The Epidemic Type Aftershock Sequence (ETAS) Model

or

"A *curious* tour of point process models for seismicity"

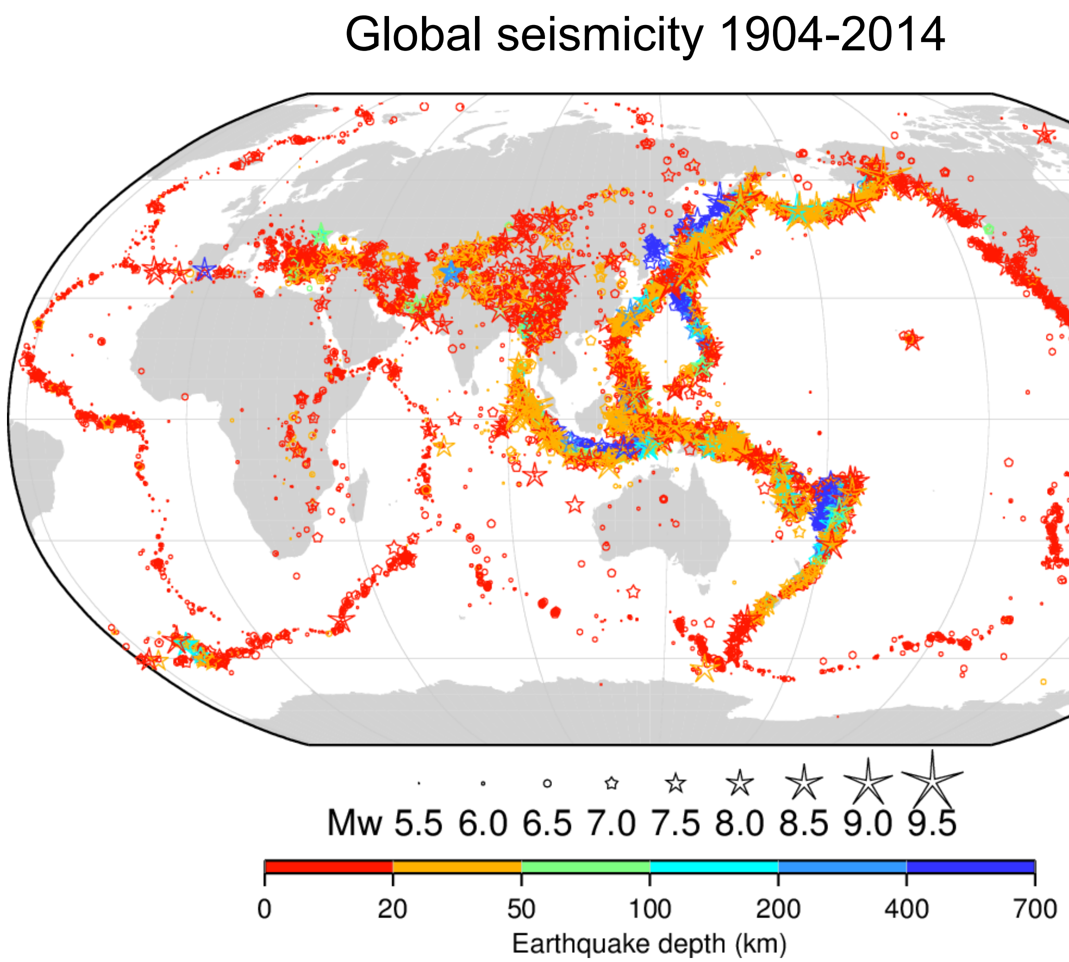
Max Werner

max.werner@bristol.ac.uk

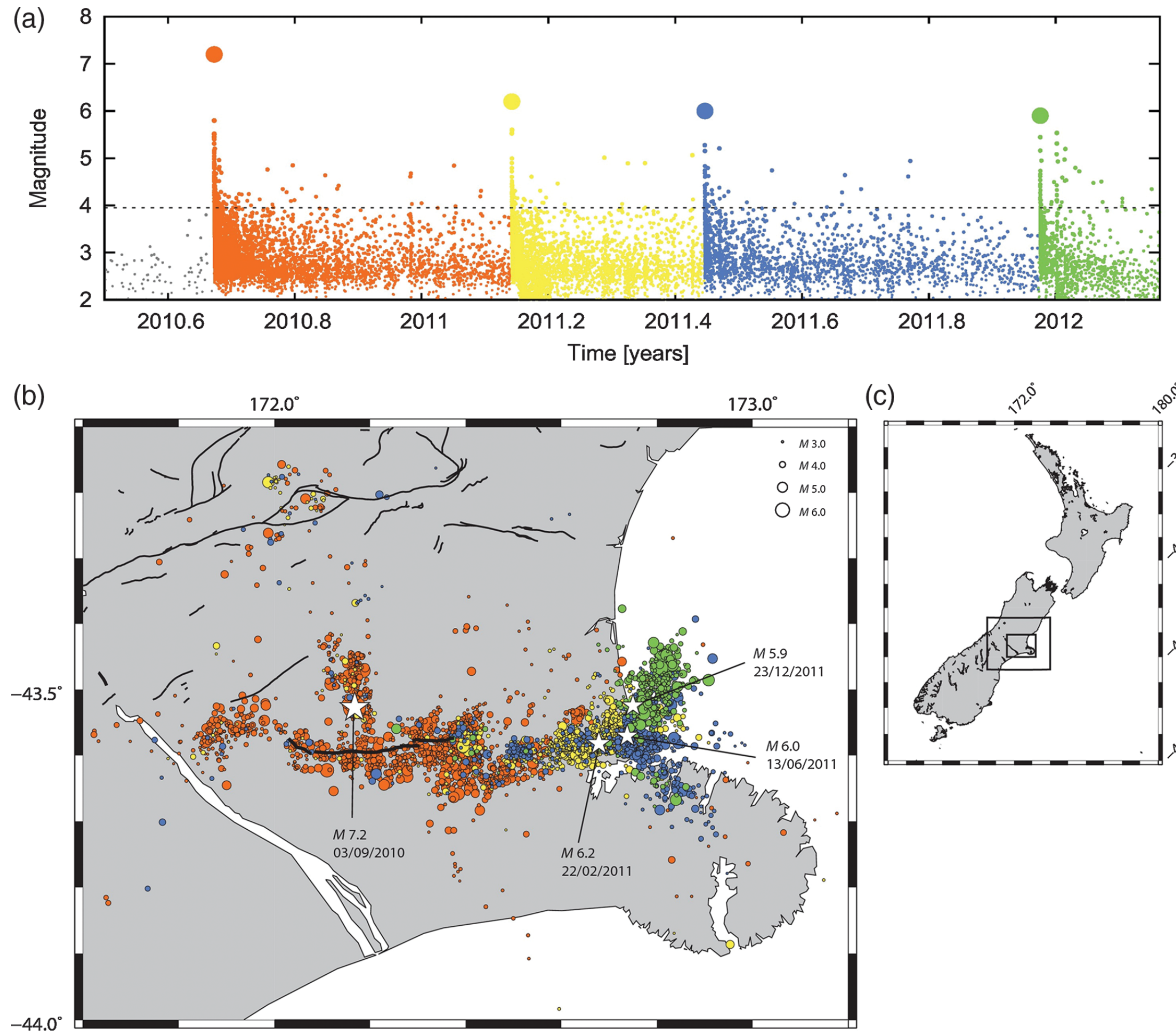
Today: a curious tour

1. Point process introduction
2. Renewal point processes of seismicity
3. Data assimilation (*illustration*) for renewal models
4. The ETAS model
5. ETAS model forecast evaluations
6. What's next for ETAS modelling?

Earthquakes as point patterns



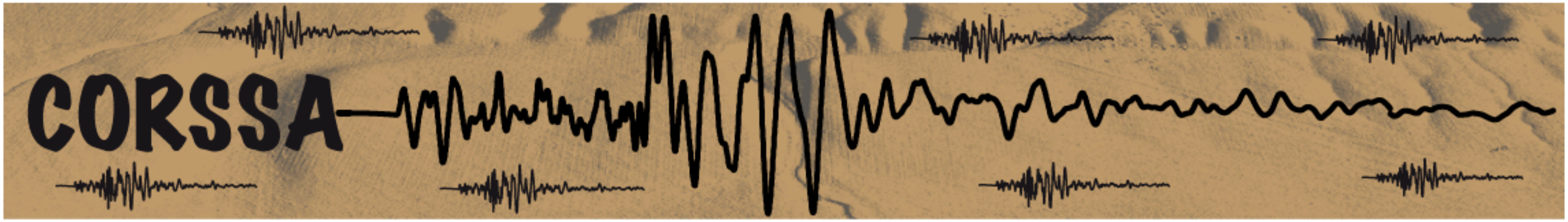
Earthquakes as point patterns



- Complex spatio-temporal aftershock patterns
- Omori law of aftershocks
- Gutenberg-Richter distribution of magnitudes
- Greater aftershock productivity of larger earthquakes
- Complex spatial patterns

University of BRISTOL

www.corssa.org



CORSSA: the Community Online Resource for Statistical Seismicity Analysis

Statistical seismology is the application of rigorous statistical methods to earthquake science with the goal of improving our knowledge of how the earth works. Within statistical seismology there is a strong emphasis on the analysis of seismicity data in order to improve our scientific understanding of earthquakes and to improve the evaluation and testing of earthquake forecasts, earthquake early warning, and seismic hazards assessments. Given the societal importance of these applications, statistical seismology must be done well. Unfortunately, a lack of educational resources and available software tools make it difficult for students and new practitioners to learn about this discipline. The goal of the Community Online Resource for Statistical Seismicity Analysis (CORSSA) is to promote excellence in statistical seismology by providing the knowledge and resources necessary to understand and implement the best practices.

CORSSA covers a variety of themes:

1. Introductory Material
2. Introduction to Basic Features of Seismicity
3. Statistical Foundations
4. Understanding Seismicity Catalogs and Their Problems
5. Models and Techniques for Analyzing Seismicity
6. Earthquake Predictability and Related Hypothesis Testing

Dear reader,

As of March 2016, we are on indefinite hiatus. In the meantime, have a look and enjoy what's here. Happy learning, see ya around!

Part 1:
Point processes

- Definition
- Conditional intensity function
- Likelihood function

Purpose of point process modelling

- Class of statistical models well suited for analysing collections of discrete events (point patterns) irregularly arranged in space and time
- To understand structural features of the arrangement of events - without modelling the underlying (e.g. physical) mechanism
- To forecast future occurrences

Definition(s) of a Point Process

- "probabilistic rules for scattering points in space & time and assigning marks" (mine)

- A random measure N specifying the number of points, N(A), in any compact set A in S, where S is the domain of the point process (e.g. T x R^2). The (counting) measure N is non-negative, integer-valued and finite on any finite subset of S. (e.g. Peng, 2003)
- Full measure theoretic definitions: Daley & Vere-Jones (2003)

Conditional Intensity Function (CIF)

- "instantaneous rate of occurrence of events at time t given the observed history H_t"
- "Probability of an event in a tiny interval"
- "instantaneous hazard rate"

$$\boxed{\lambda(t \mid H_t) = \lim_{dt \to 0} \frac{\Pr\{N(t, t+dt) > 0 \mid H_t\}}{dt}}$$

- Once specified, the CIF ***completely*** specifies the point process! It's the single most important object of a point process.
- Generalises (pretty well) to higher dimensions.

