INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/709-17

Second Workshop on Non-Linear Dynamics and Earthquake Prediction

22 November - 10 December 1993

Goals and Optimal Decisions in Earthquake Prediction

G.M. MOLCHAN

Russian Academy of Sciences
International Institute of Earthquake Prediction
Theory and Mathematical Geophysics
Moscow 113556
Russian Federation

GOALS AND OPTIMAL DECISIONS IN EARTHQUAKE PREDICTION

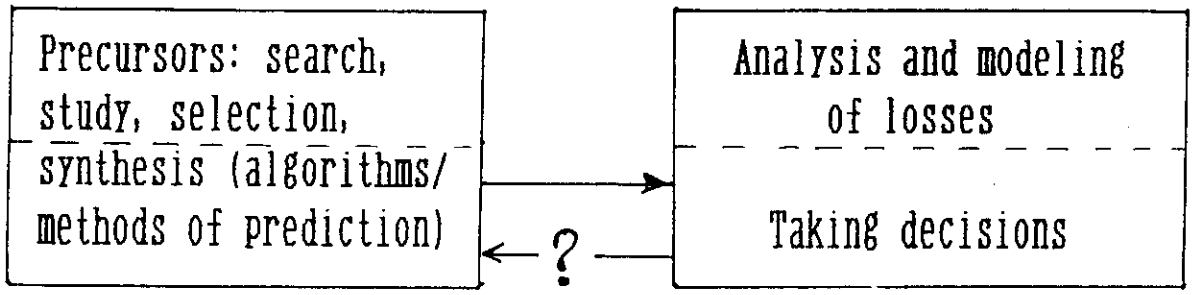
G.M. MOLCHAN

Inst. of Earthquake Prediction Theory
and Mathematical Geophysics
Academy of Sci. of the USSR

Prediction problem on the whole

Geophysical part

"Economical" part
(or goals on research stage)



Problem: modeling and optimization of losses

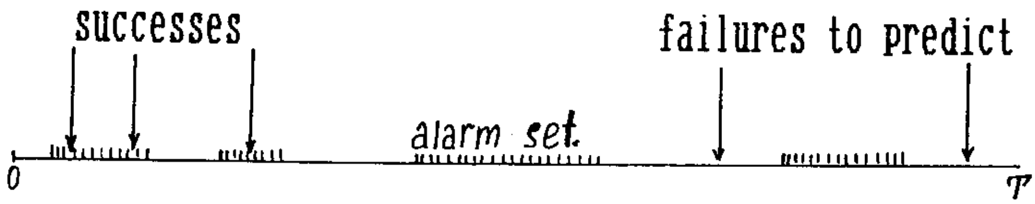
Goals: to understand
- how are synthesized precursors best ?
- which statistical properties of geophysical information J(t) should be analyzed for prediction?

The simplest optimization problem in prediction

Prediction policy:

$$\pi(t \mid \mathcal{J}_t) = \begin{cases} 1 & \text{alarm} \\ 0 & \text{no alarm} \end{cases} \text{ in } (t, t+\Delta) \text{ with probablity } \begin{matrix} p(\mathcal{J}_t) \\ 1 - p(\mathcal{J}_t) \end{matrix}$$

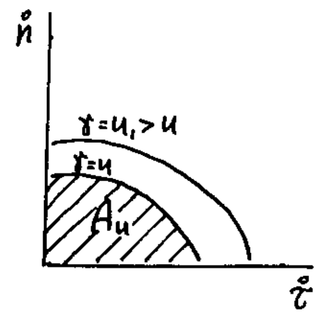
(The typical prediction policy: $p(J_t) = 1$ or 0)
J is geophysical information at moment t:
(state of geophysical fields in some intervals $(t-t_i,\ t-h_i)$,
catalogs of events etc.)



Long-term characteristics (errors) of prediction:

rate of failure-to-predict $\quad \mathring{n} = \dfrac{\#\{\text{failures-to-predict}\}}{\#\{\text{all strong events}\}}, \quad T \gg 1$

