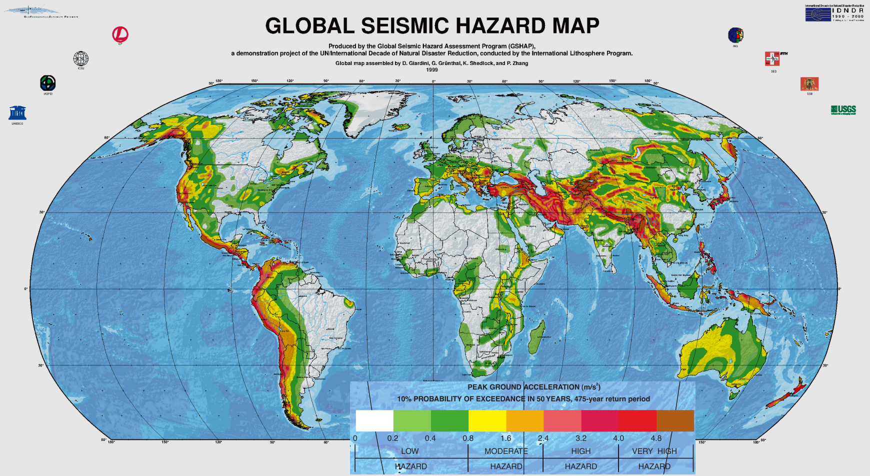
# Hazard, Risk and Prediction

Stefan Hergarten

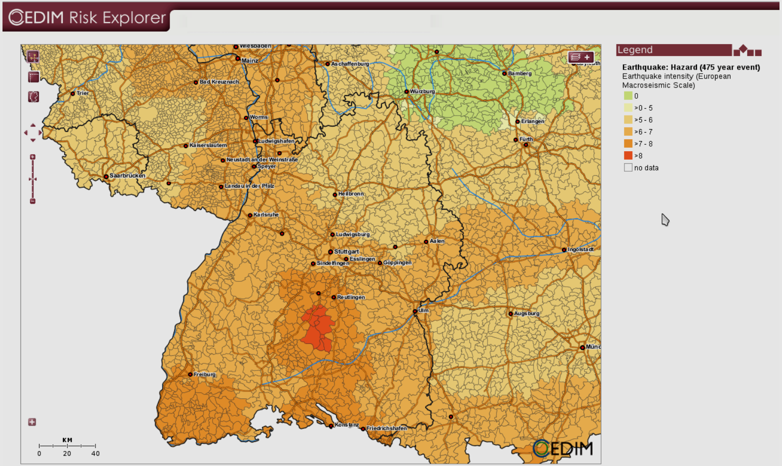
Institut für Geo- und Umweltnaturwissenschaften
Albert-Ludwigs-Universität Freiburg

