united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

H4.SMR/1330-14

" Sixth Workshop on Non-Linear Dynamics and Earthquake Prediction"

15 - 27 October 2001

Fractal and Chaotic Properties of Earthquakes and Seismicity Pattern Decomposition

Christian Goltz

University of Kiel
Institute of Geophysics
Kiel, Germany

Reference for most of part I & II:

C. Goltz, Fractal and Chaotic Properties of Earthquakes, Springer, 1997

Ref. for Part III:

C. Goltz, Decomposing spatio-temporal seismicity patterns, Natural Hazards and Earth System Sciences, Vol 1, No. 1, 1 - 10, 2001

Fractal and Chaotic Properties of Earthquakes and Seismicity Pattern Decomposition

Sixth Workshop on
Non-Linear Dynamics
and
Earthquake Prediction
Trieste, October 2001

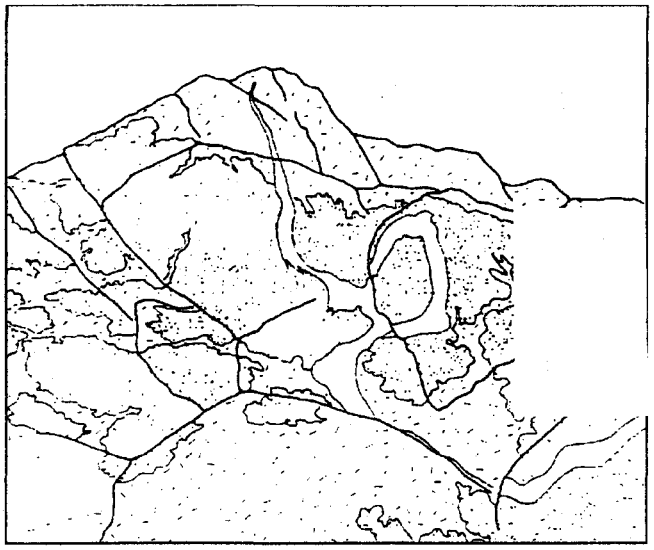
Christian Goltz
Geophysics, Kiel University

- Fractals
  - Definitions
  - Fractal Analysis (numerical)
  - Pitfalls
  - Earthquakes as Fractals
  - Reasons for Fractality
  - (Temporal Variation of Fractal Parameters)
- Chaos
  - Definitions
  - Non-linear Time Series Analysis
  - Analysis of Model Dynamics
  - Analysis of Earthquake Data
- Pattern Decomposition
  - Motivation and Problems
  - A novel approach: PCA
  - Incorporating Time
  - Sample Application
  - Conclusion / Outlook

Fractal Basics

Practical Definition of a fractal:

- statistical self-similarity = scale invariance = no characteristic length



(from Bebien et al., 1987)

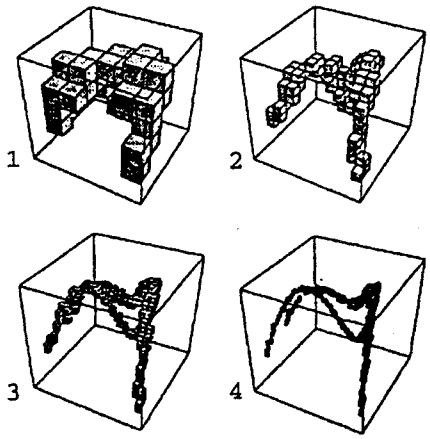
- has fractal (non-integer) dimension:

Euclidean: point = 0, line = 1, plane = 2 etc.
More general (Capacity, Box-Counting Dim.):

$$D_0 = -\lim_{r \to \infty} \frac{\log N(r)}{\log(1/r)}$$

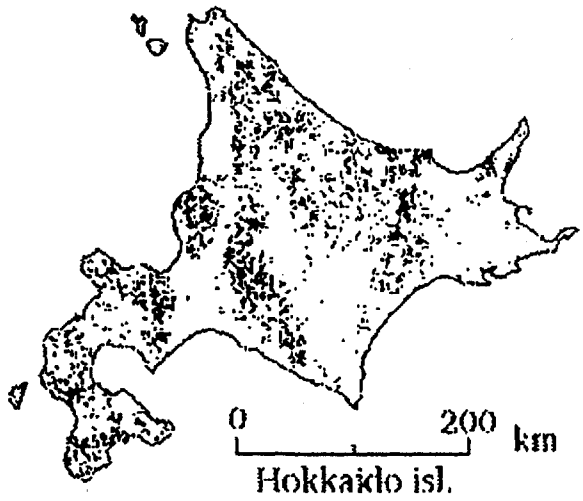
where $N(r)$ is the number of non-empty `boxes' in a covering with boxes of size r.
For e.g. a line: $N(r) \propto r \Rightarrow D_0 = 1$.

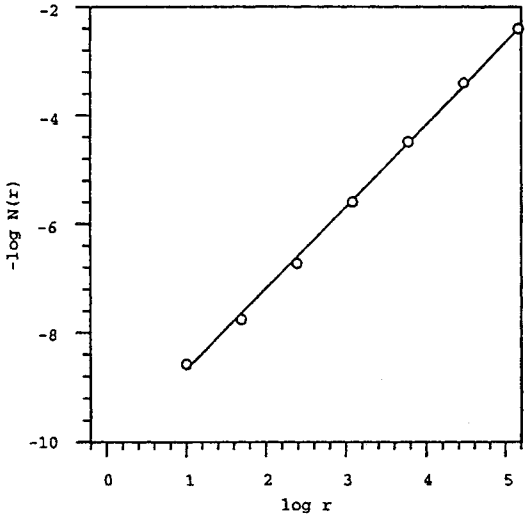
$N(r) \propto (1/r)^{D_0}$, i.e. a power-law relation.



(after Abarbanel et al., 1993)

We may already see: At every resolution, we learn nothing about how many events are contained in every `box'. To learn about clustering (probabilities), more detailed description is needed.





Such an analysis only shows that the events form a fractal distribution (D_0=1.57 ± 0.02). We know nothing about the obvious inhomogeneity of landslides. But:

- Scaling region Δ r = 2.73 - 174.57 km (Important!)
- Only one scaling exponent over Δ r (-"-)

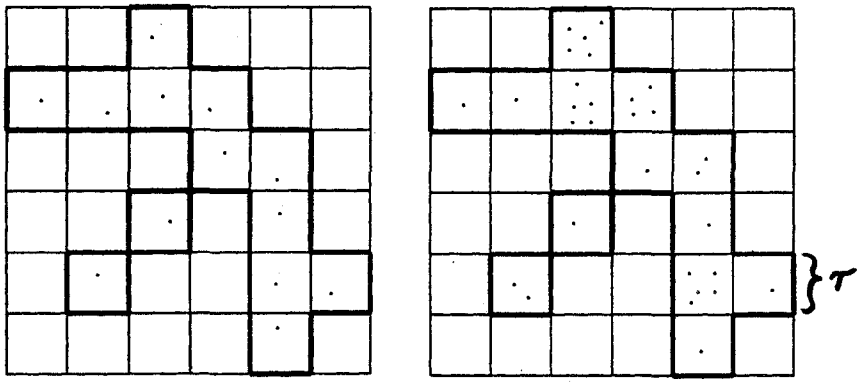
We would not detect anisotropy in D (sect-affinity) like this!

Take into account heterogeneity:

Information Dimension D_1:

Information = $I(E) \equiv - \log P(E) = \log 1/P(E)$

(Note: P = 1 ⇒ I = 0; P = 0 ⇒ I = ∞)



The average information is $I(r) = \sum_{i=1}^{N(r)} -P_i(r) \log P_i(r)$

Thus similar to D_0:

$$D_1 = \lim_{r \to 0} \frac{I(r)}{\log(1/r)}$$

Hence $I \propto (1/r)^{D_1}$ i.e. a power-law relation between information and size of boxes denotes a fractal.

Homogeneous case (P_i = const.): $D_0 = D_1$

D_1 is not very popular, but introduced here due to its relationship with the Configuration Entropy H(r):

Simply, H(r) = - I(r), i.e. $H(r) = - D_1 \log r$!
That is $H(r) = \sum p_i(r) \log p_i(r)$, where p_i is the probability to find i events in a certain cell.
Thus D_1 describes the scaling of information (a measure of the point set), H(r) describes information only at specific scales.
Here, I want to vary *r*, however, to compare the information at different scales. At r^*, H(r) has its maximum H^*. H^* is the max. information (entropy, ~~information~~ disorder) contained in a distribution at all scales.

Correlation Dimension D_2:

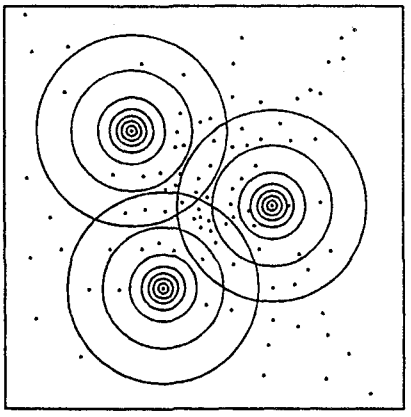
Determined by Grassberger-Procaccia algorithm (sphere-counting), not a fixed grid is used, but the scaling property is determined from the spatial correlation of points by first getting the correlation function

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1,i\neq j}^{N} \Theta(r - \|x_i - x_j\|)$$

where N = no. of (reference) points and
Θ is 0 if argument < 0, 1 otherwise.

Then,

$$D_2 = \lim_{r\to 0} \frac{\log C(r)}{\log r}$$



i.e. when a power-law relation exists between C(r) (no. of points in spheres, mass) and r, the object is fractal with dimension D_2.
Dense regions are weighted even more, thus good for small data sets. Disadvantages: Boundary effects, slow.

Why power laws?

Scale-invariance of f(x) means that

$$f(x) \propto f(\lambda x) \text{ for all } \lambda$$

Therefore, if f(x) is fractal there exists a function $C(\lambda)$ such that

$$f(x) = C(\lambda)\, f(\lambda x)$$

Differentiating w.r.t. x and eliminating $C(\lambda)$

$$f'(x)/f(x) = \lambda\, f'(\lambda x)/f(\lambda x)$$

Substituting x = 1 and integrating w.r.t. λ

$$f(x) = f(1)\, x^{\alpha}\,,\ \alpha = f'(1)/f(1)$$

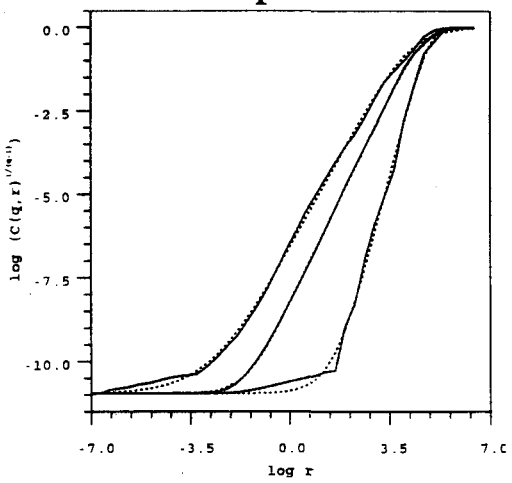
Thus: Power functions are the only scale-invariant functions.

Fractal Anomalies

Possible deviations from the described ideal scaling behaviour are:

- Limited scaling region

  $r \to 0$ and $N \to \infty$ of the previous definitions are not possible with real (finite) data sets. Due to finite resolution, the log-log plot saturates at low r; it also saturates at high r due to the finite number of points:



One must make sure that Δr spans at least ten orders.

- Multiscaling (≠ Multifractal!)

  The *log-log* plot may show piece-wise linear regions instead of only one exponent over Δr. This might indicate that e.g. two underlying mechanisms are involved which scale differently at small and large scales (bi-fractality) or also indicate sampling problems. Such behaviour, however, was not observed here. Note that undetected multiscaling will render results for D meaningless, because e.g. the linear regression will average over the different slopes.

- Self-affinity resp. anisotropy of D

  Profiles and time series are self-affine, i.e. scale differently in vertical and horizontal direction. For topography, the physical reason is gravity which only works in z-direction while erosion and tectonic forces work in all directions.

Multifractals

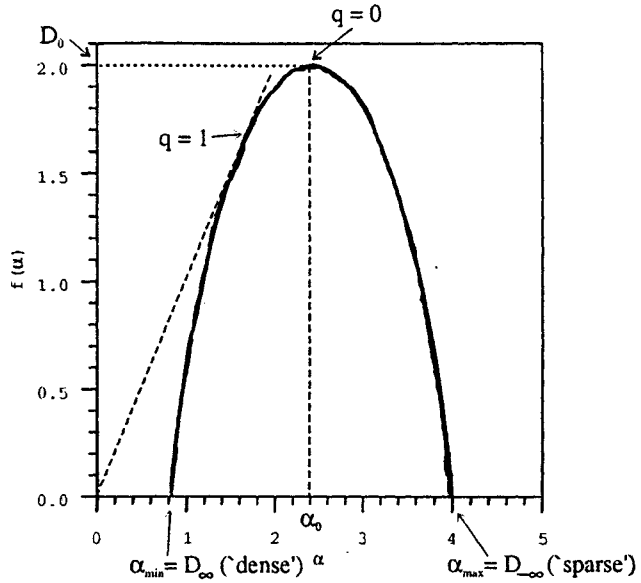
In a sense another deviation, but also a much more advanced concept are multifractals:
The previous fractal dimensions all increasingly emphasise the dense areas within a fractal. If the fractal is heterogeneous, even a comparison between the first three generalised dimensions can tell us only very little about the inhomogeneous structure. In fact, there are infinitely many fractal dimensions of the form:

$$D_q = \frac{1}{q-1} \lim_{r \to 0} \frac{\log\left(\sum_i \{P_i(r)\}^q\right)}{\log r}$$

Note that the qth moments of the probabilities introduced before are used here to emphasise dense areas ($q>0$) and sparse areas ($q<0$). E.g. D_2 may be obtained for $q=2$. Especially $q<0$ is new here: We also want to characterise the areas where few events happen. The limits D_∞ and $D_{-\infty}$ thus describe the seismologically important areas of most and least intensive occurrence of earthquakes.

As all differently clustered areas must obey scaling laws of their own, a multifractal can be said to be a fractal of fractals. Multifractality is not a property of a point set (the support) itself, but of measures such as their spatial density, information etc. (`derived measures') or the size of events at the points (e.g. seismic energy, landslide size).

One may plot D_q vs. q to obtain the spectrum of generalised dimensions but here I prefer the $f(\alpha)-\alpha$ curve, the multifractal spectrum. The latter has a direct physical meaning and may be obtained from the D_q-curve by a transformation (Legendre).



α may be said to be a local scaling index; f is the fractal dimension of all points (areas) with local scaling exponent α. Hence, f tells how many `α-areas' there are! Furthermore, $\alpha_{min} = D_{\infty}$, $\alpha_{max} = D_{-\infty}$, $f(\alpha_0)=D_0$, $f(\alpha_{min})$ tells how many most strongly clustered areas there are, $f(\alpha_{max})$ tells how many least populated areas there are. Thus, one might expect these values to be extremely sensitive to changes in seismicity patterns.

For a monofractal, the $f(\alpha)-\alpha$ curve should collapse into the single point $f = \alpha = D_0$. In practice, due to numerical effects, almost anything will seem multifractal though; all the above constraints must be taken into account, a comparison with synthetic data sets of the same overall properties should be performed and $\Delta\alpha$ should be big enough to trust multifractality. Especially the calculation of high order (negative) moments is critical.

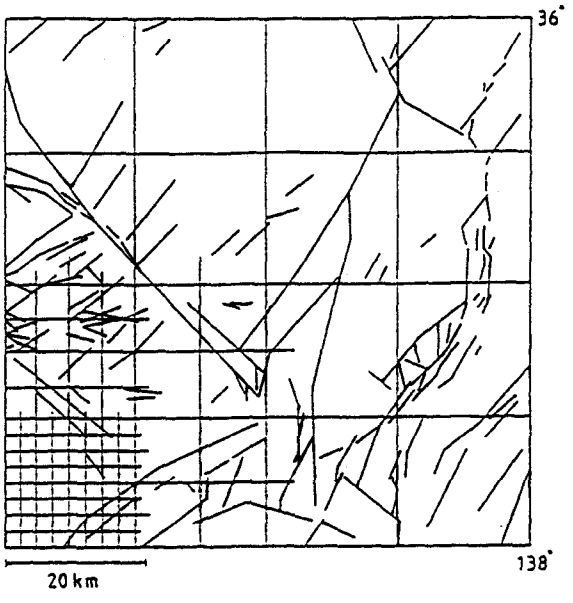
What's fractal about EQs?

