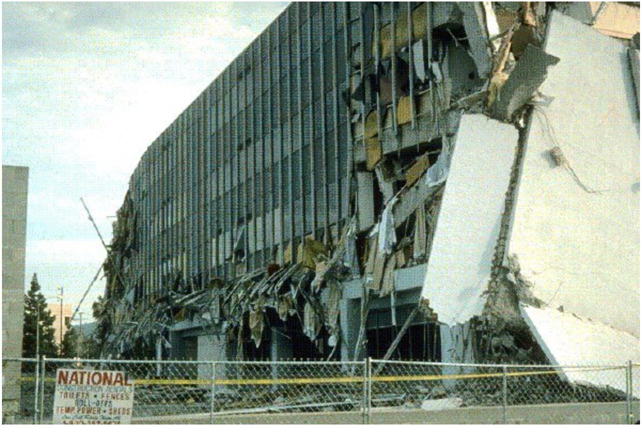
Evaluation of Earthquake Forecast Models Using Goodness-of-Fit Methods for Spatial-Temporal Point Processes

Robert Clements

July 30, 2009

Why Goodness-of-fit methods are important

- useful for evaluating the behavior of earthquake models
  - can suggest improvements
  - can yield better forecasts
- earthquakes cost money (insurance, repairs), lives, and cause damage



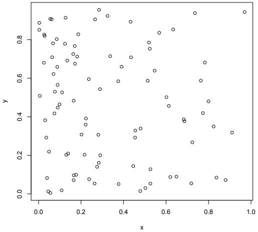
The Northridge earthquake caused an estimated $40 billion in damages

Part I

Background

Point processes

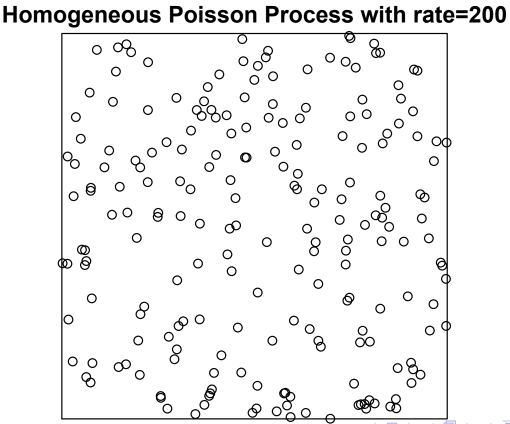
- random collections of points in some compact subset S of $\mathbb{R}^n$
- (locally) σ-finite: finitely many points within any bounded set
- simple: points at distinct locations



Poisson processes

- for any set B, the number of points in B, $N(B)$, follows a Poisson distribution
- no clustering, no inhibition
- the homogeneous (stationary) Poisson process
  - completely spatially random
  - any clustering is by chance
- the inhomogeneous (non-stationary) Poisson process
  - Poisson process with rate as a function of time (or space)

Example



The conditional intensity (Brémaud, 1981)

- the conditional intensity $\lambda(x, y, m, t)$ is defined as the frequency with which events are expected to occur in a space S around a specific point, time and mark (for marked point processes), conditional on the prior history H_t
- λ uniquely characterizes a simple point process

$$\boxed{\lim_{\Delta x, \Delta y, \Delta m, \Delta t \downarrow 0} \frac{E[N\{(x, x+\Delta x)x(y, y+\Delta y)x(m, m+\Delta m)x(t, t+\Delta t)\}|H_t]}{\Delta x \Delta y \Delta m \Delta t}}$$

