The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2060-10

Advanced School on Non-linear Dynamics and Earthquake Prediction

28 September - 10 October, 2009

Phenomenology of Earthquake Occurrence
Part II: Quantification of Observations
Part III: Statistical Inferences

Antoni M. Correig

Universitat de Barcelona
Spain

ton.correig@am.ub.es

Strada Costiera 11, 34151 Trieste, Italy - Tel.+39 040 2240 111; Fax +39 040 224 163 - sci_info@ictp.it

PHENOMENOLOGY OF EARTHQUAKE OCCURRENCE

Antoni M. Correig

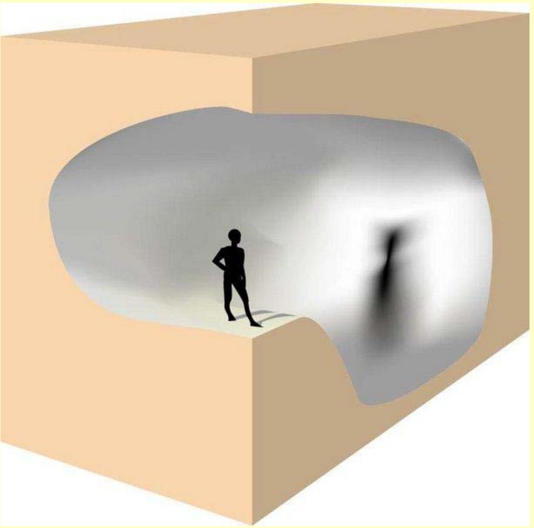
Universitat de Barcelona

Coping with the lack of fundamental equations

“It became clear for me that it is unrealistic to have a hope for the creation of a pure theory [of the turbulent flows of fluids and gases] closed in itself.
Due to the absence of such a theory we have to rely upon the hypotheses obtained by processing of the experimental data.“

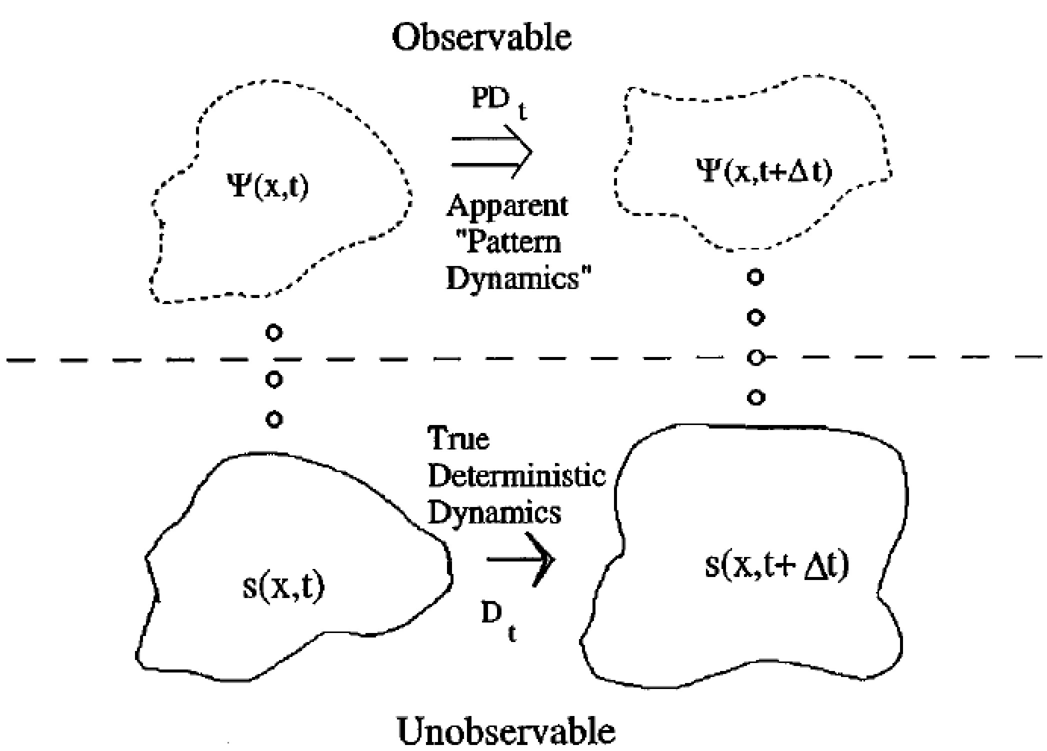
A. Kolmogorov.

SEISMOGENESIS



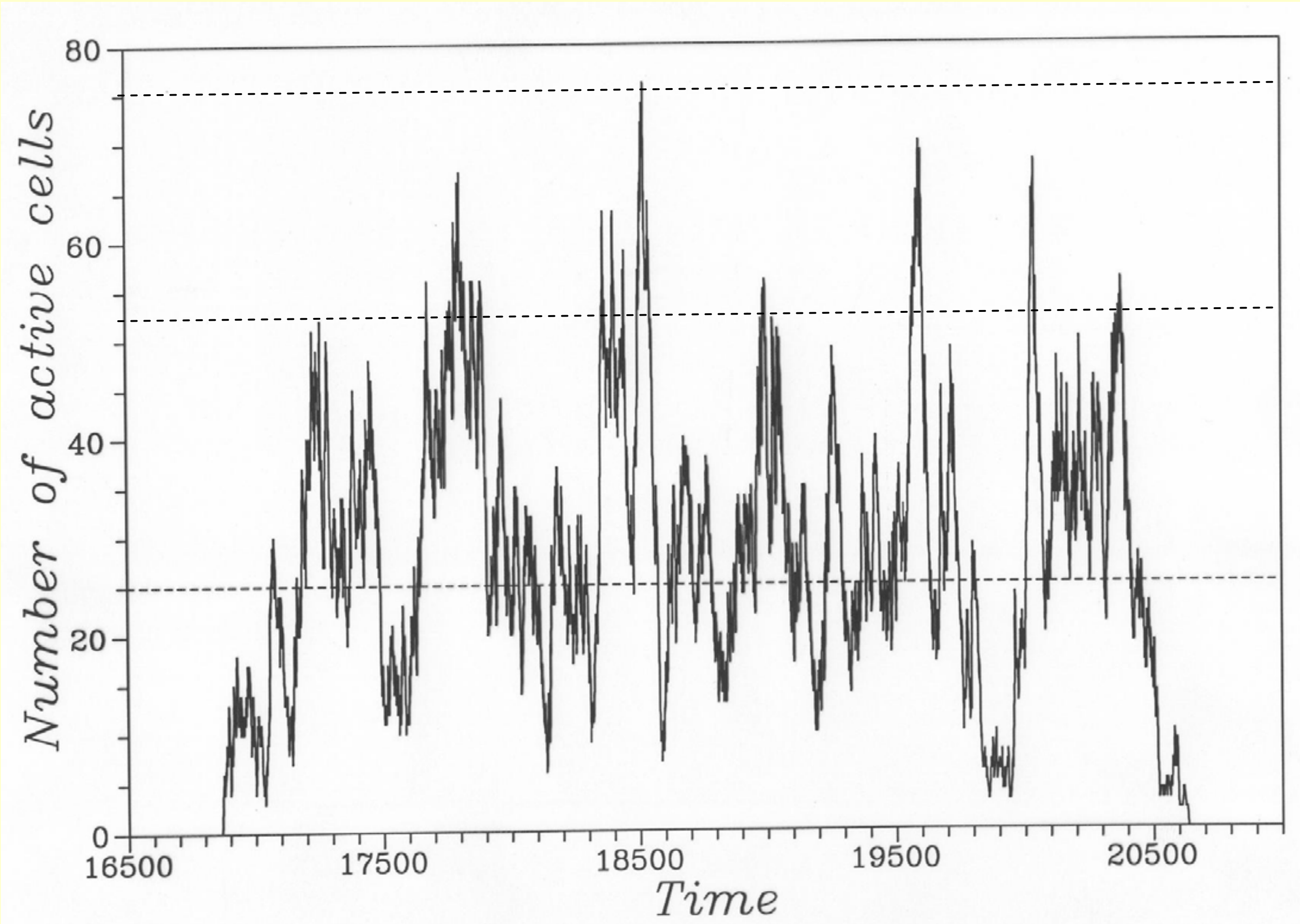
To get some insight on earthquake occurrence, is similar to deduce the reality of the external world from the imperfect shadows projected on the wall of the cave of Plato’s myth.

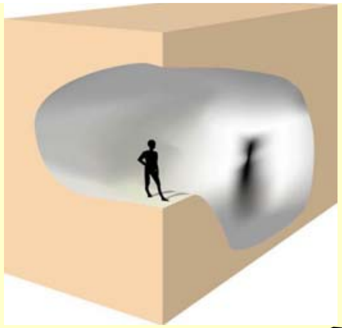
$$D_t\{s(\mathbf{x},t)\} \rightarrow s(\mathbf{x},t+\Delta t).$$



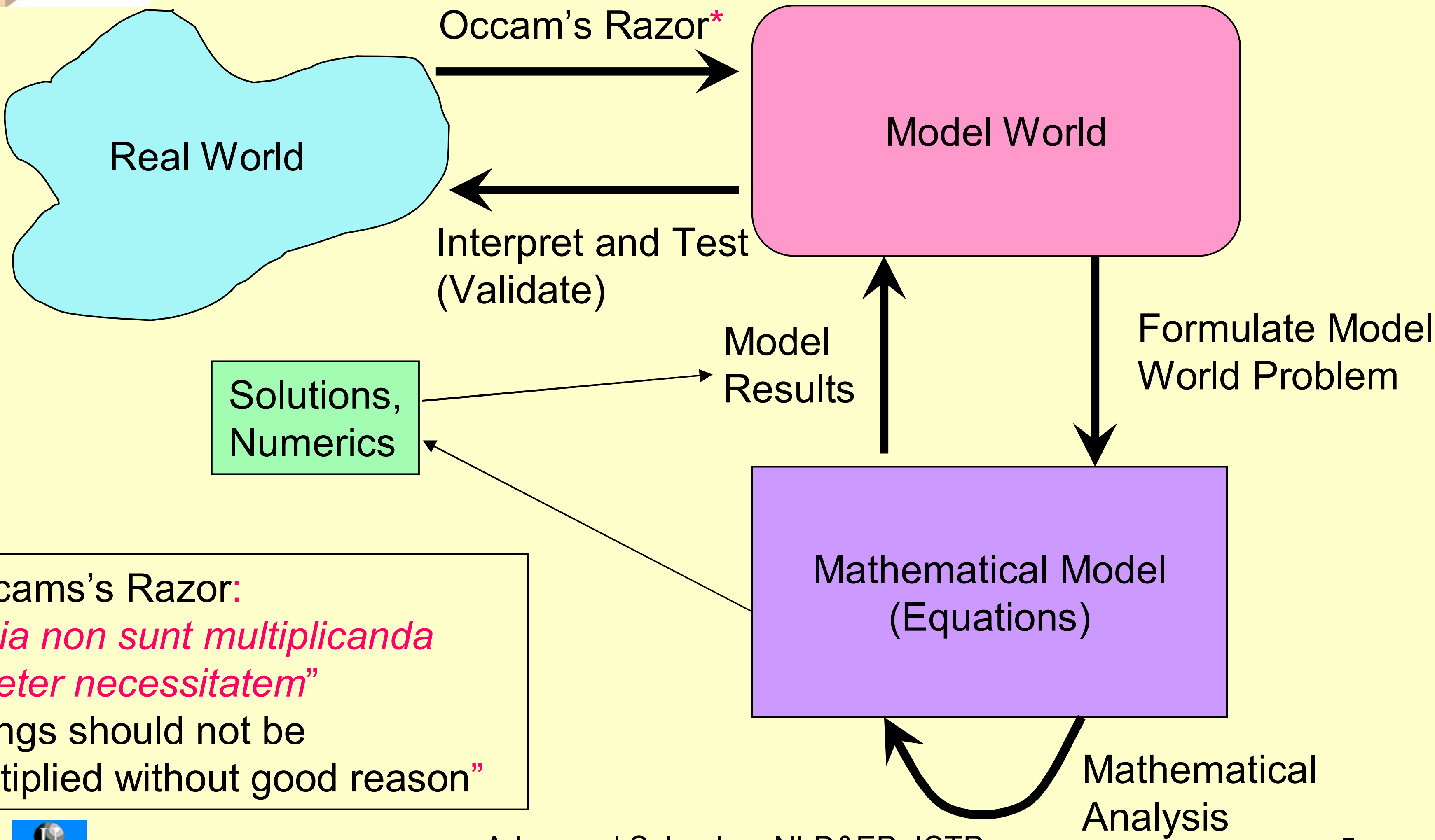
(Rundle *et al*., 2000)

SEISMOMETER: A DISCRIMINATOR



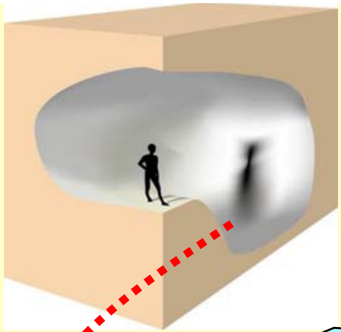


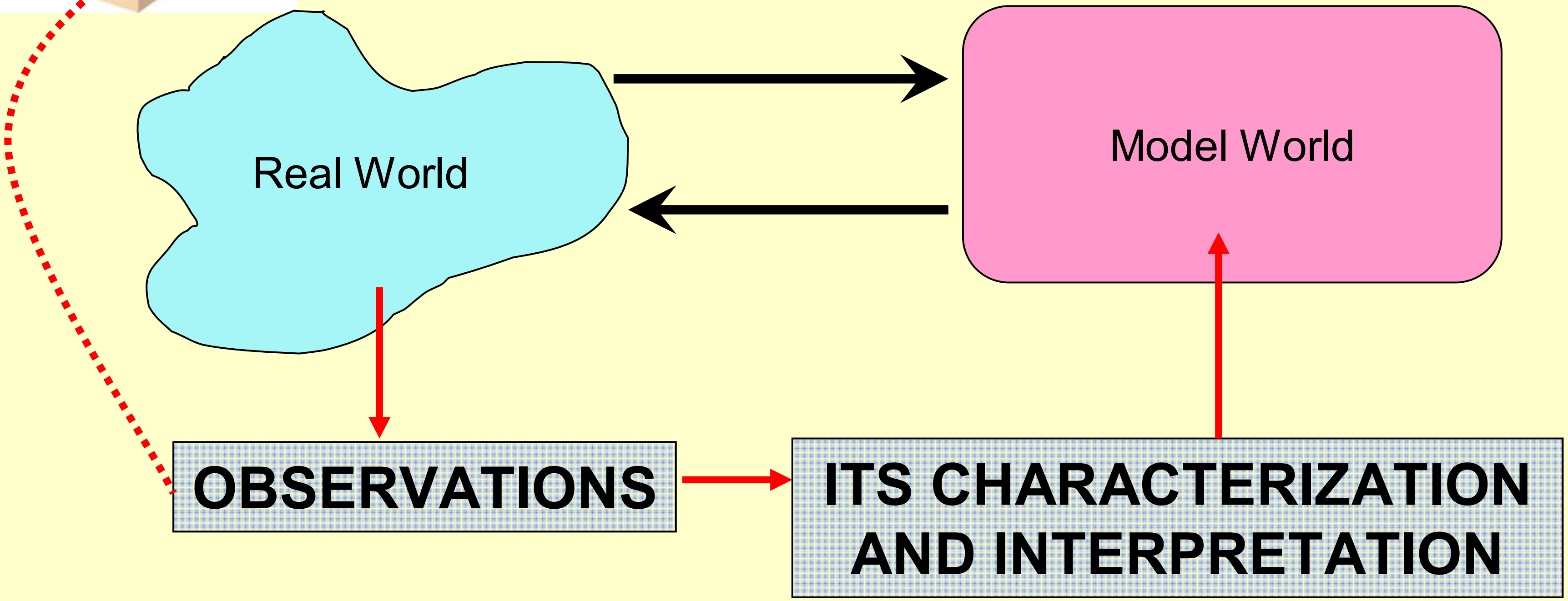
The Modeling Process



*Occams's Razor:
"*Entia non sunt multiplicanda praeter necessitatem*"
"Things should not be multiplied without good reason"

The Modeling Process



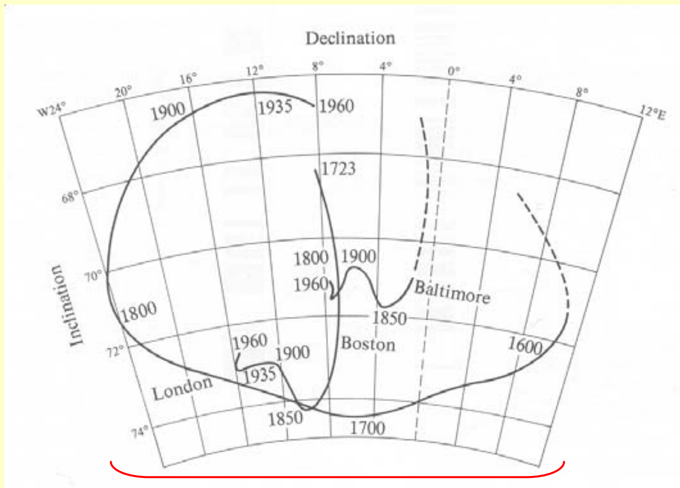


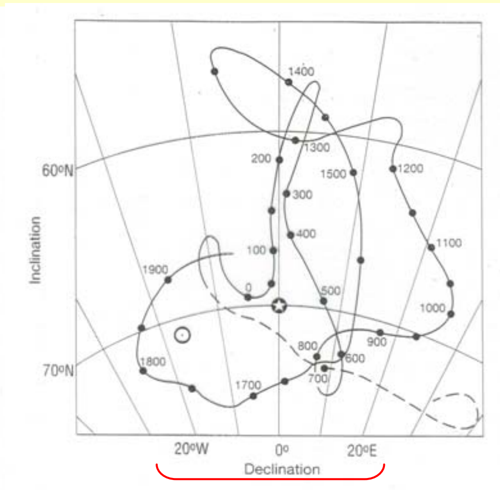
STATISTICS: the art and science of making sense of discrepant data

THE RULES OF THE GAME

- ❖ Typical problem of statistics: to detect a real systematic “effect”, and measure its size, when the data are affected by errors and/or by natural variability of the phenomena in question.
- ❖ Techniques for dealing with these situations must depend on methods of *describing* patterns of variability as well as methods of interpreting discrepant data.
- ❖ Roughly speaking, the descriptive patterns are supplied by *probability* theory, and the interpretative, inferential, conclusion-drawing parts make up *statistical* theory.

EXAMPLE.
GEOMAGNETISM:
SECULAR VARIATION





PART 1: CHARACTERIZATION OF OBSERVATIONS

OBSERVATIONS: THE SEISMIC CATALOG

seismic source + earthquake coordinates

Seismic source: - strength M_0 or M_w

- spatial orientation of slip (usually not accounted for)

Earthquake coordinates: - hypocentral location

- origin time

UNIVERSAL LAWS

QUANTIFICATION OF OBSERVATIONS

Basic and robust facts of earthquake phenomenology:

- **Gutenberg-Richter law (strength)**

$$\log \dot{N}(m) = -bm + \log \dot{a}$$

(0.8 ≤ b ≤ 1.2)

- **Omori's law (temporal occurrence)**

$$\dot{n}(t) = \frac{k}{(t+c)^p}$$

(p ~ 1)

- **The spatial distribution of epicenters in fault's systems is fractal.**

GUTENBERG-RICHTER LAW

FREQUENCY OF EARTHQUAKES IN CALIFORNIA*

By B. GUTENBERG and C. F. RICHTER

[Bull. Seism. Soc. Am., *34*: 185-188 (1944)]

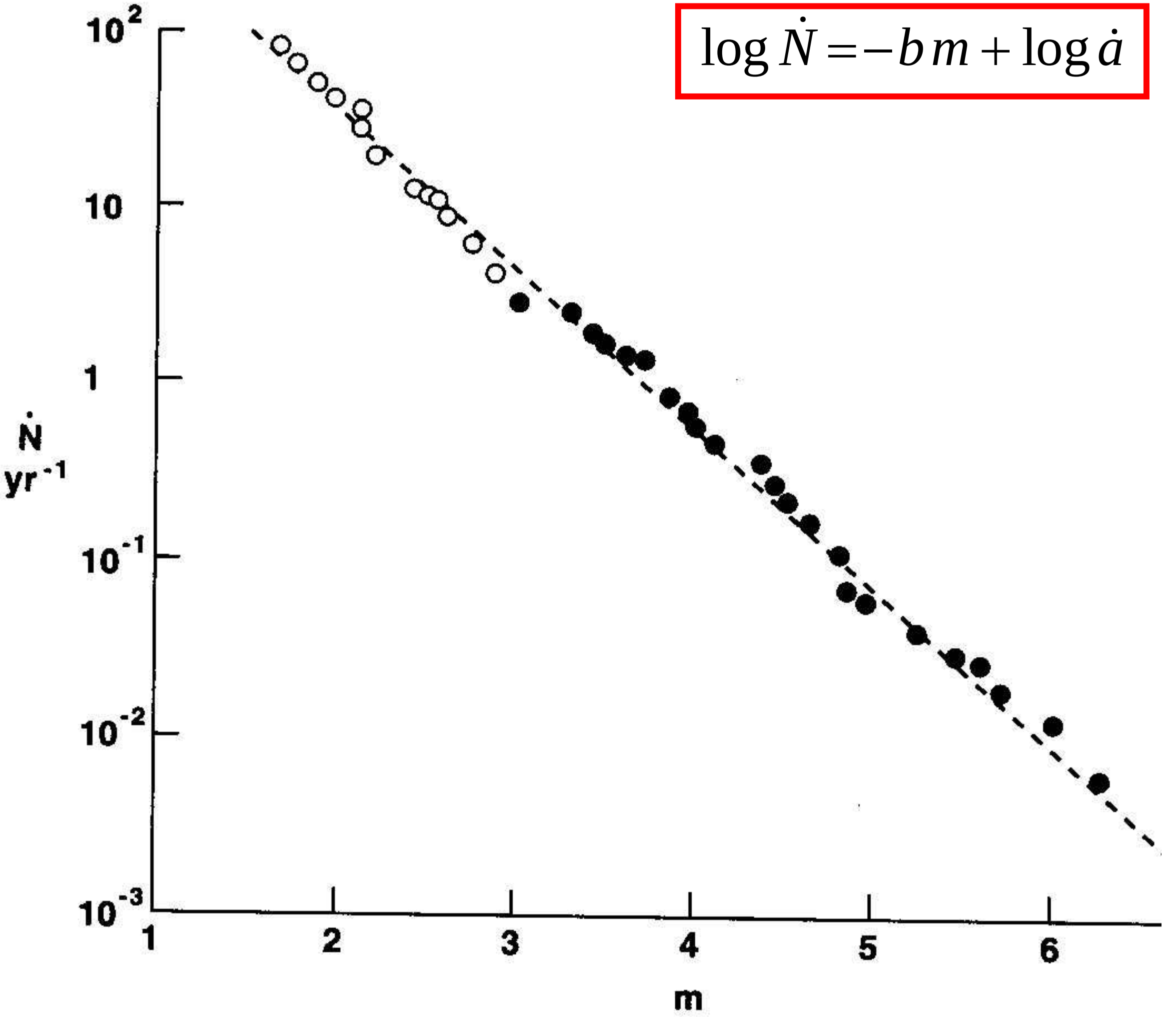
Listing is complete for magnitude 4 and higher; for magnitude 3.5, various estimates indicate that the count is roughly 20 per cent too small. Supposing

$$\log N = a + b\,(8 - M) \qquad (1)$$

a least-squares solution gives

$$a = -2.04 \pm 0.09, \qquad b = 0.88 \pm 0.03 \qquad (2)$$

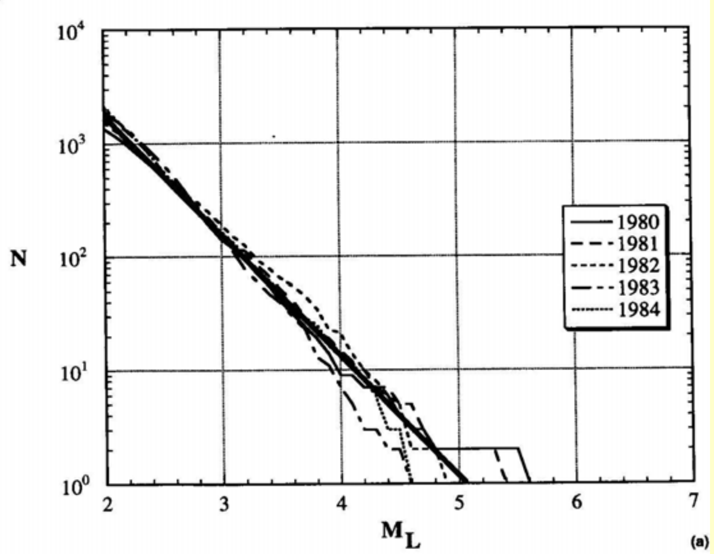
SEISMICITY. 1

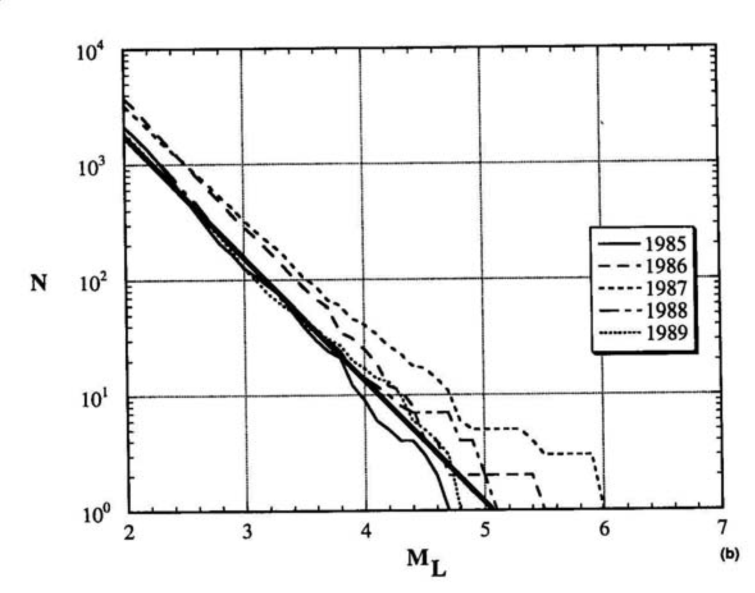


$$\log \dot{N} = -b\,m + \log \dot{a}$$

Figure 4.4. The cumulative number of earthquakes per year $\dot{N}$ occurring in the Memphis–St. Louis (New Madrid, Missouri) seismic zone with magnitudes greater than *m* as a function of *m* (Johnston and Nava, 1985). The data are for the period 1816–1983. The open circles represent instrumental data and the solid circles historical data. The dashed line represents (4.1) with $b = 0.90$ ($D = 1.80$) and $\dot{a} = 2.24 \times 10^3$ yr^{-1}.