## Worldwide Earthquake Hazard



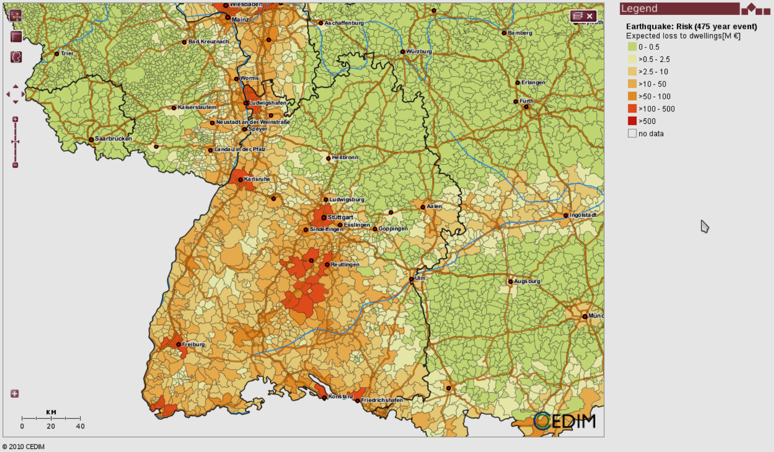

Source: Global Seismic Hazard Assessment Program

## Regional Earthquake Hazard



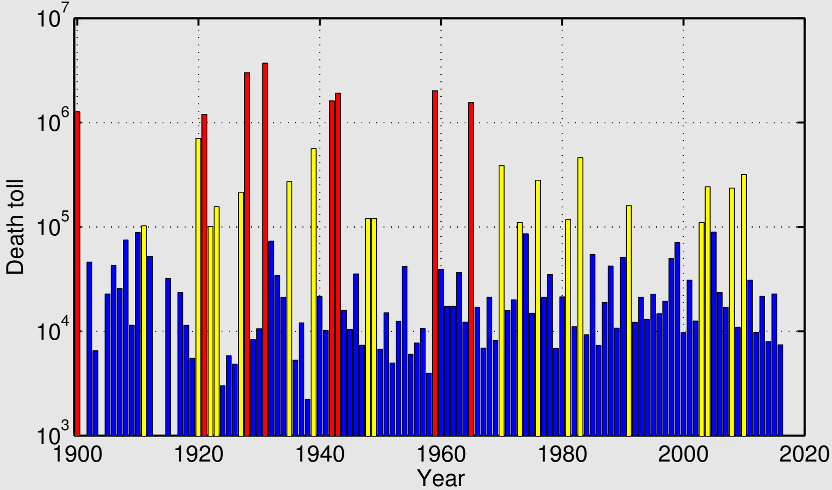

Source: CEDIM Risk Explorer (KIT / GFZ Potsdam)

## Regional Earthquake Risk

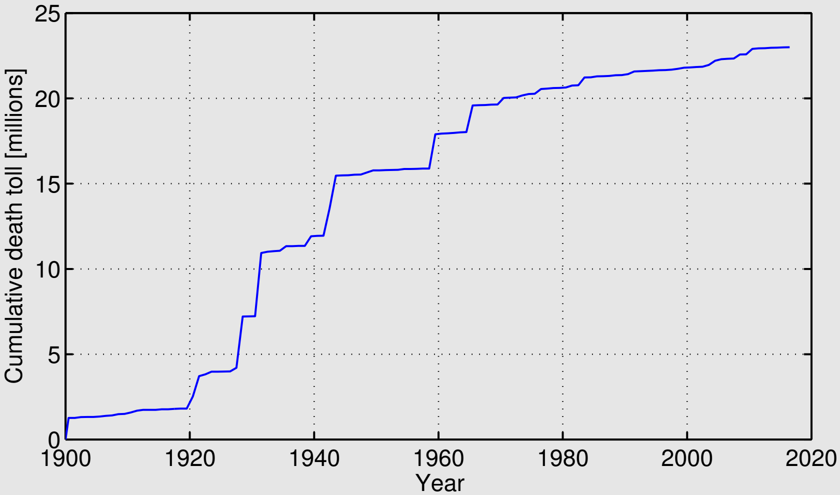



Source: CEDIM Risk Explorer (KIT / GFZ Potsdam)

