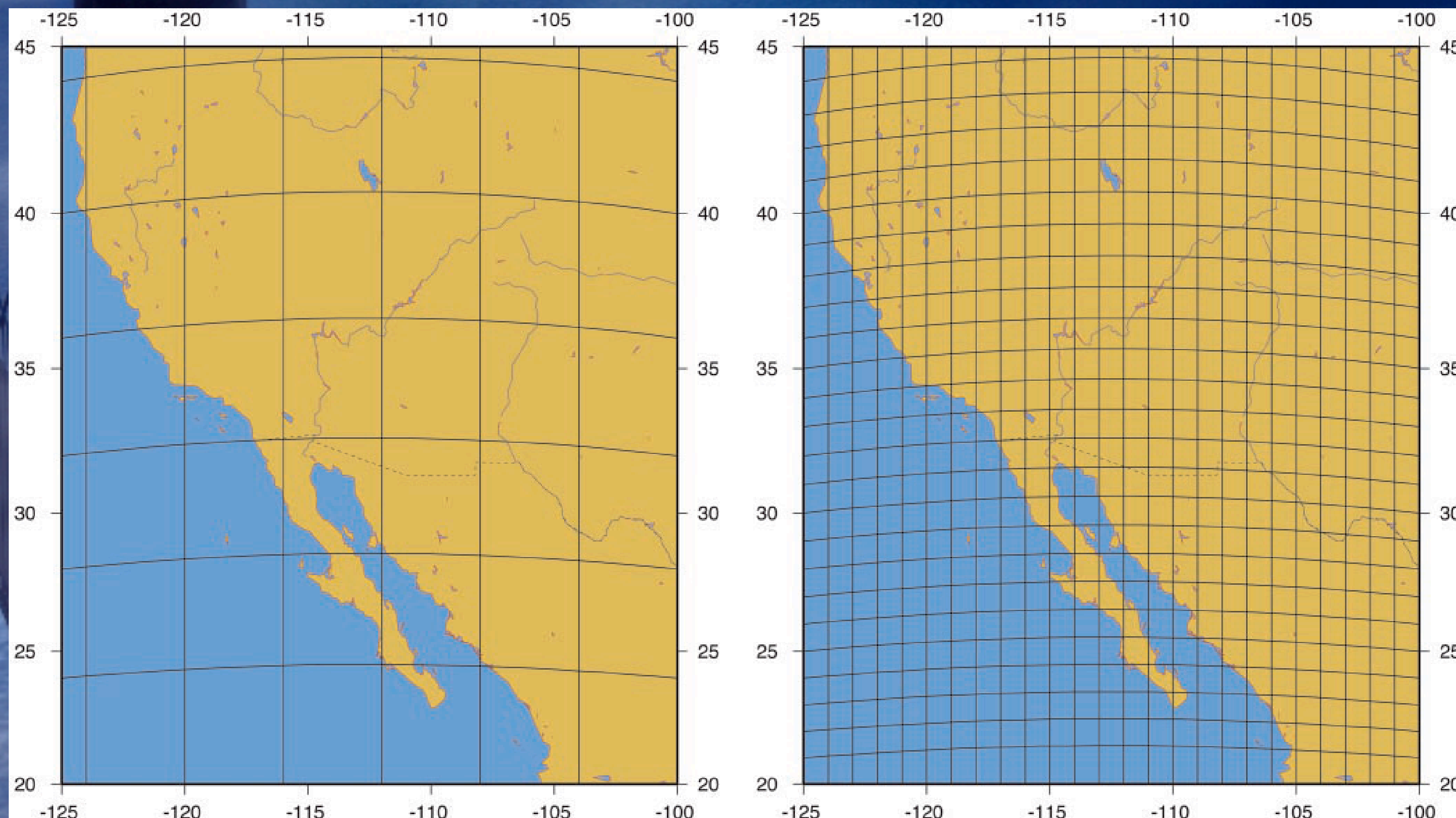
Networks, records, causality and earthquakes

Joern Davidsen
Peter Grassberger
Maya Paczuski

GRL (2006)
PRE (2008)

Background: Unified Scaling Law for Earthquakes (Bak et al. 2002)



$$P_{S,L}(T) \propto T^{-\alpha} f(T L^{d_f} S^{-b})$$

Motivations for this work

How to detect causal features (leading to e.g. clustering) in data sets making minimal a priori assumptions?

A mathematical theory of records where the record variable and time are put on an equal footing.

This gives a well defined acausal null model.

Application to Seismicity.

Large earthquakes can be dangerous

What's the point?

- earthquakes & records go together naturally:
- Earth quakes are so devastating!
- Will the next one be a record?
- This is NOT subject of present talk.
- What relationship is meant then?

Epistemology

- Use **ONLY** relations between events.
- Do not impose any scales.
- This includes NOT selecting large events as being more important than the others to begin with.
- A sparse but indefinite number of causal predecessors (a network of earthquakes).

Two approaches so far…

- Two approaches: networks of earthquakes and aftershocks (M. Baiesi & MP: PRE (2004); Nonlin. Proc. Geophys. (2005)
- Earthquakes as a record breaking process.

Our definition of records

- Assume there was an earthquake some time ago in Erice.
- then in Sidney ...
- then in Calgary ...
- then in Cologne ...
- then in Rome ...
- then in Palermo...
- then again in Erice…
- Each successive event in this sequence is considered a ”record”, as seen from Erice: a record in closeness

But why should Erice (or any other place) be special?

- Because we assumed that there was one in Erice before.
- ONE earthquake might happen by chance, but TWO?
- Why not? If the first happened a million years ago?
- And if the epicenter is next street?
- Hmm – what about five million years?
- And if the epicenter is under the next building? And why five million years? Why not two? or ten?
- —→ Successive records in closeness have more chance to be causally connected than earthquakes chosen at random

Formally

A,B,C, . . .: space-time events

Question: under which conditions is it likely that pair of events (A,B) are causally related?

(not via a chain of dependencies)

Finite memory: close in time

Finite interaction range: close in space

Finite speed of propagation: not too close in time, unless very close in space

—→ all depends on model details

Least Model Dependent Approach

- Assume $t_A < t_B$.
- Then A,B are likely to be causally related, if there was no other event C at intermediate time $t_A < t_C < t_B$, which was also closer to A than B, or no C such that

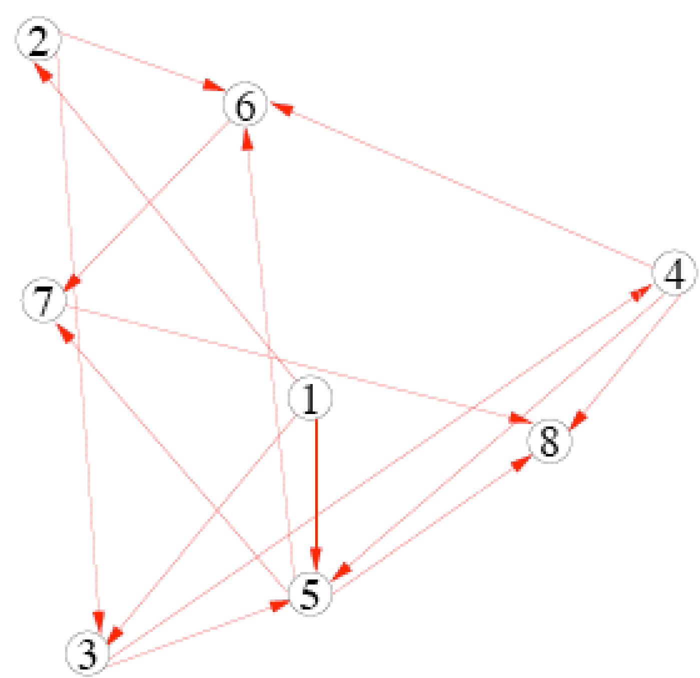
$$d_{AC} \equiv |x_C - x_A| < |x_B - x_A| \equiv d_{AB}.$$

- In that case: B is new record in closeness to A, or B is "recurrence" of A
- Draw directed link between any two events A,B, if B is recurrence of A
- $- \rightarrow$ directed causal network of recurrences or records

An example

- Each node has an in-degree k_{in} and an out-degree k_{out}
- Zero'th order assignment of causes and effects.
- Record property is preserved under addition of new nodes with time.

