Online forecasting and offline optimization: daily earthquake forecasts by using a model that requires heavy computation

Jiancang Zhuang

Institute of Statistical Mathematics
Tokyo

Presentation for Japan CSEP Workshop, 2009-05-27, ERI, Univ. Tokyo

Outlines

PART I. Review: online forecast and offline optimization structure for using ETAS model forecasting seismicity in California t

1. *Space-time ETAS model*
2. *online forecast and offline optimization structure*
3. *Output examples of ETAS model.*

PART II. Another model with heavy computation: local likelihood ETAS model

1. *Local likelihood ETAS models*
2. *Initial results from local likelihood ETAS models*
3. *Realization of local likelihood ETAS models*

PART I. Review: online forecast and offline optimization structure for using ETAS model forecasting seismicity in California

Space-Time Epidemic Type Aftershock Sequence (ETAS) model

§ Seismicity rate = "background" + “Triggered seismicity":

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i)\right]$$

§ Time distribution:
the Omori-Utsu law

$$g(t) = \frac{p-1}{c}\left(1+\frac{t}{c}\right)^{-p}, \quad t>0$$

§ Spatial location distribution of children:

$$f(x,y;m) = \frac{q-1}{\pi D e^{\gamma(m-m_c)}}\left(1+\frac{x^2+y^2}{D e^{\gamma(m-m_c)}}\right)^{-q}, \quad q>1$$

§ productivity: mean number of children

$$\kappa(m) = A e^{\alpha(m-m_c)}, \quad m \ge m_c$$

Space-time ETAS model

- Conditional intensity

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i)\right]$$

- Likelihood function

$$\log L = \sum_{(t_i,x_i,y_i,m_i)\in[0,T]\times A\times M} \log \lambda(t_i,x_i,y_i,m_i) - \int_0^T \iint_A \int_M \lambda(t,x,y,m)\,dt\,dx\,dy\,dm$$

Space-time ETAS model

- Time varying seismicity rate (conditional intensity or stochastic intensity) at event j

$$\lambda(t_j, x_j, y_j, m_j) = s(m_j)\left[\mu(x_j, y_j) + \sum_{i:t_i<t} \kappa(m_i) g(t_j - t_i) f(x_j - x_i, y_j - y_i)\right]$$

Contribution from background seismicity

$$\frac{s(m_j)\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j, m_j)}$$

Contribution from the *i*-th event

$$\frac{s(m_j) g(t_j - t_i) f(x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j, m_j)}$$

Space-time ETAS model

- For each event j

Pr{event j is from background} $\qquad \varphi_j = \dfrac{s(m_j)\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j, m_j)}$

Pr{event j is from i} $\qquad \rho_{ij} = \dfrac{s(m_j)g(t_j - t_i)f(x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j, m_j)}$

- Conventional Declustering

  Pr{event j is from background}$= 0$ or 1

Thinning method

- For each event j

Pr{event j is from background} $\varphi_j = \dfrac{s(m_j)\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j, m_j)}$

Pr{event j is from i} $\rho_{ij} = \dfrac{s(m_j)g(t_j - t_i)f(x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j, m_j)}$

Stochastic declustering: Set event j to be a background event or a child of event 1, 2, …,j-1, according to probabilities φ_j or $\rho_{1j}, \rho_{2j}, \ldots, \rho_{j-1,j}$ respectively

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m)=s(m)\left[\mu(x,y)+\sum_{i:t_i<t}\kappa(m_i)g(t-t_i)f(x-x_i,y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

How to estimate background seismicity?

How to estimate clustering parameters?

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

How to estimate background seismicity?

How to estimate clustering parameters?

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\,g(t-t_i)\,f(x-x_i, y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

Kernel, spline, tessellation, histogram, …

How to estimate background seismicity?

How to estimate clustering parameters?

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

Kernel, spline, tessellation, histogram, …

How to estimate background seismicity?

How to estimate clustering parameters?

Maximum likelihood estimate if background seismicity **μ** is know

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m) = s(m)\left[\mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

Kernel, spline, tessellation, histogram, …

How to estimate background seismicity?

?

How to estimate clustering parameters?

Maximum likelihood estimate if background seismicity $\boldsymbol{\mu}$ is know

Estimation problems

- Time varying seismicity rate (conditional intensity or stochastic intensity)

$$\lambda(t,x,y,m)=s(m)\left[\mu(x,y)+\sum_{i:t_i<t}\kappa(m_i)g(t-t_i)f(x-x_i,y-y_i)\right]$$

How to estimate time-free total seismicity $\lambda(x,y)$?

How to estimate background seismicity?

How to estimate clustering parameters?

Kernel, spline, tessellation, histogram, …

Kernel, spline, tessellation, histogram, …, with each event weighted by φ_j

Solution—estimating parameters and background rate simultaneously

Algorithm:

1. Assume an initial background rate.
2. Using MLE to estimate parameters in the clustering structures.
3. Using the assumed background and estimated clustering parameters to evaluate φ_j.
4. Using φ_j to get a better background rate.
5. Update the background rate by this better one.
6. Go to Steps 2 to 5 until results converge.

φ_j: Estimate of probability that event j is of background

Simulation algorithm

- Generate the background catalog with the estimated background rate, recorded as Generation 0.

- For each event, in the last simulated generation, generate its children, with their occurrence times, locations and magnitude from the pdf.s as assumed in the model, where the number of children is a Poisson random variable with a mean of the productivity function.