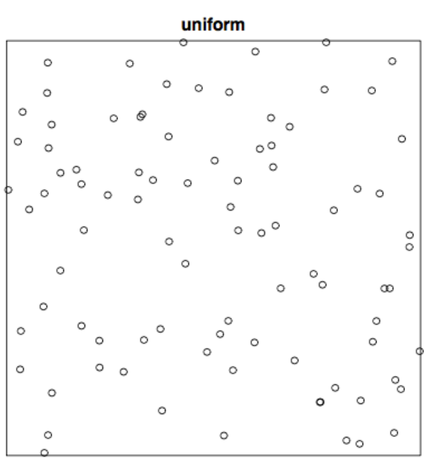
Example: Homogeneous Poisson Process

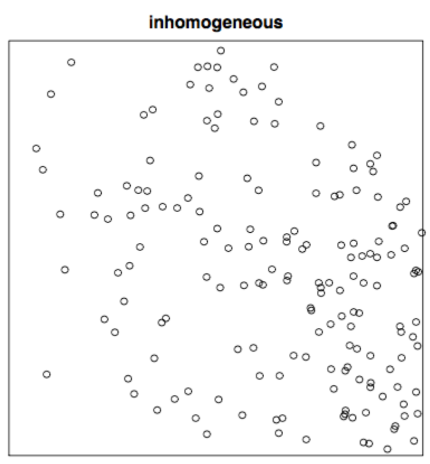
- CIF: $\lambda(t \mid \mathrm{H}_t) = \lambda$
- Independent of history and time
  - complete randomness and lack of memory
- Exponentially distributed waiting times $P(\tau > u) = \exp(-\lambda u)$
- An often-used benchmark
- Number of events in finite interval (S,T] given by discrete Poisson distribution function with parameter λ(T-S)
- Generalises directly to space and space-time process

Example: Inhomogeneous Poisson Process

- CIF: $\lambda(t \mid \mathrm{H}_t) = \lambda(t)$
- Independent of history but depends on time (or location)
  - Often used when there is an apparent trend, seasonality, concentrations
- Number in finite interval (S,T] given by discrete Poisson distribution function with parameter:

$$\Lambda(S,T) = \int_S^T \lambda(t)\,dt$$





Likelihood Function

- Probability of the data under *any* point process model can be (heuristically) derived by multiplying the likelihood of observing events at each t_i and of observing no events elsewhere:

$$L_{(S,T]}(N;t_1,\ldots,t_N) = e^{-\int_S^T \lambda(t)dt} \prod_{i=1}^{N} \lambda(t_i \mid \mathrm{H}_{t_i})$$

- Generalizes (fairly well) to higher dimensions
- Once the CIF is specified, the likelihood can be calculated, providing access to full machinery of likelihood-based inference!

Parameter estimation

- If the parameters of the point process are unknown, one can maximise the log-likelihood to estimate parameters:

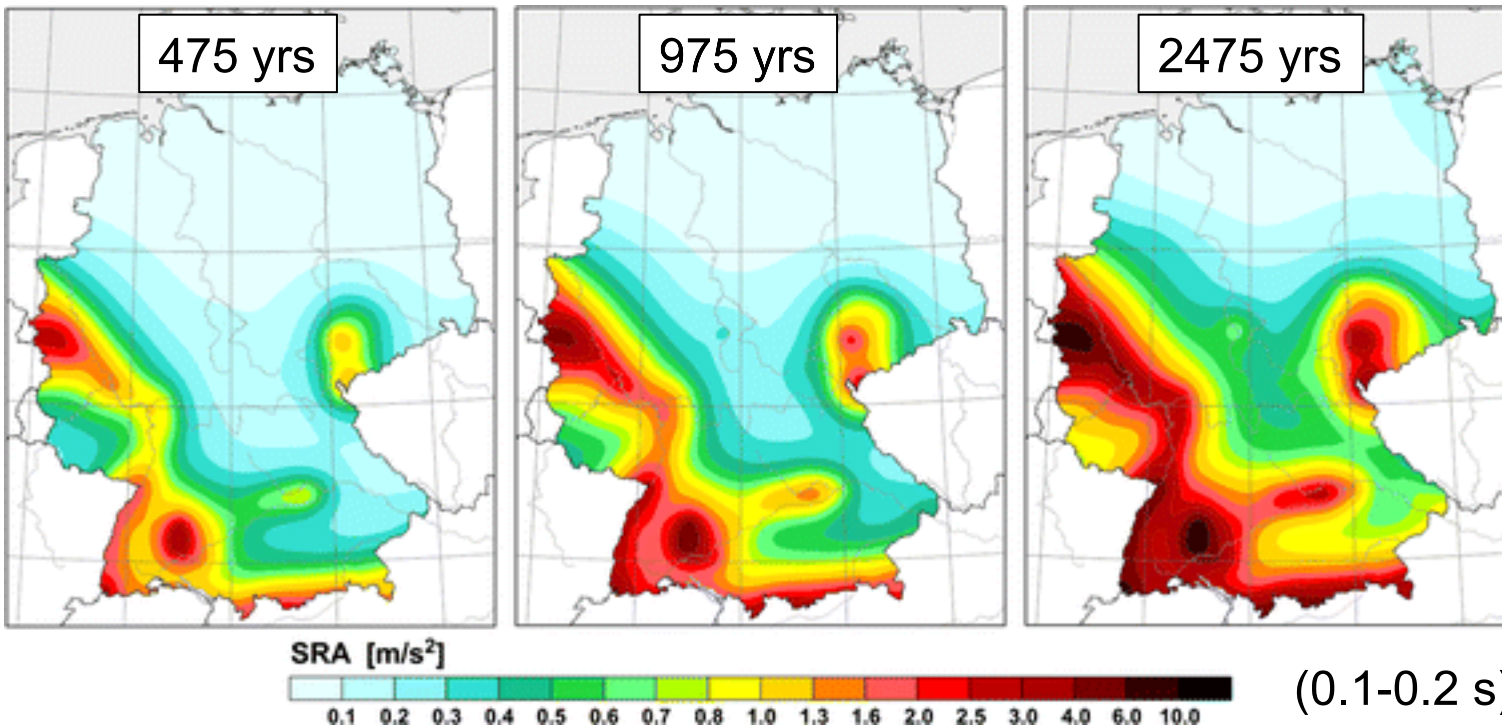
$$\hat{\theta} = \arg_{\theta} \max \log L(N;S,T;\theta)$$

Part 2

Renewal point process models of seismicity

Poisson Process

- Memoryless, uniformly random hazard rate $\lambda(t \mid \mathrm{H}_t) = \lambda$
- Exponential inter-event time distribution $P(\tau > u) = \exp(-\lambda u)$
- Basis of many seismic hazard maps, e.g. Germany 2016:



(0.1-0.2 s)

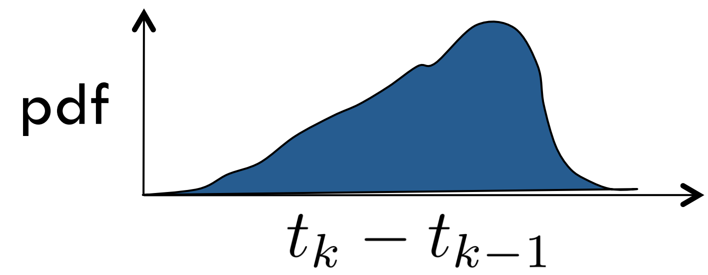
Renewal Processes

- Probability of next event depends on time since last event

$$p(t_{k+1}|t_k) = p(t_{k+1} - t_k)$$

$$\lambda(t) = \frac{p(t - t_{k-1})}{1 - F(t - t_{k-1})}$$

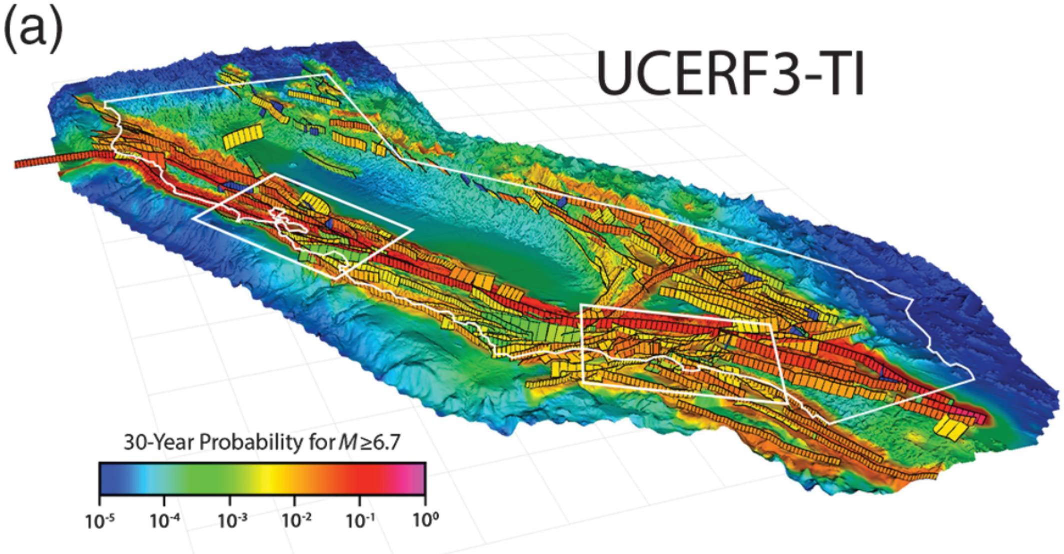
- Popular choices:
  - Lognormal
  - Brownian passage time
  - Weibull
  - ...



- Fit renewal model to earthquake recurrence data from paleoseismology (e.g. digging trenches)
- Often motivated physically by time required to build up energy/stress for next big one ("elastic rebound", Reid, 1910)

e.g. Zhuang et al. 2011, www.corssa.org for a review

Uniform California Earthquake Rupture Forecast (version 3)



Time-independent (TI) probabilities that certain locations in California will participate in one or more M≥6.7 earthquake ruptures during a 30 year interval.

Field et al. (2014, 2015)

Part 3

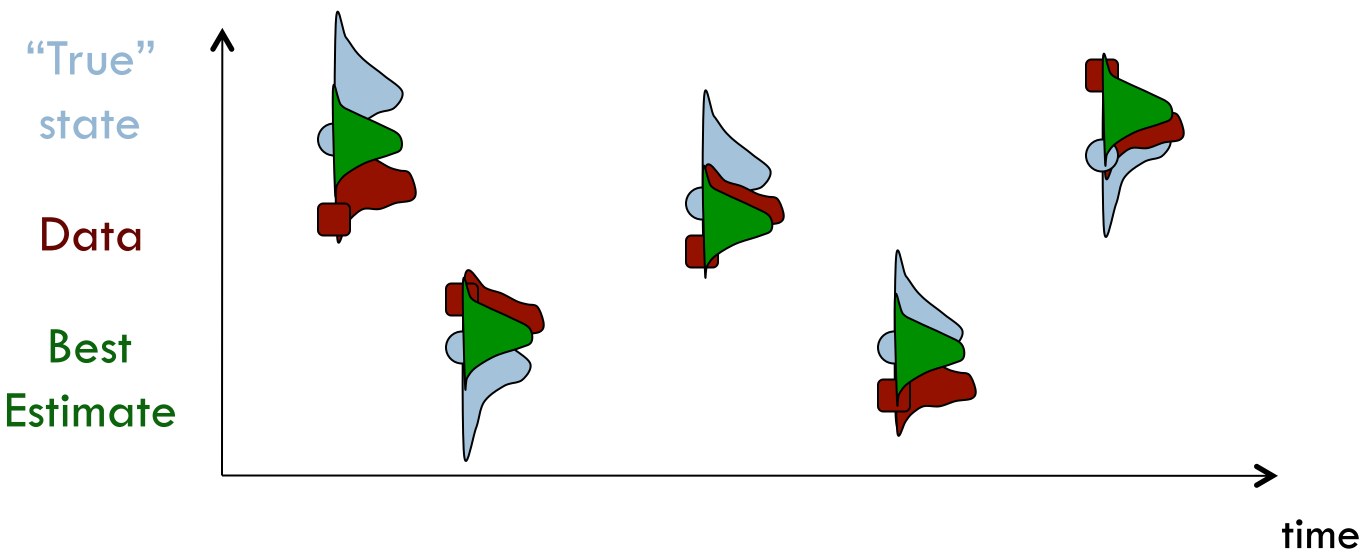
Sequential data assimilation for renewal processes

Forecasting based on data assimilation

- Goal:
  - Develop earthquake forecasting based on data assimilation
    - Taking into account uncertainties in measurements
    - To perform model inference (i.e. likelihood-based selection)
- Challenges
  - Statistical models, based on stochastic point-processes (few models based on partial differential equations)
  - (Strongly) non-Gaussian distributions
  - Long-term memory (spatio-temporal clustering)
- Strategy
  - ‘Start’ with a simple model of seismological relevance:
    - 1d temporal renewal point-process
    - Uncertainties in the observed occurrence times
  - Continue with more complex models (e.g. ETAS)…?