E.g. eight events $A_1 \ldots A_8$ with $t_1 < t_2 < \ldots t_8$:



A null model: random processes give acausal networks

- iid distributed events: the joint pdf

$$\rho_n(\mathbf{x}_1, t_1; \ldots \mathbf{x}_n, t_n) = \prod_{i=1}^{n} \rho_1(\mathbf{x}_i, t_i)$$

- single-event distributions factorize:

$$\rho_1(\mathbf{x}, t) = \rho_x(\mathbf{x})\rho_t(t)$$

→ simplest theory of records

→ explicit results for any acausal network: significant statistical differences in the real data set are due to causality!

Notice:

- Distributions of spatial & temporal distances from recurrence-defining event $A_0 = (\mathrm{x}_0, t_0)$ are often more useful; if no integration over A_0 is made, then distance distributions also factorize, with single-distance distributions $\mu_l(l)$ and $\mu_t(t)$. The latter still depend on A_0, in general.

- Assumptions of stationarity / translation invariance are not needed, if distances are transformed according to:

$$l \to \xi = \int_0^l dl' \mu_l(l') \;, \qquad t \to \tau = \int_0^t dt' \mu_t(t') \;. \tag{1}$$

- The new ("canonical") variables ξ, τ are – just as l and t – random variables. Their densities are either constant or step functions with height 1, such that the integrals over the densities are either finite or infinite. Which case is realized depends on the specific problem.

- One has thus four cases. Denoting

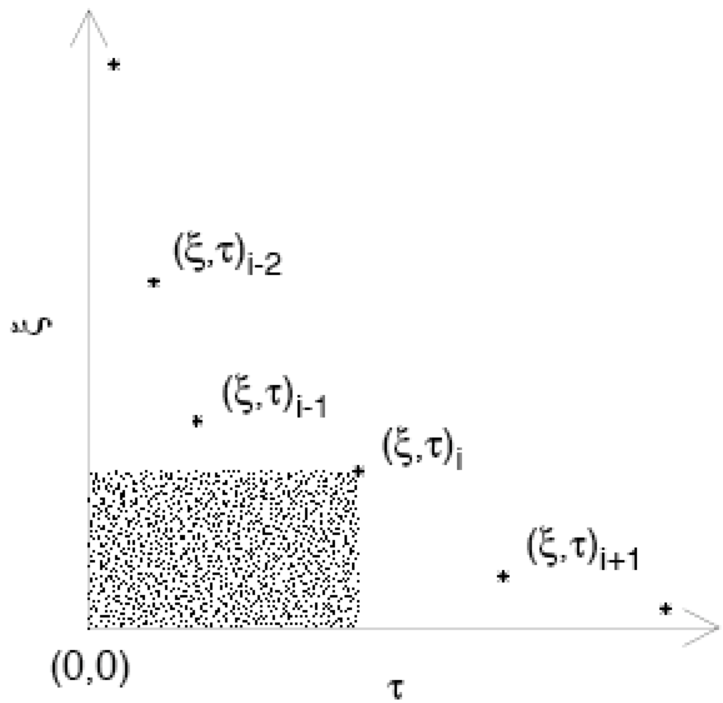
$$\lambda = \int_0^\infty dl \mu_l(l) \ , \qquad \sigma = \int_0^\infty dt \mu_t(t), \tag{2}$$

one can have

1. integrable spacial distance distributions ($\lambda < \infty$), and non-integrable temporal distance distributions ($\sigma = \infty$). This is typically the case for events happening in a finite volume, but stationary in infinite time.
2. $\lambda = \infty$ but $\sigma < \infty$. This might result from infinite space, or from events clustering in space so strongly, that their distance distribution is non-integrable. In contrast, stationarity is broken by a cutoff which renders the time differences integrable.
3. & 4. the other two cases are analogous.

A theory of records

Typical sequence of recurrences (reference event A_0 is at orign):



Event (ξ_i, τ_i) is recurrence, if no other event in shaded region;
$\longrightarrow$ density of events: $p(\xi, \tau) = e^{-\xi\tau}$.

The four different cases: distinguished by different integration regions:

- unrestricted: recurrence (record) density is equal (exactly!) to $1/\tau$ resp. $1/\xi$. Densities in original coordinates are obtained by inverting the transformation to canonical coordinates.
  E.g. stationary, space = R^D: $p_t(t) = 1/t,\ p_l(l) = D/l$, independent of event rates!

- resticted (integrable event densities): more complicated, but still analytically easy:

$$p_\tau(\tau) = \int_0^\lambda d\xi p(\xi, \tau) = \tau^{-1}(1 - e^{\lambda\tau}),$$

  and same (with $\lambda \leftrightarrow \sigma$) for ξ.
  Original coordinates, R^D : $\rightarrow$ characteristic length (time)

scales l^*, t^***, e.g.** TDl^{D-1}

$$p_l(l) = \begin{cases} \text{event density} & \textbf{for } l \ll l^*(T),\ l < R\ , \\ D/l & \textbf{for } l \gg l^*(T),\ l < R\ , \\ 0 & \textbf{for } l > R\ , \end{cases} \qquad (\text{:}$$

If N, L, T **are** $\{$ **# of events, system size, duration** $\}$**:**
$l^* = L/N,\ \ t^* = T/N$**.**

Correlations between recurrences:

$\rightarrow$ **multidimensional integrals over exponentials & step functions**

E.g.:

$$\begin{aligned} q(t,t') \equiv & \textbf{ prob. density}\{ \textbf{ two } \textit{successive} \textbf{ events at time} \\ & \textbf{distances } t > t' \textbf{ from } A_0 \textbf{ are } \textit{both} \textbf{ recurrences}\} \\ = & \ 1/t^2 \ \textbf{ plus correction terms for finite event rate } \lambda. \end{aligned}$$

Similar results for

- $p(t,t') =$ prob. density{ any two events at $t > t'$ are recurrences}
- spatial correlations
- distributions of fixed rank recurrences
- $m-$recurrences (similar to $m-$records: there are only m other events which are "better")

