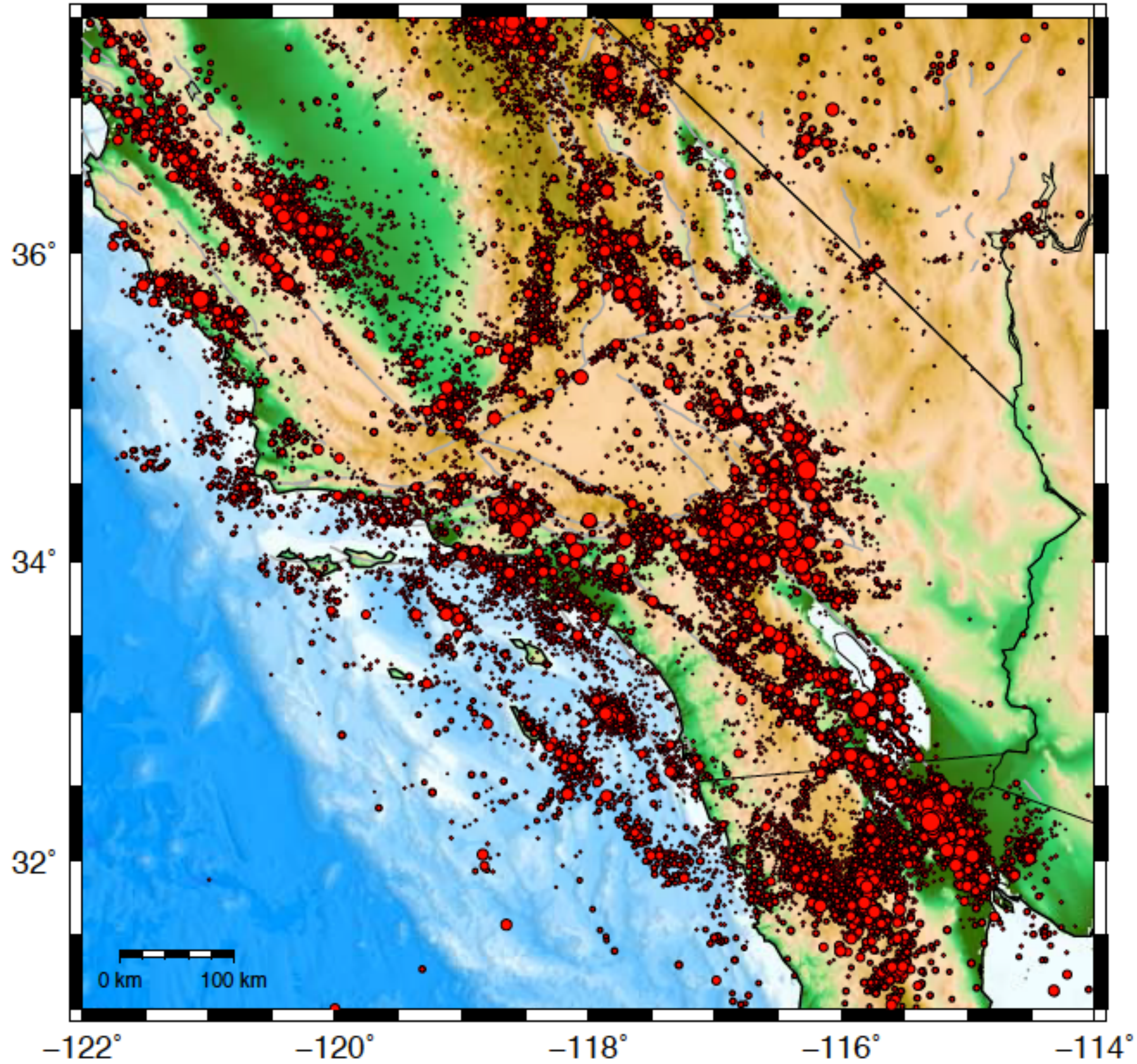


Effective Multifractal Spectra

Jan Reimann
Pennsylvania State University

Effective Randomness

- Algorithmic randomness/information theory links measure theoretic complexity to computational complexity of points.
    - Many new results/structures/questions in recursion theory.
    - Currently: calibration of the "randomness strength" of probabilistic (almost everywhere) theorems.
- A key ingredient: existence of universal objects. A universal test/semimeasure plays a role similar to that of the halting problem.
- The complexity of a "point" is then measured by comparing its local entropies with respect to the given measure and the universal measure.

The “Effective Multifractal Analysis Program”

1. Multifractal analysis studies measures instead of sets. Is there a universal object (measure) exhibiting "global" universality with respect to multifractality?

   Just like we can measure the randomness of a sequence by gauging its complexity along the universal semimeasure, can we describe multifractality by gauging the whole measure against the universal measure?

The “Effective Multifractal Analysis Program”

2. Can we use this universality to prove consistency results for estimators?

Randomness can be characterized by looking at lower bounds on complexity of finite sequences. Can we use similar characterizations to prove that an estimator (based on a finite number of observations) behaves consistent with the underlying mechanism generating the points?

The “Effective Multifractal Analysis Program”

3. Can we design new (better) estimators based on data compression methods?

Replacing Kolmogorov complexity by compressors, can one overcome difficulties by detecting dependencies in the data that causes great problems for traditional methods.

(A similar philosophy underlies MDL (in inductive inference), clustering algorithms by Cilibrasi-Vitanyi and others.)