Werner, Ide & Sornette, 2011, Nonlin. Proc. Geophys.

Sequential Data Assimilation



State evolution:

Posterior:

$$p(x_t|y_t) = \frac{p(y_t|x_t)p(x_t)}{p(y)}$$

$$p(x_t|x_{t-1})$$

random model error

model

Observation:

Forecast prior:

$$y_t = H(x_t) + \nu_t$$

$$p(x_t) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{t-1})dx_{t-1}$$

$$p(y_t|x_t)$$

random measurement error

Data Assimilation

- This is a conceptual solution only.

- Analytical solutions exist under restrictive assumptions
  - Kalman filter: Gaussian distributions, linear model
  - Kalman-Levy filter: Levy-stable distributions, linear model
- Approximations:
  - Local Gaussian: extended Kalman filter
  - Ensembles of local Gaussians: ensemble Kalman filter
  - various filters suited for different scenarios
  - Monte Carlo: non-linear model, arbitrary evolving distributions
    - Bayesian MCMC
    - Sequential MC based on sequential importance sampling

Prediction: $p(x_t|y_{1:t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})dx_{t-1}$

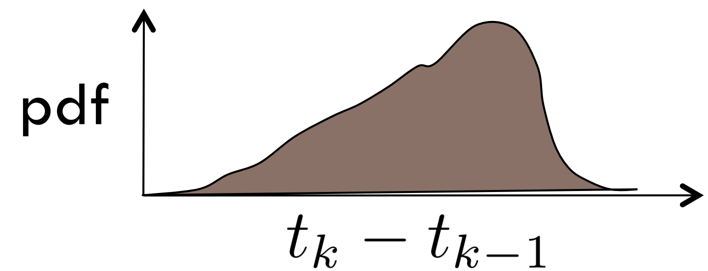
Update: $$p(x_t|y_{1:t}) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t}$$

Numerical Application

Transition kernel (model): temporal renewal process

$$p(t_{k+1}|t_k) = p(t_{k+1} - t_k)$$

e.g. log-normal distribution



Observational uncertainties in occurrence times $\quad t_k^o = t_k^t + \epsilon_k$

Uniform noise: $\quad p_\epsilon(\epsilon_k) = U(-\frac{\Delta}{2}, +\frac{\Delta}{2})$

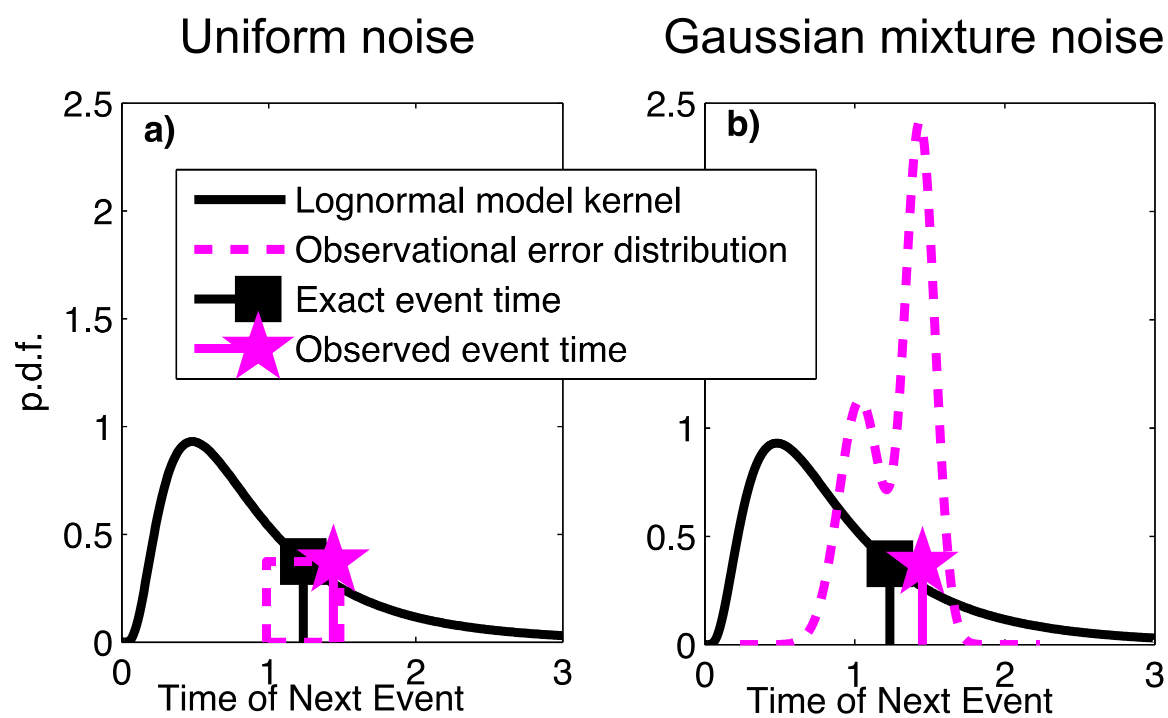
Gaussian mixture noise: $\quad p_{GM}(\epsilon_k) = \omega N(\mu_1, \sigma_1) + (1-\omega) N(\mu_2, \sigma_2)$

Methods

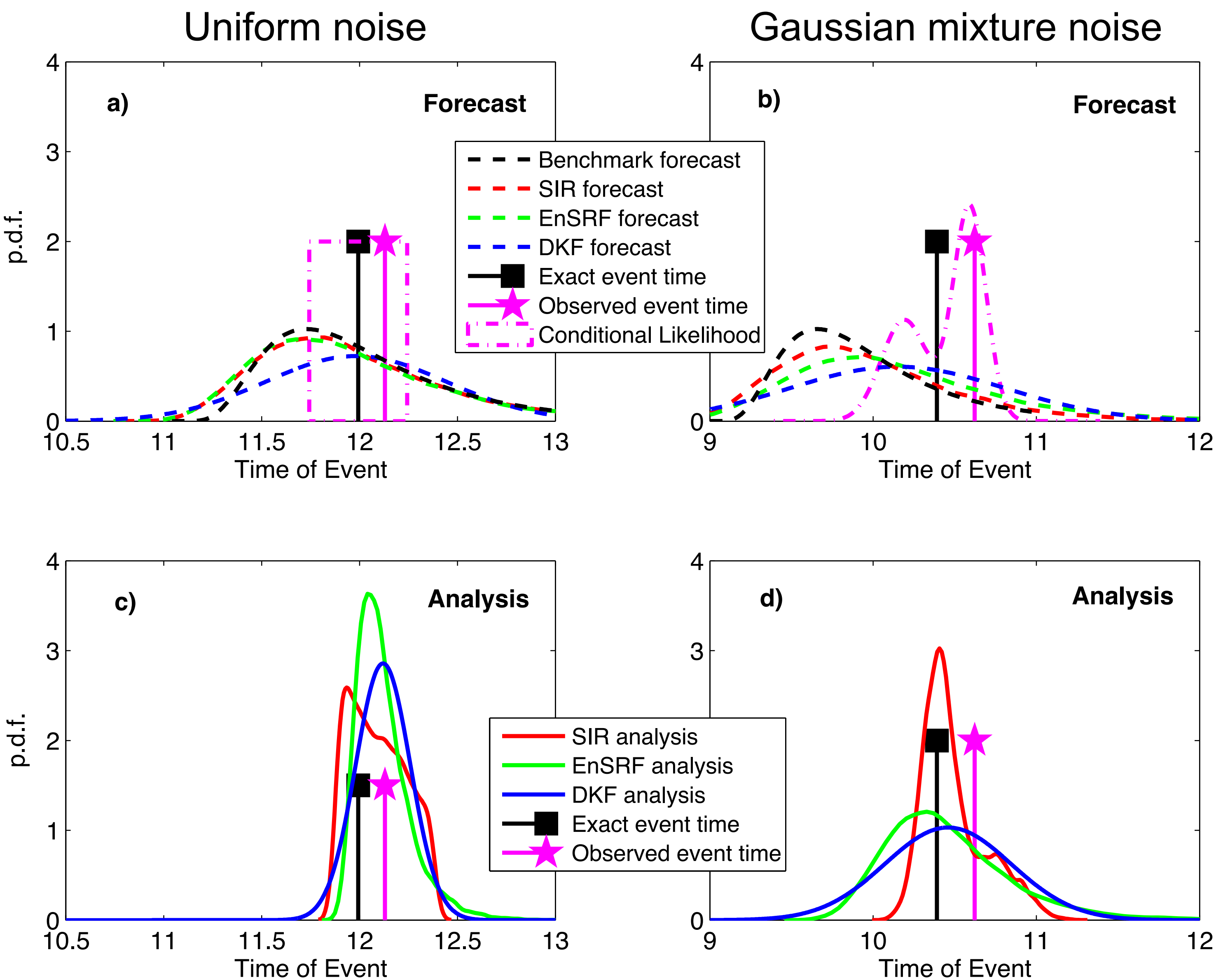
- Benchmark: Ignore observational uncertainty
- DKF: Deterministic Kalman Filter
- EnSRF: a type of Ensemble Kalman Filter (Ensemble Square Root Filter, Tippett, 2003)
- SIR: Sequential Importance Resampling Particle Filter