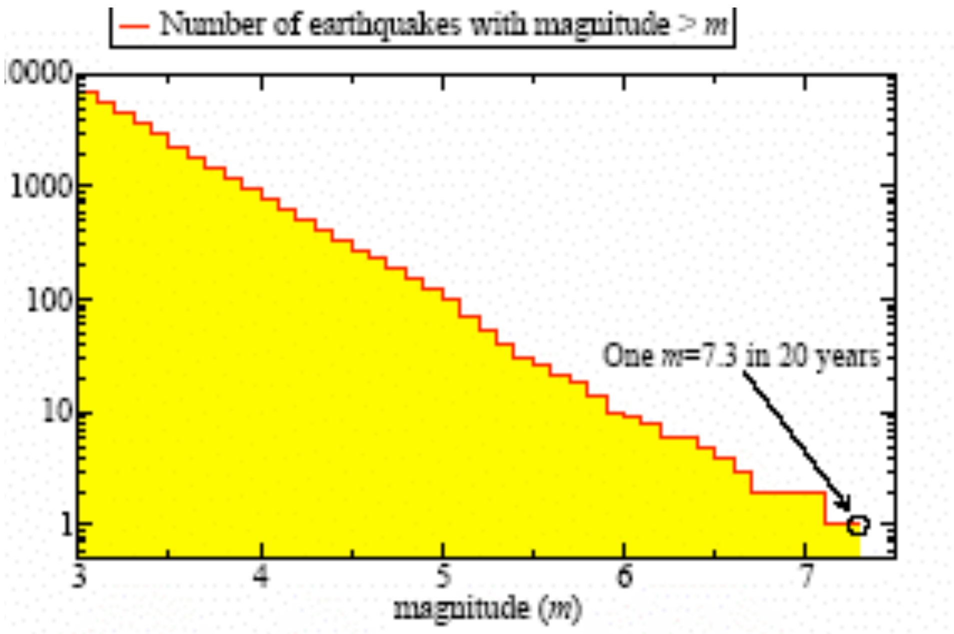
Network properties for random events in space-time

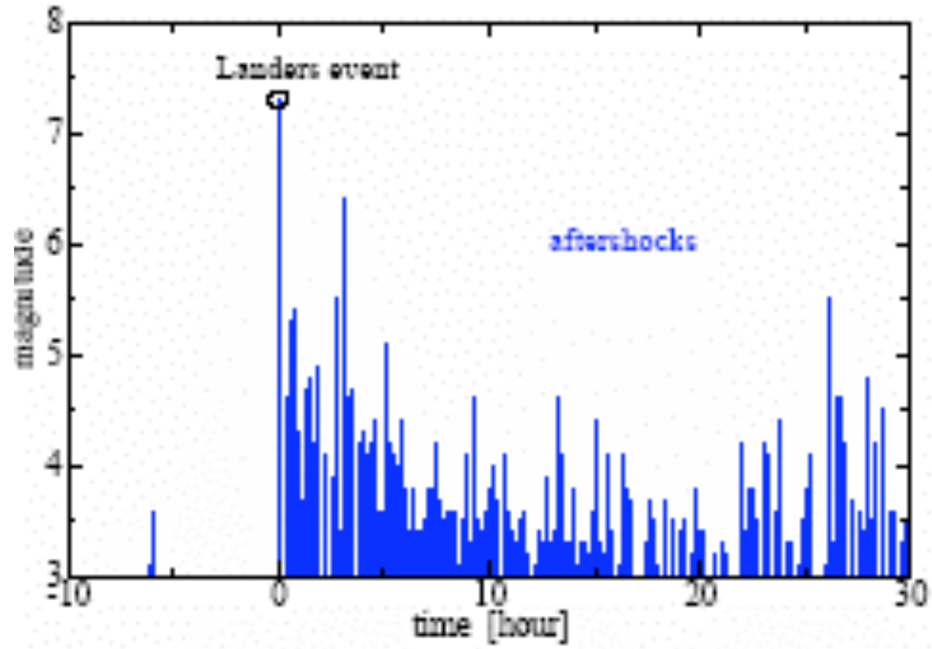
$$
\begin{aligned}
P_i(k) &= \frac{1}{i} \sum_{1<l_1<\cdots<l_{k-1}\leq i} \frac{1}{(l_1-1)\cdots(l_{k-1}-1)} \\
&= \frac{|S_i^k|}{i!} \\
&\approx \frac{(\ln i)^{k-1}}{i(k-1)!} \quad ,
\end{aligned}
$$

$$
P^{out}(k,N) \approx \frac{1}{N} \sum_{i=1}^{N-1} \frac{(\ln i)^{k-1}}{i(k-1)!} \approx \frac{(\ln(N))^k}{N\ k!}
$$

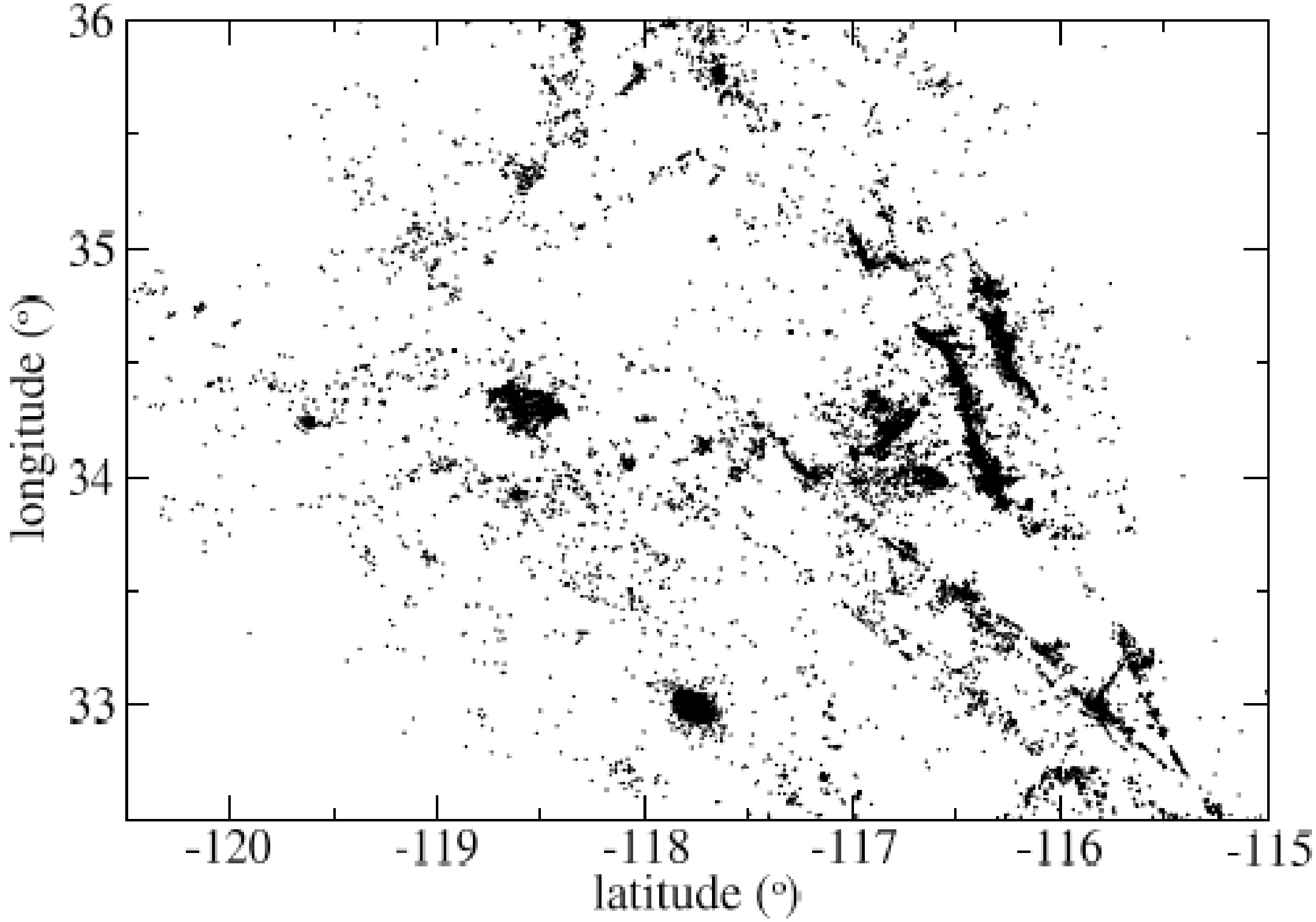
$$
\begin{aligned}
P_i(k^{in},k^{out}) &= P_{N-i}(k^{out}) P_i^{in}(k^{in}) \\
&\approx \frac{(\ln(N-i))^{k^{out}-1}}{(N-i)(k^{out}-1)!} \frac{(\ln(i))^{k^{in}-1}}{i(k^{in}-1)!}
\end{aligned}
$$