Brief overview: Fractal Dimensions

- Fractal dimensions capture certain regularities/invariants of sets that are irregular from a topological/Lebesgue measure point of view:
    - Self-similarity
    - Scaling invariance
    - Densities
    - Information/Entropy

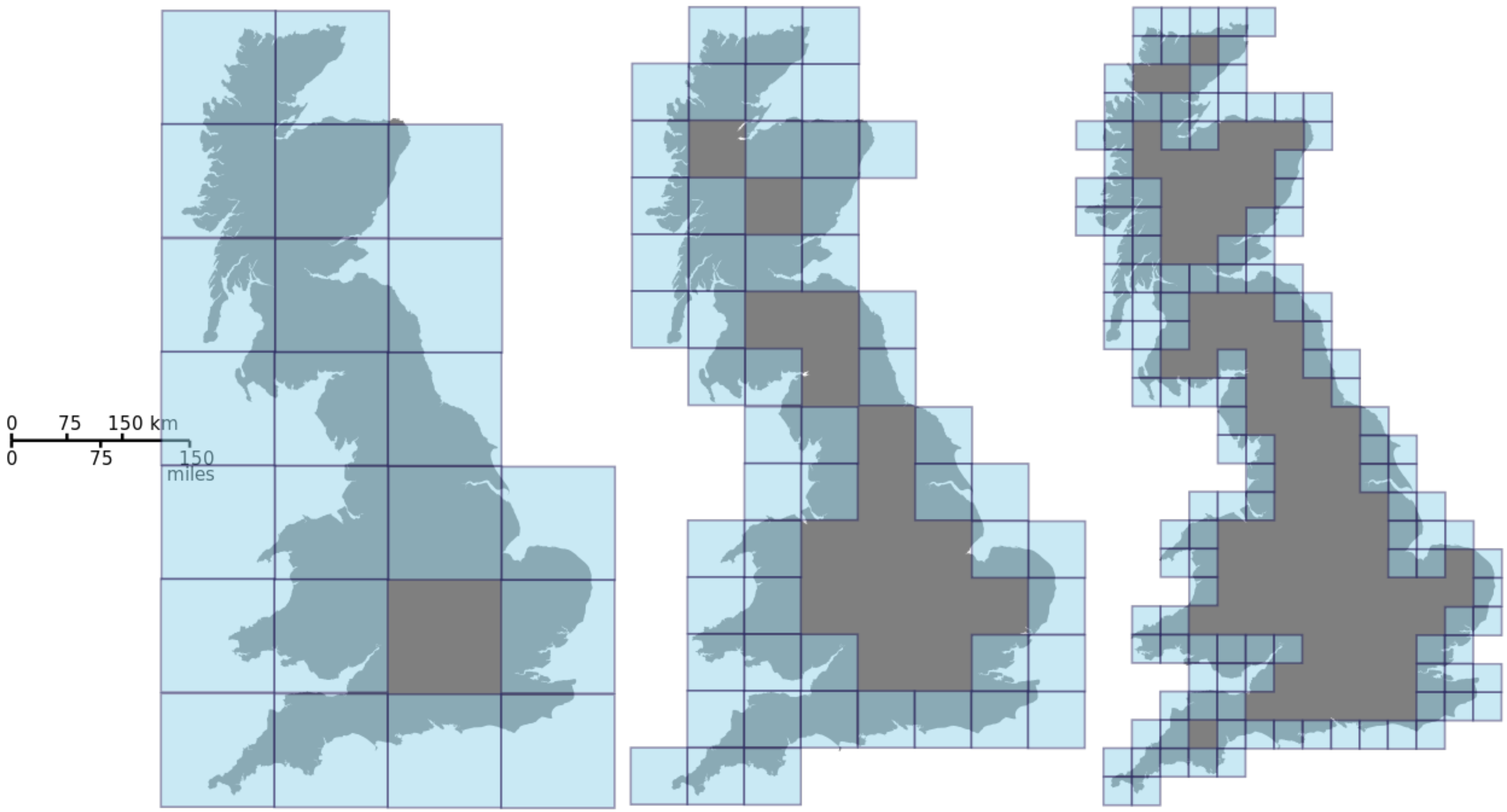
Box Counting Dimension

- Box counting dimension: Let $X \subseteq \mathbb{R}^d$ be bounded.
  - Cover $\mathbb{R}^d$ with a mesh of side length r.
  - Count the number of r-cubes containing points from X.
  - Define

$$\dim_B X = \lim_{r \to 0} \frac{\log N_r(X)}{\log r}$$

(if the limit does not exist, work with lim inf and lim sup).

Box Counting Dimension



Source: Wikipedia

Hausdorff Dimension

- Let $X \subseteq \mathbb{R}^d$.
  - r-covering: cover X with cubes C_i of side length at most r.
  - Optimize the s-dimensional measure of this covering:

$$\mathcal{H}_r^s X = \inf \{ \sum_i \mathrm{diam}(C_i)^s : \ (C_i) \text{ r-cover of } X\}$$

