M8**

- Step 2: Detect and recognize the precursors.

**Number of earthquakes have been predicted**

- Step 3: Get all your colleagues to agree and then publicly predict the earthquake through approved channels.”

Scholz, C.H., 1997. Whatever happened to earthquake prediction. *Geotimes*, **42**(3), 16-19

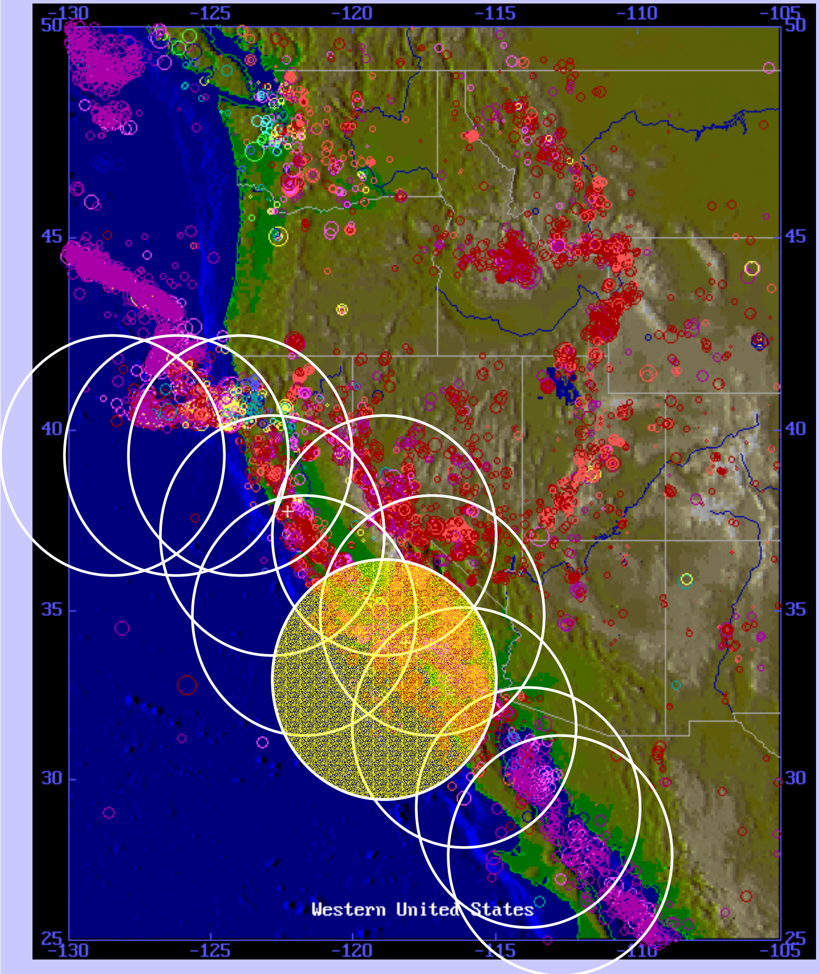
# M8 algorithm

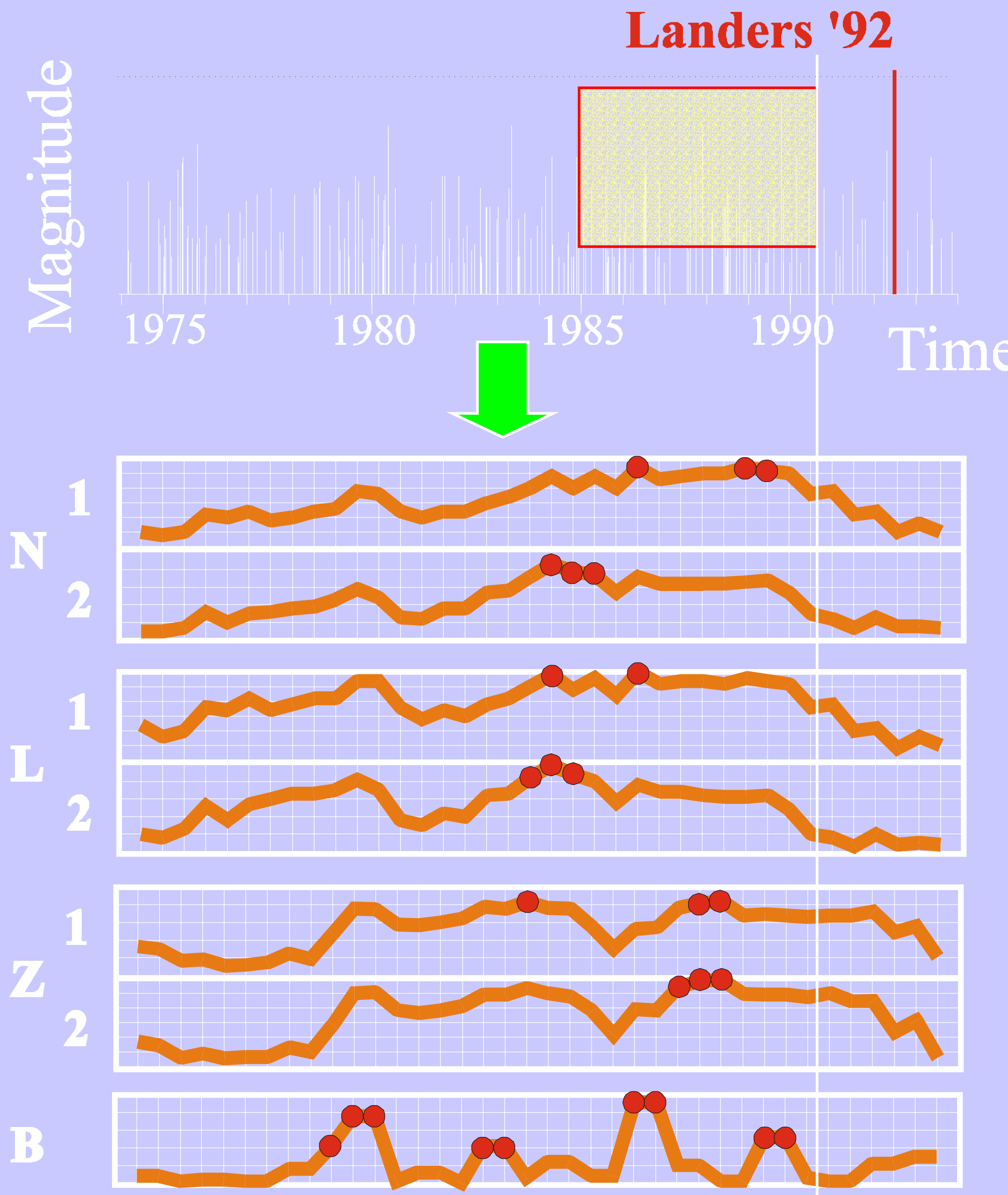
This intermediate-term earthquake prediction method was designed by retroactive analysis of dynamics of seismic activity preceding the greatest, magnitude 8.0 or more, earthquakes worldwide, hence its name.

Its prototype (***Keilis-Borok and Kossobokov, 1984***) and the original version (***Keilis-Borok and Kossobokov, 1987***) were tested retroactively at 143 points, of which 132 are recorded epicenters of earthquakes of magnitude 8.0 or greater from 1857-1983.