## Worldwide Death Toll of all Geohazards since 1900

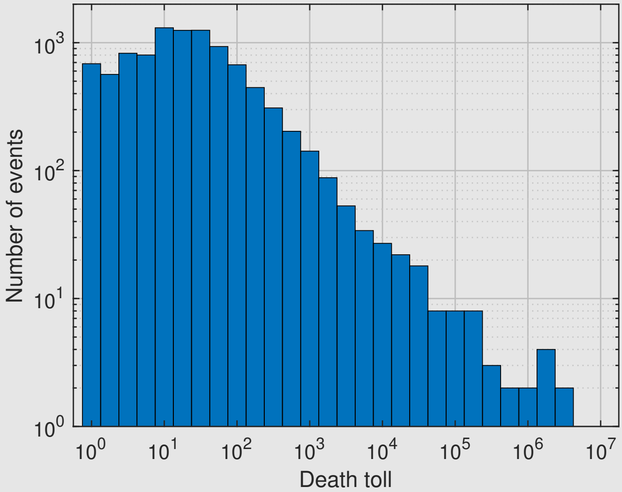


## Worldwide Death Toll of all Geohazards Since 1900

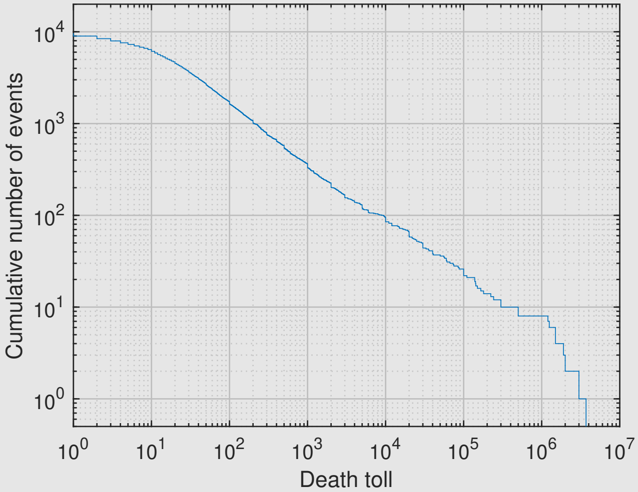


# Frequency-Magnitude Relations

## Binning



## Cumulative Frequency



# Frequency-Magnitude Relations

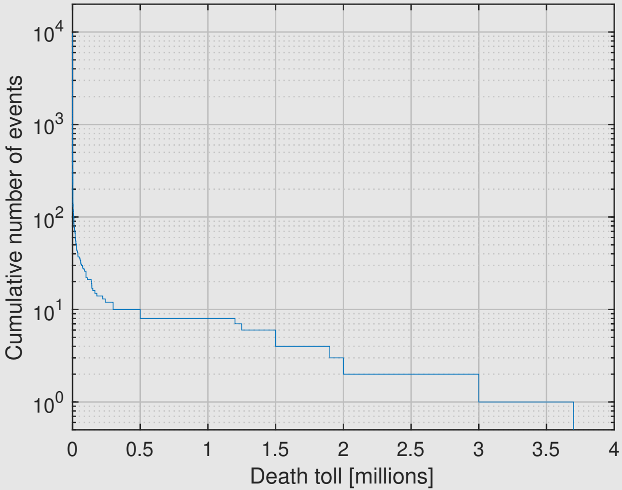
## Cumulative Frequency

$$F(s) \; = \; \text{expected number of events with sizes} \geq s$$

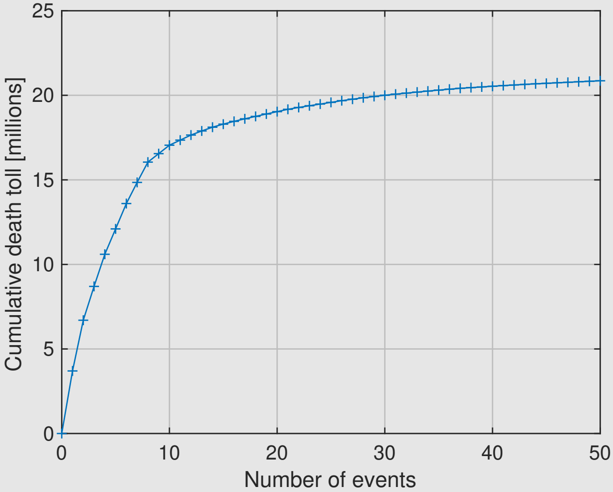
- Can be either considered for a given region (or worldwide) or per domain size (area).
- Can be either considered for a given time interval or per time.
- Often called frequency-magnitude relation.