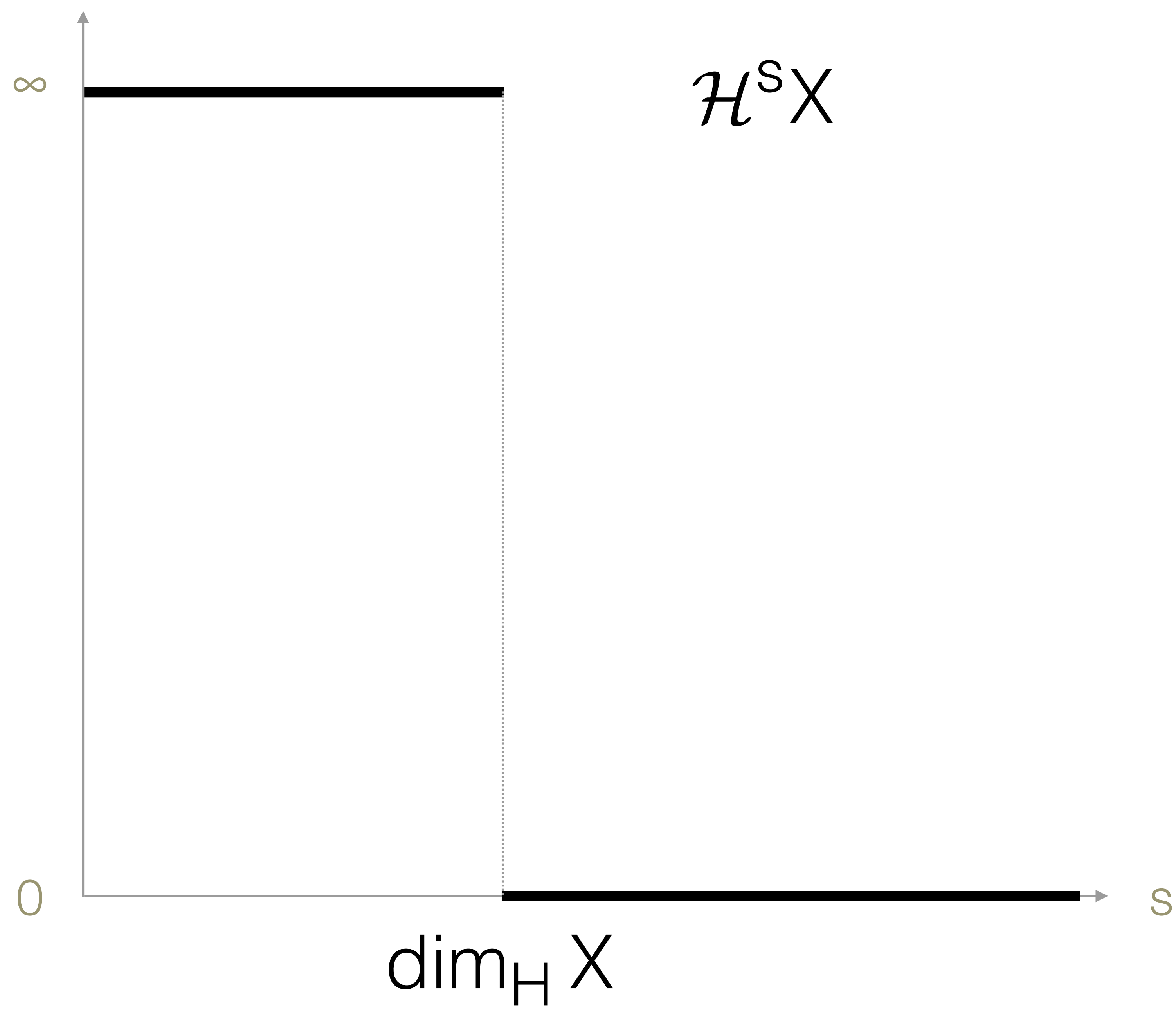
  - Define

$$\mathcal{H}^s X = \lim_{r \to 0} \mathcal{H}_r^s$$

  - The Hausdorff dimension of X is given as

$$\dim_H X = \inf\{s : \mathcal{H}^s X = 0\}$$

Hausdorff Dimension



Dimension and Information

- Eggleston, 1949:
  - Let X_p be the set of all real numbers x so that in the binary expansion of x, 1 appears with frequency p in the limit.
  - Then

$$\dim_H X_p = H(p) = -[p \log p + (1-p) \log(1-p)]$$

Algorithmic Entropy

- Kolmogorov, 1965

THREE APPROACHES TO THE QUANTITATIVE DEFINITION OF INFORMATION

A. N. Kolmogorov

Problemy Peredachi Informatsii, Vol. 1, No. 1, pp. 3-11, 1965

There are two common approaches to the quantitative definition of "information": combinatorial and probabilistic. The author briefly describes the major features of these approaches and introduces a new algorithmic approach that uses the theory of recursive functions.

- Independently by Solomonoff, 1964 —
  "*A formal theory of inductive inference*"

Kolmogorov Complexity

finitary object
(e.g. binary string)

Universal Turing machine

$$C(a) = \min \{|p| : U(p) = a\}$$

binary input

- The Kolmogorov complexity of a string is the length of a shortest possible program computing it.

- The use of a universal machine ensures that this notion is machine independent up to an additive constant:

  If we replace U by another machine M (i.e. use another effective description/coding method), then there exists a constant c so that

$$C(a) \leq C_M(a) + c$$

Algorithmic Information

- A variant, K, of Kolmogorov complexity based on prefix-free codes, resembles classical entropy in many ways:
  - K takes its largest values on strings generated by uniformly random sources:

    $K(a) \geq^{+} |a|$ (a is incompressible)

  - K is subadditive: $K(a,b) \leq^{+} K(a) + K(b) + c$
  - Symmetry of information: $K(a,b) =^{+} K(a) + K(b \mid a, K(a))$

Kolmogorov Complexity and Fractal Dimension

- Ryabko; Staiger: For any set $X \subseteq \mathbb{R}$, there exists an $x \in X$ such that

$$\liminf_n \frac{K(x|_n)}{n} \geq \dim_H X$$

that is, X contains at least one element whose lower asymptotic compression ratio is at least as high the the Hausdorff dimension of X.

