Statistics 222, Spatial Statistics.

Outline for the day:

1. Continue with day7.r.
2. Nonparametric estimation of Hawkes processes using MISD.

Background and motivation.

* History of numerous models for earthquake forecasting, with mostly failures.
(elastic rebound, water levels, radon levels, animal signals, quiescence, electro-magnetic signals, characteristic earthquakes, AMR, Coulomb stress change, etc.)

* Skepticism among many in seismological community toward *all* probabilistic forecasts.

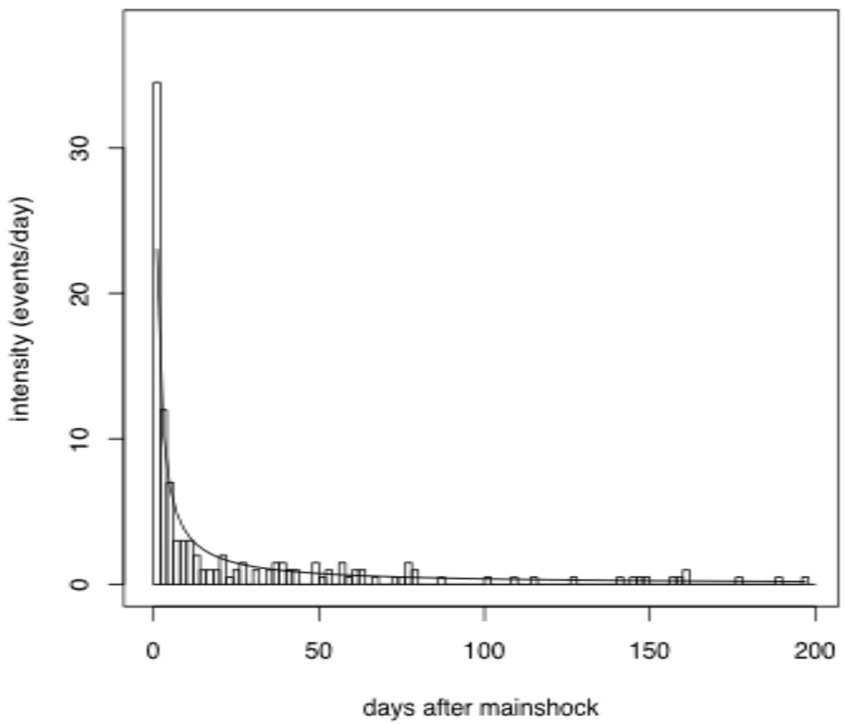
* Different models can have similar fit and very different implications for forecasts.

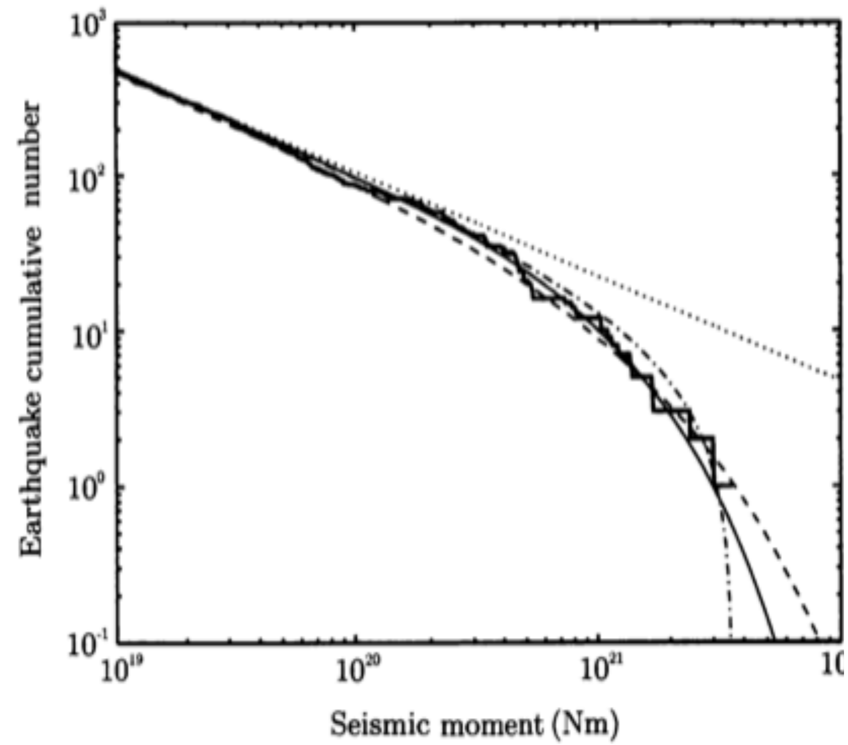
(e.g. Pareto vs. tapered Pareto for seismic moments. Fitting these by MLE to 3765 shallow worldwide events with M≥5.8 from 1977-2000,
the Pareto says there should be an event of M ≥ 10.0 every 102 years,
the tapered Pareto every 10^{436} years.
The fitted Pareto predicts an event with M≥12 every 10,500 years,
the tapered Pareto every 10^{43400} years.)