- Repeat last step until no more new event is generation. Return with all the events in all generations

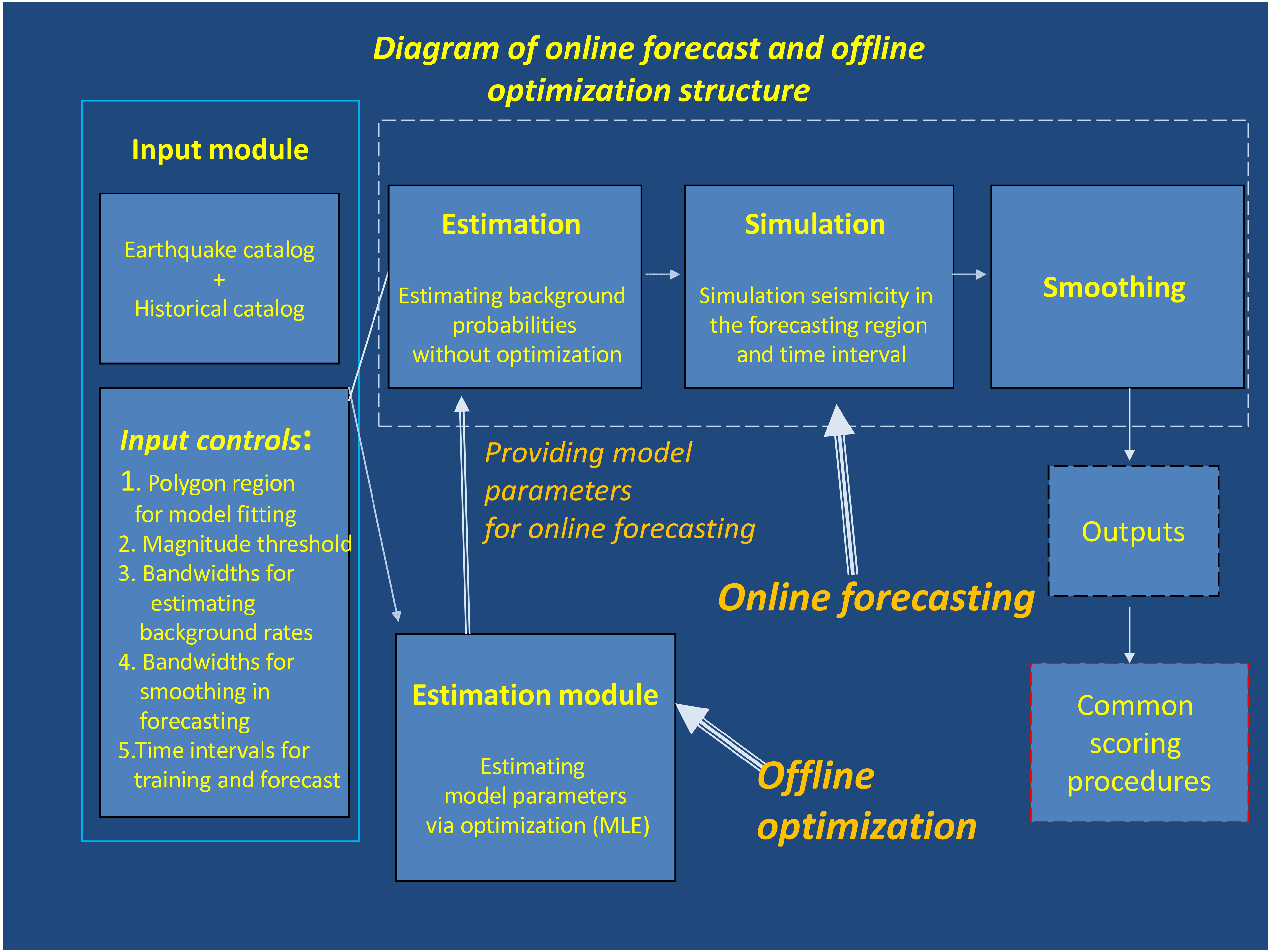
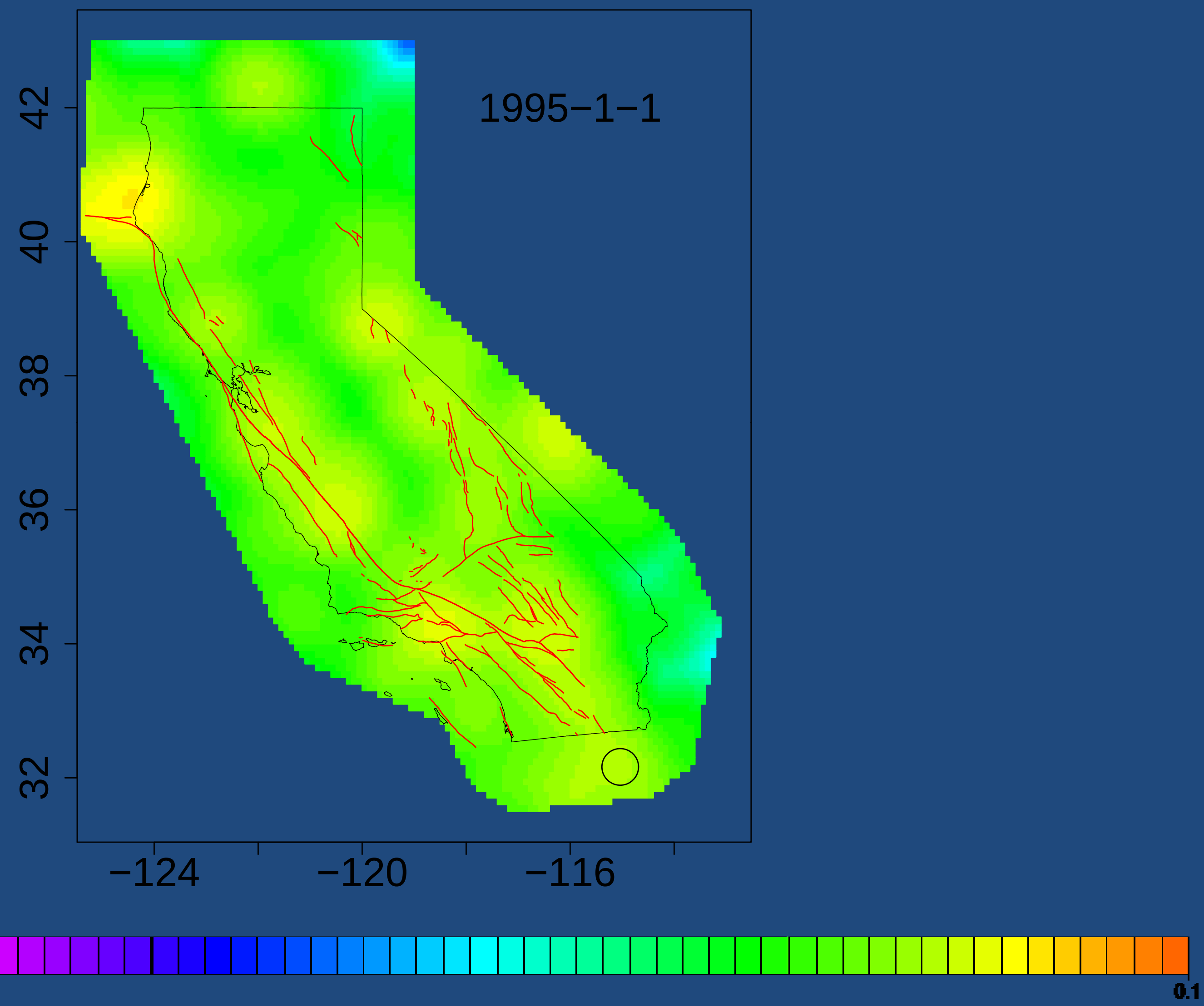
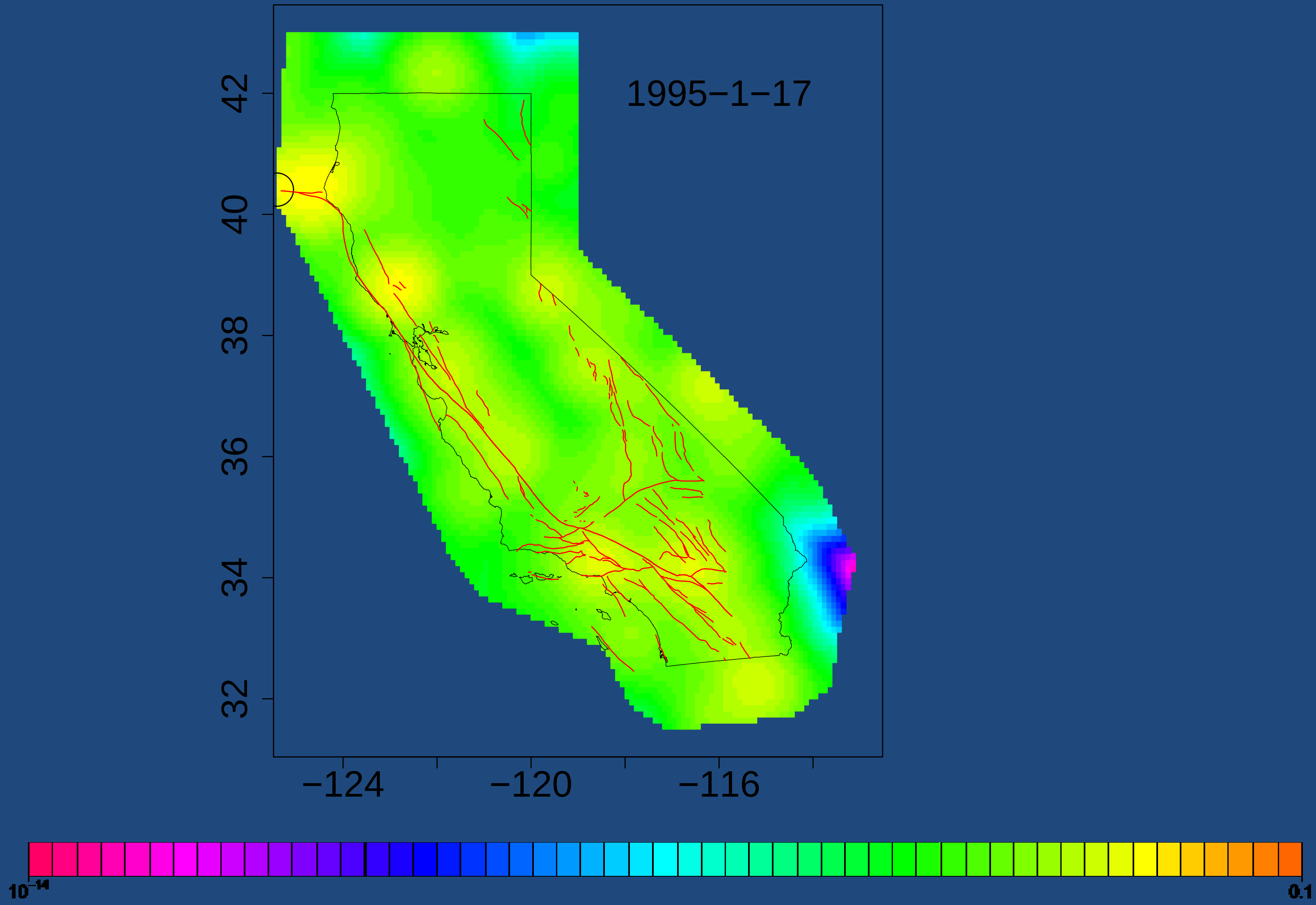