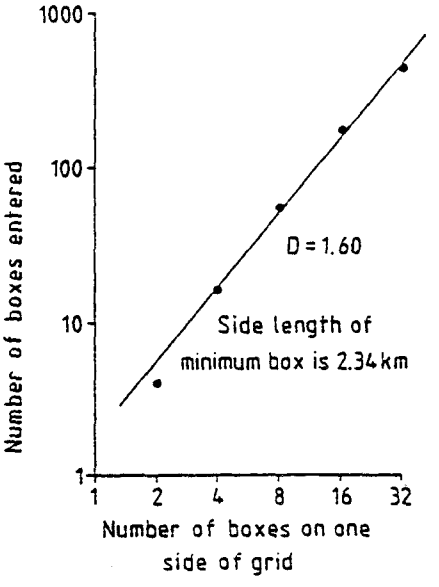
Earthquakes have so many different fractal properties that they belong to the most interesting fractal phenomena (Takayasu 90)

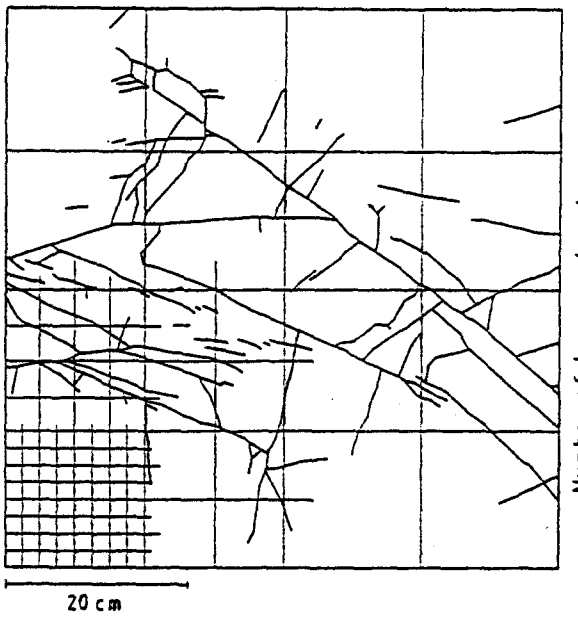
In addition to theoretical evidence for fractality of EQs (SOC e.g. Sornette and Sornette 1989, but also other slider-block models), the following empirical observations were made in the past:

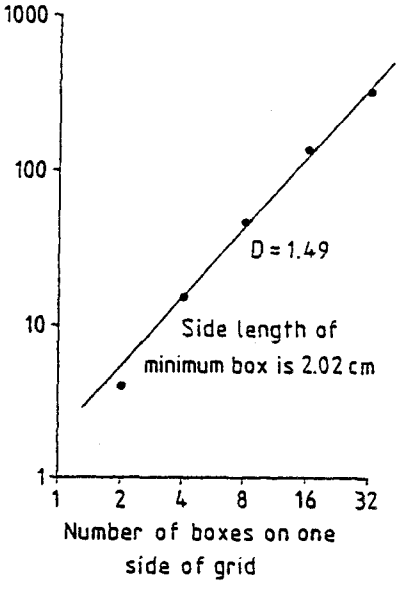
- D of regional and worldwide hypo- and epicentre distribution is 1.2 - 1.6 (e.g. Kagan and Knopoff 1980, Sadvskiy et al. 1984, Okubo and Aki 1987, Aviles et al. 1987); (multifractal: Geilikman et al. 1990, Hirata & Imoto 1991, Hirabayashi et al. 1992, Hooge et al. 1994)
- Faults and fractures (in heterogeneous media) and their distribution are fractal (e.g. Brown & Scholz 1985, Scholz & Aviles 1986, Okubo & Aki 1987, Sammis & Biegel 1987, Hirata 1989, Kagan 92)
- The Gutenberg-Richter relation $f(m) \propto m^{-D}$ is `fractal' ($\Rightarrow f(E) \propto E^{-2b/3}$ also scaling); $D = 2b$ (Aki 1981)
- The Omori-law $n(t) \propto t^{-D}$ means temporal scaling of excited aftershocks; also confirmed by laboratory experiments (AE, e.g. Scholz 1986, Hirata 1987)
- In general, the temporal distribution of EQs is (multi)fractal too (e.g. Godano & Caruso 1995) (Ito, 95)
- Spatial distribution of EQ size (seismic energy) is (multi)fractal (e.g. Hirabayashi et al. 1992, Hooge et al. 1994)

Temporal and spatial scaling together show that the EQ process is a scaling spatio-temporal process (i.e. coupled).
One important consequence: EQs not Poissonian! (cf. Mandelbrot 1983)


36°
138°
20 km


1000
100
10
1
Number of boxes entered
D = 1.60
Side length of
minimum box is 2.34 km
1
2
4
8
16
32
Number of boxes on one
side of grid


20 cm


1000
100
10
1
Number of boxes entered
D = 1.49
Side length of
minimum box is 2.02 cm
1
2
4
8
16
32
Number of boxes on one
side of grid

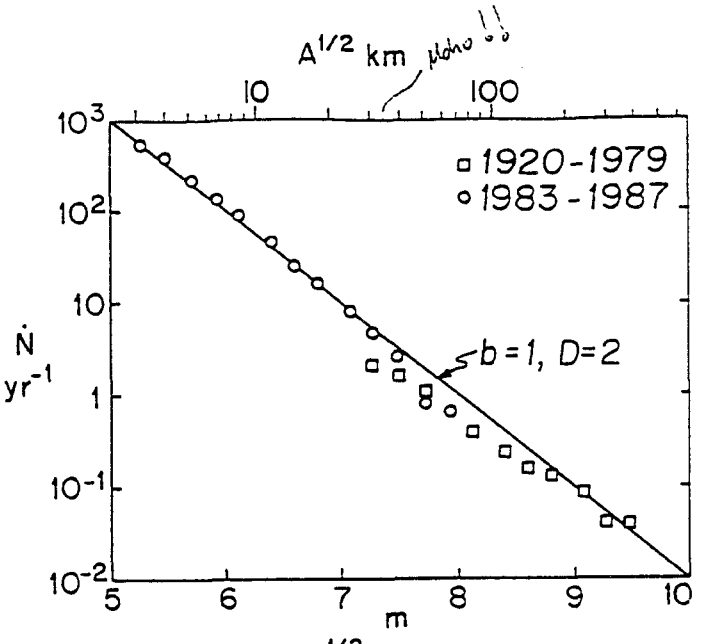


Figure 4.1. World-wide number of earthquakes per year, $\dot{N}$ with magnitudes greater than m as a function of m. The square root of the rupture area A is also given. Circles, 1983–1987 (Dziewonski *et al.*, 1989, and others); squares, 1920–1979 (Purcaru and Berckhemer, 1982). The solid line represents (4.1) with $b = 1.00$ and $\dot{a} = 10^8\,\mathrm{yr}^{-1}$.

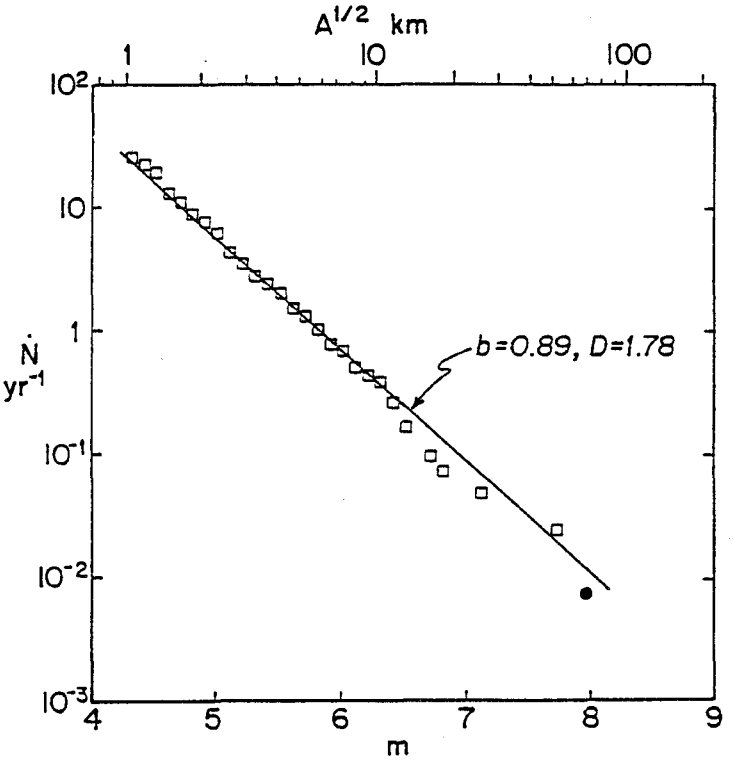
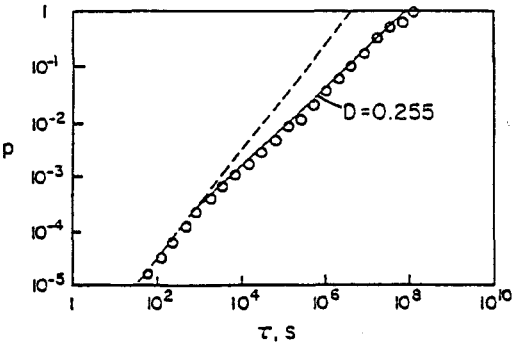
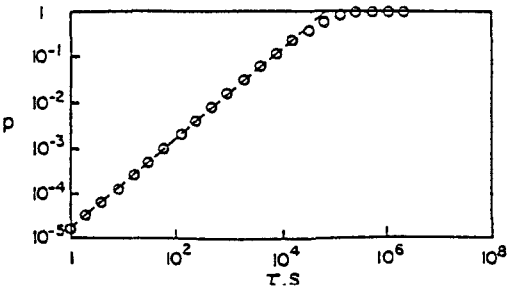


Figure 4.2. Number of earthquakes per year $\dot{N}$ with magnitudes greater than m as a function of m. Squares, Southern California 1932–1972 (Main and Burton, 1986); solid circle, expected rate of occurrence of great earthquakes in Southern California (Sieh *et al.*, 1989). The line represents (4.1) with $b = 0.89$ and $\dot{a} = 1.4 \times 10^5\,\mathrm{yr}^{-1}$.



(a)



(b)

Figure 6.3. Fractal cluster analysis of 49 earthquakes that occurred near Efate Island, New Hebrides in the period 1978–1984 (Smalley *et al.*, 1987). (a) The circles give the fraction of intervals p of length τ that include an earthquake as a function of τ. The solid line represents the correlation with (6.2) taking $D = 0.255$. The broken line is the result for uniformly spaced events. (b) The results for 49 randomly distributed events (Poisson process).

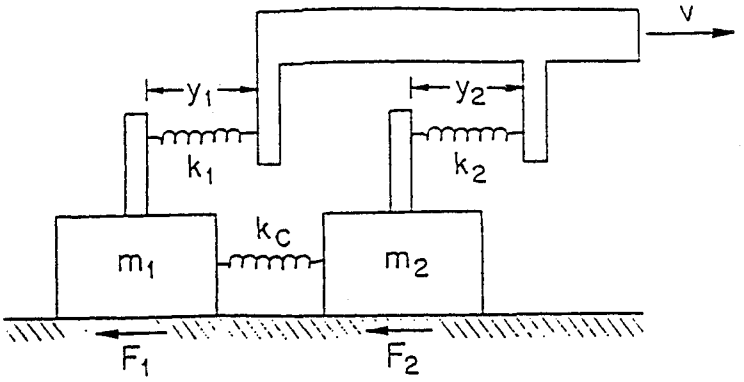


Fig. 5. Illustration of the two-block model. The constant velocity driver extends the springs until sliding of a block commences. In some cases sliding of one block induces the sliding of the second block.

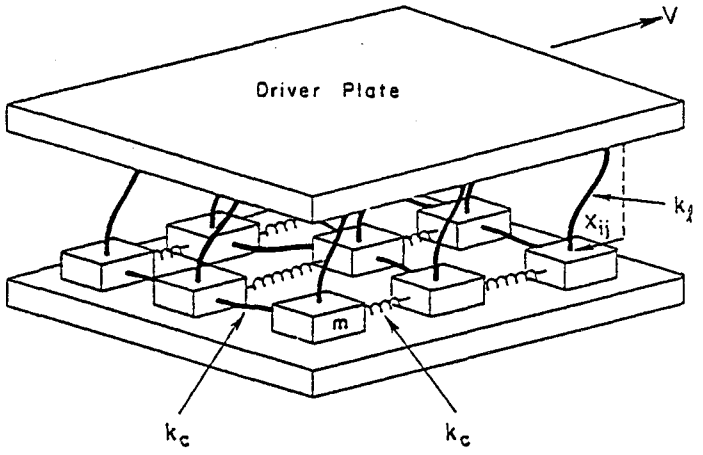


Fig. 8. Illustration of the two-dimensional slider-block model. An array of blocks each with mass m is pulled across a surface by a driver plate at a constant velocity V. Each block is coupled to the adjacent blocks with either leaf or coil springs of spring constant k_c, and to the driver plate with a leaf spring of spring constant k_l. The extension of the (i, j) pulling spring is x_{ij}.

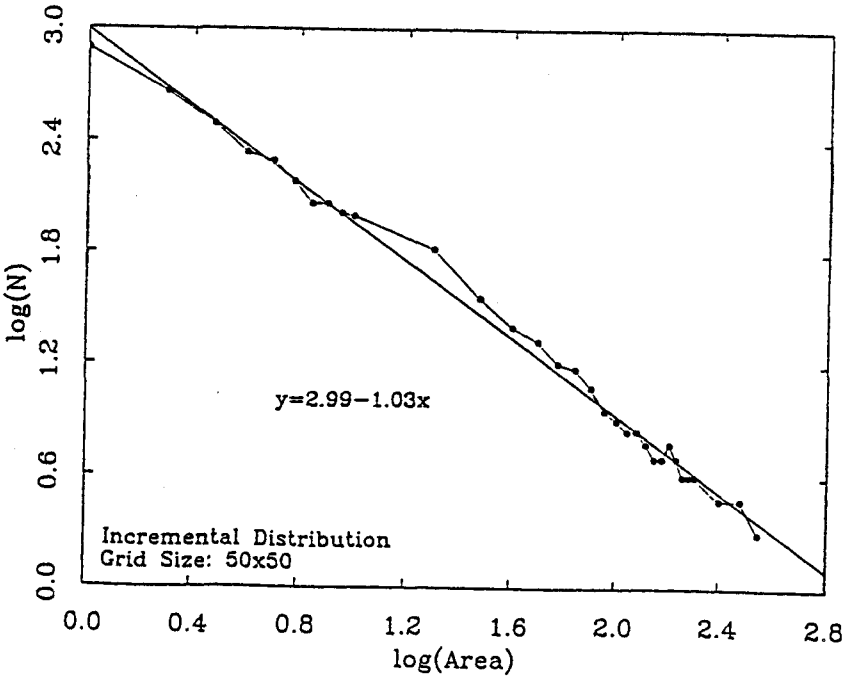


Figure 16.2. Statistics for a cellular automata model on a 50 × 50 grid. The number N of events in which a specified number A of boxes became unstable is given as a function of A.

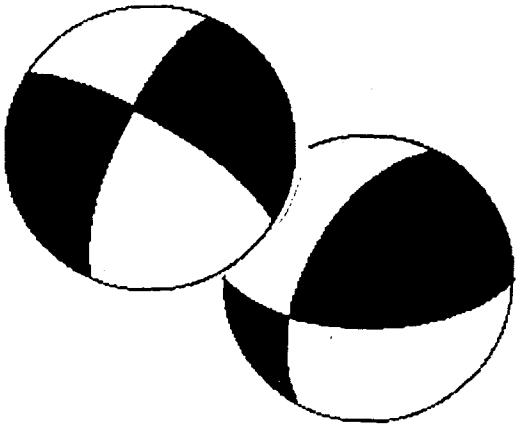
Possible Reasons for Earthquake Complexity and Consequences

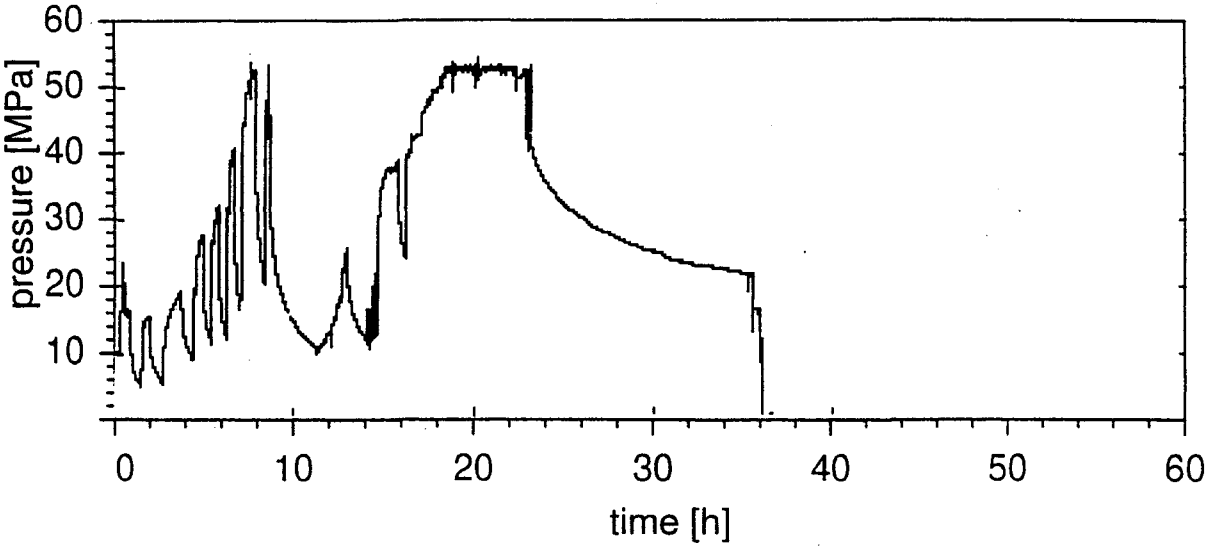
Chaos: deterministic; described by system of non-linear differential equations; exponential sensitivity to small changes in initial conditions.
Problems: systems (block models) need tuning; applicability of models to nature?
Consequences: Depending on Lyapunov exponent, only (very) short-term predictions possible.
But: tools of non-linear time series analysis might be useful for effective monitoring

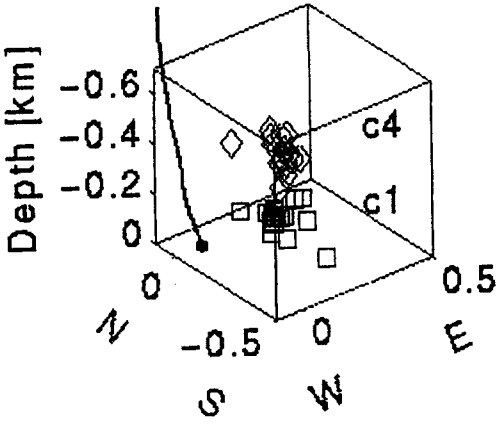
Critical Point Phenomenon/Phase Transition: divergence of correlation length -> fractal at the critical point.
More general: transition simple <-> complex (percolation threshold)
Problems: Crust is non-equilibrium system, needs tuning towards critical point. Transition of the type "fractal <-> not fractal" not observed.
Consequences: If transition happens slowly, a warning of increased hazard level might be possible.