* Model evaluation techniques and forecasting experiments to discriminate among competing models and improve them are very important.

* We also need <u>non-parametric</u> alternatives to these models.

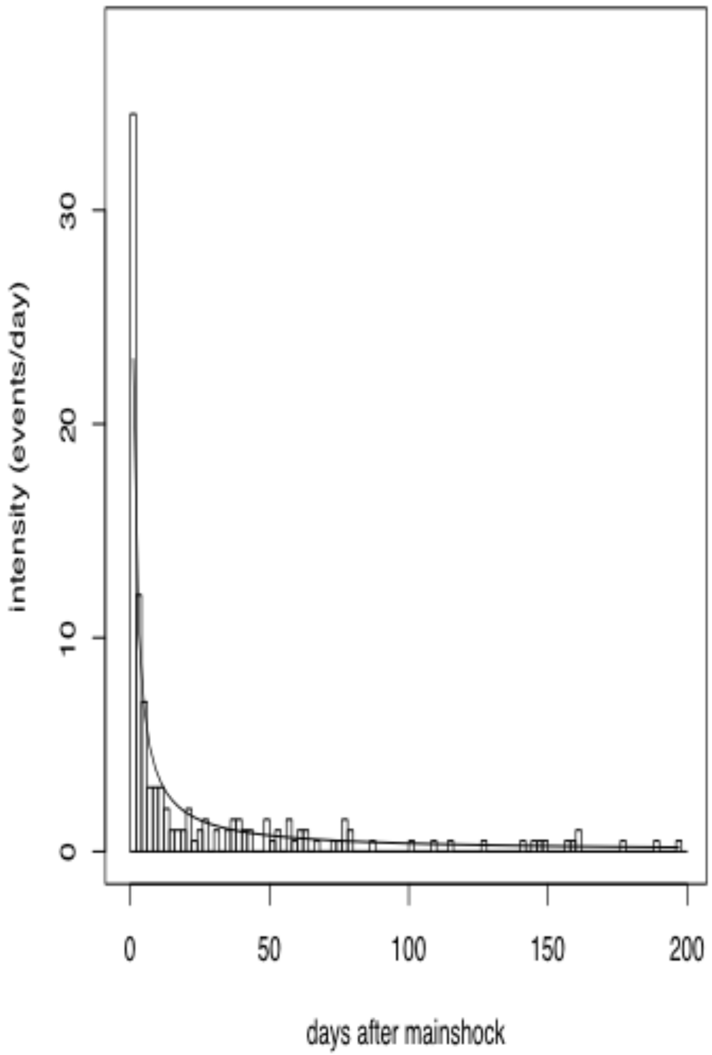
Kagan and Schoenberg (2001)