# Frequency-Magnitude Relations

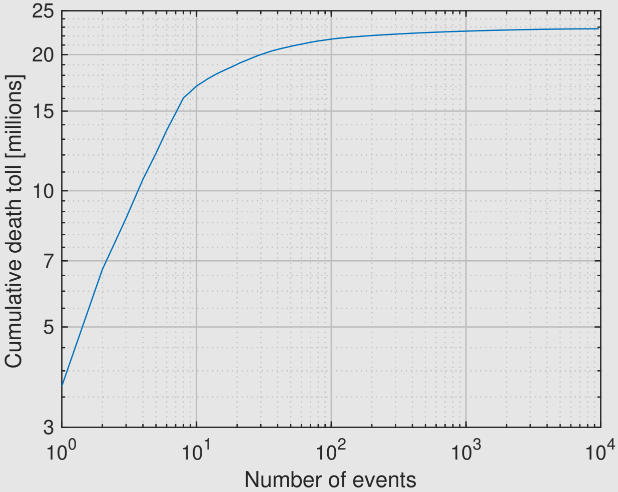
## Cumulative Frequency



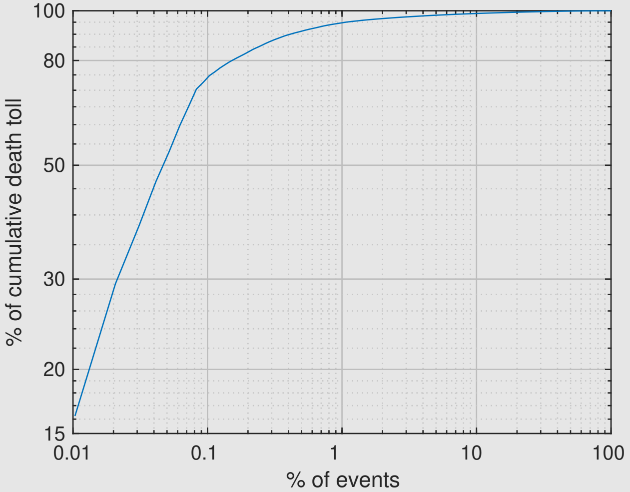
## Pareto Diagram of all Geohazards Since 1900



## Pareto Diagram of all Geohazards Since 1900



## Pareto Diagram of all Geohazards Since 1900



## Frequency Density

$$f(s) \;=\; -F'(s)$$

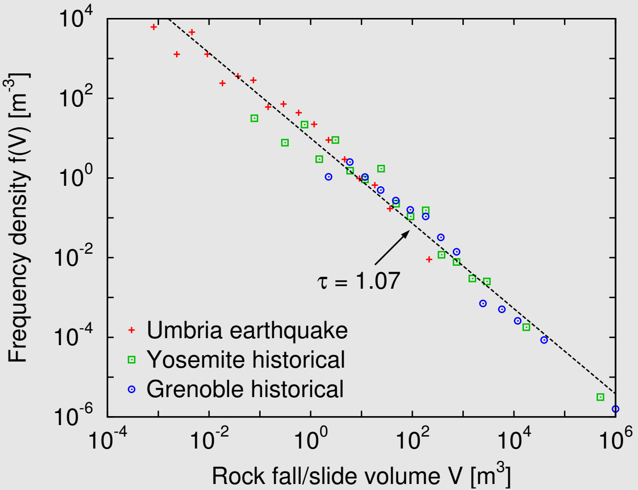
$$\Downarrow$$

$$\int_{s_1}^{s_2} f(s)\,ds \;=\; F(s_1) - F(s_2)$$

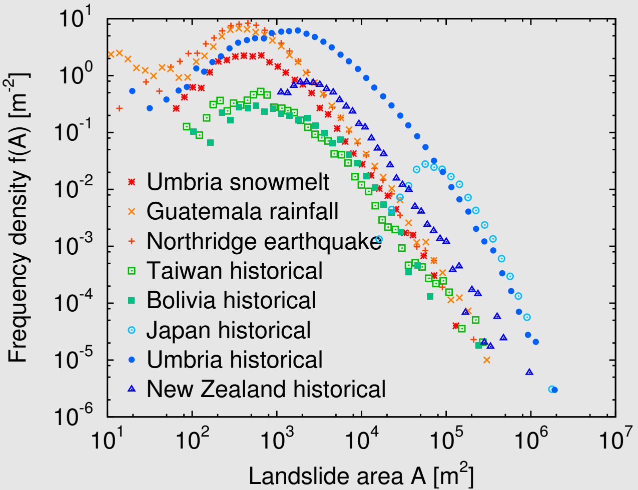
is the expected number of events with sizes between $s_1$ and $s_2$.

# Frequency-Magnitude Relations

## Frequency Density of Rockfalls



## Frequency Density of Regolith Landslides



## Cumulative Probability

$$P(s) = \frac{F(s)}{F(s_0)}$$

($s_0$ = smallest possible event size) is the probability that the size of a randomly picked event is $\geq s$.

- Often $s_0 = 0$ or $s_0 = -\infty$
- In mathematics defined as the probability that a value drawn from a random distribution is $\leq s$.