A complex spatiotemporal phenomena





Earthquakes in Space



Degree distributions for the network of records

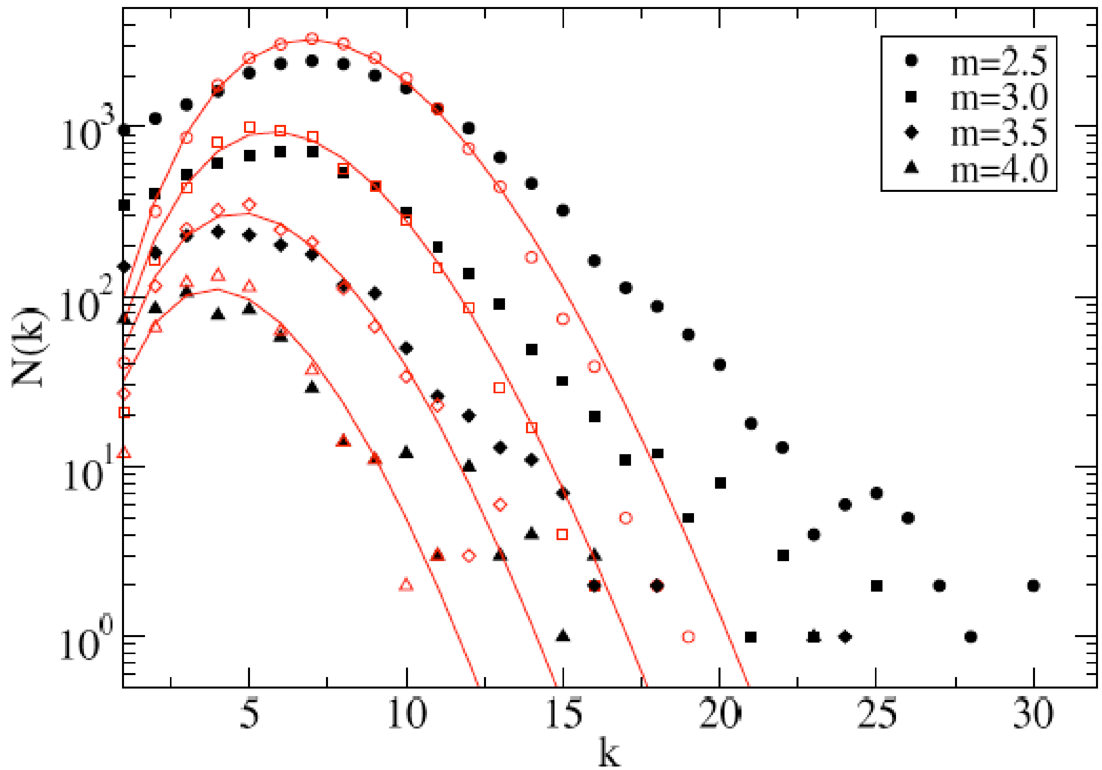


FIG. 8: (Color online) In- and out-degree histograms for different values of m. **For a given earthquake, the in-degree (out-degree) k is the number of links directed at it (originating from it) as defined in Section II.** Open (red) symbols correspond to the in-degree, filled (black) symbols correspond to the out-degree. Error bars can be estimated as $\sqrt{N(k)}$. The red lines correspond to Poisson distributions with the same respective mean and normalization.

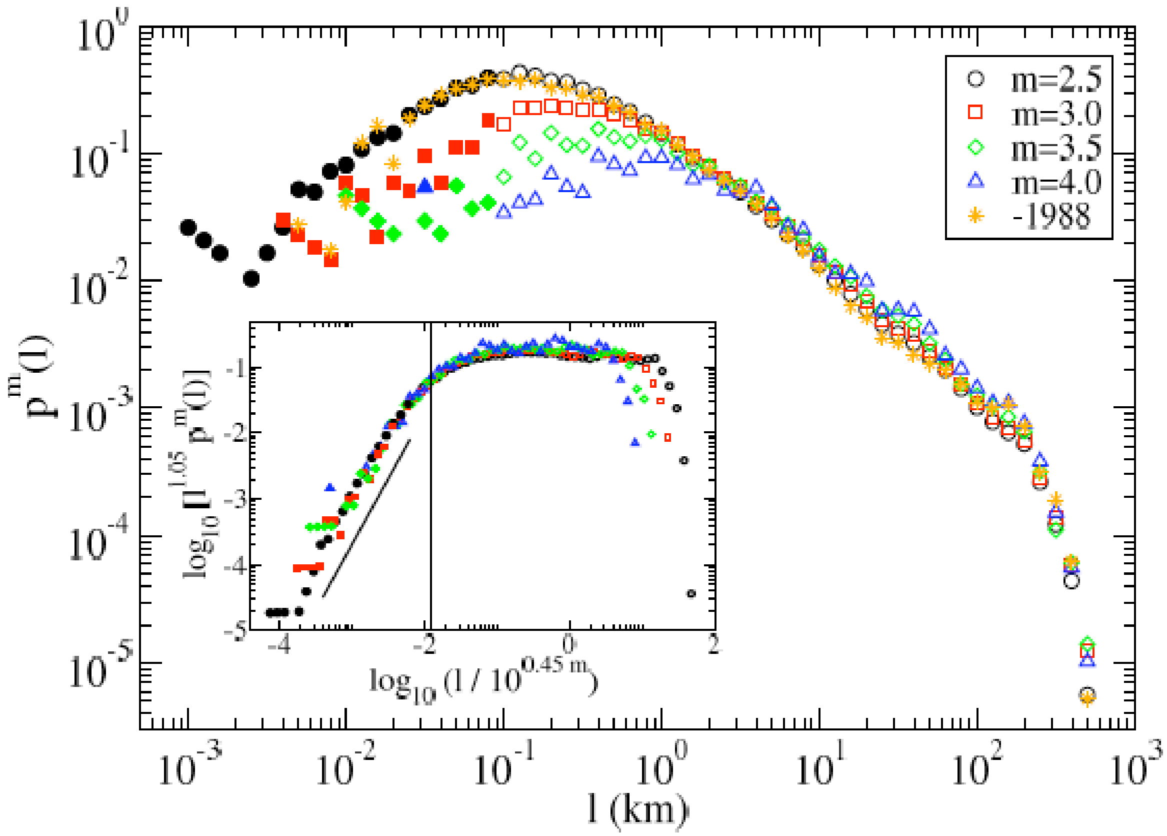
More Signatures of Causality: Recurrence length distribution is stationary and reveals a length scale: the rupture length