Diagram of online forecast and offline optimization structure
Input module
Earthquake catalog
+
Historical catalog
Input controls:
1. Polygon region for model fitting
2. Magnitude threshold
3. Bandwidths for estimating background rates
4. Bandwidths for smoothing in forecasting
5.Time intervals for training and forecast
Estimation
Estimating background probabilities without optimization
Simulation
Simulation seismicity in the forecasting region and time interval
Smoothing
Providing model parameters for online forecasting
Online forecasting
Outputs
Common scoring procedures
Estimation module
Estimating model parameters via optimization (MLE)
Offline optimization

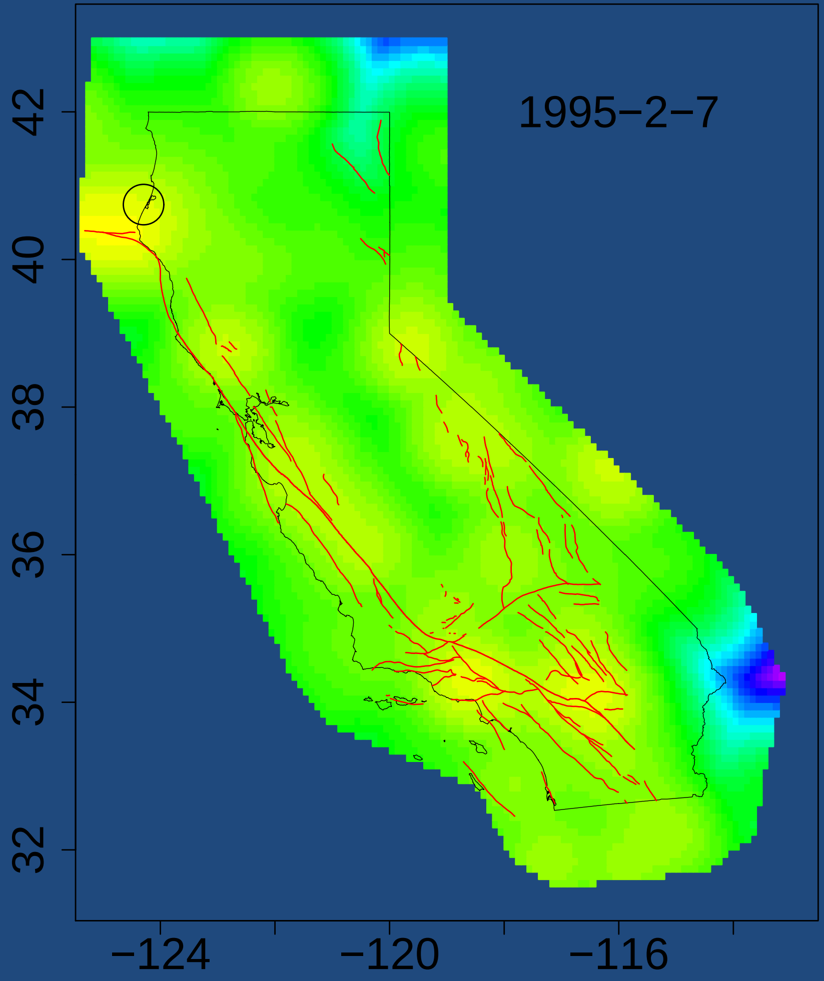
Daily forecasting for California by ETAS



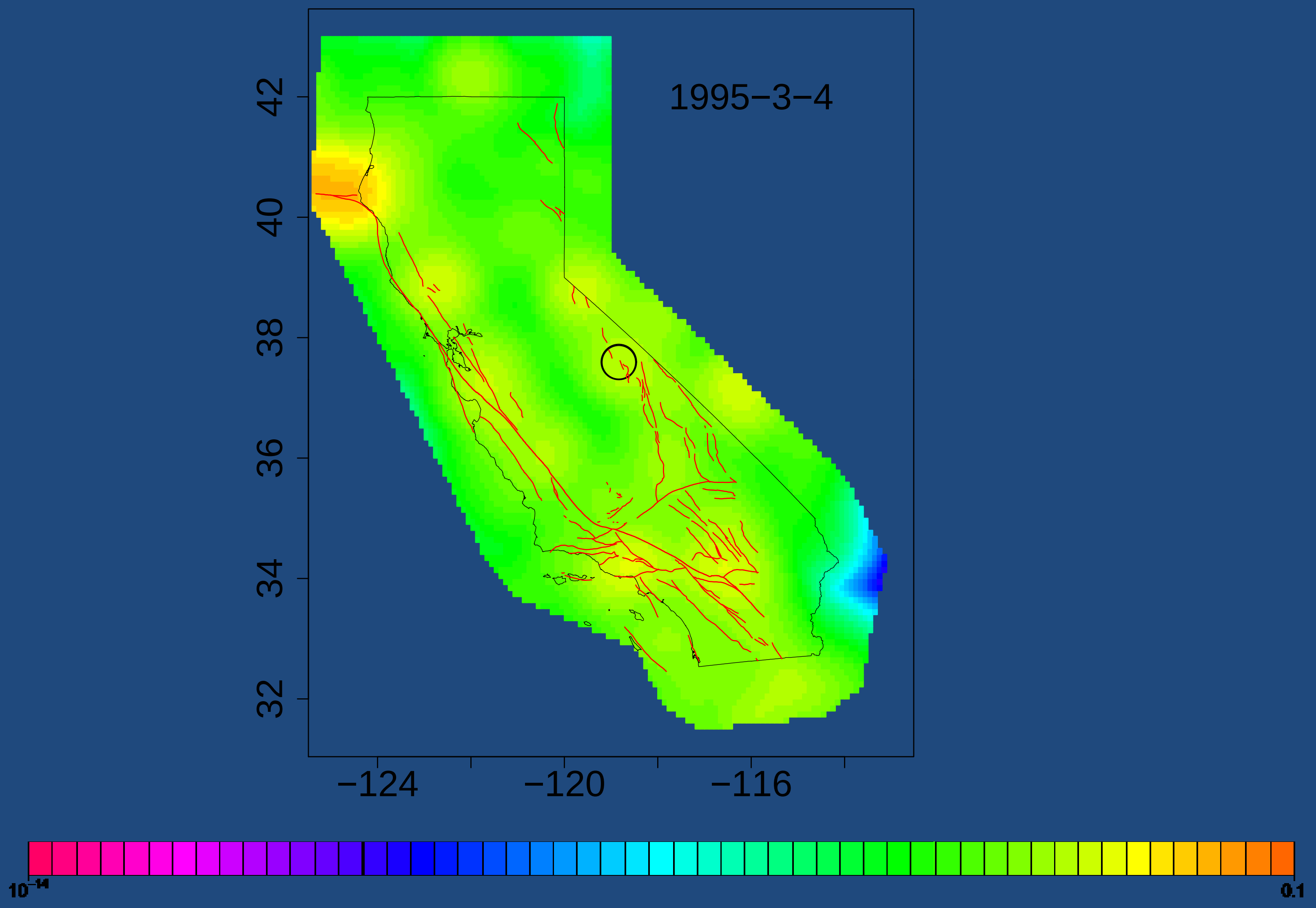
Daily forecasting for California by ETAS