Numerical experiment

Illustration of state evolution and observational error



Example forecasts and analyses



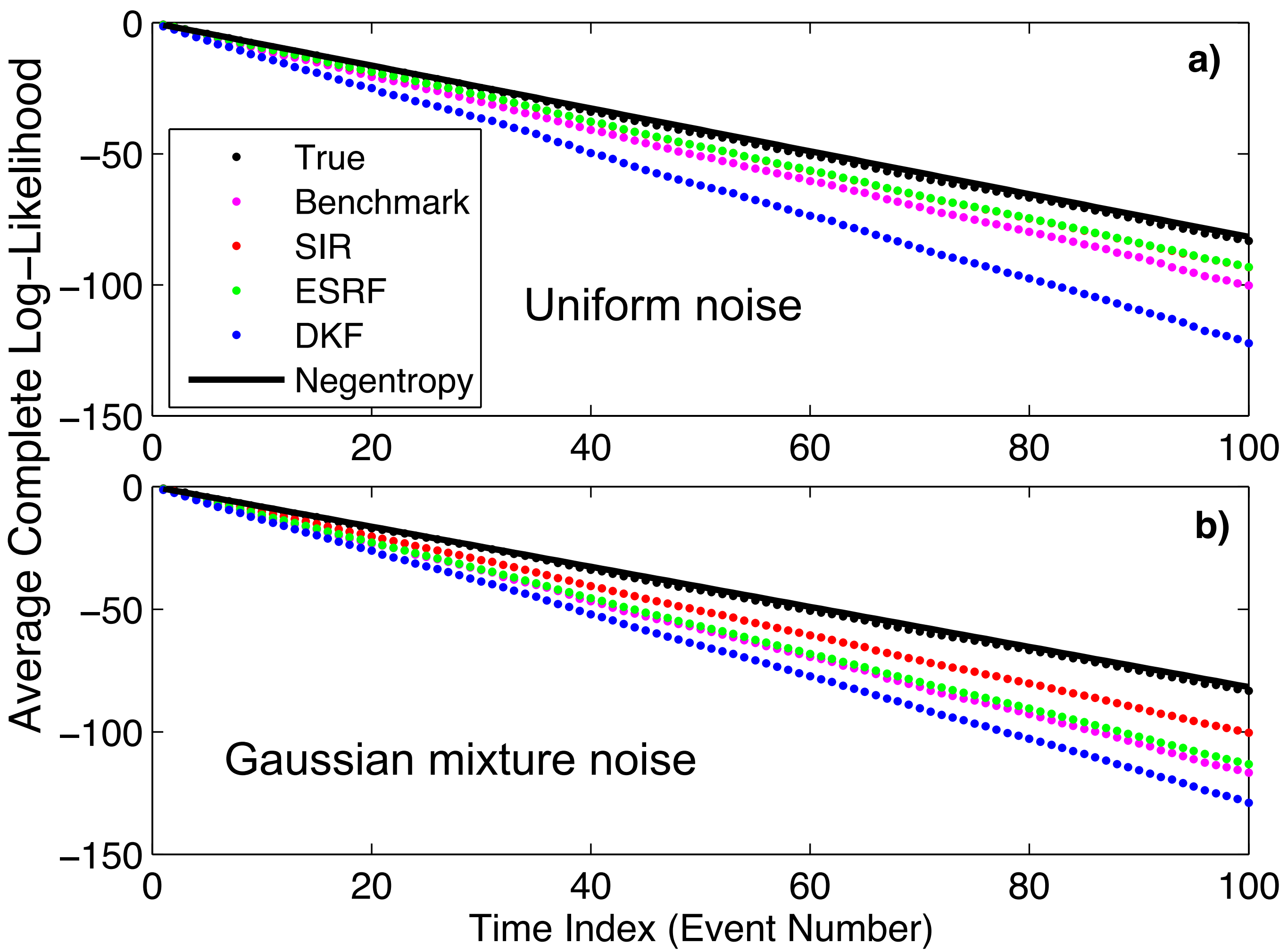
Results: complete log-likelihood scores

Uniform noise:

Particle filter and ensemble Kalman filter perform best

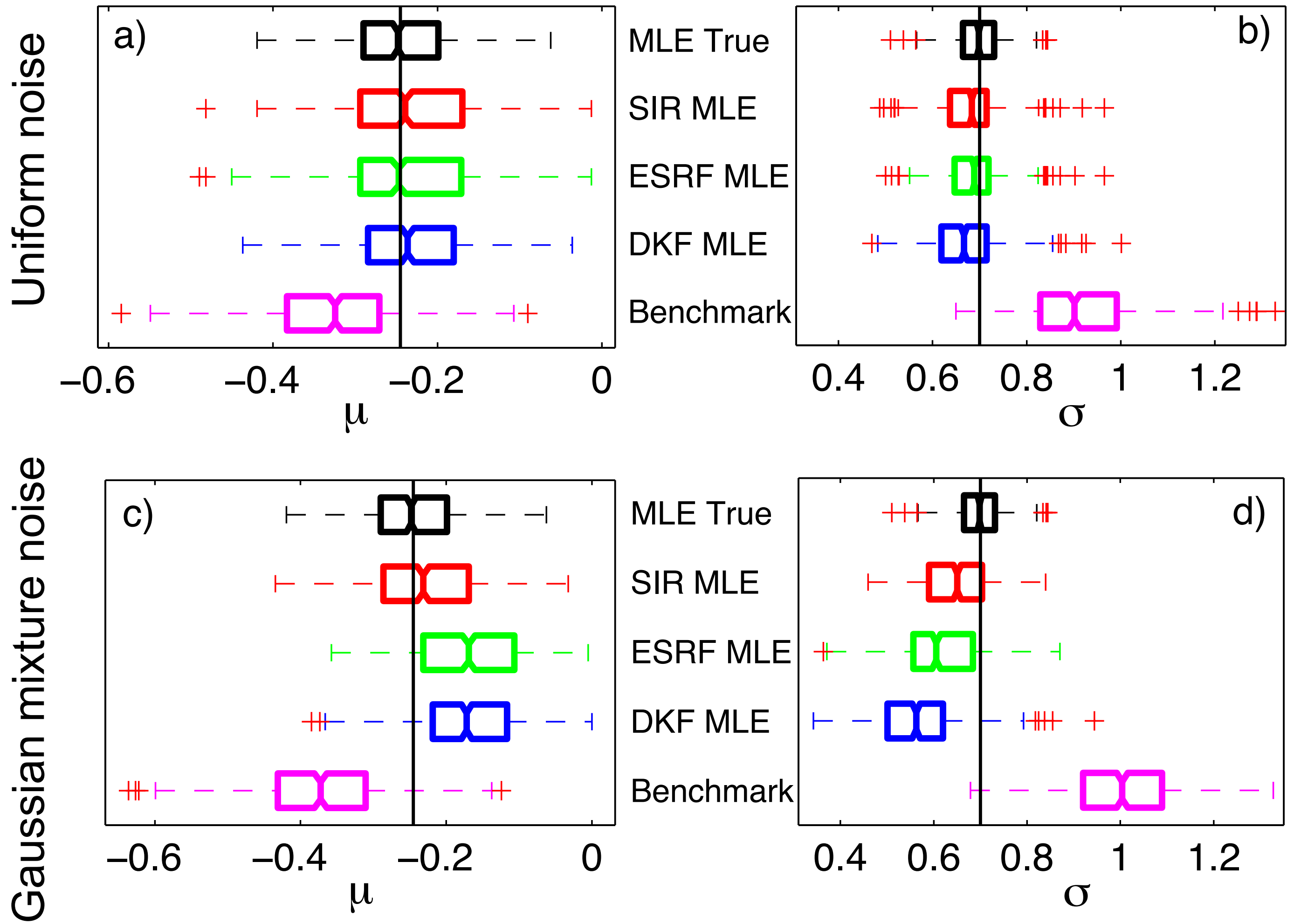
Gaussian mixture noise:

Particle filter performs best



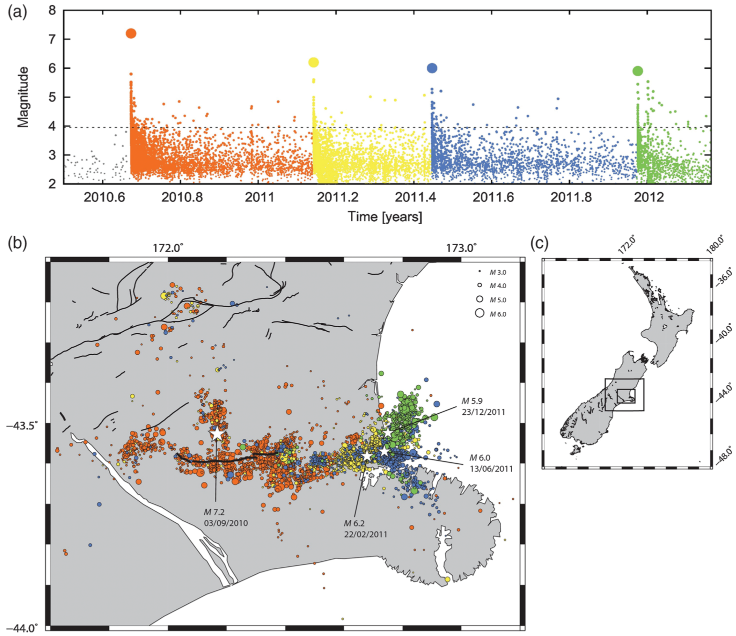
Results: parameter estimates

Particle filter produces least biased parameter estimates for complex observational noise



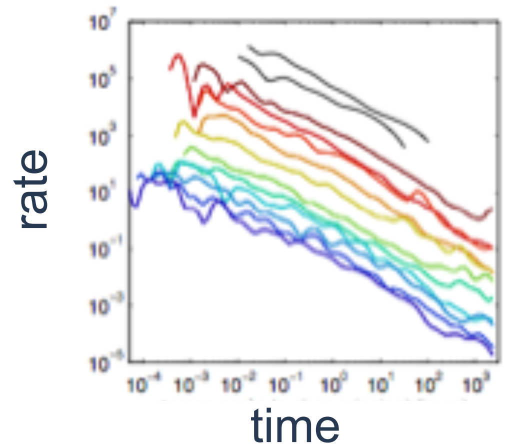
Part 4

The Epidemic Type Aftershock Sequence (ETAS) model



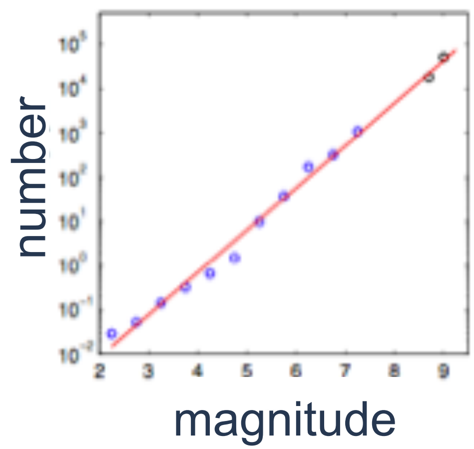
Epidemic Type Aftershock Sequence (ETAS) model

Omori law



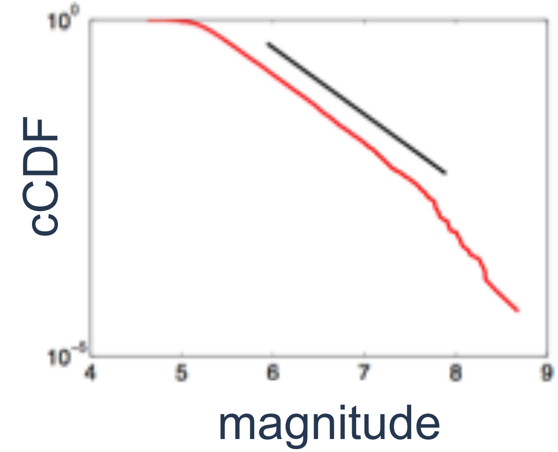
$$n(t) = \frac{\kappa}{(t+c)^p}$$

Productivity law



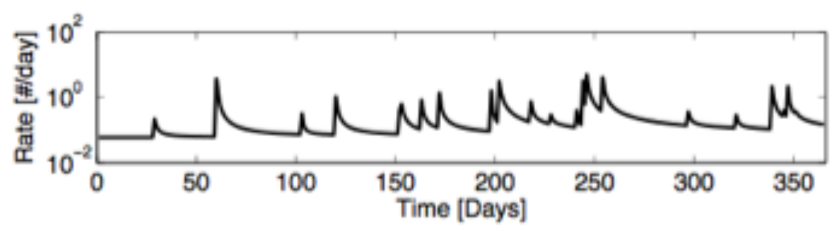
$$\kappa(m) = ke^{\alpha m}$$

Gutenberg-Richter law



$$p(m) = \beta e^{\beta m}$$

$$\lambda(t|H_t) = \mu + \sum_{i|t_i<t} \frac{ke^{\alpha m_i}}{(t-t_i+c)^p}$$



Marked spatio-temporal ETAS model

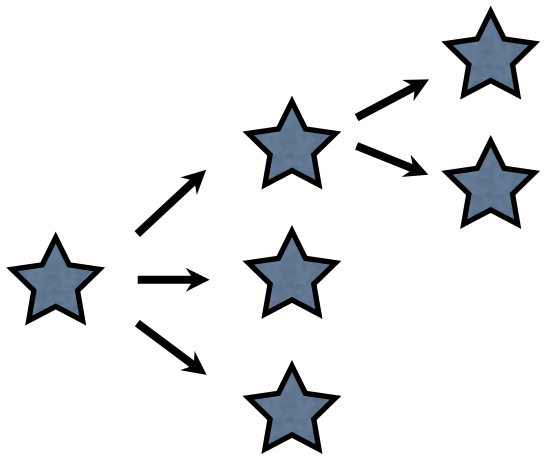
$$\lambda(t, m, \vec{r}|H_t) \quad = p(m) \quad \left[\mu(\vec{r}) + \sum_{i|t_i<t} \kappa(m_i) g(t - t_i) S\left(|\vec{r} - \vec{r_i}|; m_i\right)\right]$$

where

$$\begin{array}{rll} p(m) & \sim 10^{-bm} & \text{Gutenberg-Richter law} \\ \mu(\vec{r}) & & \text{time-independent background rate} \\ \kappa(m_i) & \sim 10^{\alpha m_i} & \text{expected number of offspring} \\ g(t) & \sim (t+c)^{-p} & \text{Omori law} \\ S\left(|\vec{r} - \vec{r_i}|; m_i\right) & & \text{Spatial distribution of direct offspring} \end{array}$$