$$l^*(m) \approx L_0 \times 10^{0.45m}$$

$$L_0 = 0.012\text{km}.$$

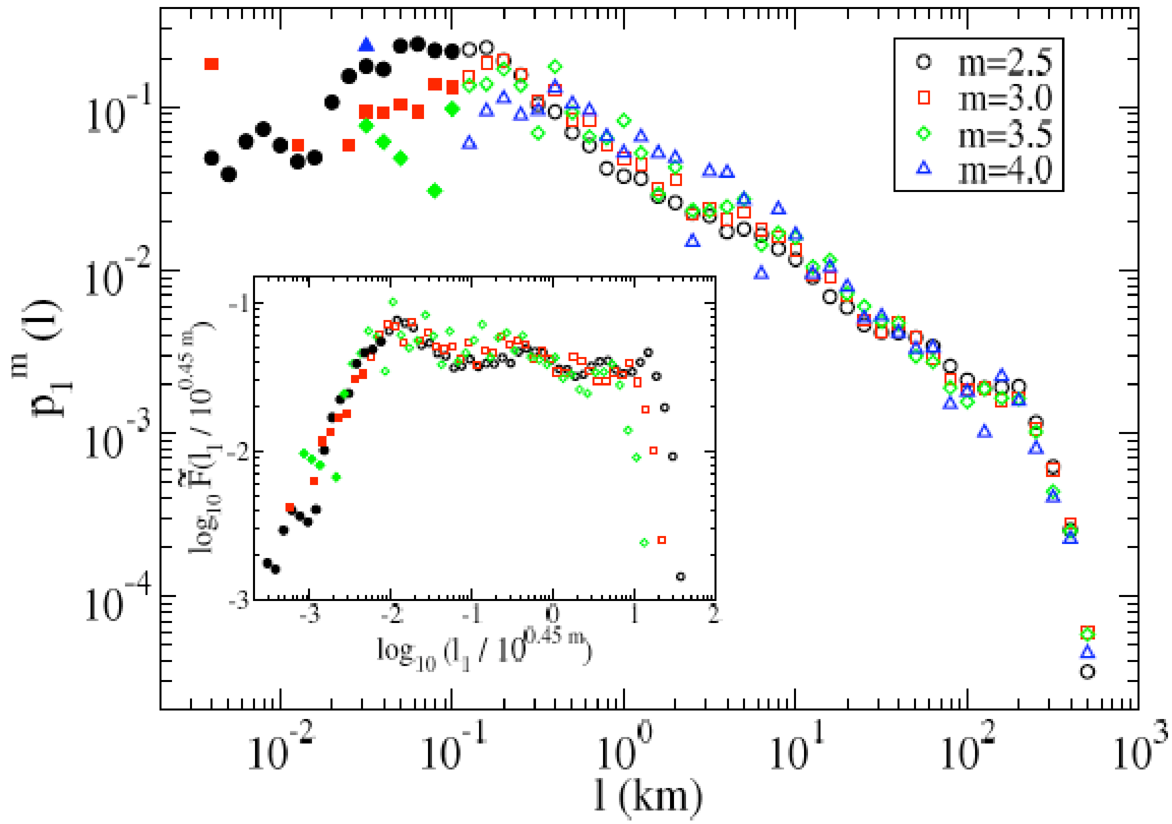
l(m)* is the position of the maximum. It is independent of T whereas in null model

$$l^*(T) \equiv (abT)^{-1/D}$$



Distribution of first recurrences also gives the same length scale

- In contrast, the null model gives a monotonic increasing function

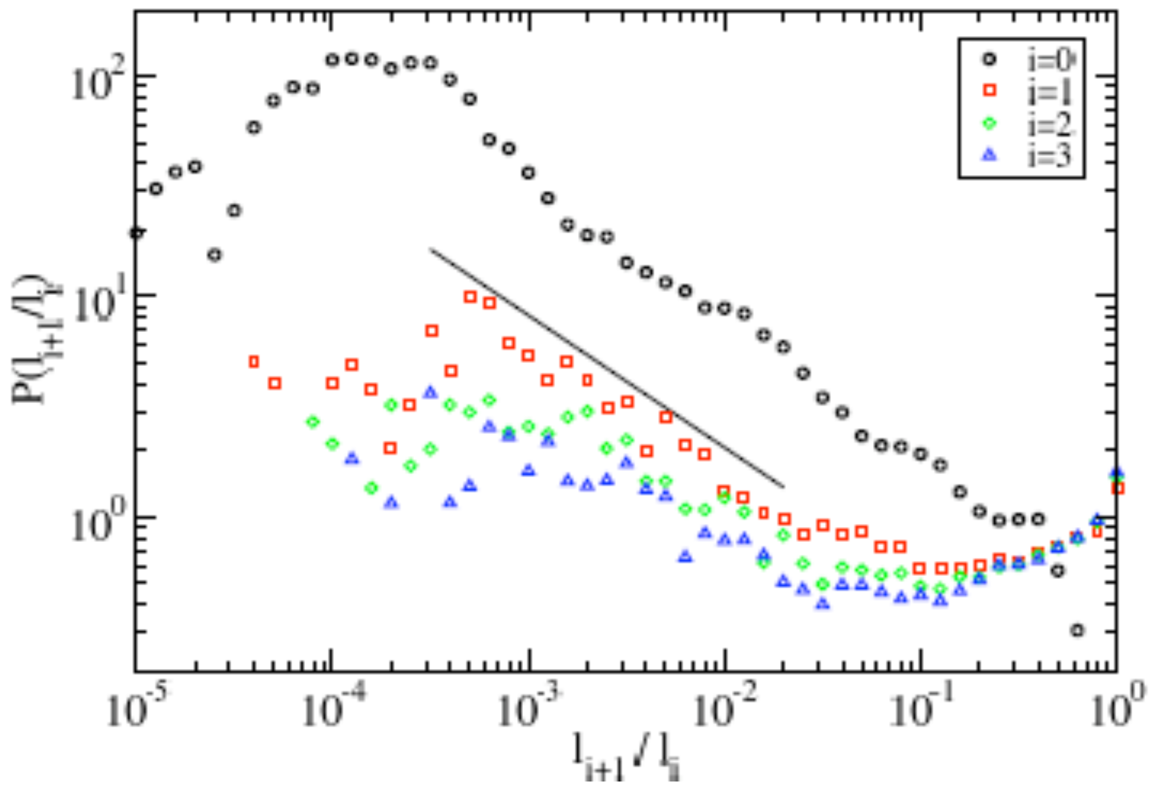


first recurrence is much more likely to happen at a typical distance of l^* than predicted by the null model. This enhancement goes along with a suppression of recurrences with $l \ll l^*$

Signatures of causality: Hierarchy of scales in the cascade of recurrences

•In contrast, the null model is independent of i and only increasing

$$q_i^m(x) = Dx^{D-1}$$



Conclusions

- Benchmark tests for models of seismicity or other processes.
- Application to other phenomena.
- Generalization to more than one event at a time??

Complexity Science at Calgary

www.phas.ucalgary.ca/complexity

IBI

UNIVERSITY OF CALGARY

Institute for Biocomplexity and Informatics

Institute for Quantum Information Science at the University of Calgary

INSTITUTE FOR SPACE RESEARCH

Located in the scenic foothills of the Canadian Rockies, the city of Calgary is a vibrant and growing metropolitan centre of approximately one million residents. In addition to enjoying the benefits of major city living, including professional sports, concerts, ballet, and opera, Calgarians are less than a 2 hour drive to Kananaskis, Banff, Lake Louise, and the Tyrrell Museum of Paleontology.