The algorithm is based on a simple physical scheme…

# General scheme

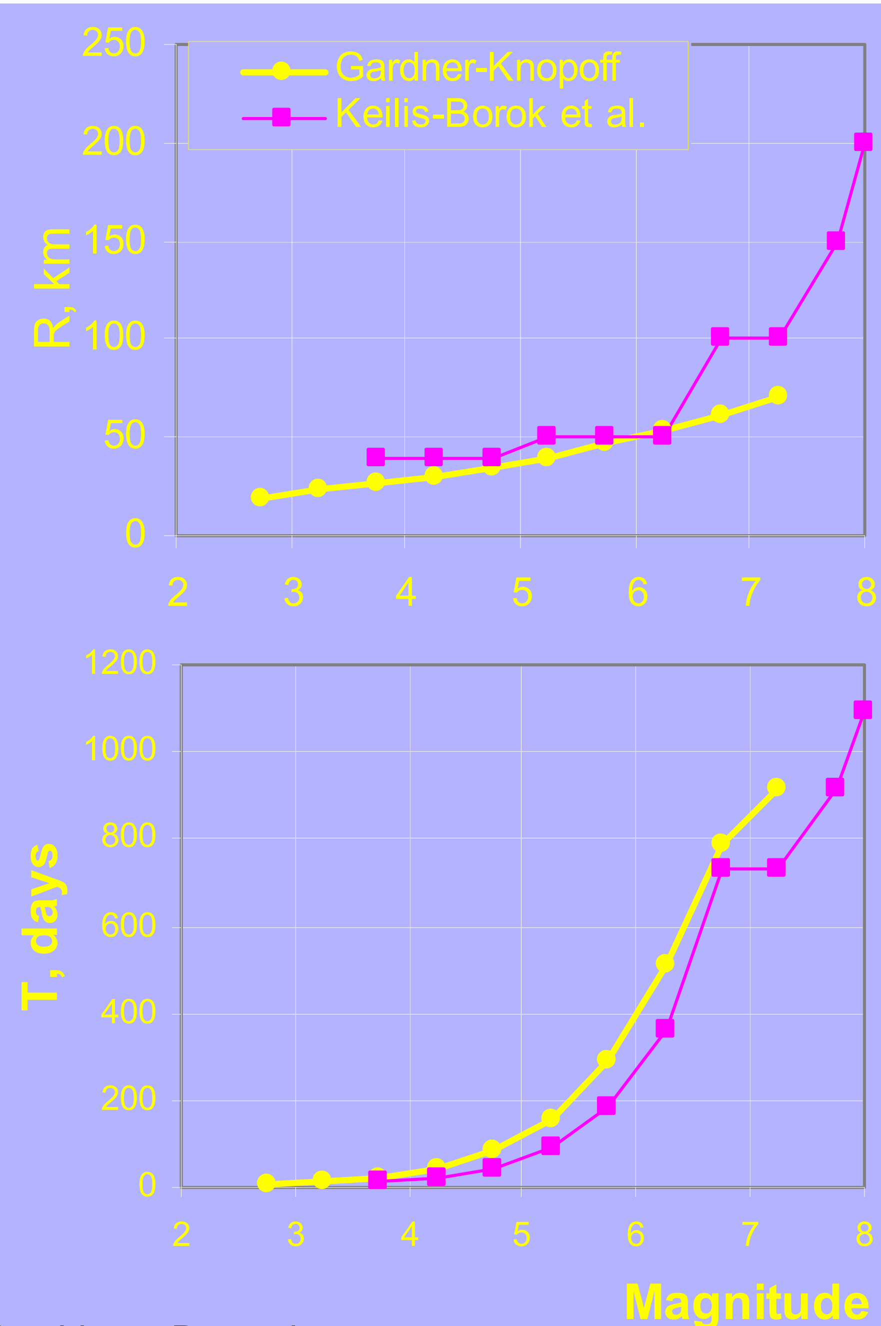


# Data

- Catalog of main shocks: $\{t_i, m_i, h_i, b_i(e)\}$, $i = 1, 2 \ldots$

Here $t_i$ is the origin time, $t_i \le t_{i+1}$; $m_i$ is the magnitude, $h_i$ is focal depth, and $b_i(e)$ is the number of aftershocks with magnitude $M_{aft}$ or more during the first $e$ days.



# Prediction aimed at magnitude $M_0$

- Prediction is aimed at earthquakes of magnitude $M_0$ and larger from the range $M_0+ = [M_0, M_0+\Delta M]$ (where $\Delta M < 1$). Magnitude scale should reflect the size of earthquake sources (accordingly, MS usually taken for larger magnitudes, while mb is used for smaller ones).
- If the data permits, we set different $M_0+$ with a step 0.5.
- Overlapping circles, with the diameter

  $D(M_0) = ( \mathit{exp}(M_0 - 5.6)+1 )^0$ in degrees of the Earth meridian,

  scan the seismic region under study.
- The sequence is normalized by the lower magnitude cutoff m = $M_{min}(\tilde{N})$, $\tilde{N}$ being the standard value of the average annual number of earthquakes in the sequence.

# Trailing averages

- Several running averages are computed for this sequence in the trailing time window (t - s, t) and magnitude range $M_0 > M_i \geq m$.
- They depict different measures of intensity in earthquake flow, its deviation from the long-term trend, and clustering of earthquakes.

The averages include:

# Rate and acceleration of activity

$N(t \mid m, s)$ - the number of earthquakes

with $M \geq m$ in time interval from (t-s) to t, i.e., the number of events of certain size per unit time,

***rate of activity*.**

$L(t \mid m, s, t_0)$ - deviation of activity from a longer-term trend over the period from $t_0$ to t:

$$L(t \mid m, s, t_0) =$$

$$N(t \mid m,s) - N(t \mid m,t\text{-}s\text{-}t_0) \times s/(t\text{-}s\text{-}t_0)$$

***i.e. differential of the rate of activity***

# Linear concentration of main shocks

$Z(t) = Z(t \mid m,M',s,\beta,\gamma) = \Sigma 10^{\beta M_i}/N^{\gamma}$ is a linear concentration of the main shocks {i} from the magnitude range $m \le M_i < M'$ and interval $t - s \le t_i < t$ estimated as the ratio of the average diameter of the source, $l \sim \Sigma 10^{\beta(M_i - \alpha)}/N$ **(**when $\beta$**=0.46)**, to the average distance, $r \sim N^{1/3}$, between them **(**that implies $\gamma$ **= 2/3)**

# Characteristic of clustering

$B(t \mid m, M', s, m_{aft}, e) = \max b_i(e, m_{aft})$ is the maximum calculated over the main shocks with $m \leq M_i < M'$ and time interval $(t-s,t)$.

# Vector of description

- Each of the functions N, L, Z is calculated twice for $m = M_{min}(\tilde{N})$, $\tilde{N} = 20$ and $\tilde{N} = 10$.
- As a result, the earthquake sequence is given a robust averaged description by seven functions: N, L, Z (twice each), and B –

N1, N2, L1, L2, Z1, Z2, B

# Criterion – abnormal values

"Very large" values are identified for each function using the condition that they exceed $Q$ percentiles (i.e., they are higher than $Q$ percent of the encountered values).

That is another local normalization of function values according to the natural empirical distribution.