SEISMICITY. 2





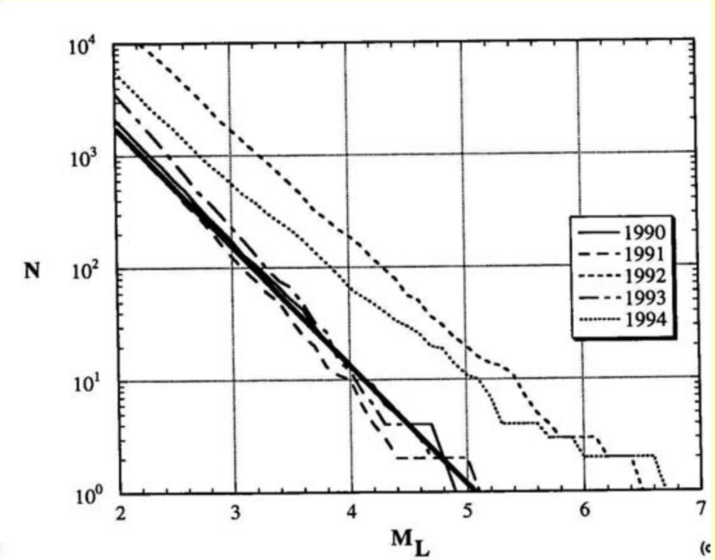
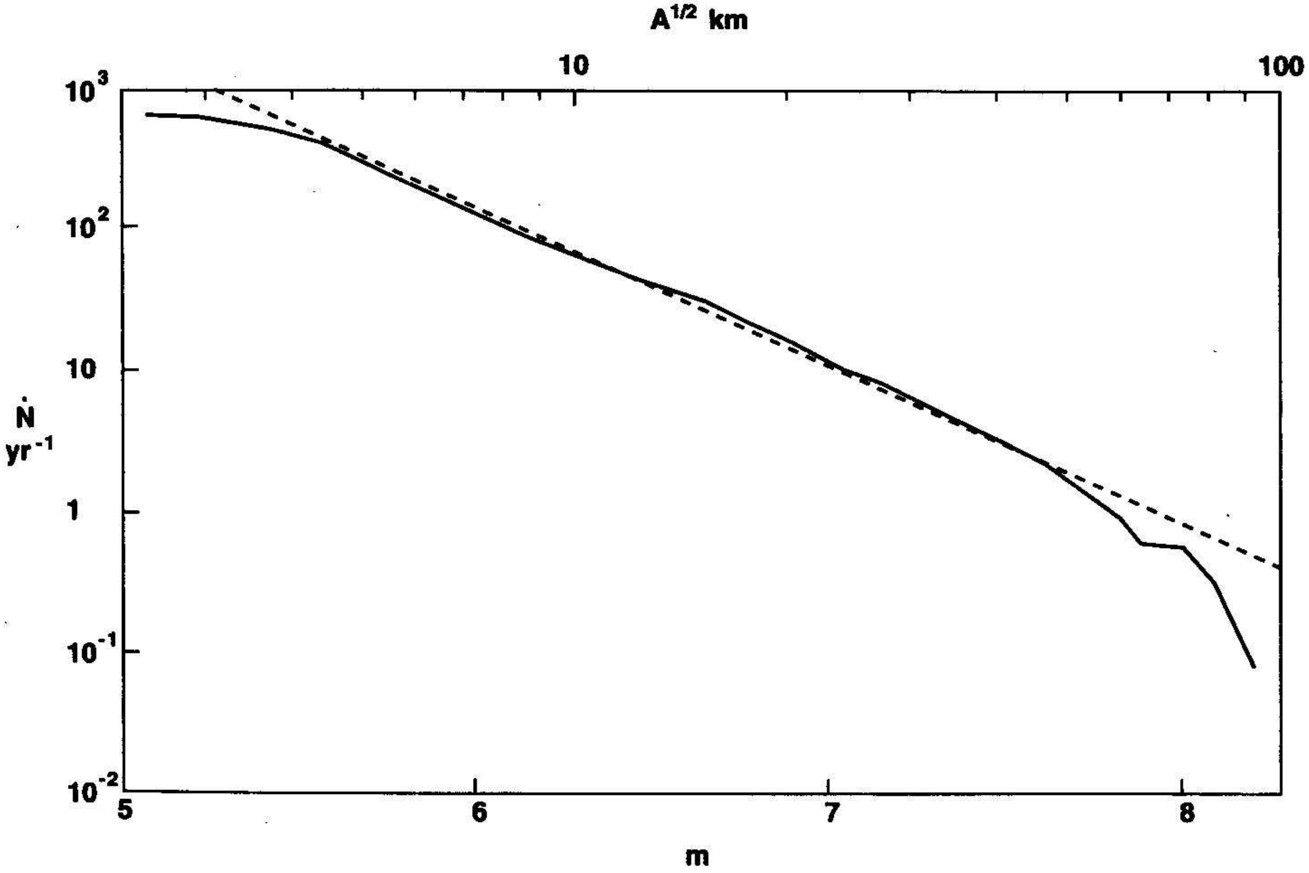


Figure 4.3. The cumulative number of earthquakes N with magnitudes greater than m for each year between 1980 and 1994 is given as a function of m; the region considered is southern California. (a) 1980–1984; (b) 1985–1989; (c) 1990–1994. The straight-line correlation is with the Gutenberg–Richter relation (4.1) with $b = 1.05$ and $\dot{a} = 2.06 \times 10^5\ \mathrm{yr}^{-1}$. The relatively large numbers of earthquakes in 1987, 1992, and 1994 can be attributed to the aftershocks of the Whittier, Landers, and Northridge earthquakes, respectively. If aftershocks are excluded, the background seismicity in southern California is nearly uniform in time.

[Turcotte (1997)]

SEISMICITY. 3



$$\log \dot{N} = -b\,m + \log \dot{a}$$

Figure 4.1. Worldwide number of earthquakes per year, $\dot{N}$, with magnitudes greater than m as a function of m. The square root of the rupture area A is also given. The solid line is the cumulative distribution of moment magnitudes from the Harvard Centroid Moment Tensor Catalog for the period January 1977 to June 1989 (Frohlich and Davis, 1993). The dashed line represents (4.1) with $b = 1.11$ ($D = 2.22$) and $\dot{a} = 6 \times 10^8\ \mathrm{yr}^{-1}$.

[Turcotte (1997)]

EXTENSION OF GUTENBERG-RICHTER LAW

As already said, originally, the G-R law was written as

$$\log N(M) = a - b(M - M_r), \; M_{lb} \leq M \leq M^{ub}$$

where N is the annual number of earthquakes of magnitude M or greater, M_{lb} and M^{ub} are the lower and higher magnitude cutoffs. The coefficient a characterizes the expected level of seismic activity in the area, and b reflects changes in the number earthquakes in successive magnitude ranges. M_r refers to a reference magnitude, and has the meaning of a characteristic size; very often M_r does not appear in the formulation.

EXTENSION OF GUTENBERG-RICHTER LAW

Kossobokov and Mazhkenov (1994) generalized G-R law to make the parameter ***b*** to be comparable in the different seismic zones, of different size.

Assume that a sequence of earthquakes is self-similar in space, and let ***N(M, L)*** be the expected annual of number earthquakes in an area of linear dimension ***L***.

To eliminate the influence of clustering consider mainshocks only.

By assuming similarity, the correspondence between ***N(M, L)*** and ***N(M)*** can be written as

$$N(M,L) = N(M)(L/l)^{C}$$

where ***N(M) = N(M, l)***, ***l*** is the characteristic length of the region and ***C*** reflects spatial similarity of set of epicenters. Rearranging:

$$\log N(M,L) = A - B(M-5) + C\log L$$

G-R LAW: PHYSICAL JUSTIFICATION

Statistical mechanics tells us that power law statistics (absence of characteristic sizes) appears when close to a critical point that defines a continuous phase transition.

Hence, formally we can assimilate the occurrence of an earthquake to a continuous phase transition order – disorder.

HOWEVER … FAILURE OF G – R LAW?

- Two different branches in the frequency size scaling relation have been reported in the literature:

- Knopoff (PNAS, 97 (2000)11880), M ~ 4.8
- Kanamori & Heaton (Geophys. Mon. 120 (2000) 147), $M_W > 4.5$ and $M_W < 2$.
- Ben-Zion & Zhu (Geophys. J. Int., 148(2002), F1-F5). $M_L \sim 3.5$.

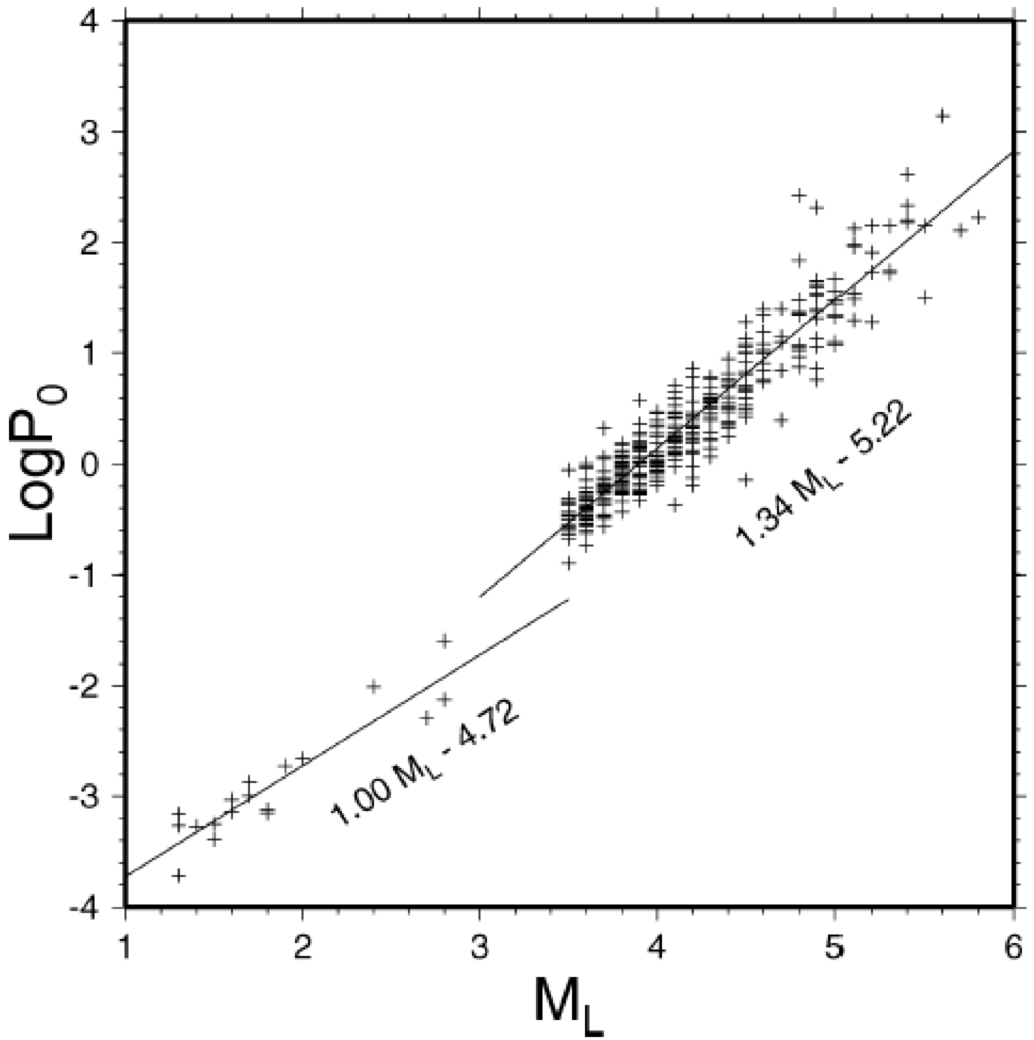
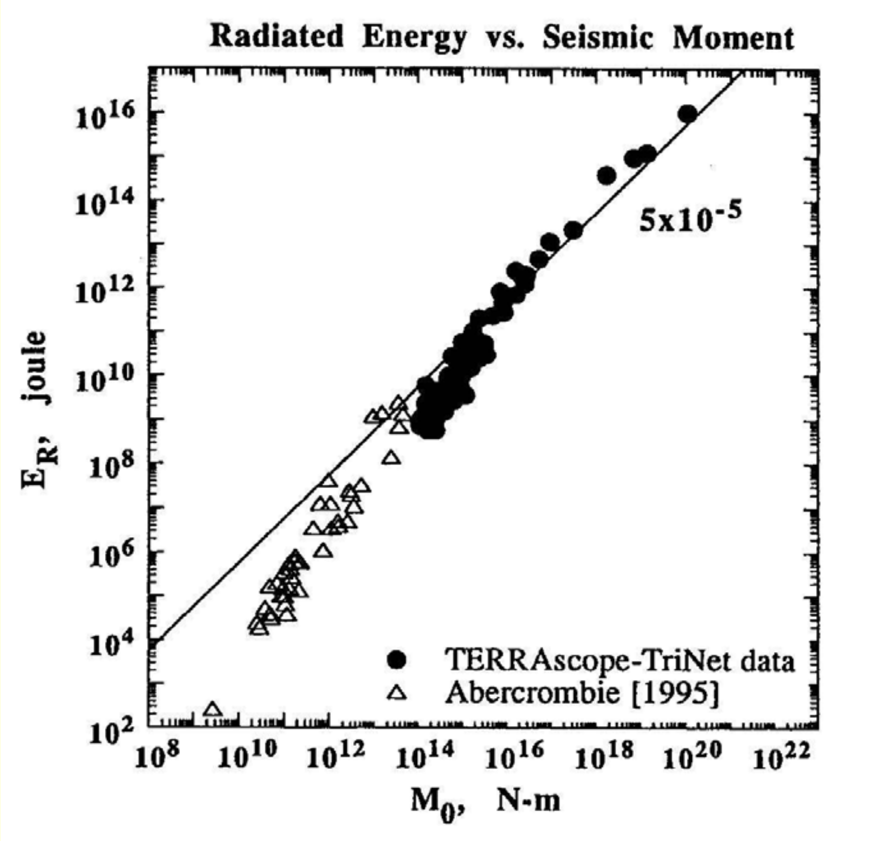


Figure 2. Linear least-squares fits done separately for the $18 M_L < 3.5$ and $418 M_L > 3.5$ events. The best-fitting parameters are summarized in Table 1. Potency values are in cm km^{-2}.

$$\log P_0 = \gamma M^2 + cM + d$$

(Ben-Zion& Zhu)



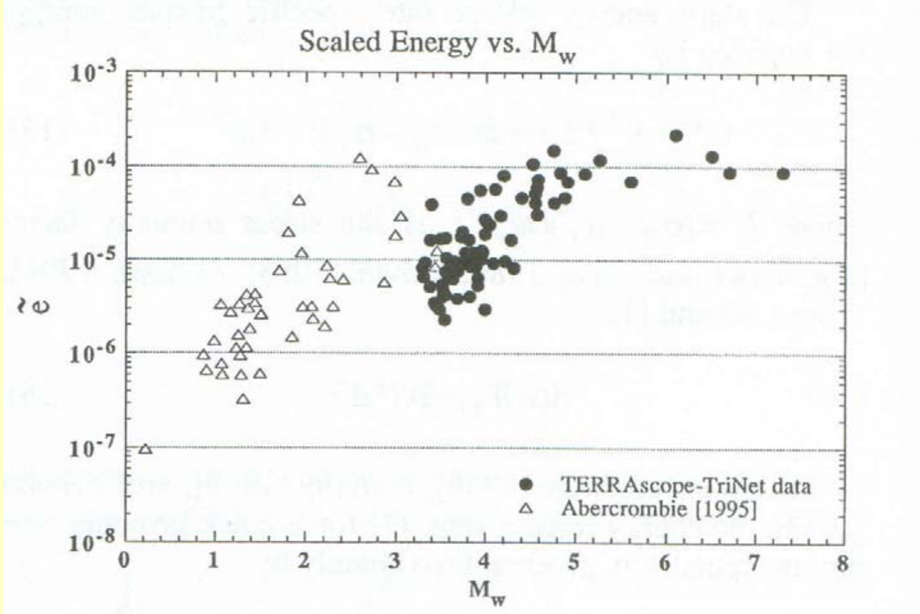


Figure 5. a). Relation between the radiated energy E_R and the seismic moment M_0. The data for large earthquakes (solid circle) are from southern California [updated from *Kanamori et al.*, 1993], and those for small earthquakes (open triangles) are taken from *Abercrombie* [1995]. b) The scaled energy, $\tilde{e}=E_R/M_0$, computed as a function of M_W. Note that the values of $\tilde{e}$ for small earthquakes are 10 to 100 times smaller than those for large earthquakes.

(Kanamori & Heaton)

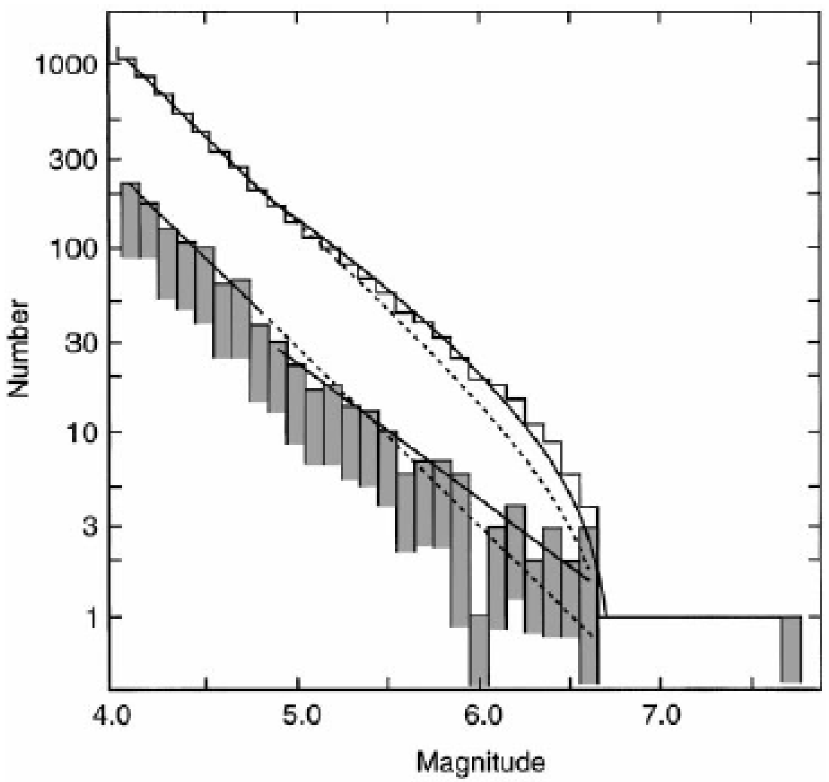


Fig. 3. Maximum likelihood fits to the cumulative and differential (shaded) distributions of the magnitudes of the complete catalog of earthquakes in the Southern California region from 7-1-44 to 3-1-90. The dashed line is the fit $4.1 \le M \le 6.6$ under the assumption that there is a single branch to the distribution. The solid curve is the fit for the two independent branches $4.1 \le M \le 4.8$ and $4.9 \le M \le 6.6$. The curvature in the cumulative fit at large magnitudes is due to the assumption of a finite upper magnitude cutoff.

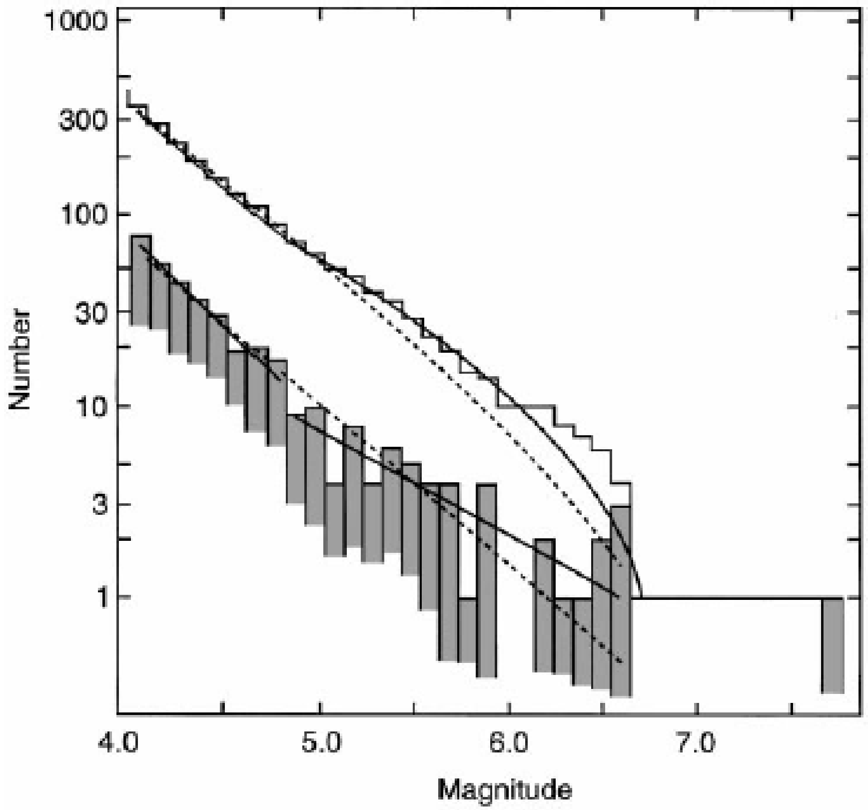


Fig. 4. Cumulative and differential (shaded) magnitude distribution of main shocks in the Southern California region. The distributions are fit by a single segment that spans the range $4.1 \le M \le 6.6$ (dashed), and by two segments that span the ranges $4.1 \le M \le 4.8$ and $4.9 \le M \le 6.6$ (solid). There is a kink in the fit to the cumulative distribution at $M = 4.8$ and a rolloff at large magnitudes because of the assumption of a finite upper magnitude cutoff.

(Knopoff, 2000)

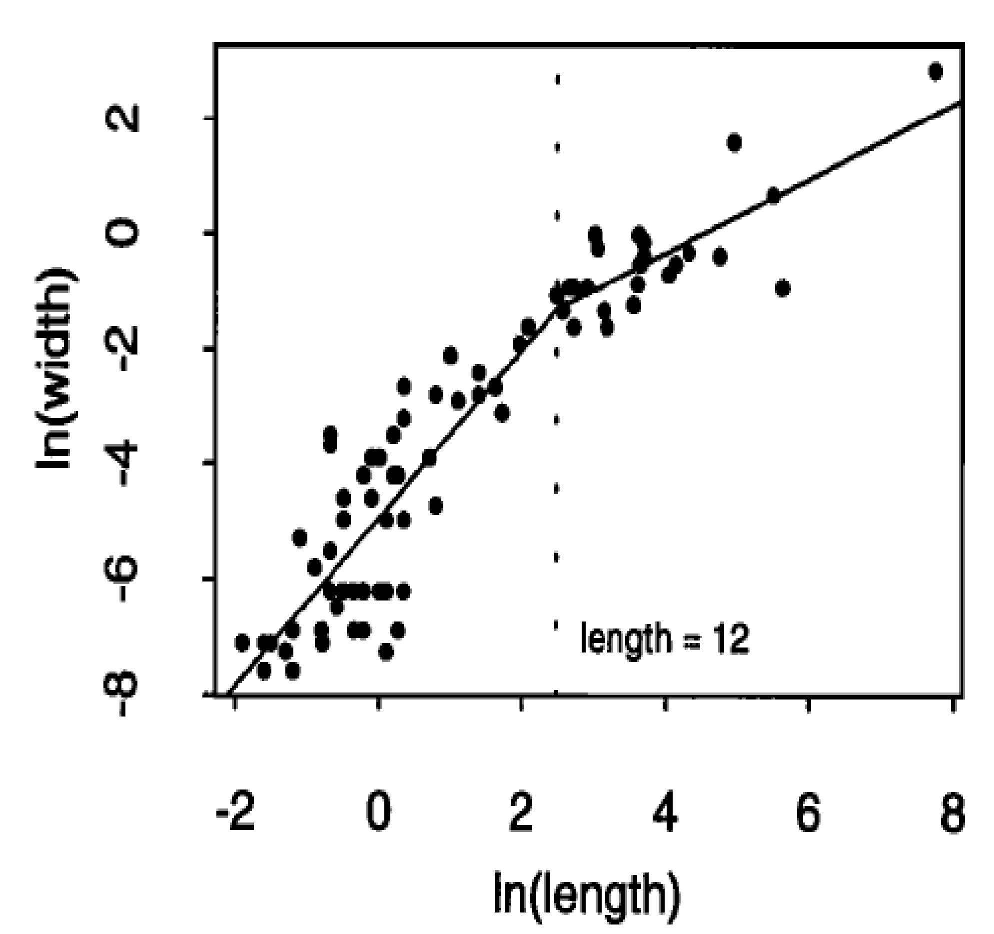
The binned distribution densities of magnitudes in both the complete and the declustered catalogs of earthquakes in the Southern California region have two significantly different branches with crossover magnitude near M = 4.8. In the case of declustered earthquakes, the b-values on the two branches differ significantly from each other by a factor of about two. The absence of self-similarity across a broad range of magnitudes in the distribution of declustered earthquakes is an argument against the application of an assumption of **scale-independence** to models of main-shock earthquake occurrence. The presumption of scale-independence for complete local earthquake catalogs is attributable, not to a universal process of self-organization leading to future large earthquakes, but to the **universality of the process that produces aftershocks**, which dominate complete catalogs.

However, the G-R law, with b a constant across a wide range of magnitudes, is indeed a valid property of the distribution density.

(Knopoff, 2000)

ONE SLOPE OR TWO? (Main et al., 1999)

In analyzing data from the Krafla fissure in the north of Iceland, a significant break of slope was find:



Fracture width (opening displacement) u versus length l. The data are shown on log-log scales using natural logarithms. The best fitting lines using a Bayesian method are shown. The location of the best fitting break of slope is indicated by the vertical dashed line.

HOWEVER ... (Main, 2000)

We examine the problem in a forward modeling mode by adding a realistic degree of statistical scatter to ideal incremental frequency-moment distributions of various commonly used forms.

Adopting a priori the assumption of a piecewise linear distribution, we find in each case apparently statistically distinct breaks of slope that are not present in the parent distribution.

These breaks of slope are artifacts produced by a combination of

(a) high-frequency noise introduced by the random statistical scatter,

(b) the more gradual natural roll-over in the cumulative frequency data near the maximum seismic moment, and

(c) a systematic increase in the apparent regression coefficient due to the natural smoothing effect of the use of cumulative frequency data.

Therefore, if there is no apparent break of slope in the incremental distribution, it is unwise to interpret the cumulative-frequency data uniquely in terms of a break in slope.

DEPARTURES OF THE G-R LAW ?

Rundle (1993) suggested that departures from simple Gutenberg-Richter scaling are apparently due, at very large magnitudes, to the finite width of the brittle layer, and at very small magnitudes to the finite thickness of the inelastic fault zone.

Gutenberg-Richter Law: the cumulative frequency dn_0/dt of events with magnitude larger than m is given by

$$dn_0 / dt = 10^{A_0} 10^{-bm}$$

A_0 characterizes the level of seismicity of the fault system, and the value of b determines the frequency of occurrence of large events relative to small ones. It is widely accepted that $b \approx 1$.

The "brittle" seismogenic part of Earth is for the most part the uppermost 10-50 km of Earth's lithosphere.

DEPARTURES OF THE G-R LAW ?

Earthquakes of intermediate size with magnitudes up to m ~ 6 have source dimensions smaller than ~ 10 km and are thus relatively uninfluenced by the finite width of the brittle lithosphere.

Large events with m ~ 7 and above have a length L that can greatly exceed the width of the brittle lithosphere. As a result, the way in which slip scales with event size changes.

The expected values for b for intermediate size events is $b_I \approx 1$ and for large events is $b_L \approx 1.5$.

The departure from strict Gutenberg-Richter scaling at large magnitudes is apparently related to the finite depth of the brittle layer. The slip in these events reached saturation when the depth of the earthquake roughly equaled the depth extent of the brittle seismogenic zone.

There may also be a departure from GR scaling at very small magnitudes. It has been observed a cumulative magnitude-frequency relation for small events that falls significantly below the rate given by GR with b ≈ 1 as m decreases below some cutoff magnitude m_c , where typically m_c is in the range of 1 to 3.

DEPARTURES OF THE G-R LAW ?