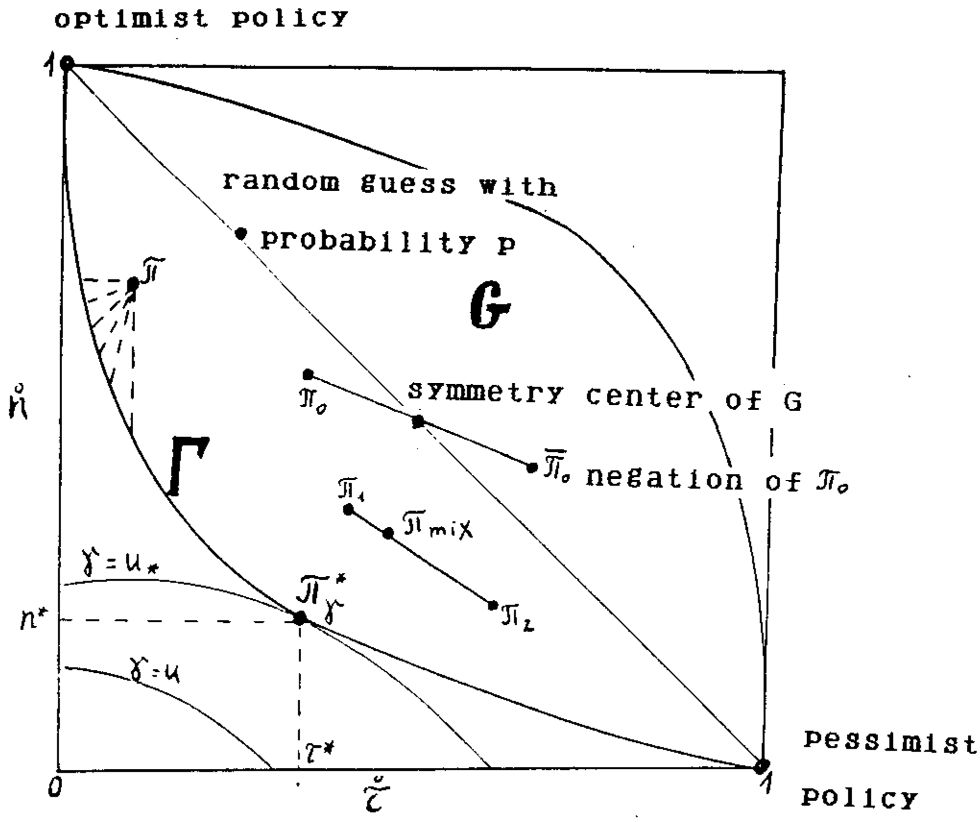
rate of alarm time $\quad \mathring{\tau} = \dfrac{\text{mes}\{\text{alarm sets}\}}{T}, \quad T \gg 1$



"Loss" function $\quad \gamma = f(\mathring{n}, \mathring{\tau})$
$A_u = \{\mathring{n}, \mathring{\tau} : \gamma < u\}$ is convex

Prediction goal: to minimize γ

ERRORS SET G FOR ALL POLICIES BASED ON J(t).



(i) G is convex: $\pi_{mix} = \begin{cases} \pi_1 \\ \pi_2 \end{cases}$ with probability $\begin{matrix} p \\ 1-p \end{matrix}$

(ii) G is set of the best policies based on J(t)

(iii) $\min_{\pi} \gamma = f(\mathring{n}^*, \mathring{\tau}^*) = u_*$

The diagram **Γ** characterizes the limit capabilities of information J(t) in earthquake prediction

⇓

Estimation of **Γ** is geophysical part of prediction problem

Examples: $(\mathring{n}, \mathring{\tau})$ for different prediction methods.

STRUCTURE OF γ-OPTIMAL POLICY, $\pi^*_\gamma(t)$

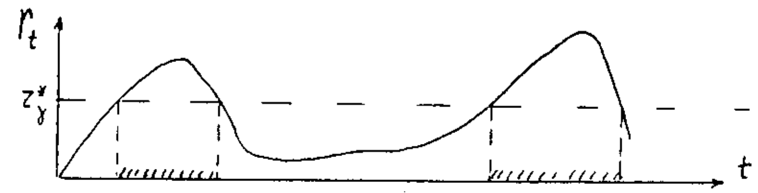
HAZARD FUNCTION (conditional intensity of events):

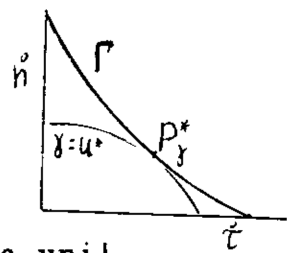
$$r_t = \Pr\{ n(t) > 0 \mid J(t) = u \} / \Delta \doteq r(u)$$

where

$$n(t) = \#\{ \text{events in } (t, t+\Delta) \}$$

If $\{ n(t), J(t) \}$ is stationary ergodic flows then
- optimal alarm set (//////////) is as follows:





where

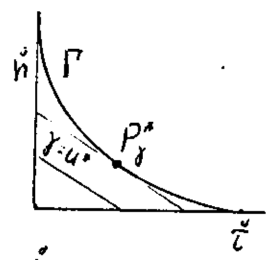
$$r^*_\gamma = -\lambda \frac{d\mathring{n}}{d\mathring{\tau}} (P^*_\gamma)$$

λ is number of events per time unit

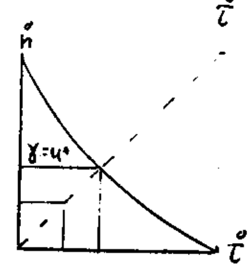
EXAMPLES

1. Linear losses : $\gamma = \alpha \lambda \mathring{n} + \beta \mathring{\tau} \quad \Rightarrow \quad r^*_\gamma = \beta / \alpha$

α is losses for one failure-to-predict
β is losses for alarm time unit in g
γ is total losses of prediction per time unit