# Rules for issuing an alarm

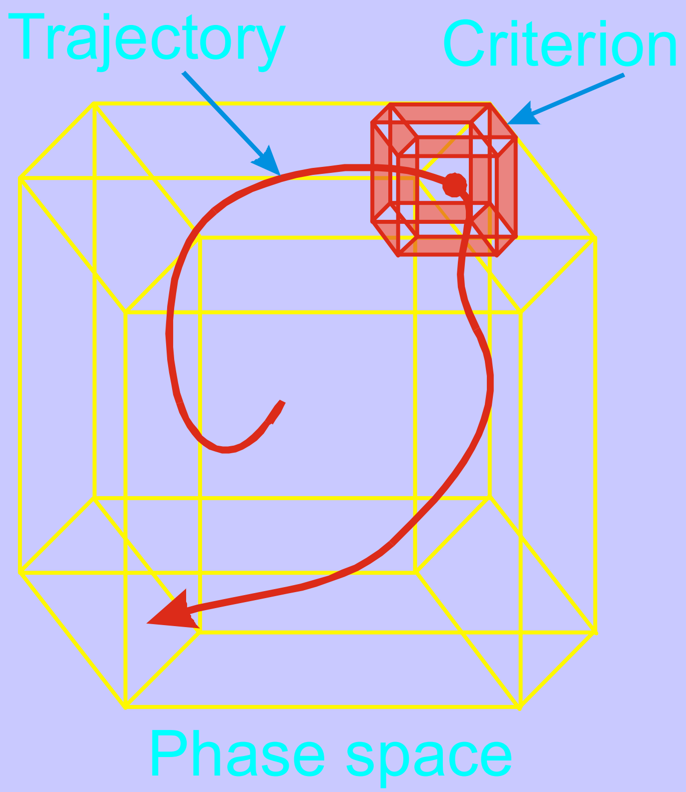
- An alarm or a TIP, "Time of Increased Probability", is declared for five years, when at least six out of seven functions, including B, become "very large" within a narrow time window (t - u, t).
- To stabilize prediction, this criterion is required for two consecutive moments, t and t+0.5 years.

In course of a forward application, the alarm may extend beyond or be terminated before five years in case the updating causes changes in determination of the magnitude cutoffs and/or the percentiles.

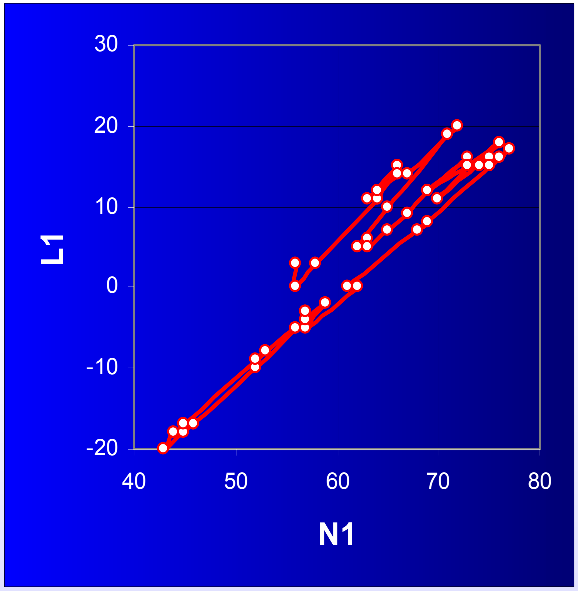
# Standard values of parameters

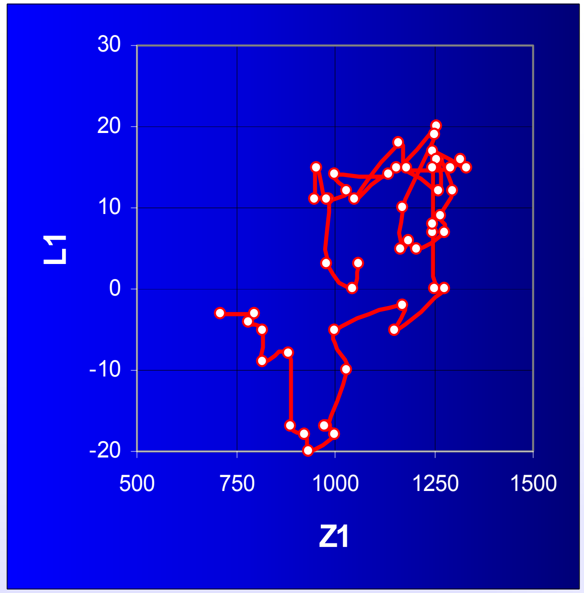
**The following standard values of parameters indicated above are prefixed in the algorithm M8: *D*($M_0$)={exp($M_0$- 5.6)+1}$^0$ in degrees of meridian (*this is 384 km, 560 km, 854 km and 1333 km for $M_0$ = 6.5, 7.0, 7.5 and 8 respectively*), *s* = 6 years, *s'* = 1 year, *g* = 0.5, *p* = 2, *q* = 0.2, *u* = 3 years, $\beta$=0.46, $\gamma$ = 2/3, and *Q* = 75% for B and 90% for the other six functions.**

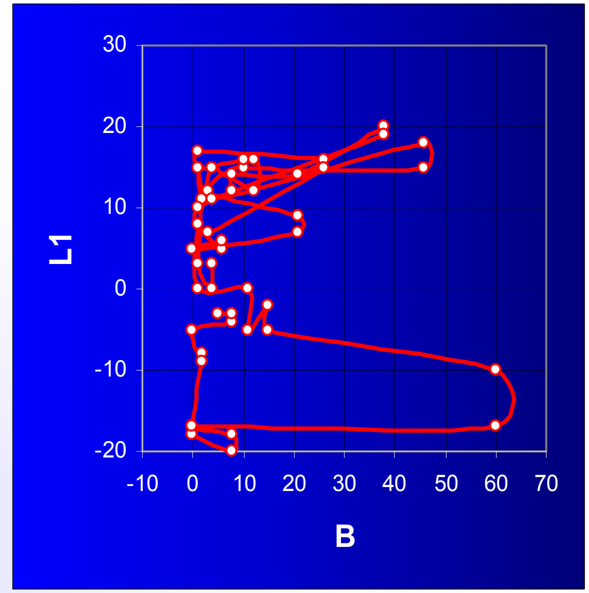
# Criterion in the phase space



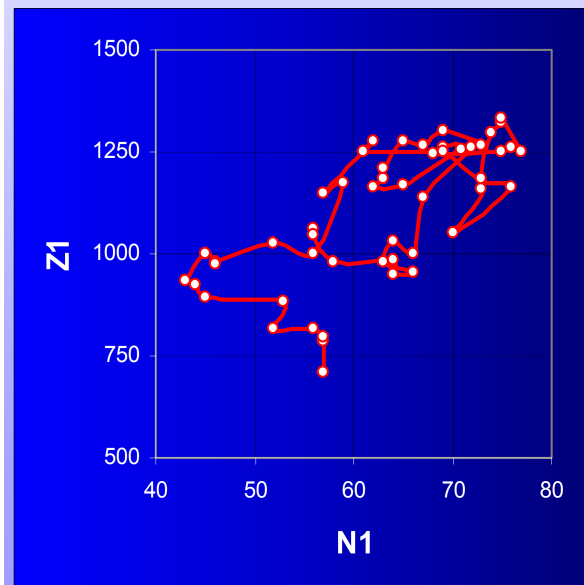

- The algorithm M8 uses traditional description of a dynamical system adding to a common phase space of rate (N) and rate differential (L) dimensionless concentration (Z) and a characteristic measure of clustering (B).
- The algorithm recognizes *criterion*, defined by extreme values of the phase space coordinates, as a vicinity of the system singularity. When a trajectory enters the criterion, probability of extreme event increases to the level sufficient for its effective provision.