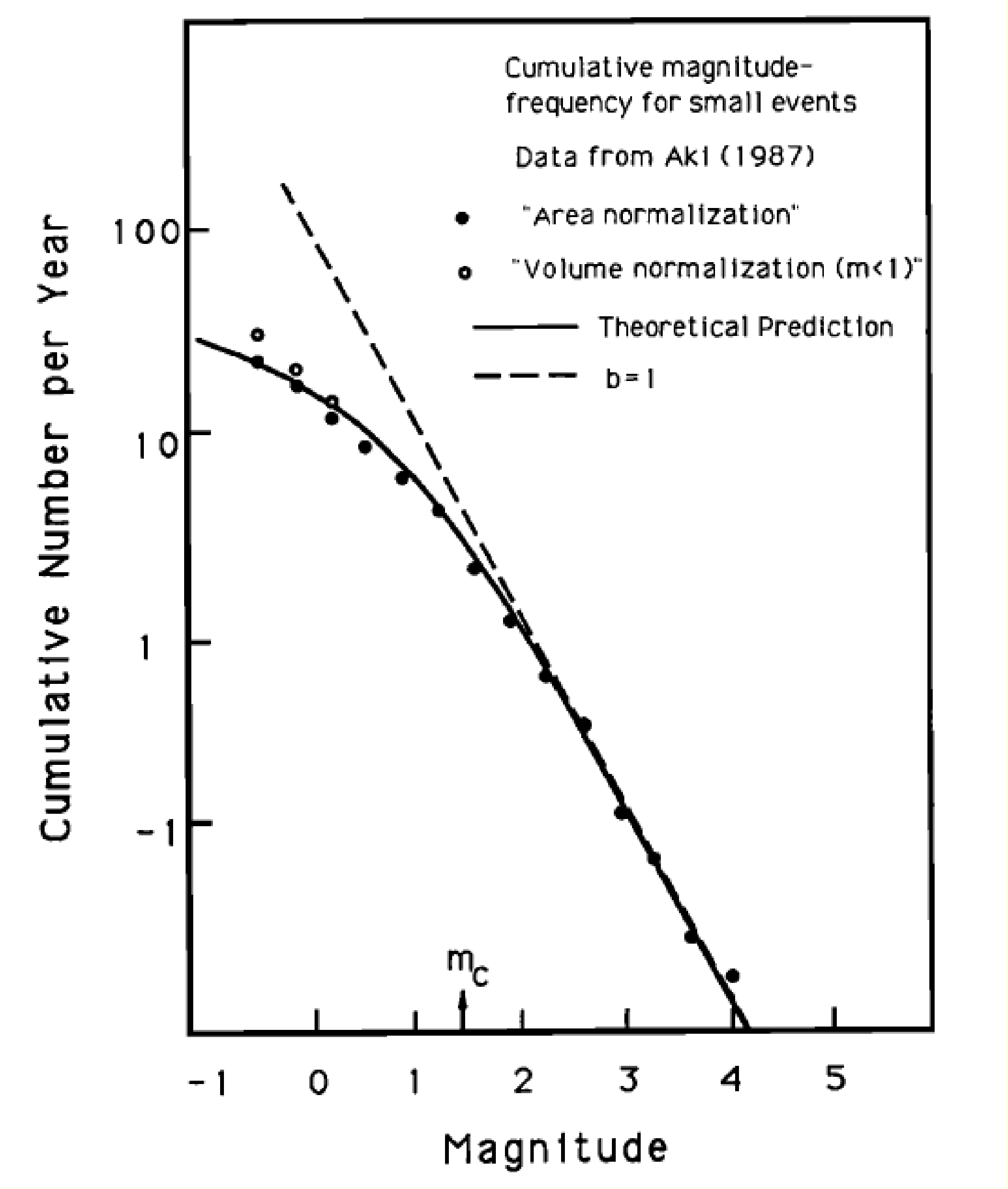
It can be argued that very small events have a fixed source size λ_c and that the moment is determined only by the slip, which is thus independent of source size. This is in contrast to intermediate size events, in which slip scales linearly with source radius, and large events, in which slip is proportional to the depth of the brittle layer and thus saturates.

Thus, in both the small and large events, slip is independent of source size, and the resulting frequency of events is less than predicted by GR.

Based on statistical mechanics grounds, Rundle recovers GR law for intermediate magnitude events and propose new relations for large and small earthquakes:

$$dn_L/dt = 10^{A_L}\, 10^{-1.5m} \qquad \text{large events}$$

$$dn_S/dt = 10^{A_I - m_c} \log\left[1 + 10^{m_c - m}\right] \qquad \text{small events}$$


Cumulative magnitude-
frequency for small events
Data from Aki (1987)
"Area normalization"
"Volume normalization (m<1)"
Theoretical Prediction
b = 1
100
10
1
-1
Cumulative Number per Year
m_c
-1
0
1
2
3
4
5
Magnitude

TRUNCATED POWER LAWS

NOISE 1/f. 1

[Dutta & Horn (1981)]

1/f noise refers to fluctuations which have spectral densities varying approximately as 1/f over a large range of frequency, f. Fluctuations with such spectra have been observed in a tremendous variety of dissimilar physical systems. Is there a universality in the underlying equations which leads to 1/f noise in many apparently unrelated systems? The shape of the power spectrum uniquely characterizes the process only if it is stationary and Gaussian (all higher-order correlations are zero).

In a generic way, 1/f noise can be contemplated as an **activated random process**. A random process with a **characteristic time** τ has a Debye-Lorentzian spectrum

$$S(\omega) \propto \frac{\tau}{\omega^2 \tau^2 + 1}$$

Any spectrum may be generated by postulating an appropriate **distribution *D(τ)* of the characteristic times** within the sample. This could arise if, for example, the sample was inhomogeneous. Then

$$S(\omega) \propto \int \frac{\tau}{\omega^2 \tau^2 + 1} D(\tau) d\tau.$$

NOISE 1/f. 2

In particular, if

$$D(\tau) \propto \tau^{-1} \qquad for \qquad \tau_1 \leq \tau \leq \tau_2,$$

then

$$S(\omega) \propto \omega^{-1} \qquad for \qquad \tau_2^{-1} << \omega << \tau_1^{-1}$$

If, as in many physical processes, τ is thermally activated,

$$\tau = \tau_0 \exp(E / kT),$$

Then the required energy distribution is

$$D(E) = \text{const. for} \qquad kT \ln(\tau_1 / \tau_0) \leq E \leq kT \ln(\tau_2 / \tau_1)$$

The problem of justifying a 1/f spectrum has now been shifted to one of motivating the **required energy distribution**. Very often, specially in natural phenomena, several intervals can be found behaving as $1/f^{\alpha}$, each one with a different value of α.

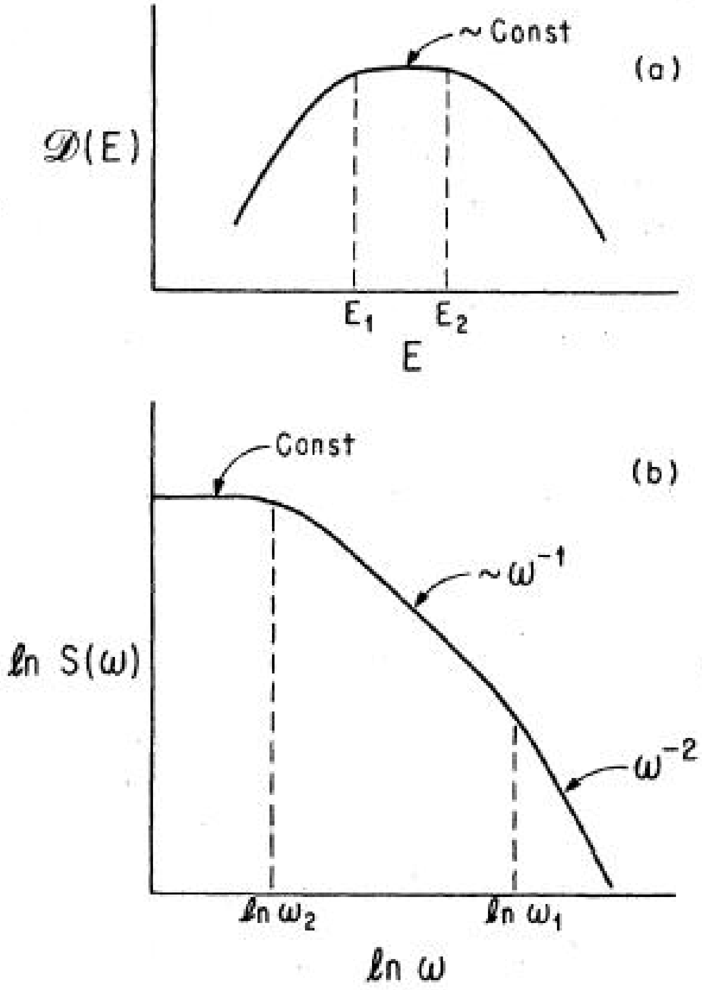


FIG. 5. (a) "flat" energy distribution. (b) frequency spectrum resulting from this distribution

VARIABILITY OF B-VALUE ON STRESS REGIMES? 1

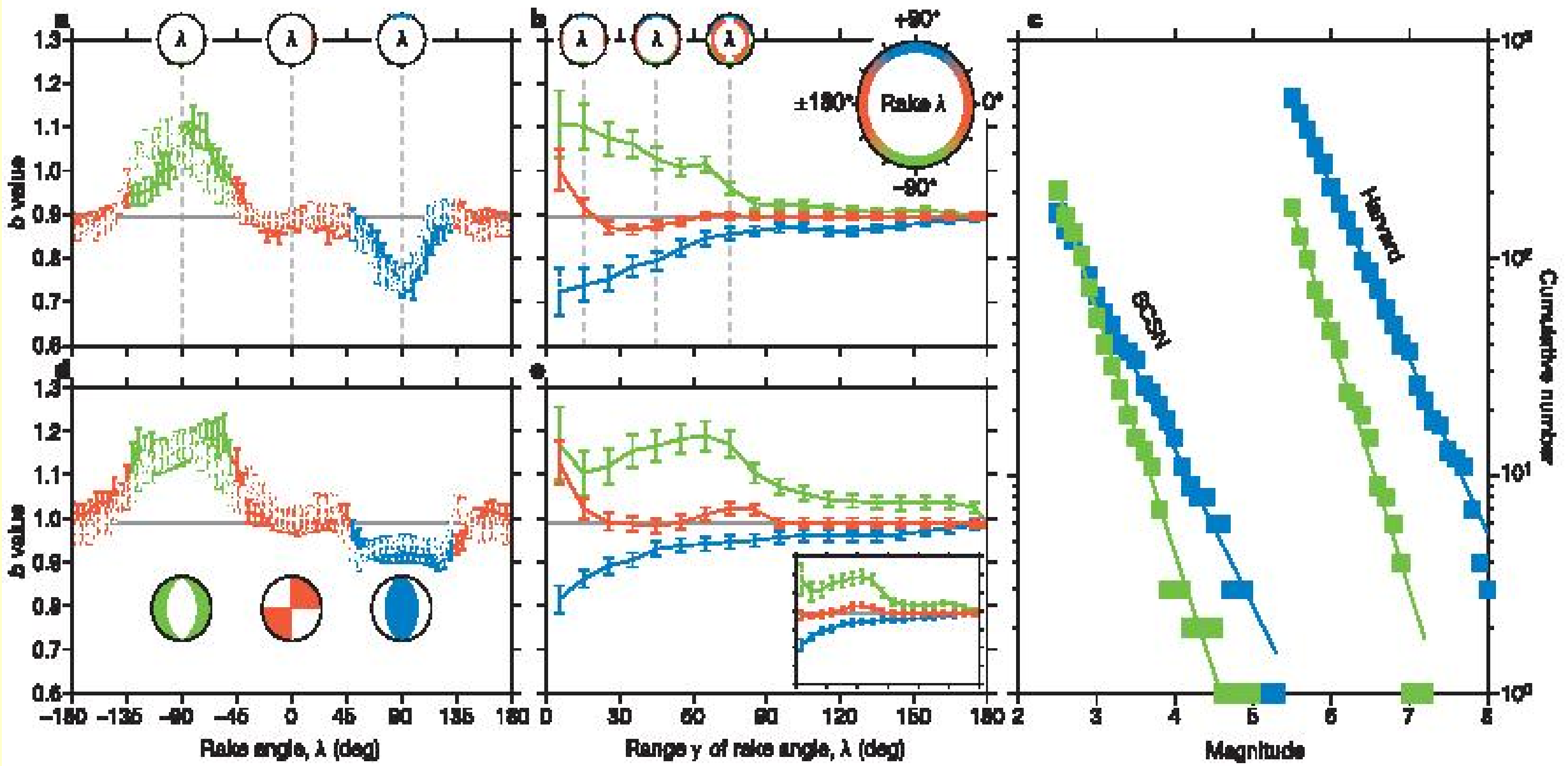


Figure 1 | Plots of *b* against rake angle and range of rake angle. **a**, b_λ plot of events in southern California, from the SCSN catalogue. In all frames the green, red and blue lines (solid, first plane; outlined, second plane) mark the *b* values of mainly normal, strike-slip and thrust events, respectively; the grey line marks the average *b* value and the vertical bars indicate the standard error[7]. (Solid bars are used for samples with $N \geq 200$, and dashed bars for samples with $200 > N \geq 100$). The circles at the top of the frame show the rake λ used for computing the *b* values $\lambda = -90° \pm \gamma$, $\lambda = 0° \pm \gamma$ and $\lambda = 90° \pm \gamma$, $\gamma = 20°$. **b**, b_γ plot of events in southern California. The circles at the top of the frame show the range of rake λ used for computing the *b* values ($\gamma = 15°$, $45°$ and $75°$). Inset, circle explaining the rake values and the corresponding colours of the classes of events. **c**, Frequency–magnitude distributions for pure normal (green) and pure thrust (blue) events of the SCSN and Harvard catalogues ($\gamma = 5°$). **d**, As **a** for the Harvard catalogue. **e**, As **b** for the Harvard catalogue. Inset, as main panel but considering only the rake of the first nodal plane.

Schorlemmer *et al.* (2005)
Nature, 437, 539.

VARIABILITY OF B-VALUE ON STRESS REGIMES? 2

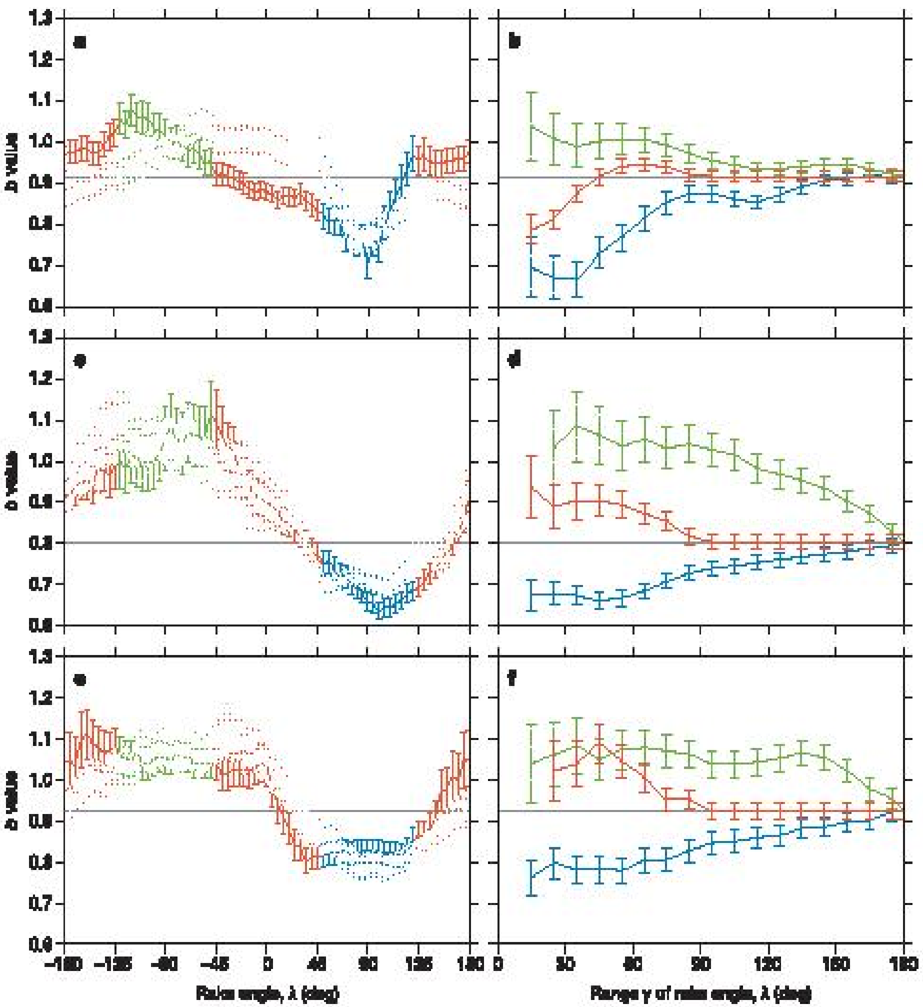


Figure 2 | b_λ plots and b_γ plots for the NCSN catalogue (a, b), the NEID Kanto-Tokai catalogue (c, d) and the NEID F-Net catalogue (e, f). In all frames the green, red and blue lines (solid, first plane; outlined, second plane) mark the *b* values of mainly normal, strike-slip and thrust events, respectively; the grey line marks the average *b* value and the vertical bars indicate the standard error[7]. (Solid bars are used for samples with $N \geq 200$, dashed bars for samples with $200 > N \geq 100$).

Schorlemmer *et al.* (2005)
Nature, 437, 539.

VARIABILITY OF B-VALUE WHEN TO CLOSE TO A CRITICAL POINT ?

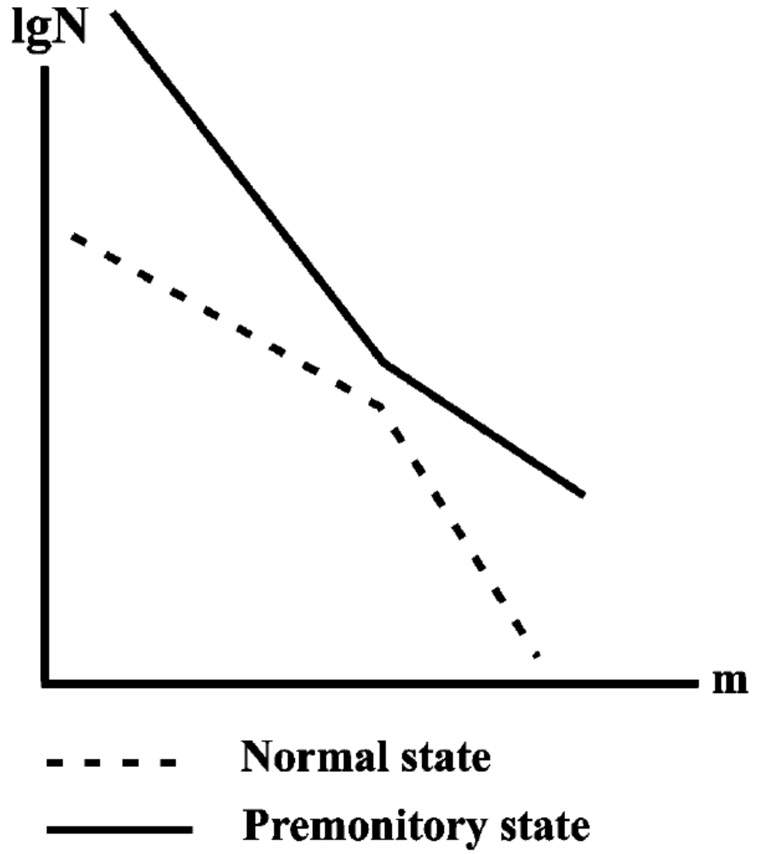


Figure 10 Premonitory changes of the Gutenberg-Richter relation. *N* is the occurrence rate of earthquakes, *m* is the magnitude (logarithmic measure of energy). The dashed line represents time far from a strong earthquake; the solid line represents time close to it.

CHARACTERISTIC EARTHQUAKES

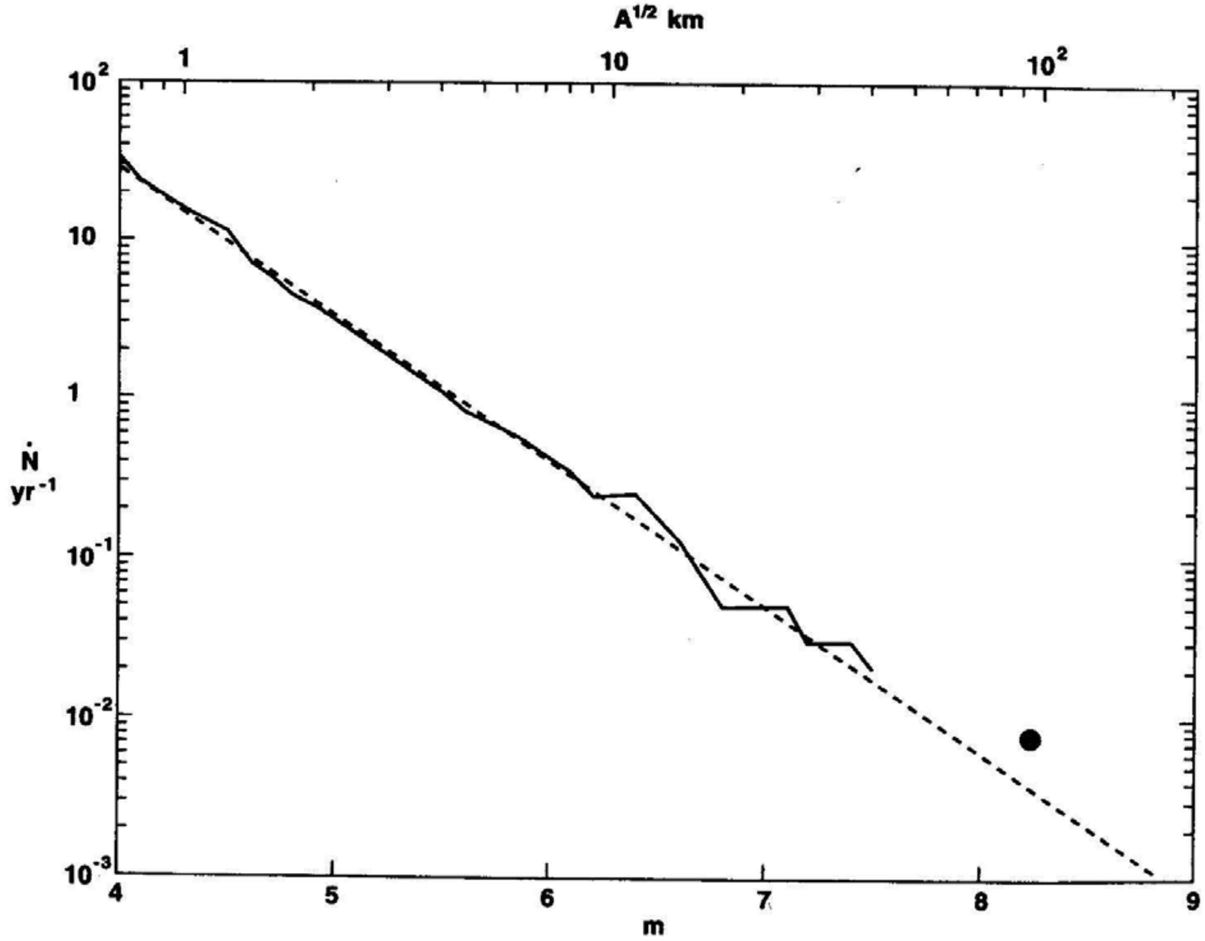
It has been found that fault systems with highly irregular geometry, which have many offsets and branches, display ‘‘power-law’’ statistics over the whole range of observed magnitudes.

On the other hand, the available data indicates that fault systems with more regular geometry (presumably generated progressively with increasing cumulative slip) display power-law distributions only for small events, which occur in the time intervals between roughly quasiperiodic earthquakes of a much larger ‘‘**characteristic**’’ size which rupture the entire fault.

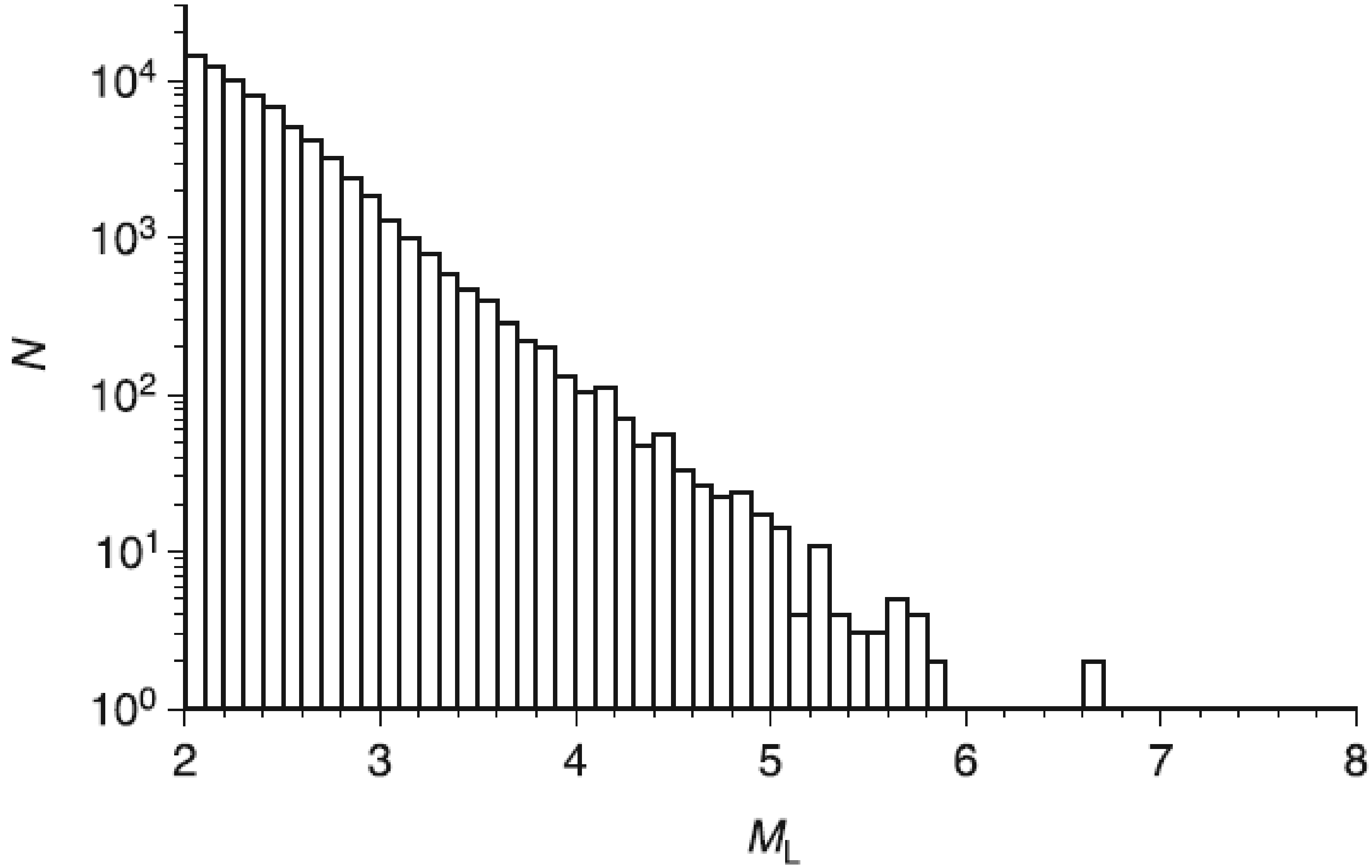
There are practically no observed earthquakes of intermediate magnitudes on such geometrically regular fault systems. Distributions of this type are called ‘‘**characteristic earthquake**’’ distributions.

SEISMICITY. 4

Figure 4.2. Number of earthquakes per year $\dot{N}$ occurring in southern California with magnitudes greater than m as a function of m. The solid line is the data from the southern California earthquake network for the period 1932–1994. The straight dashed line is the correlation with (4.1) taking $b = 0.923$ ($D = 1.846$) and $\dot{a} = 1.4 \times 10^5$. The solid circle is the observed rate of occurrence of great earthquakes in southern California (Sieh *et al.*, 1989).