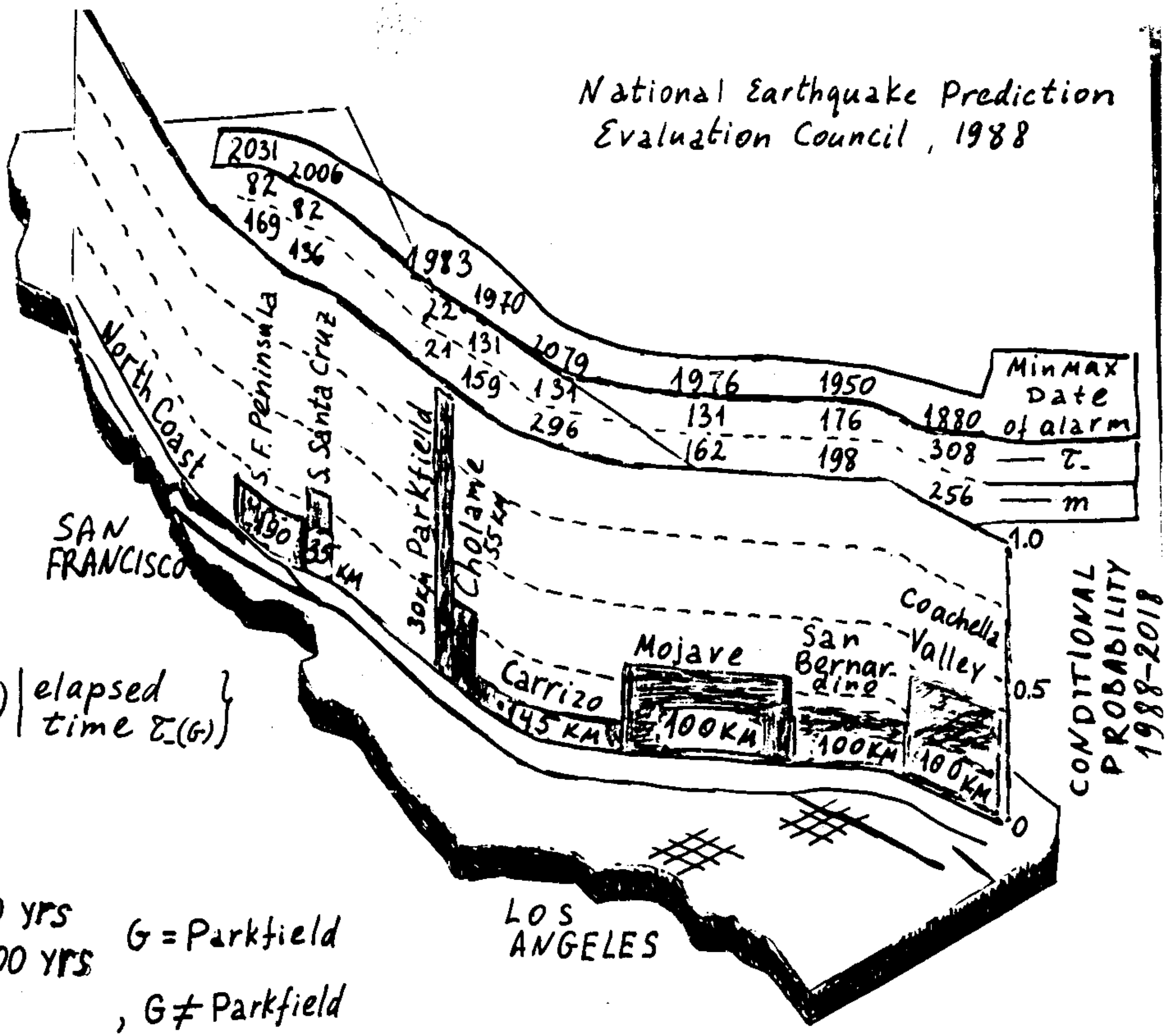


2. Minimax policy: $\gamma \Rightarrow \min$
   where $\gamma = \max(\mathring{n}, \mathring{\tau})$ $\Leftrightarrow$ $\mathring{n} = \mathring{\tau} \Rightarrow \min_{\pi}$



CN algorithm (Keilis-Borok, Rotwain) gives: $\mathring{n} = \mathring{\tau} = 20 - 25\%$

CONDITIONAL PROBABILITY OF MAJOR EARTHQUAKES ALONG SEGMENTS OF THE SAN ANDREAS FAULT 1988-2018



Log-Normal model $F(x)$ with

$m_G = 22 - 296$ yer

$\frac{\sigma_G}{m_G} = 0.24 - 0.6$

$$R = Pr\{t_i \in (1988, 2018) \mid \text{elapsed time } \tau_-(G)\}$$

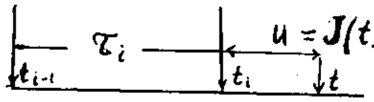
$$R = \begin{cases} > 0.9, & \tau_- < 800 \text{ yrs} \quad G = \text{Parkfield} \\ \simeq 0.5, & \tau_- > 4000 \text{ yrs} \\ < 0.45, & \tau_- > 0, \quad G \neq \text{Parkfield} \end{cases}$$