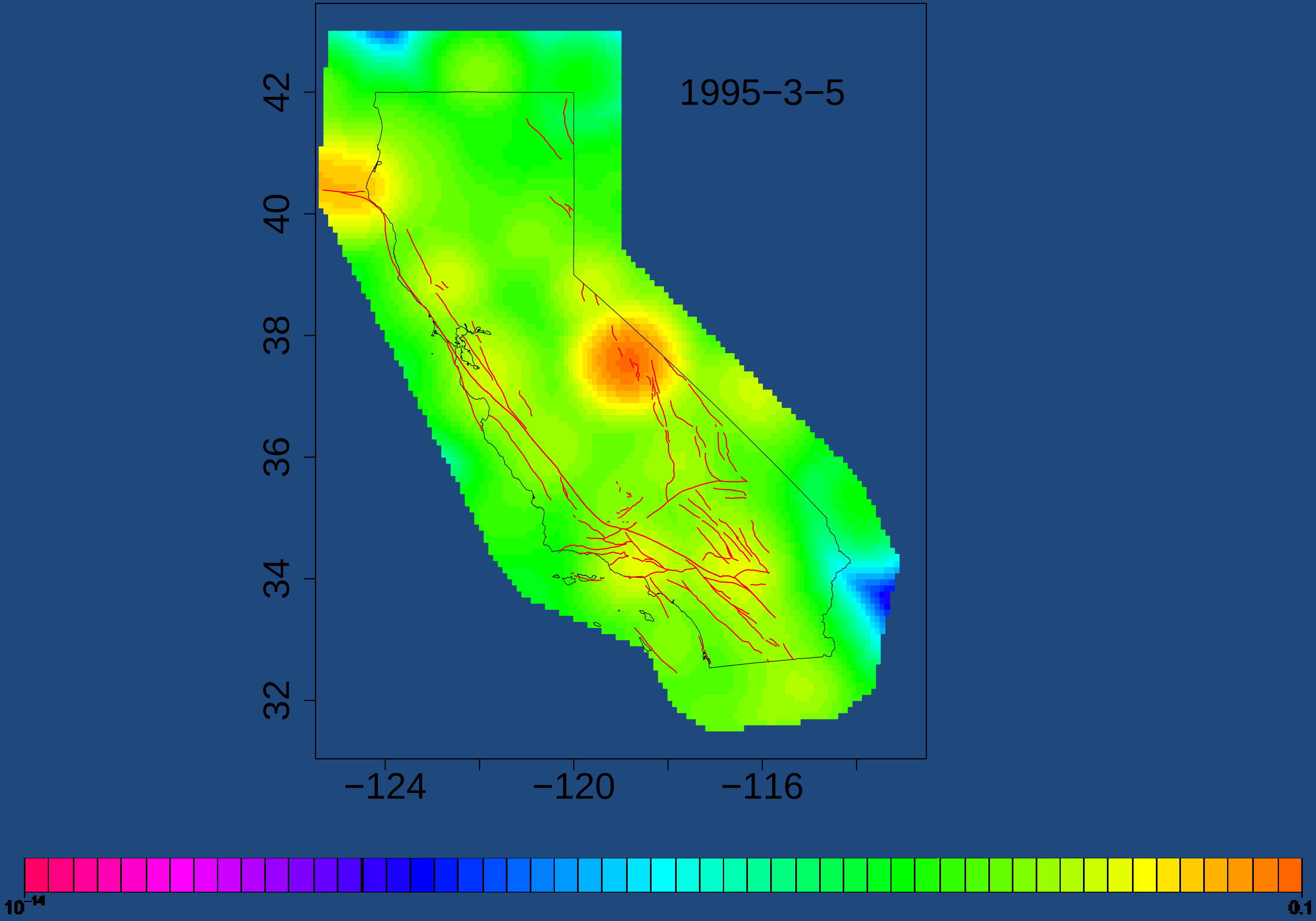
Daily forecasting for California by ETAS