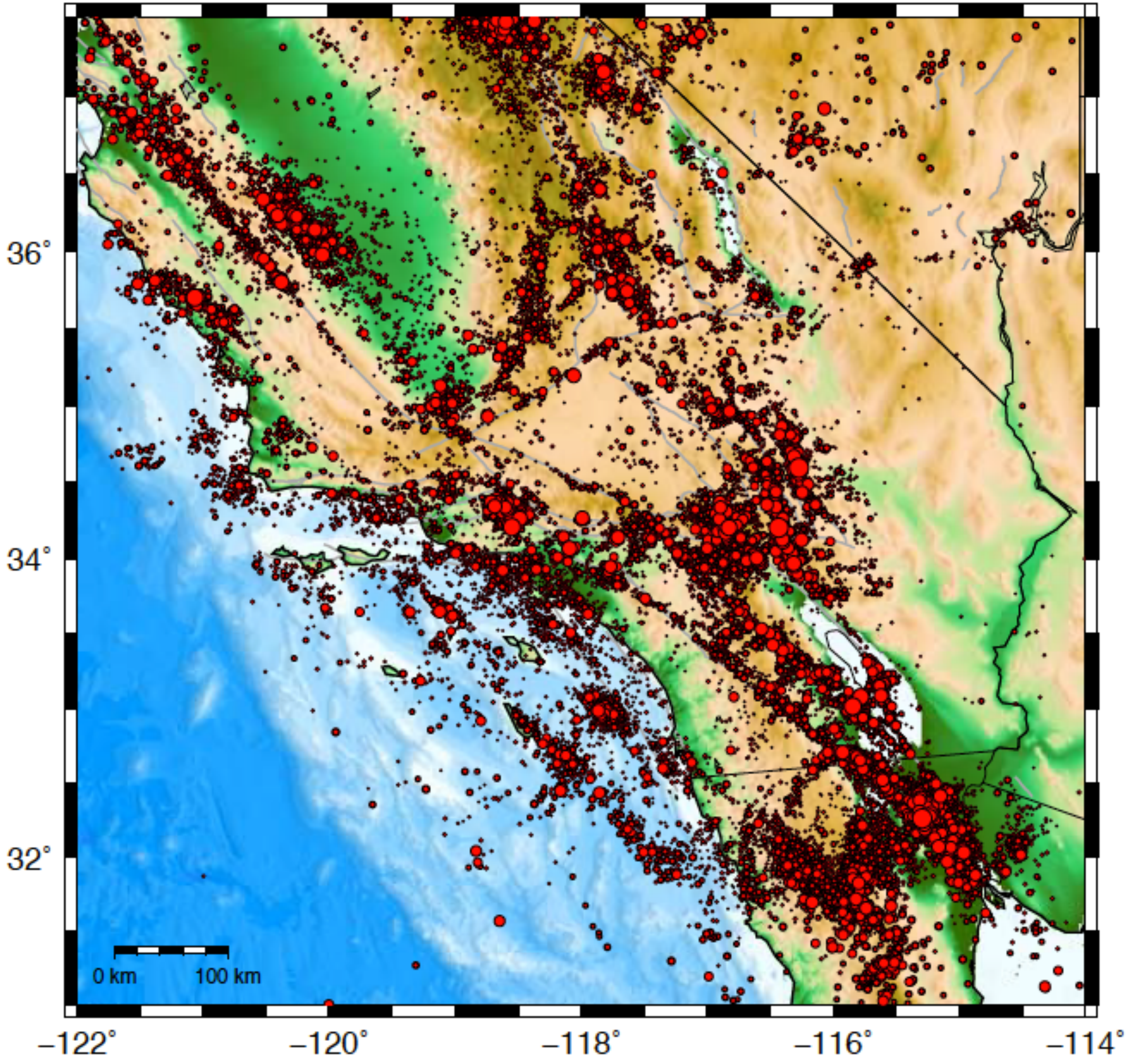
- If the set X is easily definable in the sense that it is a union of effectively closed sets (i.e. sets whose complements can be effectively enumerated), then we can actually characterize the dimension of X via compression ratios of its members:

$$\dim_H X = \sup \{ \liminf K(x|_n)/n : x \in X \}$$

[Lutz; Staiger; Hitchcock]

Measures as Fractals

- While fractal dimensions are useful in capturing geometric invariants of sets, one often encounters fractality of a more complicated, “layered” structure:
  - Consider the distribution of earthquakes and assume it is determined by an underlying dynamics/measure.
  - This measure seems to be supported on a fractal-like set (due to the mechanics of the fracturing process).
  - But there is more to it: Clustering of earthquakes gives different densities/dimensions to different regions.



Hauksson-Shearer-Yang catalog of southern CA earthquakes 1981-2011

Measures as Fractals

- Multifractal spectra try to capture these variations by studying

  (1) the local scaling behavior of a measure at a given point,

  (2) the global (average) scaling behavior of balls.

- For many measures the two aspects are closely related

  → Multifractal Formalism

Observing Multifractal Measures

- An important practical aspect is how to compute multifractal spectra?
- For physical data, there will only be a finite number of observations.
- Several estimators have been introduced, most famously the Grassberger-Procaccia algorithm.
- To show that an algorithm is consistent, one usually has to assume some underlying dynamics or probabilistic process that produces the data.
- Algorithmic Information Theory seems to be a natural and very general framework for this.

relate the complexity of a finite pointset to the complexity of the measure.

Dimension Distribution of a Measure

- Let μ be a Borel probability measure on $\mathbb{R}^N$ with compact support. The Hausdorff dimension distribution of μ is defined as

$$\mu_{\dim}([0,t]) = \sup\{\mu(D) : \dim_H D \leq t, \ D \text{ Borel}\}$$

(This extends to a probability measure on [0,N]. A similar concept can be defined for packing dimension.)

- Example: $\underline{\mu_{\dim}(\text{Lebesgue}) = \delta_N}$

Measures with this property
are called exact dimensional

Dimension Distribution

- It turns out the dimension distribution of μ is given by the μ-distribution of the effective Hausdorff dimension.