## Probability Density

$$p(s) \;=\; -\,P'(s)$$

$$\Downarrow$$

$$\int_{s_1}^{s_2} p(s)\,ds \;=\; P(s_1) - P(s_2)$$

the is the probability that the size of a randomly picked event is between $s_1$ and $s_2$.

# Frequency-Magnitude Relations

## Pareto Distribution

$$P(s) = \left(\frac{s}{s_0}\right)^{-b}$$

$$\Downarrow$$

$$\begin{aligned} p(s) &= -P'(s) \\ &= b\, s_0^b\, s^{-b-1} \\ &= \frac{b}{s_0}\left(\frac{s}{s_0}\right)^{-(b+1)} \end{aligned}$$

**Exponential Distribution**

$$P(s) \;=\; e^{-\lambda(s-s_0)}$$

$$\Downarrow$$

$$p(s) \;=\; \lambda\, e^{-\lambda(s-s_0)} \;=\; \lambda\, P(s)$$

## Pareto Distribution vs. Exponential Distribution

| Slope in diagram | Pareto Distribution | Exponential Distribution |
| --- | --- | --- |
| axis scaling | | |
| cumulative probability / frequency | | |
| probability / frequency density | | |
| logarithmically binned numbers | | |

## Expected Value of Pareto Distribution

$$\bar{s} = \int_{s_0}^{\infty} p(s)\, s\, ds = \begin{cases} \frac{b}{b-1}\, s_0 & \text{for } b > 1 \\ \infty & \text{for } b \leq 1 \end{cases}$$

## Expected Value of Exponential Distribution

$$\bar{s} = \int_{s_0}^{\infty} p(s)\, s\, ds = s_0 + \frac{1}{\lambda}$$

## Kolmogorov-Smirnov Test

Simplest test whether

- a given sample might come from a given statistical distribution
- two given samples might come from the same (unknown) statistical distribution

Properties:

The Kolmogorov-Smirnov test

- does not rely on a certain statistical distribution
- is not very sensitive

## Simple Example

3 farms produce apples with different colors at the following probabilities:

| | Red | Green | Blue |
|---|---|---|---|
| farm 1 | 0.6 | 0.3 | 0.1 |
| farm 2 | 0.3 | 0.4 | 0.3 |
| farm 3 | 0.5 | 0.4 | 0.1 |

In a shop we find a sample of 3 red apples, 2 green apples, and 1 blue apple without declaration. Which is the most likely source of this sample?

## Concept

Starting point:

- Sample of $n$ elements $s_1, \ldots, s_n$ from a given distribution.
- Probability density $p(s)$ depends on unknown parameters $\lambda_1, \ldots, \lambda_k$.

Task: Find the most likely values of $\lambda_1, \ldots, \lambda_k$.

Likelihood of the parameter set $\lambda_1, \ldots, \lambda_k$
= probability density for the given sample:

$$L(\lambda_1, \ldots, \lambda_k) \; = \; \prod_{i=1}^{n} p(s_i)$$

Find $\lambda_1, \ldots, \lambda_k$ that maximizes $L(\lambda_1, \ldots, \lambda_k)$.

## Technical Implementation

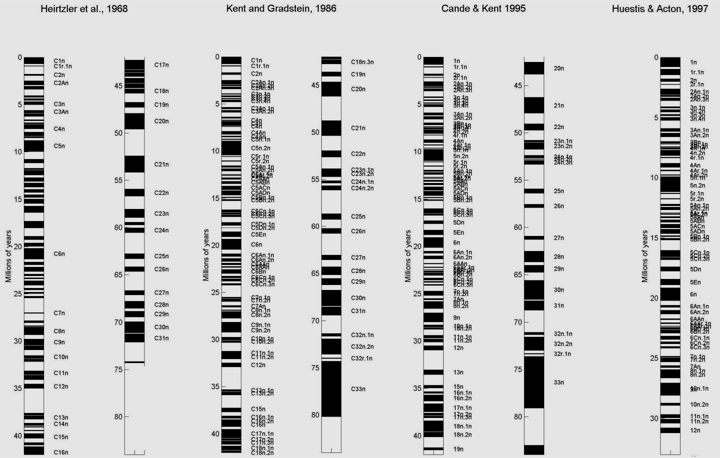
Minimize

$$-\ln L(\lambda_1, \dots, \lambda_k) \;=\; -\sum_{i=1}^{n} \ln p(s_i)$$

either numerically or by the condition(s)