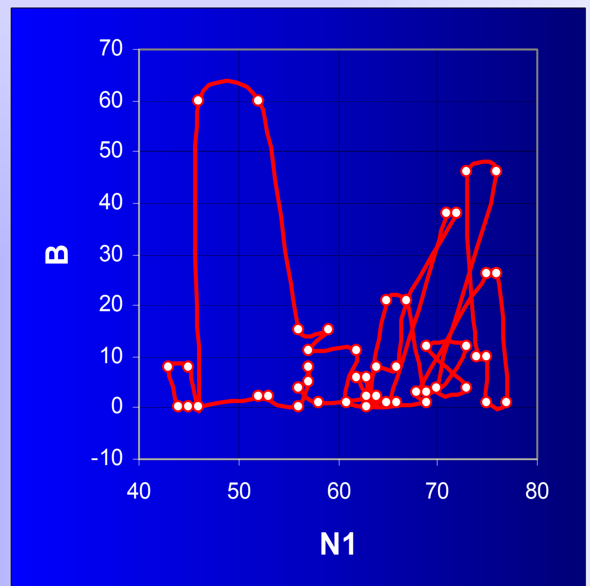


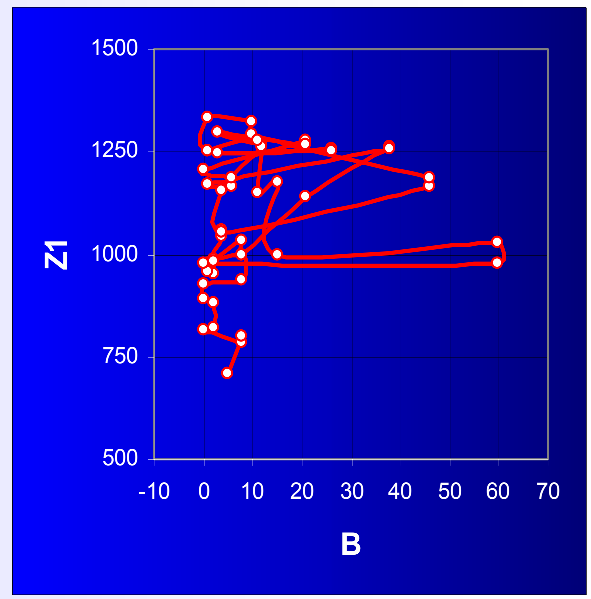




# Trajectory in CI#116, Central California



# M8 algorithm performance

- Retrospectively (*Keilis-Borok and Kossobokov, 1990*) the standard version of the algorithm was applied to predict the largest earthquakes (with $M_0$ ranging from 8.0 to 4.9) in 14 regions.

**25 out of 28 predicted in 16% of the space-time considered.**

- Modified versions in 4 regions of lower seismic activity predicted

**all the 11 largest earthquakes in 26 % of the space-time considered.**

# Worldwide performance of earthquake prediction algorithms M8 and M8-MSc: Magnitude 8.0 or more.

| Test period | Large earthquakes Total | Predicted by M8 | Predicted by M8-MSc | Measure of alarms,% M8 | Measure of alarms,% M8-MSc | Confidence level, % M8 | Confidence level, % M8-MSc |
|---|---|---|---|---|---|---|---|
| 1985-present | 11 | 9 | 7 | 34.9 | 18.0 | 99.80 | 99.90 |
| 1992-present | 9 | 7 | 5 | 30.2 | 15.3 | 99.55 | 99.39 |

The significance level estimates use the most conservative measure of the alarm volume accounting for empirical distribution of epicenters.

**To drive the achieved confidence level below 95%, the Test should encounter four failures-to-predict in a row.**

# Second approximation prediction method

The algorithm for reducing the area of alarm (*Kossobokov, Keilis-Borok, Smith, 1990*) was designed by retroactive analysis of the detailed regional seismic catalog prior to the Eureka earthquake (1980, M=7.2) near Cape Mendocino in California, hence its name abbreviated to MSc.

Qualitatively, the MSc algorithm outlines such an area of the territory of alarm where the activity, from the beginning of seismic inverse cascade recognized by the first approximation prediction algorithm (e.g. by M8), is continuously high and infrequently drops for a short time. Such an alternation of activity must have a sufficient temporal and/or spatial span.

The phenomenon, which is used in the MSc algorithm, might reflect the second (possibly, shorter-term and, definitely, narrow-range) stage of the premonitory rise of seismic activity near the incipient source of main shock.

# Given a TIP…

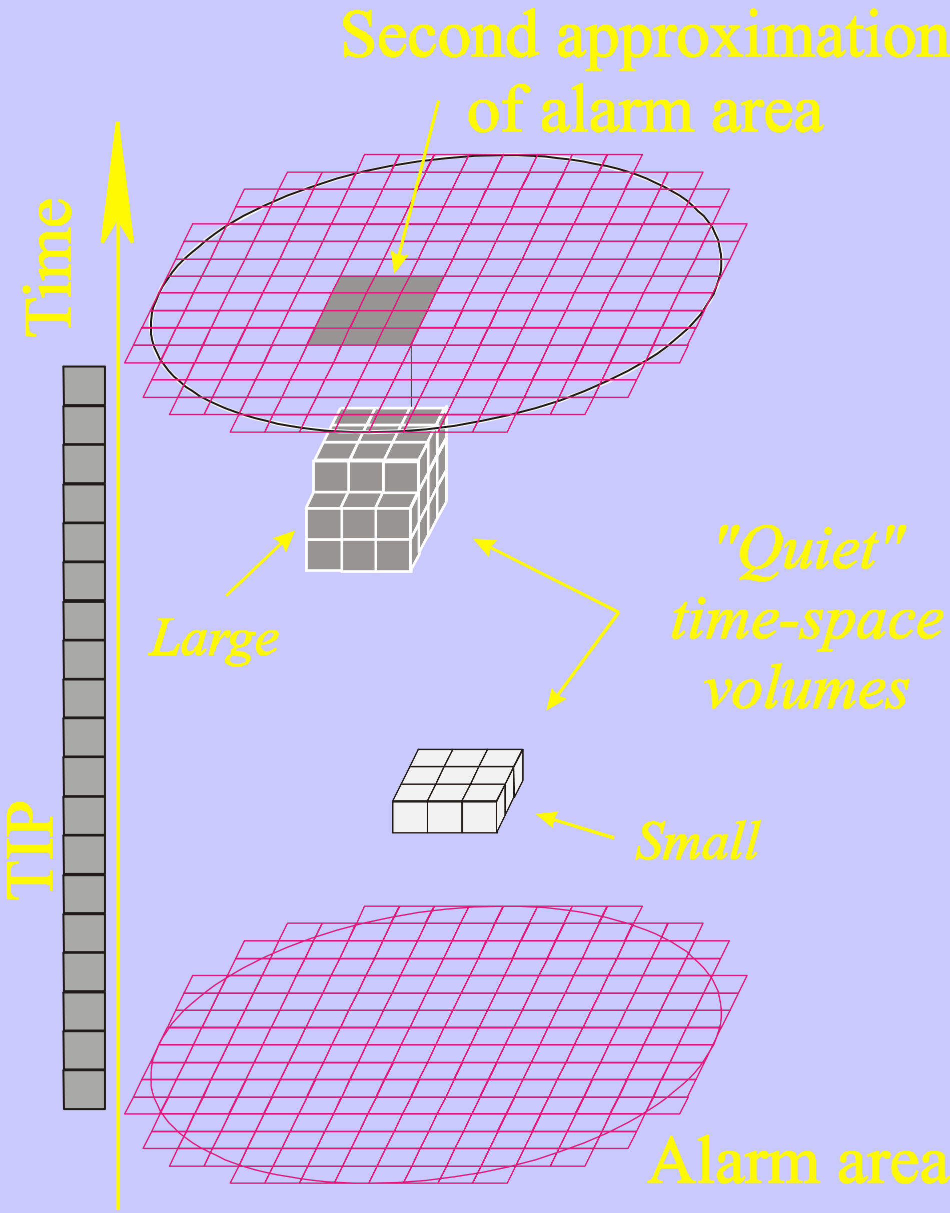
- Given a TIP diagnosed for a certain territory U at the moment T, the algorithm attempts to find within U a *smaller* area V where the predicted earthquake can be expected.
- The algorithm requires a reasonably complete catalog of earthquakes with magnitudes $M \geq (M_0 - 4)$, which is lower than the minimal threshold usually used by M8.