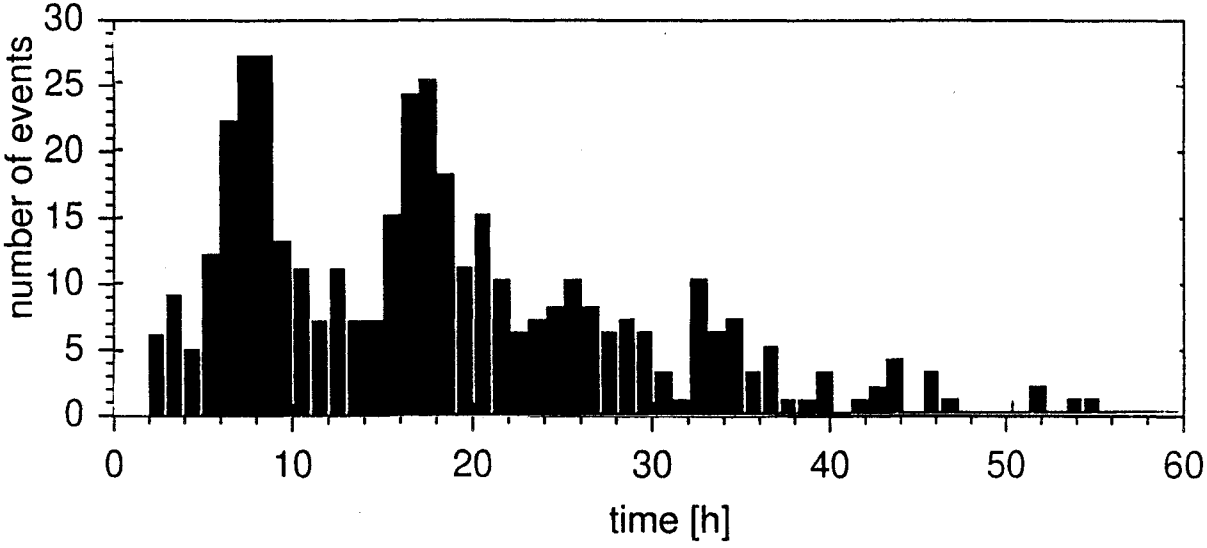
SOC: non-equilibrium state where small perturbations can cause events of all sizes (following a power-law size distribution). No tuning required.
Problems: ?
Consequences: Prediction of large events impossible because large events differ from small ones only in size!
Strong evidence also from induced seismicity.

However, it is possible, that a combination of the above mechanisms applies...









KTB / induzierte Seismizität

Zoback und Harjes 1997
Jost et al 1999

Zeitliche Variation
Temporal Variation

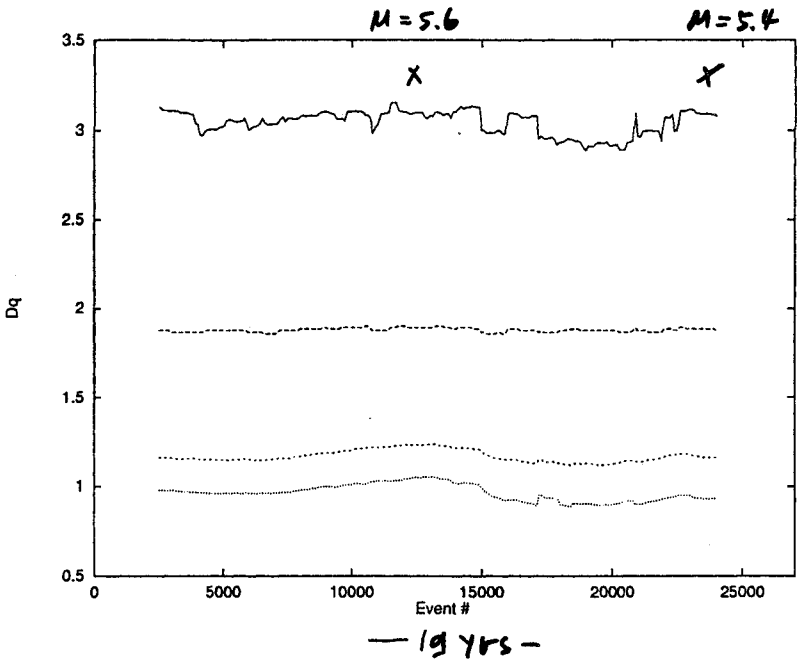
Kyoto Univ.

Data from Jan. 1976 up to m=6.9 Hyogo-ken Nanbu EQ on Jan. 17, 1995 at 34.60 ° N, 135.03 ° E (Abuyama Obs., Kyoto Univ., H. Katao). Magnitude completeness is 1.5 but was ignored for more events and inclusion of micro-seismicity. Area selected according to geophysical and numerical constraints (e.g. centered on mainshock, boundary effect for sphere count).
~ 26 500 events

For the moving analyses, a window size of 5000 events with an overlap of 2%, i.e. 100 events was selected to avoid too small data sets but achieve good temporal resolution. The values obtained are thus integral over about 3 yrs on average; they are plotted centered on each interval. The EQ history was recorded for the greater area 33° – 36° N, 134° – 137°E not to miss the influence of large EQs that happened slightly outside the area of analysis.
Several moderate to larger EQs occurred in the interval, most notably an m=5.6 event in May 1984 and an m=5.4 event in Sept. 1990.

Temporal evolution of the spatial density (epicentre distribution)

q= -2, 0, 1, 2
from top

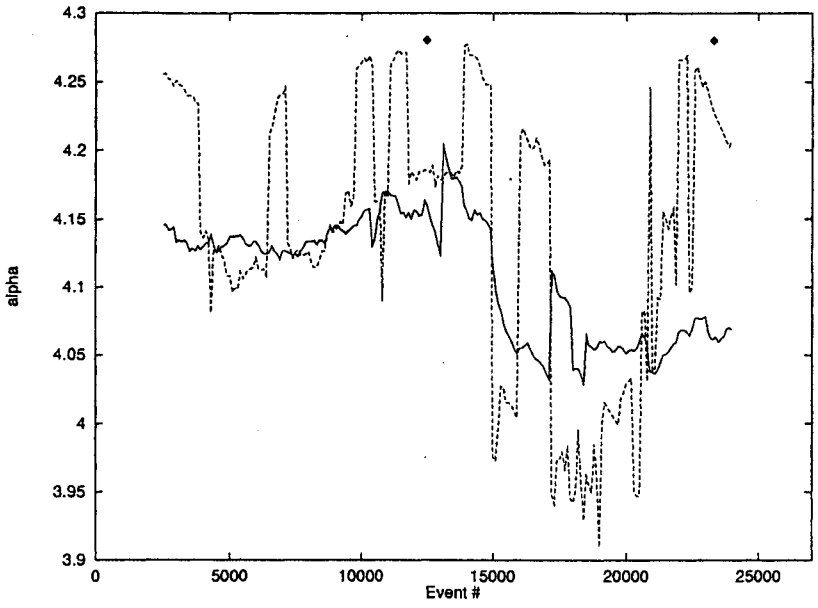


The first three generalised dimensions are not very sensitive to the seismic evolution. The sensitivity increases with increasing positive q though. D_{-2}, as also shown by Hirabayashi et al., is more sensitive. The strongest changes in seismicity occur in less active areas thus. However, it becomes clear that we need the whole spectra to completely describe the changes. There is no clear correlation with big events and no long-term trend though.

There is no way to show more than 200 $f(\alpha)$-α curves, so in the following, I plot characteristics of the spectra as time series.

α_{min} and α_{max} (most dense and most sparse areas):

solid = α_{min}

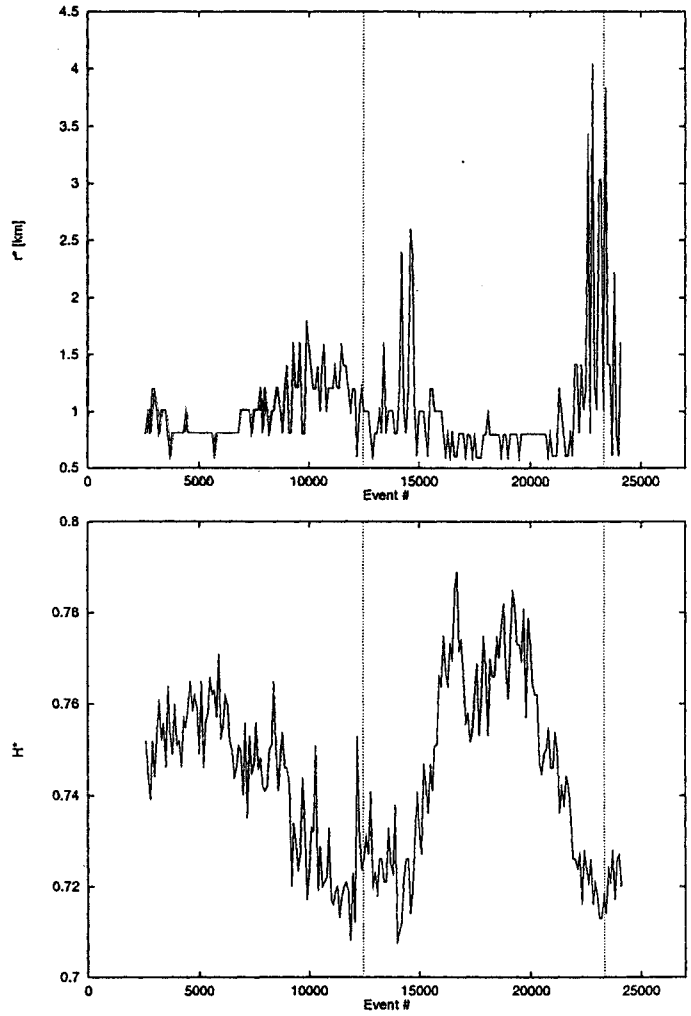


One sees a roughly synchronous behaviour of both values. Nothing points out the Kobe EQ at the end of the interval.(?)
But: There is increase before the m=5.6 event, followed by a decrease and again an increase before the m=5.4 event. The increasing trends start at about 4 resp. 3 yrs before.
Because of the anticipated large error in α_{max}, the abrupt jumps are probably not real.
Recall that

- increase in α_{max} $\Rightarrow$ extremely sparse clusters get even sparser
  decrease $\Rightarrow$ extremely sparse clusters get less sparse
- increase in α_{min} $\Rightarrow$ extremely dense clusters get less dense
  decrease $\Rightarrow$ extremely dense clusters get more dense

Thus, the simultaneous increase of both values indicates that extremal clusters get less populated. Whether overall seismicity decreases, depends on the whole shape of the spectrum.

Moving entropy analysis:



There are no seemingly random fluctuations here, but r* shows discontinuous plateaus. The 2nd peak precedes the m=5.4 or Kobe event, while the m=5.6 EQ is not reflected. Further research into r* seems promising. Better is H*: There are 2 cycles which seem to correspond to the two major EQs besides Kobe. First, the disorder decreases for about 4 yrs from a possible maximum before the m=5.6 EQ occurs at a min. of disorder (= on average most ~~homogeneous~~ clustered ~~(i.e. random)~~ epicentre distribution). Then, disorder assumes a new max. before it decreases again for about 3 yrs before the m=5.4 event.
An entropy analysis is comparatively easy and reliable.

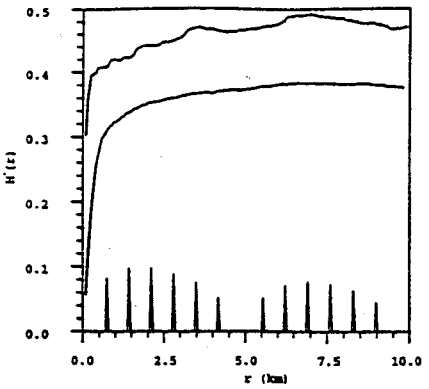


Figure 1: Examples for entropy curves for, from top to bottom, a Sierpinski triangle, a homogeneous random distribution and a regularly spaced set of grid points.

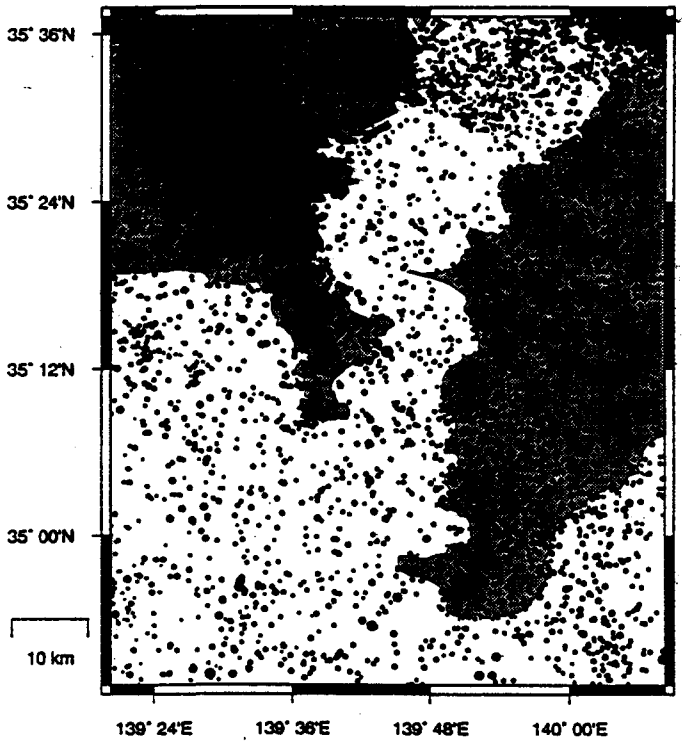


Figure 2: A subset of the NIED earthquake catalogue ("Region A") selected according to the criteria given in the text. The area is centered on the February 1992 M = 6.1 earthquake and contains 4,600 events total. Circle size corresponds to magnitude, colour to depths.

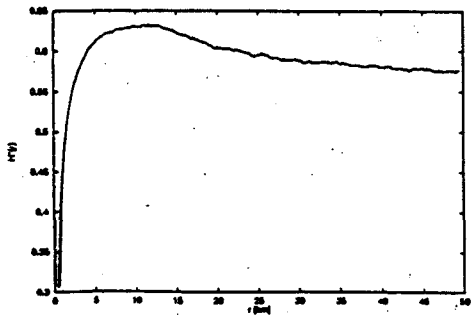
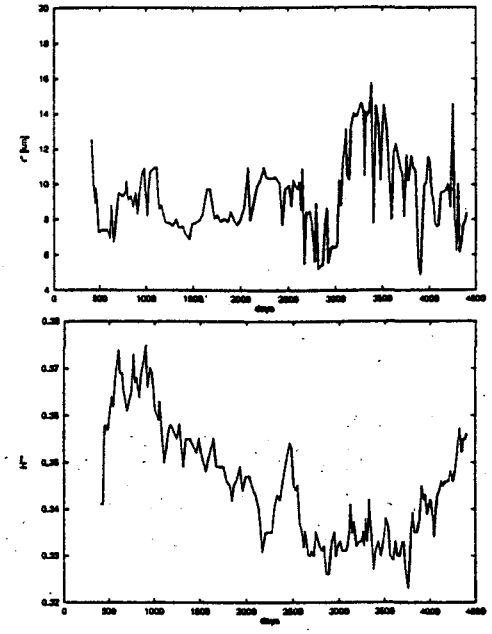


Figure 3: Entropy curve for the data from Region A, typical for a multifractal because of the hump with a well defined maximum.



crit. point

correlation length (diverges?)

(increases)

SOC

spatial clustering increases

Figure 4: Results of a moving entropy analysis of Region A. r^* is shown in the upper part, H^{***} below. Note that, especially in the case of H^{***}, the curves possess clear trends over several years.

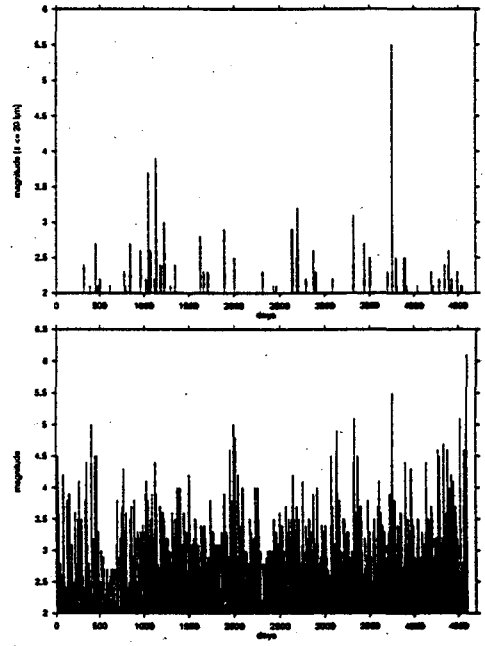
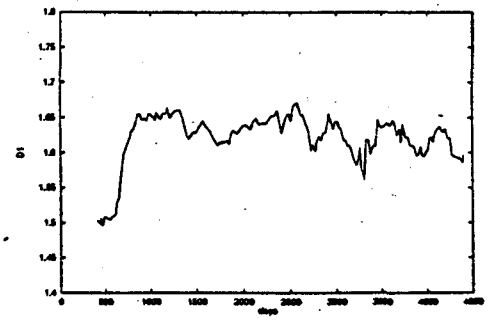


Figure 5: Magnitude versus time for Region A (Top: Only shallow events; Bottom: All events). H^{***} decreases for several years before the occurrence of large events. This correlation shows best for shallow large events. See text for r^*.



" $b = 0.9 \pm 0.15$ " (D too)

Figure 6: Result of a moving determination of the fractal information dimension D_1. The fluctuations seem random and are largely within the error.

- more case histories
- surrogate data
- 3D
- grid algor. to determin region of max. change

Chaos: Basics

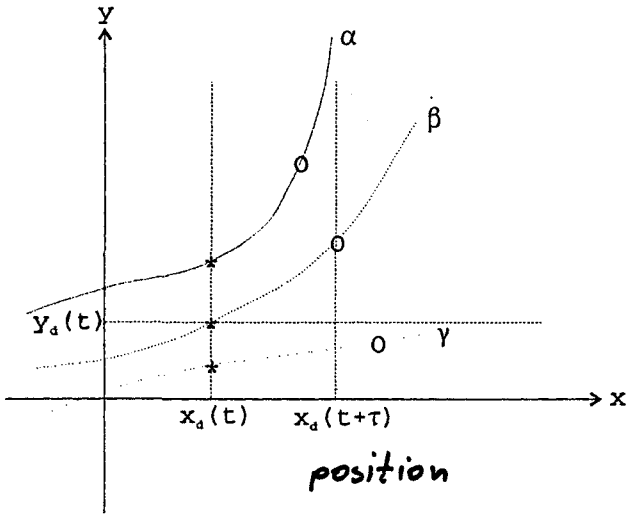
Now not seismicity patterns but time series are analysed. The main question: Is it possible to model EQ-time series (and therefore the EQ process itself) by low-dimensional deterministic models? Related: Can one unambiguously detect precursors? Are EQs chaotic and what does this mean?