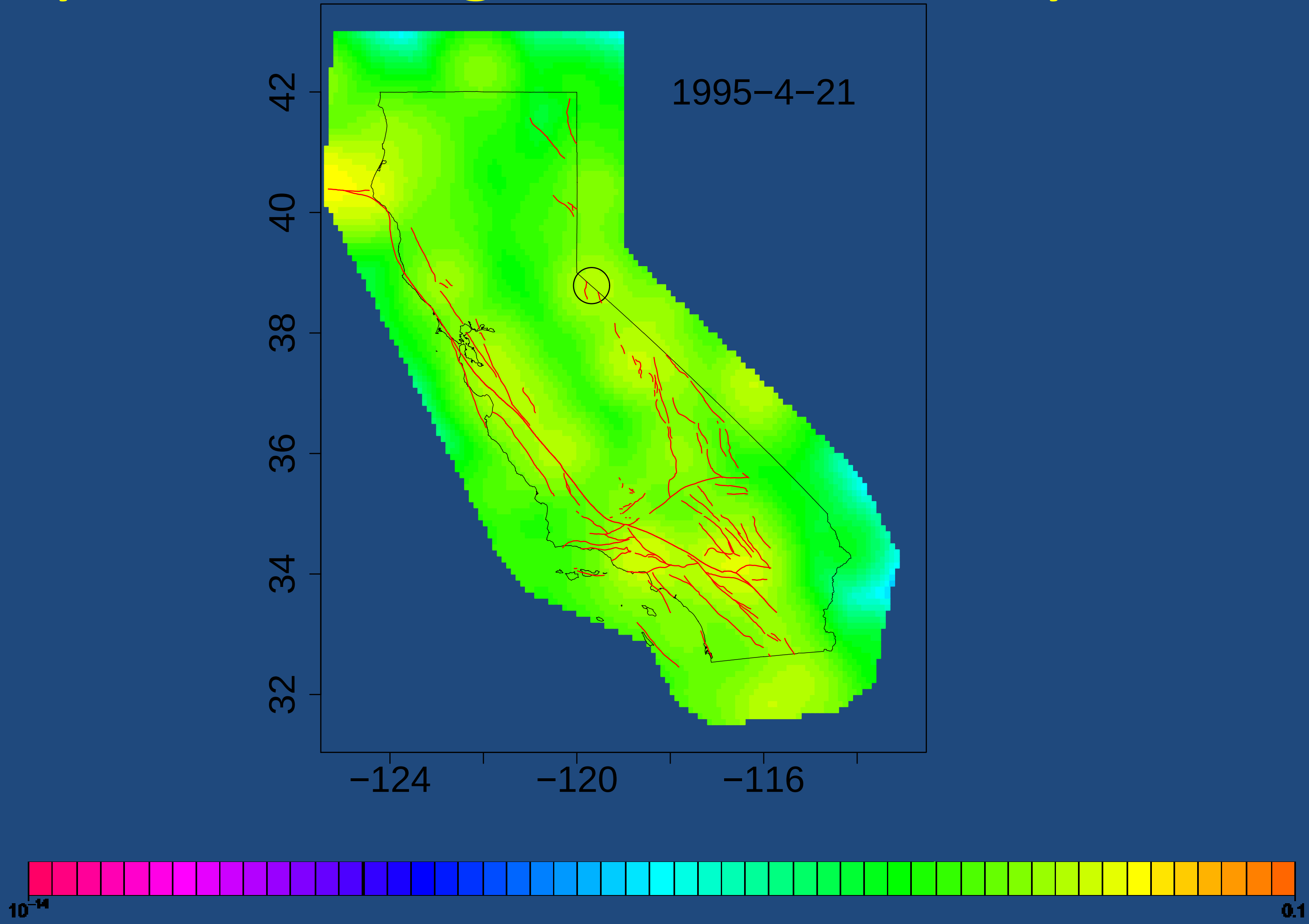
Daily forecasting for California by ETAS



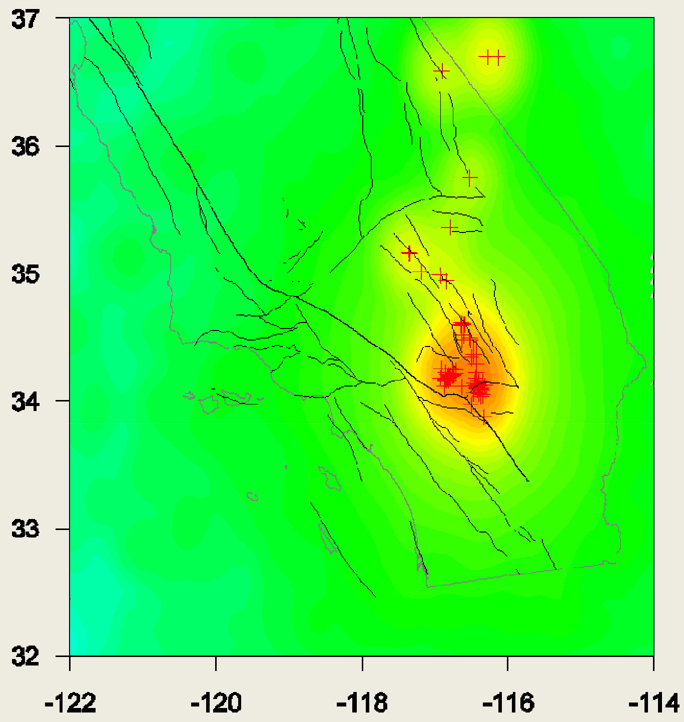
Daily forecasting for California by ETAS



