The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

2060-61

Advanced School on Non-linear Dynamics and Earthquake Prediction

28 September - 10 October, 2009

Earthquake Forecast/Prediction: Problem of Verification, Accuracy and Limitations

Vladimir Kossobokov

International Institute of Earthquake Prediction Theory and Mathematical Geophysics
Moscow
Russia
Institute de Physique du Globe de Paris

volodya@mitp.ru/volodya@ipgp.jussieu.fr

Earthquake Forecast/Prediction:

Problem of Verification, Accuracy and Limitations

IPGP
INSTITUT DE PHYSIQUE
DU GLOBE DE PARIS

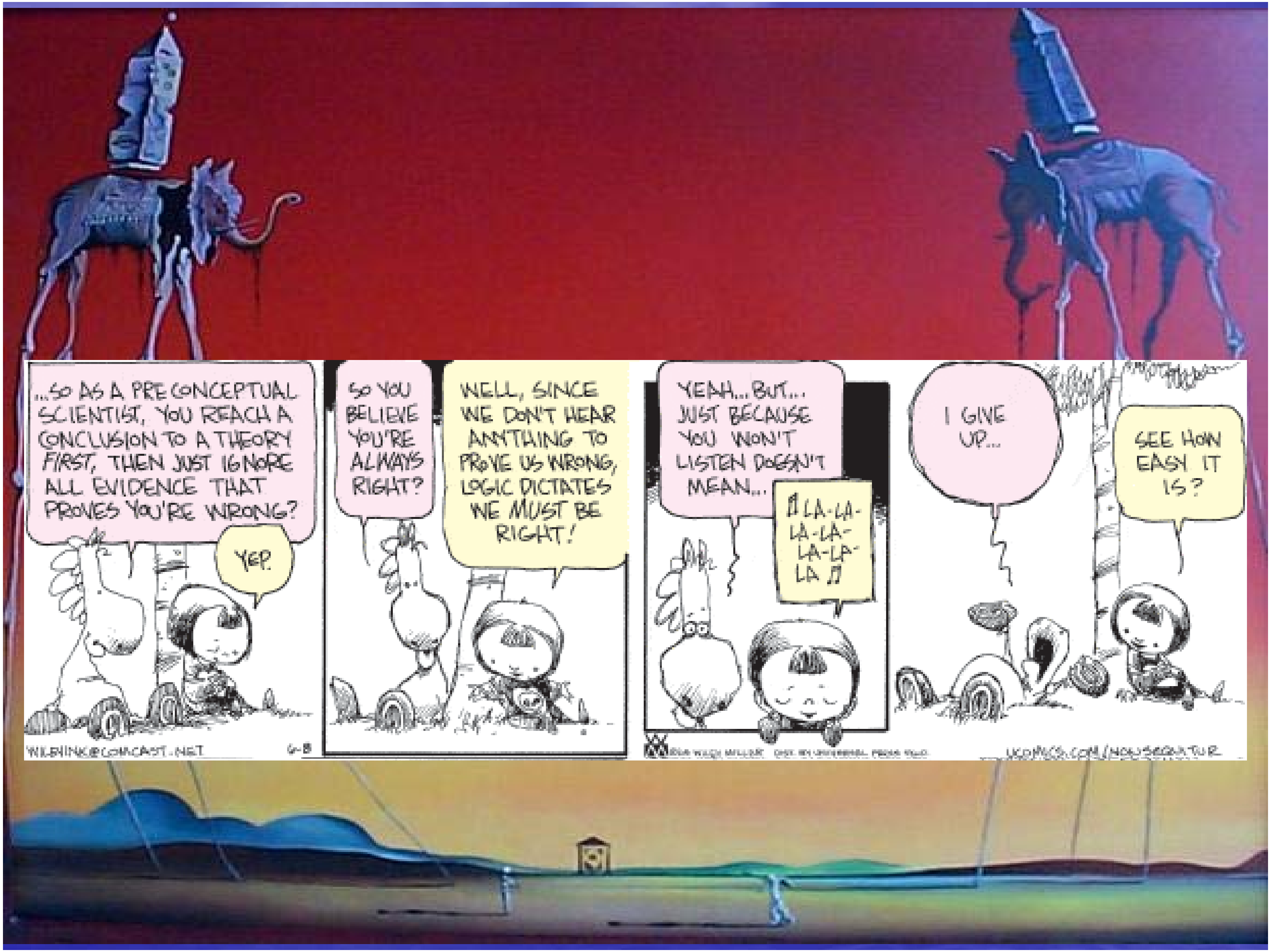
МИТП РАН
IIEPT

Vladimir Kossobokov

International Institute of Earthquake Prediction Theory and Mathematical Geophysics,
Russian Academy of Sciences,
84/32 Profsouznaya Ulitsa, Moscow 117997, 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


...SO AS A PRECONCEPTUAL SCIENTIST, YOU REACH A CONCLUSION TO A THEORY *FIRST*, THEN JUST IGNORE ALL EVIDENCE THAT PROVES YOU'RE WRONG?
YEP.
SO YOU BELIEVE YOU'RE ALWAYS RIGHT?
WELL, SINCE WE DON'T HEAR ANYTHING TO PROVE US WRONG, LOGIC DICTATES WE *MUST* BE RIGHT!
YEAH... BUT... JUST BECAUSE YOU WON'T LISTEN DOESN'T MEAN...
♫ LA-LA-LA-LA-LA-LA-LA ♫
I GIVE UP...
SEE HOW EASY IT IS?
WILEYINK@COMCAST.NET
6-8
UCOMICS.COM/NONSEQUITUR

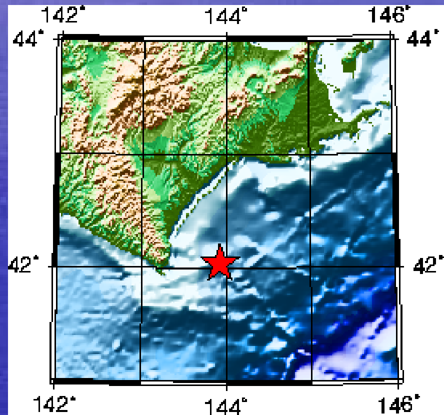
Seismology is juvenile and its appropriate statistical tools to-date may have a "medieval flavor" for those who hurry up to apply a fuzzy language of a highly developed probability theory. To become "quantitatively probabilistic" earthquake forecasts/predictions must be defined with a scientific accuracy. Following the most popular objectivists' viewpoint on probability, we cannot claim "probabilities" adequate without a long series of "yes/no" forecast/prediction outcomes. Without "antiquated binary language" of "yes/no" certainty we cannot judge an outcome ("success/failure"), and, therefore, quantify objectively a forecast/prediction method performance.

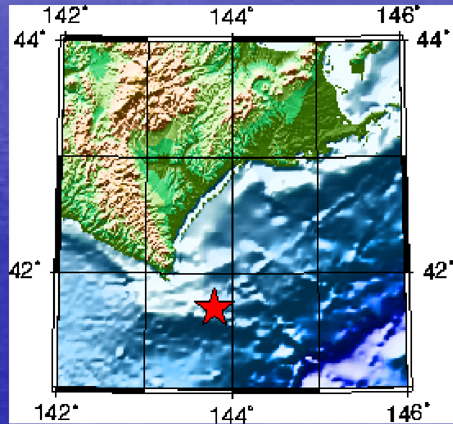
PLANETS ALIGN:
On Wednesday morning, September 24th, 2003 a lovely trio appeared in the eastern sky: **Jupiter, the crescent moon and Mercury…**

Is it a coincidence or a law?

Two days later …

防災科研Hi-net暫定処理による震源位置

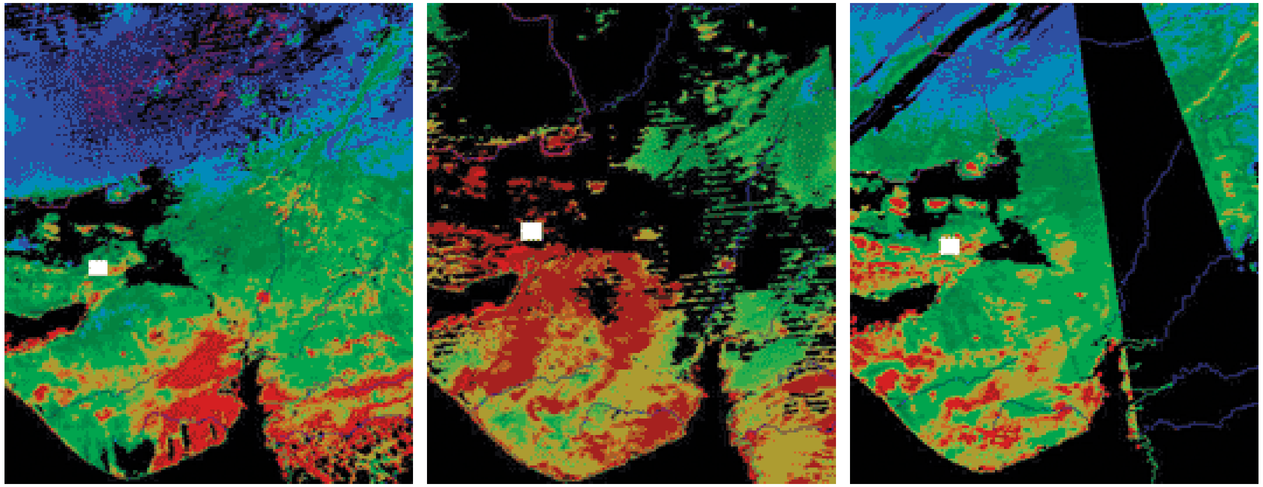




本震
2003年9月26日04時50分11秒
北緯42.0度
東経143.9度
深さ25km
M7.7

最大余震
2003年9月26日06時08分03秒
北緯41.8度
東経143.9度
深さ35km
M7.4

Isn't it a coincidence ?



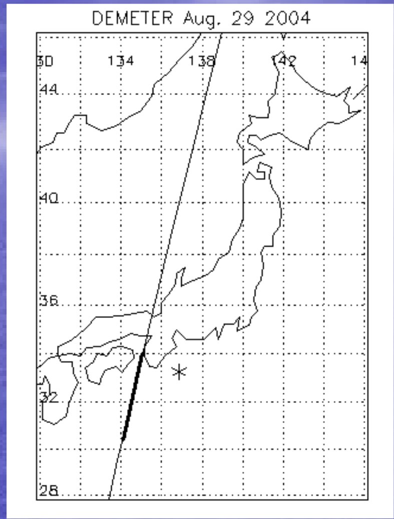
WARM BEFORE THE STORM: An earthquake killed more than 20 000 people on 26 January 2001 in the Indian state of Gujarat. NASA's Terra satellite made infrared maps of the region on 6, 21, and 28 January [left to right]. Five days before the earthquake [middle], the area near the epicenter [white square] gave off an unusual amount of infrared radiation [red]. Just two days after the quake [right], the radiation was gone.

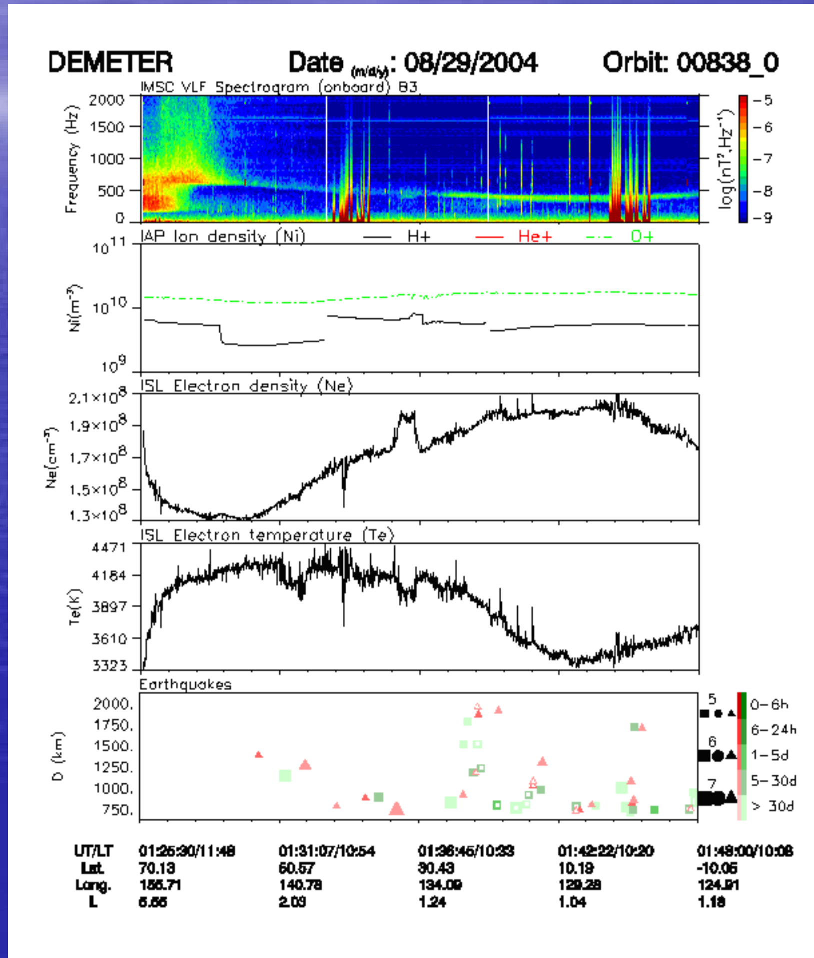
IMAGES: NASA

Isn't it a coincidence ?

"Orbit of DEMETER above Japon on August 29, 2004. The star indicates the epicenter of an earthquake of magnitude 7.1 which will occur on September 5, 2004 in the region of Kii-Peninsula (Lat=33.05N, Long=136.78E).

The thick line on the orbit corresponds to the period where an ionospheric perturbation is observed with DEMETER (next Figure)."





"From the top to the bottom the panels successively show a spectrogram of a magnetic component between 0 and 2 kHz, the ion density given by IAP, the electron density and temperature, and the earthquakes seen by DEMETER along the orbit. In this last panel, the Y-coordinate gives the distance between DEMETER and the earthquake hypocenter. The red color of the symbol which size is proportional to the magnitude shows that DEMETER is flying over the region before the earthquake. A large variation of the ionosphere parameters is observed when the satellite is above the seismic zone (in the top panel, the two bursts of interferences correspond to periods where wheels, which are used for the satellite attitude control, are desaturated)."

The explosive eruption of Asama volcano on September 01, which ash-fall covered a narrow elongated area reaching ca 250 km to Pacific Ocean seems a better alternative than the two earthquakes of M7.2 and M7.4 on September 05, 2004 in Japan, doesn’t it ?

“Earthquakes are so complicated that we must apply some Statistics.”

Seismic Roulette null-hypothesis

Consider a roulette wheel with as many sectors as the number of events in a sample catalog, a sector per each event.

- Make your bet according to prediction: determine, which events are inside area of alarm, and put one chip in each of the corresponding sectors.
- Nature turns the wheel.
- If seismic roulette is not perfect…

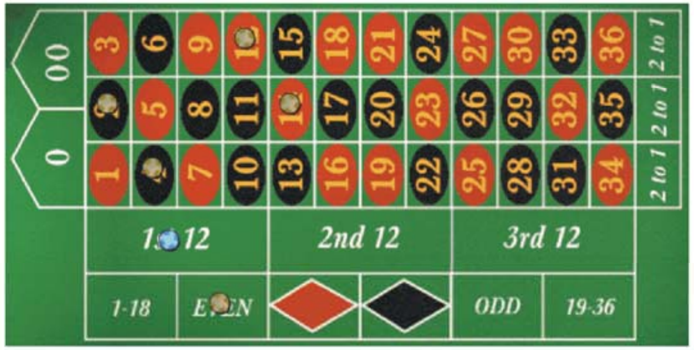
then ***systematically*** you can win! ☺