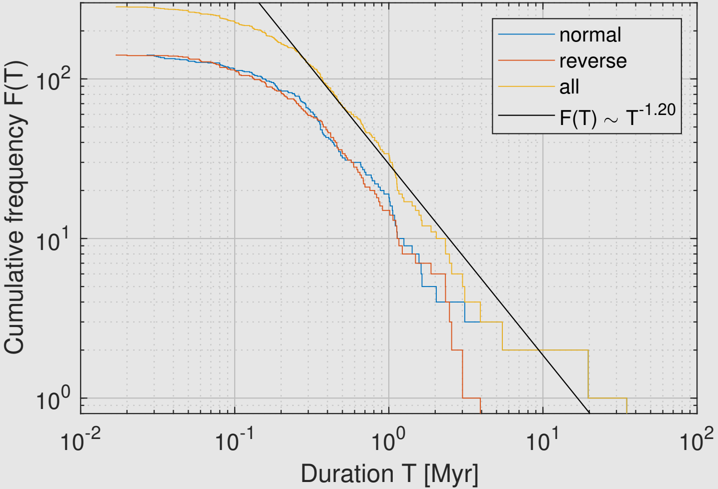
$$\frac{\partial}{\partial \lambda_i}\left(-\ln L(\lambda_1, \dots, \lambda_k)\right) \;=\; 0$$

## Magnetic Field Reversals




Quelle: Woods Hole Oceanographic Institution

## Magnetic Field Reversals



# Waiting-Time Distributions

## Theoretical Concept

$$T \quad = \quad \text{time since the last event took place}$$

Cumulative waiting-time distribution

$$\begin{aligned} P(T) \quad &= \quad \text{probability that a period of quiescence has} \\ &\quad\;\; \text{a length of at least } T \\ &= \quad \text{probability that there is no event until } T \end{aligned}$$

What is the meaning of

$$\lambda(T) = \frac{-\frac{d}{dT}P(T)}{P(T)} = \frac{\phantom{xxxx}}{P(T)} = -\frac{d}{dT}\phantom{xxxx} ?$$

# Waiting-Time Distributions

## Theoretical Concept

General solution:

$$P(T) \ = \ e^{-\int_0^T \lambda(t)\,dt}$$

Simplest situation: $\lambda(T) = \text{const}$:

$$P(T) \ = \ e^{-\lambda T}$$

For which model $\lambda(T)$ are the waiting times Pareto-distributed?

**Definition**

Assume events (or waiting-times) with a given distribution $P(s)$ and consider the expected value $\overline{s - s_0}$ for those events with $s \geq s_0$ for a given threshold size $s_0$.

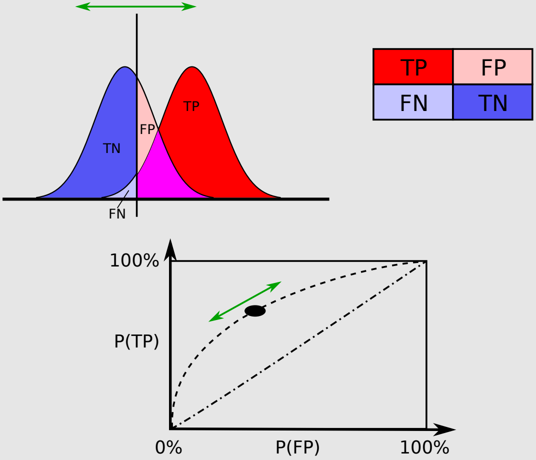
Light-tailed distribution: $\overline{s - s_0} \to 0$ for $s_0 \to \infty$

Heavy-tailed distribution: $\overline{s - s_0} \to \infty$ for $s_0 \to \infty$

Medium-tailed distribution: else

# Assessment of Predictions

## Receiver Operating Characteristic Curves




Source: Wikipedia, ©Sharpr

# Assessment of Predictions

## Receiver Operating Characteristic Curves

Normalization:

$$\begin{aligned} P(TP) + P(FN) &= 1 \\ P(TN) + P(FP) &= 1 \end{aligned}$$

$$\Downarrow$$

ROC curve only describes the test, but not the probability that an event occurs.