Epidemic-Type Aftershock Sequences (ETAS) Model

branching process model of seismicity cascades:



spontaneous earthquakes

Gutenberg-Richter law

Omori law

productivity law

spatial triggering kernel

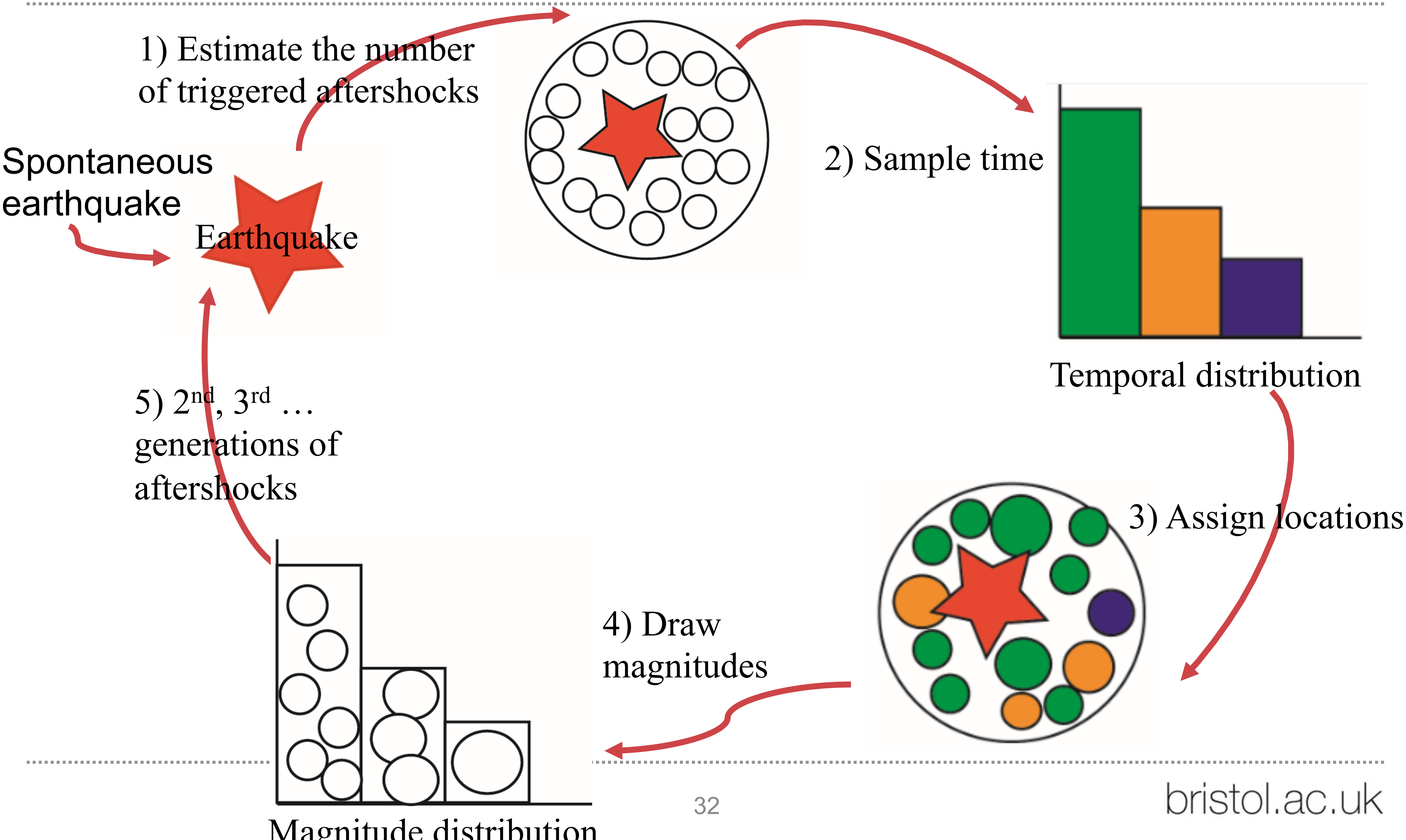
marked spatio-temporal point process

Kagan & Knopoff (1987), Ogata (1988)

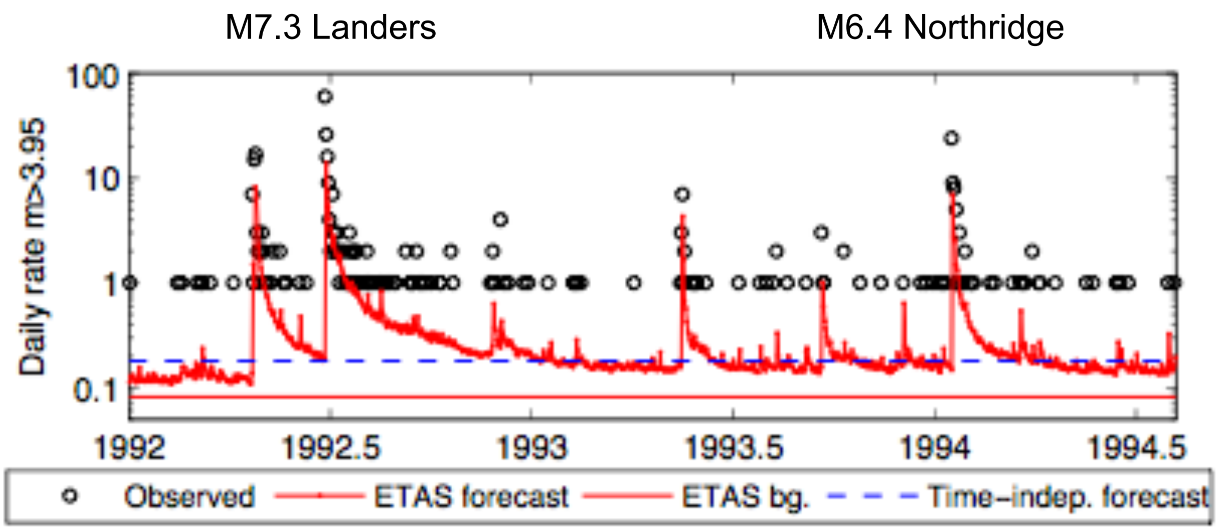
building on Hawkes (1971), Hawkes & Oaks (1974)

Epidemic Type Aftershock Sequence (ETAS) model

University of BRISTOL



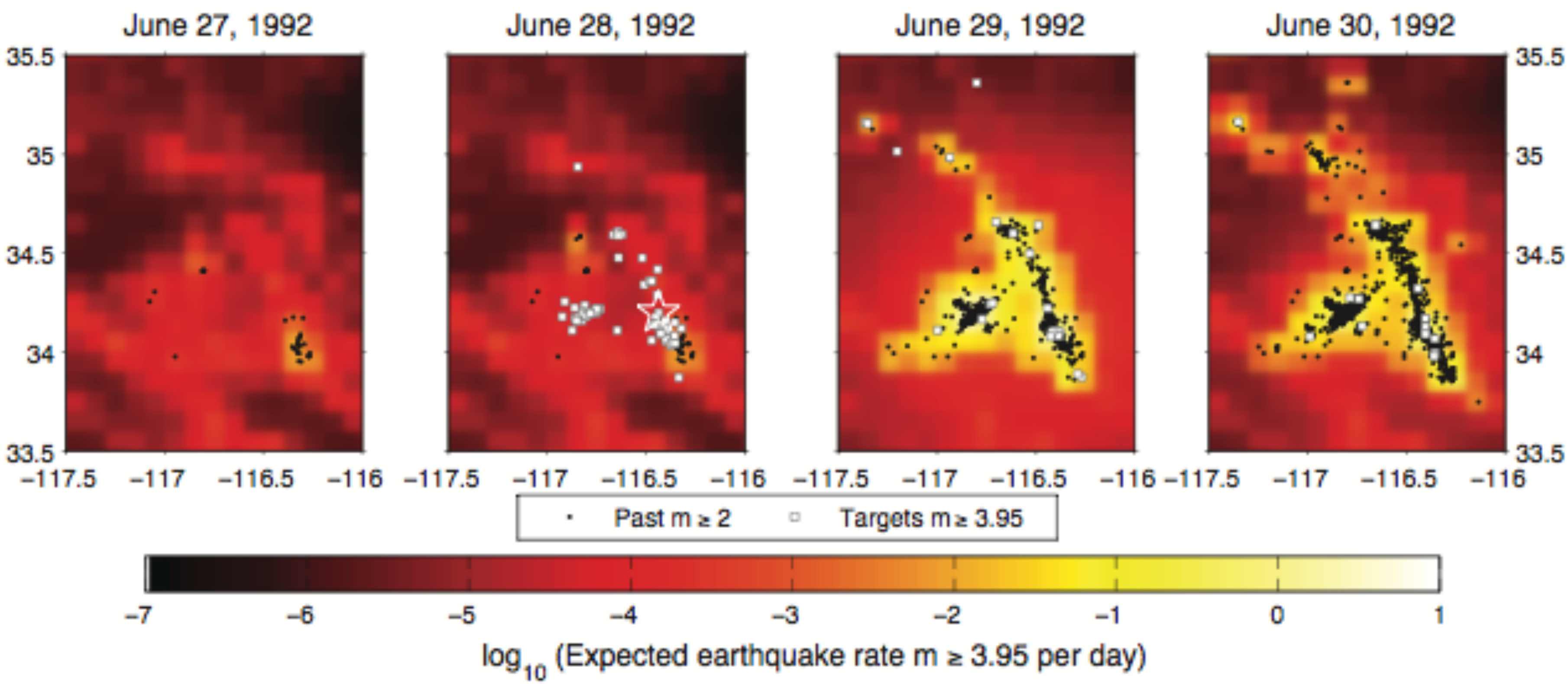
tracking earthquake cascades in time ...