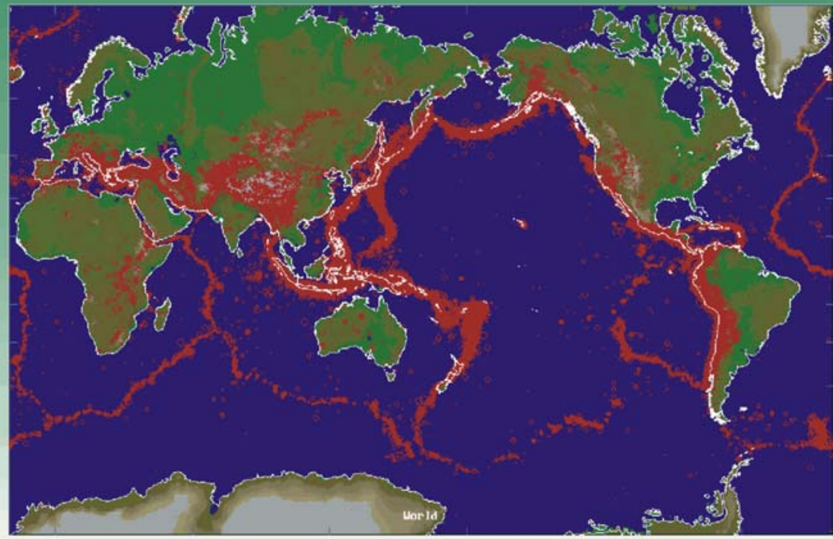
or lose … ☹

If you are smart enough to know "antipodal strategy" (Molchan, 1994; 2003), make the predictions efficient ------

and your wins will outscore the losses! ☺ ☺ ☹ ☺ ☺ ☺ ☹ ☺ ☺ ☺

Seismic Roulette





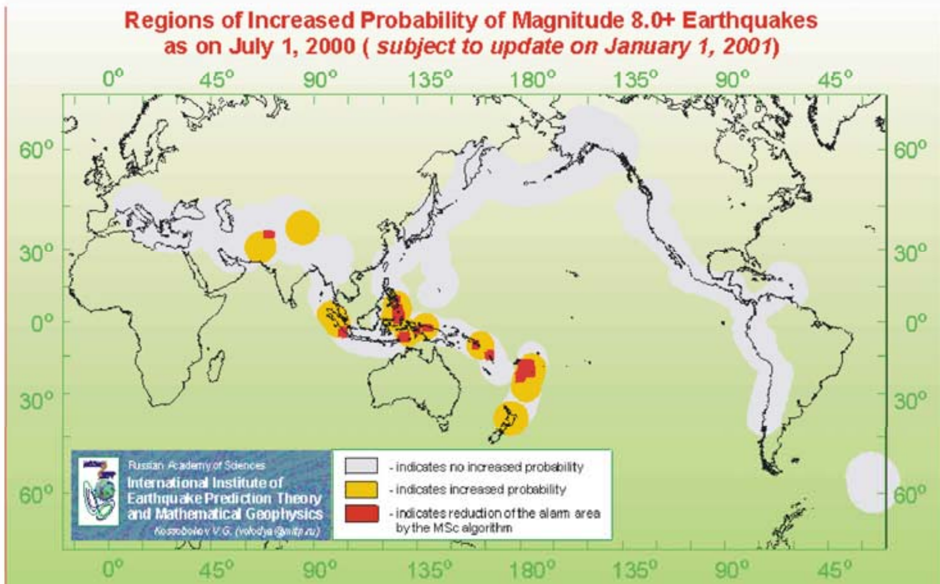
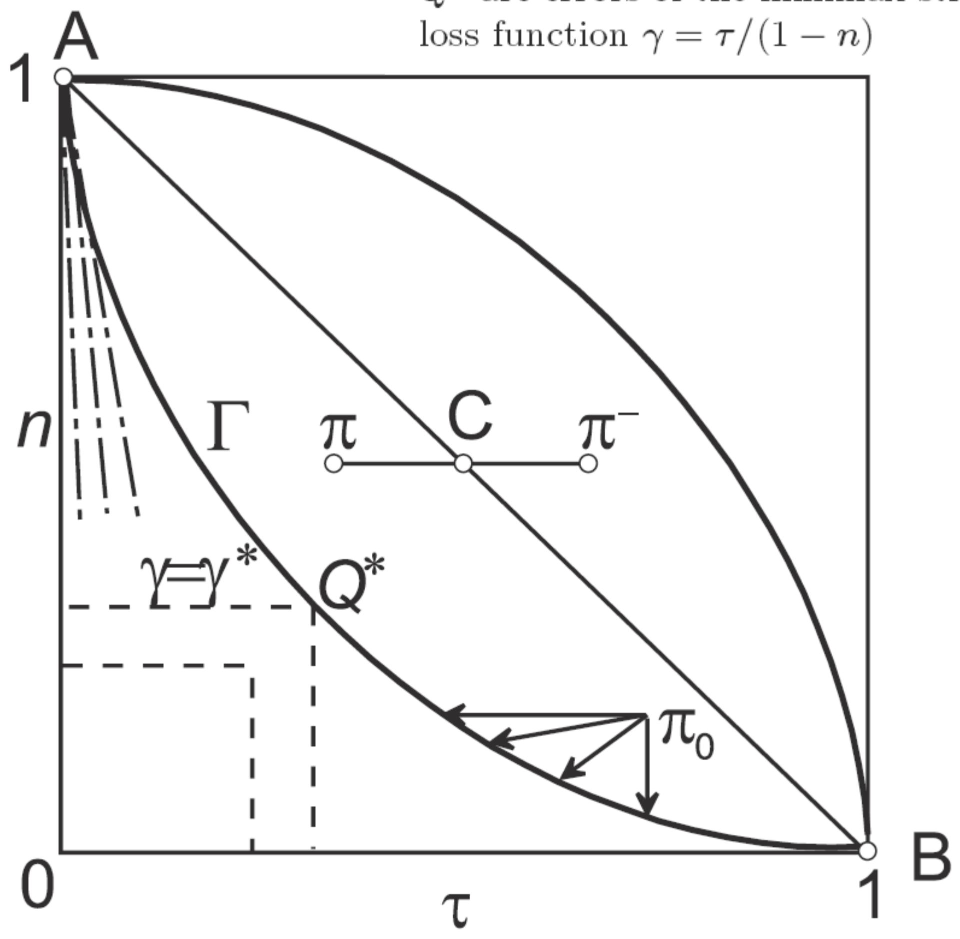


Fig. 5.1. Error set $\mathcal{E}(J)$ for prediction strategies based on a fixed type of information J. Point A corresponds to an optimistic strategy, point B to a pessimistic strategy, and the interval AB corresponds to strategies of random guess. C is the center of symmetry of $\mathcal{E}(J)$. π and π^- are a strategy and its antipodal strategy. Γ is the error diagram of optimal strategies. *Arrows* indicate a better forecast relative to the strategy π_0. *Dashed lines* are contours of the loss function $\gamma = \max(n, \tau)$. Q^* are errors of the minimax strategy, $n = \tau$. *Dash-dotted* lines are contours of the loss function $\gamma = \tau/(1-n)$

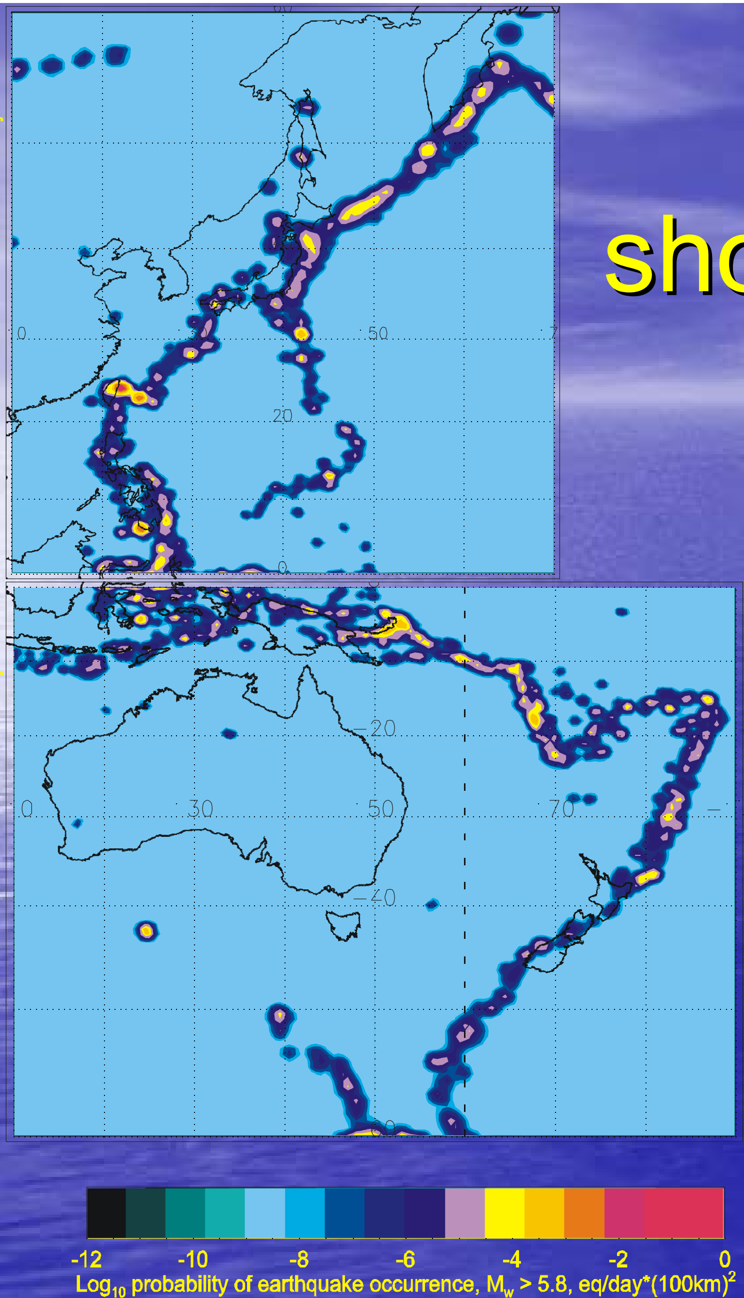


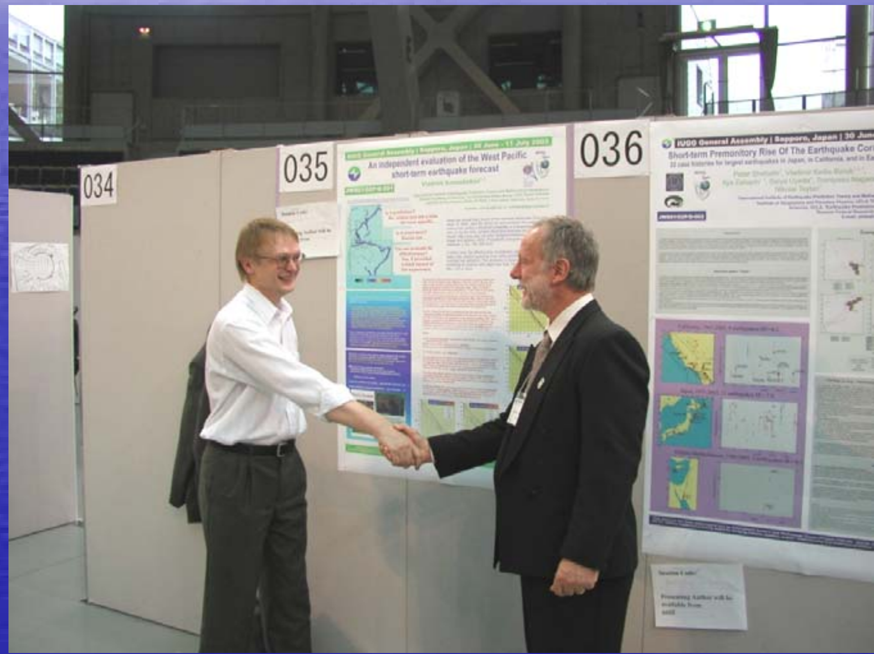
Error diagram

Molchan, G.M. Earthquake Prediction as Decision-making Problem. Pure Appl. Geoph, **149**, 233-247, 1997.

Molchan, G.M. 5. Earthquake Prediction Strategies: a theoretical analysis. In: Keilis-Borok, V.I., and A.A. Soloviev, (Editors). *Nonlinear Dynamics of the Lithosphere and Earthquake Prediction*. Springer, Heidelberg, 208-237, 2003.

West Pacific short-term forecast





Jackson and Kagan "Testable earthquake forecasts for 1999", Seism. Res. Lett., 70, 393-403, 1999

Kagan and Jackson (2000) "Probabilistic forecasting of earthquakes", Geophys. J. Int., 143, 438-453

We have analyzed the predictions arising from setting a threshold probability or a threshold probability ratio on top the daily updated Short-term forecasts for NW and SW Pacific in April 2002 - September 2004 (*http://scec.ess.ucla.edu/~ykagan/predictions_index.html*; *Kagan and Jackson, 2000. Probabilistic forecasting of earthquakes, Geophys. J. Int., 143, 438-453*) and the catalog of earthquakes for the same period and have come to the following conclusion:

The predictions based on the Yan Y. Kagan and David D. Jackson forecasts are hardly better than random guessing, when main shocks are considered, and could be used for effective prediction of aftershocks only.

The conclusion is based on the prediction outcome achieved for 218 shallow (with depth less than 70 km) earthquakes of MwHRV = 5.8 or more. According to the definition from (*Keilis-Borok et al., 1980*), there are 67 aftershocks and 151 main shocks.

The territory of West Pacific short-term forecast is coarse-grained into cells, 0.5 by 0.5 degree each. Making a “bet” on a cell C, we pay n(C), which is the number of earthquakes from the sample catalog. Each target earthquake E defines the threshold value - p(E) (or p/P(E)) - being the value of short-term probability p (or the value of probability ratio p/P) determined in advance for the day of the earthquake.
In its turn the threshold defines the minimal cost of a bet required for successful prediction of the target earthquake, N(E), which is the sum of all bets n(C) over the union of cells with p equal or above p(E) (same for the ratio p/P). The track record of the experiment provides the set of bets {N(E)} associated with target earthquakes that happened.

Denote μ being the bet sum normalized to the total sum of n(C) and ν being the number of failures-to-predict normalized to the total number of target earthquakes that happened in the course of testing. The ν vs. μ diagram characterize the effectiveness of the prediction method, e.g., random prediction performance is associated with the diagonal that connects "optimist's" {1,0} and "pessimist's" {0,1} strategies (*Molchan, G. M.. Earthquake Prediction as a Decision-making Problem, Pure Appl. Geophys., 149, 233-247, 1997*).

Given -

(1) the track record of the West Pacific short-term forecasts in the period from April 10, 2002 to September 13, 2004;

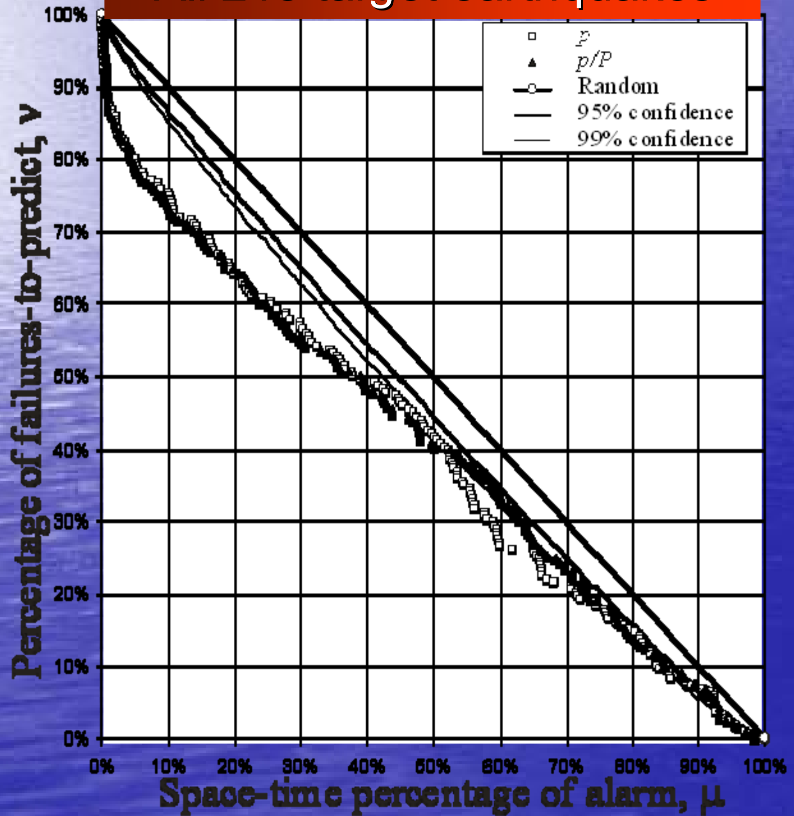
(2) the Harvard CMT catalog for the same period of time;

(3) the counts of n(C) based on the NEIC catalog of shallow earthquakes -

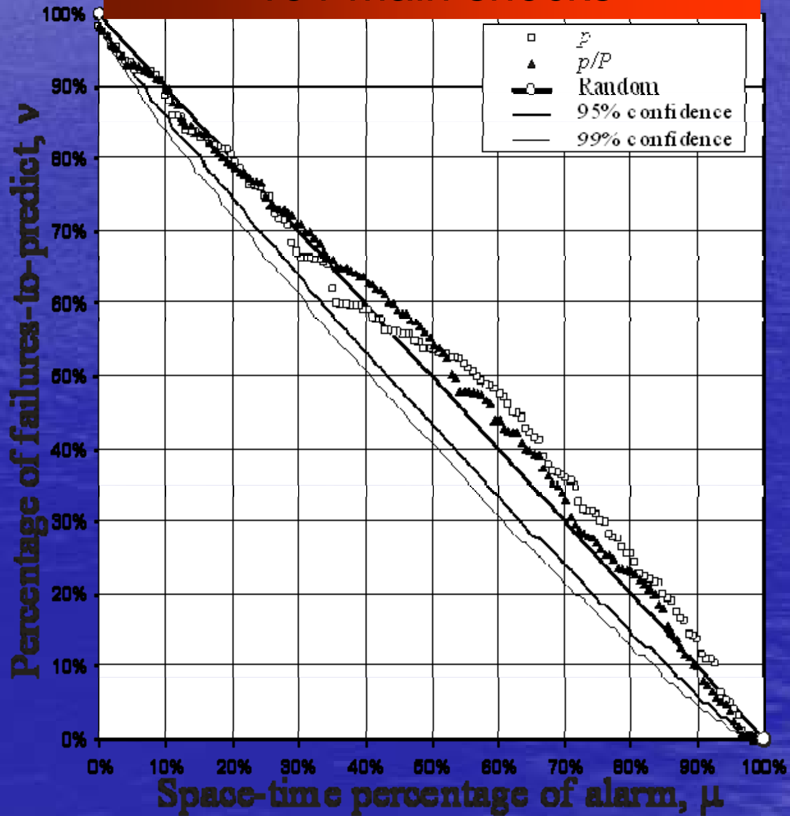
we plotted several ν vs. μ diagrams.

The two figures show the performance of predictions based on p or p/P in the test period from April 10, 2002 to September 13, 2004. The total of 218 earthquakes of magnitude Mw = 5.8 or more with the depth of 70 km or shallower occurred in the West Pacific. According to definition from (*Keilis-Borok et al., 1980*), 67 of them are aftershocks and 151 main shocks.