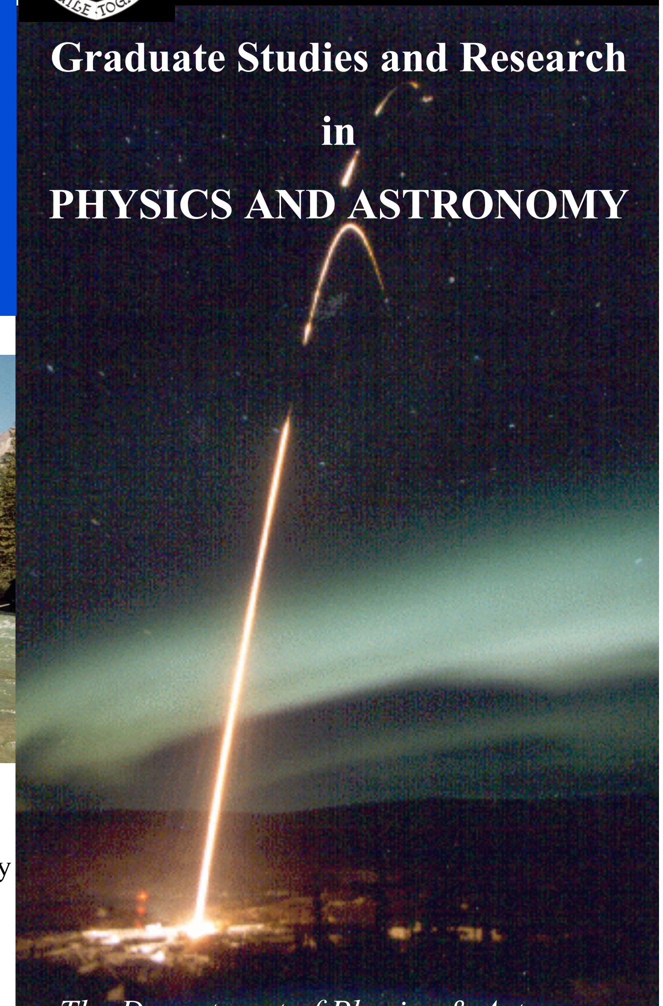
THE UNIVERSITY OF CALGARY

MO SHÙILE TOGAM SUAS

Graduate Studies and Research

in

PHYSICS AND ASTRONOMY



The Department of Physics & Astronomy is a young, dynamic department with an exciting future. We invite you to consider graduate studies in Physics, Astrophysics, Space Physics, or Medical Physics at the University of Calgary.

http://phas.ucalgary.ca

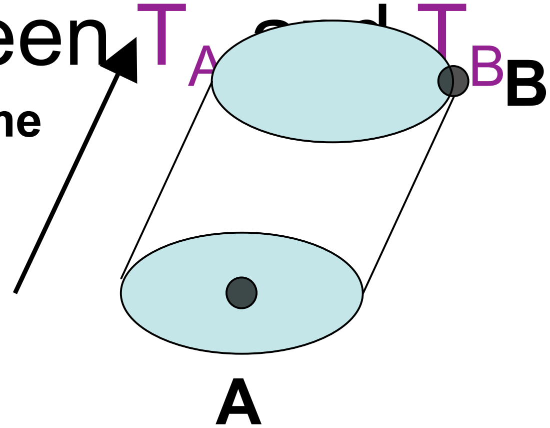
Analytic results for a random process on a fractal of dimension D

$P(T)dT$ is the probability that a recurrence occurs in the time interval $[T - dT, T]$ after the defining event. If a recurrence occurs in this interval, then this is necessarily the closest event in space during the entire interval $[0, T]$. Since the process is stationary, the probability for this to happen is dT/T, and $P(T) = 1/T$.

The probability density $P(l)$ that a recurrence occurs at a spatial distance l from the defining event is similarly obtained: To be a recurrence, an event within distance l from the defining one must be the closest in time. The chance that the closest event in time is within in a distance interval $[l - dl, l]$ is Ddl/l, and thus $P(l) = D/l$.

Earthquakes as a Record Breaking Process (Davidsen, Grassberger & MP (2005))

- Event B is a recurrence of a previous event A if no intervening event happened in the spatial disk centered on A of radius <u>AB</u> in the time interval between T_A and T_B



B is the closest in space to A up to that time; it is a *record.* Link all pairs of recur-rences.

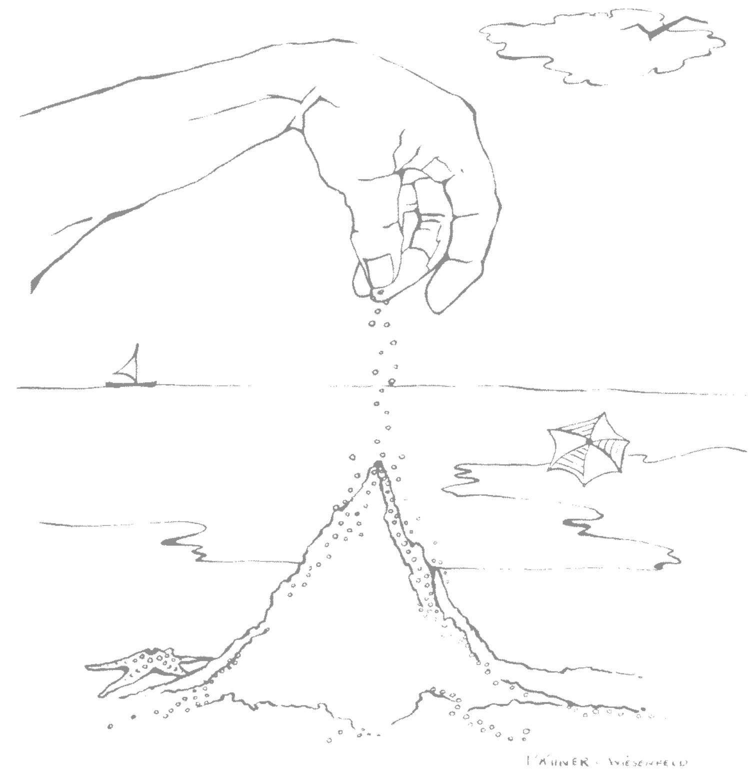
Analytic results for a random process on a fractal of dimension D

$P(T)dT$ is the probability that a recurrence occurs in the time interval $[T - dT, T]$ after the defining event. If a recurrence occurs in this interval, then this is necessarily the closest event in space during the entire interval $[0, T]$. Since the process is stationary, the probability for this to happen is dT/T , and $P(T) = 1/T$.

The probability density $P(l)$ that a recurrence occurs at a spatial distance l from the defining event is similarly obtained: To be a recurrence, an event within distance l from the defining one must be the closest in time. The chance that the closest event in time is within in a distance interval $[l - dl, l]$ is Ddl/l, and thus $P(l) = D/l$.

Self Organized Criticality

- Avalanches with a power law distribution
- Correlations over many space and time scales
- Solves ‘fine-tuning’ problems
- Fundamental parameters are emergent



Robust & universal mech-
anism ⇨ simple models

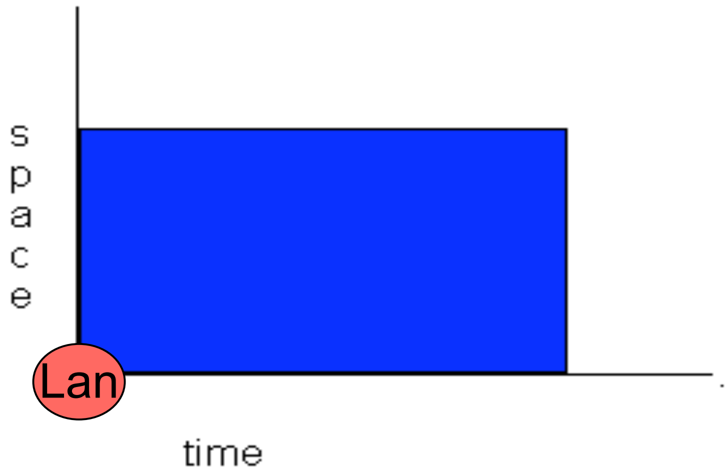
P. Bak, C. Tang and K. Wiesenfeld (’87)

A viewpoint

- Geophysicist Yan Y. Kagan of the University of California at Los Angeles agrees :

"the distinction between aftershocks and main shocks is relative." Within slowly changing continental areas, aftershocks can rumble on for centuries."