Thm: If μ is computable, then for all t,

$$\mu_{\dim}([0,t]) = \mu(\dim_{\leq t}).$$

$\dim_{\leq t}$ = all points of effective dimension ≤ t

$= \{x : \liminf_n K(x|_n)/n \leq t\}$

Local: Pointwise Dimension

- Pointwise (local) dimension of μ at x:

Local scaling behavior at x

$$Y_\mu(x) = \lim_{\delta \to 0} \frac{\log \mu B(x, \delta)}{\log \delta}$$

(If the limit does not exist work with lim inf and lim sup.)

- Thm: For μ computable and x μ-random,

$$\dim_H x = Y_\mu(x).$$

- Corollary [Young, Cutler]:

$$\mu_{\dim}([0, t]) = \mu(\{x : Y_\mu(x) = t\}).$$

Global: Generalized Renyi Dimensions

- Let $B(x,\varepsilon)$ be the N-dimensional ε-ball around x.

- For $-\infty < q < \infty$, $q \neq 1$, let

$$\theta(q) = \lim_{\varepsilon \to 0} \frac{\log \left[\int (\mu B(x, \varepsilon))^{q-1} d\mu(x) \right]}{\log \varepsilon}$$

θ measures the average scaling of the q-th moment of $\mu(B(x,\varepsilon))$

- For integer $q \geq 2$, $\theta(q)/q-1$ is called the correlation dimension of order q.