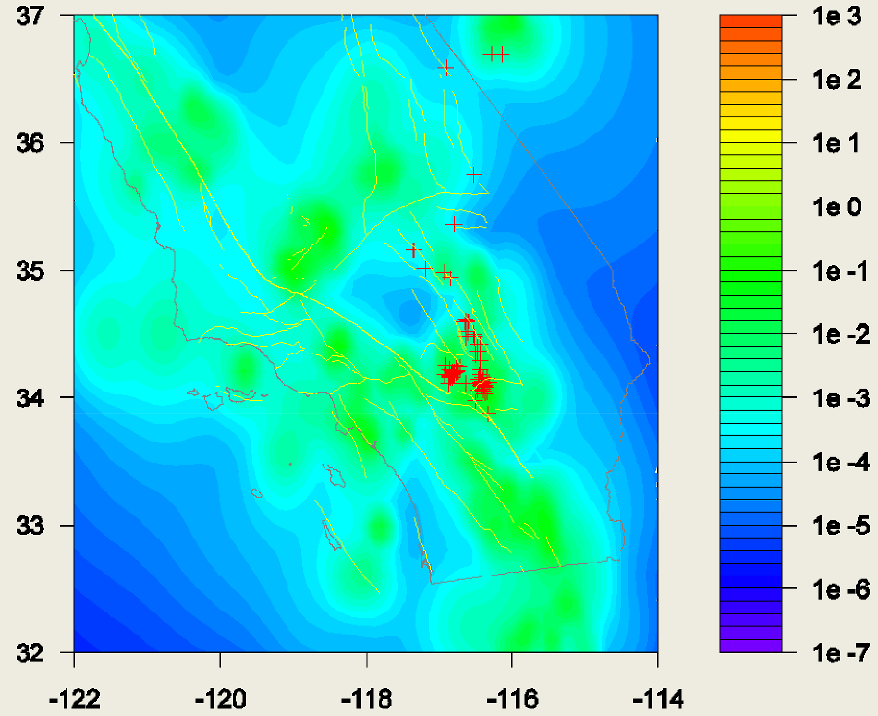
Daily forecasting for California by ETAS



Forecasting Landers aftershocks

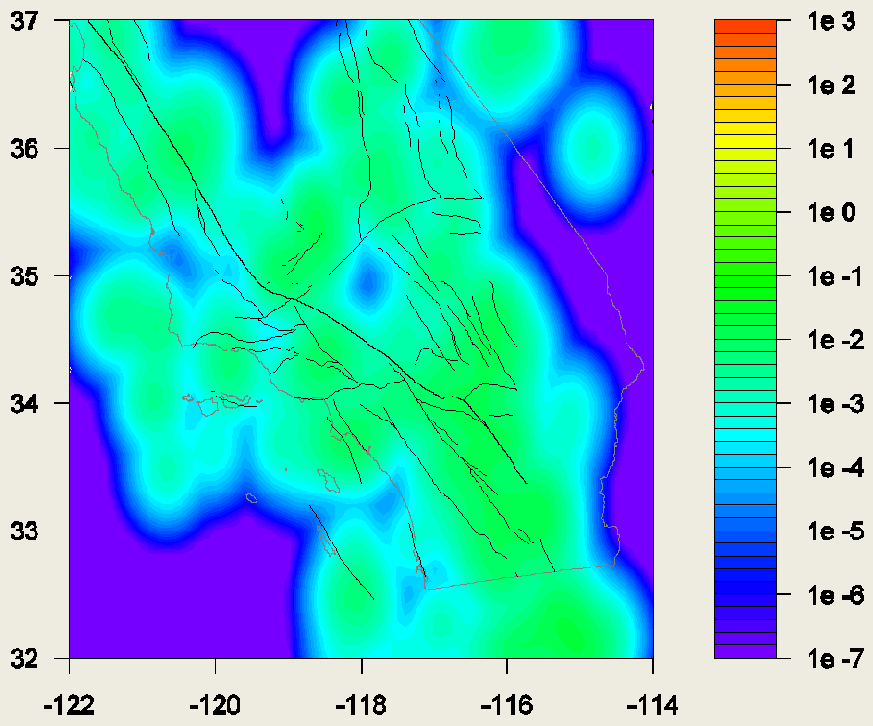


ETAS model

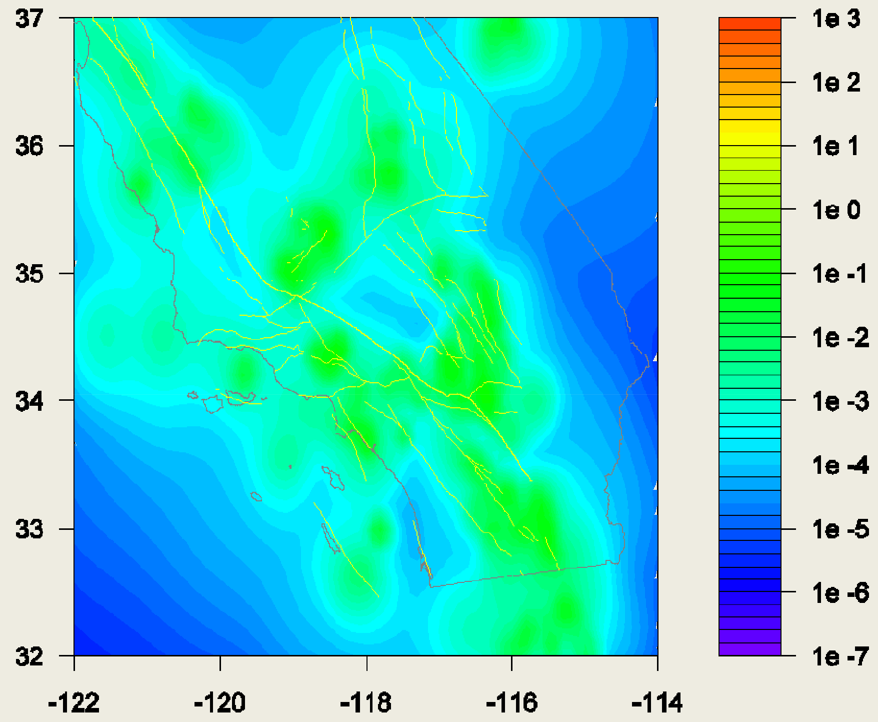


Poisson model

Forecasting for an aseismic period (2007/01/01)



ETAS model



Poisson model

Part II: Another model that needs heavy computation: Local likelihood ETAS models