Part II

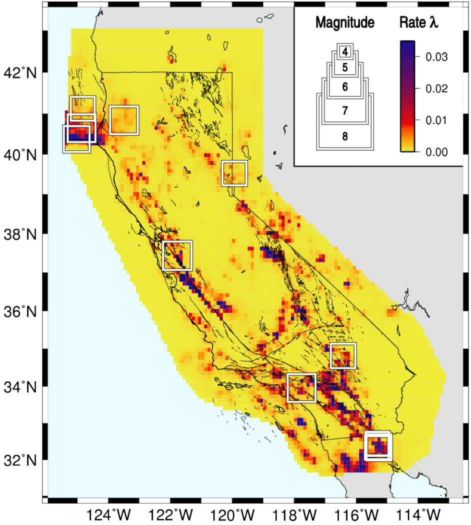
Earthquake Forecasts

CSEP

Collaboratory for the Study of Earthquake Predictability

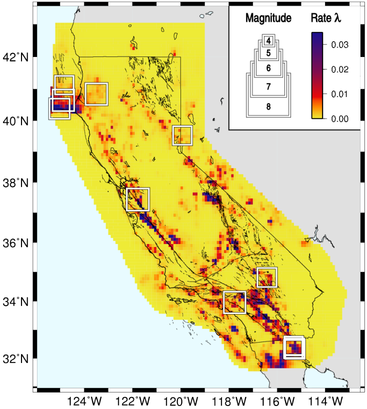
- international collaboration, with testing centers in:
  - Japan, Switzerland, New Zealand, and the US
- earthquake forecast models submitted by researchers in the field
- short-term (daily) and long-term (5-year) forecasts
- testing regions
  - California
  - Japan
  - Northwest/Southwest Pacific
  - New Zealand
  - Global

Helmstetter et al. forecast



CSEP methods (cseptesting.org)

- currently numerical summaries and error diagrams are used to test goodness-of-fit
  - the N-test (number of events)
  - the L-test (likelihood)
  - the R-test (likelihood-ratio for comparing models)
  - the Molchan-test (Molchan 1990; Molchan 1997; Zaliapin & Molchan 2004; Kagan 2009)
  - the ROC-test (Swets 1973)



Problems with CSEP's methods

- provides no information about where the model fits poorly or where the model fits well
- mainly useful for comparing models
- further tests usually necessary to draw conclusions

Goodness-of-fit of earthquake forecasts

What we propose

- we propose thinning/superposition, rescaling, pixel-based and numerical summary methods to test the goodness-of-fit of both short-term and long-term California models

Part III

Goodness-of-Fit Methods

Thinning (Schoenberg, 2003)

- for multidimensional Poisson point process models estimating the conditional intensity

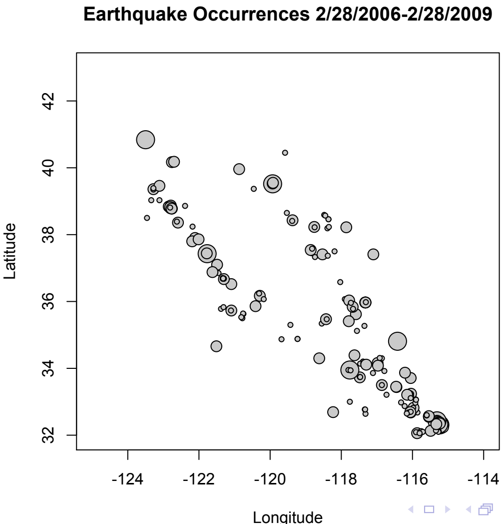
$$\lambda(x, y, m, t)$$

- keep each observation in the space S with probability

$$\frac{b}{\hat{\lambda}(x_i, y_i, m_i, t_i)} \text{ where } b = min\{\hat{\lambda}(x_i, y_i, m_i, t_i)\} \text{ over } S$$

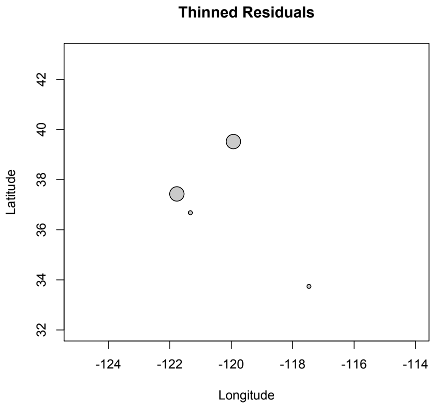
- remaining points, called *thinned residuals*, are homogeneous Poisson if model is correct
- use Ripley's K-function to test for uniformity

Example


Earthquake Occurrences 2/28/2006-2/28/2009
Latitude
42
40
38
36
34
32
-124
-122
-120
-118
-116
-114
Longitude