# The essence of MSc

- Territory U is coarse-grained into small squares of $s \times s$ size. Let (i,j) be the coordinates of the centers of the squares.
- Within each square (i,j) the number of earthquakes $n_{ij}(k)$, aftershocks included, is calculated for consecutive, short time windows, *u* months long, starting from the time $t_0$ = (T-6 years) onward, to allow for the earthquakes, which contributed to the TIP's diagnosis; here k is the sequence number of a trailing time window.

# The essence of MSc (cont.)

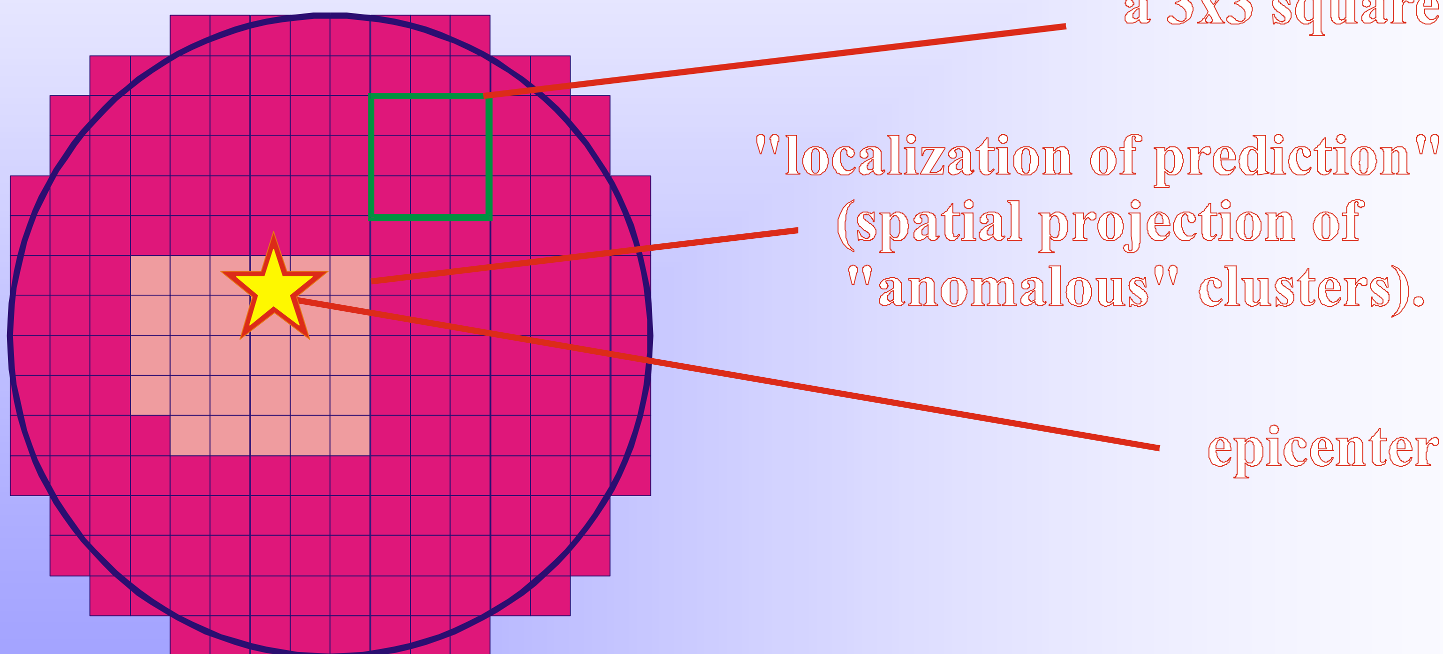
- Finally, the time-space considered is divided into small boxes (i,j,k) of the size ($s \times s \times u$).
- "*Quiet" boxes* are singled out for each small square (i,j); they are defined by the condition that $n_{ij}(k)$ is below the *Q* percentile of $n_{ij}$.
- The clusters of *q* or more quiet boxes connected in space or in time are identified.
- Area V is the territorial projection of these clusters.




The prediction is localized to a spatial projection of all recent "sufficiently large" clusters of squares being in state of "anomalous quiescence".

"Anomalous quiescence" suggests high level of seismic activity during formation of a TIP and after its declaration. "Sufficiently large" size of clusters suggests large scale correlations in recent seismicity.