Multifractal Formalism: from global to local

- μ satisfies the (strong) multifractal formalism if

$$\theta(q) = \inf_y \{qy - f(y)\}$$

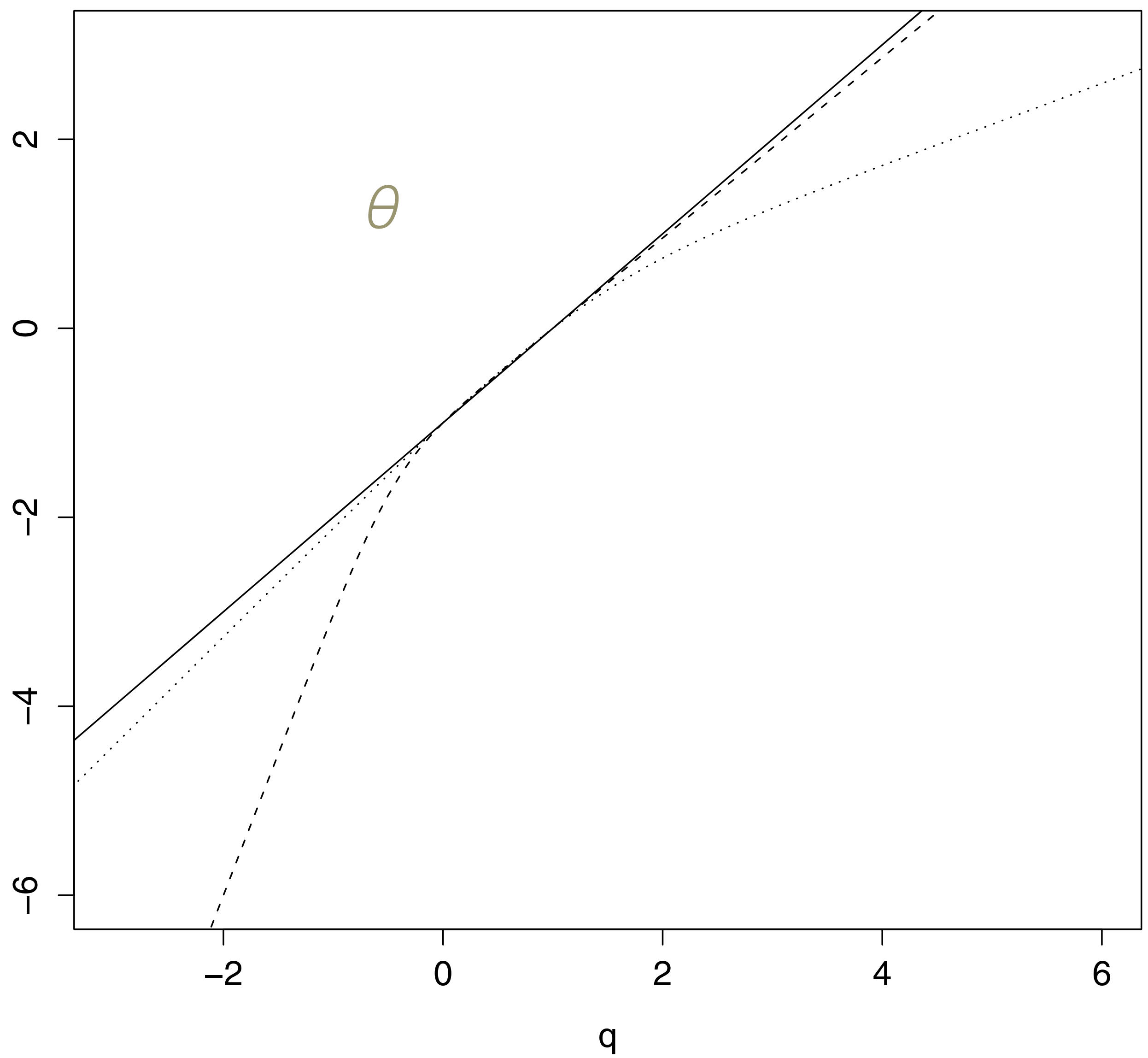
Legendre transform

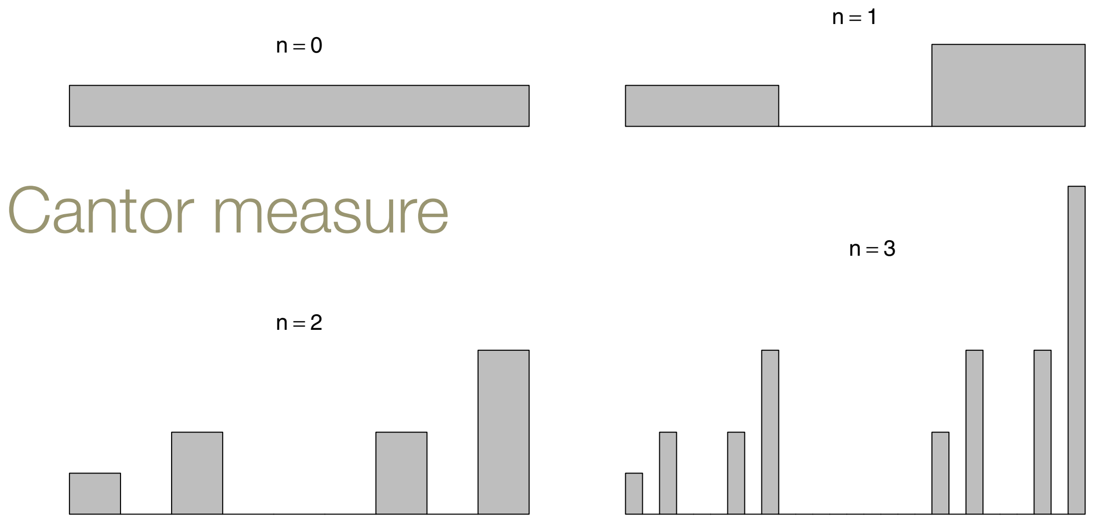
holds whenever $f(y) > 0$, where

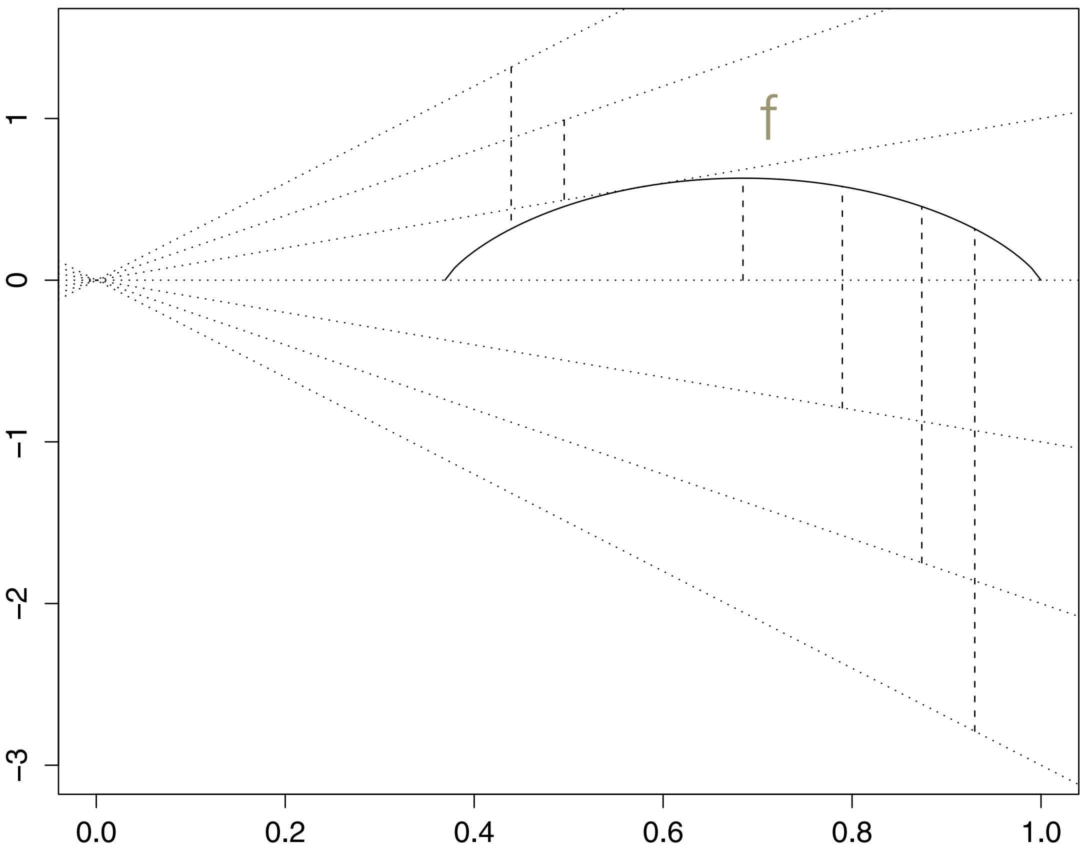
$$f(y) = \dim_H\{x : Y_\mu(x) = y\}.$$

Multifractal spectrum

Examples







Pictures from: Harte, *Multifractals* (2001)

Semimeasures

- Let $2^{\mathbb{N}}$ be the set of all infinite binary sequences.
  - For a finite string σ, [σ] denotes the basic open cylinder $\{x \in 2^{\mathbb{N}} : \sigma \subset x\}$
- A semimeasure M is a function from finite string to non-negative reals satisfying
$$M(\sigma) \geq M(\sigma 0) + M(\sigma 1).$$
- A semimeasure is enumerable if the exists an algorithm that, for input σ, enumerates the left cut of M(σ), i.e. the set $\{q \in \mathbb{Q} : q < M(\sigma)\}$.
- Levin, 1974: There exists an optimal enumerable semimeasure M*. For any enumerable semimeasure M there exists a constant c s.t.
$$c \cdot M^* \geq M$$

Semimeasures and Dimension

- The asymptotic compression ratio of a sequence is the pointwise dimension with respect to M*:

$$\liminf_n \frac{K(x|_n)}{n} = \liminf_n \frac{\log M^*(x|_n)}{n}$$

- Cai & Hartmanis, 1994: $f_{M^*}(y) = y$ for all $0 \leq y \leq 1$.