Conclusions by R-test:
Parkfield is the most dangerous region at any moment of time

EARTHQUAKE PREDICTION ON SAN ANDREAS FAULT
(National Earthquake Prediction Evaluation Council):

$R = \Pr\{\text{events within } (t,\ t+30\text{years}) \mid J(t)\}$

where $J(t)$ = (time elapsed since the last event)



1^o Value of R_t cannot be considered as a lucky choice for comparison of different segments;

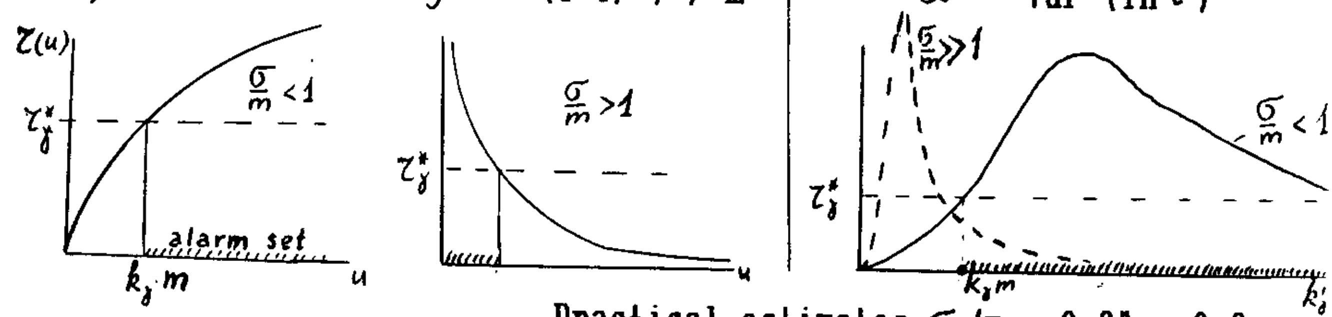
R $\leqslant .45$ constantly for all segments except Parkfield
$\geqslant .90$ in Parkfield for $0 < u < 800$ years

2^o If $F(u) = \Pr\{\tau_i < u\}$ is interevent time distribution then hazard function

$$r(u) = \frac{F'(u)}{1 - F(u)}, \quad u > 0$$

<u>EXAMPLES of F(u):</u> $E\tau = m$, $Var\ \tau = \sigma$

Gamma	Weibull	Log-Normal
$F'(u) = c\, u^{\alpha-1} e^{-\beta u}$	$F(u) = 1 - e^{-(\beta u)^\alpha}$	$F(u) = \Phi(1/\alpha \cdot \ln(u/m) + \alpha/2)$
$\beta = \alpha / m$	$\beta = \Gamma(1+1/\alpha) / m$	$\alpha^2 = Var(\ln\tau)$



Practical estimates $\sigma/m = 0.24 - 0.6$

If prediction goals are fixed (γ), then optimal decision for alarm is defined for the whole interevent time, i.e. R dynamics is not required.

Conclusions for minimax policy

- prediction results ($\dot{n}, \overset{o}{\tau}$) and normalized time threshold (k) of alarm are in weak dependence with type of distribution F when $\sigma/m < 1$;
- 5 out of 8 segments of San Andreas fault should be in state of (minimax) alarm

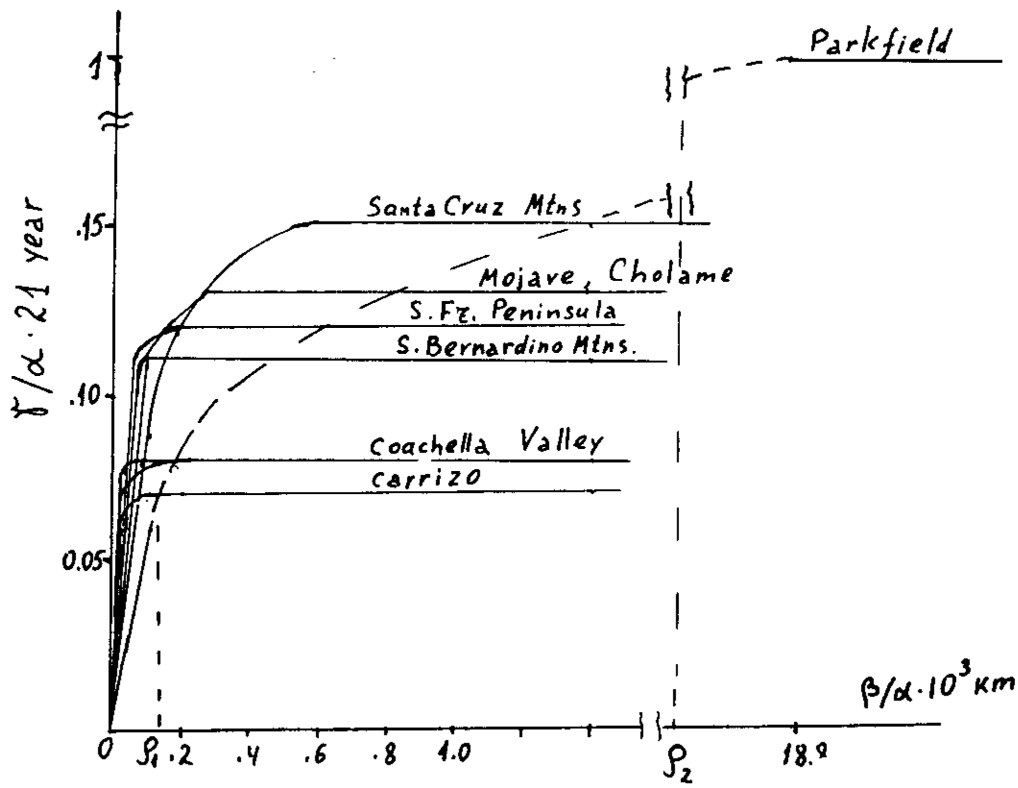
Conclusions for losses: $\gamma = \alpha_G \lambda_G \overset{o}{n} + \beta |G| \overset{o}{\tau}$