Werner et al. (2011)

tracking cascades in time and space

1992 M7.3 Landers, California, earthquake



Werner et al. (2011)

Part 5

Evaluations of ETAS forecasts

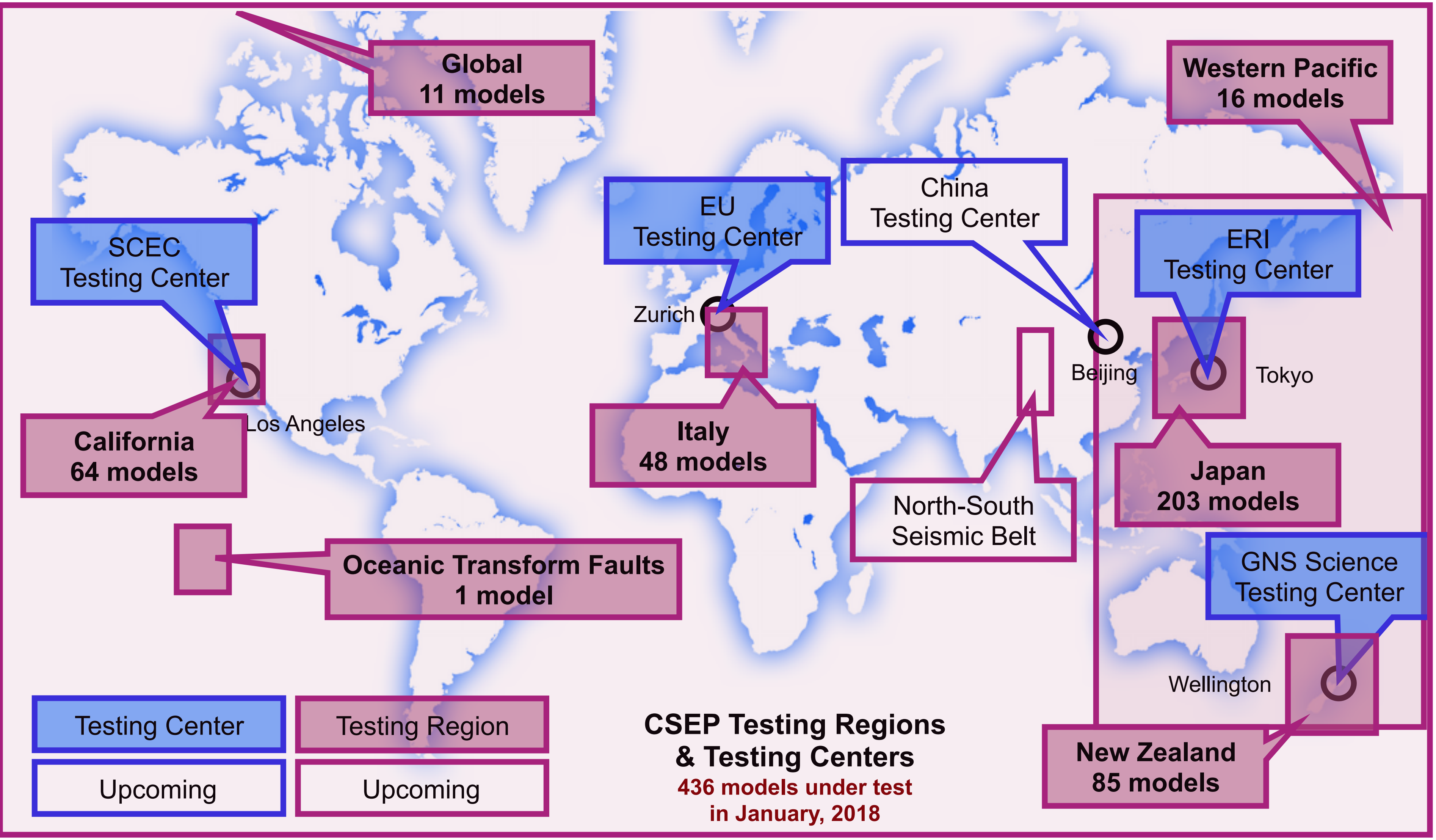
University of BRISTOL

The great paradox of science is that passionate practitioners must carefully produce dispassionate facts [J. Ravetz, *Scientific Knowledge and its Social Problems*]. Meticulous technical and normative judgement, as well as morals are necessary to navigate the forking paths of the statistical garden.

- Saltelli and Stark [Nature, 2018]

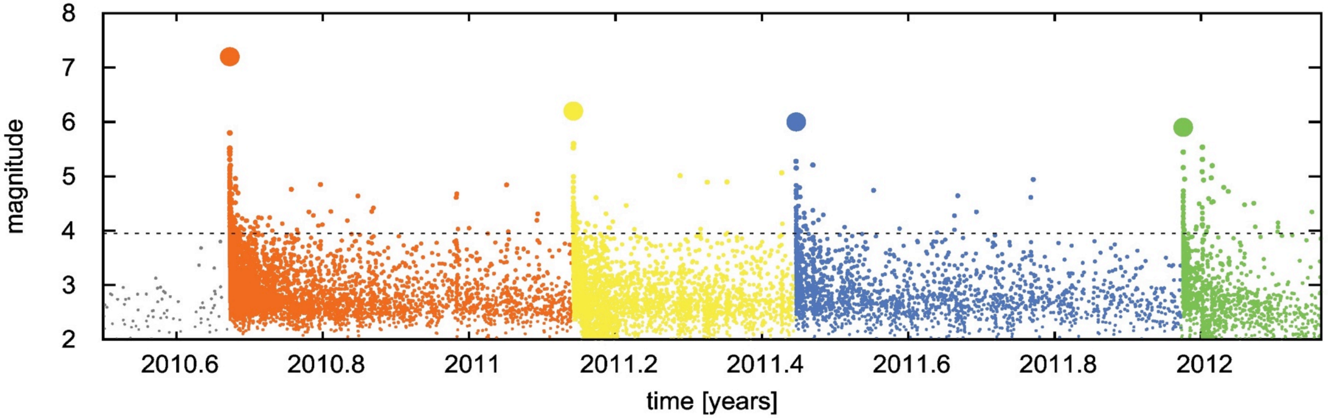
Collaboratory for the Study of Earthquake Predictability

Global platform for blind, prospective and retrospective assessment of forecasting models in a variety of tectonic environments



The Canterbury, NZ, sequence

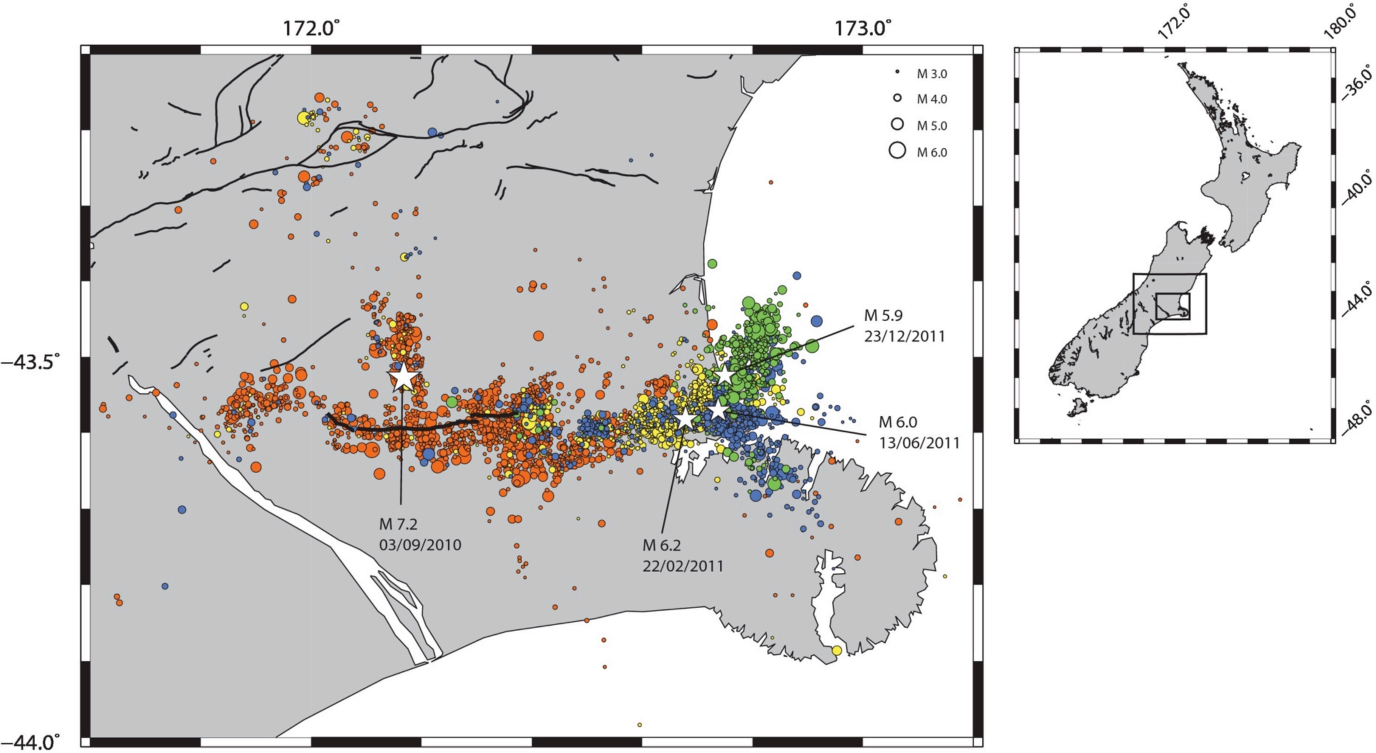
Cattania et al. (2018)

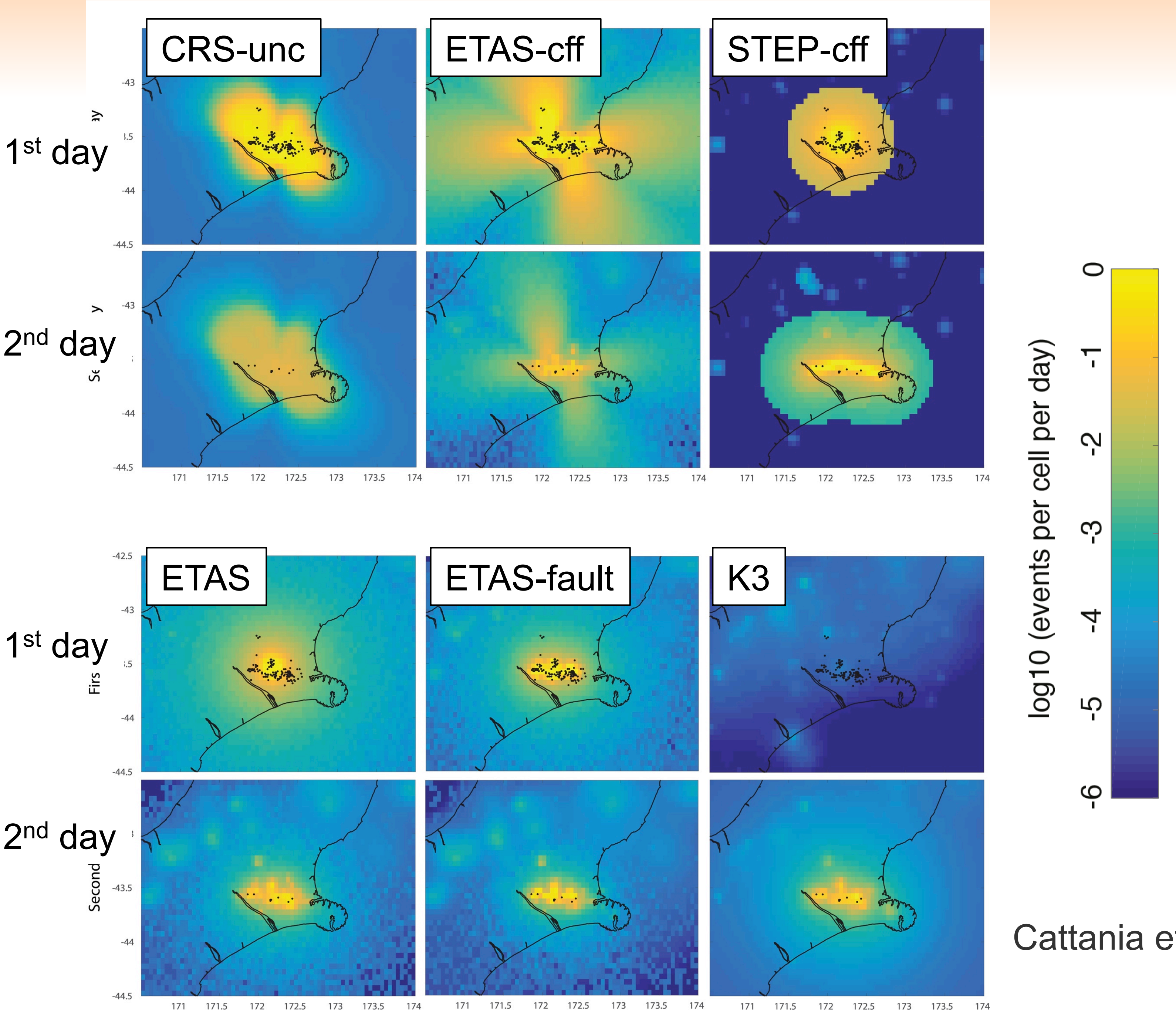


Complex:
- M7.1 Darfield
- M6.2 Christchurch
- M6.0 Christchurch
- M5.9 Christchurch