Standard Method: Space-time windows



Events in the window may not be correlated to mainshock, or events outside may be correlated.

How to estimate errors?

Observer imposes space, time, and magnitude scales.

Aftershocks are associated to only one previous event, which is also chosen by the observer.

Cannot describe swarms, remote triggering and other manifestations of seismic complexity in space-time-magnitude.

One approach: the null hypothesis & metric

- Consider a pair of events *(i,j)* with $t_i < t_j$ whose distance $r_{ij} = |\underline{r}_i - \underline{r}_j|$, and time $t_{ij} = (t_j - t_i)$. The mean number of events of magnitude within Δm of m_i is

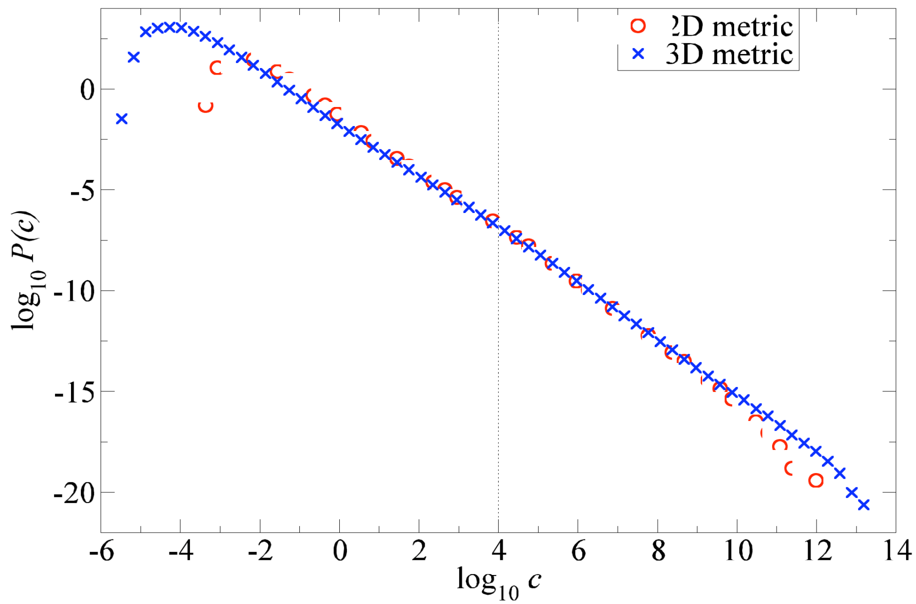
$$n_{ij} = (constant)\ (\Delta m)\ t_{ij}\ (r_{ij})^2 10^{-bm_i}$$

- if the events are occurring at random in space and time according to the Gutenberg-Richter law.

- But event *i* actually occurred relative to *j*.
- Pair is correlated if "surprise" $c_{ij} = 1/n_{ij} >> 1$

- Space and time intervals are selected by the actual sequence of events and not by the observer. "Unbiased"

Distribution of correlations between all pairs of earthquakes with m>2.5



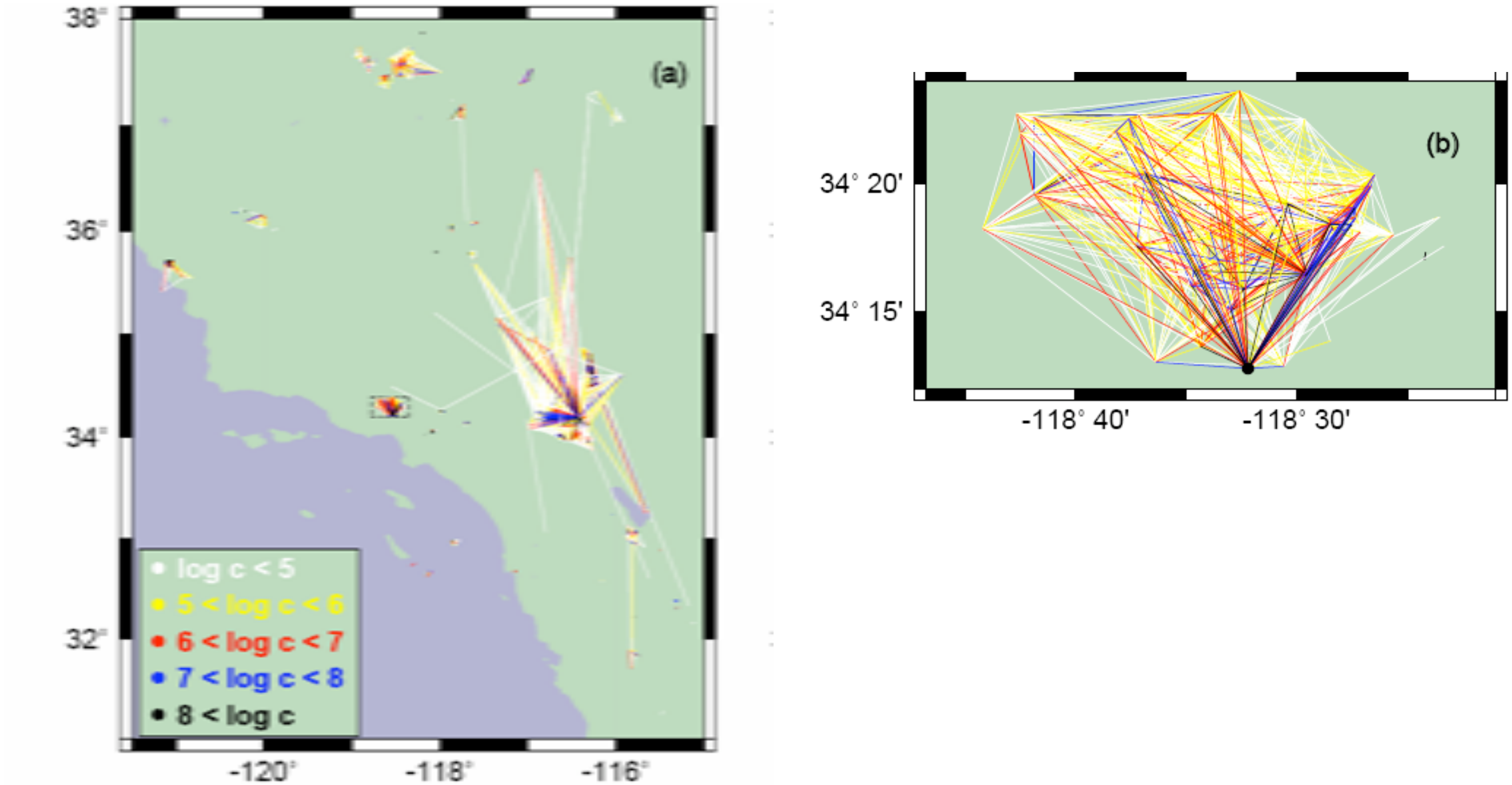
Many surprising pairs of events which could not occur by chance.

Link events with $c > c_{>}$.