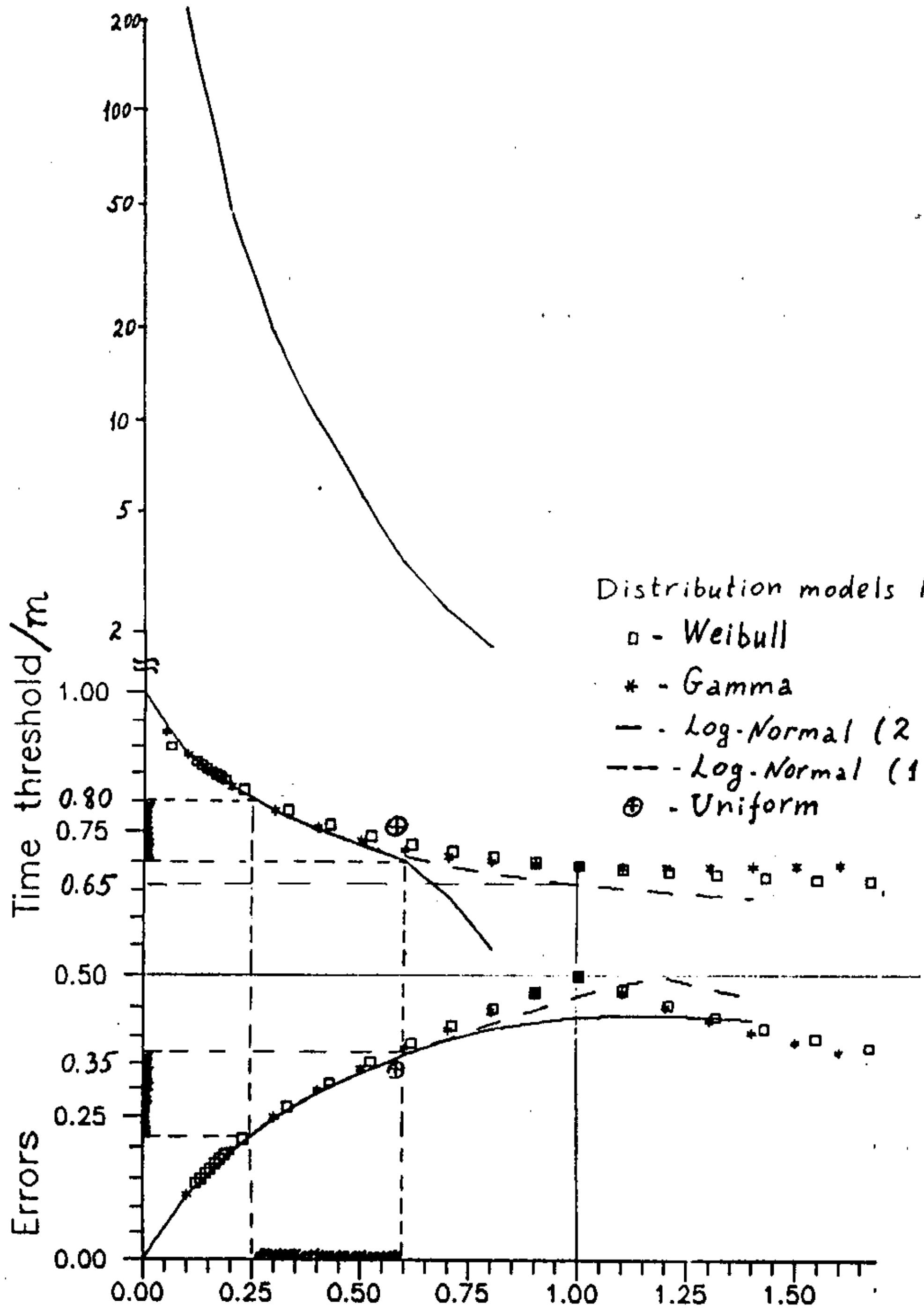
i) If $\alpha_G = \alpha$

$$\gamma(\text{Parkfield}) = \begin{cases} \min_i \gamma(G_i) \\ \max_i \gamma(G_i) \end{cases} \quad \text{for} \quad \begin{matrix} \beta/\alpha < \rho_1 \\ \beta/\alpha > \rho_2 \end{matrix}$$