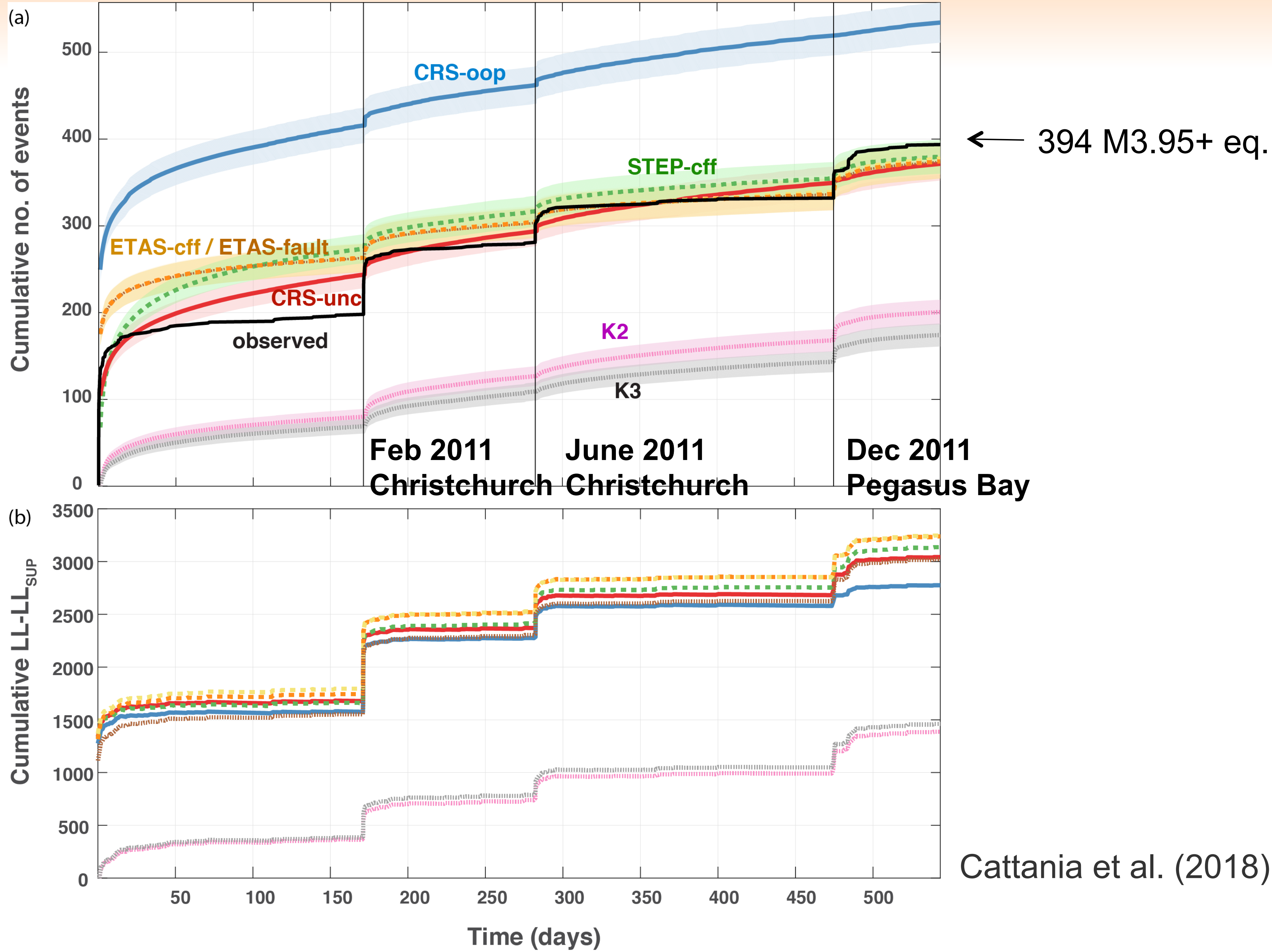
Devastating:
- Over 180 deaths
- $10-15 billion USD

Raised hazard:
- Gerstenberger et al. (2014), Earthquake Spectra



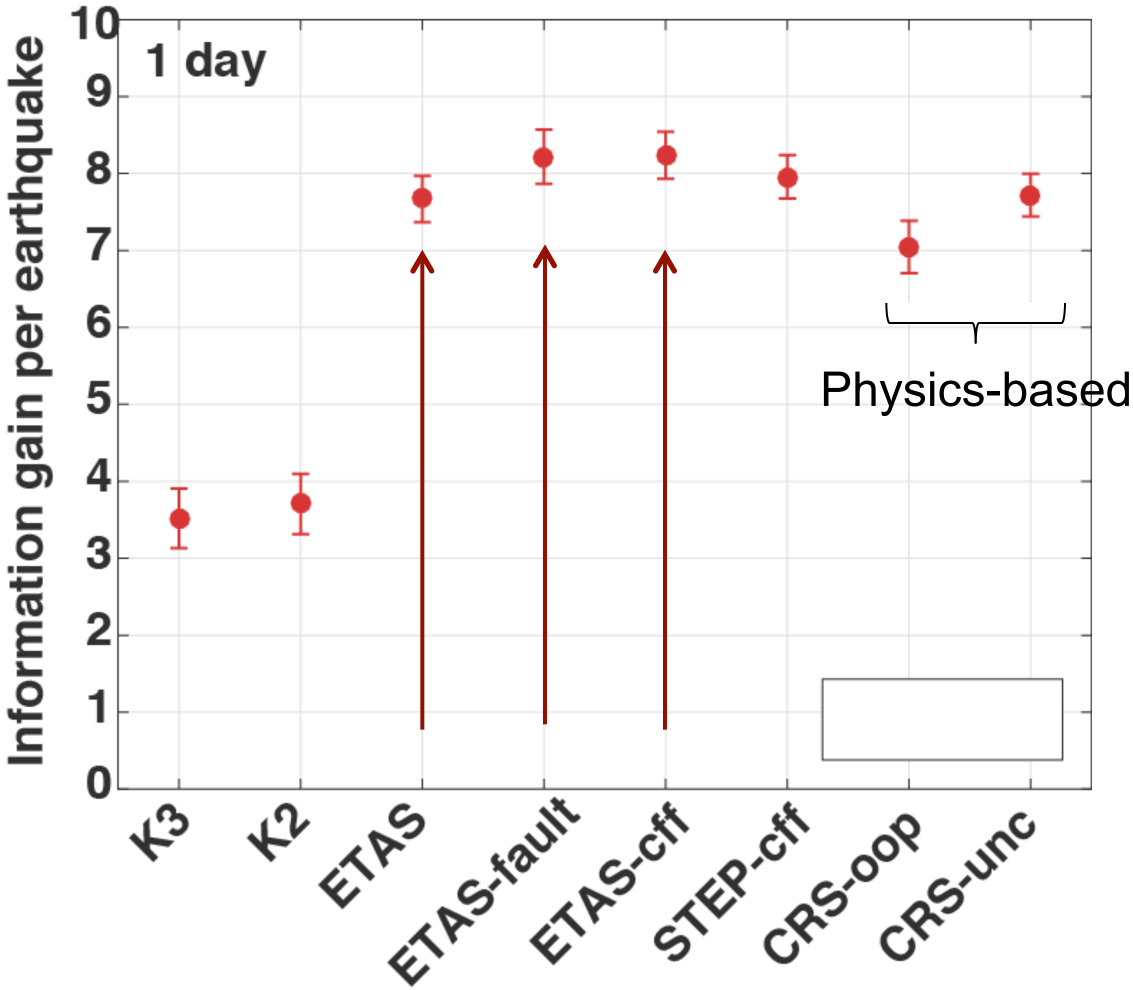


Cattania et al. (2018)



Cattania et al. (2018)

Model ranking of 1-day forecasts



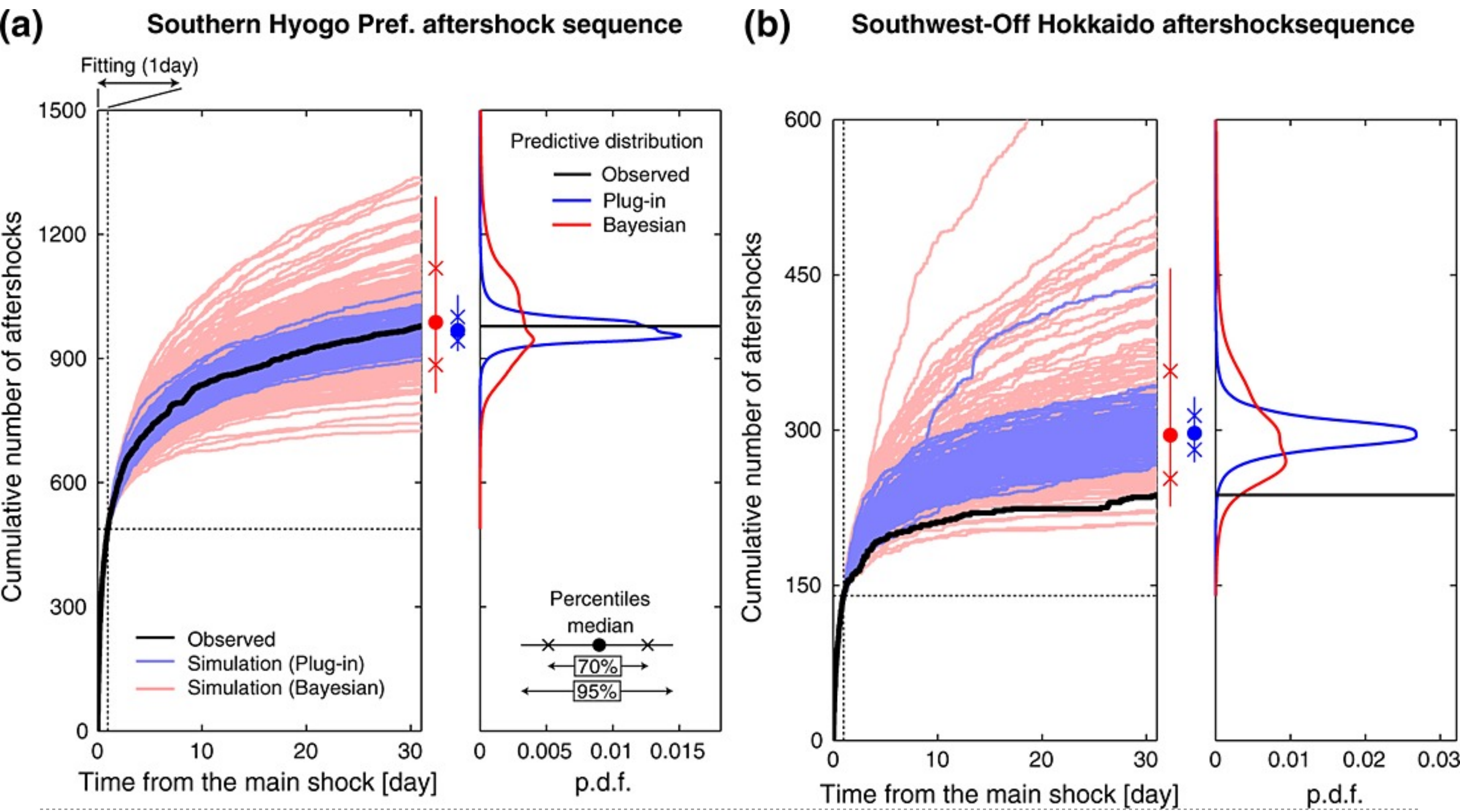
- **Probability gains per earthquake up to 3,000 over time-independent models**
- **Substantial improvements of physics-based models over the past**
- **Hybrids (statistical/ physical) are even more informative**
- **But: ETAS-fault performance about equal**

Part 6

What's new for ETAS modelling?