Chaos = low-dimensional (< about 6 degrees of freedom) deterministic chaos (high-dimensional chaos = noise for numerical purposes): Although completely deterministic, very limited predictability. Time series look like noise (optically and to linear methods like autocorrelation and Fourier analysis).

The main part in nonlin. analysis is the attempt to reconstruct the dynamics in phase space from the measured scalar signal. Then fractal analysis is applied to the obtained attractor which is made up of the trajectory of the system with time. The trajectory describes the systems state (phase) at every time.

How can one obtain the phase space structure from a one-dimensional time series?

An example of a two-dimensional phase space for the motion of a particle in one dimension:



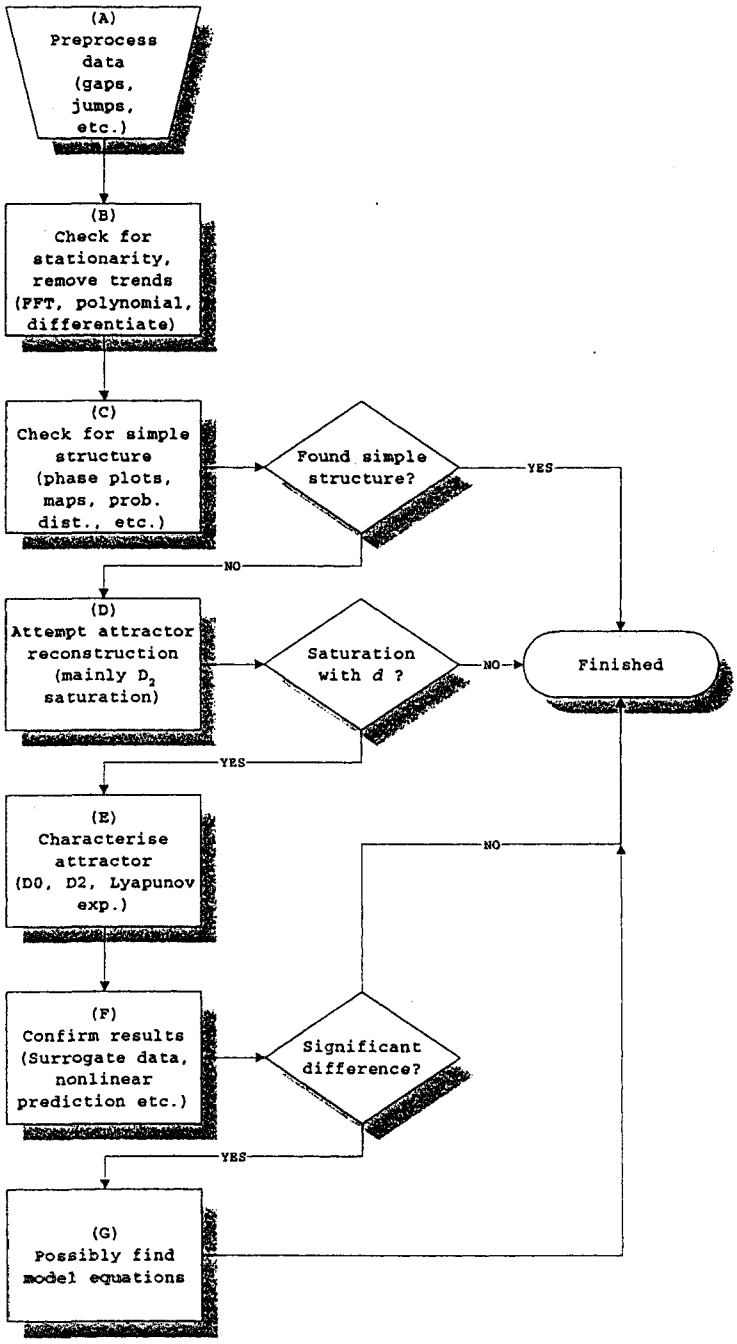
e.g. velocity

*: at time t
o: at time t + τ

By only measuring x(t), one can still reconstruct the original trajectory by delay-time embedding.

One gets the correct embedding dimension *d* by embedding the time series in successively higher *d* and checking for saturation in, usually, D_2. D_2 becomes constant at that *d* where the attractor is fully unfolded (Looking at a plane (D_2=2) in *d*=1, we get D_2=1; at *d*=2, D_2=2; but at *d*=3, still D_2=2, i.e. D_2 has saturated and we know that the object is only two-dimensional).

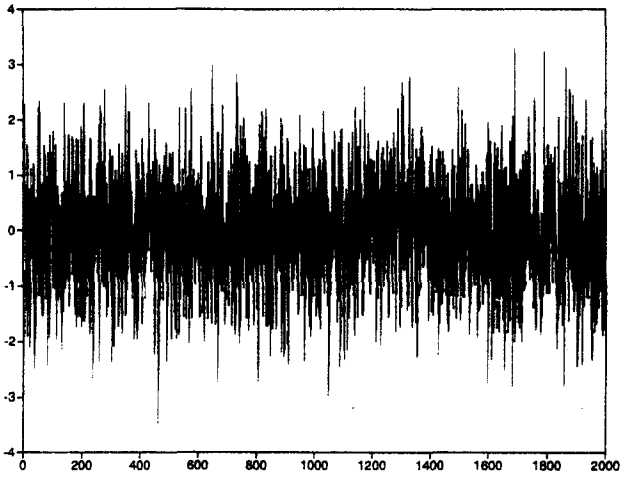
Only if deterministic structure in signal, D_2 saturates at all: Major idea to test for structure in seemingly random signals. In addition, if D_2 is non-integer, the determinism is caused by chaotic dynamics.



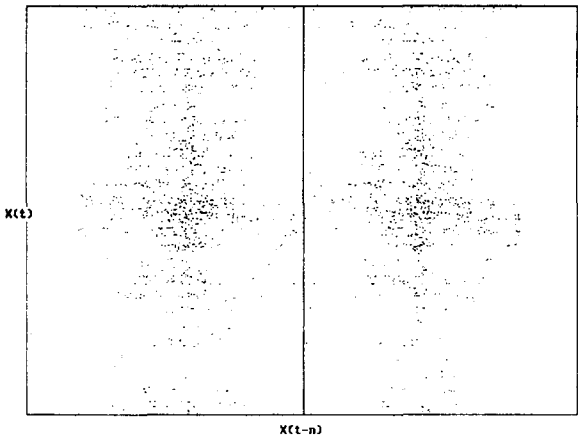
Will become more clear in the following model applications:

Model Dynamics

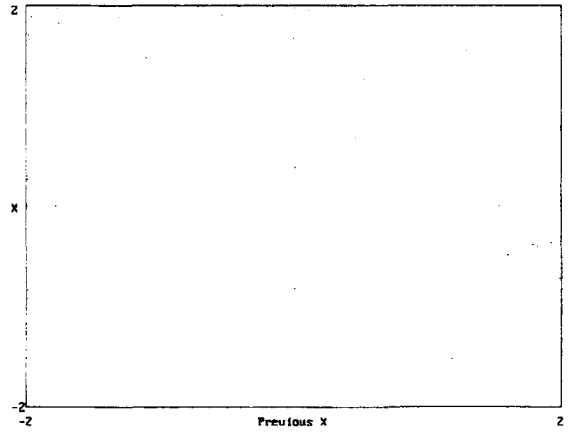
Example model 1:



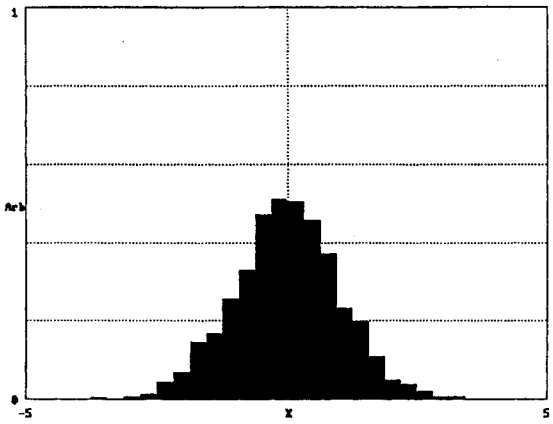
Following the outlined strategy, first some simple tests:



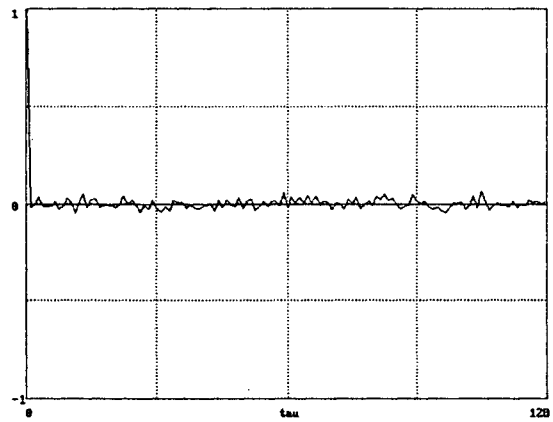
Stereographic plot of an embedding into two dimensions with delay time 1 (i.e. x(t) vs. x(t-1): No apparent structure



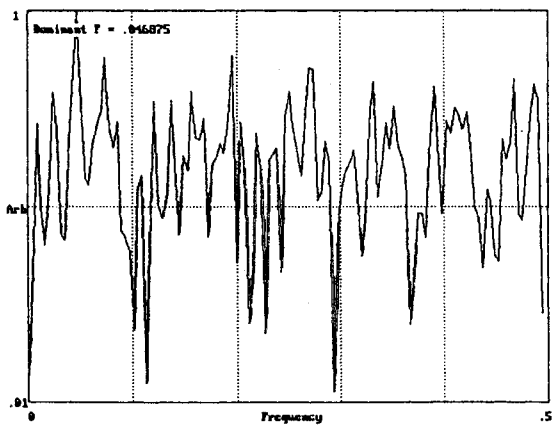
A return map where x' = 0.5 is plotted vs. the previous time when this was fulfilled. This is another way to unveil structure: No structure either.



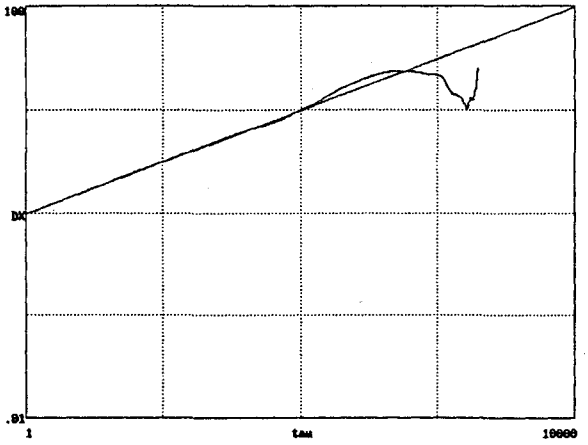
The probability histogram looks Gaussian: Noise.



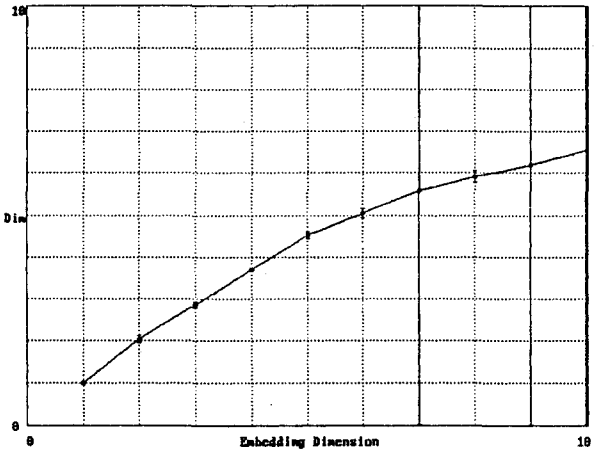
The autocorrelation function drops to zero abruptly and remains there: No correlation in data: Uncorrelated noise.



The power spectrum is broad (no periodicity) and flat (no 1/f or coloured noise: White noise.



The integrated signal resembles a random walk (Brownian motion) as witnessed by a Hurst coefficient of 0.49: White noise.

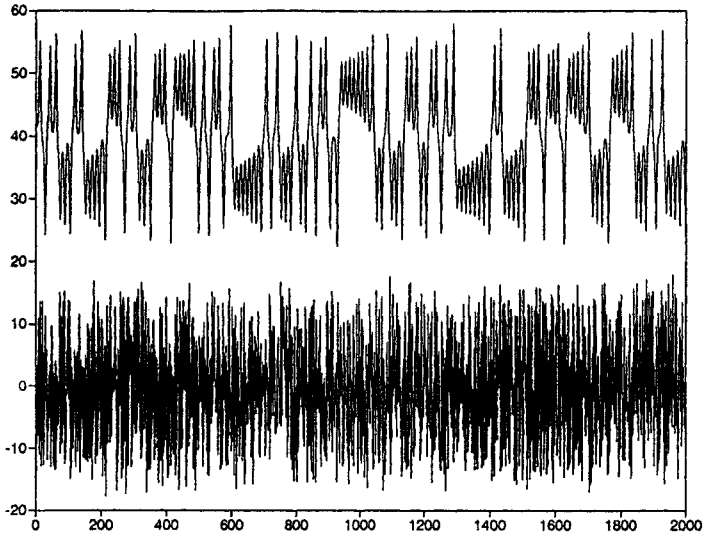


Finally, there is no saturation of D_2 up to d = 10. Also the large error bars indicate poor scaling properties.

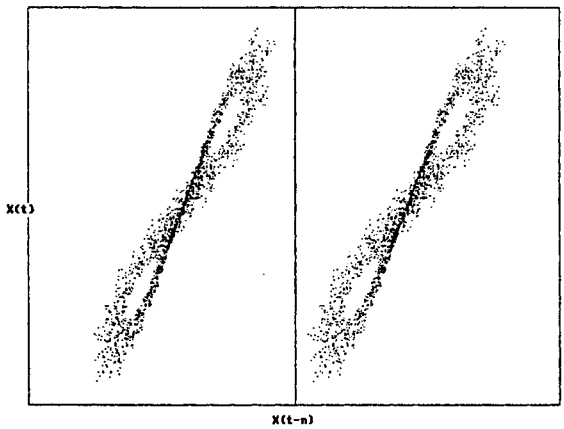
One may conclude that, at least for practical purposes, the data is uncorrelated (white) noise. Such data can not be modelled by deterministic models.

The data really represents 2000 points with a Gaussian distribution of zero mean and a standard deviation of unity.

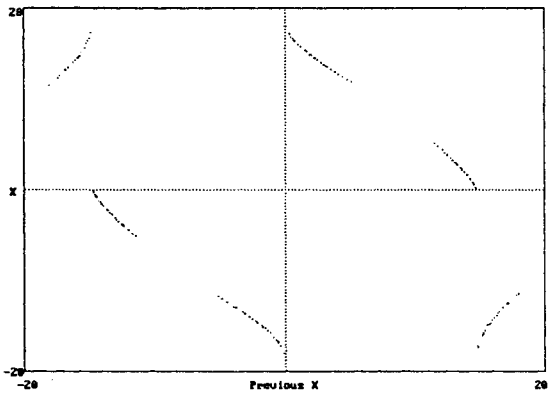
Example model 2:



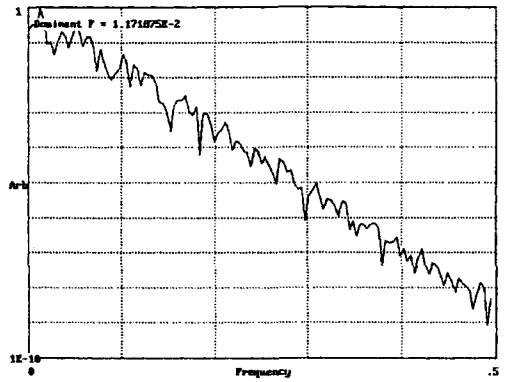
shuffled:
same prob. distr,
diff. power spectrum
and corr. function



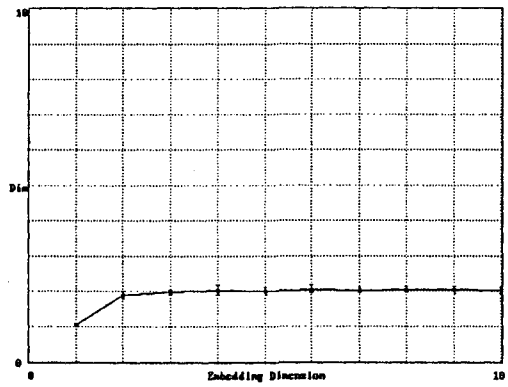
The x(t) vs. x(t-1) plot already shows a well defined object: There is some determinism.



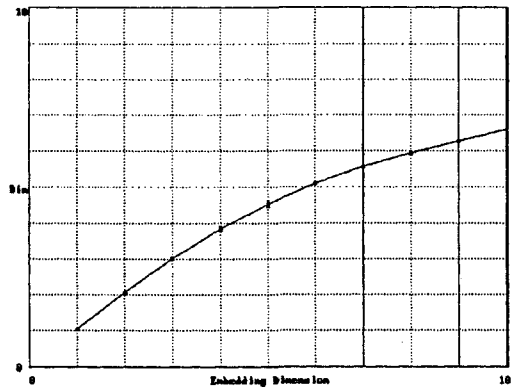
The return map shows a very simple, almost linear structure. We may thus expect the attractor to have a fractal dimension of about 2.



The power spectrum is broad, i.e. no predominant peaks due to periodicity and shows clear power-law behaviour: Good candidate for chaos.



D_2 shows very clear saturation at about 2; also the estimated error for D_2 is very small: Together with the previous tests, this must be chaos.



When embedding the randomised data, all structure is gone (if the saturation was spurious before, it would still be here now): The above result is ~~true~~ not contradicted.

As usually $n = D_2 + 1$ equations suffice to model a system with an attractor of D_2, one expects to be able to model the system with three equations.
Indeed the signal was produced by integrating the Lorenz model of three coupled first-order nonlinear differential equations with a Δt of 0.05.