ii) If $\alpha_G = \alpha \cdot m_G = \alpha / \lambda_G$ (natural model in insurance) then

$$\gamma(\text{Parkfield}) = \min_i \gamma(G_i) \quad \text{for any } \beta/\alpha$$





Parameters of the minimax strategies

MULTIPHASE ALARM MODEL

$A = \{A_i\}$ – phases (types) of alarm;

$\alpha = \{\alpha_i\}$ – prevention losses:

alarm A_i prevents losses α_i in case of strong event;

$\beta = \{\beta_i\}$ – cost of alarm maintenance:

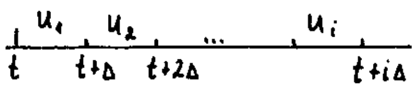
A_i requires expenses β_i per time unit

$C = [C_{ij}]$ – cost matrix for alarm phase change:

$A_i \to A_j$ requires immediate expenses $C_{ij} \geq 0$; $C_{ii} = 0$;

$C_{ij} = \infty$ if alarms transition $A_i \to A_j$ is not permitted.

LOSS FUNCTION



prediction losses for policy in time intervals Δ_i

$S_\pi(t) = u_1 + q\,u_2 + q^2 u_3 + \dots$ – total discounted losses in (t, ∞) for π

$q = \exp(-k\Delta)$ – discount factor

$k > 0$ – coefficient of efficiency for capital investment

GOALS of PREDICTION

to minimize the total discounted expected losses

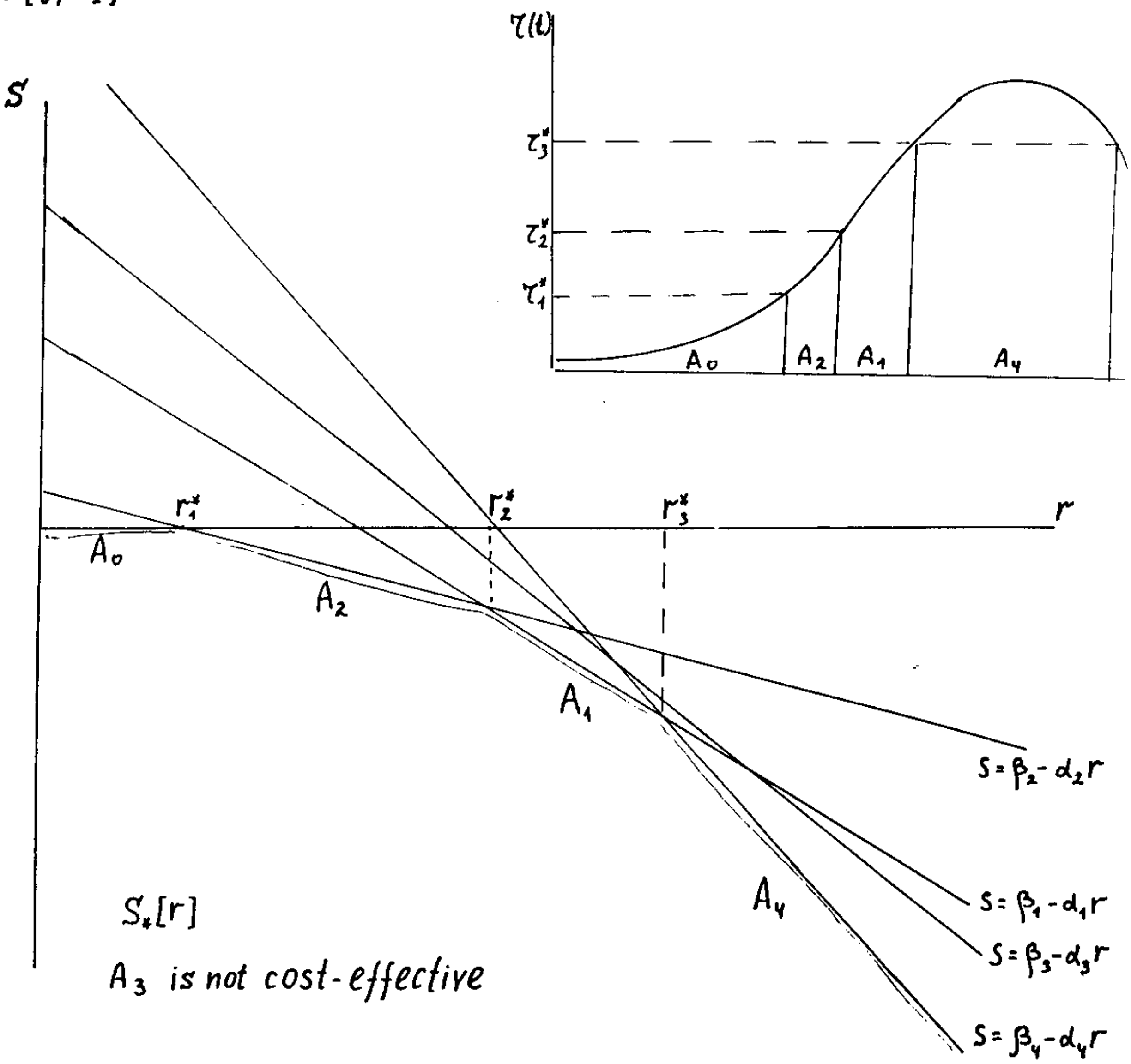
$$E\,S_\pi \Rightarrow \min_\pi$$

Unordered multiphase alarm without initial cost:

$$C_{ij} = 0\,, \quad \varepsilon_{ij} = 1$$

$$S_*^q(u,\mu) = \min_j \left(\beta_j - \alpha_j r(u)\right)\cdot\frac{\Delta}{1-q} = S_*[r]\cdot k^{-1}$$

i.e. global optimal prediction policy π_q^*, $q\in(0,1)$ coincides with local (in time) optimal policy π_0^* and does not depend on $q\in[0,1]$.