- M* is, in a certain sense, a "perfect" multifractal: All layers of pointwise dimensions are at the maximum value.

A Universal Multifractal

- Furthermore, the multifractal spectrum of any other (computable) measure can be gauged against the spectrum of M*.

- Thm: If μ is computable, then

M* pointwise dimension

$$\dim_H F_\mu(y) = \dim_H \left\{ x : \frac{\overline{\dim_H x}}{\underline{\dim_\mu x}} = y \right\}.$$

Billingsley dimension

Estimation and Stability of Multifractal Spectra

- Grassberger-Procaccia: $C_\mu(\varepsilon)$:= Probability two random, independent points x, y are no more than distance ε apart. By Fubini's Theorem,

$$C_\mu(\varepsilon) = \mu \times \mu\{(x, y) \colon \|x - y\| \leq \varepsilon\} = \int \mu B(x, \varepsilon) d\mu(x) = \theta(2)$$

- If we have only finitely many observations $x_1, \ldots, x_n$, this suggests using

$$C(n, \varepsilon) = \frac{\sum_{i=1}^{n} \sum_{j>i} 1_{\{\|x_i - x_j\| \leq \varepsilon\}}}{\binom{n}{2}}$$

as an estimator of $C_\mu(\varepsilon)$.

(Similar estimators exist for higher moments.)

Estimation and Stability of Multifractal Spectra

- Consistency of this estimator has been established in the dynamics context:
  - Denker and Keller [1986]: Smooth ergodic systems with mixing condition.
  - Pesin [1993]: Ergodic systems.
- Using the characterization of the multifractal spectrum via effective dimensions, Pesin's result can proved rather easily using the recent work on ergodic properties of ML-random sequences [Franklin et al, Bienvenu et al].

Information Distance

- In practice, the data are rarely ever independent samples (e.g. earthquake aftershocks).

- Idea: replace the use of the Euclidean distance in the GP-algorithm by an information distance.

- One such distance is based on Kolmogorov complexity [Bennet et al]:

$$EC(\sigma,\tau) = C(\sigma\tau) - \min\{C(\sigma),C(\tau)\}$$

- The effective spectrum lets us quantitatively weigh the randomness/independence deficiency of the data against the multifractal deficiency of the limit measure.

- This yields new (often easier) proofs of the consistency of this estimator in a number of settings.

Practical Issues

- Kolmogorov complexity is not computable.
- For applications, we have to approximate it with
  - compressors (Lempel-Ziv etc.)
  - string complexity functions (Lempel-Ziv, Ehrenfeucht-Mycielski, Becher-Heiber)
- Many of the consistency results still go through if we work with a normal compressor (Cilibrasi-Vitanyi):

(1)	Idempotency:	$C(aa) = C(a)$
(2)	Monotonicity:	$C(ab) \geq C(a)$
(3)	Symmetry:	$C(ab) = C(ba)$
(4)	Distributivity:	$C(ab) + C(c) \leq C(ac) + C(bc)$

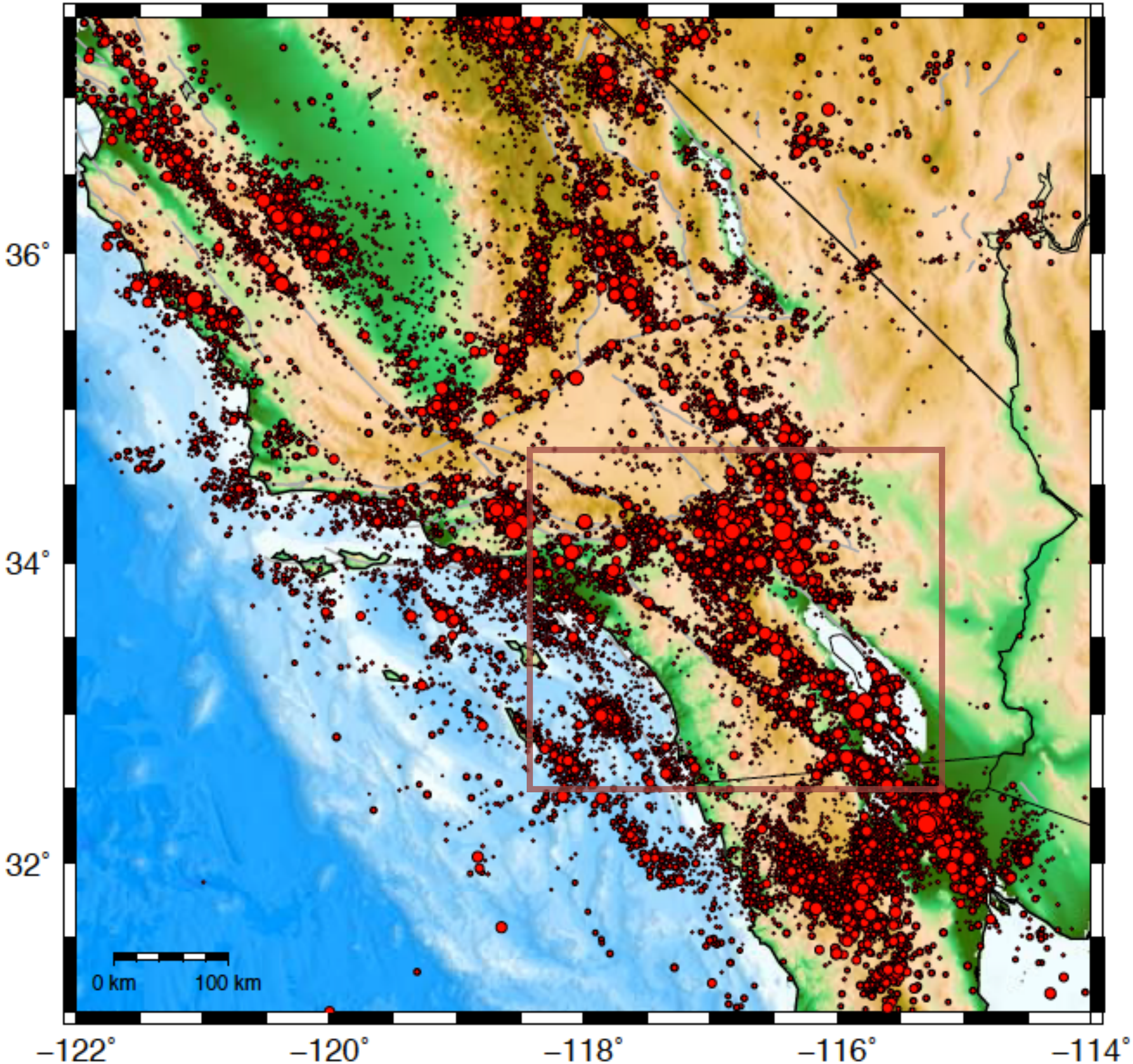
Problem: For the popular compression algorithms it has not been formally established yet that they are normal