[Turcotte (1997)]



10^4
10^3
10^2
10^1
10^0
N
2
3
4
5
6
7
8
M_L

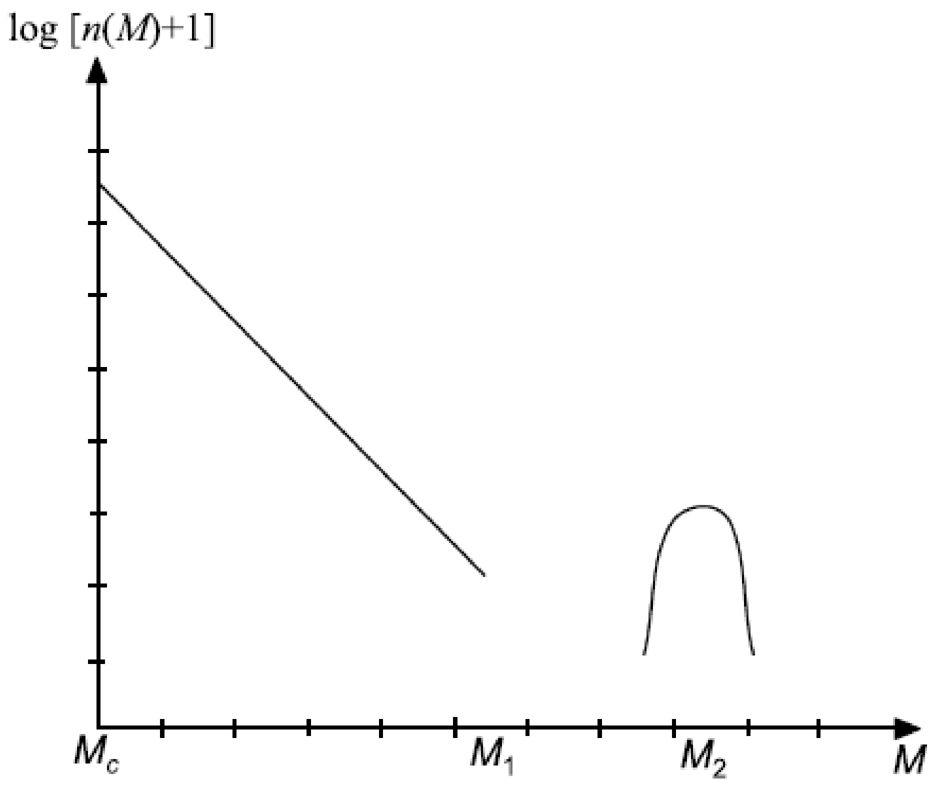
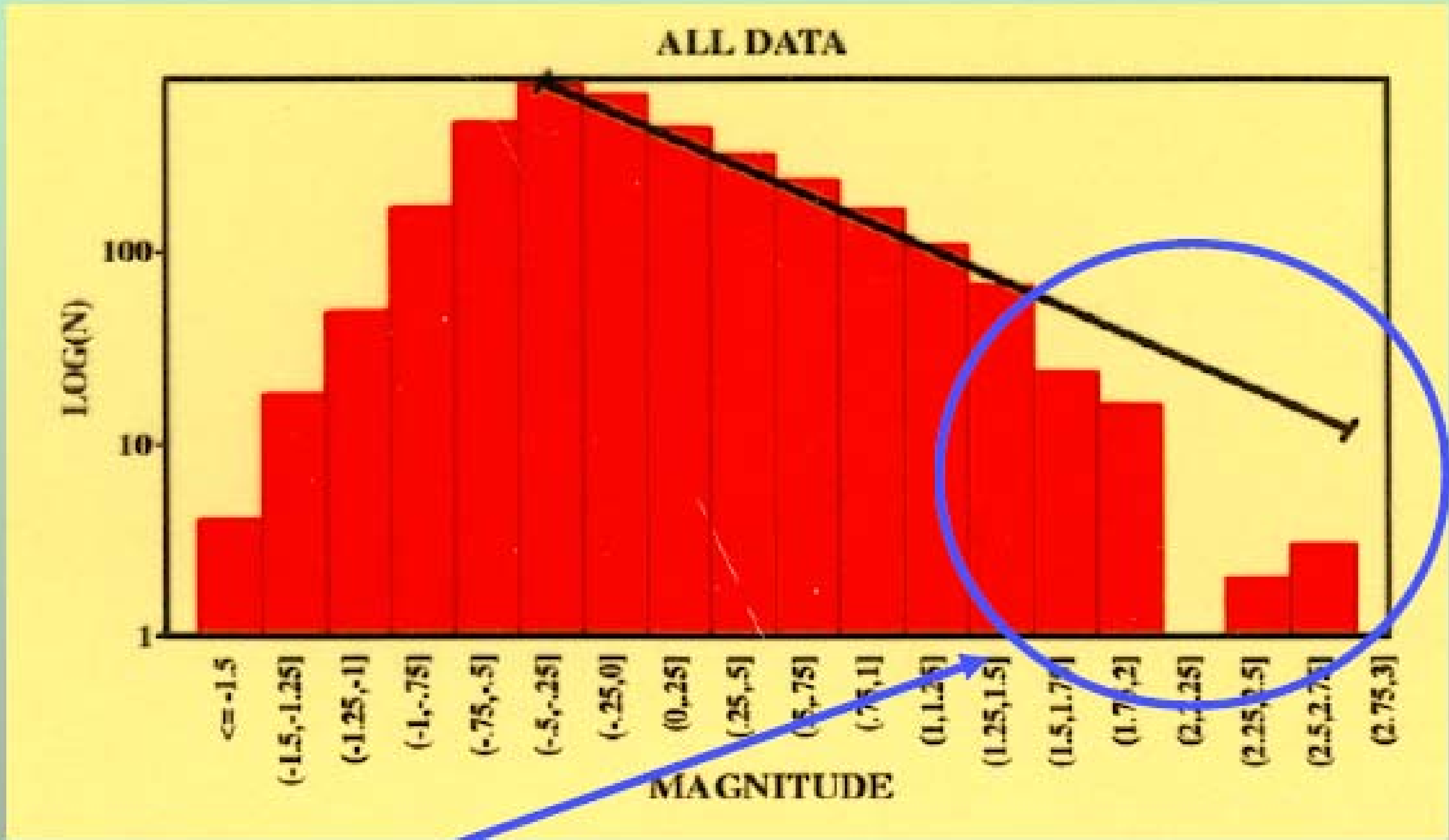


Figure 8. A schematic representation of the characteristic earthquake distribution in a discrete form for events larger than some magnitude cutoff M_c with two populations: a Gutenberg-Richter distribution of events smaller than M_1 and a peaked distribution around a large characteristic event size M_2. The region between M_1 and M_2 may be associated with a magnitude gap, while outside that region $n \gg 1$ (i.e., $n + 1 \approx n$). From *Ben-Zion* [2003].



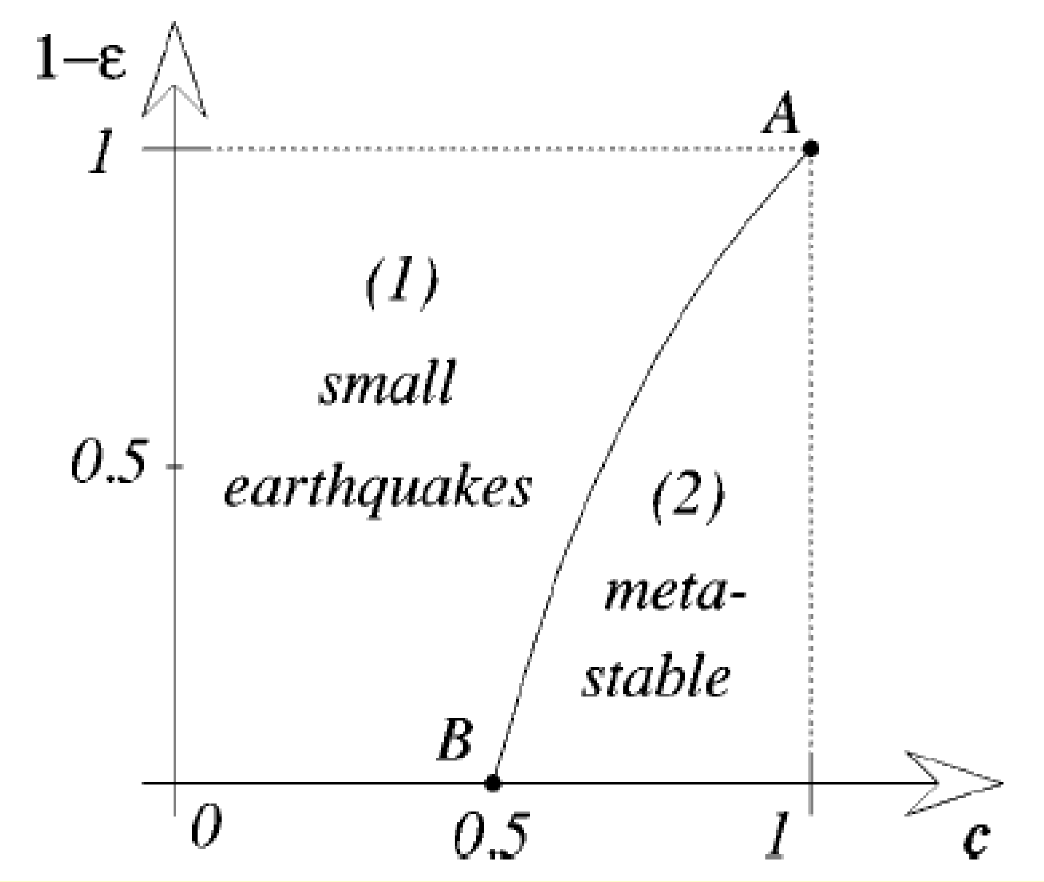
??? • An artifact due to depopulation?
- A statistical scatter?
- A real break in scaling law?

HOWEVER ...

According to Kagan (1993), evidence of the characteristic hypothesis can be explained either by statistical bias or statistical artifact.

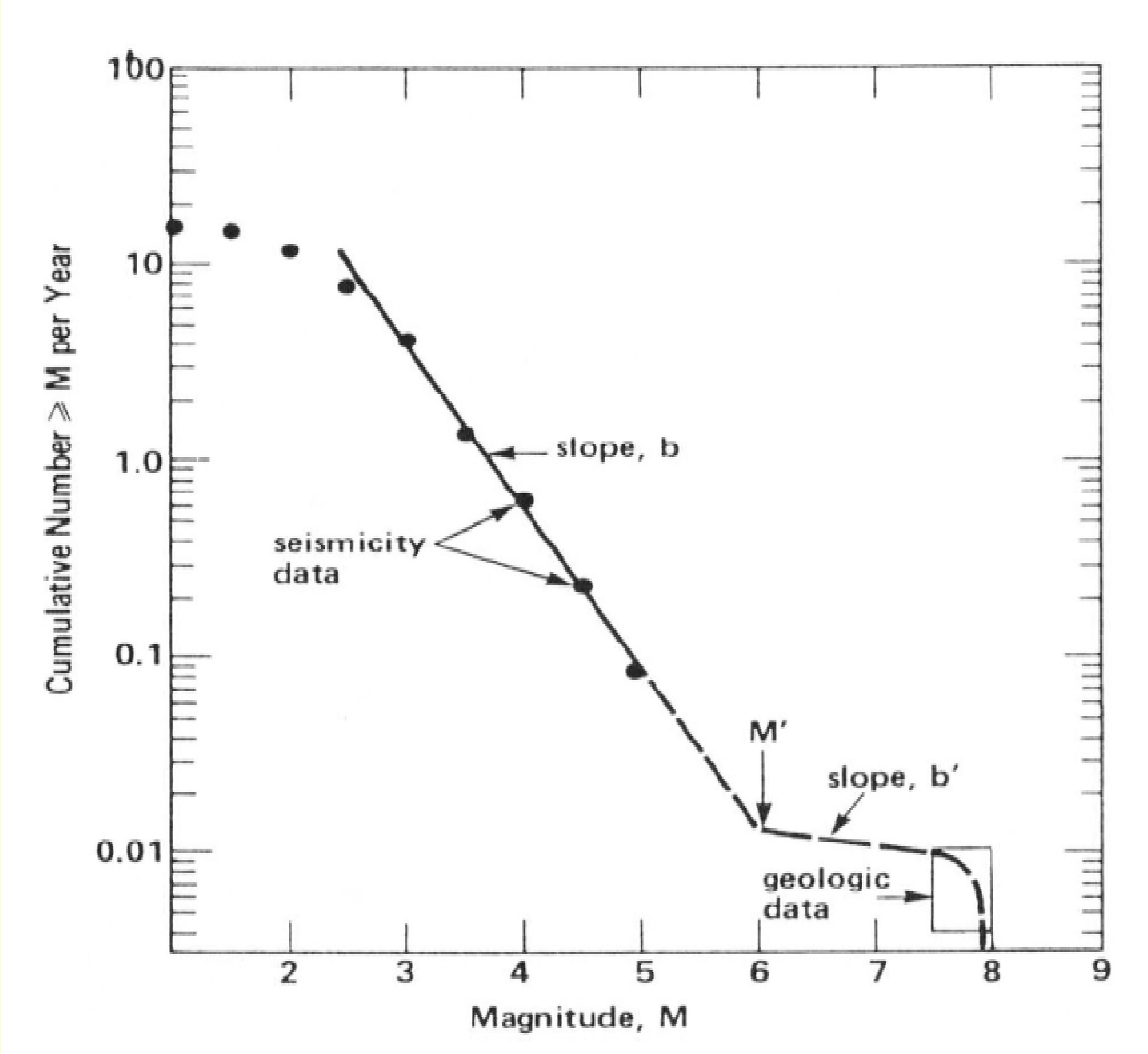
REGULAR AND CHARACTERISTIC EARTHQUAKES

Ben-Zion and Rice (1993) showed that a class of simple models of ruptures along a heterogeneous fault zone displays both types of behavior.



Schematic phase diagram of the system. There is a ‘‘coexistence’’ of two persistent stationary states called Gutenberg-Richter and runaway phases, in a finite region of parameter space, marked region ‘‘(2) metastable.’’ For region 1 given by $c < c^* = 1/(1+\varepsilon)$ (line *AB*) one finds only small avalanches, i.e., the system is always in the Gutenberg-Richter phase.

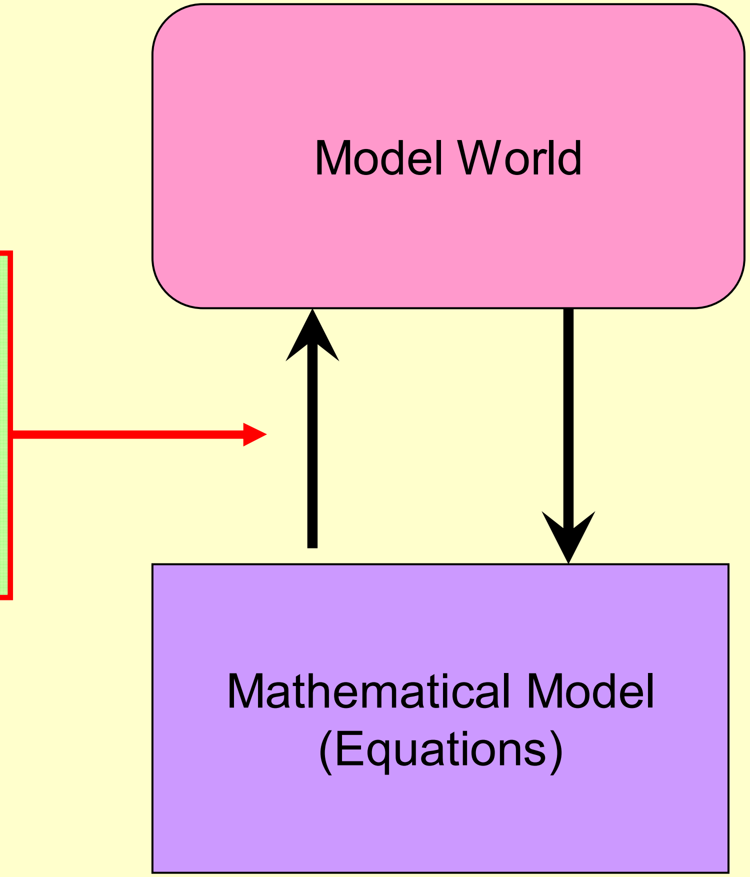
SEISMICITY. 5



PART 2: QUANTIFICATION OF OBSERVATIONS

The Modeling Process

Synthesize the statistical characteristics of observations into probability density functions



FREQUENCY SIZE STATISTICS: SEISMIC MOMENT

Kagan (2002)

Gutenberg – Richter: $$\log N(m) = a - bm \quad (1)$$

where *N(m)* is the number of earthquakes with magnitude ≥ *m*. In terms of the seismic moment *M*, *m* and *M* are empirically related as

$$m = \frac{2}{3}\log M - 6.0 \quad (2)$$

Because of the limited sensitivity of the seismographic networks a catalogue completeness threshold (observational cutoff) m_t has to be introduced, and (1) becomes

$$\log N(m) = a_t - b(m - m_t). \quad (3)$$

by combining (3) and (2) the original G-R is transformed into the **Pareto** distribution for scalar seismic moment

$$\phi(M) = \beta M_t^{\beta} M^{-1-\beta} \quad \text{for} \quad M_t \le M \quad (4)$$

where β is the index parameter of the distribution, β = 2/3 b. Note that whereas the magnitude is not a physical parameter, the seismic moment is.

FREQUENCY SIZE STATISTICS: SEISMIC MOMENT

Kagan (2002)

Simple considerations of the finiteness of seismic moment flux or of deformation energy, available for an earthquake generation, require that the Pareto distribution (4) be modified at the large size end of the moment scale: the distribution density tail have a decay stronger that $M^{-1-\beta}$ with $\beta > 1$.

This problem is solved by introducing into the distribution an additional parameter, called the *maximum* or *corner* moment (M_x or M_c).

Requirements that should satisfy the statistical distribution of seismic moment:

1. Should have an extended scale-invariant part, describing small and moderate earthquakes.
2. Should have a small number of parameters.
3. A sharp cutoff at the large event distribution tail is not warranted, since it contradicts the known behavior of dissipative physical dynamical systems. The abrupt truncation may be replaced by a soft Gaussian-like roll-off, with the penalty of introducing an additional degree of freedom.

The distribution which is a power-law form for small values of the argument and has an exponential tail is the **gamma** density distribution

$$\phi(M) = C^{-1} M^{-1-\beta} \exp(-M/M_c) \quad \text{for } M_c \le M < \infty$$

Here M_c, is the parameter that controls the distribution in the upper ranges of M and C is a normalizing coefficient:

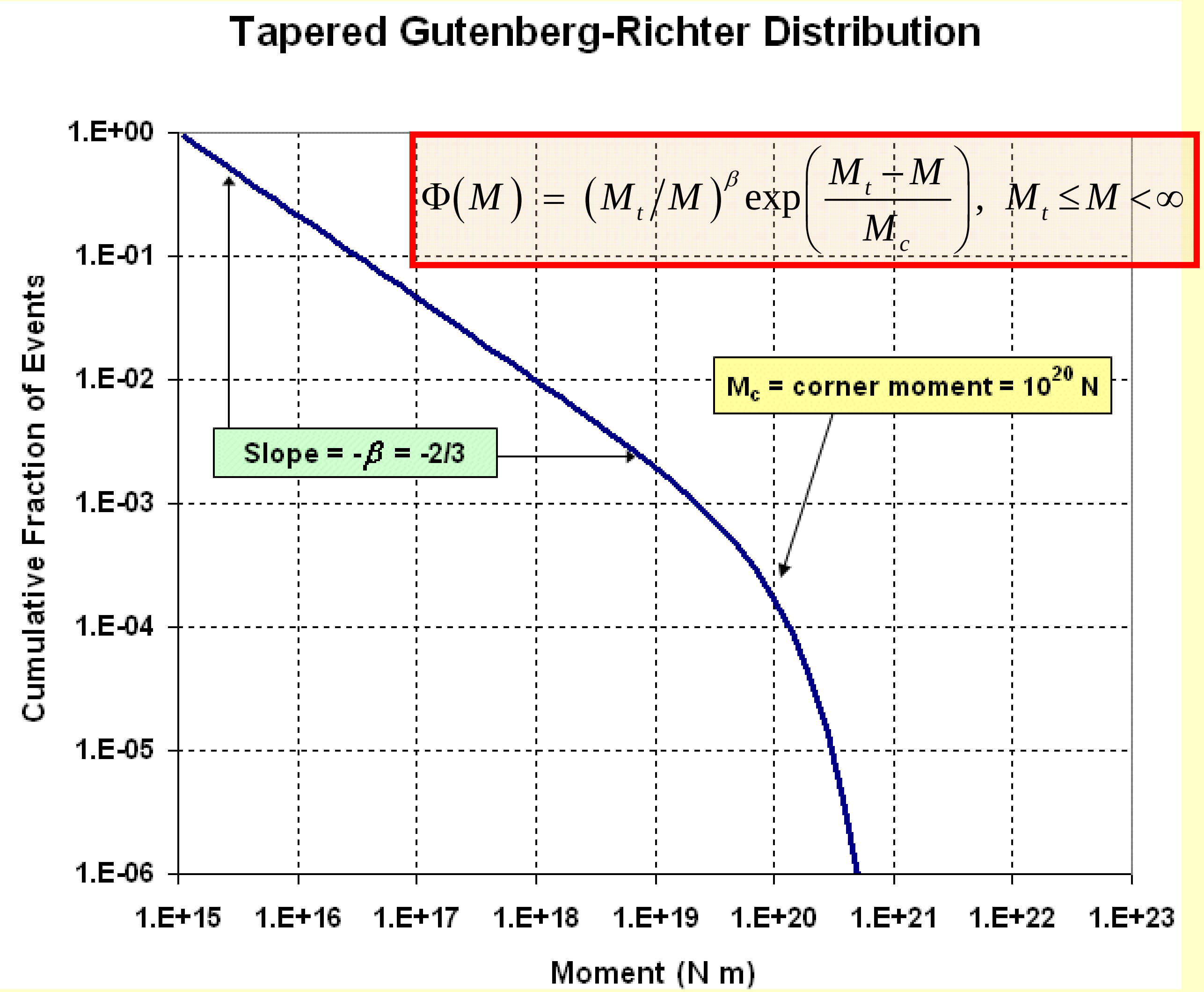
The **tapered G-R distribution** of M, **TGR**, is

$$\Phi(M) = (M_t/M)^{\beta} \exp\left(\frac{M_t - M}{M_c}\right), \quad M_t \le M < \infty$$

$M_t \le M$ is the threshold seismic moment.

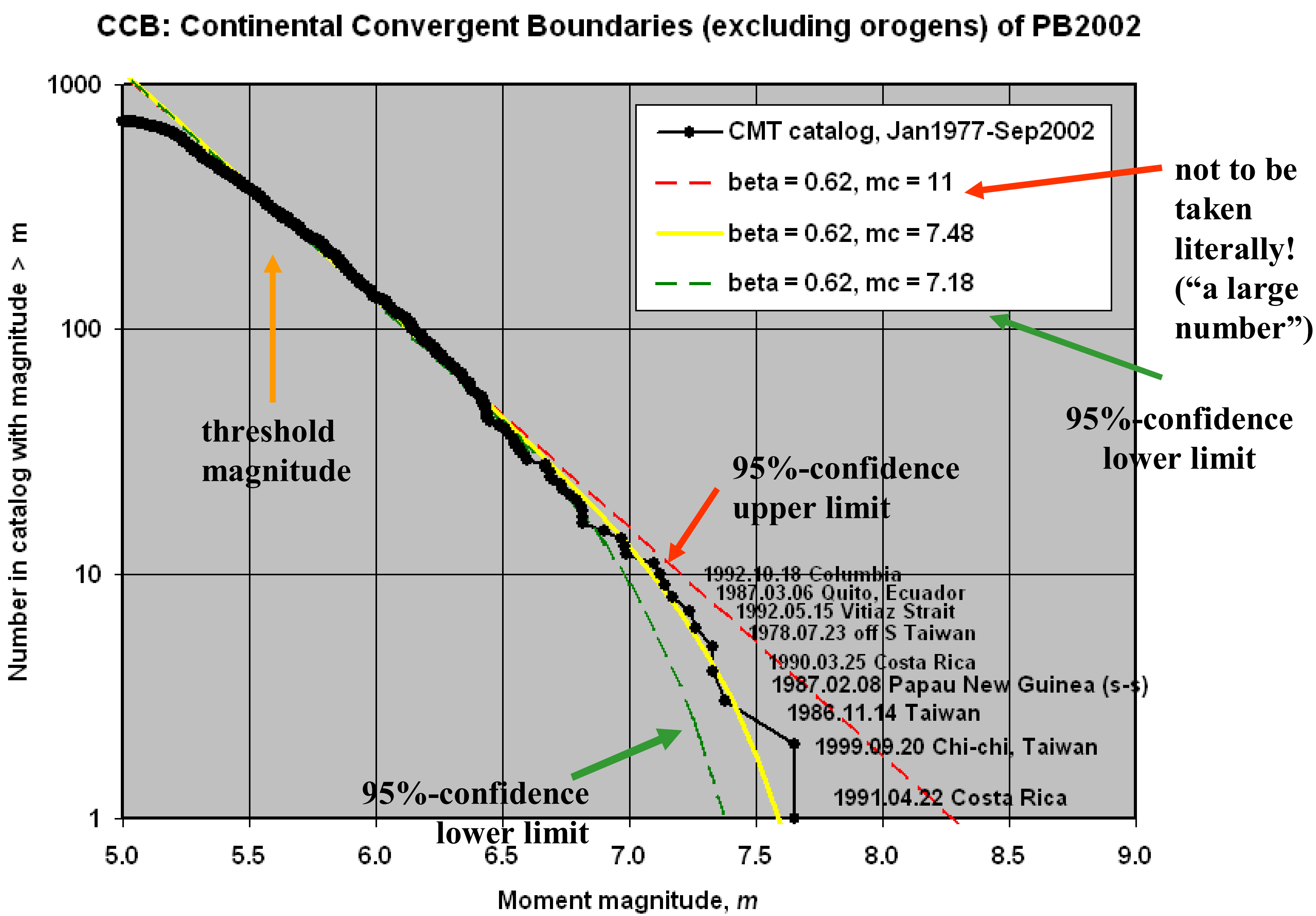
For small values of M the TGR behaves as a power law (Pareto) distribution