Example 3:

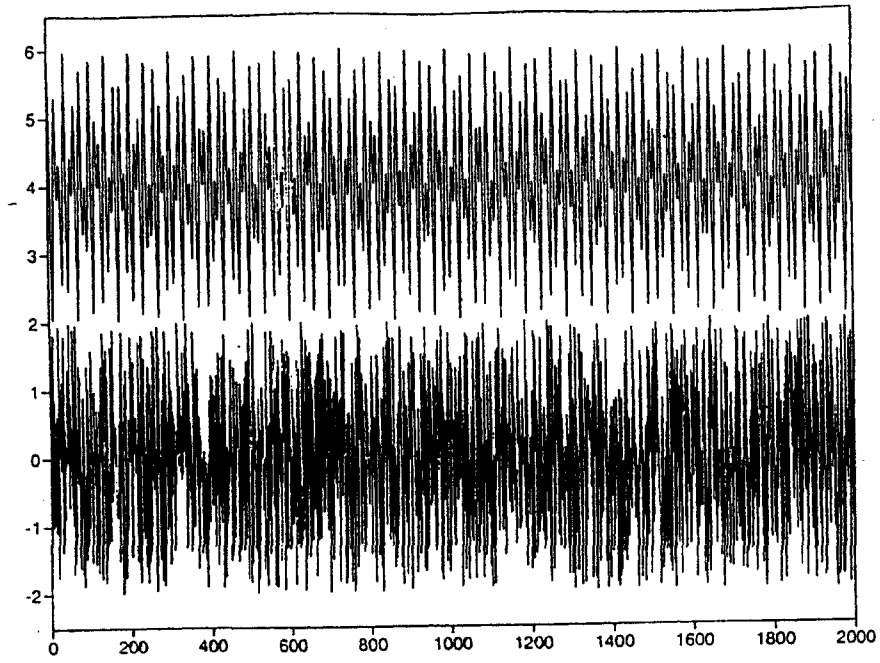
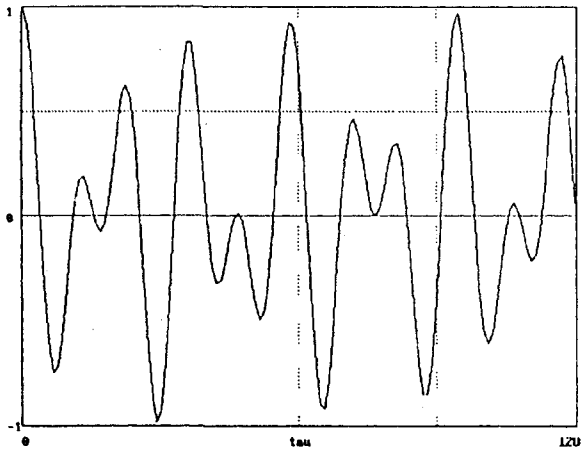
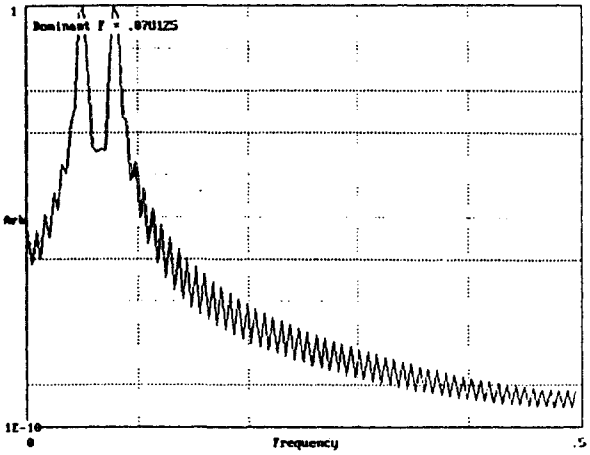


Figure 7.3: A seemingly random signal of two incommensurate frequencies (above) and its shuffled version (below)



High correlation, periodicity → we already know: no noise, no chaos.

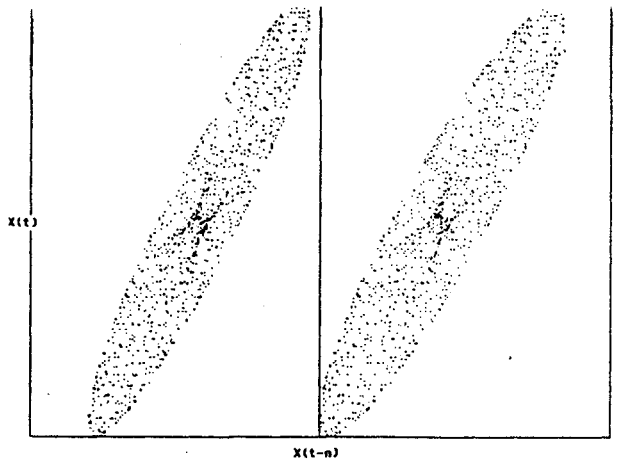
Figure 7.5: Autocorrelation function of quasi-periodic data



Confirms two predominant frequencies.

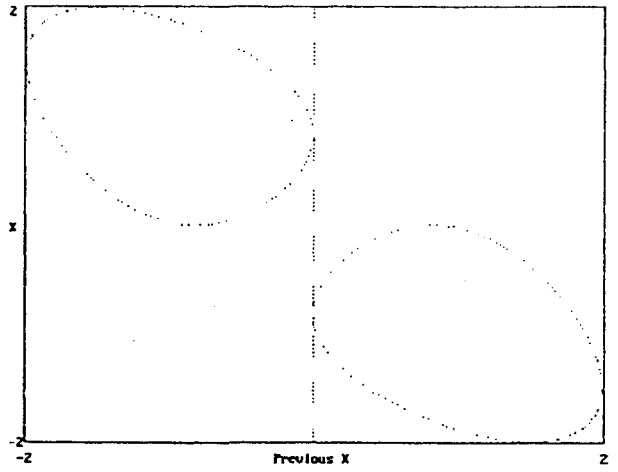
Figure 7.4: Power spectrum of quasi-periodic data

→ Boring data. But let's see what the other "tools" will show...



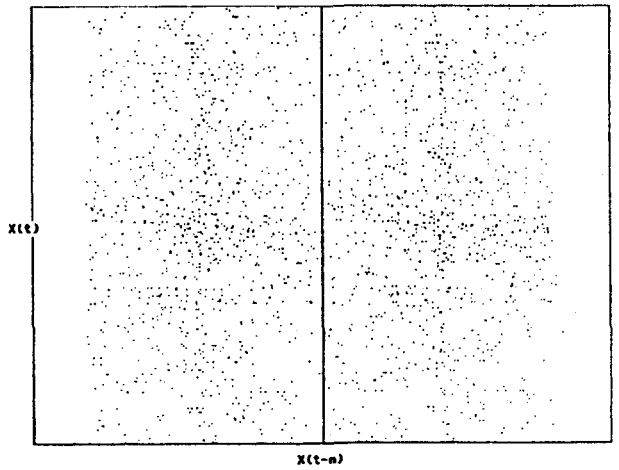
simple toroidal structure becomes visible.

Figure 7.6: Stereographic delay time plot of quasi-periodic data revealing a torus



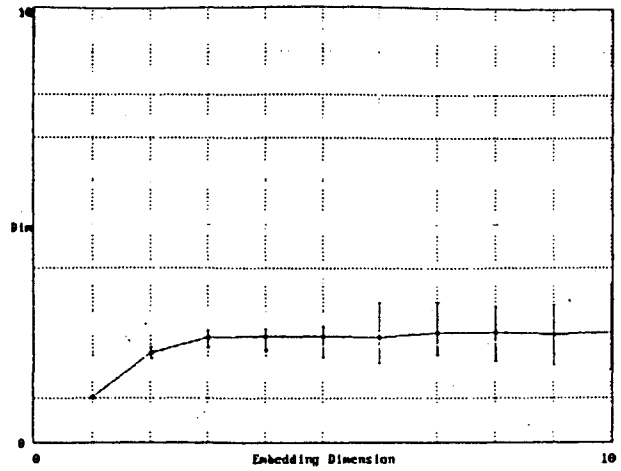
Return map ≙ cut through torus

Figure 7.7: Return map of quasi-periodic data in which $x(t)$ is plotted at positions where $x'(t) = 0.5$ (y-axis) versus the previous time where the condition was fulfilled (x-axis)



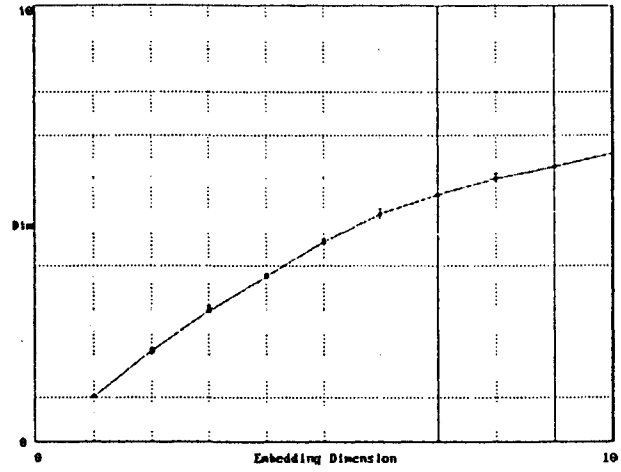
there was real determinism!

Figure 7.8: Phase space plot of shuffled quasi-periodic data: The structure seen in Fig. 7.6 has been destroyed



very clear saturation at $D_2 \sim 2$ as expected.
⇒ without the previous tests we couldn't rule out chaos.

Figure 7.9: $D_2(d)$ for embedding dimensions of 1 to 10 for quasi-periodic data



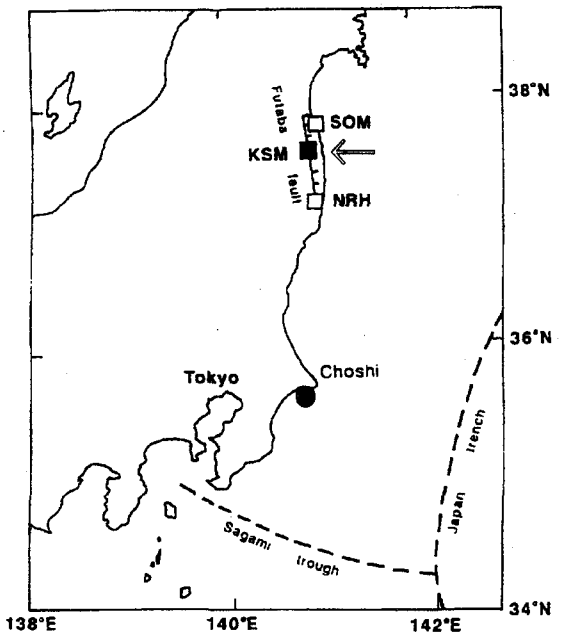
Saturation plot also shows loss of deterministic structure after randomisation of original signal.

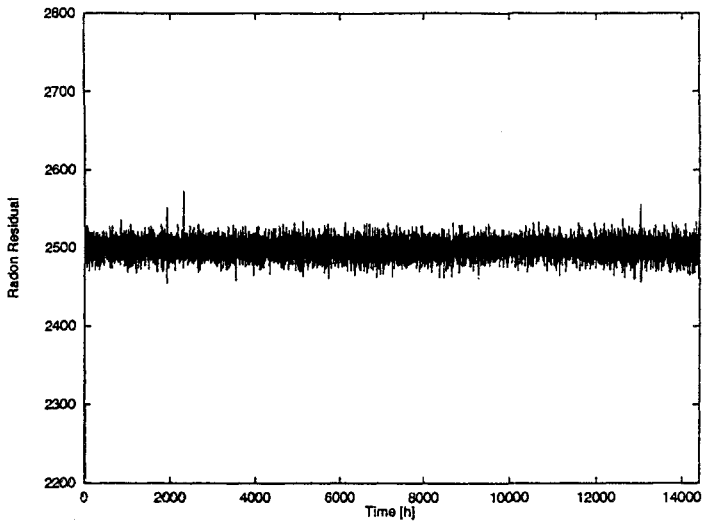
Figure 7.10: $D_2(d)$ for embedding dimensions of 1 to 10 for shuffled quasi-periodic data for confirmation of structure in the original data

Application to EQ data

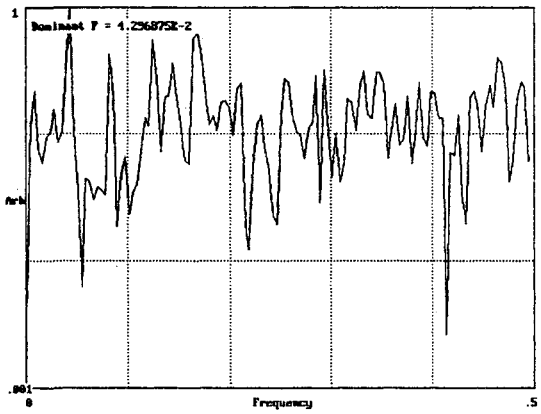
Radon is the only radioactive gas emitted from the surface of the crust and thought to be predominantly governed by stress changes in the underlying basement rock. Changes in radon concentration in groundwater with time are considered a promising EQ precursor (e.g Wakita 1982).

The data is from the Laboratory for EQ Chemistry at the Univ. of Tokyo (G. Igarashi) and was measured at site KSM at the Futaba fault in NE Japan.

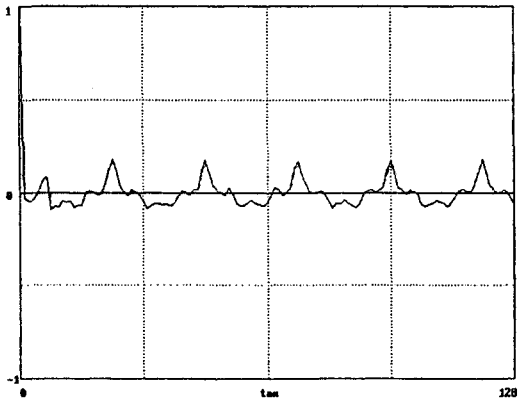




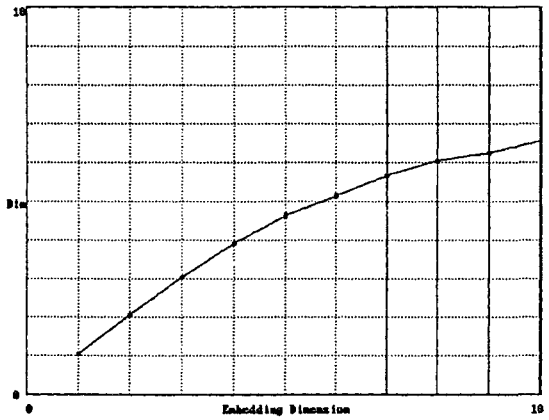
This is the non-harmonic residual of about 14500 hourly values from Nov. 1985 on after seasonal and other periodic trends have been removed. As the y-scale compares to the one of the original data, it may be seen that the residual has small amplitude and variation.



The power spectrum is typical for white noise: Broad and flat.



The autocorrelation function confirms this: It drops to zero quickly and remains there except for a periodic artifact from detrending.



Finally, there is no asymptotic behaviour of D_2. Thus there is no low-dimensional structure in the radon data.

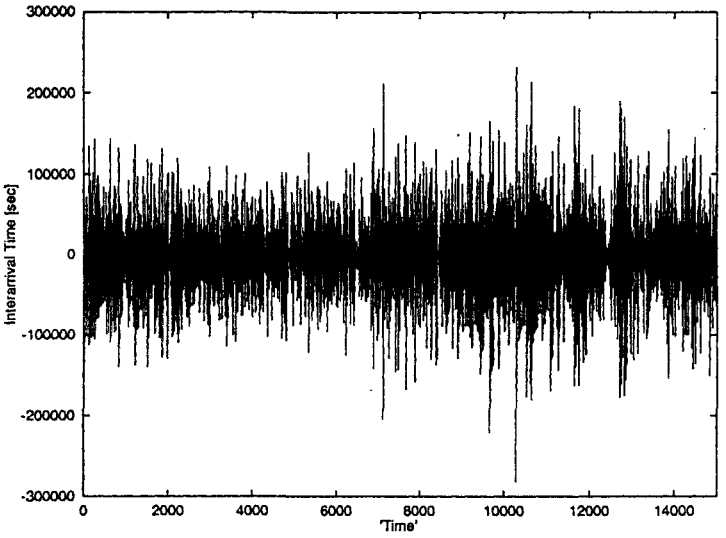
For all practical purposes, one must assume that the radon signal as observed here represents noise.

The result is not surprising as almost all geophysical and other experimental time series, including strain data as observed at Yamasaki fault by Kyoto Univ. (K. Watanabe), show this behaviour. Possible reasons are that the dynamics might be too high-dimensional, that there might be dynamic noise (noise which does not simply add but (nonlinearily) interact) or, finally, that the data really possesses no determinism.

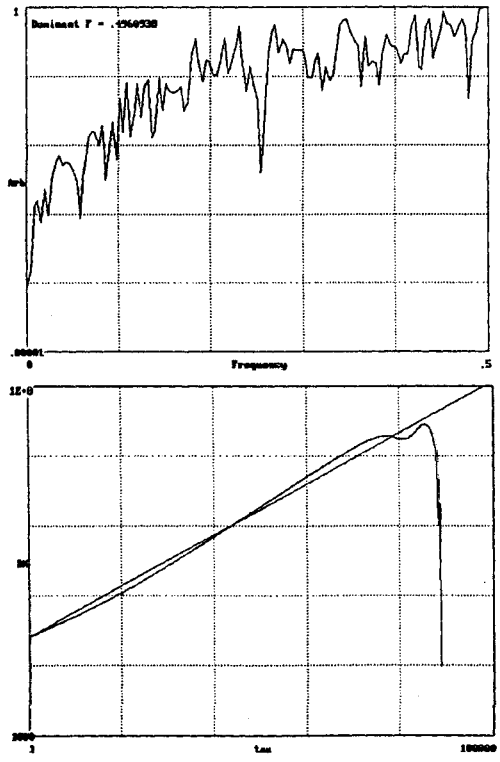
Consequence: There is no way to deterministically model the dynamical systems underlying these signals. Thus one can not obtain a reliable criterion to distinguish anomalous from normal behaviour in that one can not monitor any deviations. Unphysical statistical criteria must be used like in the past.

One can not say, however, that the occurrence of EQs is noise-like (e.g. Poissonian), because we have no proof that the time series capture EQ dynamics to a sufficient degree.