Example

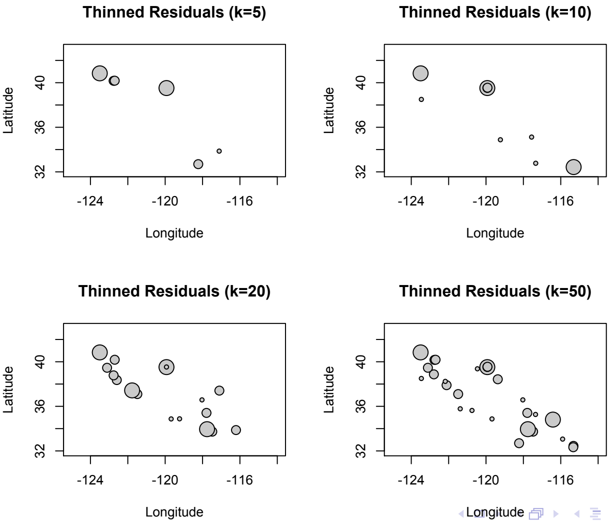
Thinned using the Helmstetter et al. forecast model



Problems with thinning

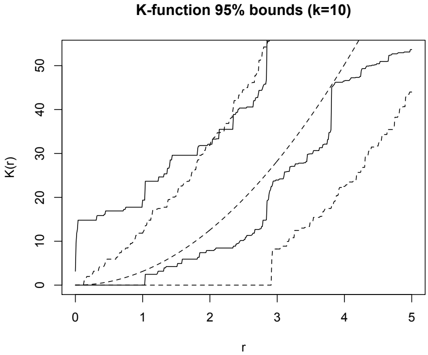
- if b is very small, we end up with few points
  - can instead use probability $\frac{k}{\hat{\lambda}(x_i, y_i, m_i, t_i) \sum_{i=1}^{N(S)} \frac{1}{\hat{\lambda}(x_i, y_i, m_i, t_i)}}$
- how do we choose k to get the most power from this method?

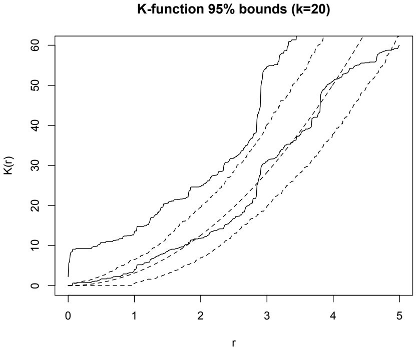
Thinned residuals (k=5, 10, 20, 50)



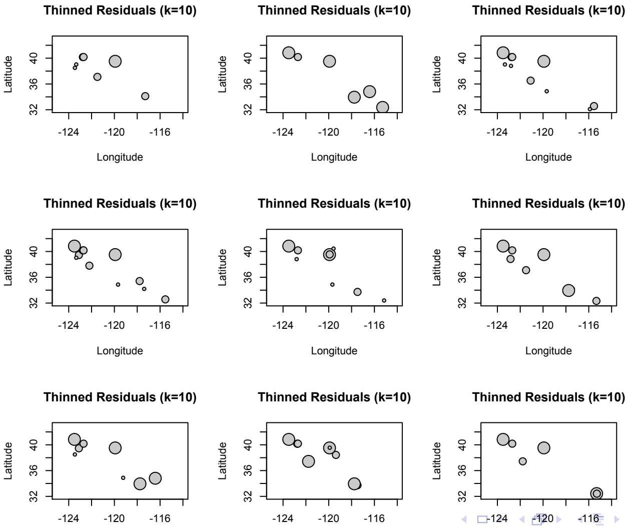
Ripley's K-functions (k=10, 20)

Solid lines - thinned residuals
Dashed lines - homogeneous Poisson





Looking for trouble spots



Summary

- current numerical summary methods used to evaluate earthquake forecasts lack helpful information for improving the models
- other methods, such as thinning, may be more useful
- goodness-of-fit methods need more work

Thank you for attending!

Contact info:
Robert Clements
Department of Statistics, UCLA
clements@stat.ucla.edu