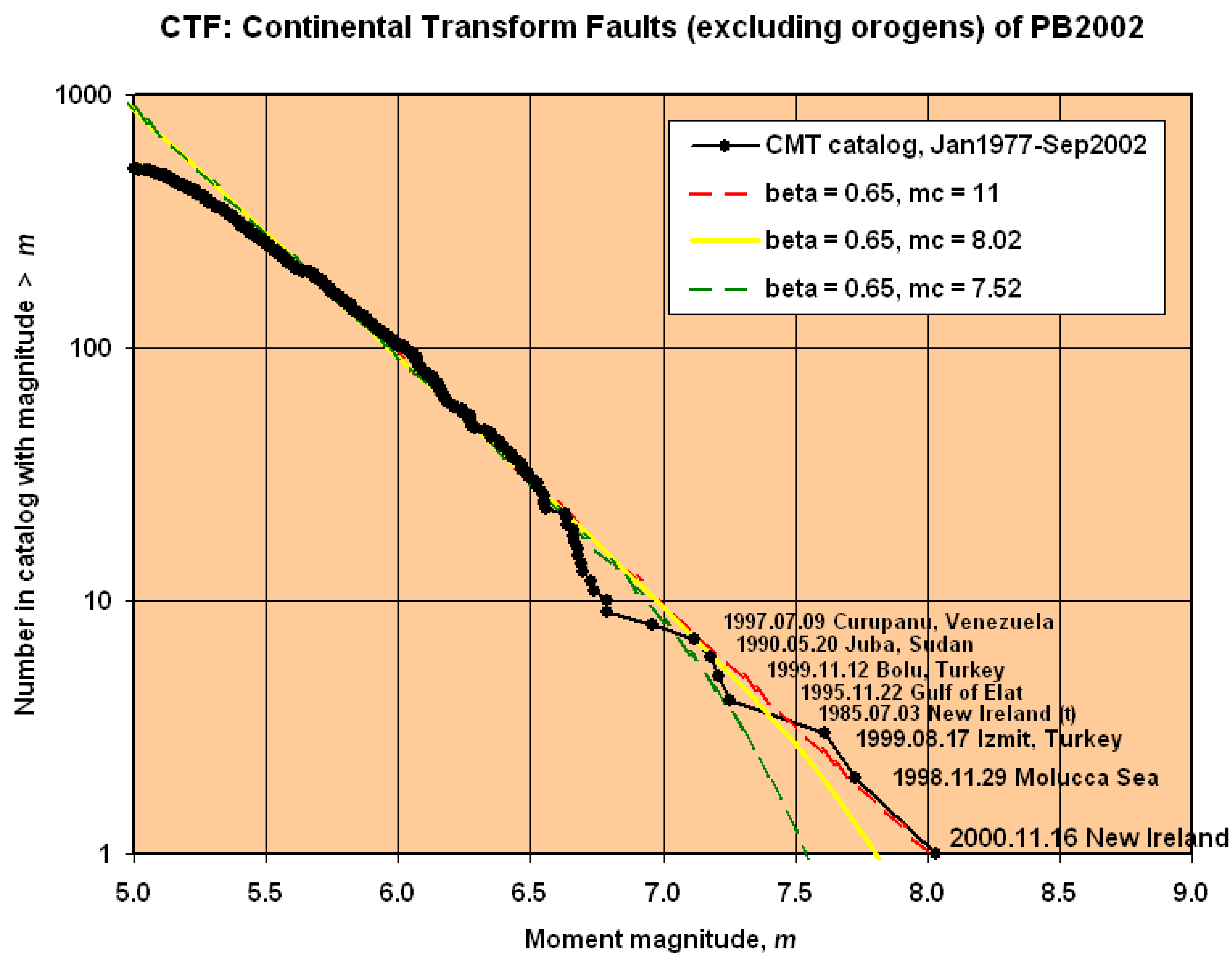
$$\Phi(M) = (M_t/M)^{\beta}$$



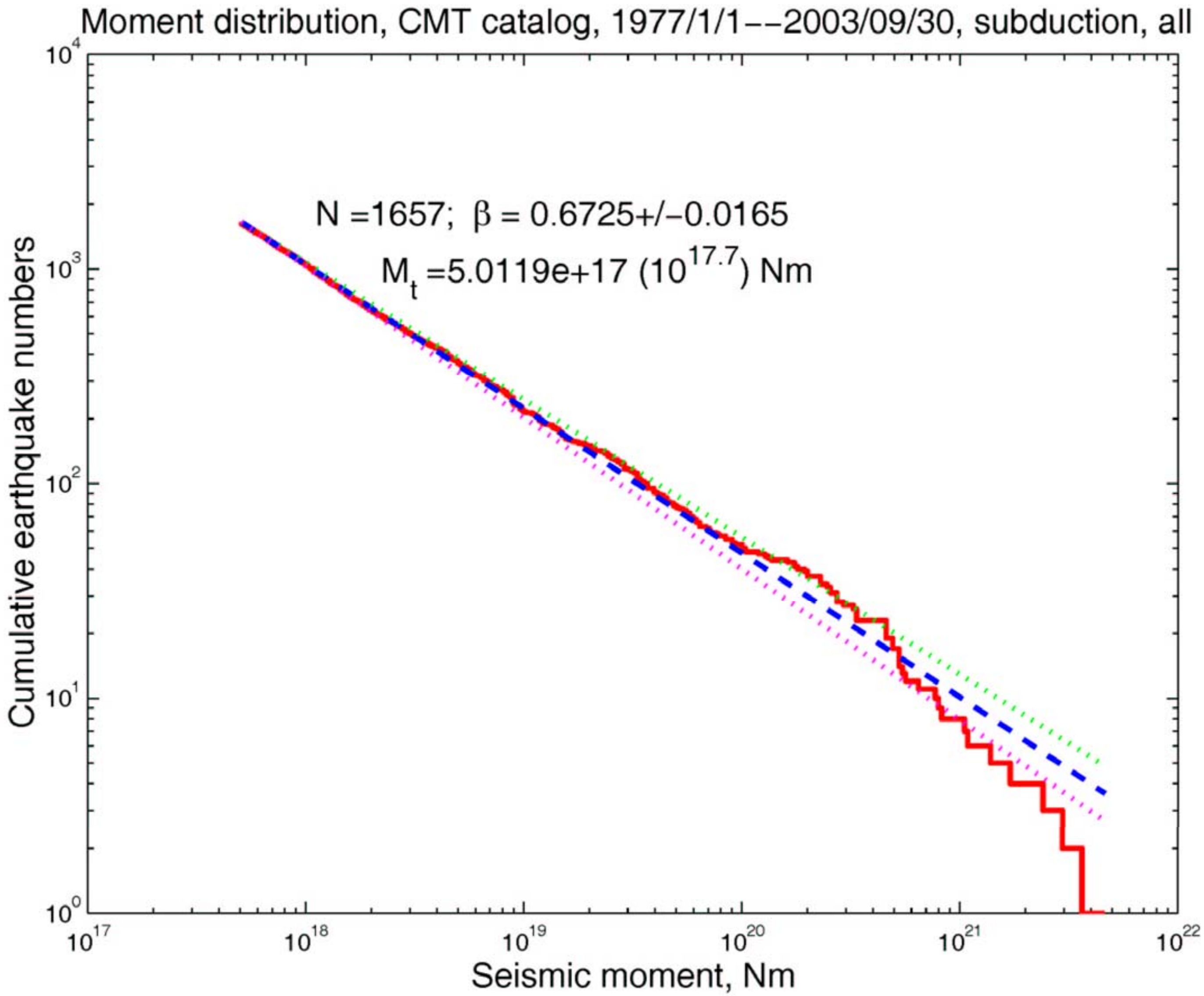
Advantages of this distribution:

- ✓**Simple (only one more parameter than G-R);**
- ✓**Has a finite integrated moment (unlike G-R) for $\beta < 1$;**
- ✓**Fits global subcatalogs slightly better than the gamma distribution.**


CCB: Continental Convergent Boundaries (excluding orogens) of PB2002
CMT catalog, Jan1977-Sep2002
beta = 0.62, mc = 11
beta = 0.62, mc = 7.48
beta = 0.62, mc = 7.18
not to be taken literally! (“a large number”)
95%-confidence lower limit
threshold magnitude
95%-confidence upper limit
1992.10.18 Columbia
1987.03.06 Quito, Ecuador
1992.05.15 Vitiaz Strait
1978.07.23 off S Taiwan
1990.03.25 Costa Rica
1987.02.08 Papau New Guinea (s-s)
1986.11.14 Taiwan
1999.09.20 Chi-chi, Taiwan
1991.04.22 Costa Rica
95%-confidence lower limit
Number in catalog with magnitude > m
1000
100
10
1
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
Moment magnitude, m



CTF: Continental Transform Faults (excluding orogens) of PB2002
1000
100
10
1
Number in catalog with magnitude > m
5.0
5.5
6.0
6.5
7.0
7.5
8.0
8.5
9.0
Moment magnitude, m
CMT catalog, Jan1977-Sep2002
beta = 0.65, mc = 11
beta = 0.65, mc = 8.02
beta = 0.65, mc = 7.52
1997.07.09 Curupanu, Venezuela
1990.05.20 Juba, Sudan
1999.11.12 Bolu, Turkey
1995.11.22 Gulf of Elat
1985.07.03 New Ireland (t)
1999.08.17 Izmit, Turkey
1998.11.29 Molucca Sea
2000.11.16 New Ireland


Moment distribution, CMT catalog, 1977/1/1--2003/09/30, subduction, all
N =1657; β = 0.6725+/-0.0165
M_t =5.0119e+17 (10^17.7) Nm
Cumulative earthquake numbers
Seismic moment, Nm
10^0
10^1
10^2
10^3
10^4
10^17
10^18
10^19
10^20
10^21
10^22

Omori's Law

Aftershocks occur with a pattern that follows Omori's law, or more correctly the modified Omori's law, and is an empirical relation for the temporal decay of aftershock rates. In 1894, Omori published his work on the aftershocks of earthquakes, in which he stated that aftershock frequency decreases by ***roughly the reciprocal of time after the main shock***, or, more accurately***,***

$$n(t) = \frac{k}{c+t}$$

where:

- *k* is the amplitude, and
- *n(t)* is the rate of earthquakes measured in a certain time *t* after the main shock,
- *c* is the "time offset" parameter.

The modified version of Omori's law, now commonly used, was proposed by Utsu in 1961.

$$n(t) = \frac{k}{(c+t)^p}$$

$$n(t) = \frac{k}{(c+t)^p}$$

p modifies the decay rate and typically falls in the range 0.7–1.5.

The rate of aftershocks is proportional to the inverse of time since the mainshock.

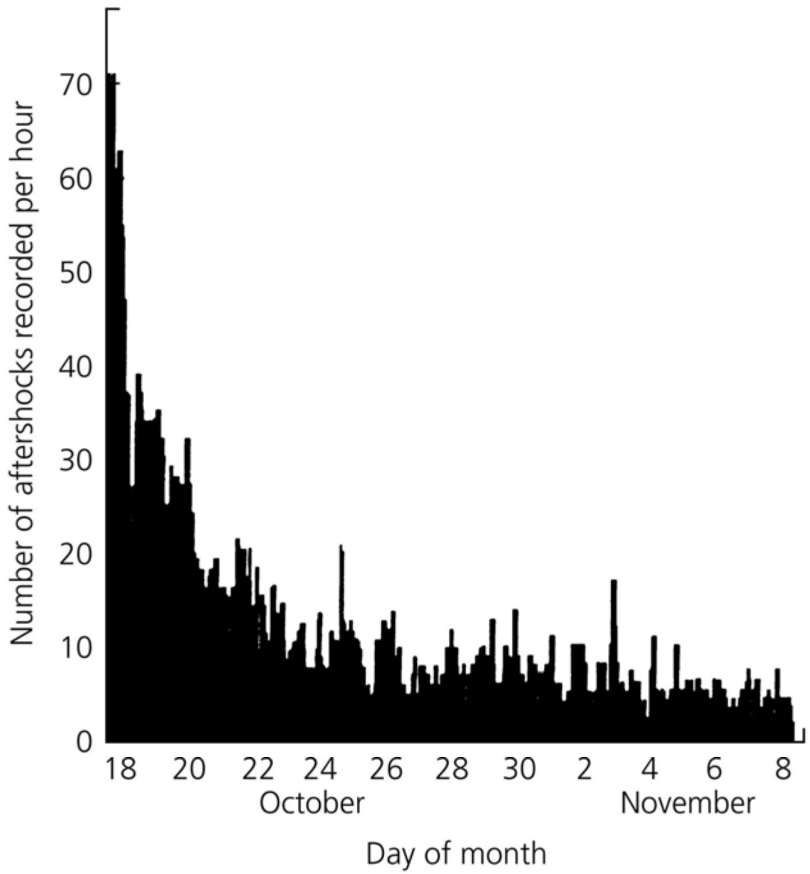
These patterns describe only the mass behavior of aftershocks; the actual times, numbers and locations of the aftershocks are 'random', while tending to follow these patterns.

As this is an empirical law, values of the parameters are obtained by fitting to data after the mainshock occurred and they have (not yet) physical basis/meaning. (Omori, 1984; Utsu et al, 1995).

OMORI'S LAW

Figure 4.7-8: Aftershocks following the 1989 Loma Prieta earthquake.

Magnitude	Number	Effect
5	2	Damaging
4	20	Strong
3	65	Perceptible
2	384	Not felt
1	1855	Not felt
<1	2434	Not felt
Total	4760	



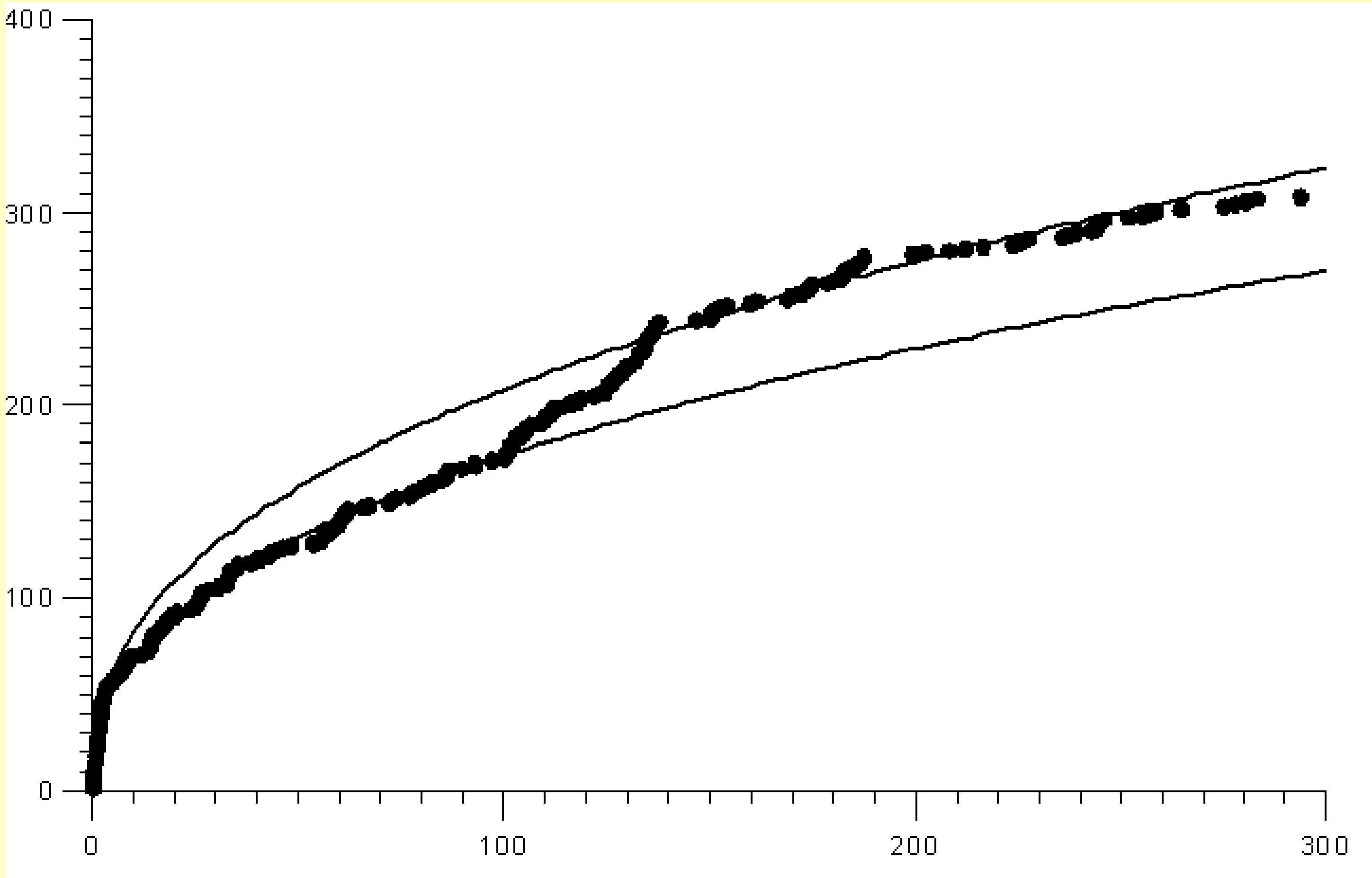
4760 aftershocks of the Loma Prieta earthquake had been recorded by noon on November 7, 1989. The diminishing number of aftershocks with time is typical for large California earthquakes.

AFTERSHOCK TIME SERIES. 1

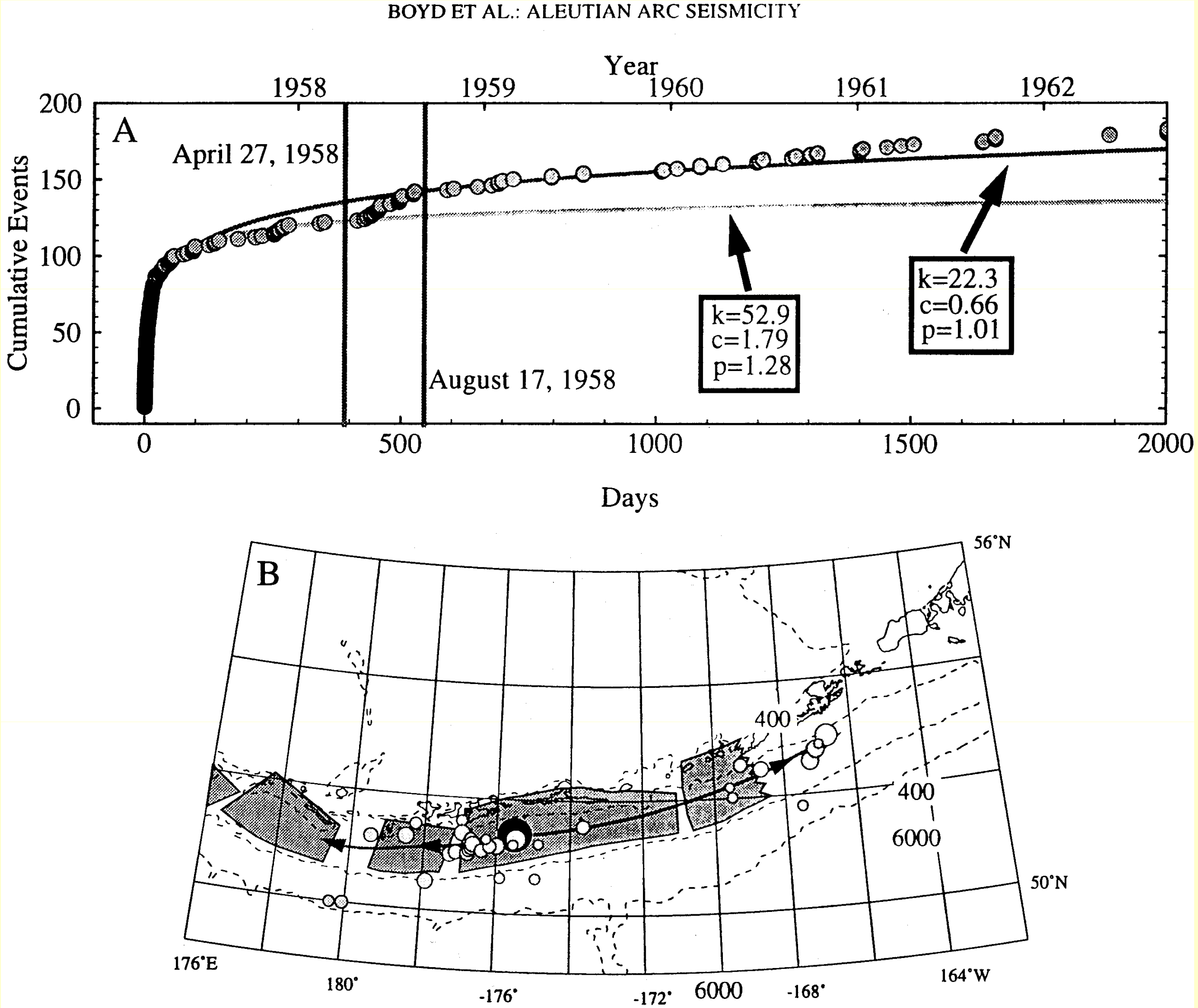
- The geometrical structure of the aftershock time series is composed of a **relaxation process** (Omori's potential law) and **fluctuations,** that can be positive (accelerations) or negative (decelerations).
- (A possible pattern could emerge: in the observed aftershock time series, positive fluctuations dominate over the negative ones.)

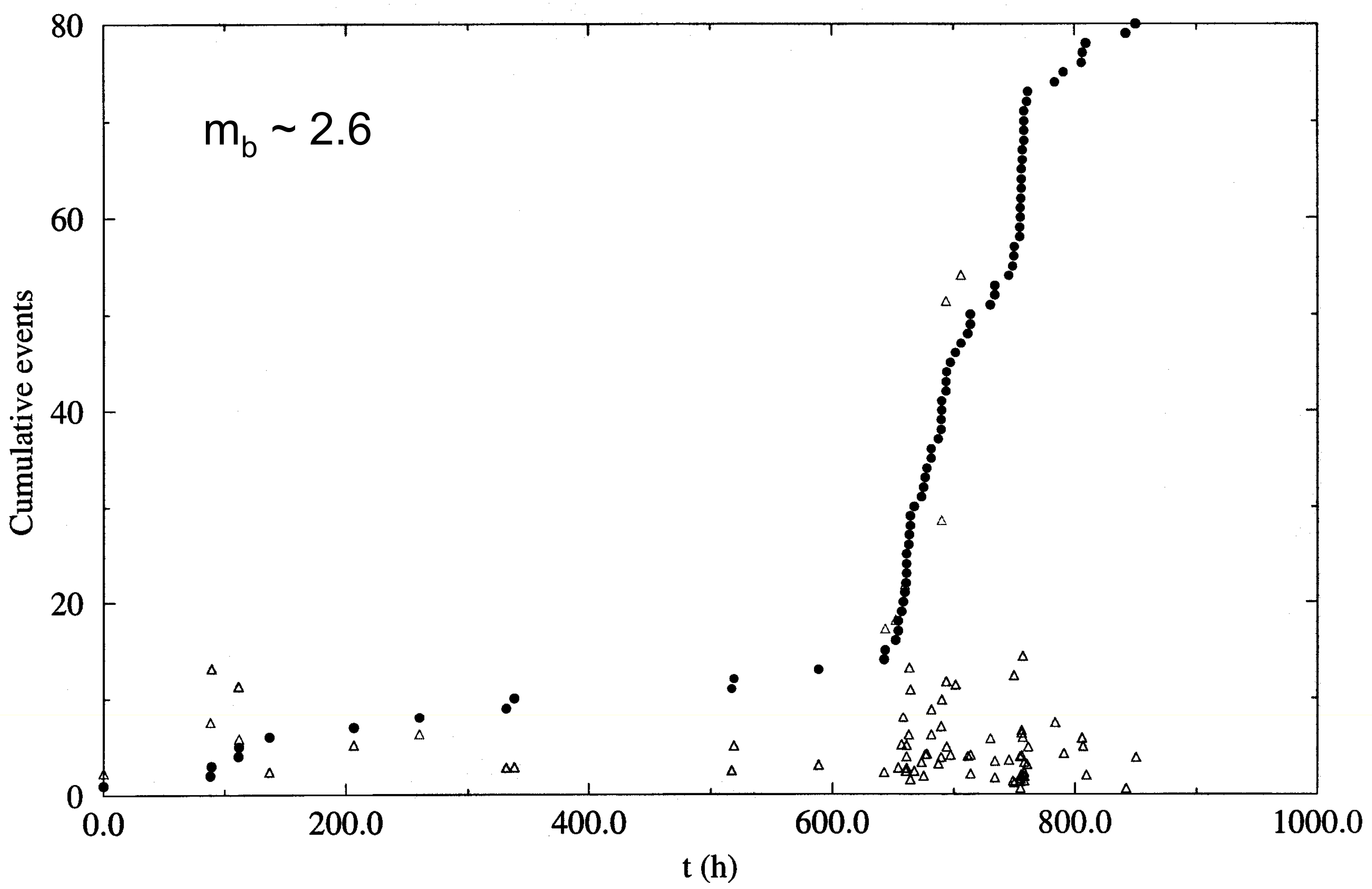
HOWEVER …

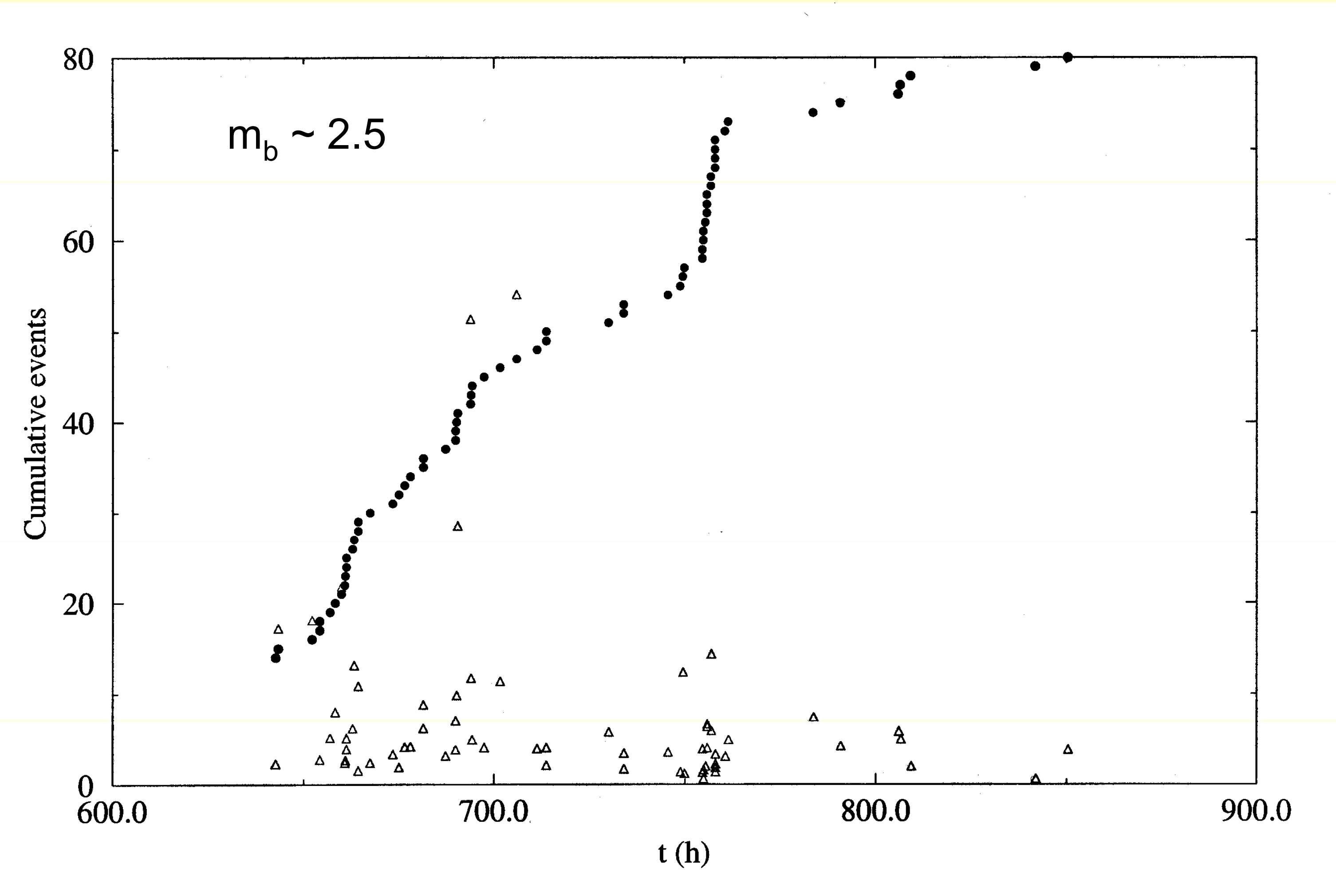
- Fluctuations are too large
- Model parameters are time dependent (back to equilibrium?)



(Correig *et al*., 1997)


BOYD ET AL.: ALEUTIAN ARC SEISMICITY
Year
1958
1959
1960
1961
1962
A
April 27, 1958
August 17, 1958
Cumulative Events
200
150
100
50
0
k=52.9
c=1.79
p=1.28
k=22.3
c=0.66
p=1.01
0
500
1000
1500
2000
Days
B
56°N
50°N
400
400
6000
6000
176°E
180°
-176°
-172°
-168°
164°W


$m_b \sim 2.6$
Cumulative events
0
20
40
60
80
0.0
200.0
400.0
600.0
800.0
1000.0
t (h)



$m_b \sim 2.5$
Cumulative events
t (h)
0
20
40
60
80
600.0
700.0
800.0
900.0

AFTERSHOCK TIME SERIES. 2

- A lot of mean field models have been constructed to deal with aftershock, some of them based on physical grounds (Yamashita (2003), GJI, 152, 20-33) and some other, based on the validity of universal laws of O and G-R (Helmstetter & Sornette (2003), 107(B10), 2237, doi: 10.1029/2001JB001580, 2002). None of them, however, account for fluctuations or time variation of model parameters.
- Moreover, every model is able to explain some specific features, but not other.

AFTERSHOCK TIME SERIES. 3

- In some way, aftershocks are similar to precursory activity :

  - not always appear
  - not always repeat

Spatial epicenter/hypocenter distributions

- Distribution of distances between earthquakes characterizes their fractal structure. Using these distributions we determine fractal dimension of fault system.
- For planar faults the dimension (D) is 2.0.
- For shallow seismicity D=2.2, for deep earthquakes D=1.6; this means that deep event clusters are separated by aseismic regions.

UNIFIED LAW: TEMPORAL STATISTICS

Bak *et al.* (2002) attempted an explanation of **interevent time** of earthquake occurrence by unifying the observations on

- statistics of earthquakes,
- the geometrical fractal structure of hypocentral locations
- the fractal structure displayed by faults, considered all of them as a result of a dynamical process.

The underlying philosophy was (Corral, 2004)

- Do not bother about the tectonic environment.
- Do not bother about aftershocks and foreshocks, all are equally treated.
- Do not bother about temporal heterogeneity.

THE END OF SEISMIC CYCLE?