OPTIMAL PREDICTION

If

$$\pi(t) = f[J(t), \pi(s), s<t]$$

$$Pr\{n(t)>0 \mid J(t)=u, J(s), s<t\} = r(u)\Delta$$

$$Pr\{J(t+\Delta) \mid J(t)=u, J(s), s<t\} = P_{u,v}$$

(Markov Property)

then

i) the optimal expected losses $S_*(U,i)$ under the initial conditions $J(t) = u$ and $\pi(t-\Delta) = i$ are given as solution of the following equation

$$(*) \quad S(u,i) = \min_j \left(c_{ij} + \beta_j \Delta - r(u)\cdot\Delta\cdot\alpha_j + q \sum_v P_{u,v} S(v,j)\right) \equiv T_q S$$

ii) the equation $S = T_q S$ has a unique solution S_*^q :

$$S_0^q = 0 \; ; \; S_{n+1}^q = T_q S_n^q \longrightarrow S_*^q \; :$$

$$\max | S_n^q - S_*^q | < L q^n (1-q)^{-1}$$

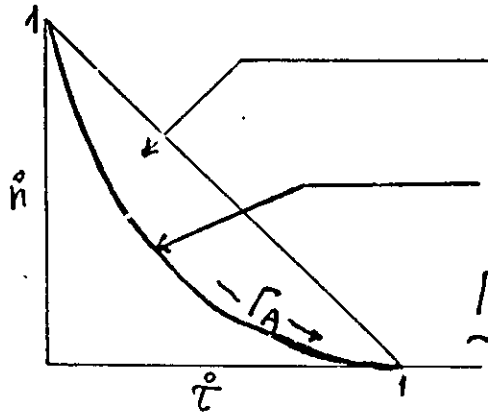
OPTIMAL DECISION:

$\pi(t) = A_j$ under condition $J(t) = u$, $\pi(t-\Delta) = i$
if the index j relizes the minimum in (*)

iii) if $q = 1$ (no discount factor) then S_N^q is optimal expected losses in the interval $(0, N\Delta)$.

PREDICTION ALGORITHMS COMPARED

p - internal parameters of the algorithm *A (or precursor)*

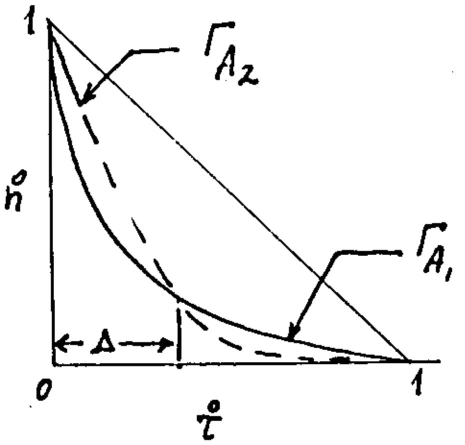


$\pi_A(p)$ - strategies of the algorithm

$\mathcal{T}_A(p)$ - best strategies of the algorithm

Γ_A - curve line of errors characterises predictive ability of the algorithm *(or precursor)*

(estimate of Γ_A , dim p, strong events statistics) - empiric characteristic feature of predictive ability of the algorithm.



Δ- area where A_1 is better than A_2

- Conclusion does not depend on the choic losses.