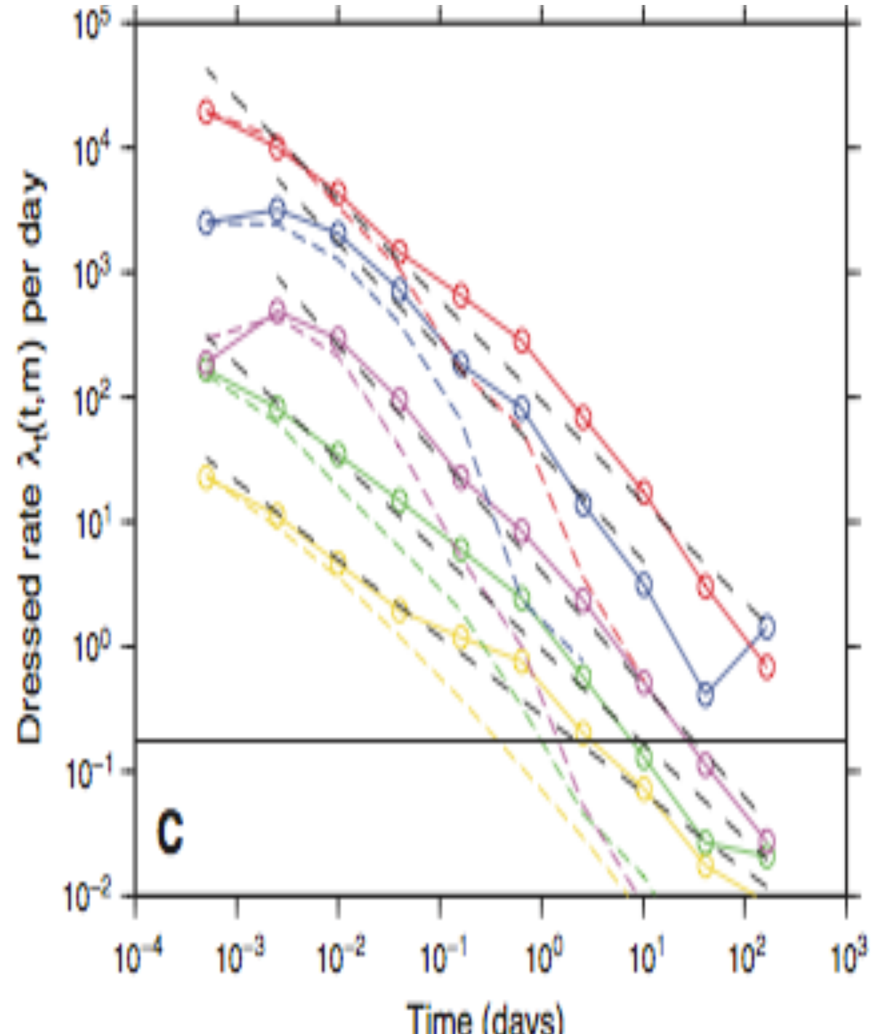


* We also need <u>non-parametric</u> alternatives to these models.

Temporal activity described by modified Omori Law: $K/(u+c)^p$



Marsan and Lengliné (2008)



Nonparametric estimation of Hawkes and ETAS processes.

Let **x** mean spatial coordinates = (x,y).
Hawkes processes have $\lambda(t,\mathbf{x}) = \mu(\mathbf{x}) + K \sum_i g(t-t_i, \mathbf{x}-\mathbf{x}_i)$.

- An ETAS model may be written

$$\lambda(t, \mathbf{x}|\mathcal{H}_t) = \mu(\mathbf{x}) + K \sum_{i:t_i<t} g(t - t_i, \mathbf{x} - \mathbf{x_i}, m_i),$$

  with triggering function

$$g(t-t_i, \mathbf{x}-\mathbf{x_i}, m_i) = \exp\{a(m_i-M_0)\}(t-t_i+c)^{-p}(||\mathbf{x}-\mathbf{x_i}||^2+d)^{-q}.$$

with e.g. $g(u, \mathbf{x} ; m_i) = (u+c)^{-p} \exp\{a(m_i-M_0)\} (||\mathbf{x}||^2 + d)^{-q}$.