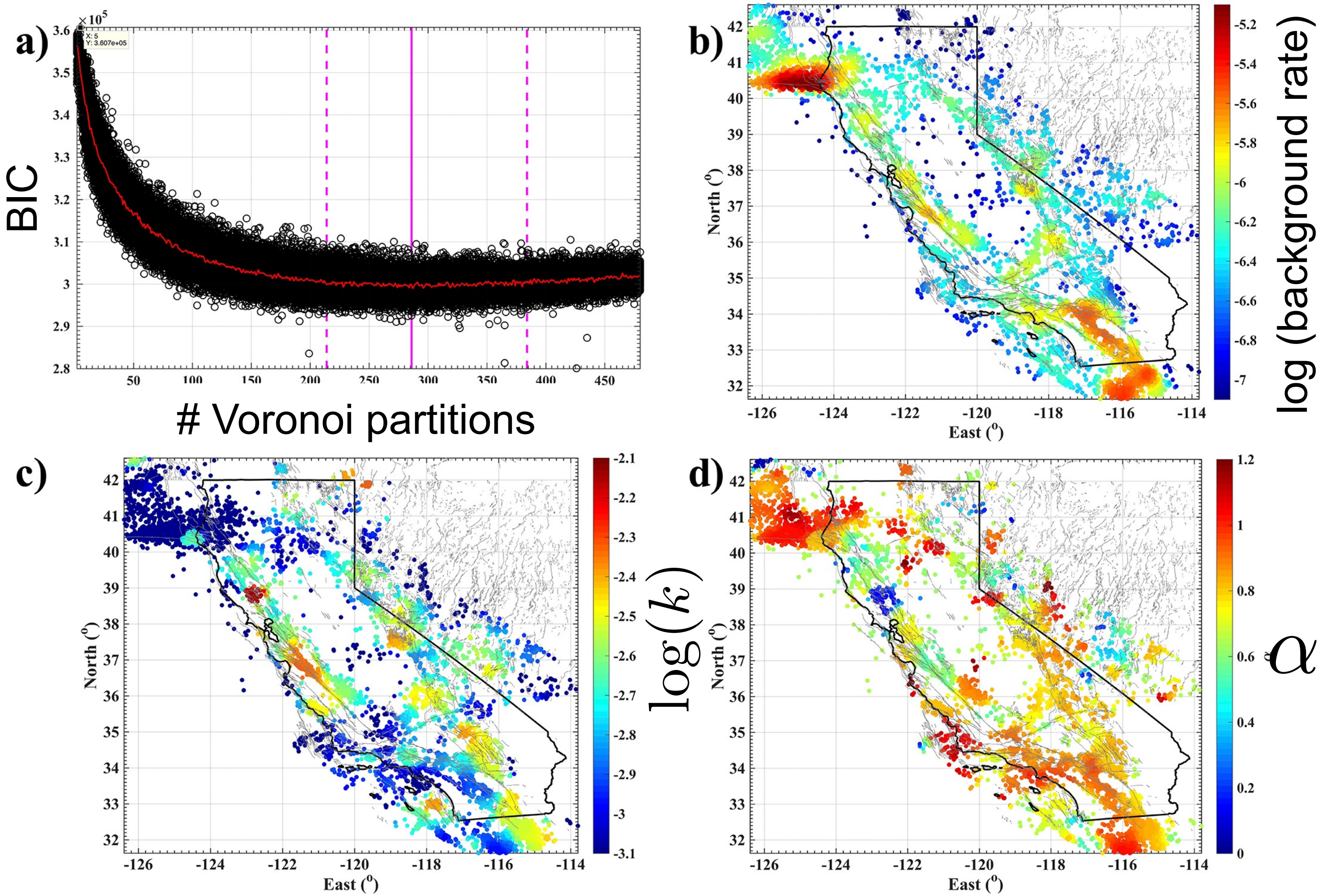
1. Bayesian forecasting (parameter uncertainty)
2. Spatially variable parameters
3. Spatio-temporal clustering in fault systems

Bayesian forecasting



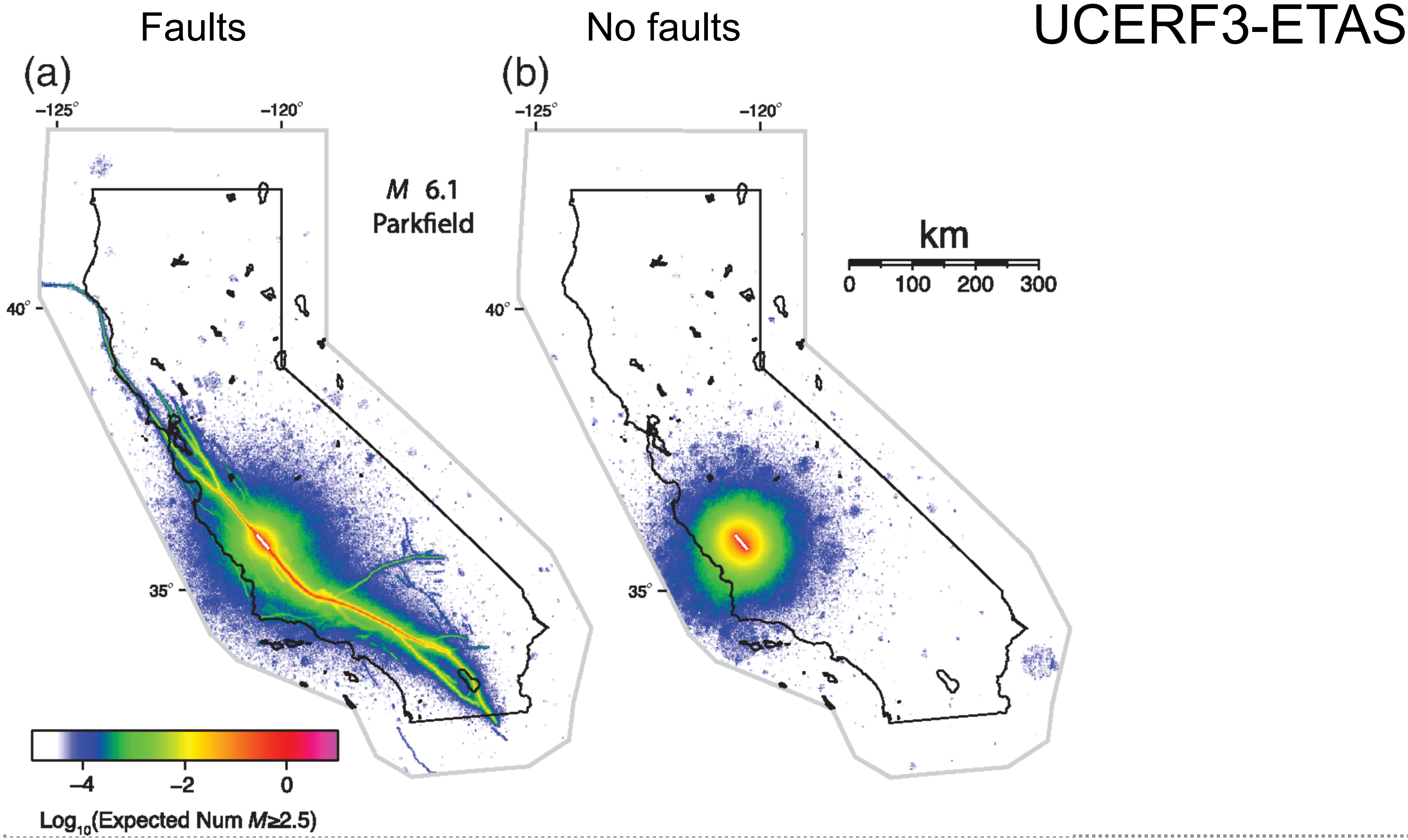
Omi et al., 2013, 2014, 2015, 2016

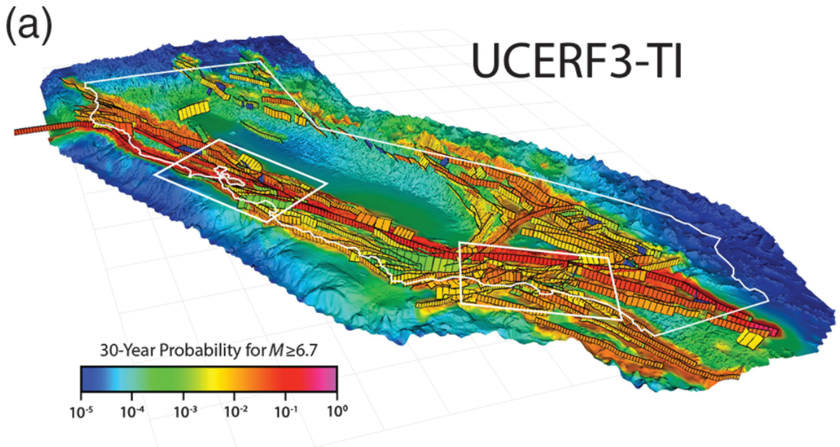
Spatial parameter variability



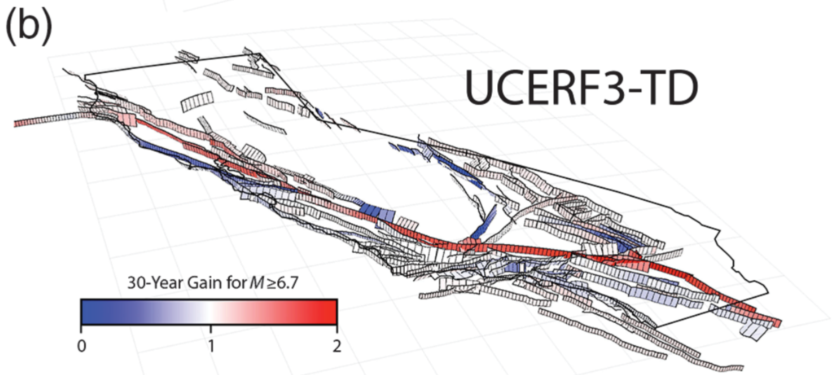
Nandam et al. 2017

Clustering in fault systems

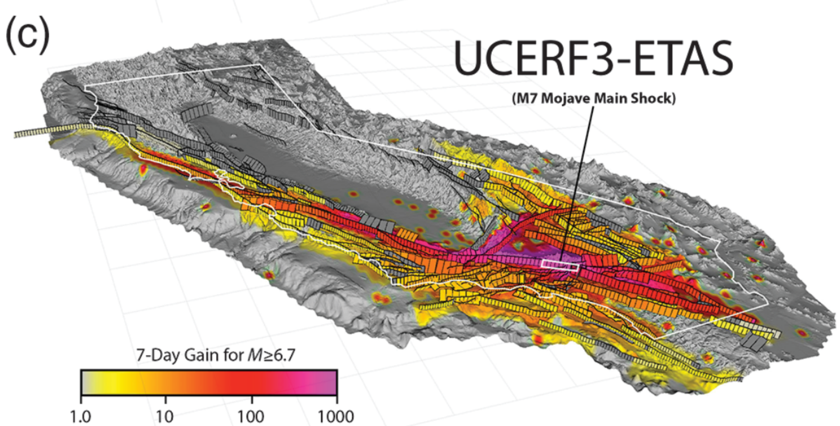




Time-independent (TI) probabilities that certain locations in greater California will participate in one or more M≥6.7 earthquake ruptures during a 30 year interval.



Time-dependent (TD) participation probability gains relative to TI for M≥6.7 fault ruptures during the next 30 year interval.



Spatio-temporal clustering in fault networks:

7-day probability gain over TI after M7 scenario earthquake

Field et al. (2017)

How data assimilation could help:

- Include observational (and parameter) uncertainty in ETAS estimation and forecasting
- Incorporate additional information into ETAS models, e.g. self-consistent coupling of renewal models (faults) and ETAS (clustering)
- Provide *real-time* updating algorithms for public operational earthquake forecasting by government agencies
- Relate ETAS model to physical models of faults...

End of the curious tour

Thank you!