This is different with the `time series' of EQ intervals, which is directly derived from the previously introduced EQ catalogue. →*



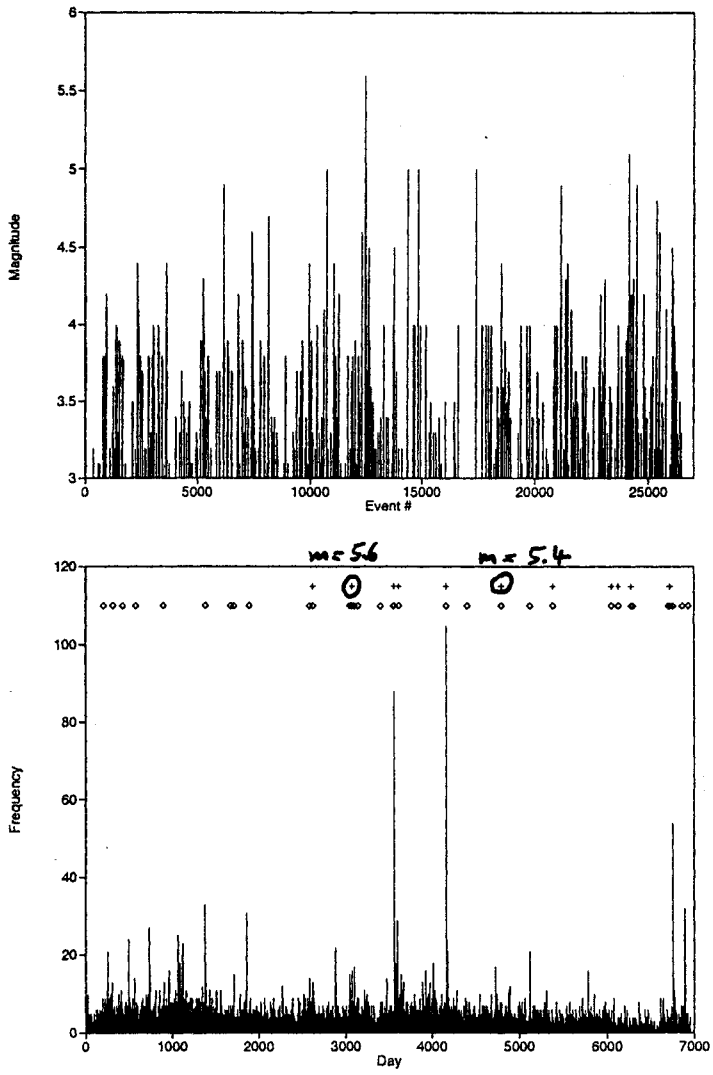
This is the first derivative of the first 15000 (of about 28000) EQ intervals. The inter-arrival time fluctuation is used to obtain a zero-mean time series instead of spikes.
Following the introduced strategy of nonlinear analysis:



The power spectrum is puzzling as it first follows an increasing power law and then seems to settle on a plateau for high *f*. It is clearly not white noise though.

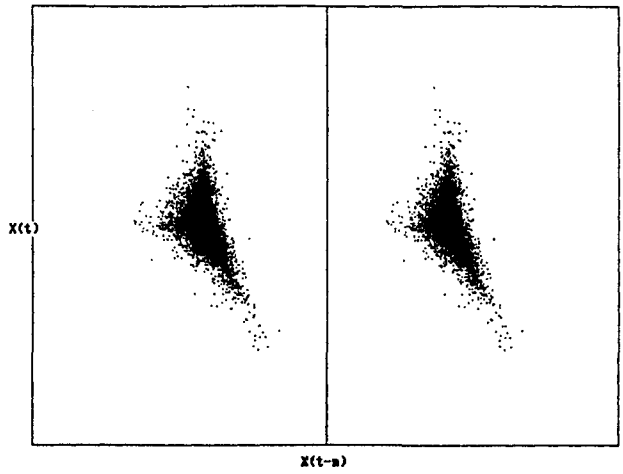
The latter is confirmed when obtaining a Hurst exponent of 0.73 which implies correlation.

The history in the area of analysis *

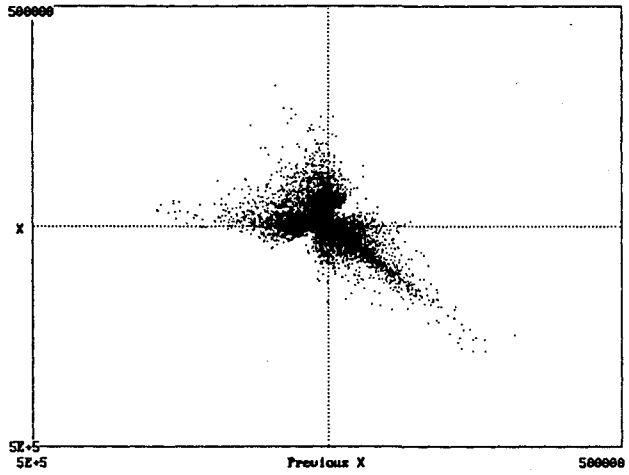


Indicated are all events of m>4.5 and m>5 (crosses) within the greater area though.

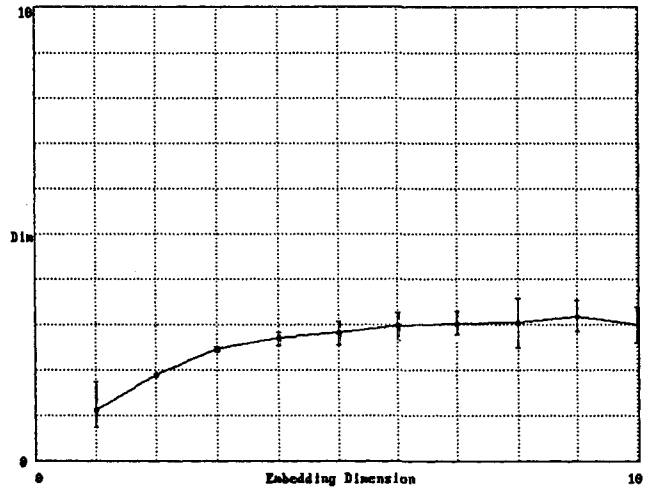
* Jan 1976 – Jan 1995
220 × 180 Km centered on the m=6.9 Kobe event of Jan 1995
(33.7° Lat. N – 35.7° Lat. N, 134.3° Lon. E – 136.3° Lon. E)
~ 26.500 events of m > 0.1
Kyoto Univ. network



The phase space plot attempt shows that the space is not filled homogeneously but that there is some structure indicating determinism.

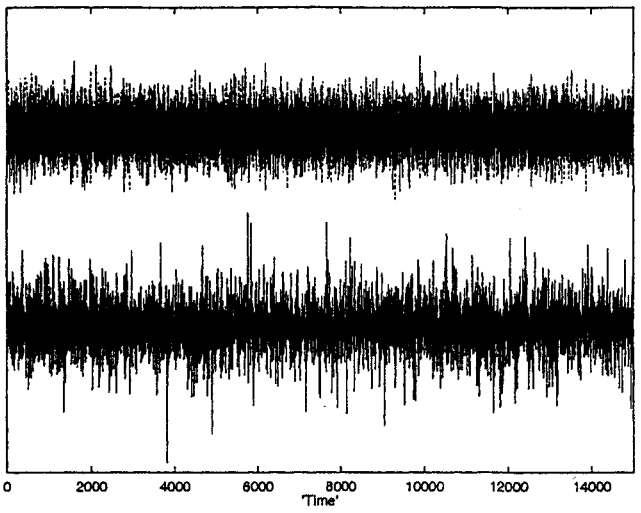


The return map confirms this in showing more detail. We can not estimate the dimension of the attractor from this though.



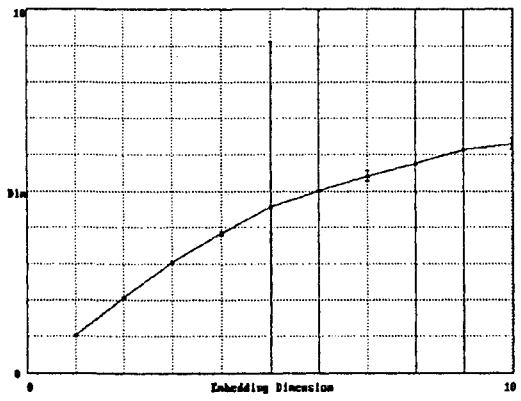
The embedding shows a clear saturation at about 3! Also the error bars for D_2 are quite small, limiting the dimensionality of the attractor to 2.5 - 3.5. Note that the data is not smoothed or filtered in any way.

The former result is surprising as it would mean that one can model the temporal occurrence of EQs with as little as probably four equations. Thus one must strictly confirm the result:

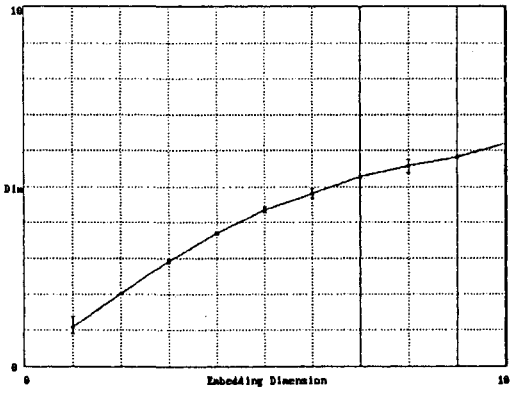


phase randomised (same power spectrum & corr. function, but diff. prob. distr.) ("softer")

time randomised



phase randomised

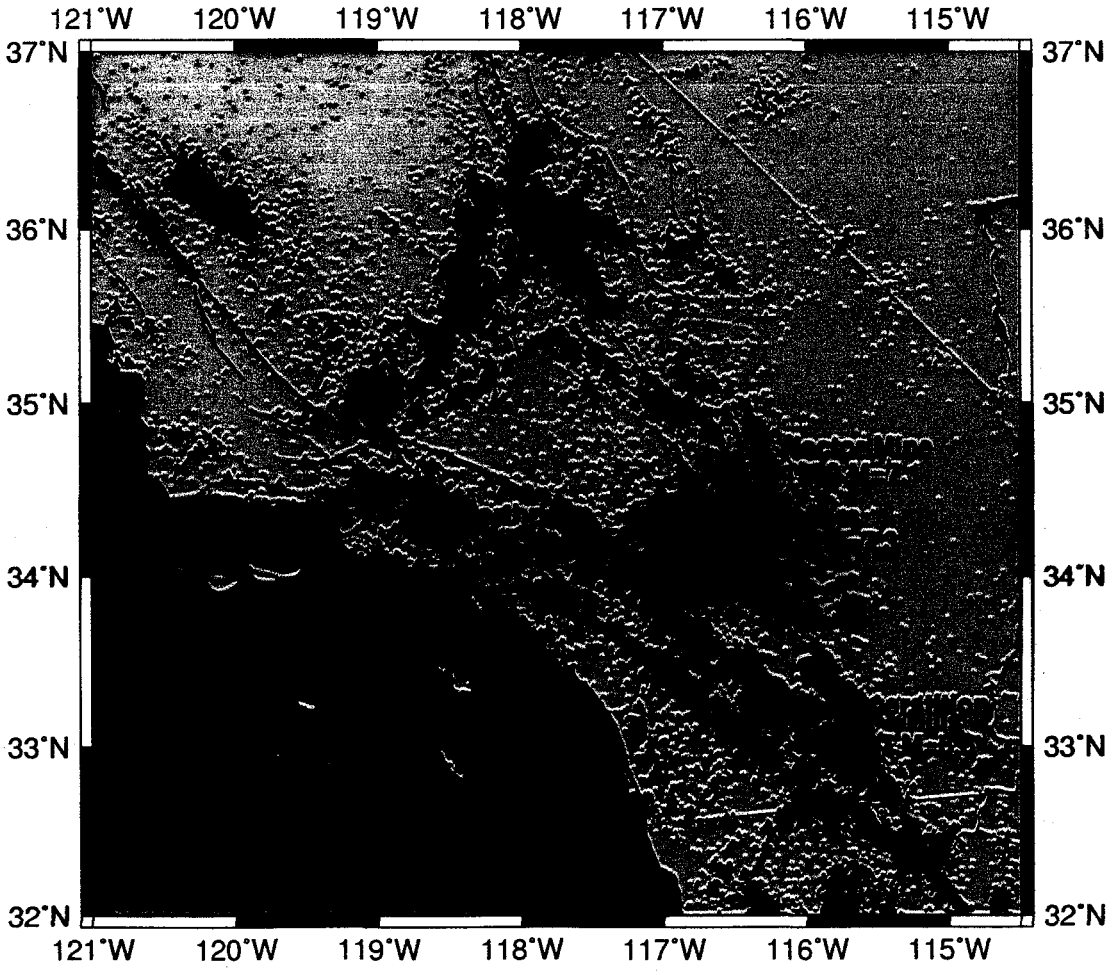


time randomised

Even the less powerful phase randomisation already completely destroys the structure in the data. Thus low-dimensional determinism in the data is ~~real~~ probably. Due to the power spectrum and the phase space and return plots, one is led to believe in chaotic dynamics.
The only argument against low-dimensional determinism in the analysed EQ inter-arrival times would be that the structure is introduced artificially by the seismological network or the evaluation process. Such an effect seems unlikely though.

Decomposing Spatio-Temporal Seismicity Patterns

- Introduction (Motivation, Problems)
- Fundamental New Approach: Principal Components Analysis
- Yes, but Spatio-Temporal? Incorporating Time
- Sample Application to Southern California
- Conclusion & Outlook



Motivation

"It is agreed that considerable information regarding the generation of large earthquakes is contained in seismicity patterns. The discussion is, however, how to extract it." M. Wyss 1999; Seismicity Patterns, their Statistical Significance and Physical Meaning; Birkhäuser

Here:
Seismicity pattern = distribution of points in space
Information = spatio-temporal evolution of patterns, specifically *earthquake rates*

Increasing quality and density of eq catalogues.

Recent developments in digital change detection techniques.

Goal

Untangle complexity of superimposed spatio-temporally coupled eq-processes.
Associate these fluctuations with the occurrence of large earthquakes.

Classical declustering methods need definition of a space-time window...
Example: popular Reasenberg (1985) algorithm,
Q determines spatial extent,
τ determines temporal extent.

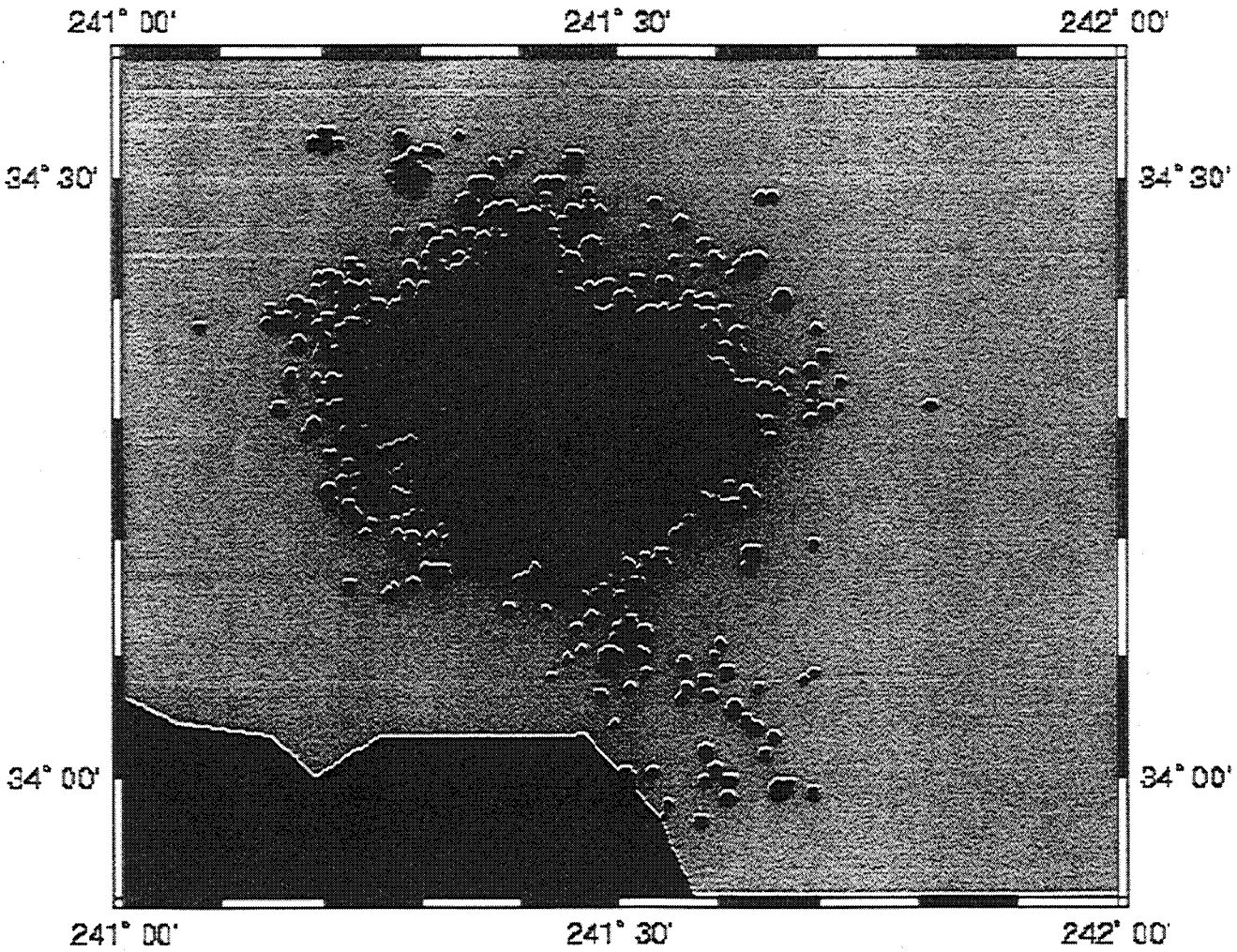
A reassuring example:

Northridge cluster

Q=20. Taumax=20.

Q=10. Taumax=10.