All 218 target earthquakes

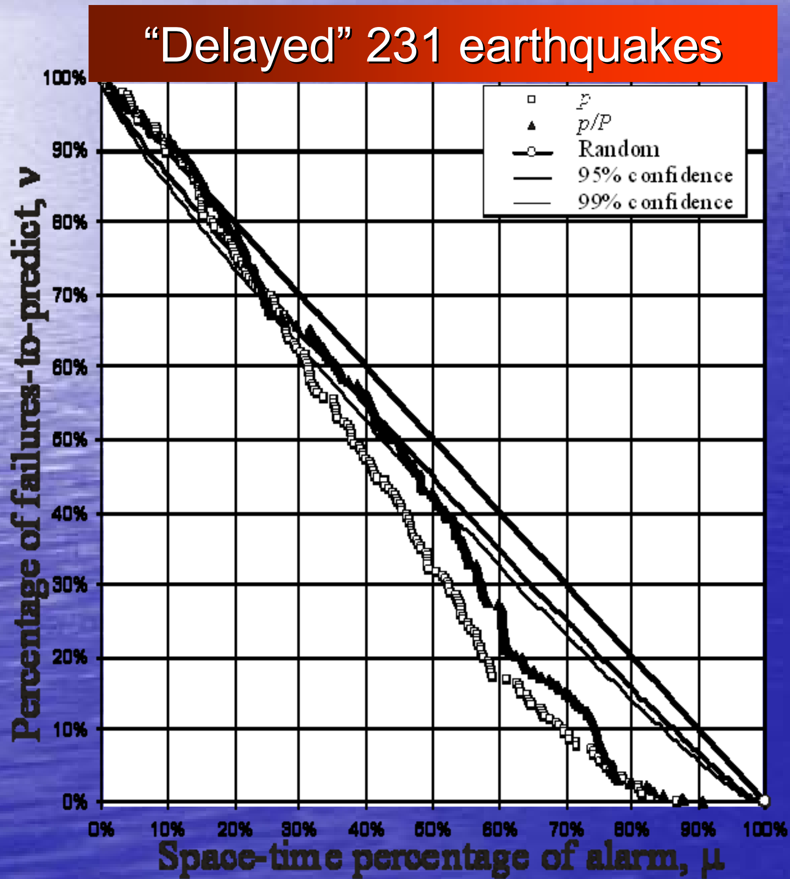


151 main shocks



The outcome of an "absurd" prediction:

The percentage of the failures-to-predict ν versus the percentage of the alerted space-time volume μ: $\{\mu_p(E), \nu_p(E)\}$ and $\{\mu_{p/P}(E), \nu_{p/P}(E)\}$ generated by "prediction" of the 231 earthquakes with magnitude MwHRV ≥ 5.8 and depth ≳ 70 km in April 10, 1992-September 13, 1994 using the *p* and *p/P* maps computed for April 10, 2002-September 13, 2004.



The observed deviation from the diagonal is about the same or larger than in the real-time applications.

Thus, we cannot reject random nature of the Jackson-Kagan "probabilistic" method and may conclude that

(i) its effectiveness for predicting large earthquakes is doubtful, and

(ii) the applicability of the underlying ETAS model is an ingrained bigotry.

Regional Earthquake Likelihood Models:
A realm on shaky grounds?

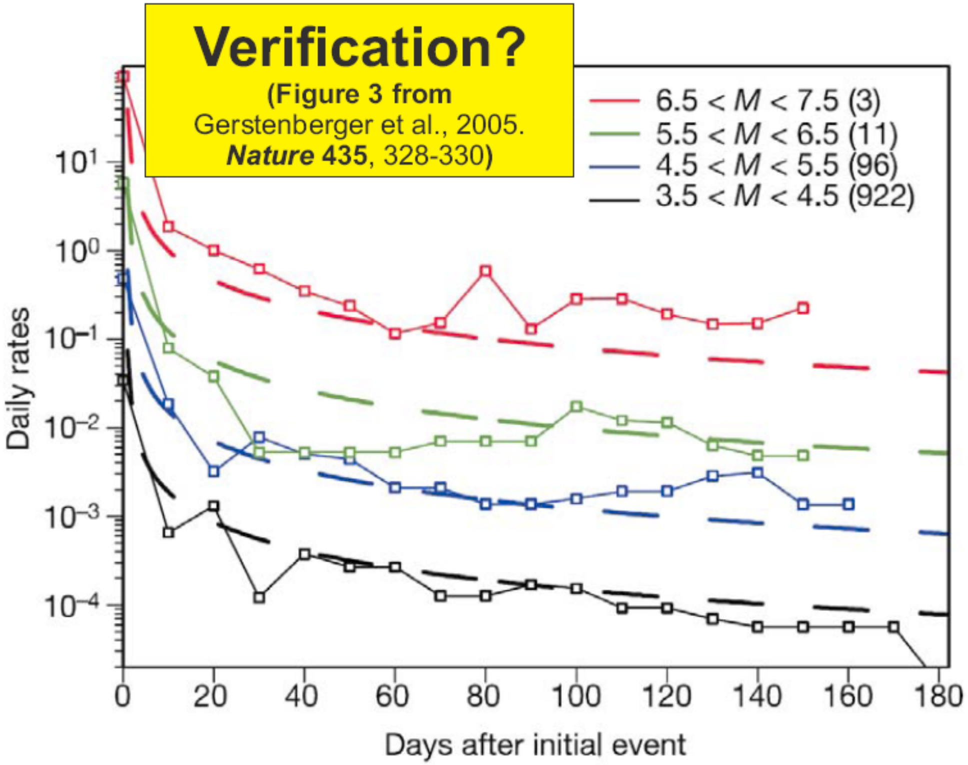
Likelihood scoring is one of the delicate tools of Statistics, which could be worthless or even misleading when inappropriate probability models are used. This is a basic loophole for a misuse of likelihood as well as other statistical methods on practice. The flaw could be avoided by an accurate verification of generic probability models on the empirical data. It is not an easy task in the frames of the Regional Earthquake Likelihood Models (RELM) methodology, which neither defines the forecast precision nor allows a means to judge the ultimate success or failure in specific cases. Hopefully, the RELM group realizes the problem and its members do their best to close the hole with an adequate, data supported choice.

On 19 May 2005, the United States Geological Survey began a public web site with forecasts of expected ground shaking for 'tomorrow' and *Nature* published the underlying work by *Gerstenberger et al*.

Gerstenberger, M. C., Wiemer, S., Jones, L. M. & Reasenberg, P. A. Real-time forecasts of tomorrow's earthquakes in California. *Nature* **435**, 328-331 (19 May 2005)

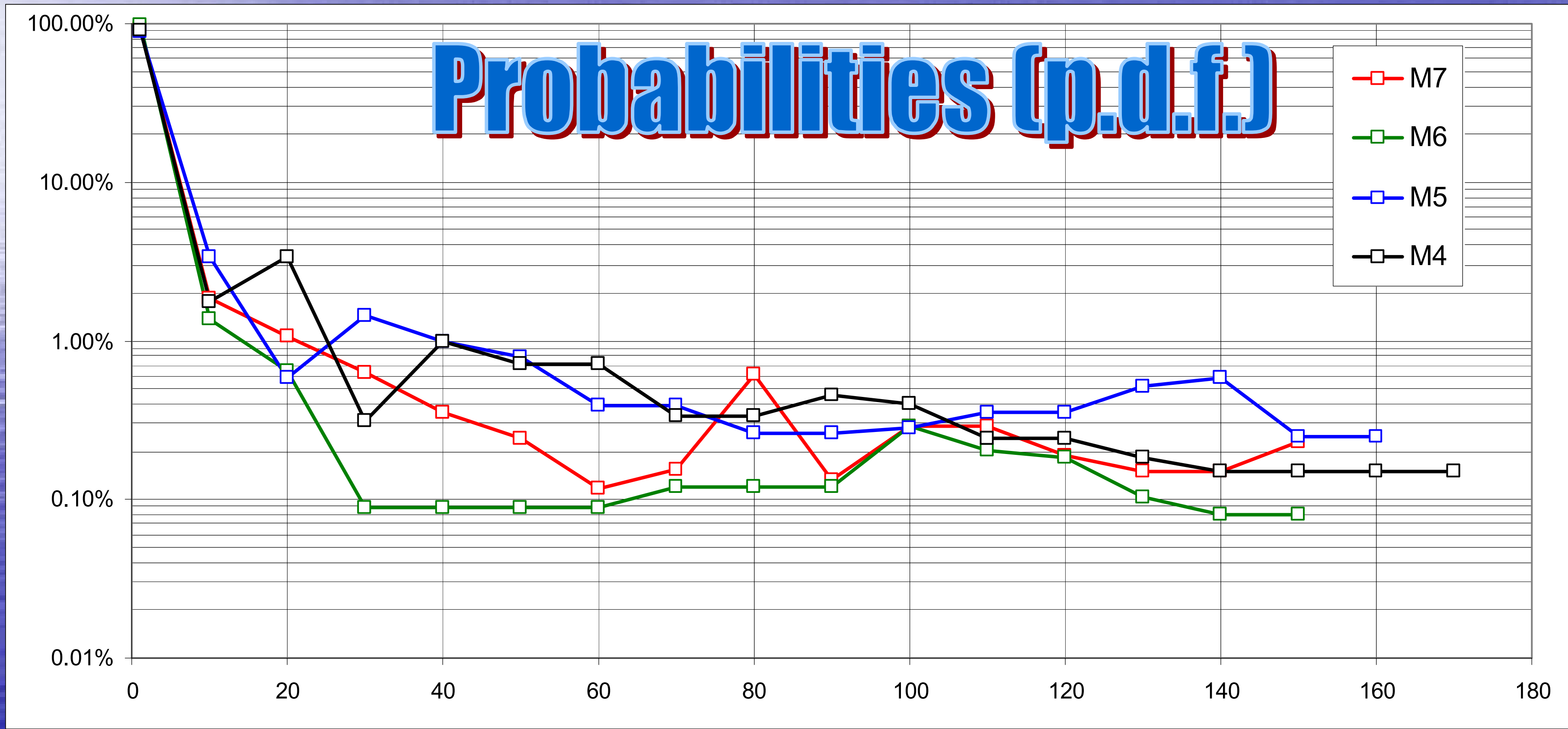
LETTERS   NATURE|Vol 435|19 May 2005

Figure 3 | Calculated and observed rates of events $M \geq 4$ in 24-hour intervals following mainshocks occurring between 1988 and 2002 in southern California. Dashed lines show the rates forecasted by the generic California clustering model (without cascades) for the mainshock magnitude (M) shown. For this test a simple circular aftershock zone implementation (solid lines) gives the observed rates of $M \geq 4.0$ aftershocks following all mainshocks with magnitude within 0.5 units of M. The aftershock zones are defined as the areas within one rupture length of the mainshock epicentre.

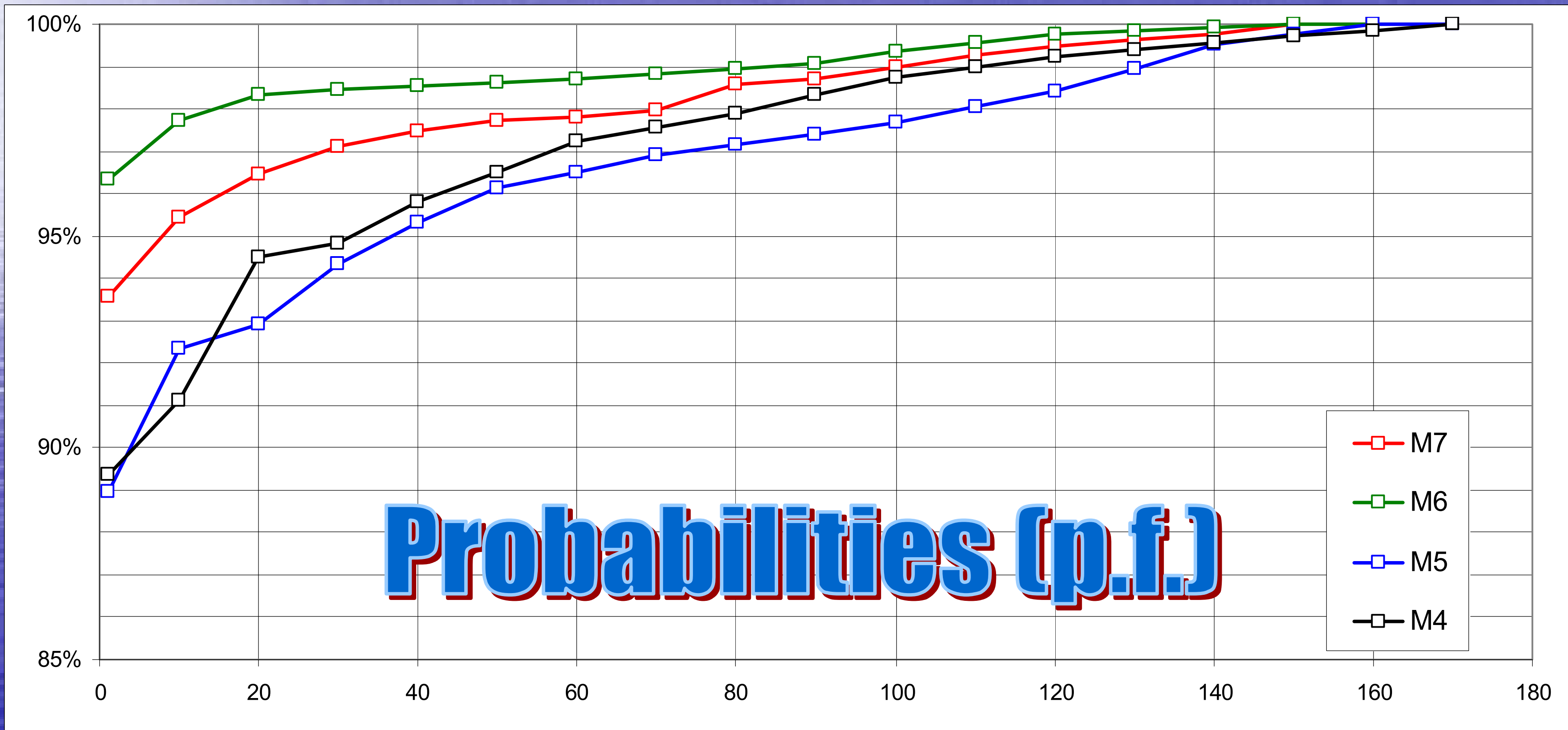


"As a first test, we verified that the generic clustering model describes the average clustering activity of California reasonably well. Using data from 1988−2002, after the period used to initially develop the model and thus independent data, we compute the average daily rate of events following an earthquake of a given size (Fig. 3)."

Probability density distribution functions of the random variable "Time after initial event" in different magnitude ranges of the initial event.



Probability distribution functions of the random variable "Time after initial event" in different magnitude ranges of the initial event.



Proof: Normalised by condition that the total integral of the p.d.f. (probability density function) increments equals 1, each of the four plots provides the minimum of positive p.d.f. increments, which are by definition either 1/N or its integer multiple (e.g., 2/N, 3/N, etc.). These are about 0.0012, 0.0008, 0.0025, and 0.0015, which values imply the sample sizes about 846, 1250, 401, and 665 or integer multiples of these values. **The probability** of a smaller value of the Kolmogoroff-Smirnoff statistic D than that for the two samples used to plot the daily rates after $5.5 < M < 6.5$ (green plot in Figure 3) event and after $3.5 < M < 4.5$ (black plot) event (which D accounts to the value $D = \max | F_{green}(t) - F_{red}(t) | \cdot (N_1 N_2/(N_1+N_2))^{1/2} \geq 2.12$)

is larger than 97%.

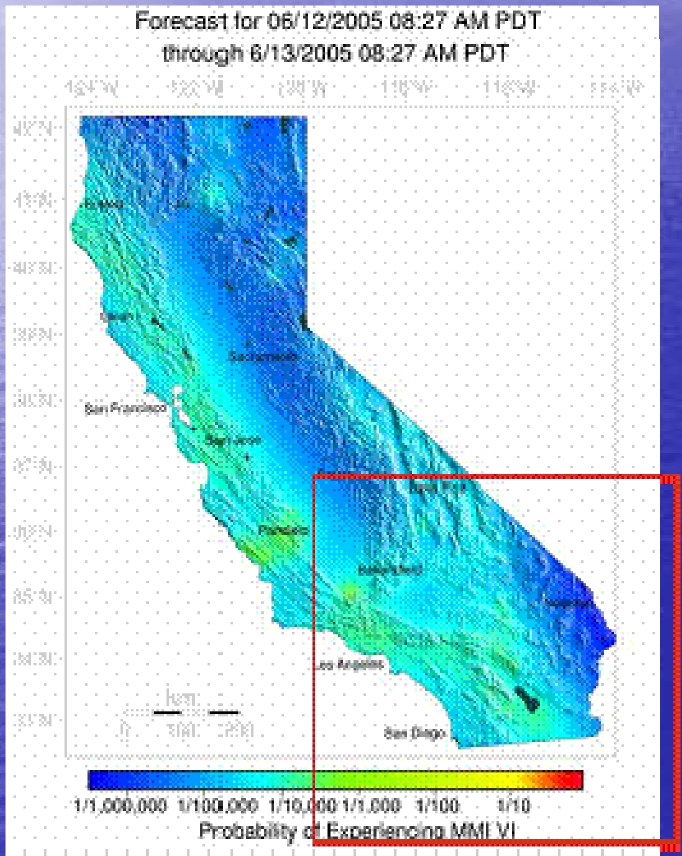
Therefore, the hypothesis that these two samples are drawn from the same distribution can be rejected at significance level of 0.03. ■

Thus, the statistics of the observed ground shaking in California, 2005-present, demonstrate that

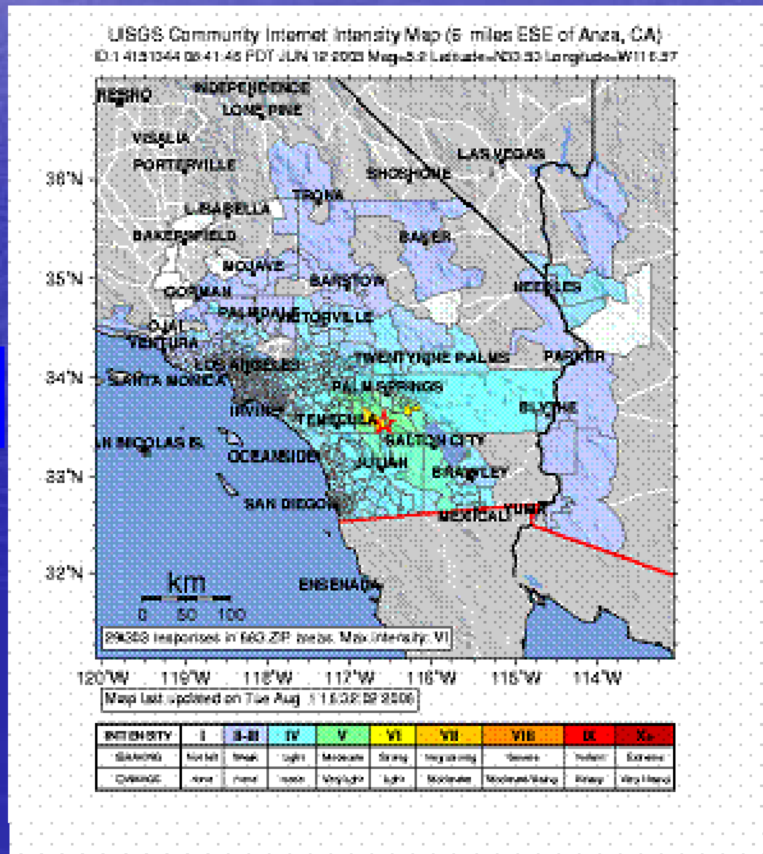
- earthquakes of Modified Mercalli intensity VI in California keep occurring in the "sky blue" areas of the lowest risk ($p<1/10000$),
- while the extent of the observed areas of intensity VI is by far less than the one expected from the calculations (currently a very crude low bound estimate of the ratio has surpassed a factor of 8.5…).

USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

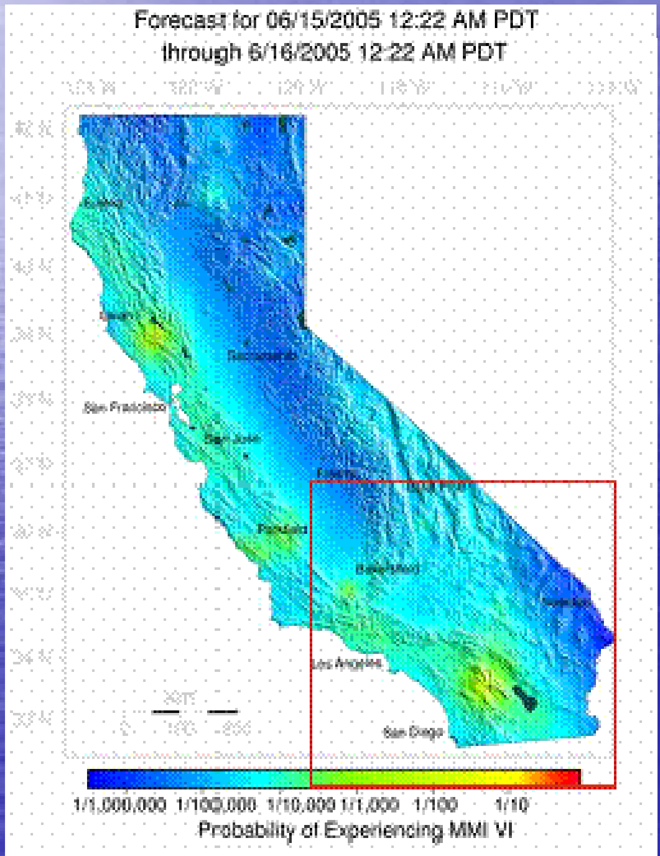


1

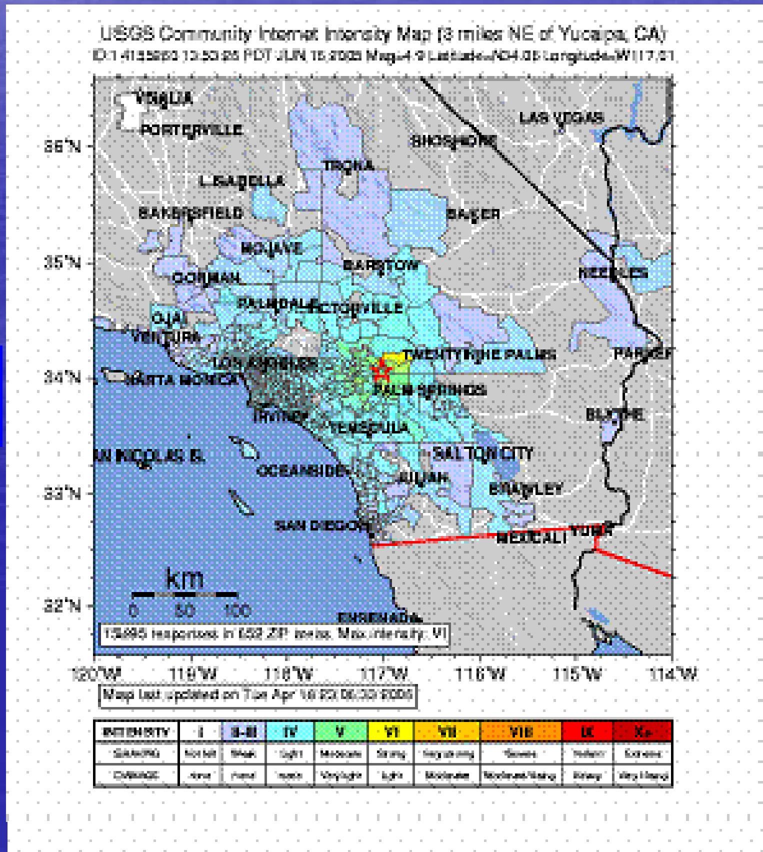


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

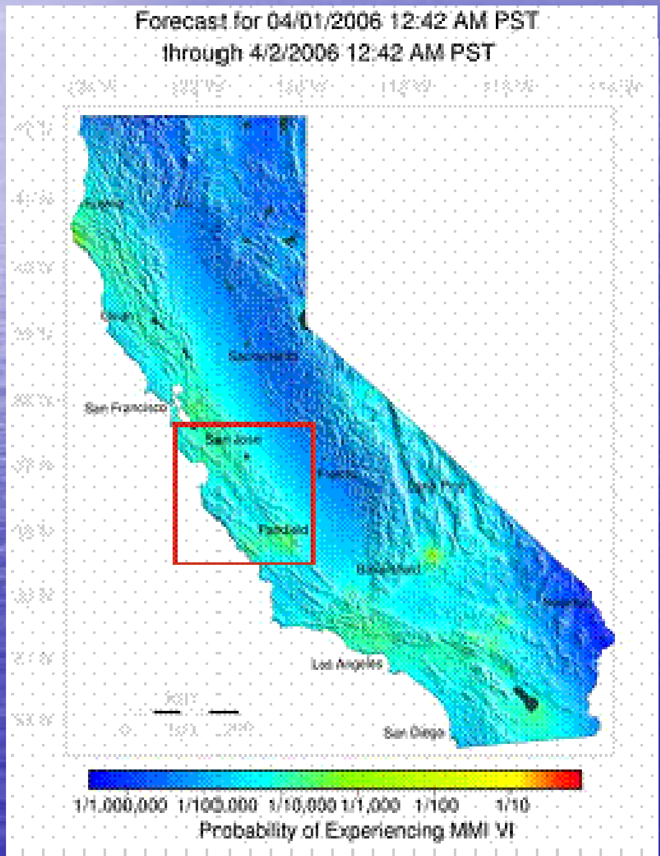


2

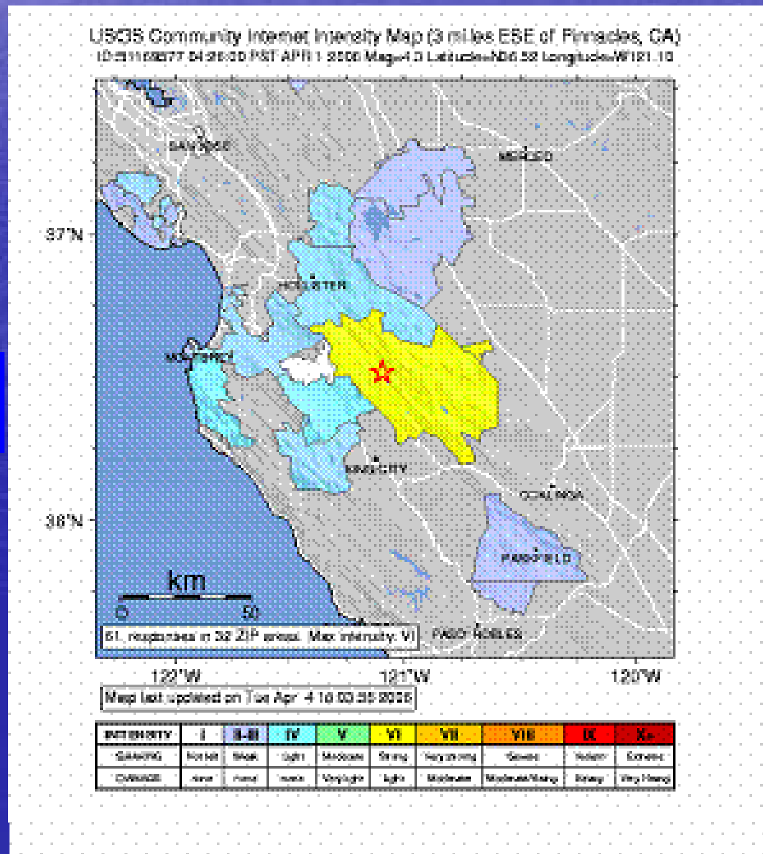


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

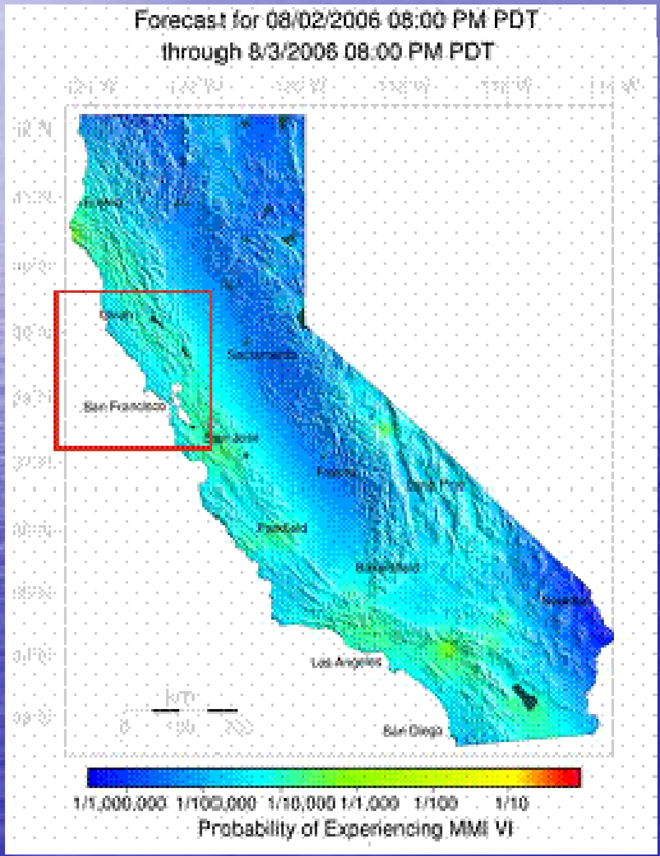


3

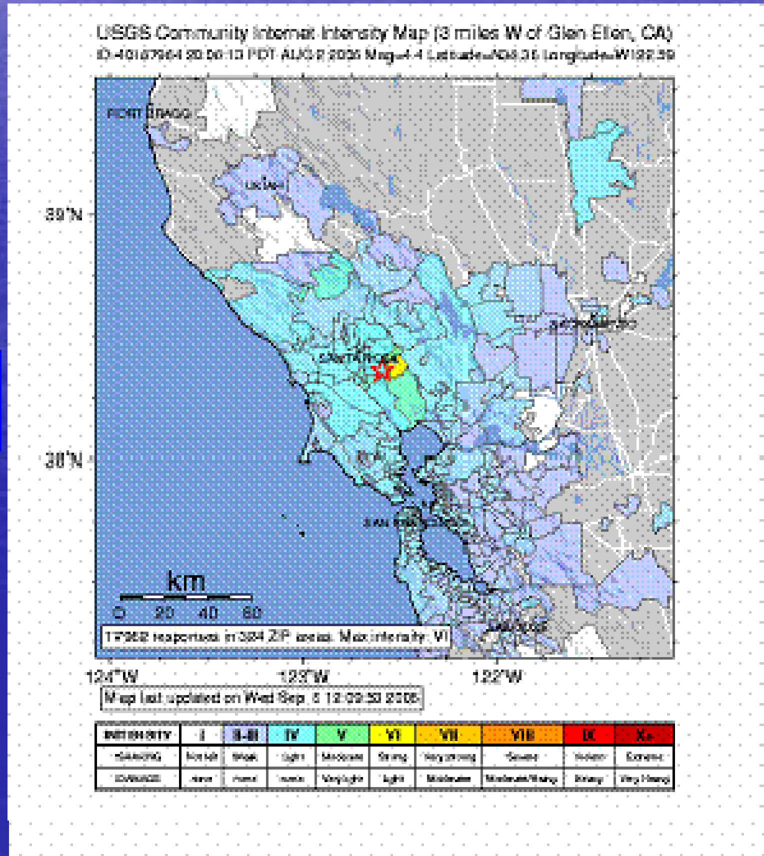


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

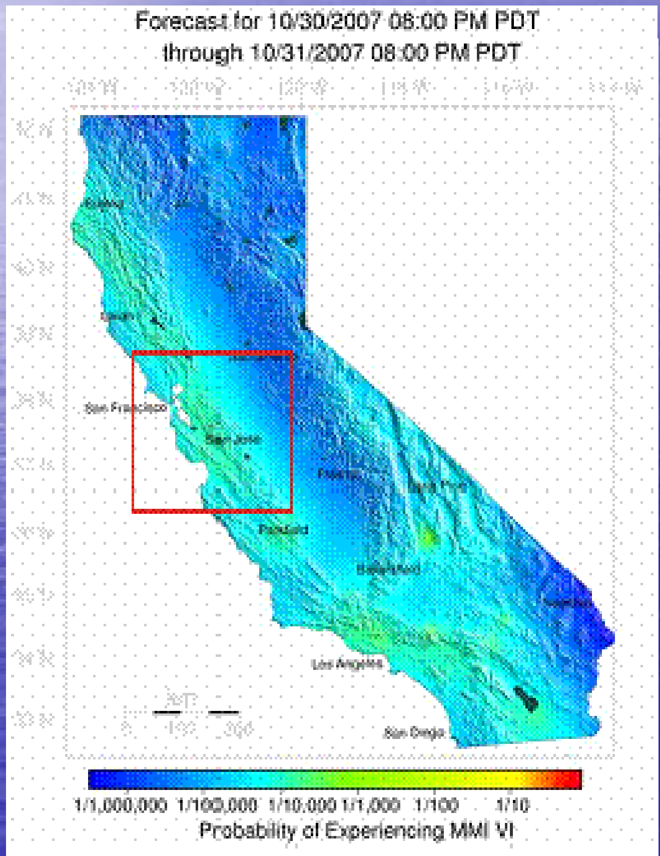


4

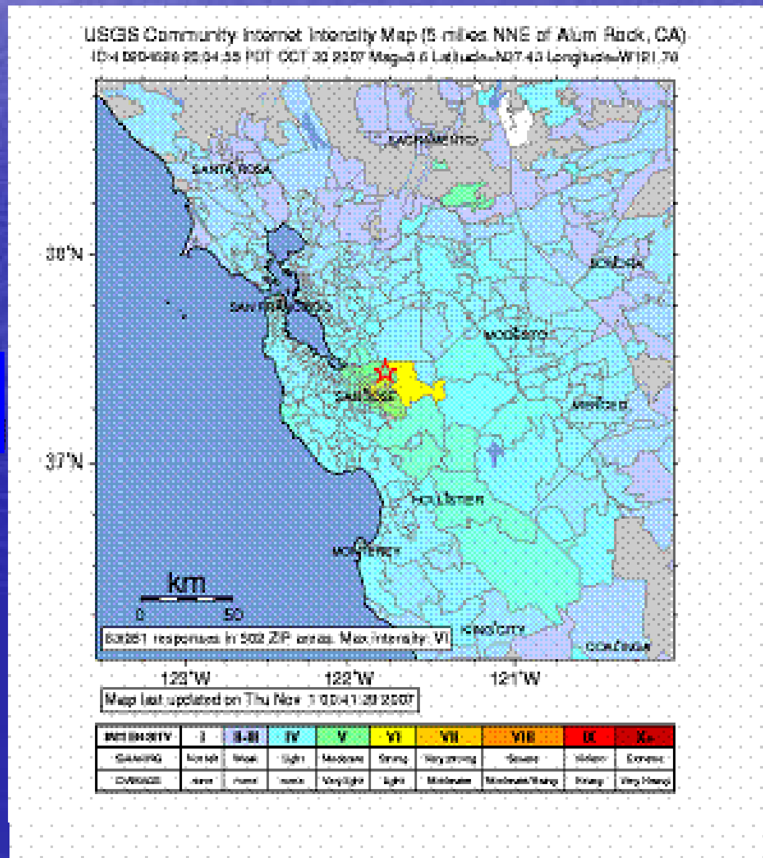


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

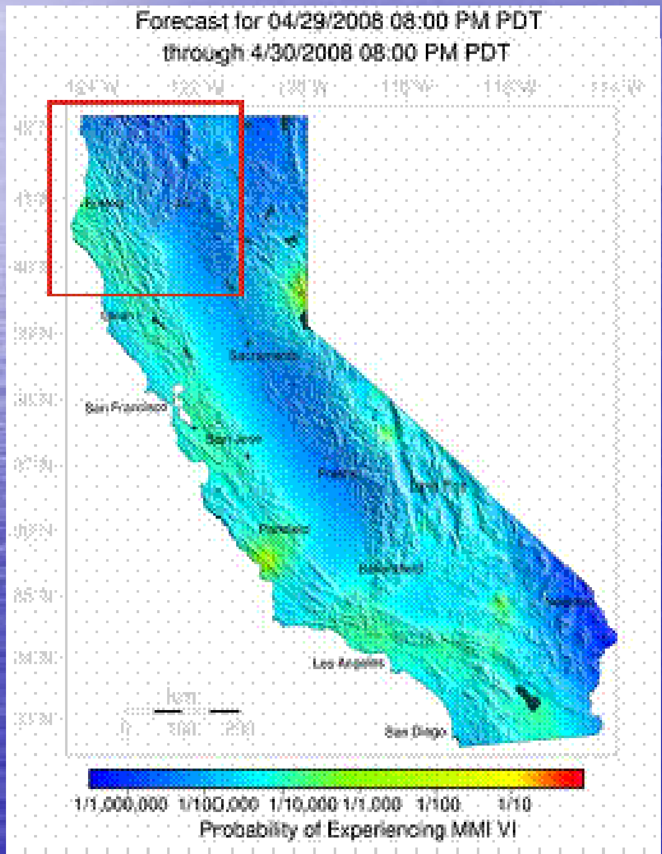


5

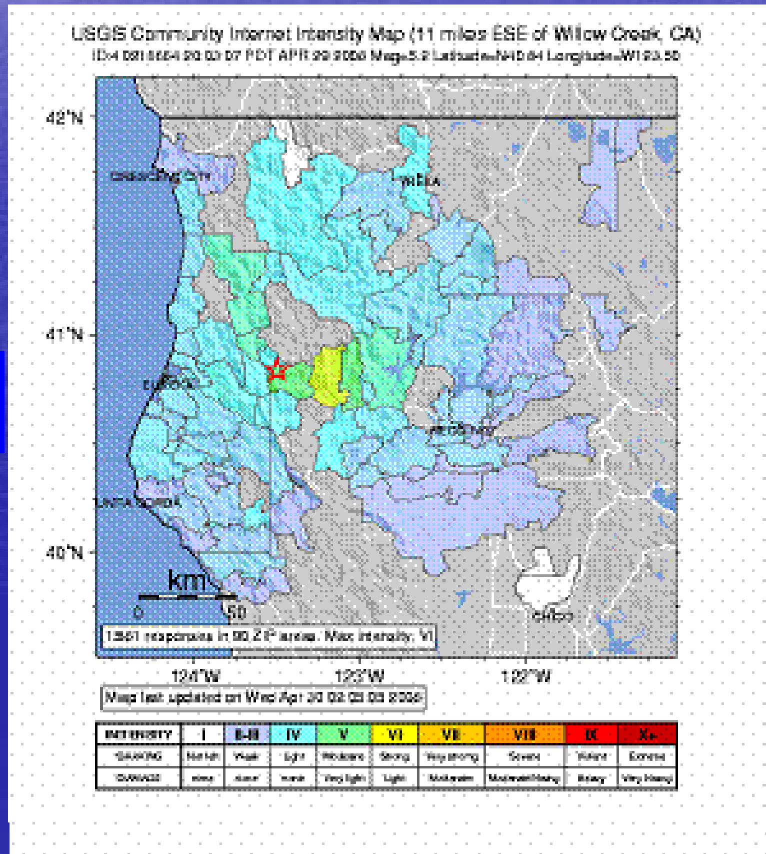


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -

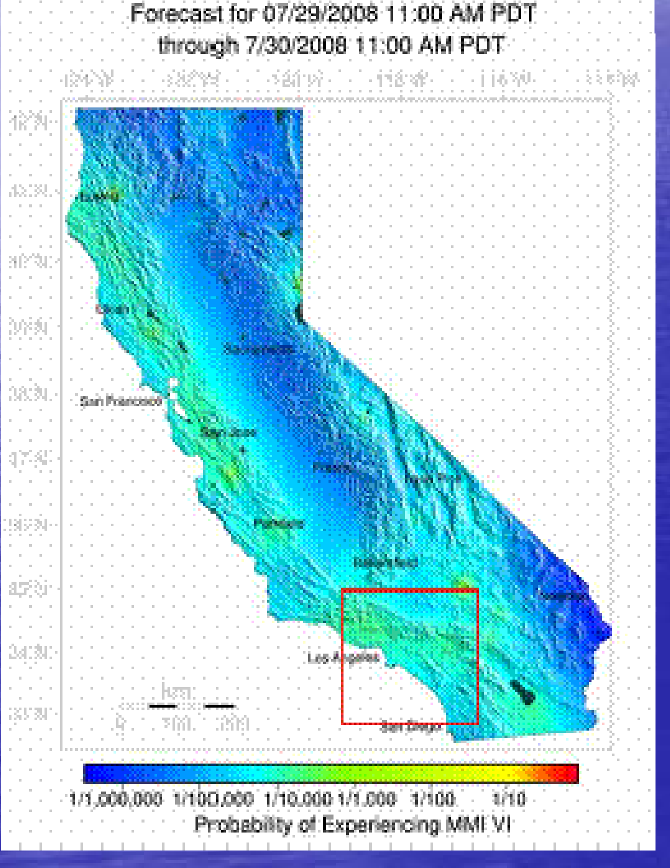