The abandon of the ingrained concept (in many seismologists' mind) of the distintion between foreshocks, aftershocks and mainshocks is an important step toward a simplification and toward an understanding of the mechanism underlying earthquake sequence.

(*Helmstetter et al., 2003*)

GLOBAL SCALING LAWS. 1

Bak et al. (2002) carried out a spatiotemporal analysis over a region with a grid with cells of dimension L x L and defined the waiting time (interevent time) as the time interval of two successive events.

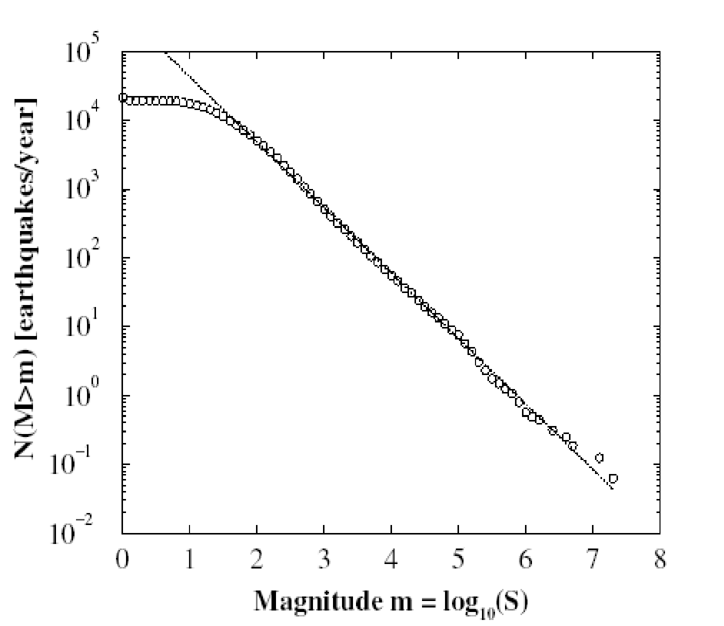
The **distribution of waiting times T, $P_{S,L}(T)$**, was measured between earthquakes occurring within a range L, whose magnitude are greater that **$m = m_c = \log(S)$,** S being the fault surface related to the energy as **$S \sim E^{2/3}$.**

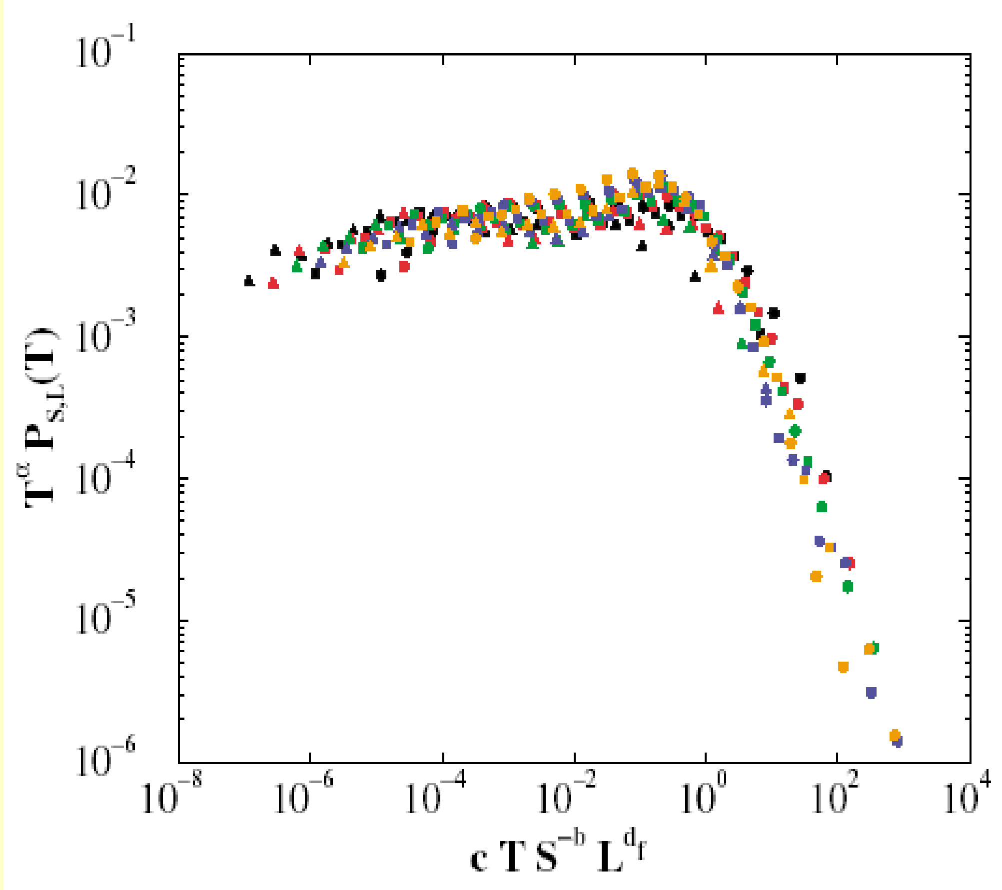
For a suitable choice on the interval exponent α, the magnitude exponent b, and the spatial dimension df, all the data collapse onto a single well-defined curve **f(x), y = f(x)**, that is: **$T^{\alpha}P_{S,L}(T) = f(TS^{-b}L^{df})$**

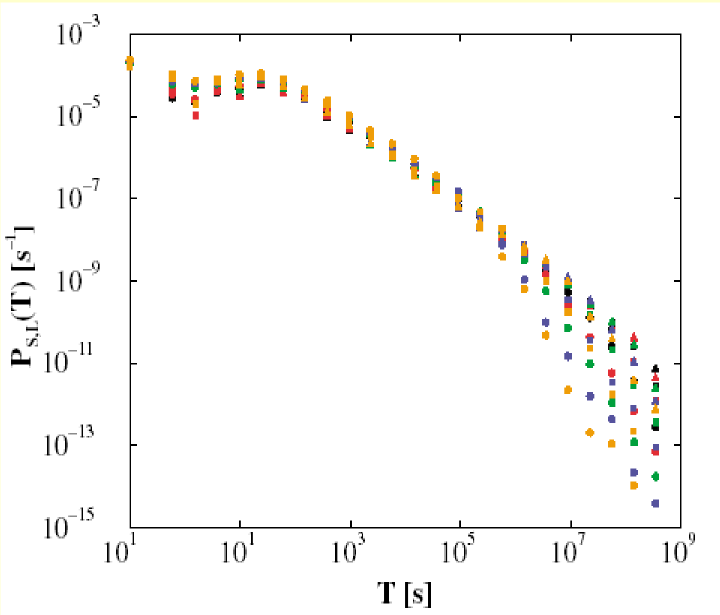
This equation expreses the **unified scaling law for earthquakes**, and consists on a constant part and a decaying part, separated by a sharp kink.

The index α ~ 1 can be identified as the **Omori-law exponent**, b ≈ 1 is the b value in the **Gutenberg-Richter law**, and **d_f** describes the 2D **fractal dimension of the epicentral distribution.**

Due to the fact that the variable **$x = TS^{-b}L^{df}$** has no absolute meaning, **there is no unique way of characterizing earthquakes as aftershocks or main shocks**.







The data with T > 38 s (left bottom) replotted with $T^{\alpha}P_{S,L}L^{df}$ as a function of the variable $x = cTS^{-b}L^{df}$, $c = 10^{-4}$. The data collapse implies a <u>unified law</u> for earthquakes. The Omori law exponent a = 1, Gutenberg-Richter value b = 1, and fractal dimension d_f = 1.2 have been used in order to <u>collapse all the data onto a single, unique curve *f(x)*.</u> The estimated uncertainty in the exponents is less than 0.2.

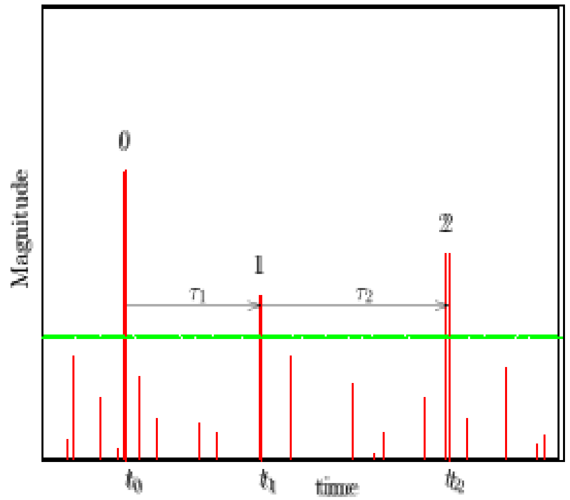
Corral (2007)

Recurrence Times

- Consider a fixed spatial region
- Consider earthquakes with magnitude larger than a threshold, $M \geq M_c$
- Compute recurrence time as the time between consecutive earthquakes

$$\tau_i \equiv t_i - t_{i-1}$$

$$i = 1, 2, 3 \ldots$$



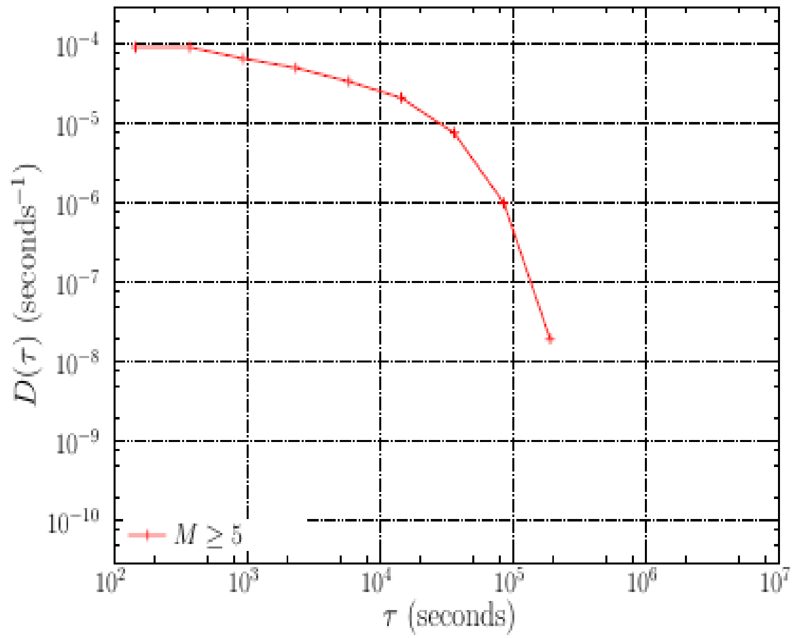
$\Longrightarrow$ Broad scale of times, from seconds to years!

Corral (2007)

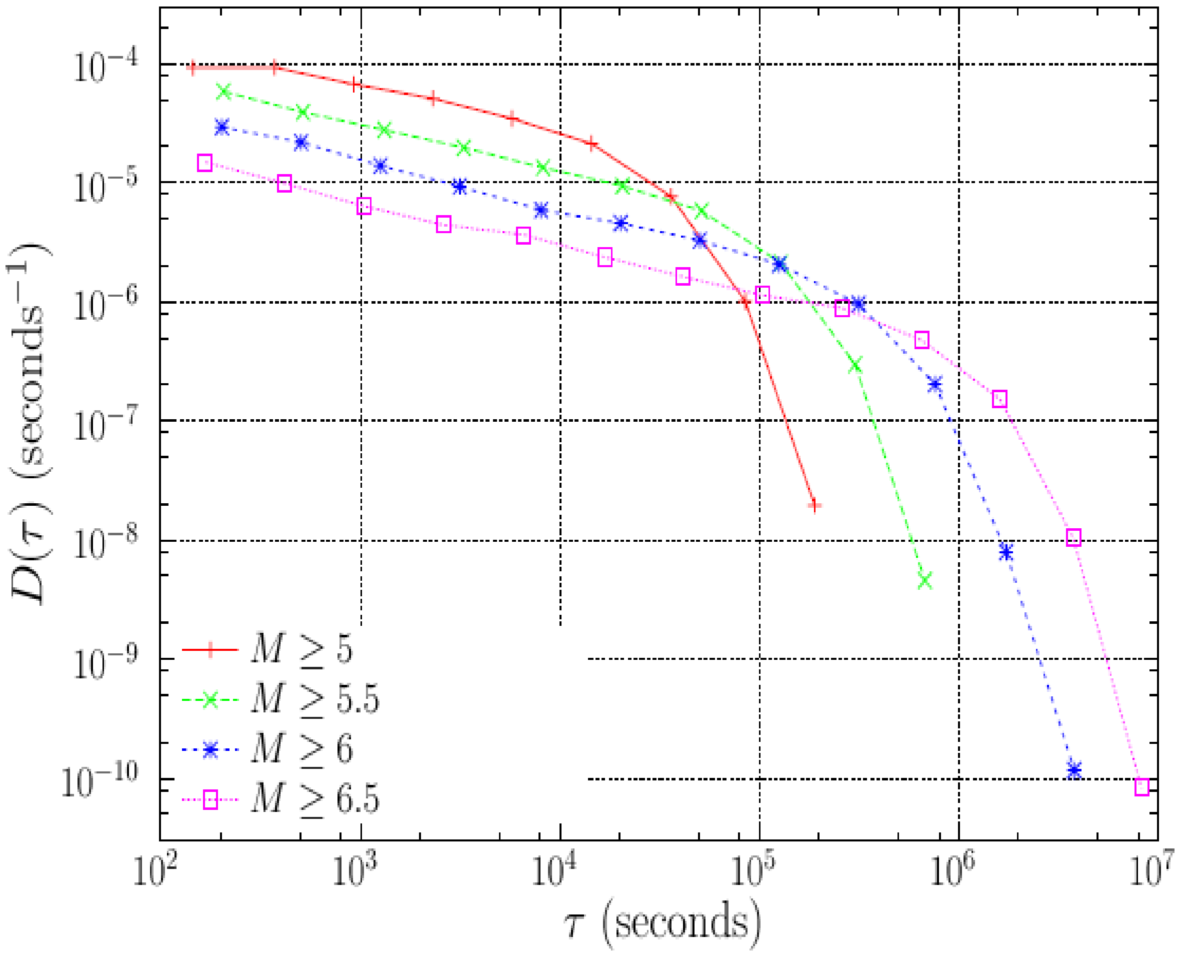
Recurrence time probability density $D_w(\tau)$: (depends on window $w = \{M \geq M_c\}$)

$$D_w(\tau) \equiv \frac{\text{Prob}[\tau < \text{recurrence time} \leq \tau + d\tau]}{d\tau}$$

For worldwide seismicity for the years 1973-2002 and $M_c \geq 5$



Corral (2007)



$D(\tau)$ (seconds^{-1})
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
$M \geq 5$
$M \geq 5.5$
$M \geq 6$
$M \geq 6.5$
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
10^{7}
τ (seconds)

Corral (2007)

Scaling: Scale transformation of the axis

$$\tau \longrightarrow R_w \tau$$
$$D_w(\tau) \longrightarrow D_w(\tau)/R_w$$

with R_w the rate of seismic activity, $R_w =$ number of earthquakes per unit time

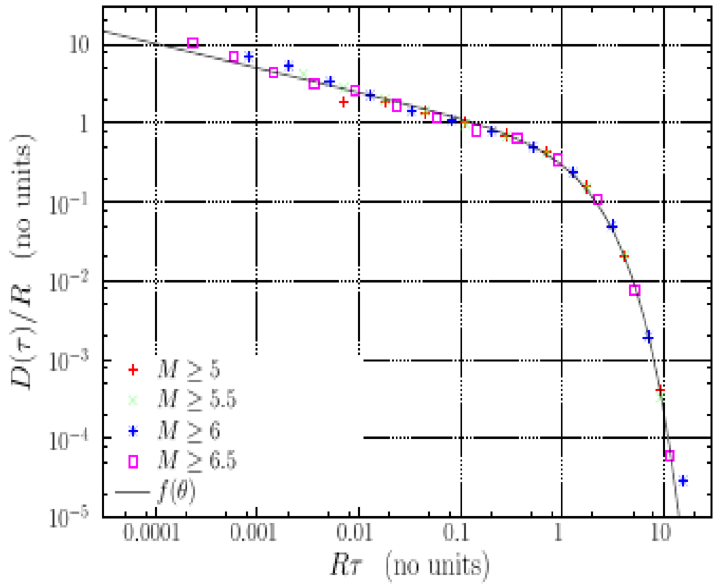
$\Longrightarrow$ Scaling law:

$$D_w(\tau) = R_w f(R_w \tau)$$

Gutenberg-Richter law:

$$R_w \propto 10^{-bM_c}$$

(usually, $b \simeq 1$)



Corral (2007)

- Combining the scaling law with Gutenberg-Richter law

$$R_w \propto 10^{-bM_c} \propto \frac{1}{E^{\beta}} \quad \Rightarrow \quad D_w(\tau) \equiv D(\tau; E) \propto E^{-\beta} f(\tau E^{-\beta})$$

with $\beta = 2b/3$

- The condition of scale invariance for two variables,

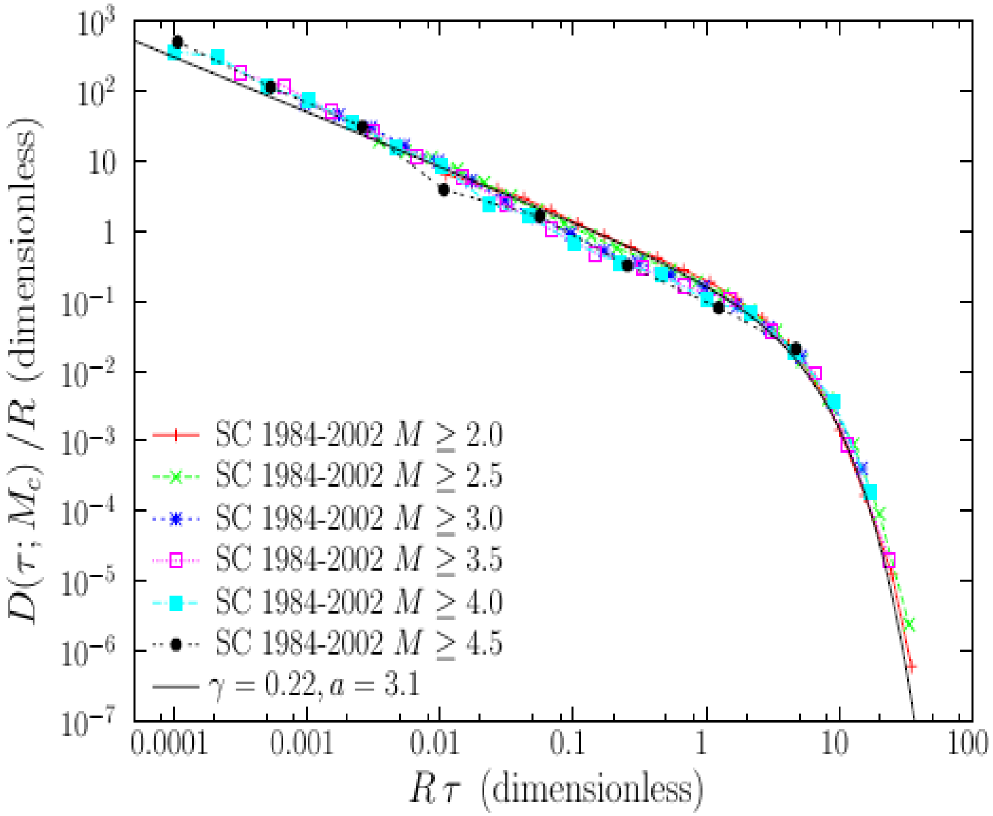
$$F(x, y) = \top[F(x, y)] \equiv c\, F(x/a, y/b),$$

leads to

$$F(x, y) = x^{\alpha} f(y/x^{\beta})$$

Corral (2007)

Southern-California seismicity, from 1984 to 2001, also fulfills a scaling law



Corral (2007)

Scaling function:

- Defining a dimensionless time, $\theta \equiv R_w \tau$

$$f(\theta) \propto \frac{1}{\theta^{0.3}} e^{-\theta/1.4} \quad \text{for worldwide seismicity, 1973-2002}$$

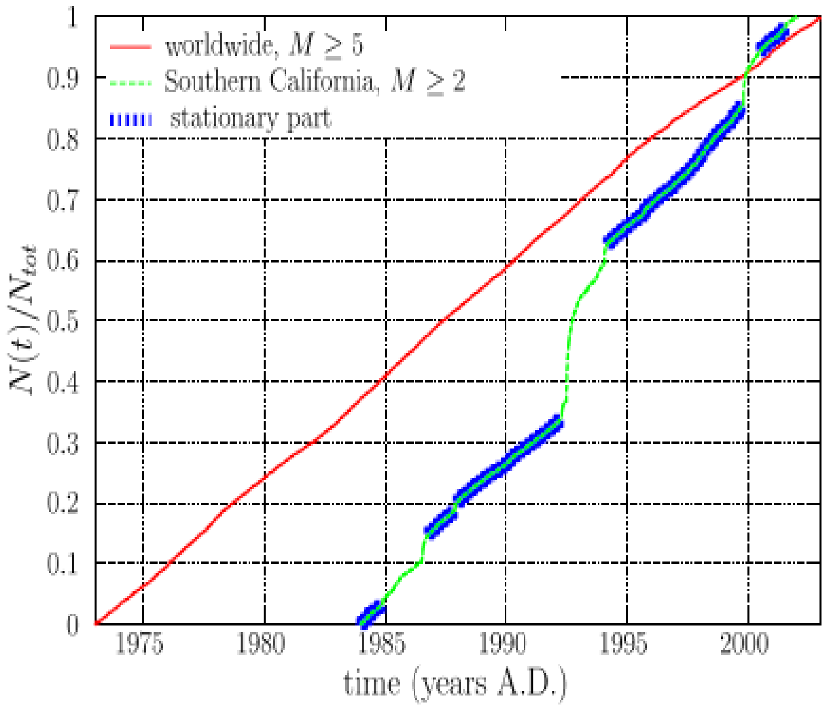
$$f(\theta) \propto \frac{1}{\theta^{0.8}} e^{-\theta/3} \quad \text{for Southern-California, 1984-2001}$$

- $f(\theta)$ is a gamma distribution
- Why are the scaling functions different?

Corral (2007)

Accumulated number of earthquakes (normalized) versus time

with $N_{tot} = 84771$ for Southern California and $N_{tot} = 46054$ for worldwide



⇒ worldwide seismicity is stationary, Southern California is not stationary

SCALING FUNCTION

Tue scaling function f(θ) can be generalized to

$$f(\theta) = \frac{C|\delta|}{a\Gamma(\gamma/\delta)}\left(\frac{\theta}{a}\right)^{\gamma-1} e^{-(\theta/a)^{\delta}},$$

which has a very general shape. Γ(.) is the gamma function. If γ and δ are positive, the former controls the shape for small θ and δ the shape at large θ; the situation is reversed if both parameters are negative; a is a scale parameter and C a normalization correction. θ is a dimensionless time.

If regions of smaller size are considered, the rate turns nonstationary, giving rise to heterogeneities in time. This is due to large earthquakes, which provoke a kind of "avalanches of earthquakes", i.e., aftershock sequences.

PROBABILITY DISTRIBUTION FUNCTION

For short times, the scaling function shows a slow variation not affecting the power law ($1/\tau$) behavior.

For long times a fast decay is obtained, which could be consistent with an exponential distribution and therefore with a Poisson process.

This PDF is relevant because

- shows a spatiotemporal occurrence of earthquakes (as in critical phenomena).
- relates interevent times with the G-R law and the epicentral distribution of earthquakes.
- is valid for all kind of events (foreshocks, mainshock, aftershocks).
- the power law tells us that immediately after an earthquake there is a high probability of return, probability that decreases in time.

OBSERVED PDF D(τ) FOR THE INTEREVENT TIME

594 Á. Corral / Physica A 340 (2004) 590–597

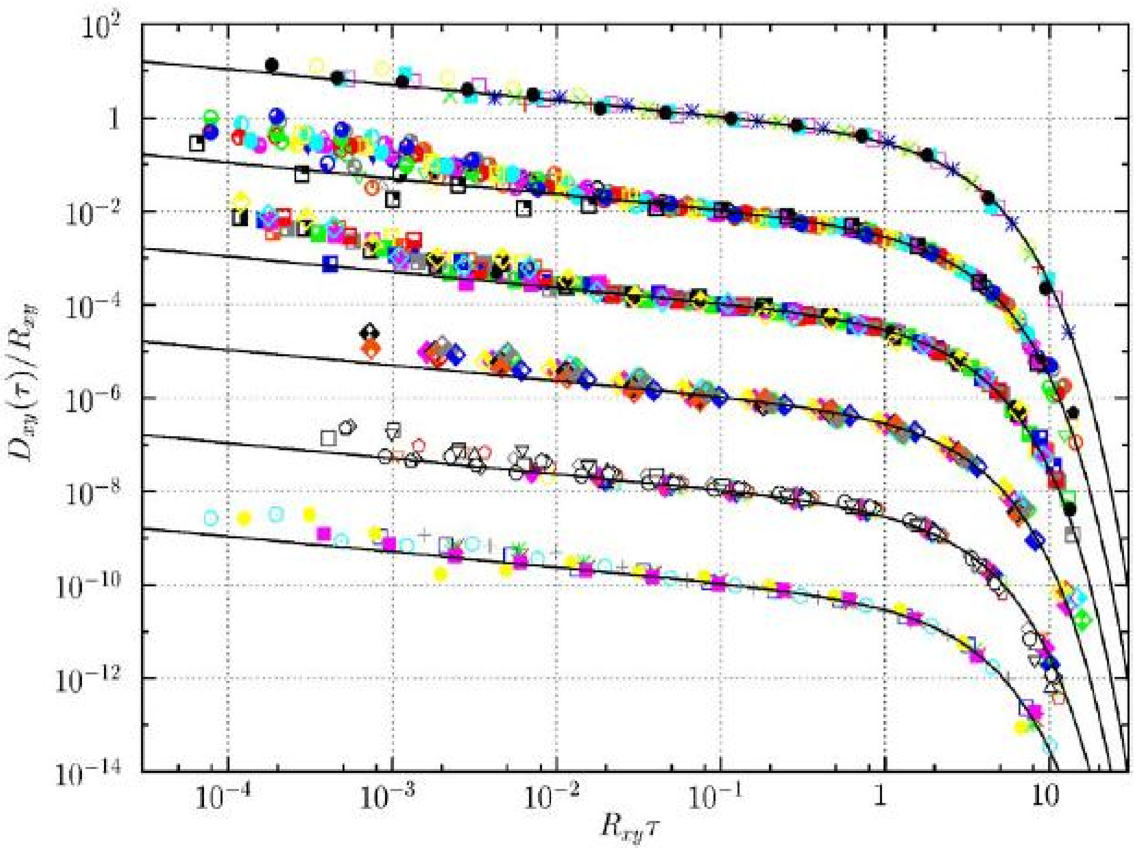


Fig. 1. Single-region recurrence-time probability densities, D_{xy}, after rescaling with the rate R_{xy}.

HOWEVER ...

The universality and interpretation of the scale funcgion has been disputed. Taking into account that seismicity consists of Poissonian mainshocks plus triggered aftershocks, and that there are some regions on Earth with independent seismicity, Molhan (2005) showed that under very general assumptions, from a mathematical point of view

. 1. The above distribution may no exist for all times.