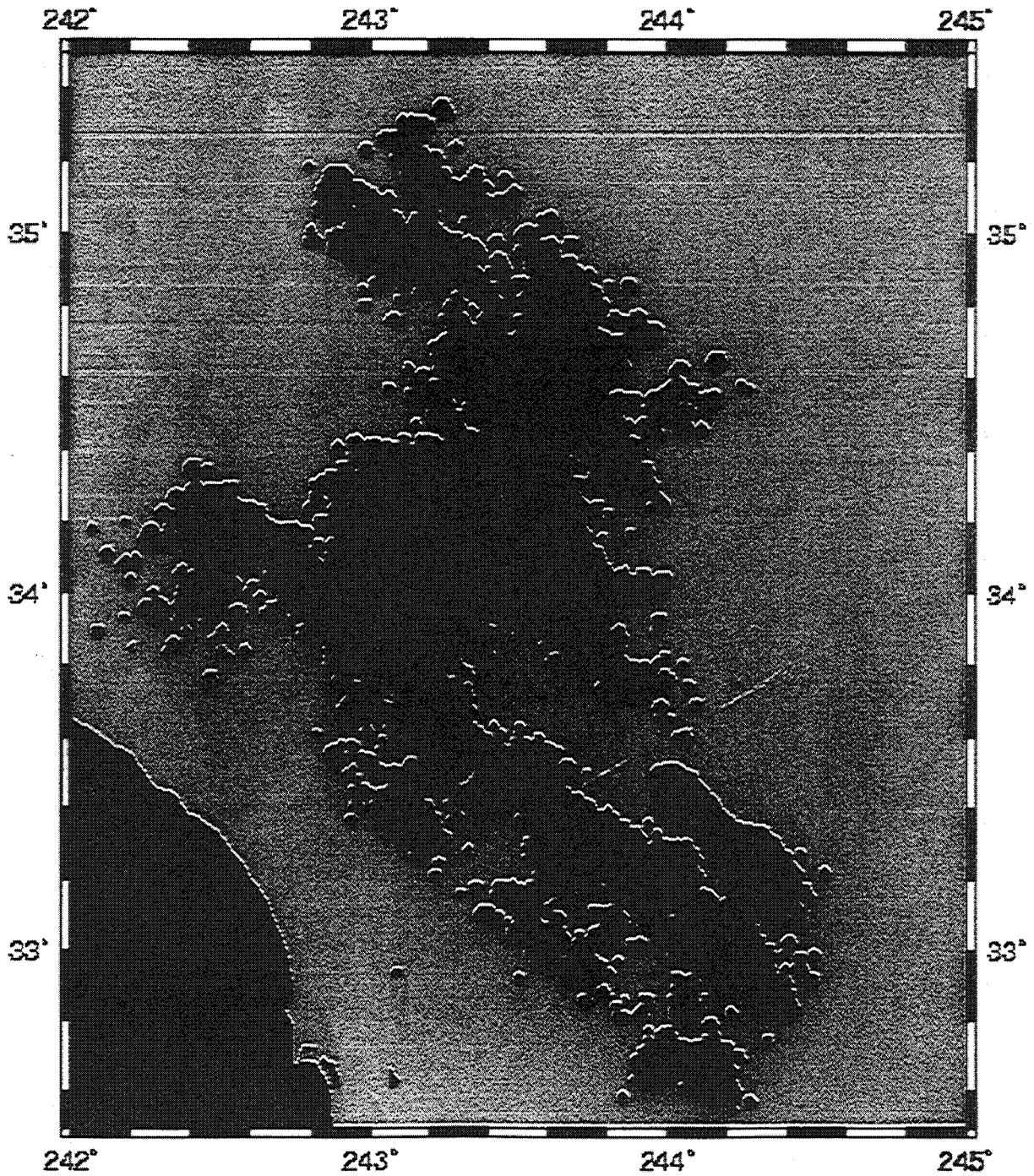
Q=2. Taumax=5.



A troubling example:

Landers cluster

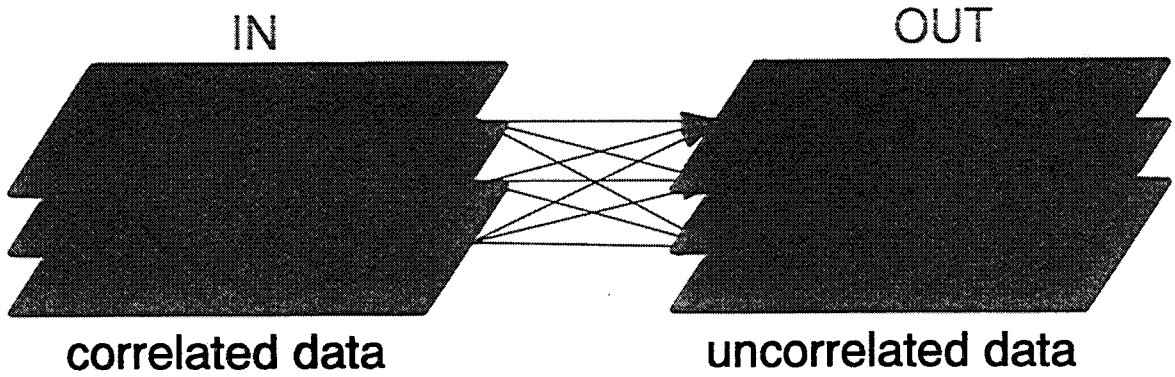
Q=20. Taumax=20.
Q=10. Taumax=10.
Q=2. Taumax=5.



Principal Components Analysis

(Karhunen-Loève-, Hotelling-Transform)

"*Input: Set of raster data sets*
Output: Mutually uncorrelated new matrices (components) explaining original variance"



- Covariance matrix **C** of original data.
- Obtain eigenvalues λ_i:

$$| \mathbf{C} - \lambda_i \mathbf{I} | = 0,$$

$i \in [1, 2, ..., n]$, **I** identity matrix.

- Obtain eigenvectors ε_i

$$(\mathbf{C} - \lambda_i \mathbf{I}) \varepsilon_i = 0$$

- Obtain component **PC** from

$$\mathbf{PC} = \mathbf{T}\ \mathbf{D},$$

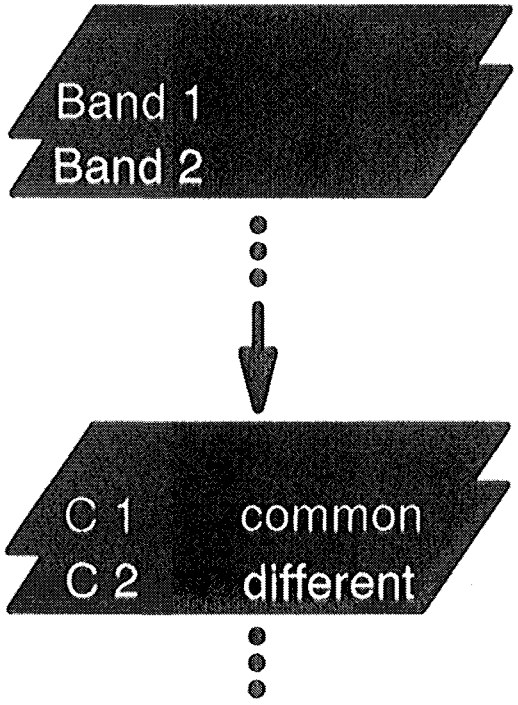
T is matrix of transposed ε_i, **D** = orig. data.

New components
- are uncorrelated (orthogonal)
- ordered by decreasing variance ("information")

C1 contains what is common in all input data sets, higher components will contain increasingly subtle differences between them!

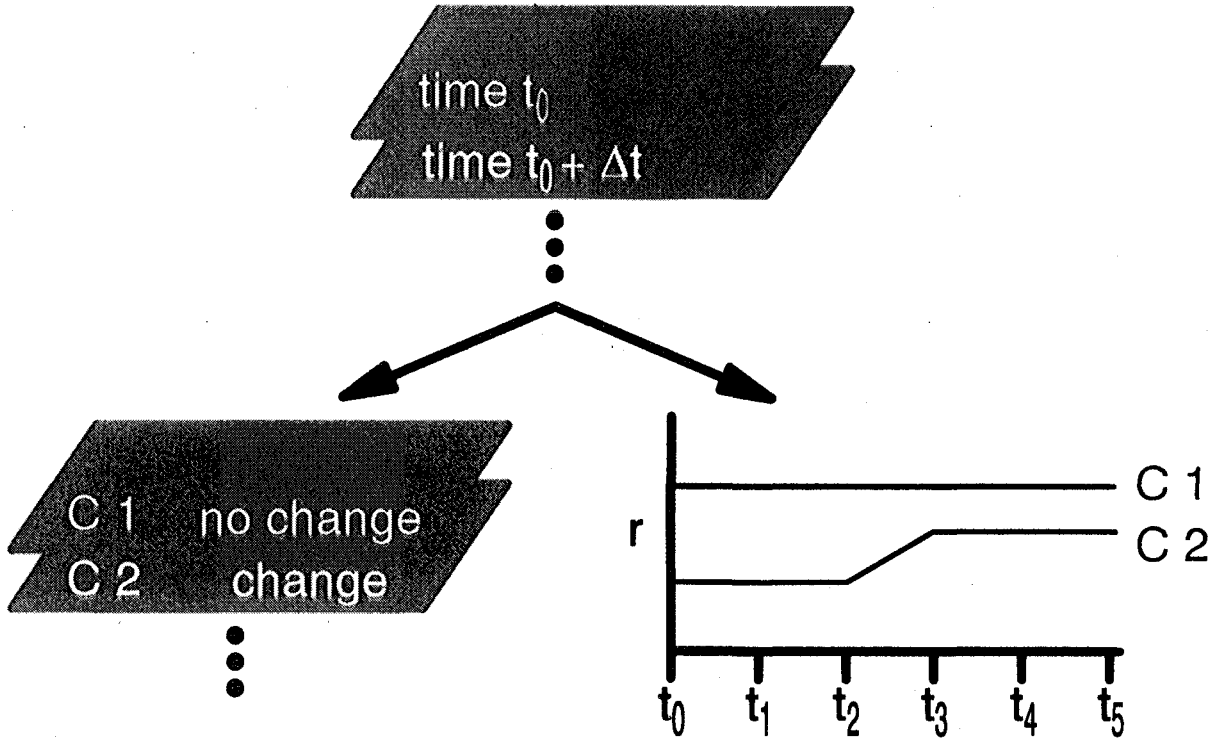
Incorporating Time

Standard PCA



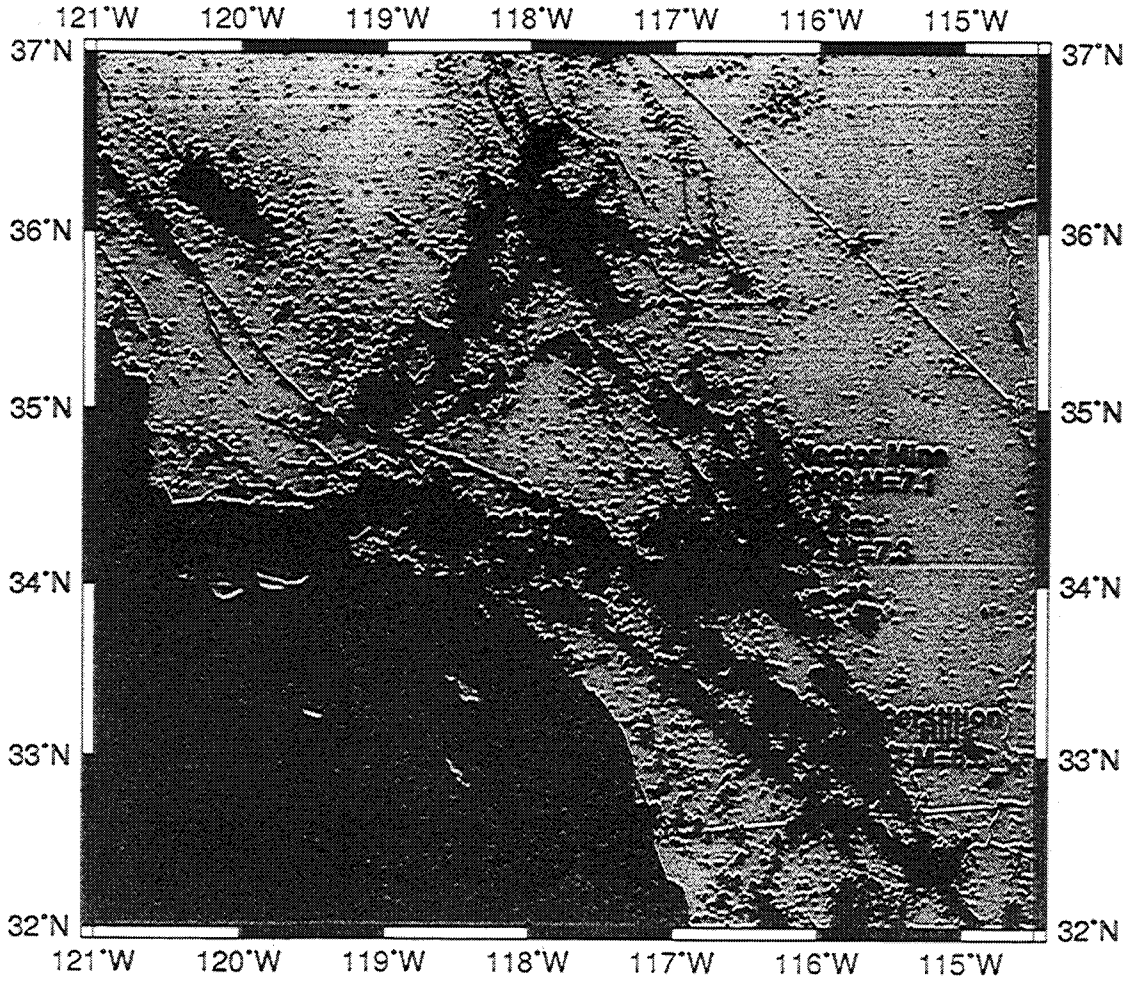
WHERE

"Change Analysis PCA"



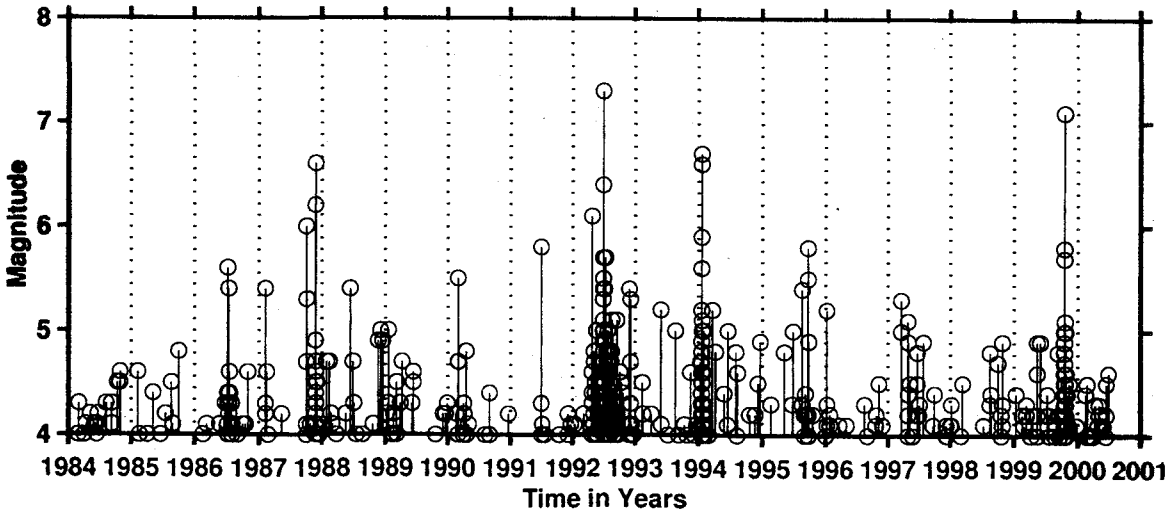
WHERE + **WHEN**

Sample Application



1984 - 1987, 1988 - 1991, 1992 - 1995, 1996 - 2000

SCEDC data 01/1984 - 07/2000
280 000 events complete for M >= 1.8



Accomodating Seismicity Patterns

Earthquake Catalogue

Select space-time volume of interest/feasible

Projection

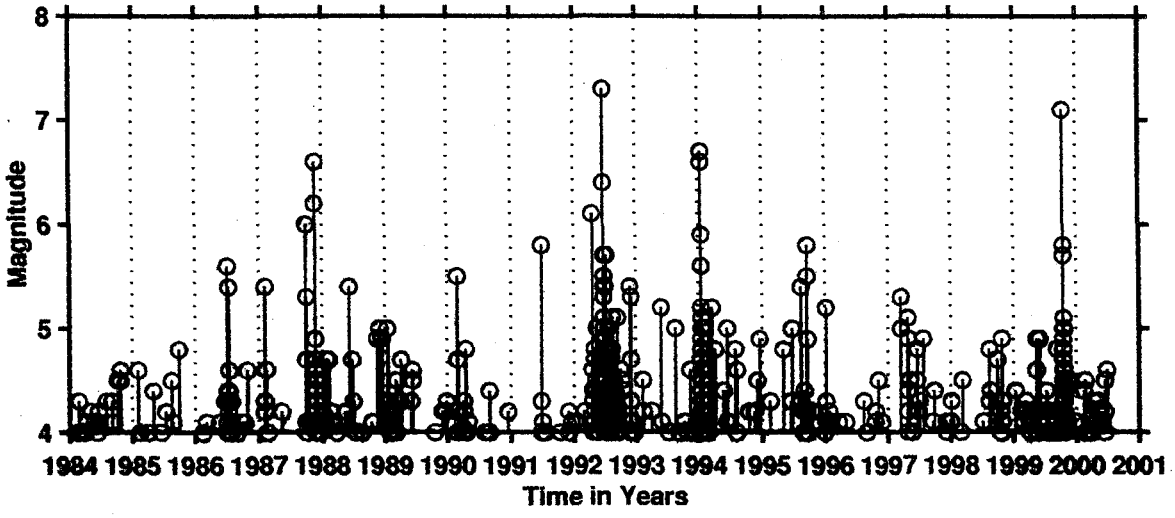
Decide on time slice duration & phase

Decide on spatial resolution of cells

Grid data by, e.g., determining cumulative number of eqs per cell (eq rates)

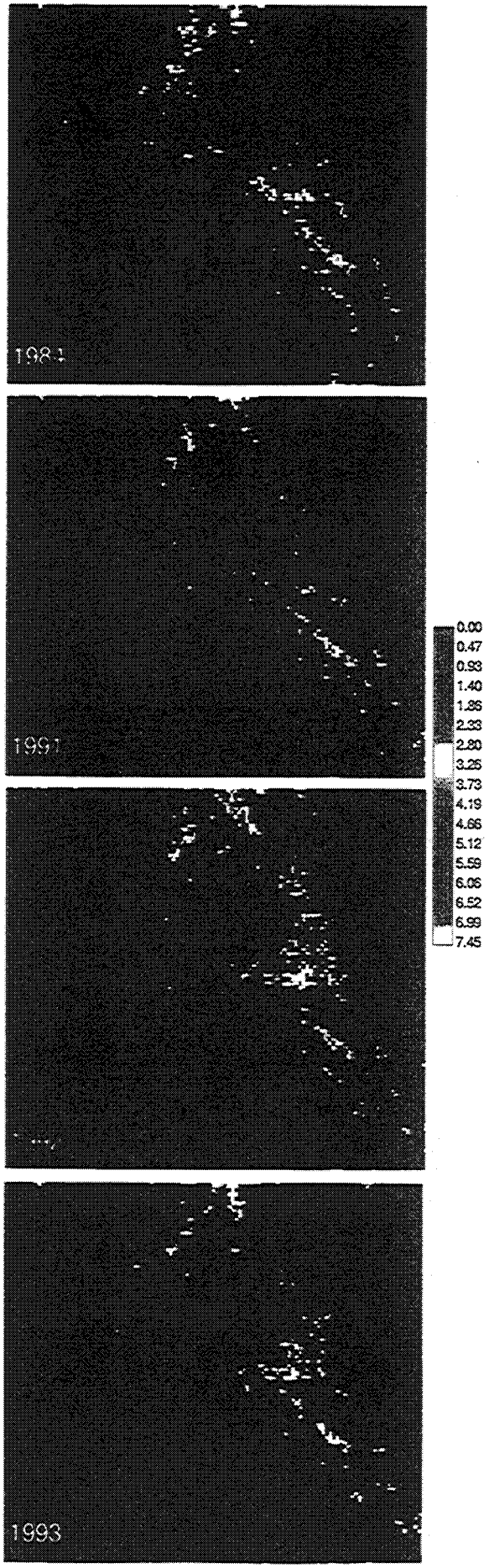
Do, e.g., log-transform to account for fractality of spatial frequency distribution of seismicity (Reduce variance due to change!)