6

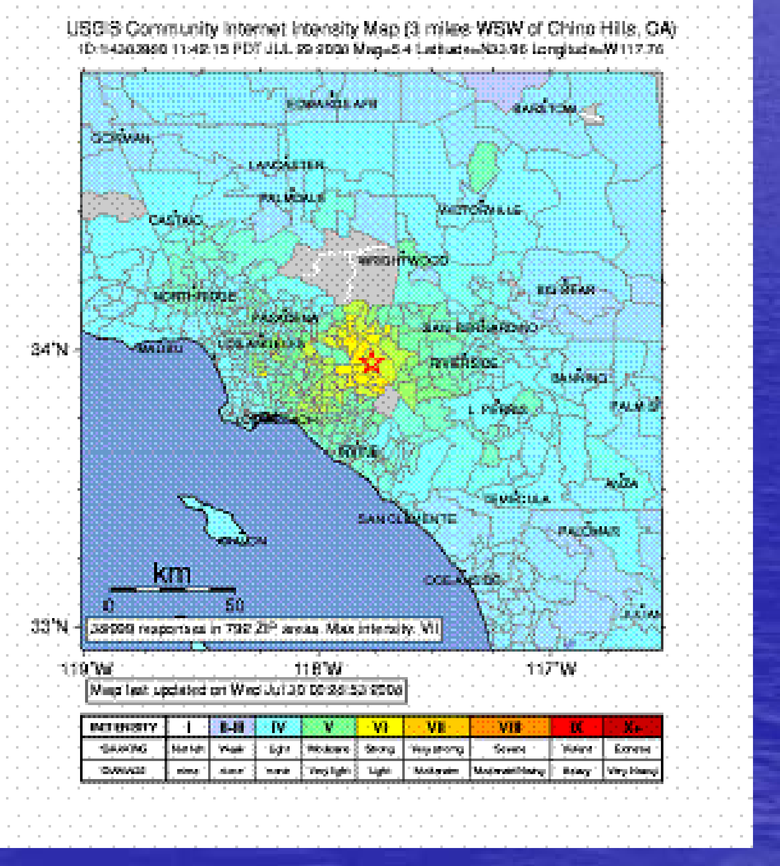


USGS Web Site Misleads Californians

Since the time of *Nature* published the work by *Gerstenberger et al* on May 19, 2005 -



7



USGS Web Site Misleads Californians

Regretfully, USGS continues delivering to the public, emergency planners and the media, a forecast product,

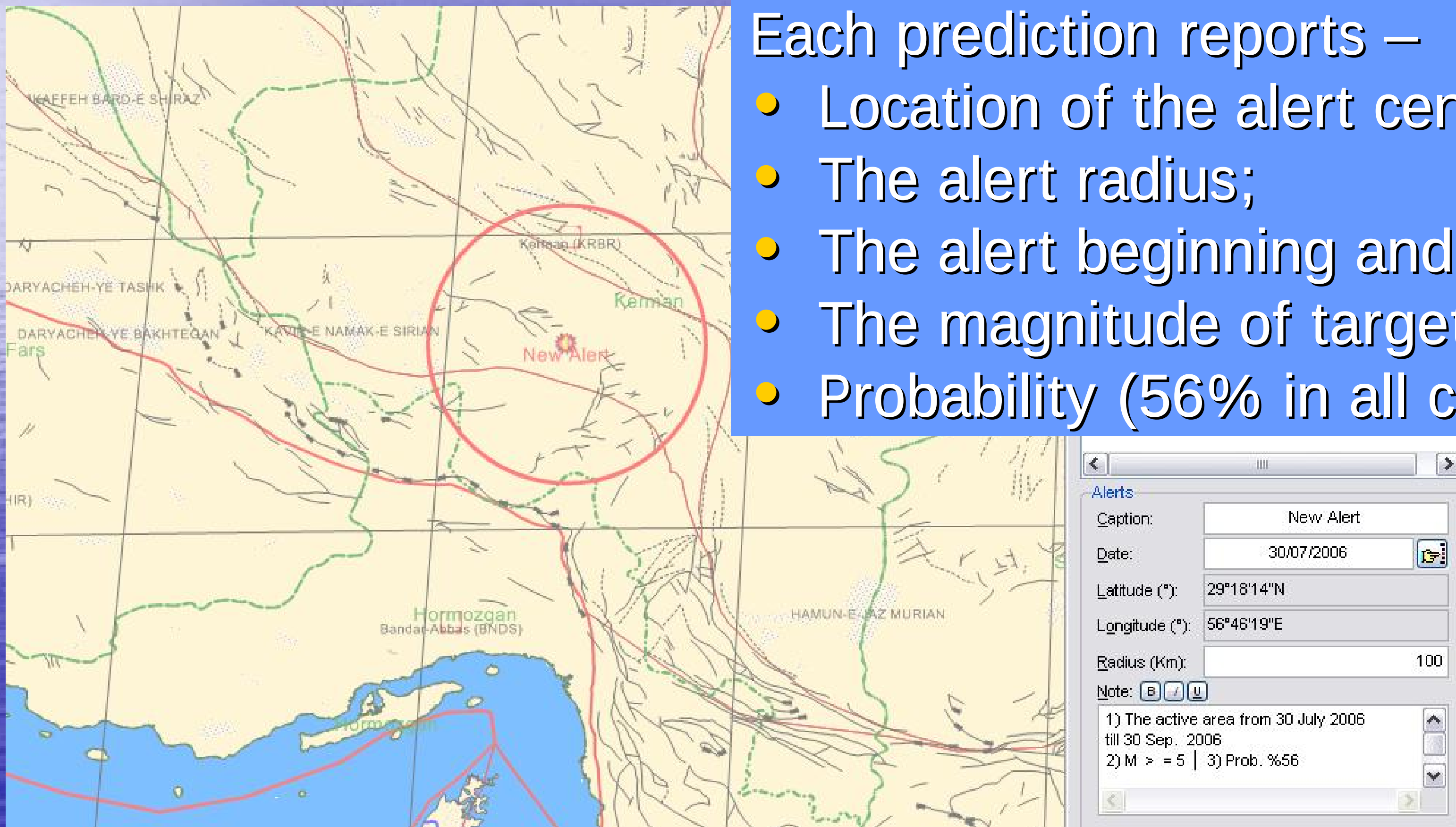
which is based on wrong assumptions, which violates the best-documented earthquake statistics in California,

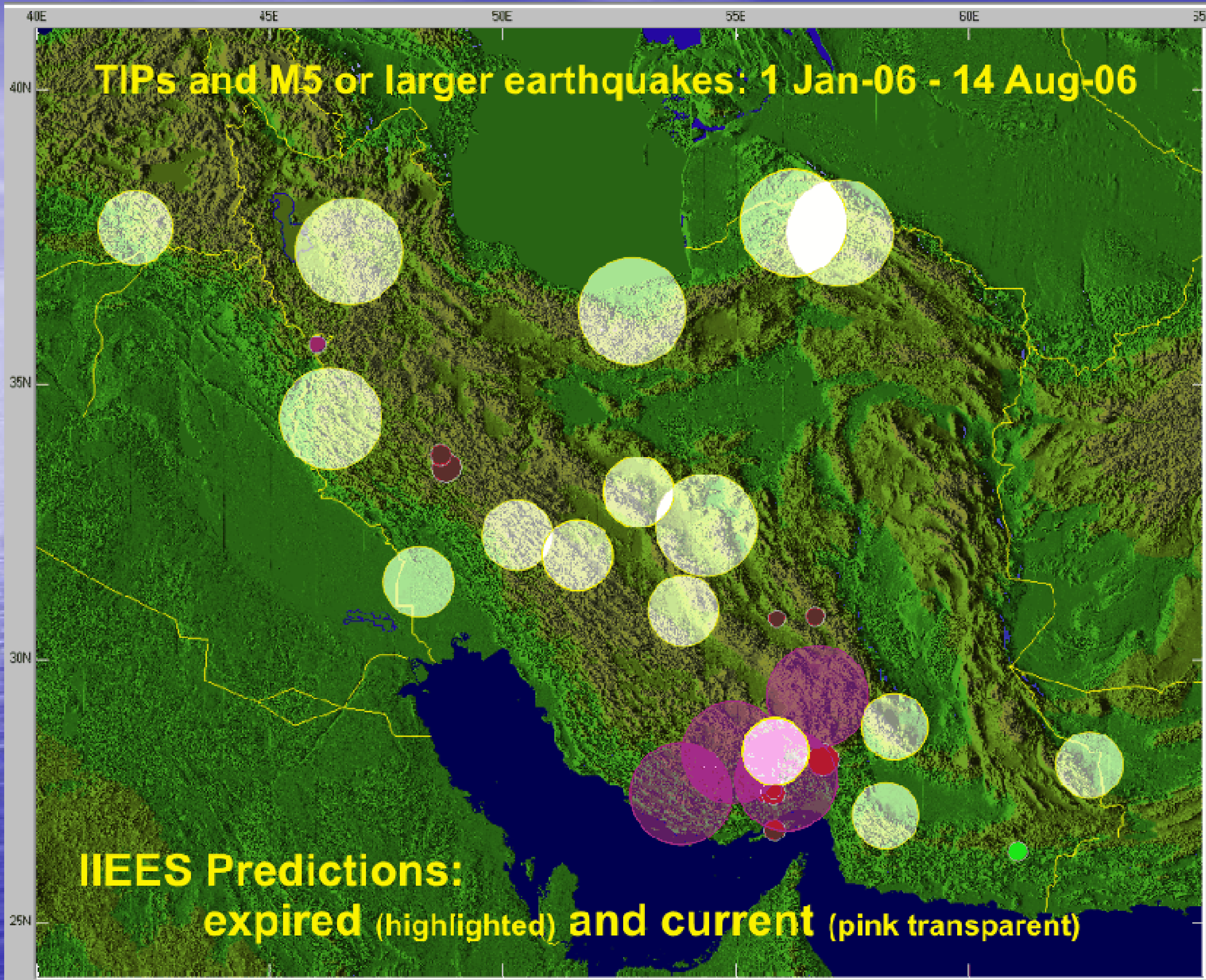
which accuracy was not investigated, and which forecasts were not tested by the authors in any rigorous way.

Recently, IIEES did set up a website of restricted access (ftp://www.iiees.ac.ir/eqprediction), which we have a chance to visit systematically since March 8, 2006.