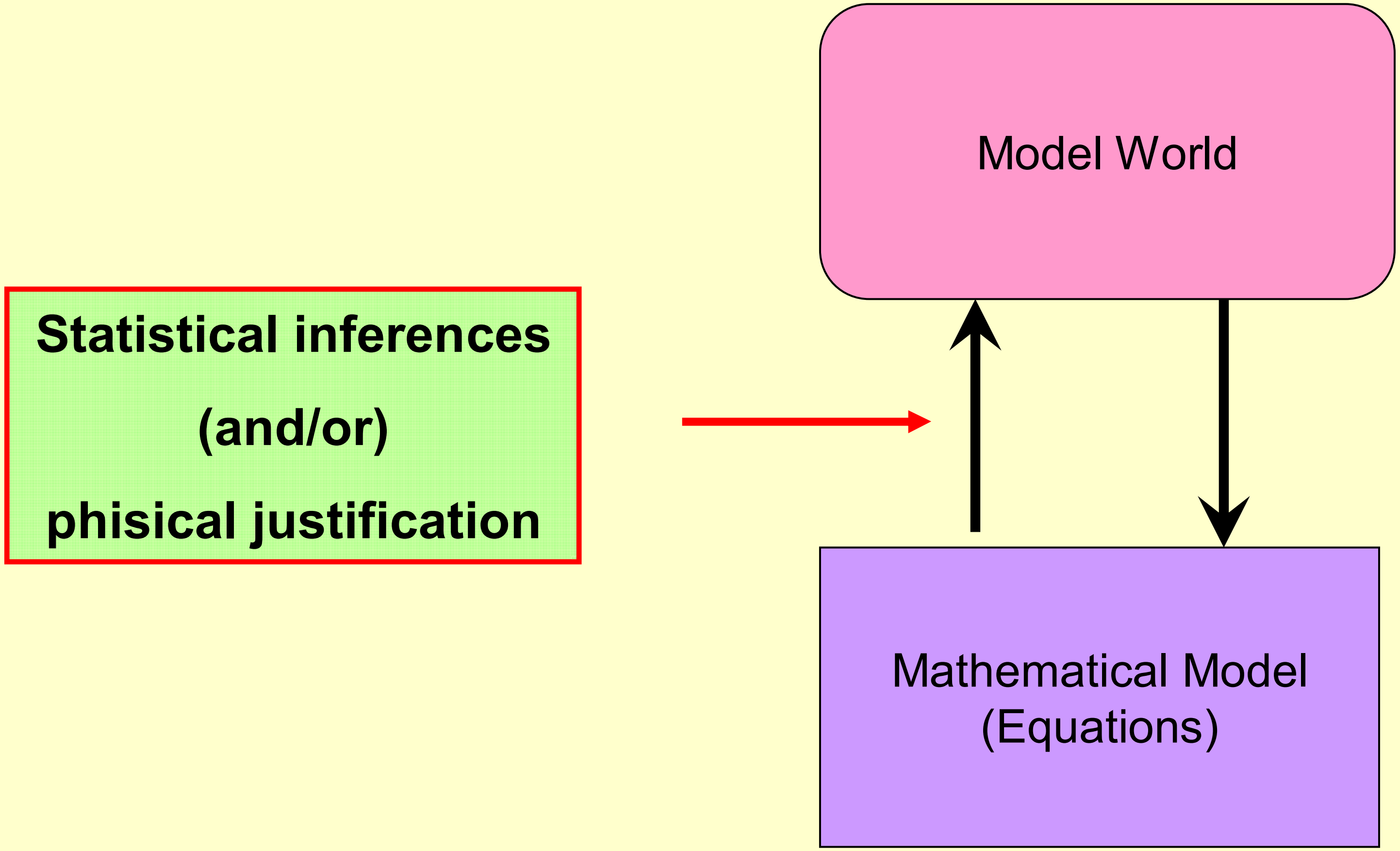
2. The only universal distribution of inter-event times in a stationary point process is exponential.

   a) This universal distribution is not valid for the relatively short intervals associated with aftershocks.

   b) The parameter *1/a* in the PDF reflects the fraction of mainshocks contained in the data.

PART 3: STATISTICAL INFERENCES

The Modeling Process



JUSTIFICATION OF THE GENERALIZED GUTERNBERG-RICHTER LAW AS A GAMMA DISTRIBUTION

INFORMATION THEORY AND ENTROPY

Jaynes (1957)

Information theory provides a constructive criterion for setting up probability distributions on the basis of partial knowledge, and leads to a type of statistical inference which is called the **maximum entropy estimate.**

If one considers statistical mechanics as a form of statistical inference rather than as a physical theory, it is found that the usual computational rules, starting with the determination of the **partition function**, are an immediate consequence of the maximum-entropy principle.

It can be shown that thermodynamic entropy and information-theory entropy appears as the same concept, so that information theory can be applied to the problem of justification of statistical mechanics.

Take the **entropy** as the starting concept and take into account the property that **a probability distribution that maximizes the entropy, subject to certain constraints, becomes the essential fact which justifies the use of that distribution for inference**.

INFORMATION Th. AND ENTROPY

Jaynes (1957)

THE PROBLEM TO BE SOLVED:

Let x_i (i = 1, 2, … n) be a discrete variable, for which the corresponding probabilities p_i are not known. All it is known is the expectation value of the function **f(x):**

$$\langle f(x) \rangle = \sum_{i=1}^{n} p_i f(x_i) \qquad (1)$$

subject to the normalization condition

$$\sum p_i = 1. \qquad (2)$$

On the basis of this information, what is the expectation value of the function **g(x)**?

INFORMATION Th. AND ENTROPY

Jaynes (1957)

According to information theory, there is a unique, unambiguous criterion for the "amount of uncertainty" represented by a **discrete probability distribution**. Following Shanon, the (positive) quantity which increases with increasing uncertainty is

$$H(p_1, \dots p_n) = -K \sum_i p_i \ln p_i \qquad (3)$$

where *K* is a positive constant. Since this is just the expression for entropy as found in statistical mechanics, it is called the **entropy of the probability distribution p_i**.

In making inferences on the basis of partial information, we must use that probability distribution which has maximum entropy subject to whatever is known. To maximize (3) subject to the constrains (1) and (2) use is made of the Lagrangian multipliers **λ**, **μ**, obtaining

$$p_i = e^{-\lambda - \mu f(x_i)} \qquad (4)$$

INFORMATION Th. AND ENTROPY

Jaynes (1957)

$$p_i = e^{-\lambda - \mu f(x_i)}$$

The constants λ, μ are determined by substituting into (1) and (2). The result may be written in the form

$$\langle f(x) \rangle = -\frac{\partial}{\partial \mu} \ln Z(\mu)$$

$$\lambda = \ln Z(\mu)$$

$$Z(\mu) = \sum_i e^{-\mu f(x_i)} \leftarrow \text{partition function}$$

APPLICATION TO SEISMICITY

Main & Burton (1984)

Let *m* stand for the moment magnitude and M_0 for the seismic moment. Consider the continuous range of magnitudes (m_c, ω), where ω is the maximum magnitude, and m_c is a lower bound. The "**missing information**" (our ignorance of the system, the amount of uncertainty) is characterized by

$$S(p) = -\int_{m_c}^{\omega} p(m)\ln\big(p(m)\big)dm$$

where ***p(m)*** **is the probability density function of magnitudes**. ***S*** **is the "information theory entropy"**, and we look for the distribution of *p* which maximizes *S* subject to the constrains

$$\int_{m_c}^{\omega} p(m)dm = 1$$

$$\int_{m_c}^{\omega} m\,p(m)dm = \langle m\rangle$$

$$\int_{m_c}^{\omega} M_0(m)\,p(m)dm = \langle M_0\rangle$$

APPLICATION TO SEISMICITY

Main & Burton (1984)

By applying the method of Lagrange multipliers we obtain

$$p(m) = \exp\{-\lambda_1 m - \lambda_2 M_0(m)\}/Z$$

where Z is the normalizing integral (the partition function)

$$Z = \int_{m_c}^{\omega} \exp\{-\lambda_1 m - \lambda_2 M_0(m)\}\, dm$$

It can be shown that

$$\langle m \rangle = -d\{\ln(Z)\}/d\lambda_1$$

$$\langle M_0 \rangle = -d\{\ln(Z)\}/d\lambda_2$$

The **cumulative form of the probability distribution** is defined by

$$P(x \geq m) = \int_{m_c}^{\omega} p(x)\, dx = N(x \geq m)/N_T$$

where *N* is the cumulative frequency distribution and N_T is the total number of events in the catalog above m_c per unit time.

APPLICATION TO SEISMICITY

Main & Burton (1984)

The **number density** (the non-cumulative distribution)

$$n(m) = -dN(x \geq m)/dm$$

is given by the **incremental probability**

$$n(m)dm = C\exp\{-\lambda_1 m - \lambda_2 M_0(m)\}dm$$

where $C = N_T/Z$

APPLICATION TO SEISMICITY

Main & Burton (1984)

Interpretation of

$$n(m)dm = C\exp\{-\lambda_1 m - \lambda_2 M_0(m)\}dm$$

The Gutenberg-Richter law constitutes a bridge that links earthquake occurrence to statistical mechanics.

In this framework, let's see that the above equation is equivalent to the Boltzmann distribution as defined in statistical mechanics.

STATISTICAL MECHANICS FRAMEWORK (MICROSCOPIC POINT OF VIEW)

STATISTICAL MECHANICS

Microstate: describes a specific detailed microscopic configuration of a system, that the system visits in the course of its thermal fluctuations. There could be several microstates that have the same properties, as for example the energy. The energy level is then said to be **degenerate**.

Macrostate of a system: refers to its macroscopic properties (such as its temperature and pressure). **A macrostate is characterized by a probability distribution on a certain ensemble of microstates**.

This distribution describes the probability of finding the system in a certain microstate as it is subject to thermal fluctuations.

EQUILIBRIUM ENTROPY

Boltzmann postulated that, for an isolated system in equilibrium, the entropy **S** is related to the number **Ω** of equiprobable microstates through

$$S = k \ln \Omega$$

k is a unit conversion factor with units of [energy]/[termperature]. The entropy would be unitless except for the fact that we measure temperature and energy with different scales.

STATISTICAL MECHANICS

NON-EQUILIBRIUM ENTROPY

The entropy for M equally-likely states is $S(M) = k \log M$. In this case, the probability of each state is $p_i = 1/M$. If we write

$$S(M) = -k \log(1/M) = -k \langle \log(p_i) \rangle$$

we get an appealing generalization for the counting entropy for cases where p_i is not constant:

$$S = -k \langle \log p_i \rangle = -k \sum_i p_i \log p_i .$$

This is the correct generalization of entropy to systems out of equilibrium.

STATISTICAL MECHANICS

When a system is in thermodynamic equilibrium at a temperature T, the probability $P(E_i)$ that a system will occupy a state with energy E_i is

$$P(E_i) = \frac{1}{q} e^{-E_i/kT} \quad , \quad q = \sum_i e^{-E_i/kT}$$

where q, the **partition function** (the total number of microstates), is a normalization constant, introduced so that

$$\sum_i P(E_i) = 1.$$

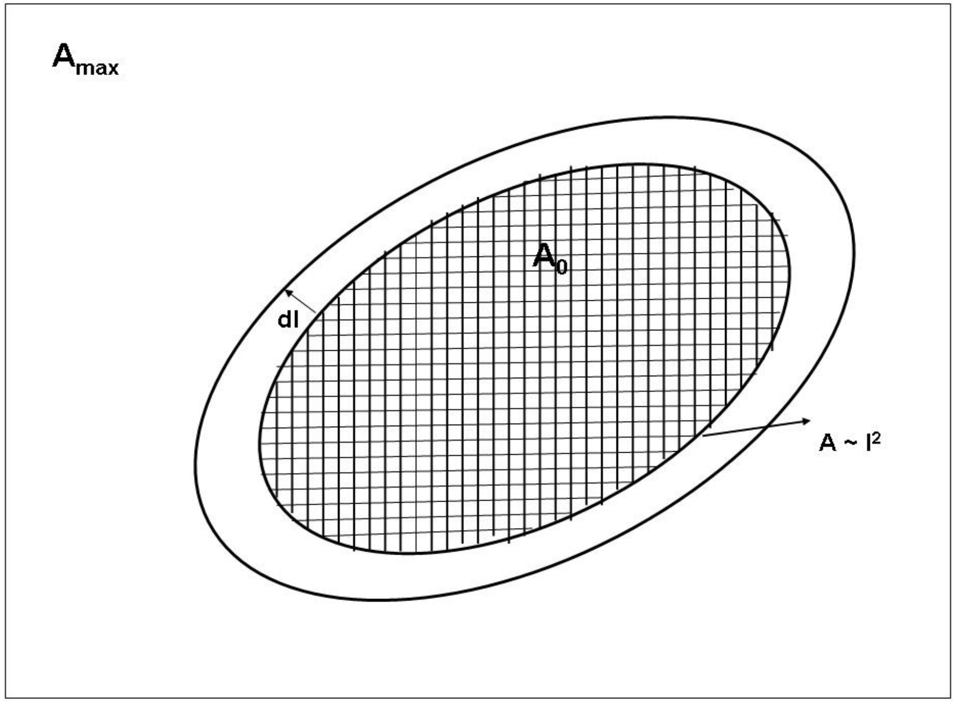
In the presence of degeneracy, if $g(E_i)$ is the number of states having the same energy E_i, then the probability that a system has an energy E_i occupying any other of the $g(E_i)$ states is

$$P(E_i) = \frac{1}{q} g(E_i) e^{-E_i/kT} \quad , \quad q = \sum_i g(E_i) e^{-E_i/kT}$$

APPLICATION TO SEISMICITY

Main & Burton (1984)

Consider a physical model of a fault and apply the method of statistical mechanics to its localized elements. These elements may be as small as the lattice constant of the predominant crystal or may be related to inhomogeneities such as joints or bedding planes. Assume that the elements A_0 (the **microstates**) are small enough to warrant a continuous approach.



Consider an arbitrary area $A \sim l^2$ which ruptures during an event on the fault plane A_{max}. Assuming a constant strain drop (so that the model is self similar), we may take the fault movement $s \sim l$. Taking into account that

$$M_0 = \mu A s = \left(\mu / \bar{\sigma}\right) \Delta W$$

$$M_0 \propto l^3$$

APPLICATION TO SEISMICITY

Main & Burton (1984)

Let E_r be an energy level that can be filled in g ways. The discrete frequency F of state transition is

$$F_r = g_r \exp\left\{-\beta'\left(E_r - E_{r'}\right)\right\}$$

where $\Delta W = \left(E_r - E_{r'}\right)$: change in strain energy $\propto$ M_0

and β' depends on $\langle E \rangle$

degeneracy g(l): On a planar fault, **take** $g(l) = A_{\max}/A(l)$, so that in the continuous case the density D(l) of degenerate states is

$$D(l)dl = g(l) - g(l+sdl) = \left(2A_{\max}/l^3\right)dl$$

The continuous number density *n(l)* is

$$n(l)dl = D(l)dl \exp\left(-\beta M_0(l)\right), \quad \beta M_0 = \beta'\Delta W$$

APPLICATION TO SEISMICITY

Main & Burton (1984)

Taking into account that

$$M_0 \propto l^3,\ m = (\log M_0 - A)/B,\ B = \beta/\ln 10,\ b = (2/3)B,\ D(l) \propto l^{-\nu}$$

the above expression $n(l)dl = D(l)dl\exp(-\beta M_0(l))$ is identical to

$$n(m)dm = C\exp\{-\lambda_1 m - \lambda_2 M_0(m)\}dm$$

In terms of the seismic moment M_0 and defining $\lambda_1 = b \ln(10)$ and $\lambda_2 = \beta$

$$n(M_0)dM_0 = const\, M_0^{-5/3} e^{-\beta M_0} dM_0$$

The form of this distribution can be interpreted as a Boltzmann distribution of energy via $\exp(-\beta M_0)$, multiplied by a geometric factor $\boldsymbol{M_0^{-5/3}}$, which results in another exponential in terms of the magnitude. The geometric term corresponds to the G-R law, and the Boltzmann term to the roll-off often observed at higher magnitudes/seismic moments. → **gamma distribution**

ENERGY

Main & Al-Kindy (2002)

In the framework of statistical mechanics, Boltzmann energy distribution reads

$$p(E)dE = \frac{g_E \exp(-E/\theta)}{Z} dE \; , \quad \theta = kT$$

where θ, known as ***temperature factor,*** has dimensions of energy. k is the Boltzmann constant and T is the absolute temperature.

In earthquake population the degeneracy occurs because energy is related to the surface rupture area ***A*** by ***E ~ A***$^{3/2}$. For earthquakes, the density distribution of the degeneracy has the power law form

$$g_E = E^{-B-1} \Big/ E_0^{-B}$$

where E_0 is a scaling constant.This equation corresponds to the Gutenberg-Richter law for the equivalent magnitude where typically the exponent $B \approx 2/3$.

ENERGY

By combining the above equations the incremental probability is found to be

$$p(E)dE = \alpha E^{-B-1}\exp(-E/\theta)dE, \ \alpha = E_0^B Z^{-1}$$

and the partition function

$$Z = \int_{E_{\min}}^{E_{\max}} \left(\frac{E}{E_0}\right)^{-B} \exp\left(-\frac{E}{\theta}\right) d\ln E$$

THE "TECTONIC" OR "TERM" TEMPERATURE

As already seen, the Boltzmann distribution of energy is given by

$$p(E)dE = \alpha E^{-B-1} \exp(-E/\theta) dE, \; \theta = kT \qquad (1)$$

On the other hand, a good approximation for the PDF for the observed seismic moment distribution of earthquakes is provided by the tapered G-R distribution (Kagan, 1991))

$$\Phi(M) = (M_t/M)^{\beta} \exp\left(\frac{M_t - M}{M_c}\right), \; M_t \le M \le \infty \qquad (2)$$

As the seismic moment is proportional to the energy $\left(E_s = M_0 \cdot 10^{-4.8}\right)$, by comparing (1) and (2) a **"tectonic" temperature *θ*** can be defined, with ***θ* proportional to the maximum event size**, *i.e.*, to the (maximum) released seismic energy. *k* has the same dimensions but is not equivalent to Boltzmann's constant, since it applies to a macroscopic system.

SEISMICITY - ENERGY AND ENTROPY

Main & Al-Kindy (2002)

The expectation value of the system ***energy*** is the first moment

$$\langle E \rangle = \int_{E_{\min}}^{E_{\max}} E\, p(E)\, dE$$

And the ***entropy*** is defined by

$$S = -\int_{E_{\min}}^{E_{\max}} \ln p(E)\, p(E)\, dE$$

The G-R law is recovered as θ → ∞, preserving a finite energy through a finite E_{max}.

ENERGY AND ENTROPY

Main & Al-Kindy (2002)

From the above equations and for a system in a near-critical state it can be shown

$$S = S_0 + B\langle \ln E \rangle, \; S_0 = ct.$$

In a subcritical state with finite positive *θ* we would expect the local slope to be greater than B, and in a supercritical state we would expect a local slope less than *B*. Corollaries of the above equation for finite, positive θ are

$$\langle E \rangle \propto \theta^{1-B}, \quad \partial S / \partial \langle E \rangle = \varepsilon \text{ (a small positive number)}$$

For a system in a near critical state, with significant fluctuations in *θ* at a constant B-value, we would expect:

1) A strong positive correlation between <*E*> and *θ*
2) A week positive correlation between *S* and <*E*>.
3) A strong positive correlation between *S* and <*lnE*> of slope *B*

ENERGY AND ENTROPY

Main & Al-Kindy (2002)

METHODOLOGY:

Start from a **seismic catalog**.

The scalar moment for the catalog is converted to seismic energy using standard relationships.

Calculate the expectation value <*E*> from the individual summed energies.

Calculate *S* using a discrete version of the definition $S = -\int_{E_{\min}}^{E_{\max}} \ln p(E)\, p(E)\, dE$

after appropriate normalization (to ensure unit total probability).

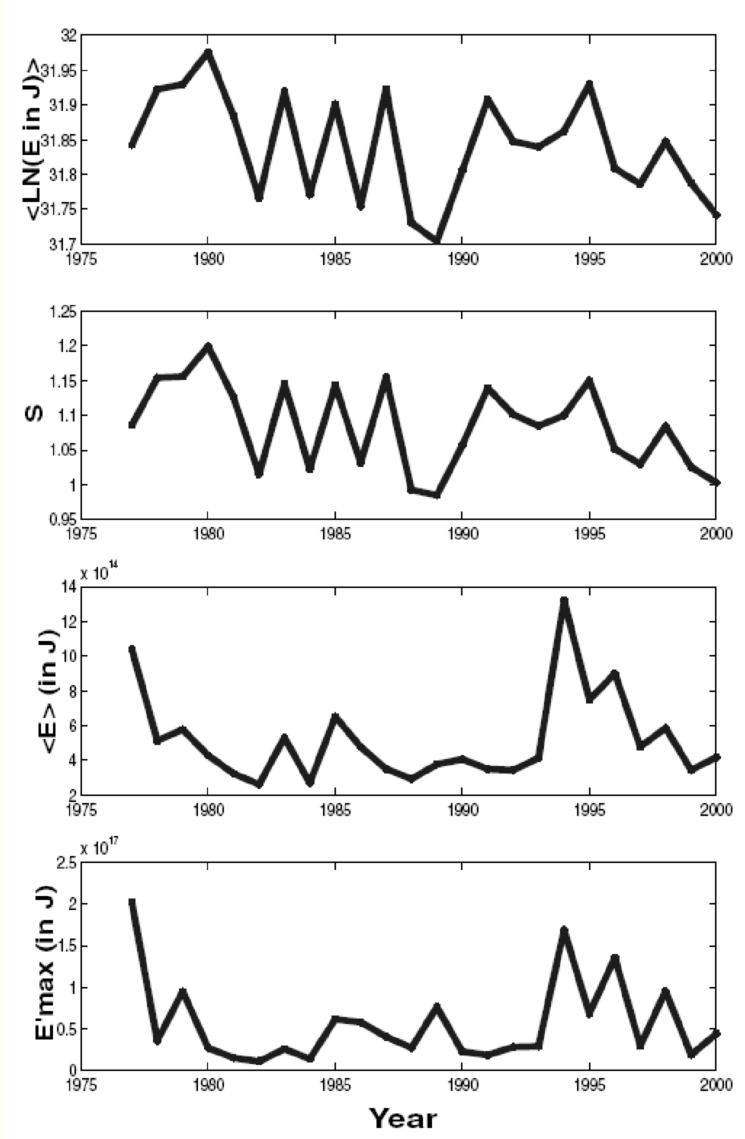
The results confirm that global seismicity is in near critical state, with large fluctuations in mean energy occurring due to small changes in entropy.

The proximity to the critical point indicates that the predictability of the system may be finite, but low.

Temporal variation in <lnE>, entropy S, mean annual energy E'_{max} and maximum energy from the global occurrence of earthquakes from the CMT catalogue for the time period 1977– 2000. Energy units are Joules. Criteria 1) – 3) are preserved:

A strong positive correlation between *S* and *<lnE>* of slope *B*

$$S = S_0 + B\langle \ln E \rangle$$



A week positive correlation between *S* and <*E*>.

$$\partial S / \partial \langle E \rangle = \varepsilon \ll 1$$

$$(\varepsilon \propto 1/\theta)$$

A strong positive correlation between <*E*> and θ (~ E_{max})

$$\langle E \rangle \propto \theta^{1-B}$$

ENERGY AND ENTROPY

Main & Al-Kindy (2003)

Taking into account that

$$p(E)dE = \frac{g_E \exp(-E/\theta)}{Z} dE \,, \qquad g_E = E^{-B-1} / E_0^{-B}$$

It can further be shown that

$$S = \ln Z + B\langle \ln(E/E_0) \rangle + \langle E \rangle / \theta$$

from which

$$\frac{1}{\theta} \sim \frac{\partial S}{\partial \langle E \rangle}$$

This relation demonstrates formally that θ is an equivalent temperature term for the system.

THERMODYNAMIC FRAMEWORK (MACROSCOPIC POINT OF VIEW)

ENERGY AND ENTROPY

There is an important caveat to the application of thermodynamics to earthquake statistics.

The thermodynamic formulation is strictly based on an internal energy $U = \langle E \rangle$,

But it is not possible to determine U for the earthquake problem: we do not have independent information on the strain energy distribution in the Earth because we cannot independently measure the stress.

Therefore applications of thermodynamics to Earthquake systems assume, explicitly or implicitly, that the distribution of radiated energy is related to that of internal energy.

This implies that large earthquakes are more likely to occur in time periods when the internal strain energy is also high.

STATISTICAL MECHANICS AND THERMODYNAMICS

STATISTICAL MECHANICS: Let $\Omega = \Omega(E)$ be the number of microstates comprising a macrostate, a statistical weight for nonequilibrium macrostates.

The equilibrium of an isolated system is defined (Boltzmann) in terms of the **entropy *S***, **a measure of the disorder of a system in a given macrostate**, as

$$S_{\text{equil}}(E) = k \ln \Omega(E)$$

As *S* is a function of the energy *E*, a definition of the absolute temperature *T* is

$$\left(\frac{\partial S}{\partial E}\right) = \frac{1}{T}$$

THERMODYNAMICS: Entropy of a system is defined as a summation of "heat supplied" divided by its "temperature" [*Clausius*, 1865]. If a certain small amount of heat *Q* is supplied quasi-statically to a system with an absolute temperature of *T*, then the entropy of the system will increase by

$$dS = \frac{\delta Q}{T}, \ \Delta S = \int \frac{dQ}{T}$$

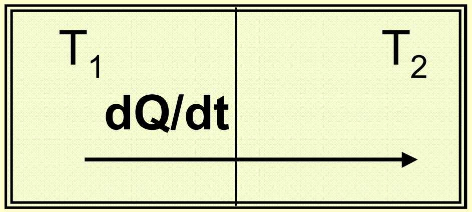
THERMODYNAMIC ENTROPY

One could choose instead a rescaled entropy in microscopic terms such that

$$S' = \ln \Omega, \;\; \Delta S' = \int \frac{dQ}{kT}.$$

This is a rather more natural form because this rescaled entropy exactly corresponds to Shanon’s information entropy.

The rate of entropy production in a system not in equilibrium is



$$\Delta S_{\text{thermo}} = \frac{Q}{T} \;\; \rightarrow \;\; \frac{dS}{dt} = \left(\frac{1}{T_2} - \frac{1}{T_1} \right) \frac{dQ}{dt} > 0,$$

where the inequality sign expresses Clausius principle of increase of entropy.

MAXIMUM ENTROPY PRODUCTION

The probability distribution which maximizes the information entropy is the true probability distribution, with respect to the testable information prescribed.

The myriad of processes that transform energy, that result in the motion of mass in the atmosphere, in oceans, and on land, processes that drive the global water, carbon, and other biogeochemical cycles, all have in common that they are irreversible in their nature.

Entropy production is a general consequence of these processes and measures their degree of irreversibility. The **proposed principle** of maximum entropy production (MEP) states that **systems are driven to steady states in which they produce entropy at the maximum possible rate given the prevailing constraints**.

THE LAW OF MAXIMUM ENTROPY PRODUCTION
A system will select the path or assemblage of paths out of available paths that maximizes the entropy at the fastest rate given the constraints

MAXIMUM ENTROPY PRODUCTION

Whereas the Second Law says that the world acts to minimize potentials, it does not say which out available paths it will take to do this. This is the question the Law of Maximum Entropy Production answers.

(Note: The Law of Maximum Entropy Production (LMEP) does not contradict or replace the second law. It is another law that is in addition to it).

Swenson (2000) pointed out that the answer to this question, as above, was that it will "select the path or assemblage of paths out of available paths that minimizes the potential or maximizes the entropy at the fastest rate given the constraints".

MAXIMUM ENTROPY PODUCTION (natural systems)

Main & Naylor (2008)

The goal is the quantification of the entropy production of earthquake occurrence from the energy budget to test whether natural seismicity is consistent with a state of maximum entropy production.

Our observable is the magnitude (preferable M_W) as reported in the seismic catalog, for the interval of magnitudes for which the catalog is complete, and then converted to seismic energy using standard relationships, as for example that of Kanamori (1977): the minimum estimate of the strain energy drop, known as the seismic wave energy E_S, can be estimated as

$$E_S = 1.5 M_W + 11.8, \quad M_W = \frac{2}{3}\log\left(M_0\right) - 10.7$$

The radiated energy E_S is a finite fraction of the total energy change ΔQ during an earthquake, $E_S = \eta \Delta Q$, where the seismic efficiency $0 < \eta < 1$, and is related to the ratio of stress drop to mean stress by

$$\eta = 0.5 \Delta\sigma / \langle\sigma\rangle .$$

MAXIMUM ENTROPY PODUCTION (natural systems)

Main & Naylor (2008)

In practice neither ΔQ –the difference between the initial and final strain energy- nor the temperature term can at the present be estimated. We have thus to relay on **numerical simulations of models of seismicity**.

To investigate the temperature term and other aspects of entropy production a model system can be used, as for example the Olami-Feder-Cristensen (OFC) multiple spring-block slider model. The OFC model is very useful because it is the simplest numerical model for earthquake dynamics that reproduces the observed G-R distribution of small and intermediate–magnitude earthquakes as an emergent property.