Eureka 1980, M7.2 earthquake

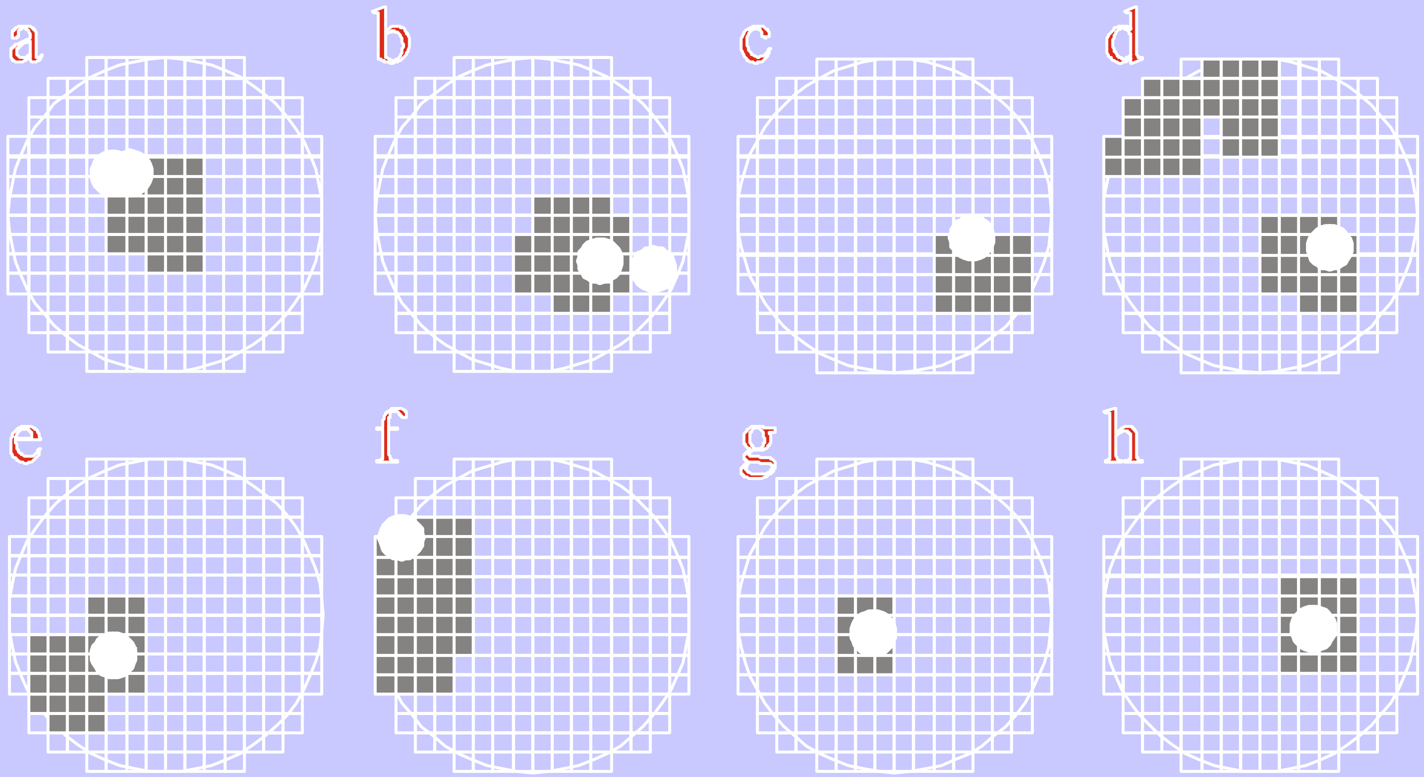


# The standard version of MSc

- *The standard values of parameters* adjusted for the case of the 1980 Eureka earthquake are as follows:

$u$ = 2 months, $Q$ = 10%, $q$ = 4, and $s = 3D/16$, $D$ being the diameter of the circle used in algorithm M8.

(a) Santa Cruz Is, 11/28/1985 & 12/21/1985; (b) New Guinea, 02/08/1987 & 10/16/1987; (c) Costa Rica, 04/22/1991; (d) Landers, CA, 06/28/1992; (e) Guam, 08/08/1993; (f) Fiji, 03/09/1994; (g) Shikotan Is, 10/04/1994; (h) Samoa, 04/07/1995.



# MSc vs. Activity

MSc outscores simple alternatives of narrowing down the area of first approximation alarm –

- Nonempty Cells (NeC);
- Most Active Cells (MAC) that contain (a) 1/8, (b) 1/4, (c) 1/3 of the recent seismic activity.

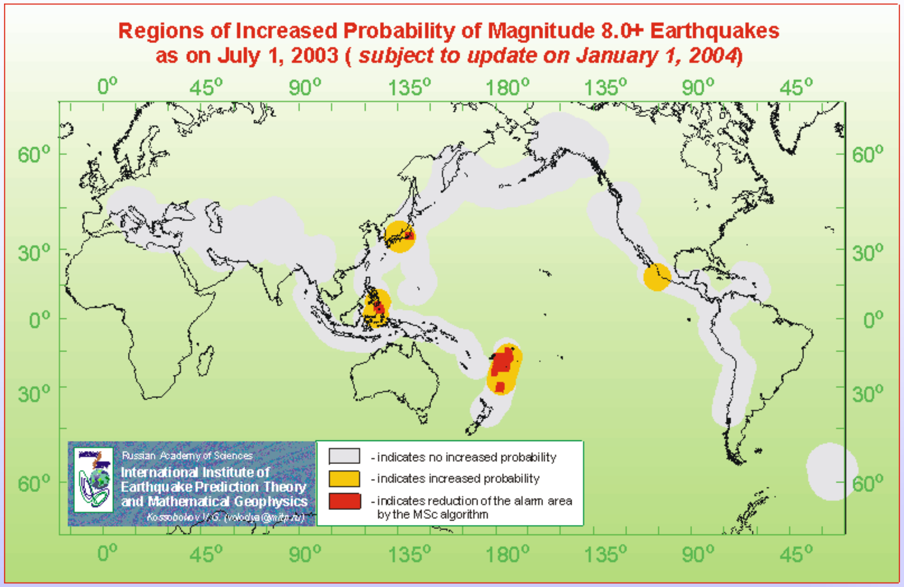
The same number of correct localizations, as obtained with MSc, is achieved also by MAC(1/3), which narrows down the alarm area to 28%, while MSc outperforms it with 14%.

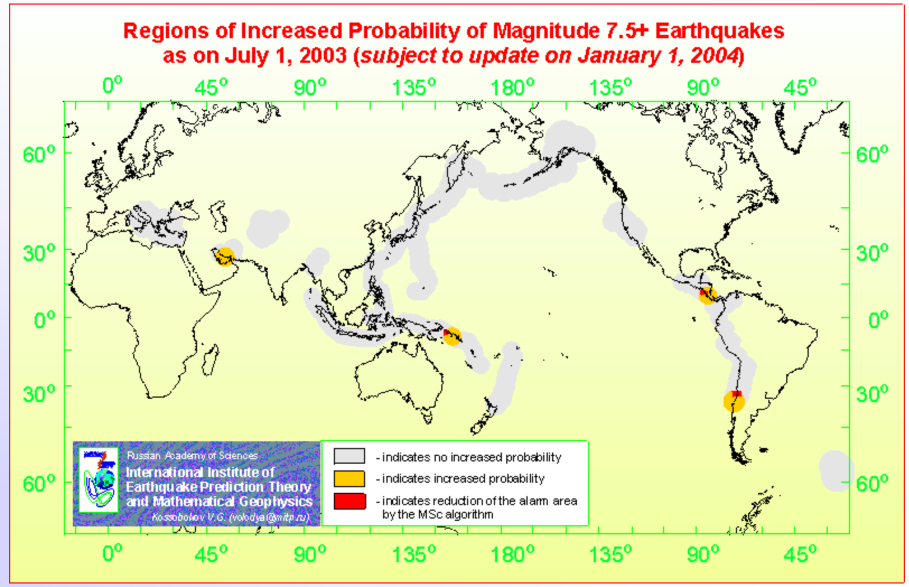
By 1992 all the components necessary for reproducible real-time prediction, i.e., an unambiguous definition of the algorithms and the data base, were specified in publications

- Algorithm M8 (*Keilis-Borok and Kossobokov, 1984, 1987, 1990*) was designed by retroactive analysis of seismic dynamics preceding the greatest (M≥8) earthquakes worldwide, as well as the MSc algorithm for reducing the area of alarm (*Kossobokov,Keilis-Borok, Smith, 1990*)
- The National Earthquake Information Center Global Hypocenters Data Base (*US GS/NEIC GHDB, 1989*) is sufficiently complete since 1963.
- This allowed a systematic application of M8 and MSc algorithm since 1985.

# Real-time prediction of the world largest earthquakes

( http://www.mitp.ru or http://www.phys.ualberta.ca/mirrors/mitp )




Although the M8-MSc predictions are intermediate-term middle-range and by no means imply any "red alert", some colleagues have expressed a legitimate concern about maintaining necessary confidentiality. Therefore, the up-to-date predictions are not shown here, although available on web-pages of restricted access provided to about 125 members of the Mailing List.