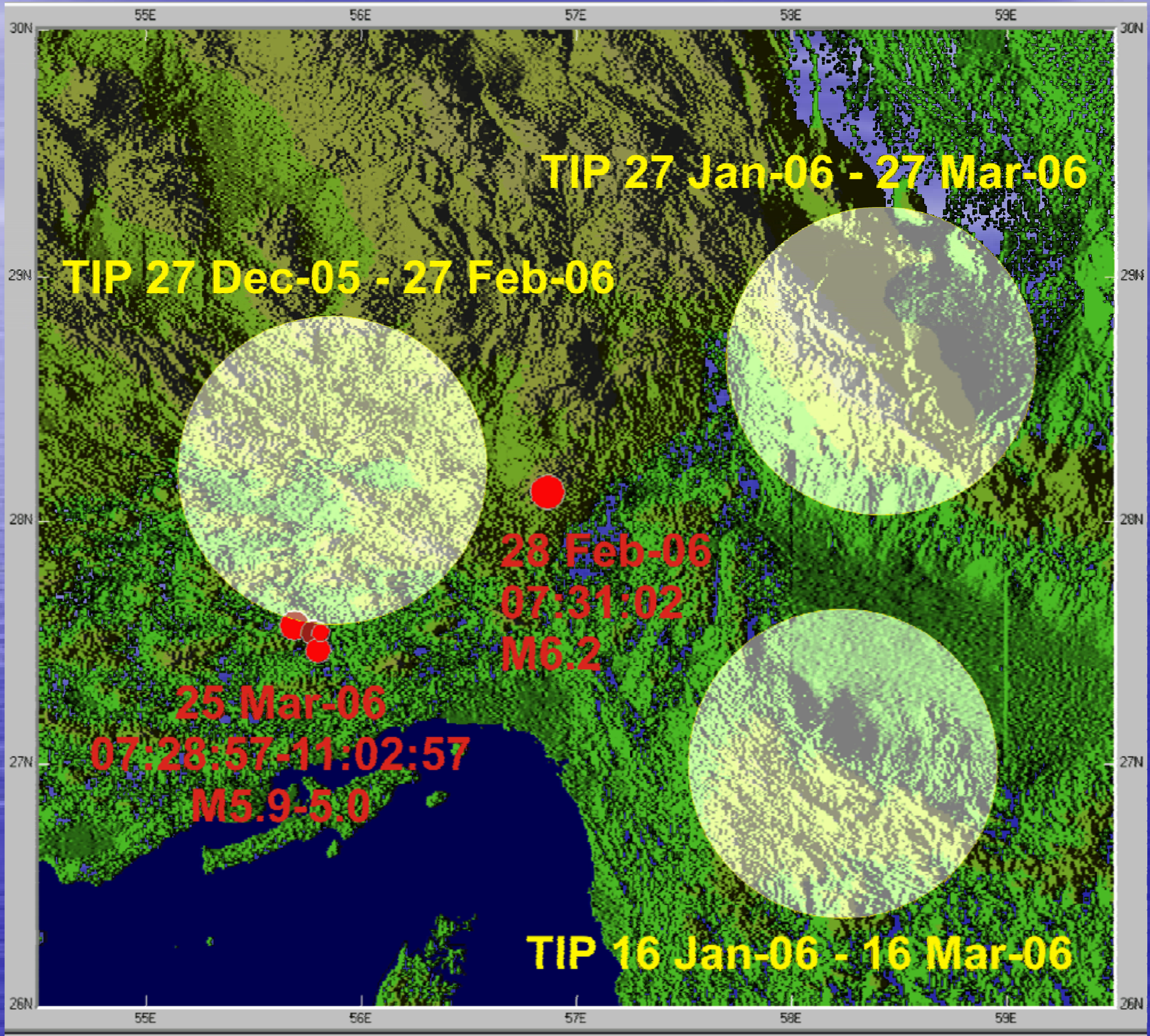
Each prediction reports –

- Location of the alert center;
- The alert radius;
- The alert beginning and end;
- The magnitude of target event;
- Probability (56% in all cases)




40E
45E
50E
55E
60E
40N
35N
30N
25N
TIPs and M5 or larger earthquakes: 1 Jan-06 - 14 Aug-06
IIEES Predictions:
expired (highlighted) and current (pink transparent)

Zoom to alerts and target earthquakes in the South

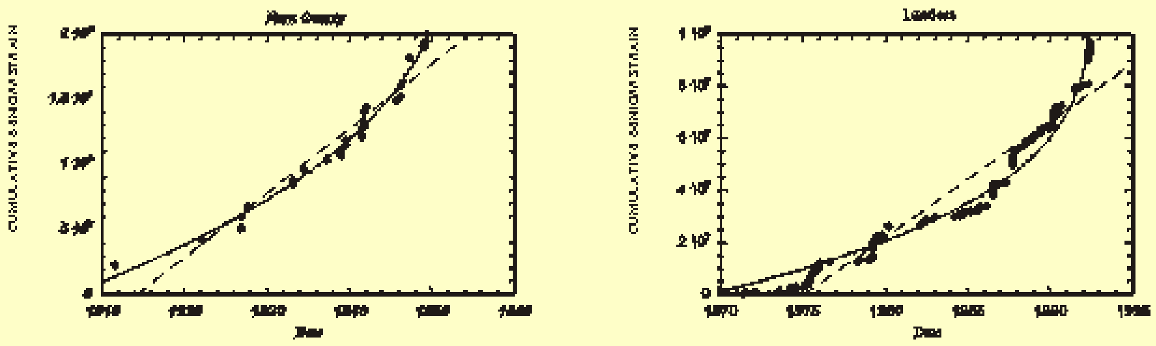


IIEES predictions:

- We continuously observe no success;
- Evidently, this highly contradicts the expected number $P \cdot N = 56\% \cdot 21 = 11.76$ (presumably, P is an estimate of probability of success);
- The IIEES predictions are misleading and their dissemination to the public, emergency planners and the media should not be done;
- The underlying theory is either erroneous or applied in a wrong way.

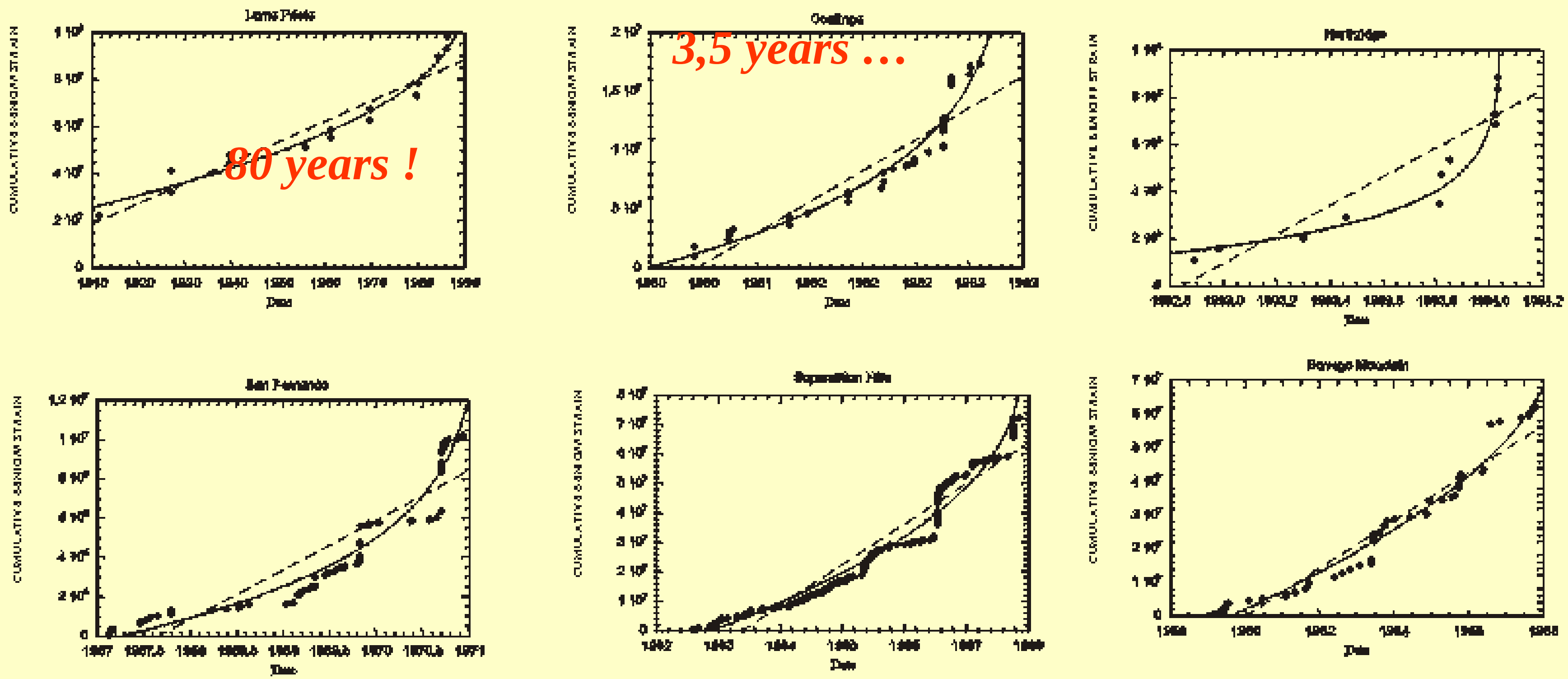
Other evident cases of misuse of Statistics

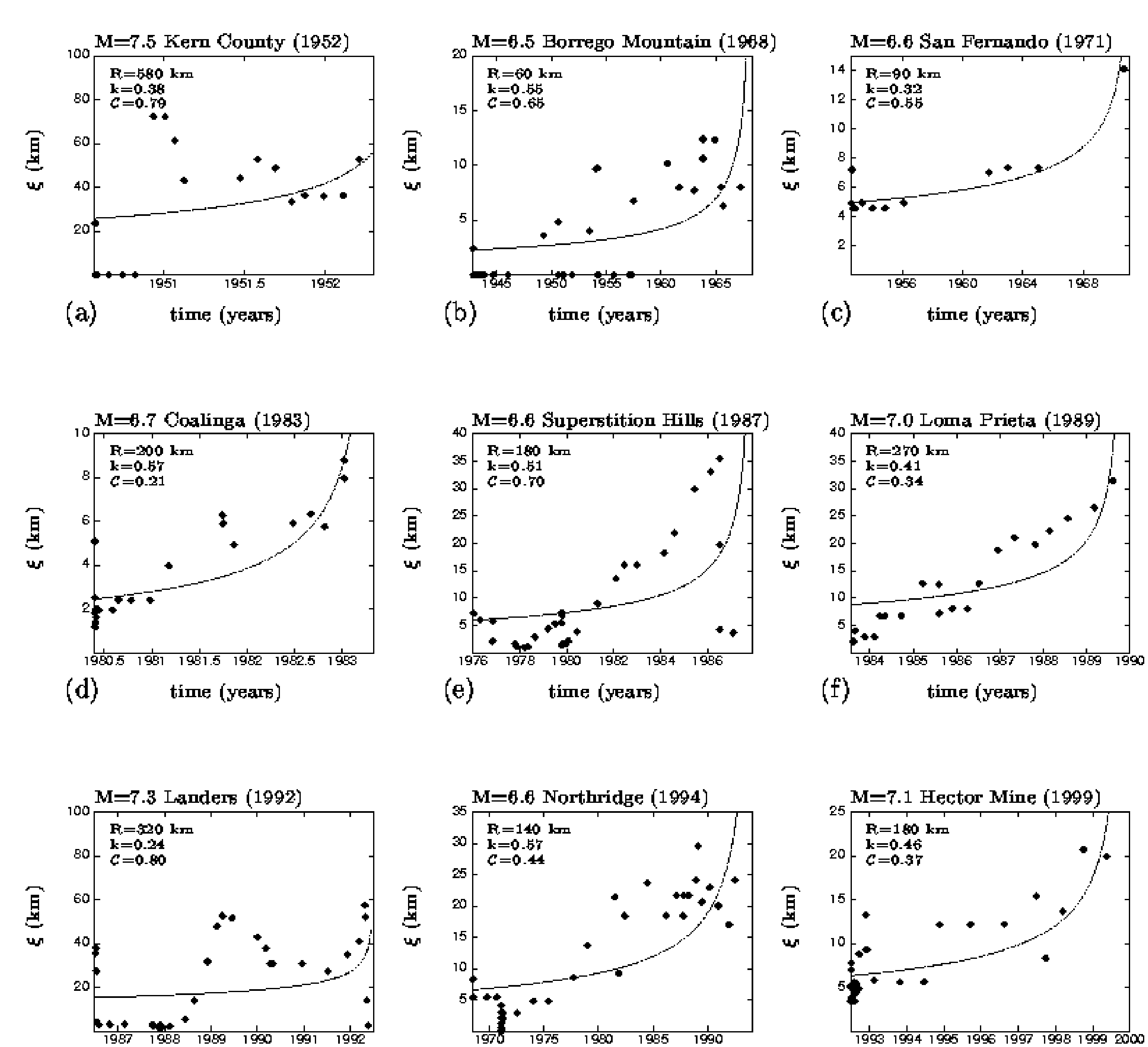
Bowman, Ouillon, Sammis, Sornette, & Sornette, 1998



It is still unclear if "the best fit" is random… ?

<u>Free parameters</u>: dT, Mc, aftershocks





Zoller, Hainzl & Kurths, 2001

Does "the best fit" fit the data at all ?

Conclusion?

"See how easy it is?"

Should seismologists continue to ignore evident features of seismic activity in favor of "old-good" fallacies?

Verified “Precursors”

- The simple seismicity patterns – Σ and “burst of aftershocks” – were given unambiguous reproducible definitions and their predictive value was validated by the prospective worldwide tests. **However,** it is not clear yet whether some single simple premonitory pattern may compete in performance with prediction algorithms that combine several traits describing the dynamics of seismic region at the approach of a large earthquake.

"Undue precision of computations is the first symptom of mathematical illiteracy"

A.N.Krylov, (1863-1945)

a famous Russian mathematician

The accuracy of an earthquake prediction method is essentially predefined by the accuracy of the data available, which is far from ideal. The unavoidable natural difficulties in observing seismic events as well as in correlating them with other geophysical phenomena and fields complicates the design and testing of a new generation of earthquake prediction technique.

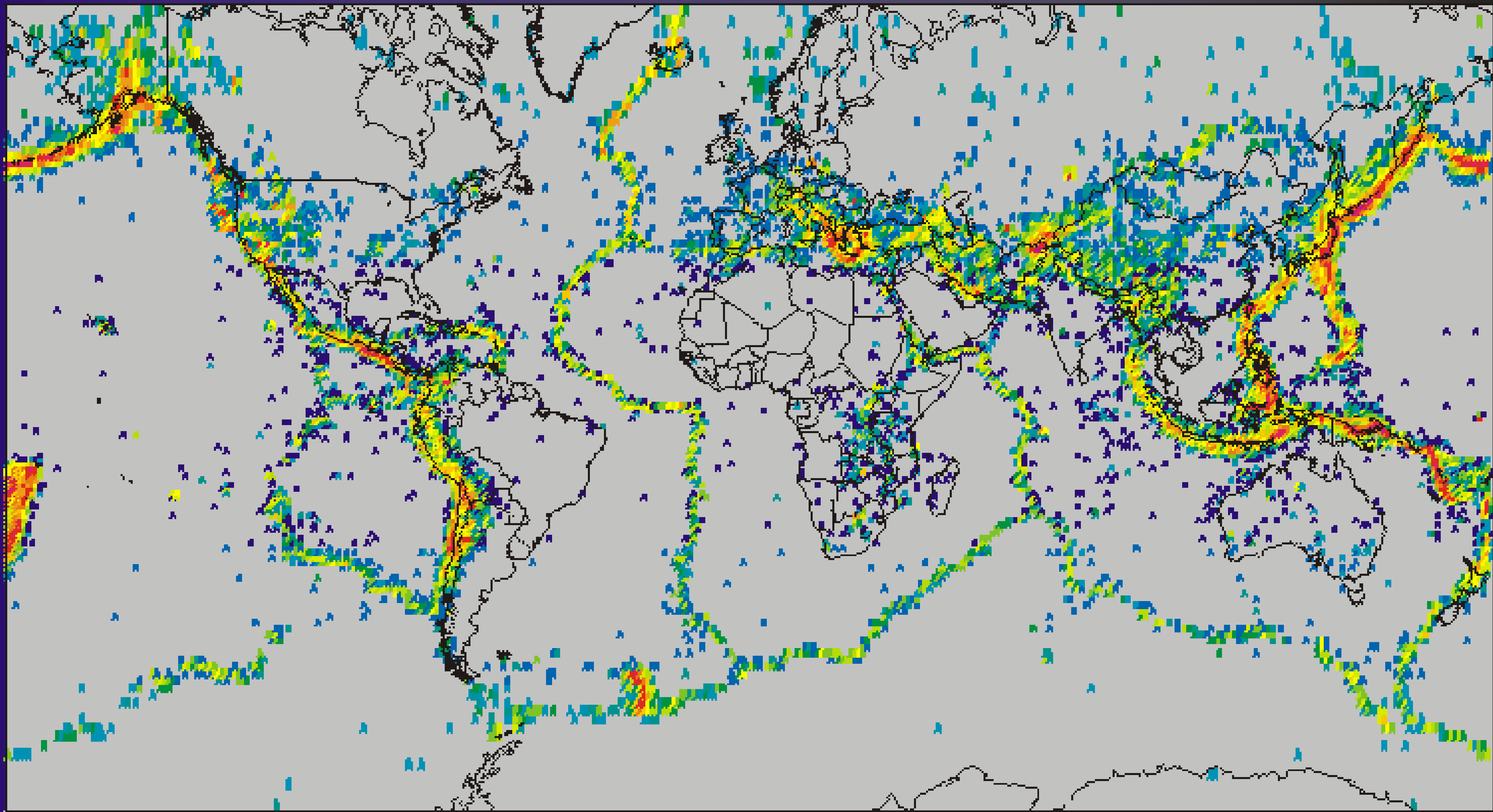
The accumulated case-histories of predicted and not predicted earthquakes provide us unique and so far very limited information that may help understanding the ultimate limits of seismic predictability.