Example: COVID-19 test at an incidence of 500 per 100000, $P(TP) = 99\,\%$, $P(FP) = 1\,\%$. What does a positive test mean?

## Receiver Operating Characteristic Curves

Probabilities if $q =$ probability of occurrence:

| | alarm | no alarm |
|---|---|---|
| event | $q\,P(TP)$ | $q\,P(FN)$ |
| no event | $(1-q)\,P(FP)$ | $(1-q)\,P(TN)$ |

## Analysis of Benefit and Cost

$$
\begin{aligned}
R &= q\,P(TP)\,(L - M + C) + q\,P(FN)\,L \\
&\qquad + (1 - q)\,P(FP)\,C \\
&= q\,L - q\,(M - C)\,P(TP) + (1 - q)\,C\,P(FP)
\end{aligned}
$$

where

$R$ = total risk

$L$ = loss caused by an event

$M$ =

$C$ =

# Earthquake Prediction

## Debate in the Nature Magazine 1999

Topic: Is the reliable prediction of individual earthquakes a realistic scientific goal?

Extent: 26 contributions over 7 weeks

Outcomes concerning 4 levels of predictability:

| Level | Target | Consensus |
|---|---|---|
| 1 | time-independent hazard | yes |
| 2 | time-dependent hazard | |
| (a) | earthquake cycle | no consensus |
| (b) | clustering of earthquakes | yes |
| 3 | intermediate (1–10 yr) to short-term ($< 1$ yr) forecasting | not possible in the near future |
| 4 | deterministic prediction | no |

# Earthquake Prediction

## Earthquake Precursors

Two groups of precursors:

Seismic precursors: spatial and temporal pattern of seismicity

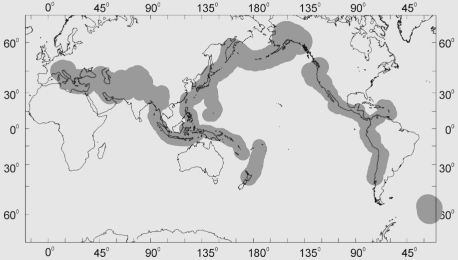
- Starting from analysis of foreshocks
- Several approaches; rather address intermediate-term (1–10 yr) forecasting than prediction

Non-seismic precursors: all other changes in the crust and the atmosphere that could announce an earthquake

- Gas emissions
- Water level changes in wells
- Electromagnetic signals
- ...

## The M8 Algorithm

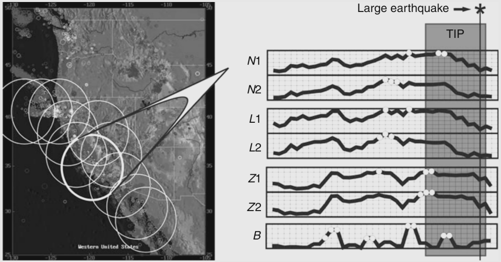
- Aims at forecasting earthquakes with $M \geq 8.0$.
- Based on a retrospective analysis of the seismic patterns prior to earthquakes with $M \geq 8.0$.
- Considers 262 overlapping circles of 668 km radius in the regions where such earthquakes occurred.




Source: Molchan & Romashkova, Geophys. J. Int., 2010

## The M8 Algorithm

- Derives 7 functions from the seismic activity in each circle.




Source: Ismail-Zadeh & Kossobokov, Encyclopedia of Solid Earth Geophysics, 2011

- Gives an alert (time of increased probability, TIP) for a 5-year period based on these functions.

## The M8 Algorithm

- Modified version: M8-MSc (Mendocino Scenario)
- Performance over a 25 year period:

| Version | Captured events | Total alarm |
|---|---|---|
| M8 | 13 out of 18 | 32.93 % |
| M8-MSc | 10 out of 18 | 16.78 % |

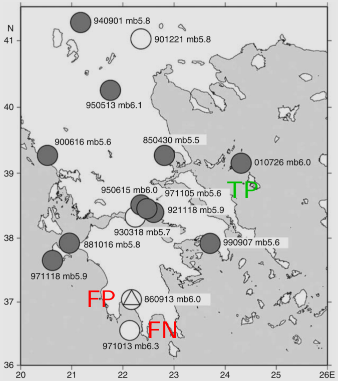
Source: Ismail-Zadeh & Kossobokov, Encyclopedia of Solid Earth Geophysics, 2011

## The VAN Method

- Developed by P. Varotsos, K. Alexopoulos, and K. Nomikos in the 1980s.
- Still the only non-seismic method that is continuously applied for short-term forecasting (some weeks).
- Based on transient variations in the electric potential measured dipoles of buried electrodes.