These ETAS models were introduced by Ogata (1998).

Instead of estimating g parametrically, one can estimate g nonparametrically, using the method of Marsan and Lengliné (2008), which they call Model Independent Stochastic Declustering (MISD).

David Marsan

Extending Earthquakes' Reach Through Cascading

David Marsan* and Olivier Lengliné

Earthquakes, whatever their size, can trigger other earthquakes. Mainshocks cause aftershocks to occur, which in turn activate their own local aftershock sequences, resulting in a cascade of triggering that extends the reach of the initial mainshock. A long-lasting difficulty is to determine which earthquakes are connected, either directly or indirectly. Here we show that this causal structure can be found probabilistically, with no a priori model nor parameterization. Large regional earthquakes are found to have a short direct influence in comparison to the overall aftershock sequence duration. Relative to these large mainshocks, small earthquakes collectively have a greater effect on triggering. Hence, cascade triggering is a key component in earthquake interactions.

Earthquakes of all sizes, including aftershocks, are able to trigger their own aftershocks. The cascade of earthquake triggering causes the seismicity to develop complex, scale-invariant patterns. The causality of "mainshock A triggered aftershock B," which appears so obvious i A happens to be large, must then into a more subtle "mainshock C1, which triggered C2, ..., wh B." This has paramount consec physical mechanism that cause gering (static or dynamic str fluid flow, afterslip, etc.) canno by looking at aftershocks th directly triggered by the mains over, if indirect triggering is imp overall aftershock budget (*1*–*3*), then direct triggering must be confined to spatial ranges and times shorter than the size of the total

Laboratoire de Géophysique Interne et Tectonophysique, CNRS, Université de Savoie, 73376 Le Bourget du Lac, France.