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

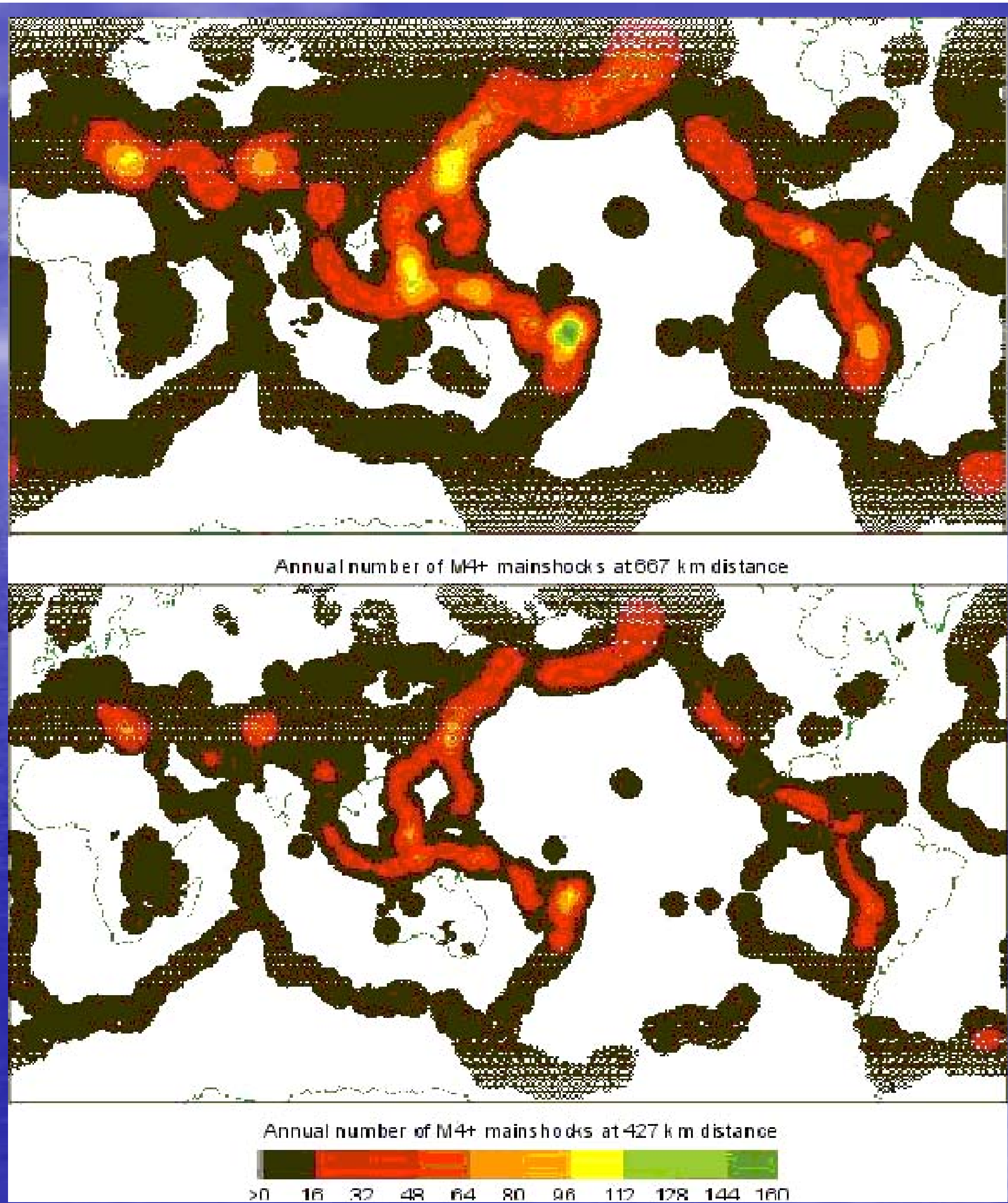
- Moreover, the Gutenberg-Richter law suggests limiting magnitude range of prediction to about one unit. Otherwise, the statistics would be essentially related to dominating smallest earthquakes.

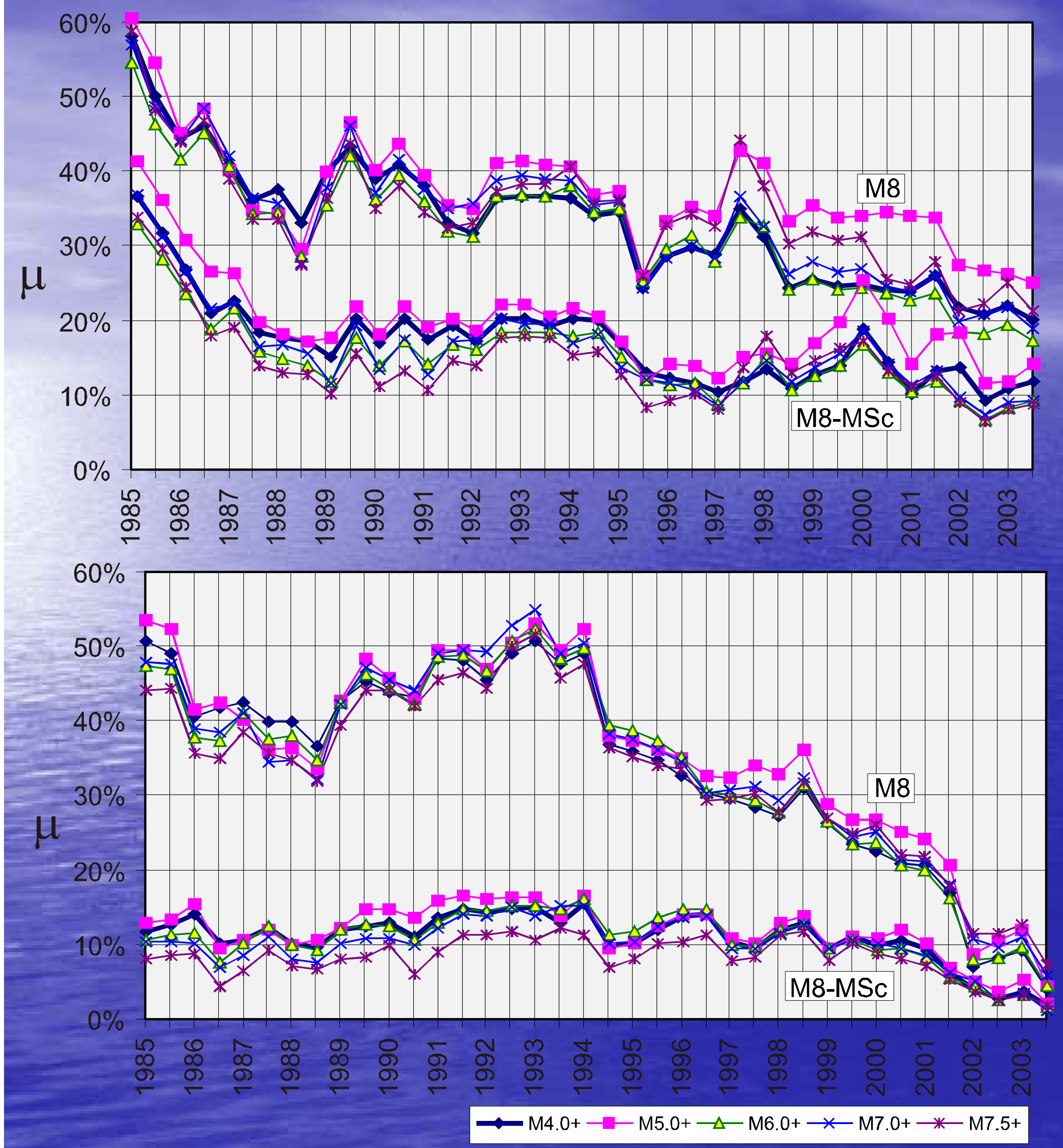


Average annual number of magnitude 4.0 or greater earthquakes at a 1°×1° cell (*normalized to its area on equator*)

Earthquakes are rare events. Therefore, the application of the M8 algorithm is limited to the areas where reported earthquakes are large enough in number.

The color on the maps signifies the annual average number of earthquakes with magnitude 4 or larger in the 667-km (above) and 427-km (below) circles centered at the point.

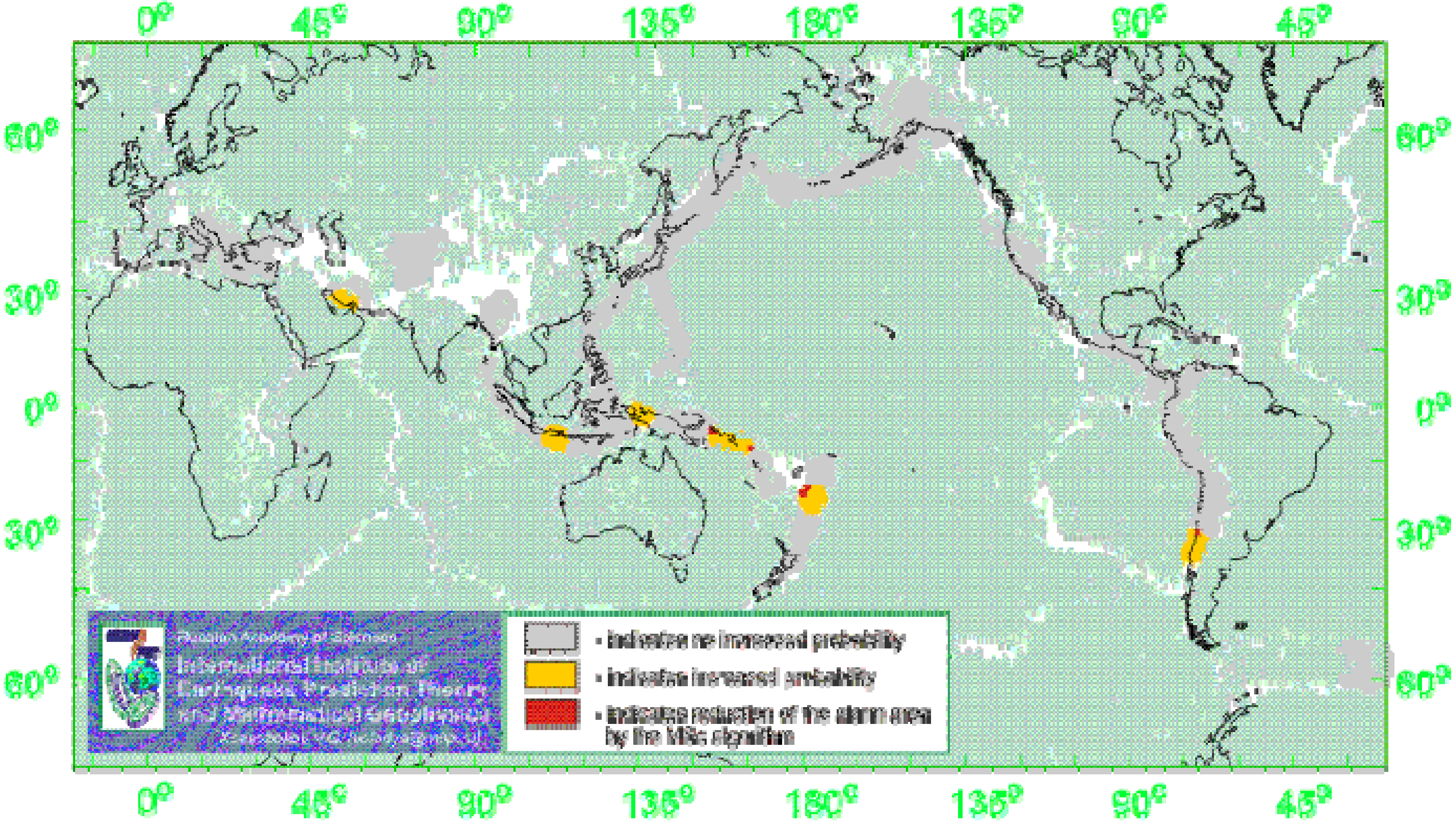




The percentage of alerted area as a function of time for M8.0+ (above) and M7.5+ (below).

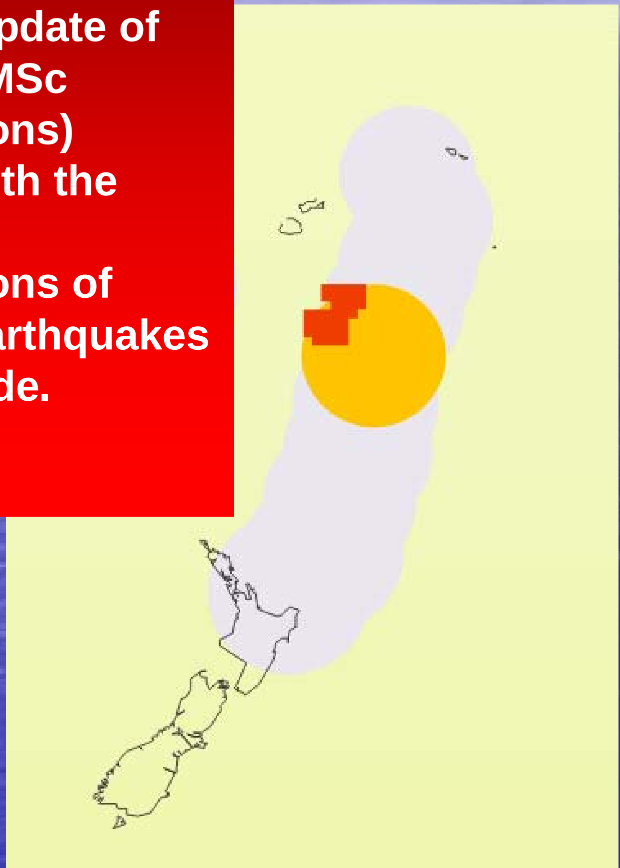
The obtained estimates are based on the counts of magnitude 4 or more and 5 or more earthquakes in the period from 1964 through 1984, while the counts of magnitude above 6.0, 7.0, and 7.5 in 1900-1984

Regions of Increased Probability of Magnitude 7.5+ Earthquakes as on July 1, 2002 (*subject to update on January 1, 2003*)



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

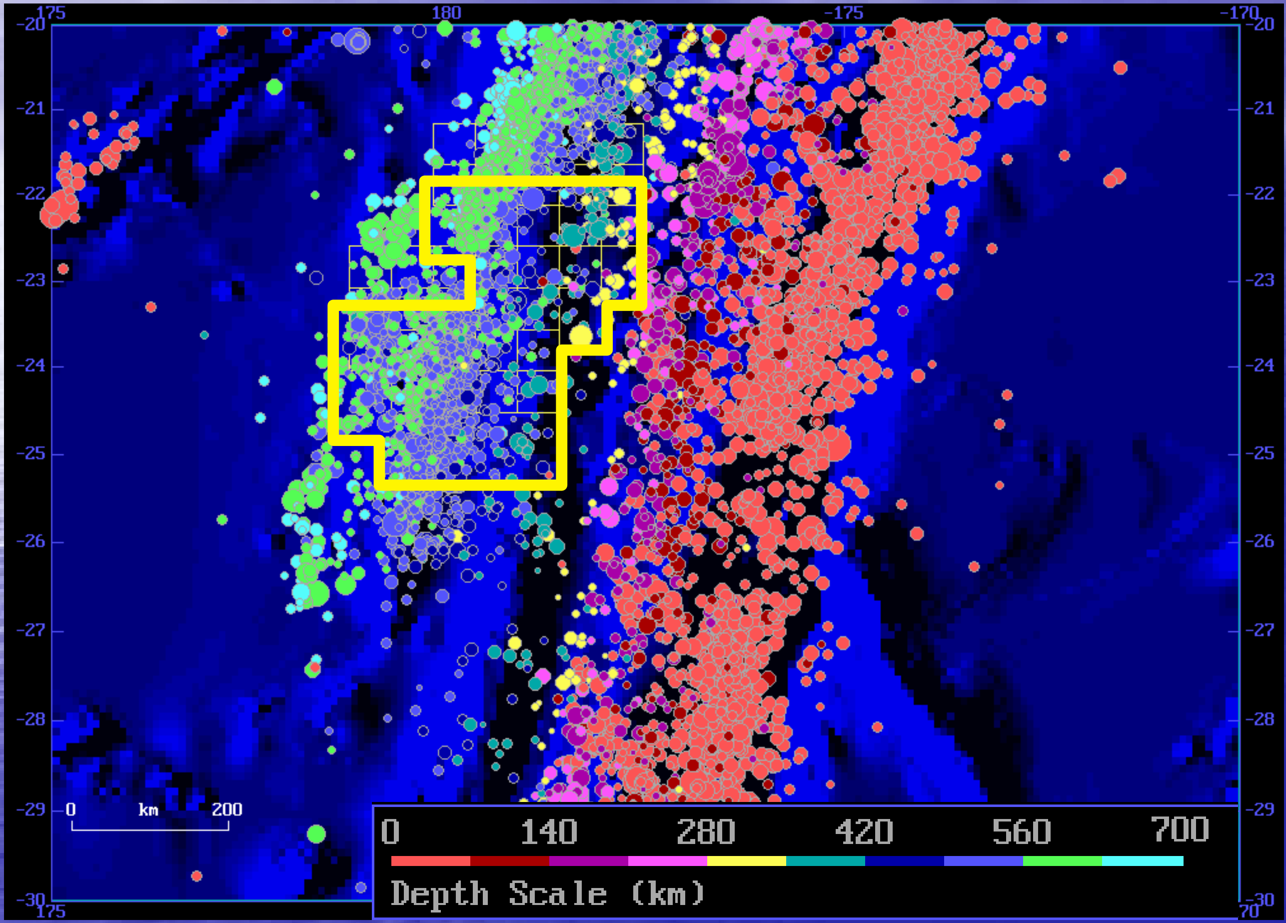
What was predicted...



- 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

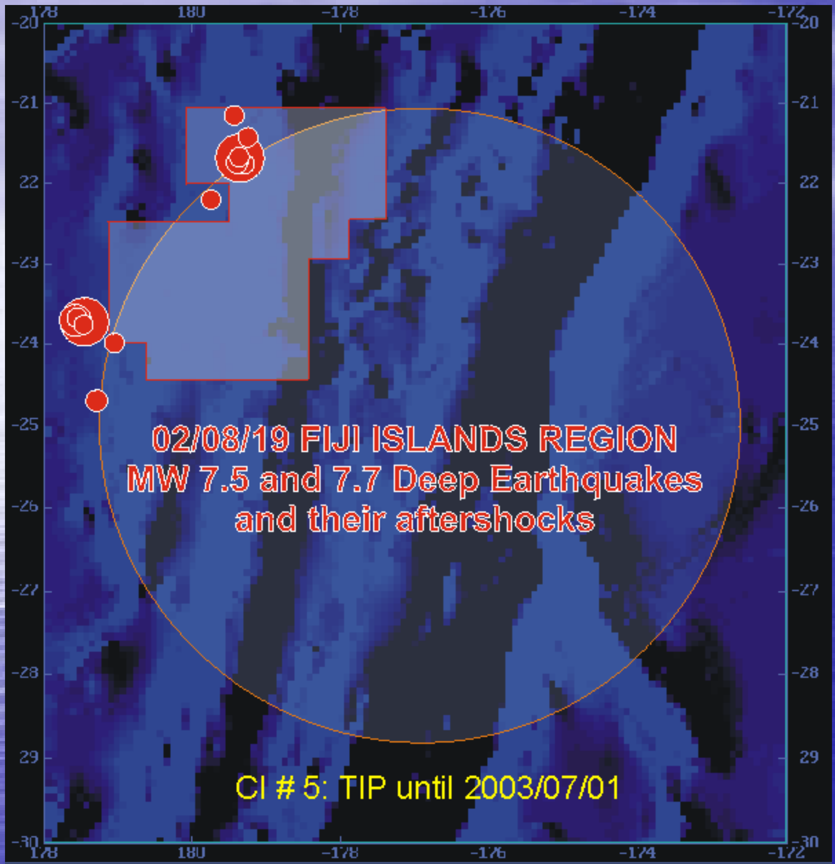
  24.52S - 21.16S and 178.76E - 177.53W.