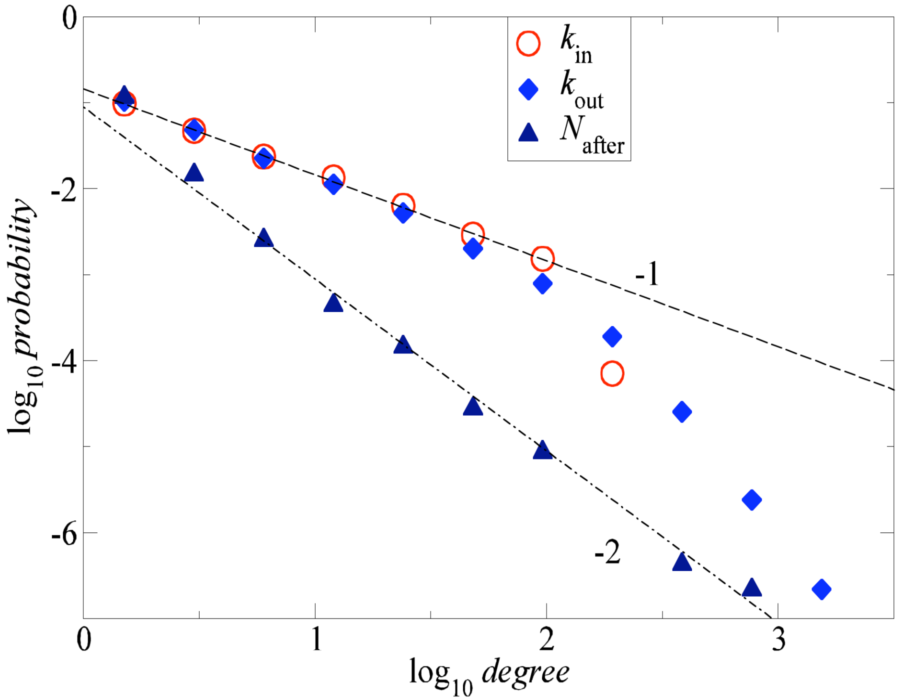
Massive data reduction with small errors.

A sparse network gives a renormalized description of seismicity.

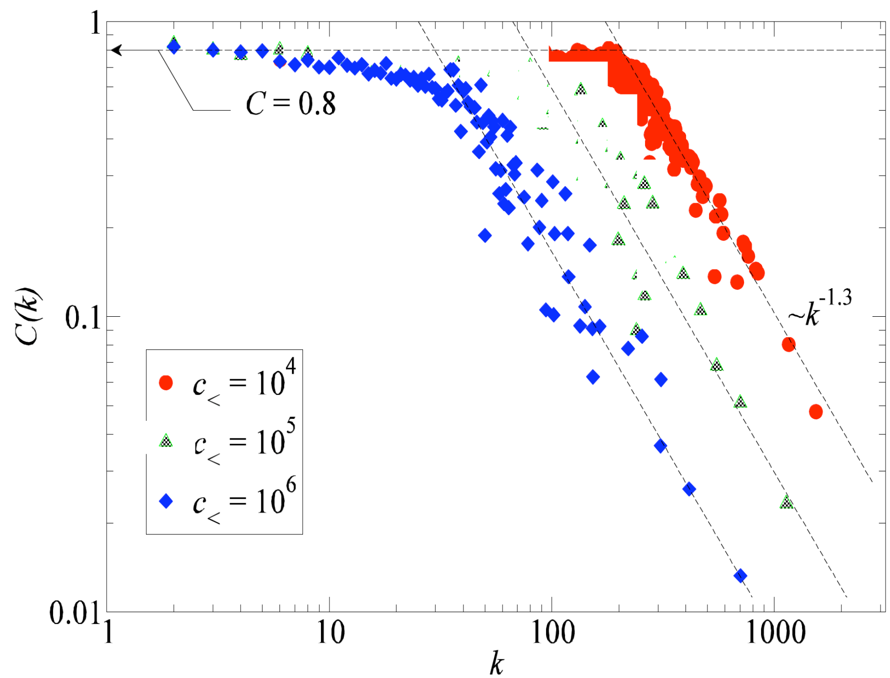
Complex network of earthquakes and aftershocks (Marco Baiesi and MP 2004)



Some universal properties of the network

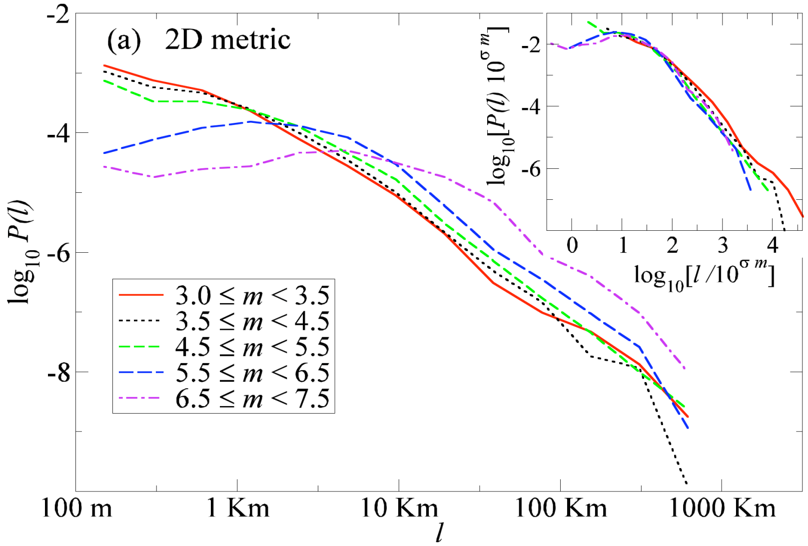


Scale free



Highly clustered

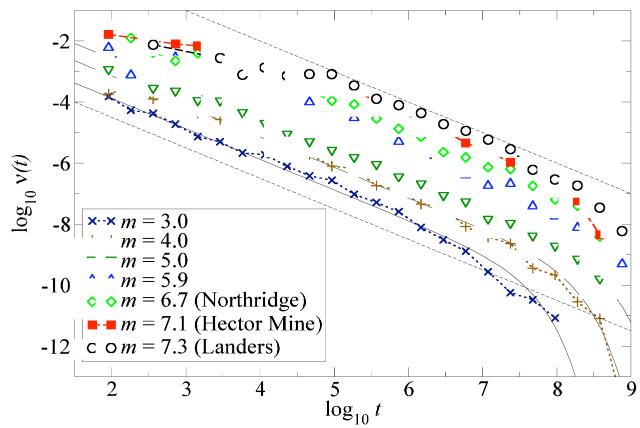
Distances between correlated events

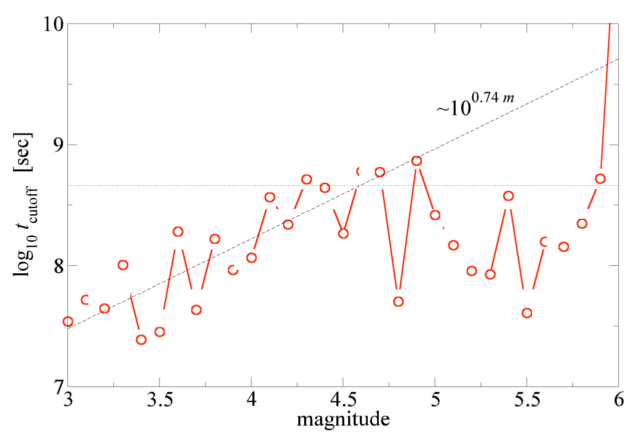


- Typical distance depends on magnitude of main-shock, *m*
- Power law tail for all *m*
- Can be rescaled onto a universal curve, inde-pendent of *m*
- $l^* \sim 10^{\sigma m}$ with σ=0.37

New scaling law for seismicity. Contradicts theory of finite aftershock zones.

Omori Law ***(1/t+A)*** for aftershock rates





- Aftershocks for earthquakes of **ALL** magnitudes, m, obey Omori law
- Rate ~ $t^{-1}\, e^{-t/t_{cutoff}}$
- $t_{cutoff} \sim 10^{\,5.25+0.74m}$ sec

…… time span of catalogue