As well, all relevant parameters for the calculation of the Maximum Energy Production can be calculated analytically.

MAXIMUM ENTROPY PODUCTION (natural systems)

Main & Naylor (2008)

Main and Naylor conclude that:

- Many observations in natural and model seismicity are consistent with the hypothesis of maximum entropy production at steady state, including complexity, broad-band scale invariance, the occurrence of spatially characteristic Earthquakes, and low but finite seismic efficiency and stress drop.
- When implemented in a numerical model entropy production is maximized in a state of self-organized sub-criticality, with b ≈ 1, also consistent with observation.
- The results are consistent with entropy production as a thermodynamic driver for domain formation and self-organized (sub) criticality in natural and model seismicity.

MAXIMUM ENTROPY PRODUCTION (natural systems)

Observations of natural phenomena (climatology, oceanic circulation, earthquake occurrence) agree with the principle of Maximum Entropy production. And? Which insight do we gain?

Maximum entropy production appears as a consequence of the general form of statistical inference, valid for non equilibrium systems and for irreversible processes (in some way, the equivalent of the arrow of time).

However, MEP is not the only consequence of the statistical inference. Other two properties, apparently disconnected, can also be derived:

- The **fluctuation theorem** (established for a variety of non-equilibrium systems)
- The emergence of the paradigm of **self-organized criticality**.

PRECURSORY ACTRIVITY: A SELF ORGANIZING PROCESS TROUGH DIFFERENT STRUCTURES.

EACH NEW STRUCTURE WOULD BE CHARACTERIZED BY A NET VARIATION OF ENTROPY, DEPENDING ON IF MORE OR LESS ORDERED.

WHICH IS THE INFORMATION CONTRENTS OF EACH STRUCTURE?

COULD IT BE RETRIEVEDS FROM THE PRECURSORY ACTIVITY (i.e. FROM THE SEISMIC CATALOG?)

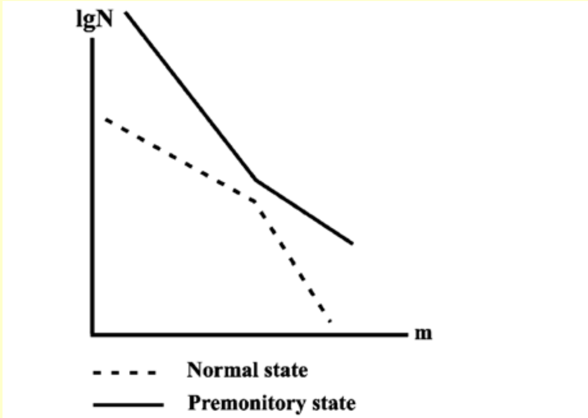


Figure 10 Premonitory changes of the Gutenberg-Richter relation. N is the occurrence rate of earthquakes, m is the magnitude (logarithmic measure of energy). The dashed line represents time far from a strong earthquake; the solid line represents time close to it.

EXAMPLE: the change of the G-R law implies a change in the PDF, and thus of the entropy.

$$S = -\int_{E_{\min}}^{E_{\max}} \ln p(E)\, p(E)\, dE$$

BASIC TYPES OF PREMONITORY PHENOMENA

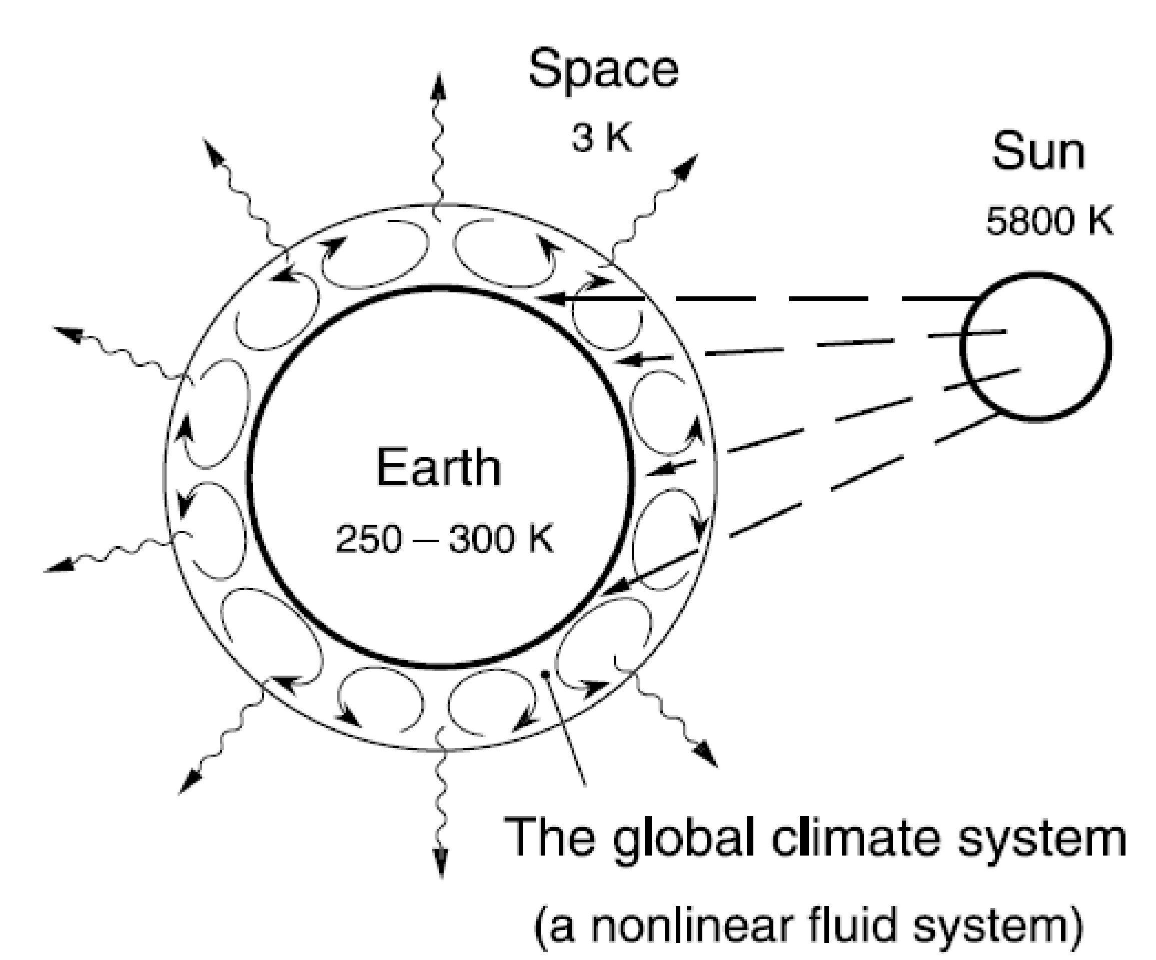
The approach of a strong earthquake is indicated by (some of) the following

changes in the basic characteristics of seismicity:

a. Rise of seismic activity.
b. Rise of irregularity in space and time.
c. Reversal of territorial distribution of seismicity.
d. Transformation of magnitude distribution.
e. Rise of earthquake clustering in space and time.
f. Rise of the earthquake correlation range.
g. Accelerated stress-release

MEP IN THE CLIMATE SYSTEM

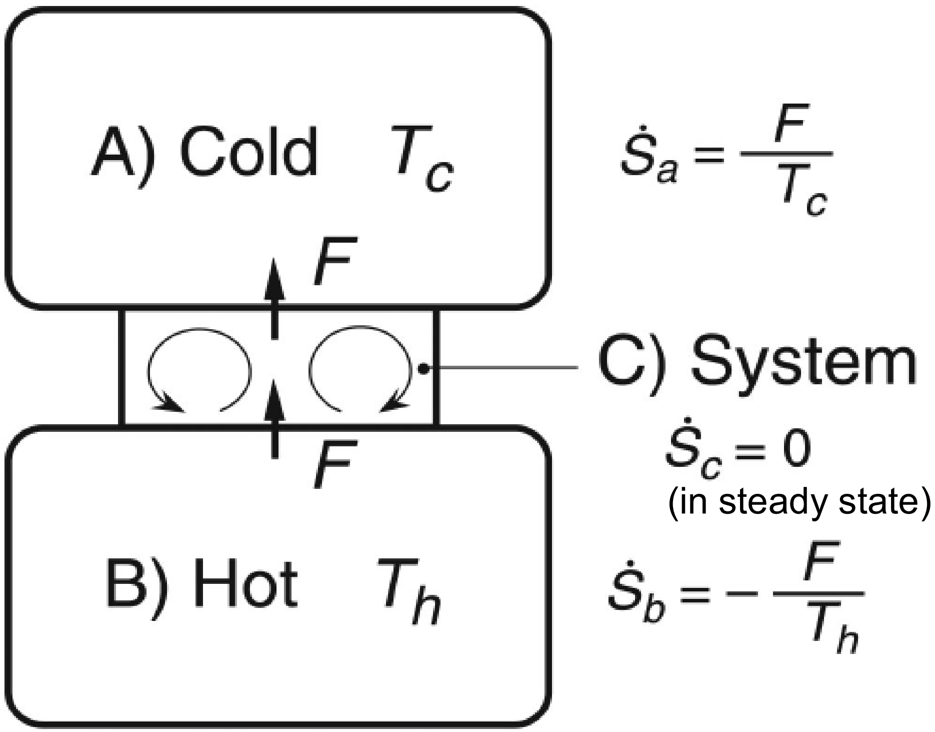
Ozawa et al. (2003)



A schematic of energy transport processes in the planetary system of the Earth, the Sun, and space. The Earth receives the shortwave radiation from the hot Sun and emits longwave radiation into space. The atmosphere and oceans act as a fluid system that transports heat from the hot region to cold regions via general circulation.

MEP IN THE CLIMATE SYSTEM

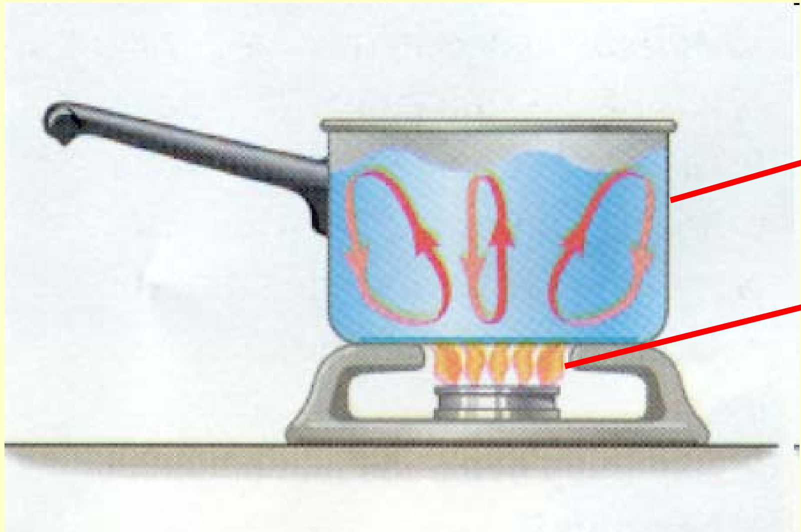
Ozawa et al. (2003)



$$\dot{S}_{\text{whole}} = \dot{S}_a + \dot{S}_b + \dot{S}_c = \frac{T_h - T_c}{T_h T_c} F \geq 0$$

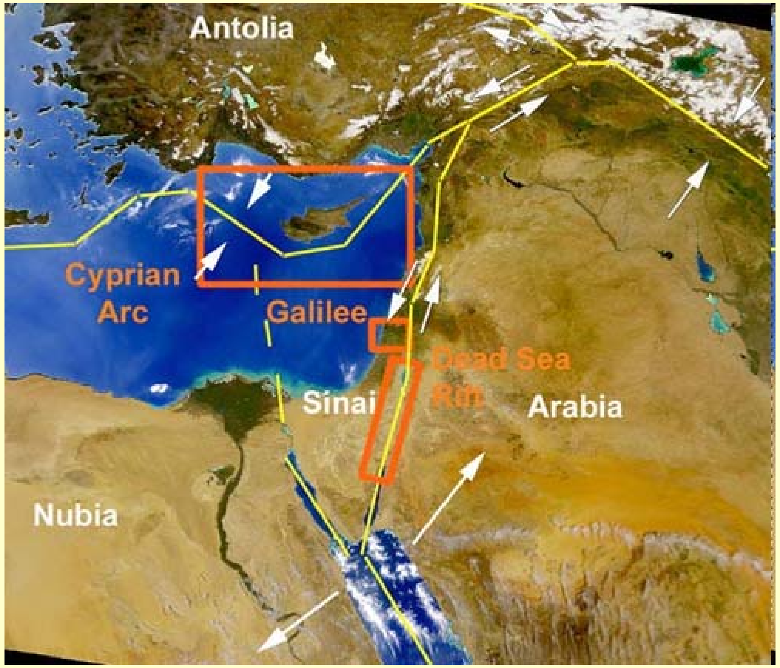
$$= \text{maximum} \quad \text{(for a nonlinear system)}$$

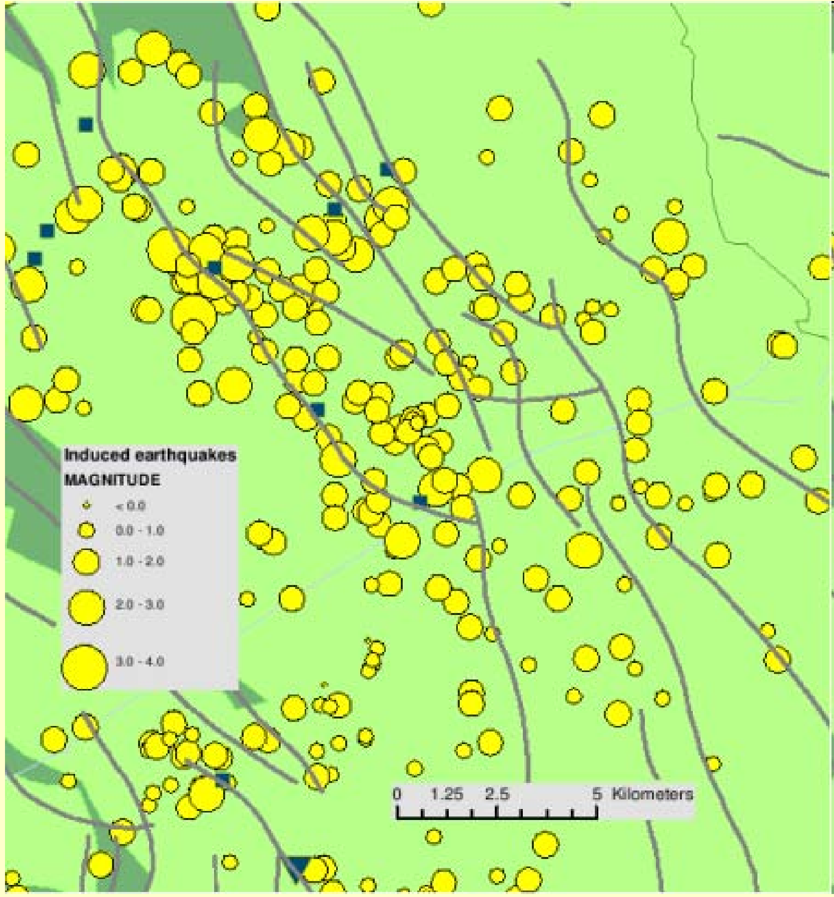
A schematic of heat transport through a small system (C) between two thermal reservoirs with different temperatures (A, cold and B, hot). By the heat transport from hot to cold, entropy of the whole system increases. In the case of a fluid system in a supercritical condition, the rate of entropy production tends to be a maximum among all possible states.



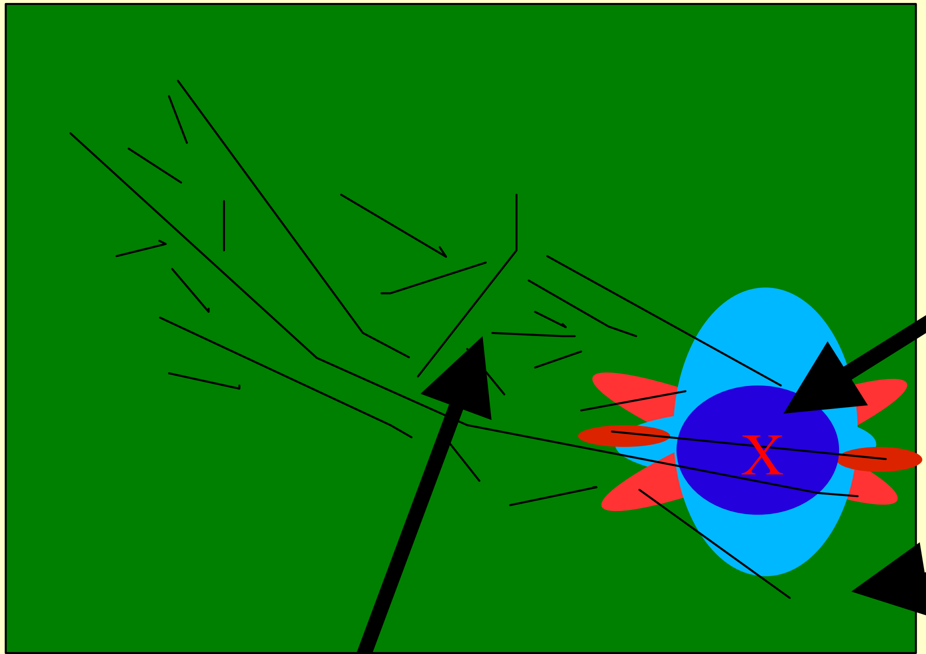
Patterns of seismic activity

Strain energy accumulation due to plate motions





DYNAMICS OF FAULT SYSTEM



Complicated interactions between faults due to stress transfer during Eqs

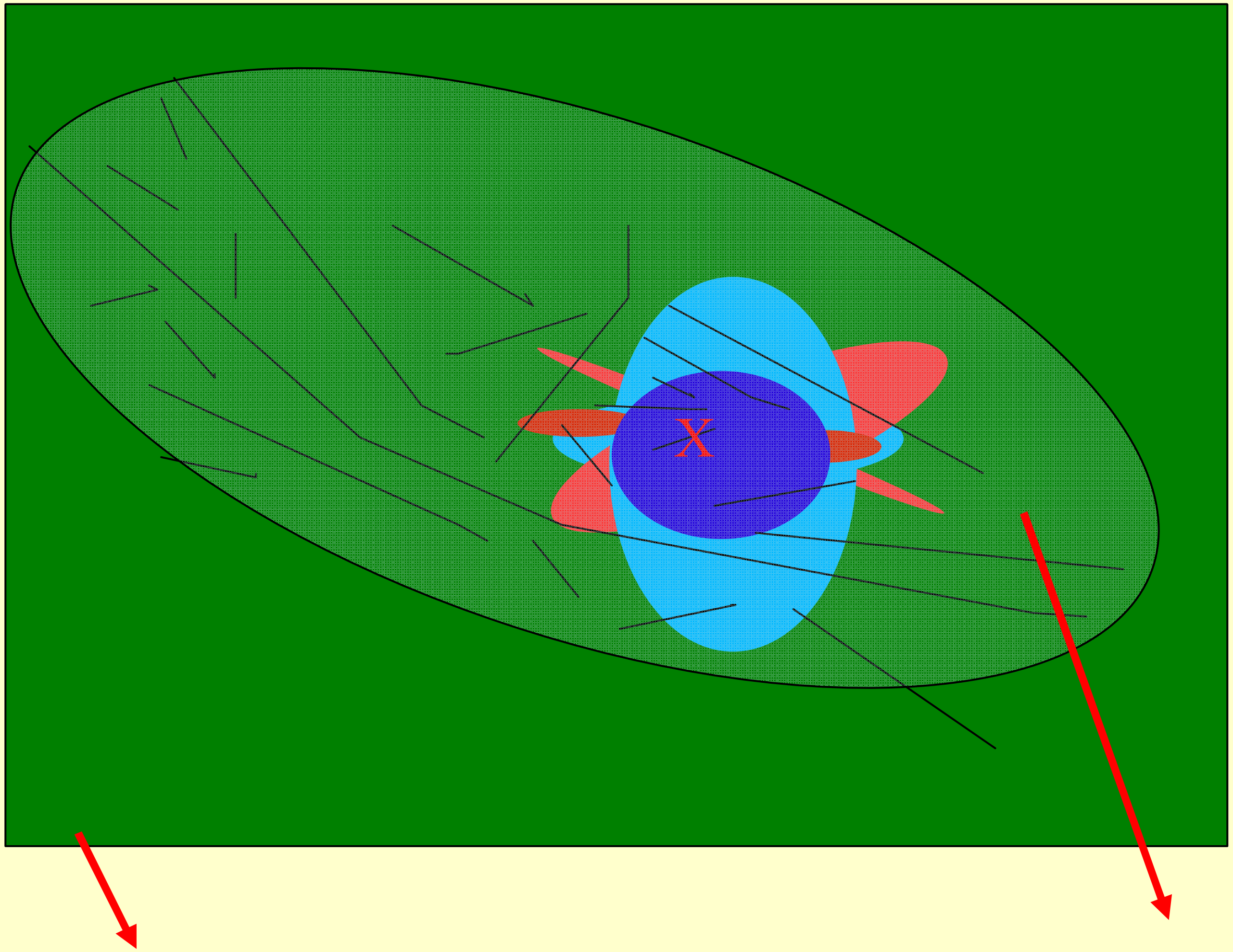
Nonlinear Rheology

Multi-Fractal fault heirarchy

Earthquake Fault systems are COMPLEX:

- Many degrees of freedom
- Strongly coupled spatial and temporal scales
- Nonlinear dynamical equations & constitutive laws
- Multi-physics: mechanical, chemical, thermal, fluids, (EM?)

D. Weatherley
QUAKES & AccESS
3rd ACES Working Group Meeting
Brisbane, Aust. 5th June, 2003.



Surrounding system (environment)

“Turbulent” system

TO BE CONTINUED ...

Study the possibility of entropy to quantify precursory activity:

1. Macroscopic (global) evolution of them system from the point of view of the energy and/or the temporal evolution of the PDF. Apply the MEP Law for the evolution of the system at the fault system.
2. Study of the entropy contents of the different precursors (ordered structures). → emergent structures, not predicted by traditional entropy considerations.

ORDER OUT OF DISORDER

Entropy is usually related to **near-equilibrium** states, and because of the second law, an ordered structure cannot be generated from a disordered one.

The lithosphere, like Rayleigh-Benard convection, constitutes a **far-from-equilibrium system**, driven by external forces (the relative motion of the plates).

For **far from equilibrium systems**, continuously sustained by a continuous input of stress, the environmental increasing stress can cause breakdowns and jumps in behavior: the system will explore all possible ways to reduce the conflict. In a way compatible with the 2nd law, the gradients encourage the system to self-organize to an ***ordered*** state since this actually increases the rate of entropy production and thus stress reduction.

The greater the energy flows, the greater the order (and information) generated becomes. This order requires the system to be **dissipative**, that means that energy must be expended (wasted) to create the visible order or information (emergent structures) from the disorder. The energy used to create the order is that accumulated in the lithosphere.

ENTROPY

In physics, the second law of **thermodynamics** is formulated within a strict context of processes that result in energy changes. The fundamental physical measures associated with the 2nd law are temperature ***T***, head ***Q*** and (thermodynamic) entropy ***S***, related by

$$dS = dQ/T$$

Statistical mechanics identifies this macroscopic measure with the number ***Ω*** of microscopically defined states accessible to the system by the relation

$$S \equiv k \ln \Omega$$

Thus defined, thermodynamic entropy has strong formal similarities to information entropy

$$H(p) = -\sum_i p_i \log p_i$$

where ***i*** ranges over the possible states of the system and p_i is the probability of finding the system in state ***i***. These formal similarities leads to the notion of "entropy" as a measure of macro-level disorder.

ENTROPY VARIATION AND INFORMATION IN SELF-ORGANIZATION PROCESSES

Clausius entropy refers to isolated systems exchanging neither energy nor matter with the environment. According to the second law

$$dS/dt \geq 0$$

For non isolated systems we must distinguish two terms in the entropy change ***dS***: ***d***d_eS, the transfer of entropy across the boundaries, and d_iS, the entropy produced within the system by irreversible processes (the only ones that produce entropy):

$$dS = d_iS + d_eS, \quad d_iS \geq 0, \quad dS \geq 0$$

Open systems could conceivably evolve to non-equilibrium Steady State with

$$d_eS \leq -d_iS, \quad dS \geq 0$$

Entropy increase at the micro level is sufficient to ensure entropy increase in the overall system even in the presence of self-organization.

Order may be created from disorder and equilibrium is no longer the only attractor of the system.

REFERENCES

- Bak, P., Chistensen, K., Danon, L. and Scanlpon, T. (2002). Unified scaling law for earthquakes, *Phys. Rev. Lett*., 88, 178501.
- Ben-Zion, Y. and Rice, J.R. (1993) *J. Geophys. Res*., 98, 14,109.
- Corral, A. (2004). Long-term clustering , scaling, and universality in the temporal occurrence of earthquakes, *Phys. Rev. Lett*., 92, 108501.
- Corral, A. (2007). Statistical properties of seismic catalogs. http://einstein.uab.es/acorralc/
- Correig, A.M., Urquizu, M., Vila, J. and Manrubia, S. (1997). The aftershock series of event 18 February 1996. An interpretation in terms of Self Organized Criticality, *J. Geophys. Res.*, 102, 27,407-27,420.
- Dutta & Horn (1981). Low-frequency fluctuations: 1/f noise, *Rev. Mod. Phys.,* 53, 497.
- Helmstetter, A., Sornette, D. and Grasso, J.R. (2003). Mainshocks are aftershocks of conditional foreshocks: How do foreshocks statistical properties emerge from aftershock laws, *J. Geophys. Res*., 108, 2046.
- Jaines, E.T. (1957). Information theory and statistical mechanics, Phys. Rev., 106, 620-630.

REFERENCES

- Kagan, Y.Y. (1991). Seismic moment distribution, Geophys. J. Int., 106, 123-134.
- Kagan, Y.Y (1993). Statistics of characteristic earthquakes, *Bull.Seism.Soc. Am*., 83, 7-24.
- Kanamori, H. (1977). The energy release in great earthquakes, *J. Geophys. Res*., 82, 2981-2987.
- Main, I. (2000). Apparent breaks in scaling in the earthquake cumulative frequency-magnitude distribution: fact or artifac?, *Bull. Seism. Soc. Am*., 90, 86-97.
- Main, I.G. and Al-Kindy, F.H. (2002). Entropy, energy and proximity to criticality in global earthquake population, *Geophys. Res.Lett*., 29, 25.
- Main, I.G. and Burton, P.W. (1984). Information theory and the earthquake frequency-magnitude distribution, *Bull. Seism. Soc. Am*., 74, 1409-1426.
- Main, I.G., Leonard, T., Papasoulitotis, O., Hatton, C.G. and Meredith, P.G. (1999). One slope or two? Detecting statistically significant breaks of slope in geophysical datq, with application to fracture scaling relationships, Geophys. Res. Lett., 26, 2801-2804.

REFERENCES

- Main, I.G., O'Brien, G. and Henderson, J.R. (2000). Statistical physics of earthquakes: comparison of disteibution exponents for source areas and potential energy and the dynamic emergence of log-periodic energy quanta, *J. Geophys. Res.,* 105, 6105-6126.
- Omori (1894) "On the aftershocks of earthquakes," *Journal of the College of Science, Imperial University of Tokyo*, vol. 7, pages 111–200.
- Ozawa, H., Ohmur, Lorenz, R.D. and Pujol, T. (2003). The second law of thermodynamics and the global climate system: a review of the maximum entropy production principle, *Rev. Geophys*., 41, 1018.
- Rundle, J.B. (1993). Magnitude-frequency relations for earthquakes using a statistical mechanics approach, *J. Geophys. Res*., 98, 21,943-21,949.
- Swenson, R. (2000). Spontaneous Order, Autocatakinetic Closure, and the Development of Space-Time. *Annals New York Academy of Sciences*, vol. 901, pp. 311-319, 2000.
- Utsu, T., Ogata, Y. ,and Matsu'ura, R.S. (1995) "The centenary of the Omori formula for a decay law of aftershock activity," *Journal of Physics of the Earth*, vol. 43, pages 1–33.