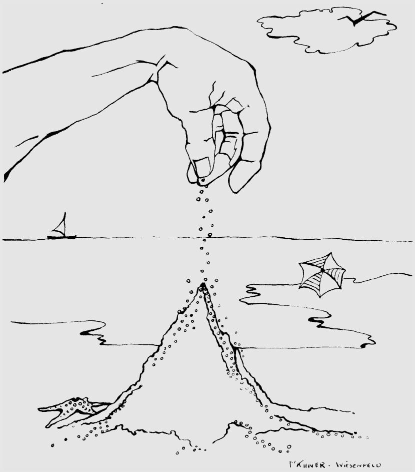
## The VAN Method

- Apparently reasonable performance, but only a few systematic tests.



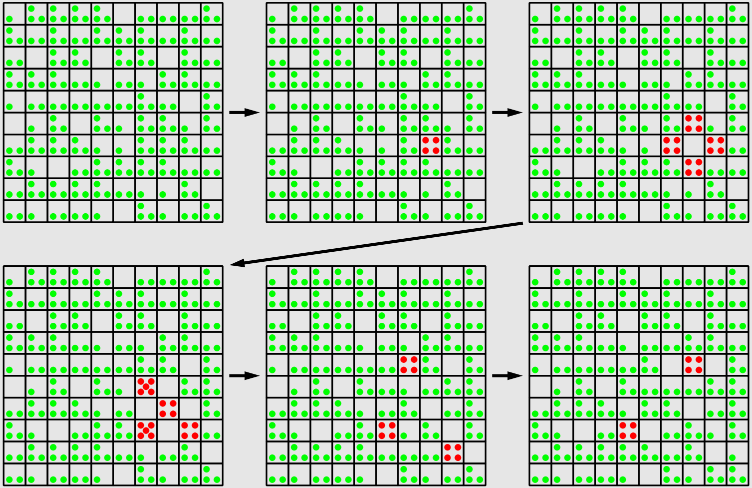

Source: Uyeda et al., Encyclopedia of Solid Earth Geophysics, 2011

## The Bak-Tang-Wiesenfeld (BTW) Model

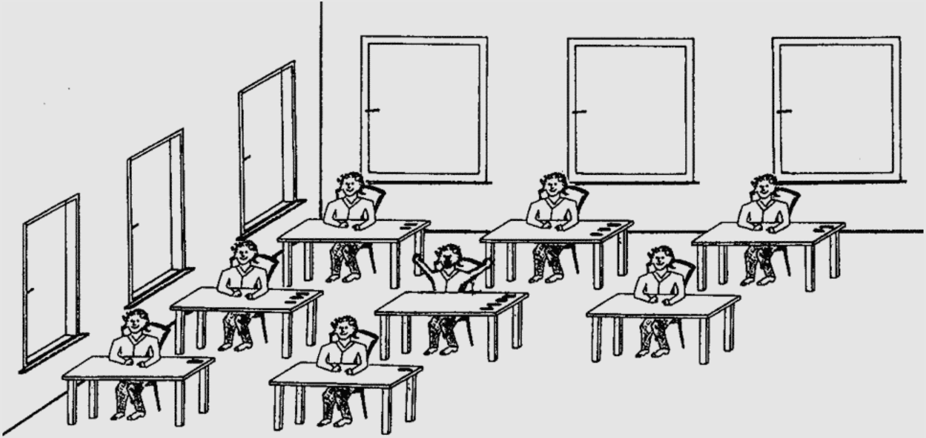



Source: Bak, How Nature Works

## The Bak-Tang-Wiesenfeld (BTW) model



## The Bak-Tang-Wiesenfeld (BTW) Model



Source: Bak, How Nature Works

## The Olami-Feder-Christensen (OFC) Model