Conclusions

1. At the research stage of prediction it is necessary to estimate errors diagrams $(\hat{n}, \hat{\tau})$ proceeding from various types of information $J(t)$.
   Estimating of $(\hat{n}, \hat{\tau})$ errors for different prediction algorithms is to be used for this problem.

2. The task of seismostatistics in general case is to estimate the hazard function $r(u)$ and the transition probability matrix $[p_{u,v}] = \mathcal{P}$ for information J on sequential time intervals. ($r(u)$ and $\mathcal{P}$ depend not on time but on the discrete state of seismic process!)
   Aki's method of $r(u)$ estimating for a set of precursors overestimates $r(u)$, as soon as it is based on the assumption of precursors conditional independency.
   To analyze P systematic descriptions of strong shock preparation processes are very important.

3. There is only one hazard function threshold r^* in the simplest prediction problem, here we can hope for the prediction stability.
   In real prediction problems there should be several hazard function thresholds. In complicate cases we must know $r(u)$ and $p_{u,v}$ in details.

   So compromises are desirable.

REFERENCES

Aki, K., (1989). *Ideal probabilistic earthquake prediction*, Tectonophysics, 169, 197-198.

Ellis, S. P., (1985). *An optimal statistical decision rule for calling earthquake alerts*, Earthquake Prediction Research, 3, 1-10.

Feng, De Yi, Jing Ping Gu, Ming Zhou Lin, Shao Xie Xu, and Hue Jun Yu, (1985). *Assessment of earthquake hazard by simultaneous use of the statistical method and the method of fuzzy mathematics*, PAGEOPH, 126, 982-997.

Gusev, A. A., (1976). *Indicator earthquakes and prediction*, in Seismicity and Deep Structure of Siberia and Far East (in Russian), pp. 241-247, Nauka, Novosibirsk.

Howard, R. A., (1960). *Dynamic Programming and Markov Processes*, 136 pp., New York.

Hawkes, A,G., (1971). *Spectra of some self-exciting and mutually exciting point processes.* Biometrica, 58, 83-90.

Kagan, Y., (1973). *On a probabilistic description of the seismic regime.* Fizika Zemli, 4, 110-123.

Kagan, Y.Y., and D.D.Jackson, (1991). *Seismic Gap Hypothesis: Ten Years After.* J.Geophys. Res. 96, 21419-21431.

Kantorovich, L. V., and V. I. Keilis-Borok, *Economics of Earthquake Prediction,* Proceedings of UNESCO conference on seismic risk, Paris, 1977.

Keilis-Borok, V.I., and I.M.Rotwain, (1990). *Diagnosis of Time of Increased Probability of strong earthquakes in different regions of the world: algorithm CN.* Phys. Earth and Planet. Inter., 61, 57-72.

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

Khokhlov, A.V. and V.G.Kossobokov, (1992). *Seismic flux and major earthquakes in the North-Western Pacific* (in Russian). Doklady RAN, 325, N 1, 60-63.

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

Molchan, G. M., (1990). *Strategies in strong earthquake prediction,* Phys. Earth and Planet. Inter., 61, 84-98.

Molchan, G.M., (1991). *Structure of optimal strateies in earthquake prediction,* Tectonophysics, 193, 267-276.

Molchan, G. M., (1992). *Models for optimization of earthquake prediction* (in Russian) in Problems in earthquake prediction and interpretation of seismologia data, Vychislitelnaya Seismologiya, 25, pp.7-28, Nauka, Moscow.

Molchan, G. M., and Y. Y. Kagan, (1992). *Earthquake prediction and its optimization,* J. Geophys. Res., 97, 4823-4838.

Narkunskaya, R. S., and M. G. Shnirman, (1993). *On an algorithm of earthquake prediction*, in Computational Seismology and Geodynamics, 22/23 (English Transl.), Am. Geophys. Union.

Nishenko, S. P., (1989). *Circum-Pacific earthquake potential*: 1989-1999, 126 pp., *USGS*, Open-file report 89-86.

Ross, M., (1970). *Applied Probability Models with Optimization Applications*, 198 pp., San Francisco.

Ogata, Y., (1988), *Statistical Models for Earthquake occurrences and Residual Analysis for Point Proceses*, J. American Statistical Association, 83, No 401, Applications, 9-26.

Sadovsky, M. A. (Ed.), (1986). *Long-Term Earthquake Prediction Methodology* (in Russian), 128 pp., Nauka, Moscow.

Sobolev, G. A., T. L. Chelidze, A. D. Zavyalov, L. B. Slavina, and V. E. Nicoladze, (1991). *Maps of expected earthquakes based on combination of parameters*, Tectonophysics, 193, 255-265.

Vere-Jones, D., (1978). *Earthquake prediction – a statistician's view*, J. Phys. Earth, 26, 129-146.

Working Group on California Earthquake Probabilities, (1988). *Probabilities of large earthquakes occuring in California on the San Andreas fault*, U.S. Gcol. Surv, Open File Rep., 88-398, 66pp.