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 –

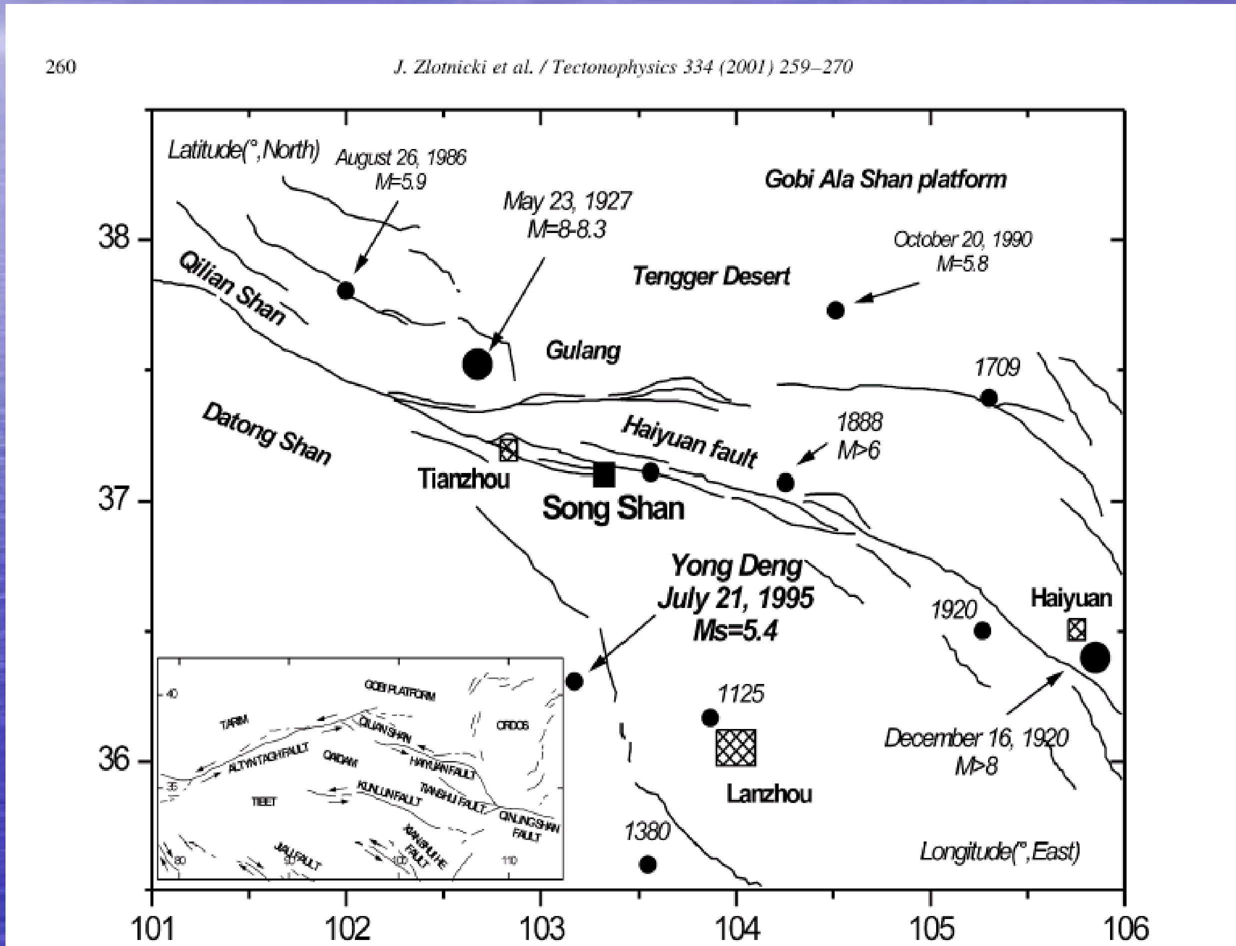
  FIJI ISLANDS REGION and
  SOUTH OF FIJI ISLANDS.

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.

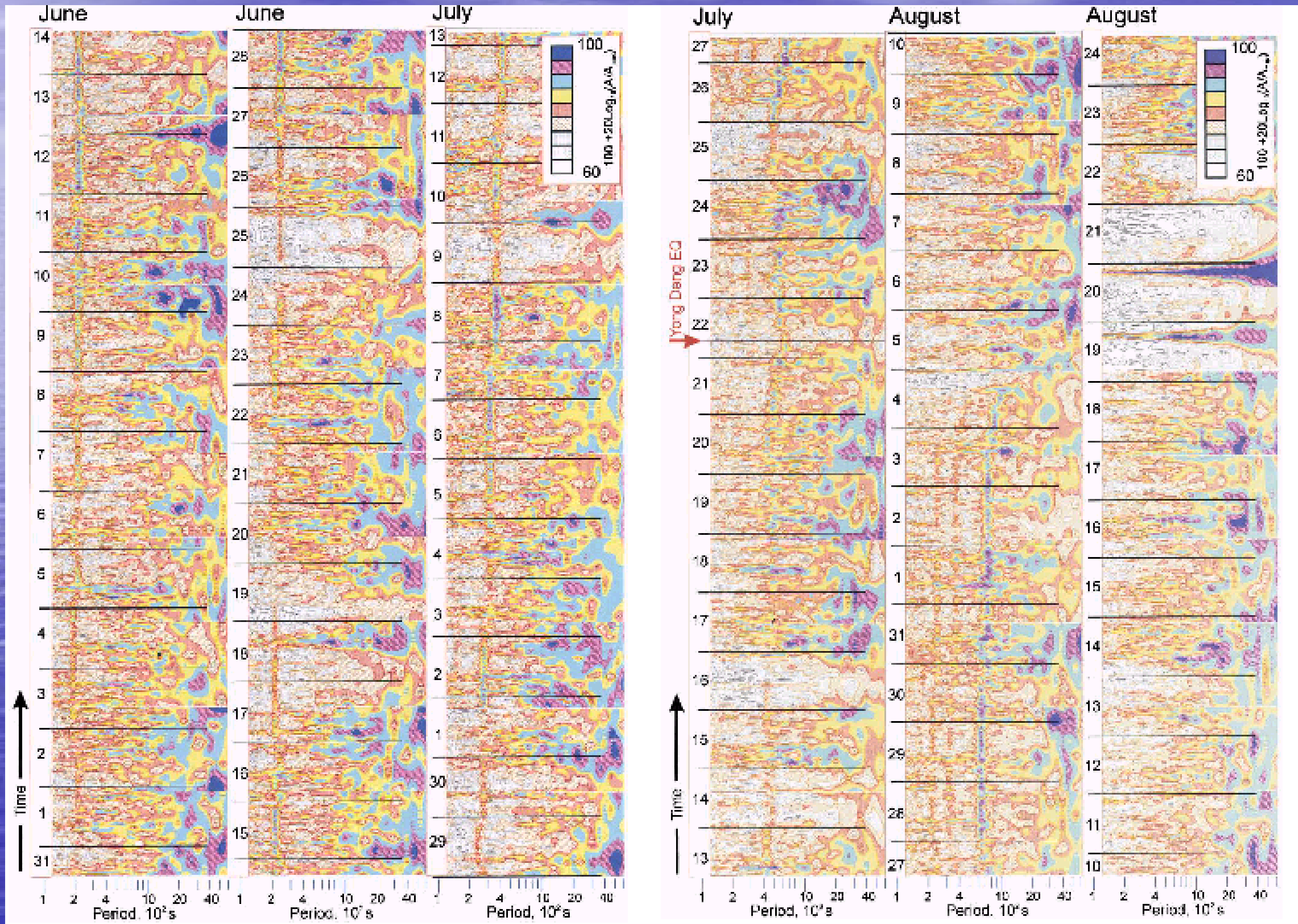
Thus, the accuracy achieved by M8 and MSc algorithms in the on-going Global testing is intermediate in time domain and varies from middle to exact in space domain.

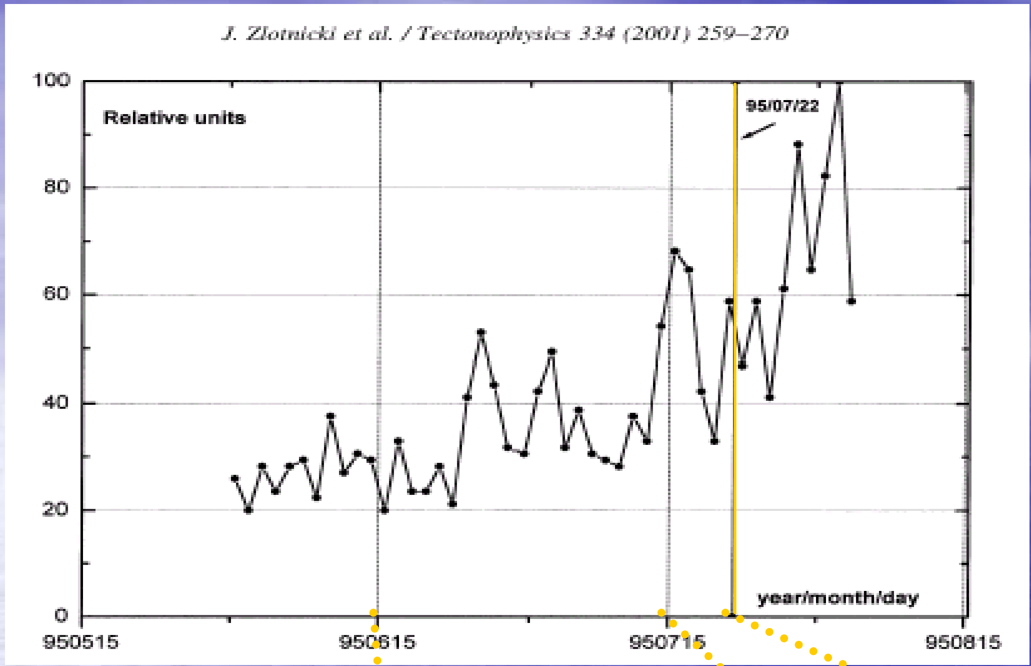
In some cases, the accuracy could be improved by making use of additional short-term monitoring of seismic activity and, perhaps, other geophysical fields in the alerted area of investigation.

One case-study of electromagnetic record about the site of 21 July 1995, M5.7 Yong Deng, China, earthquake in Tibet



FTAN diagram of the resistance observed on NS 250-m line

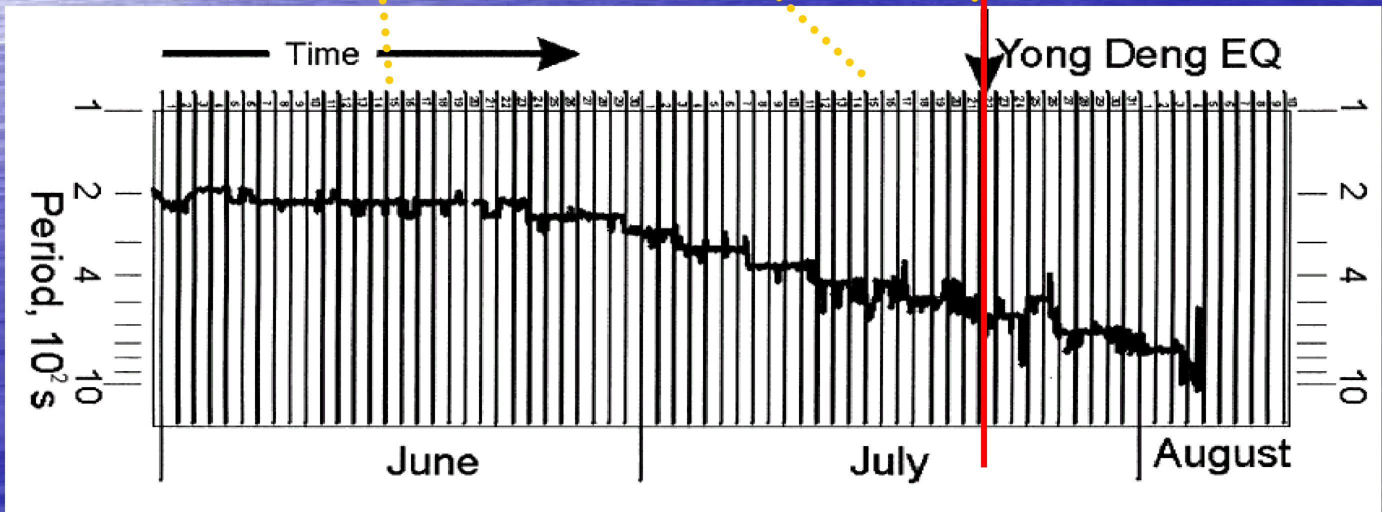


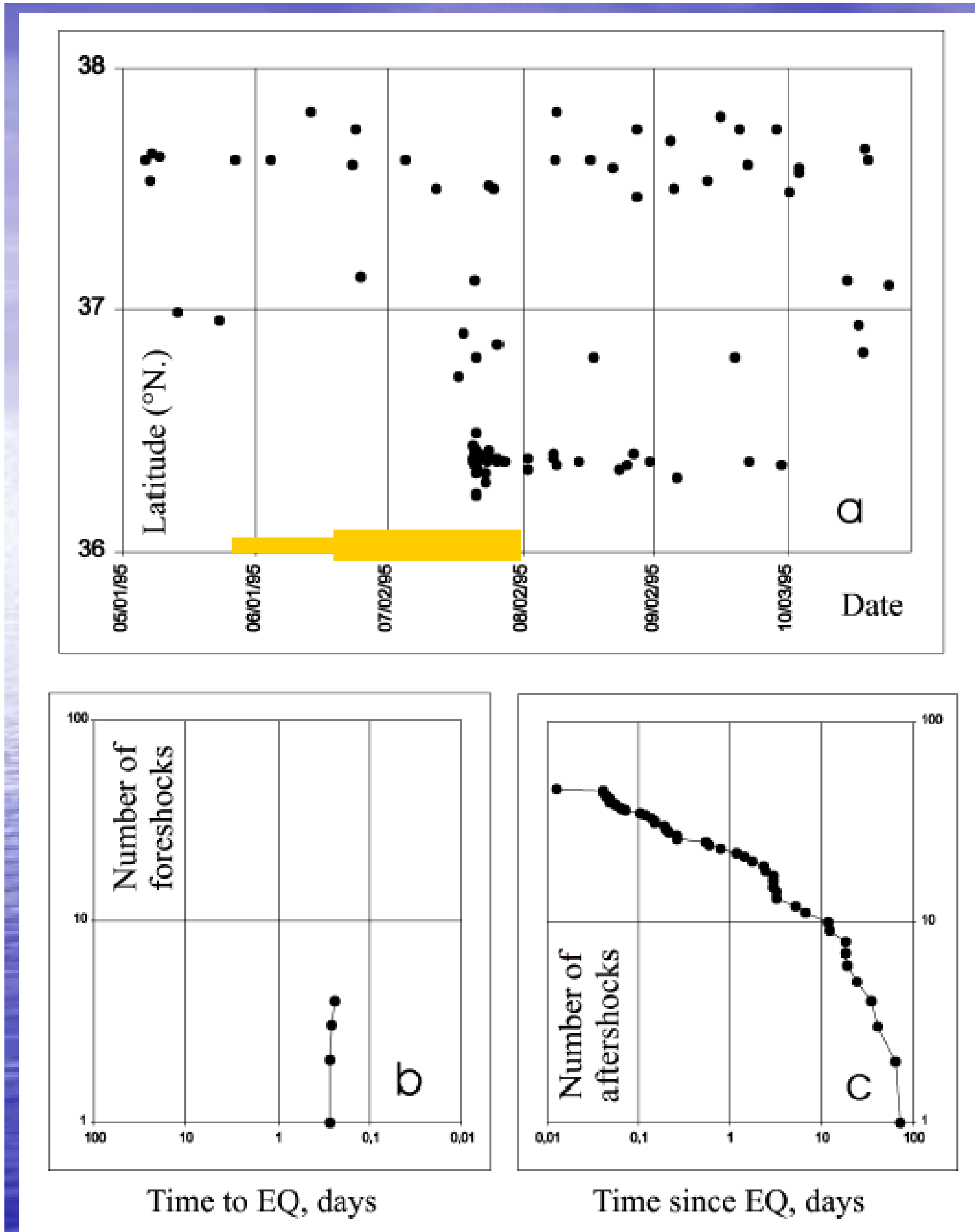


Evolution of the ULF signal

- Intensity and
- Period

GROW





Evolution of the ULF signal

- The 1995 Yong Deng earthquake occurred in less than 100 km from the instrument at the time of characteristic ULF and/or its power decay on component directed at the epicenter.
- The appearance of the ULF signal accompanied with a rise of seismic activity on adjusting segment of Haiyuan fault system.
- The characteristic ULF collapsed just before aftershocks fast disappeared (exponentially).

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, RTP, 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.