# Worldwide performance of earthquake prediction algorithms M8 and M8-MSc: Magnitude 8.0 or more.

| Test period | Large earthquakes: Total | Predicted by M8 | Predicted by M8-MSc | Measure of alarms,%: M8 | Measure of alarms,%: M8-MSc | Confidence level, %: M8 | Confidence level, %: M8-MSc |
|---|---|---|---|---|---|---|---|
| 1985-present | 11 | 9 | 7 | 34.9 | 18.0 | 99.80 | 99.90 |
| 1992-present | 9 | 7 | 5 | 30.2 | 15.3 | 99.55 | 99.39 |

The significance level estimates use the most conservative measure of the alarm volume accounting for empirical distribution of epicenters.

**To drive the achieved confidence level below 95%, the Test should encounter four failures-to-predict in a row.**

# Worldwide performance of earthquake prediction algorithms M8 and M8-MSc: Magnitude 7.5 or more.

| Test period | Large earthquakes Total | Predicted by M8 | Predicted by M8-MSc | Measure of alarms,% M8 | Measure of alarms,% M8-MSc | Confidence level, % M8 | Confidence level, % M8-MSc |
|---|---|---|---|---|---|---|---|
| 1985-present | 43 | 27 | 16 | 42.0 | 12.0 | 99.52 | 99.99 |
| 1992-present | 31 | 16 | 9 | 40.0 | 12.6 | 87.16 | 98.81 |

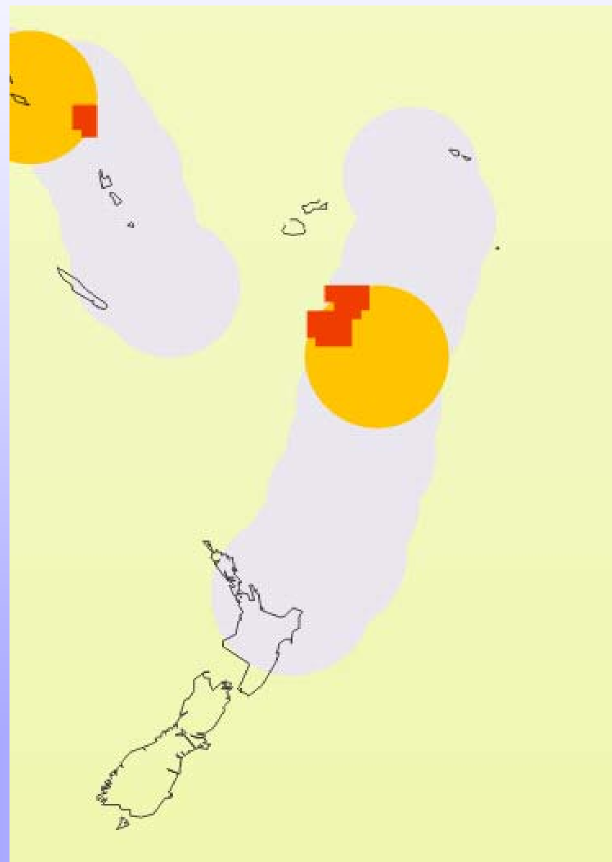
The significance level estimates use the most conservative measure of the alarm volume accounting for empirical distribution of epicenters.

**The prediction for M7.5+ is less effective than for M8.0+.**

**Nevertheless, we continue testing the algorithms for this and smaller magnitude ranges.**

# What was predicted…

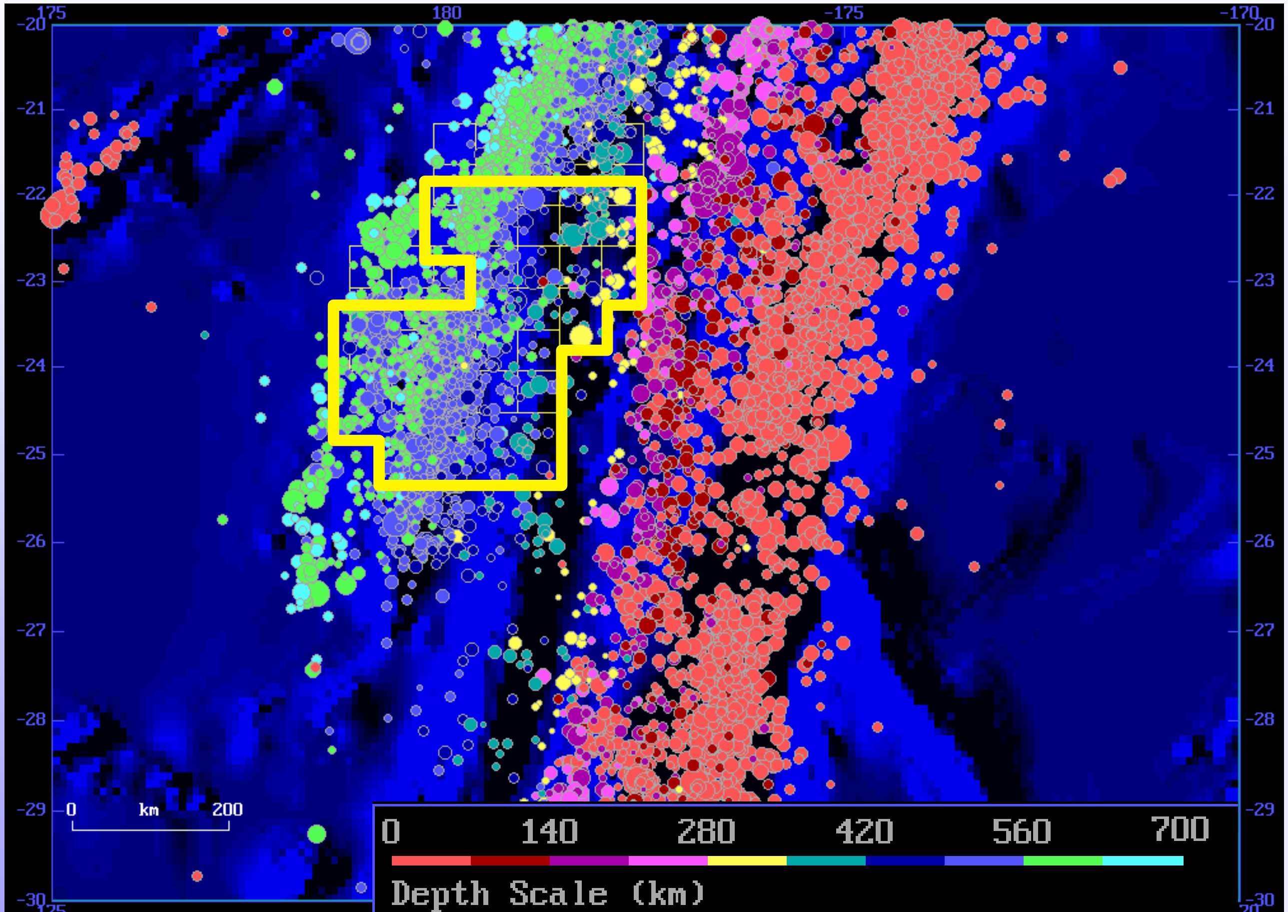
**Sent on Monday, July 15, 2002 (Subject: The 2002b Update of the M8-MSc predictions) along with the updated predictions of major earthquakes worldwide.**



- Earthquake(s) with magnitude 7.5 or more will occur in CI #5 (yellow) during the time period from July 2002 through July 2003.
- In the second approximation the MSc algorithm has identified the area (red) that stretch between