$ts = ts(\ \Delta x, \Delta X, \Delta y, \Delta Y, \Delta Z, t_0, \Delta t, \Delta T, [\Delta M, ...])$



PCA is a straight-forward method from linear multivariate statistics, there are no ambiguities or tunable parameters. Only ts() allows tuning within numerical limits.

Producing Time Slices (Δt = 1a)



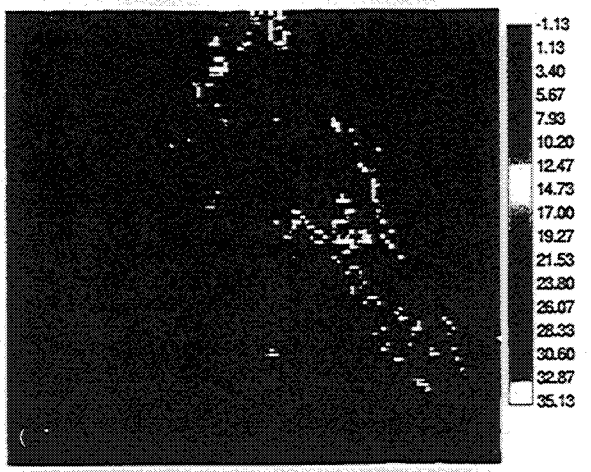
"a quiet year"

"before Landers"

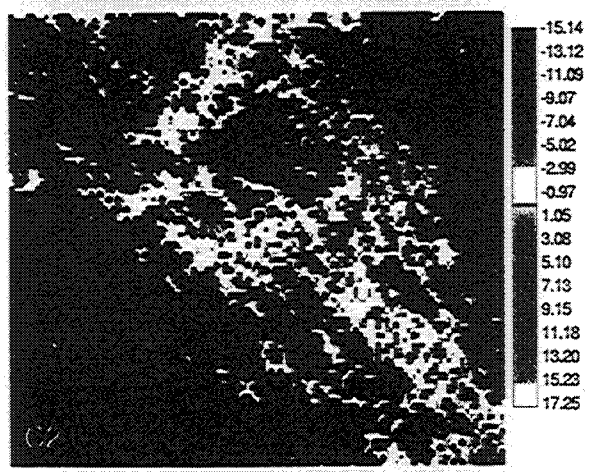
"Landers"

"after Landers"

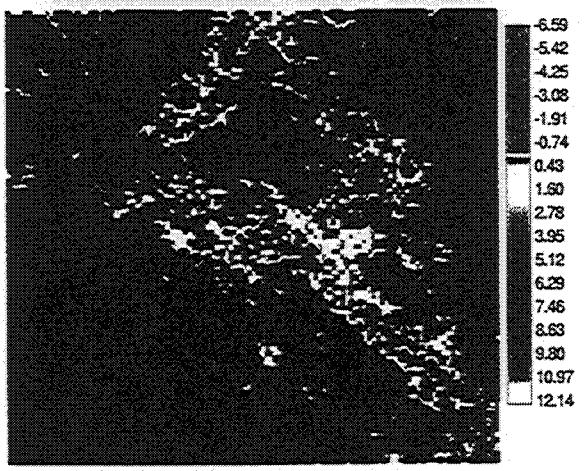
Results of PCA: Spatial



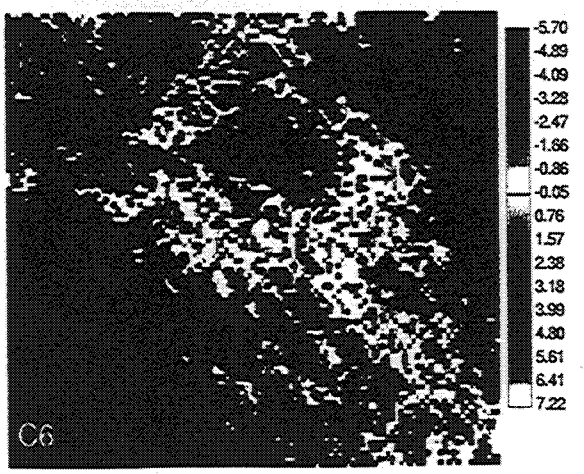
First (Principal) Comp.:
"Background Seismicity"
(61.5 % of total variance)



2nd:
First Comp. of (most dramatic) Change.
Landers, Hector Mine, Northridge.
(10.7 %)

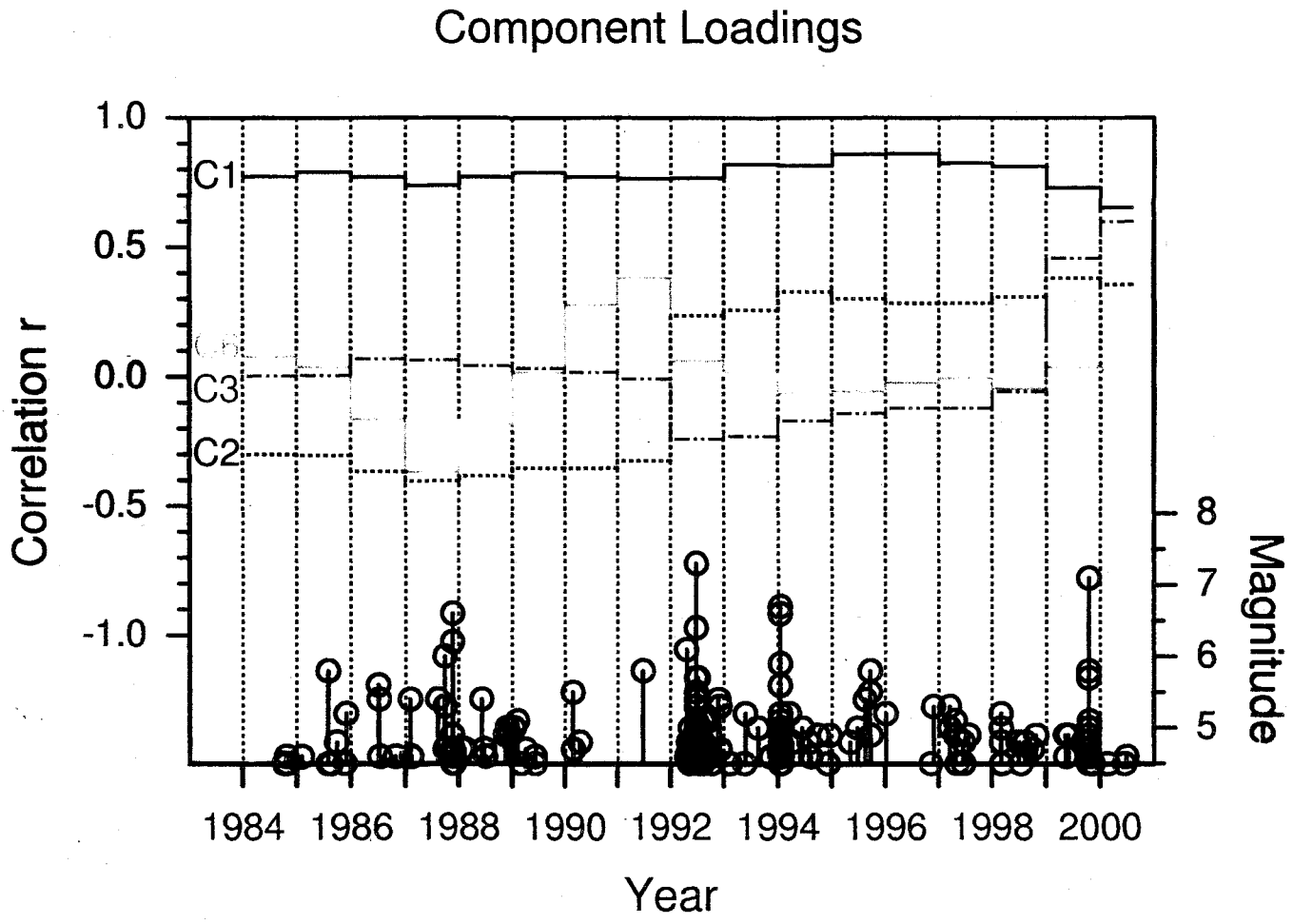


3rd:
2nd Comp. of (less dramatic) change.
Hector Mine
(4.5 %)

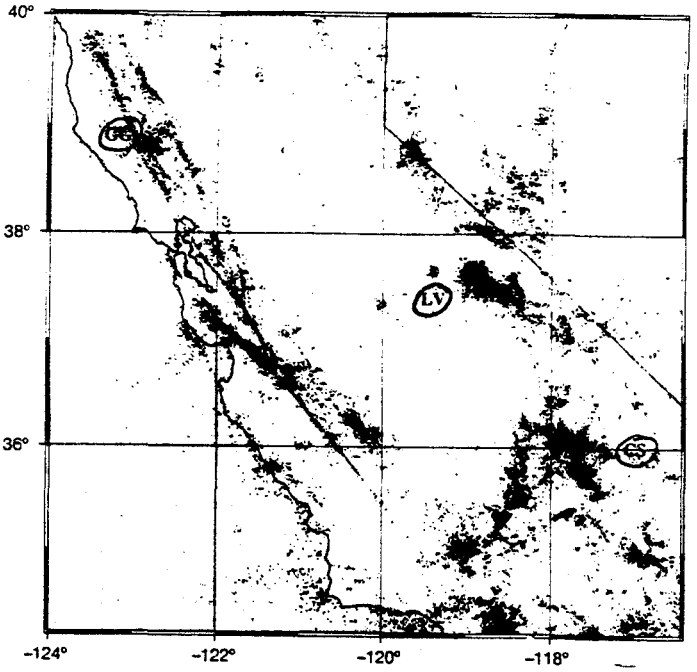


6th:
Subtle, distributed changes ...
(2.6 %)

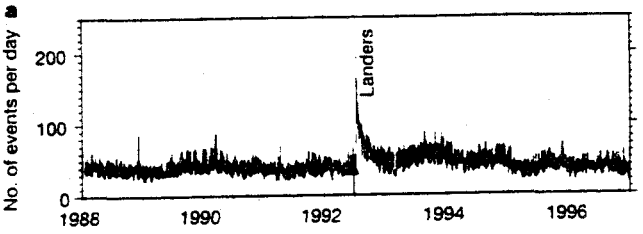
Results of PCA: Temporal



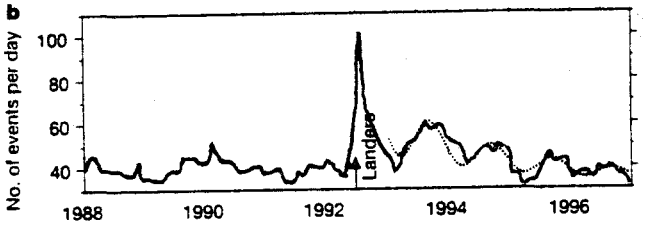
Gao et al, nature, 406, 500 - 504, 2000



geothermal/ volcanic



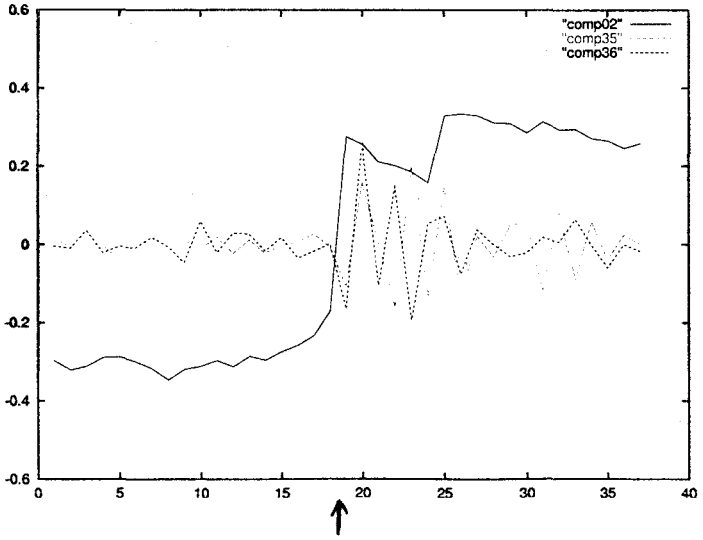
declustered (!?)



running mean + fit

PCA
(exactly same data, NOT declustered)

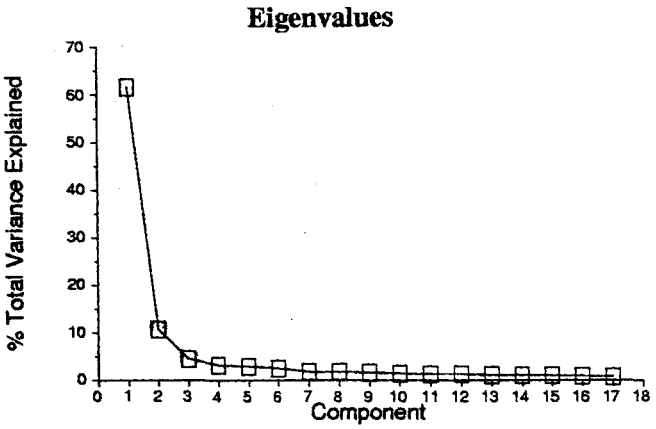
Δt = 3 months

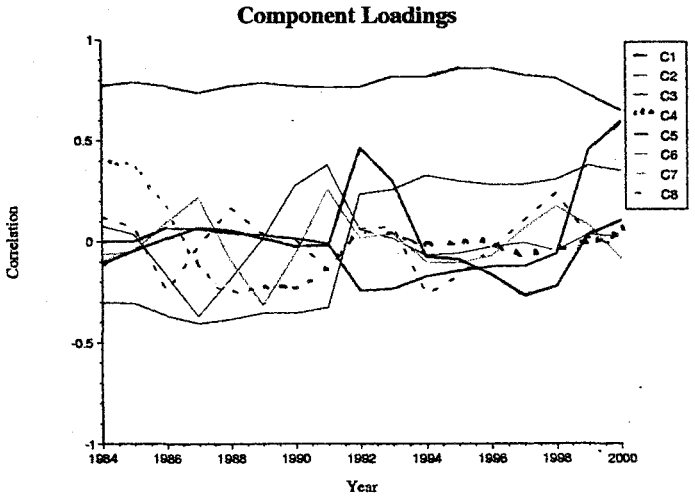


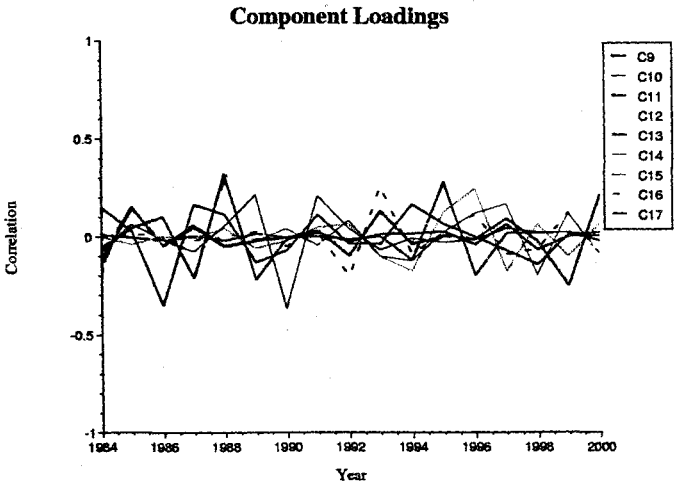
Conclusion & Outlook

- PCA can be applied to seismicity patterns (as a change analysis tool)
- It is possible to decompose distinct clusters as well as spatially distributed processes
- Very subtle changes can be detected
- Processes can be localised in space as well as time
- The method is simple and unambiguous
- Should be suitable for monitoring / detection of precursory patterns / network changes
- Suitable as a base / augmentation for other algorithms
- Supersedes classical declustering...

- Case studies, check precursor detection ability
- Check vs. surrogate data
- Try other measures in addition to eq rates
- Try oblique transform / ICA
- Postprocess results: Treshold / Classify / FCC ...


Eigenvalues
% Total Variance Explained
Component


Component Loadings
Correlation
Year
C1
C2
C3
C4
C5
C6
C7
C8


Component Loadings
Correlation
Year
C9
C10
C11
C12
C13
C14
C15
C16
C17