*To whom correspondence should be addressed. E-mail: david.marsan@univ-savoie.fr

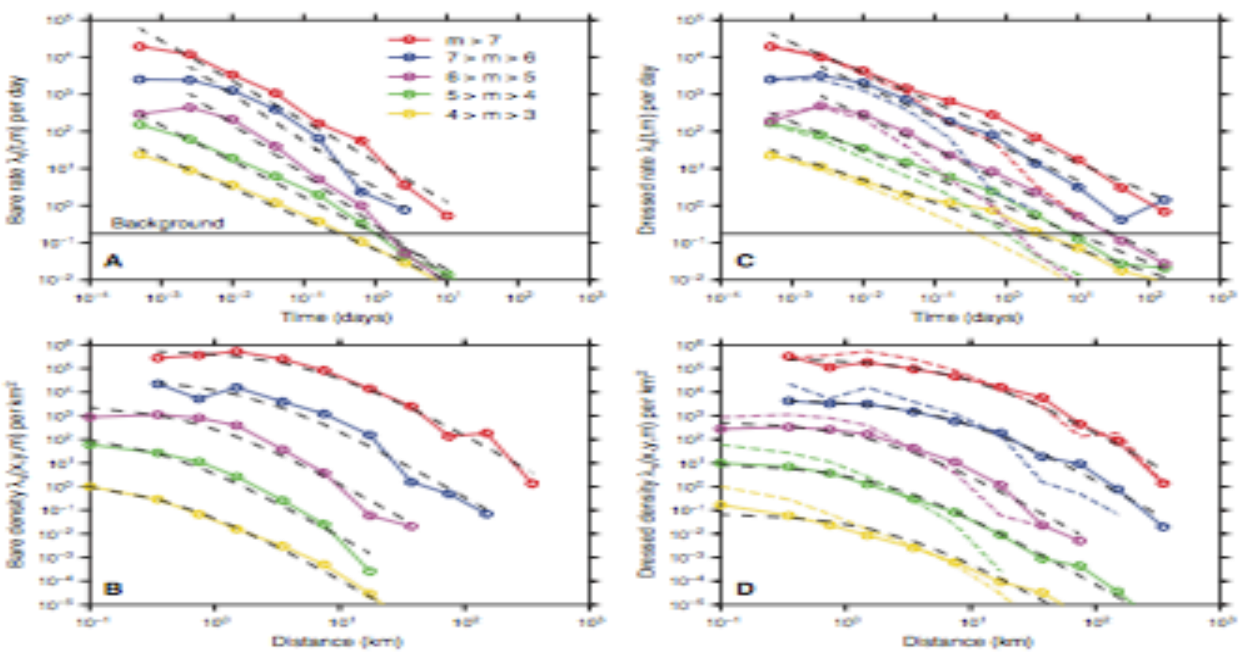


Fig. 1. Estimated rates and densities for California. (**A** and **B**) Bare kernels; (**C** and **D**) dressed kernels. The best power laws for the temporal rates $\lambda_t(t, m)$ and the best $[1 + (r/L)]^{-x}$ laws for the densities $\lambda_r(x, y, m)$ are shown as black dashed lines. The background temporal rate $\lambda_{0,t}$ [black horizontal line in (A) and (C)] is computed as $\sum_{i=1}^{n} w_{0,i}/T$. In (C) and (D), the dressed kernels (continuous lines) are compared to the bare ones (color dashed lines). The densities λ_r have been vertically shifted for clarity.

Model Independent Stochastic Declustering

- The method of Marsan and Lengliné (2008):

$$\lambda(t, m, x, y|\mathcal{H}_t) = \mu(x, y) + \sum_{j:t_j<t} \kappa(m_j) g(t - t_j) f(x - x_j, y - y_j),$$

- Maximizes the expectation of the complete data log-likelihood and assigns probabilities that a child event i is caused by an ancestor event j.

Expectation Step

$$p_{ij} = \frac{g(u)f(x, y)}{\mu(x, y) + \sum g(u) f(x, y)},$$

$$p_{ii} = \frac{\mu(x, y)}{\mu(x, y) + \sum g(u) f(x, y)}.$$

Nonparametric estimation of Hawkes and ETAS processes.

Gordon et al. (2017) let the triggering function, g,
depend on *magnitude*, *sub-region*, *distance*, and *angular separation* from the location (x, y) in question to the triggering event.

$$\lambda(t, m, x, y|\mathcal{H}_t) = \mu(x, y) + \sum_{j:t_j<t} \kappa(m_j)g(t - t_j)f(x - x_j, y - y_j; \phi_j, m_j),$$

Josh Gordon

Expectation Step

$$p_{ij} = \frac{g(u)f(x, y, \phi, m)}{\mu(x, y) + \sum g(u)f(x, y, \phi, m)},$$

$$p_{ii} = \frac{\mu(x, y)}{\mu(x, y) + \sum g(u)f(x, y, \phi, m)}.$$

Maximization Step

$$h(r, \theta, m)_{k,\ell,q} = \frac{\sum_{C_{k,\ell,q}} p_{ij}}{\Delta r_k \Delta\theta_\ell \Delta m_q \underbrace{\sum_{i=1}^{N}\sum_{j=1}^{i-1} p_{ij}}_{\#\text{ of Aftershocks}}},$$

- $C_{k,\ell,q} = \left\{(i,j) \middle| \delta r_k \le r_{ij} \le \delta r_{k+1},\ \delta\theta_\ell \le \theta_{ij} \le \delta\theta_{\ell+1},\ \delta m_q \le m_j \le \delta m_{q+1}, i > j\right\}$ is the set of indices of all pairs of events that fall within the bins specified by the multidimensional histogram density estimator for *magnitude*, *distance*, and *angular separation* $h(r, \theta, m)$.
- κ and g are maximized similarly