Dimension of Fault Systems

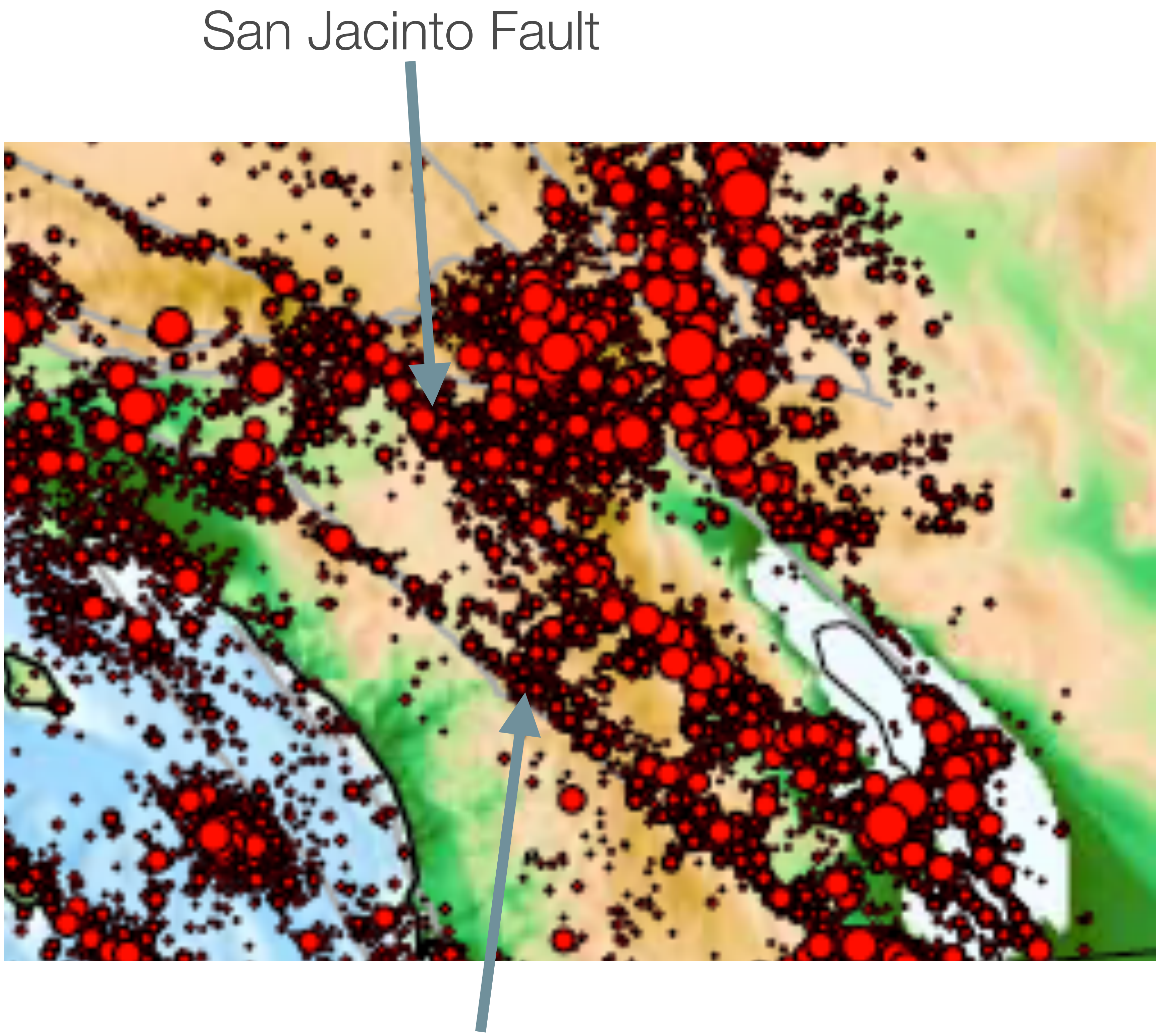
- Fractal dimensions of fault systems have been investigated by many authors.
  [Eneva (1996), Goltz (1997), Ciccotti & Mulargia (2002), Libicki & Ben-Zion (2005), Kagan (2007), Molchan & Konrod (2009)]

- However, the problems with dependencies in the data led to a wide variety of results.
  - Kagan, 2007: "practically any value for the correlation dimension can be obtained…"

- Using a compression-based estimator, initial results seem to indicate that it can account for intrinsic dependencies better than traditional methods.



Hauksson-Shearer-Yang catalog of southern CA earthquakes 1981-2011



Hauksson-Shearer-Yang catalog of southern CA earthquakes 1981-2011