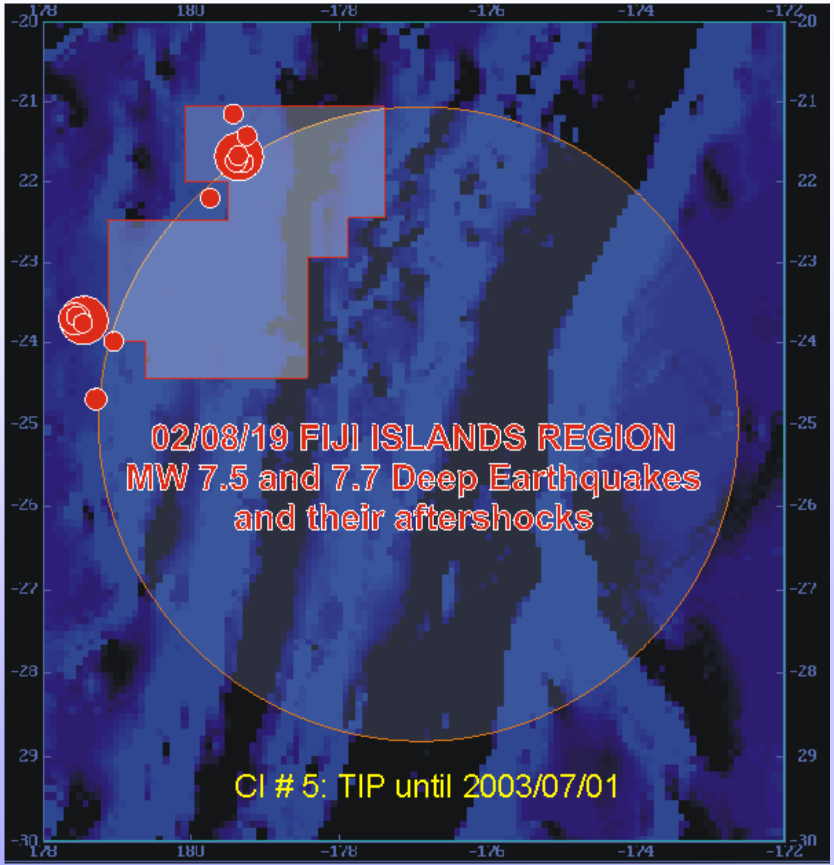
  **24.52S - 21.16S and 178.76E - 177.53W.**

# What was predicted…




- The position of the M8-MSc alarm that narrow down substantially the prediction area **suggested the occurrence of the great deep earthquakes** (**depth of about 240-700 km**).

# What happened…




- EARTHQUAKES:
  - Origin times -
    - 2002/08/19 11:01:01
    - 2002/08/19 11:08:25 ;
  - Coordinates –
    - 21.80S 179.49W
    - 23.85S 178.41E;
  - Depths - 586.8 and 693.7 km;
  - Magnitudes –
    - MwGS (MeGS)
    - 7.5 and 7.7 (7.7 and 7.4);
  - F-E Regions –

FIJI ISLANDS REGION and
SOUTH OF FIJI ISLANDS.

The two August 19 main shocks mark both northern and southern edges of the prediction area. Does it mean that sometimes exact prediction is not possible? This reduction of the uncertainty provides probability gain of more than 25.

# Conclusions – The Four Paradigms

Statistical validity of predictions confirms the underlying paradigms:

- Seismic premonitory patterns exist;
- Formation of earthquake precursors at scale of years involves large size fault system;
- The phenomena are similar in a wide range of tectonic environment…
- … and in other complex non-linear systems.

# Conclusions – Seismic Roulette is not perfect

Are these predictions useful?

- Yes, if used in a knowledgeable way.
- Their accuracy is already enough for undertaking earthquake preparedness measures, which would prevent a considerable part of damage and human loss, although far from the total.
- The methodology linking prediction with disaster management strategies does exist (*Molchan, 1997*).

# *Kofi Annan:*

**Introduction to Secretary-General's Annual Report on the Work of the Organization of United Nations, 1999 - A/54/1**

*"More effective prevention strategies would save not only tens of billions of dollars, but save tens of thousands of lives. Funds currently spent on intervention and relief could be devoted to enhancing equitable and sustainable development instead, which would further reduce the risk for war and disaster. Building a culture of prevention is not easy. While the costs of prevention have to be paid in the present, its benefits lie in a distant future. Moreover, the benefits are not tangible; they are the disasters that did NOT happen."*

# Conclusions – Implications for Physics

- The predictions provide reliable empirical constrains for modeling earthquakes and earthquake sequences.
- Evidence that distributed seismic activity is a problem in statistical physics.
- Favor the hypothesis that earthquakes follow a general hierarchical process that proceeds via a sequence of inverse cascades to produce self-similar scaling (*intermediate asymptotic*), which then truncates at the largest scales bursting into direct cascades (*Gabrielov, Newman, Turcotte, 1999*).

# What are the Next Steps?

- The algorithms are neither optimal nor unique (CN, SSE, Vere-Jones "probabilistic" version of M8, etc.). The accuracy could be improved by a systematic monitoring of the alarm areas and by designing a new generation of earthquake prediction technique ("Seismic Reversal" - SR, ROC, Accord, etc.).

... and an obvious general one -

- More data should be analyzed systematically to establish reliable correlations between the occurrence of extreme events and observable phenomena.

# Some References


- *Global Hypocenters Data Base CD-ROM, 1989. NEIC/USGS, Denver, CO. and its PDE and QED updates)*
- *Healy, Kossobokov, and Dewey, U. S. Geol. Surv. OFR 92-401, 1992.*
- *Keilis-Borok, and Kossobokov. Proc. 27th Geological congress, 61, Moscow: Nauka, 56-66, 1984.*
- *Keilis-Borok and Kossobokov. Phys. Earth Planet. Inter.* ***61:****73-83, 1990*
- *Keylis-Borok and Malinovskaya. J. Geophys. Res.* ***69****: 3019-3024, 1964.*
- *Kossobokov. User Manual for M8. In Healy, J.H., Keilis-Borok, V.I., and Lee, W.H.K. (Eds), Algorithms for earthquake statistics and prediction. IASPEI Software Library, 6. Seismol. Soc. Am., El Cerrito, CA, 1997.*
- *Kossobokov, Keilis-Borok and Cheng. Phys. Rev. E, 61: 3529-3533, 2000*
- *Kossobokov, Keilis-Borok, Romashkova, and Healy. Testing earthquake prediction algorithms: Statistically significant real-time prediction of the largest earthquakes in the Circum-Pacific, 1992-1997. Phys. Earth and Planet. Inter.,* ***111****, 3-4: 187-196, 1999.*
- *Kossobokov, Keilis-Borok, and Smith. J. Geophys. Res.,* ***95****: 19763-19772, 1990.*
- *Romashkova and Kossobokov, Comput. Seismol.,* ***32****: 162-189, 2001.*
- *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. Keilis-Borok & Soloviev (Eds) Springer, Heidelberg, 141-207, 2003.*
- *Harte, D., L. Dong-Feng, M. Vreede, D. Vere-Jones, 2003. Quantifying the M8 algorithm: reduction to a single critical variable and stability results. New Zealand Journal of Geology & Geophysics, 2003,* ***46****: 141-152*

*This presentation and research have been supported in part by the Abdus Salam International Centre for Theoretical Physics, International Science and Technology Center (Project 1538-00), by INTAS (grant #0748), and by James S. McDonnell Foundation (the 21st Century Collaborative Activity Award for Studying Complex Systems, project "Understanding and Prediction of Critical Transitions in Complex Systems".*