Motivations

1. *Seismicity changes from places to places. It is important to make a model adapted to such changes.*
2. *Fitting the space-time ETAS model for a small region with a few events makes the estimation and forecasts unstable.*
3. *Realization of local likelihood models is simple and straightforward.*

Local likelihoods for point processes

- For each point (x,y), take the kernel function h(x,y), the local likelihood is (Zhuang, 2006)

$$\log LL(x,y) = \sum_i h(x_i - x, y_i - y) \log \lambda(t_i, x_i, y_i) - \int_0^T \iint_S h(\xi - x, \zeta - y) \lambda(u, \xi, \zeta) \, d\xi \, d\zeta \, du$$

Maximum local likelihood estimate (MLLE)

$$\hat{\boldsymbol{\theta}}(x,y) = \arg\max_{\boldsymbol{\theta}} \log LL(x,y)$$

Computation method of maximum local likelihood estimates for the ETAS model

Choice of kernel functions: 2D-version step-wise kernel function

$$h(x, y) = \sum_k W_k \, \mathbb{1}_{R_k}(x, y) \tag{4}$$

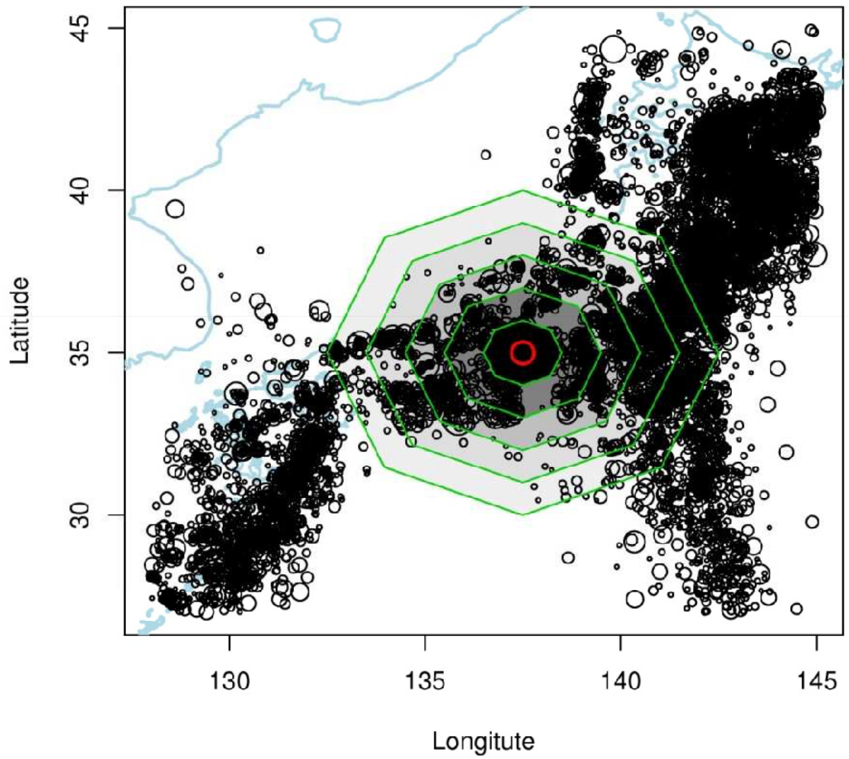
R_i: disjoint polygon rings around $(0, 0)$, $i = 1, \cdots, K$.

Likelihood

$$\log LL(x, y) = \sum_k W_k \log L(R_k + \{(x, y)\}) \tag{5}$$

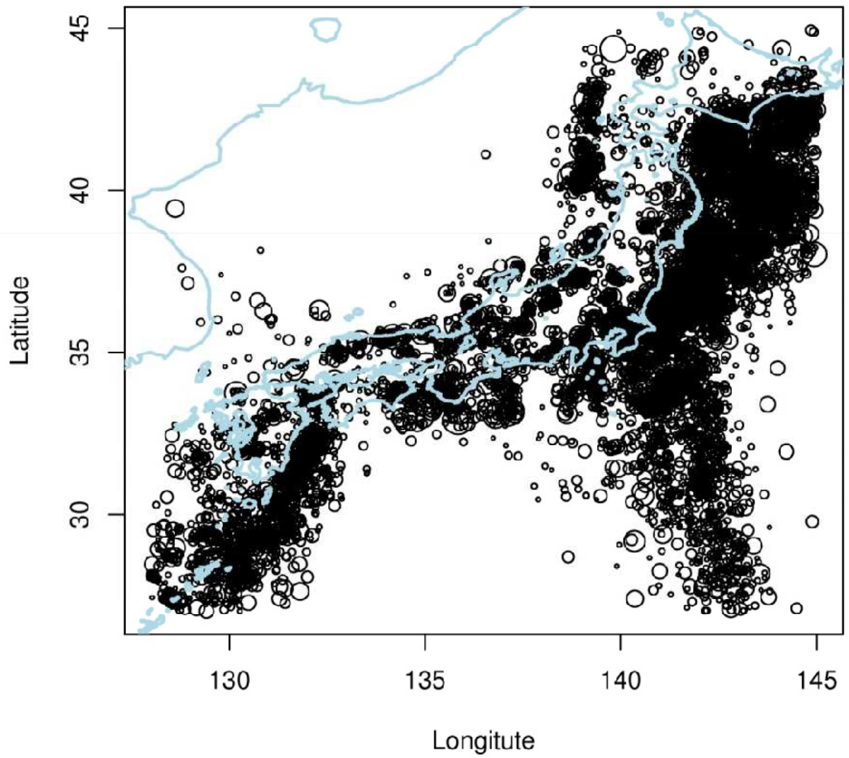
$L(S_k)$: is the likelihood for the observations on the space-time range $T \times R_k$.

Polygon-ring kernel functions

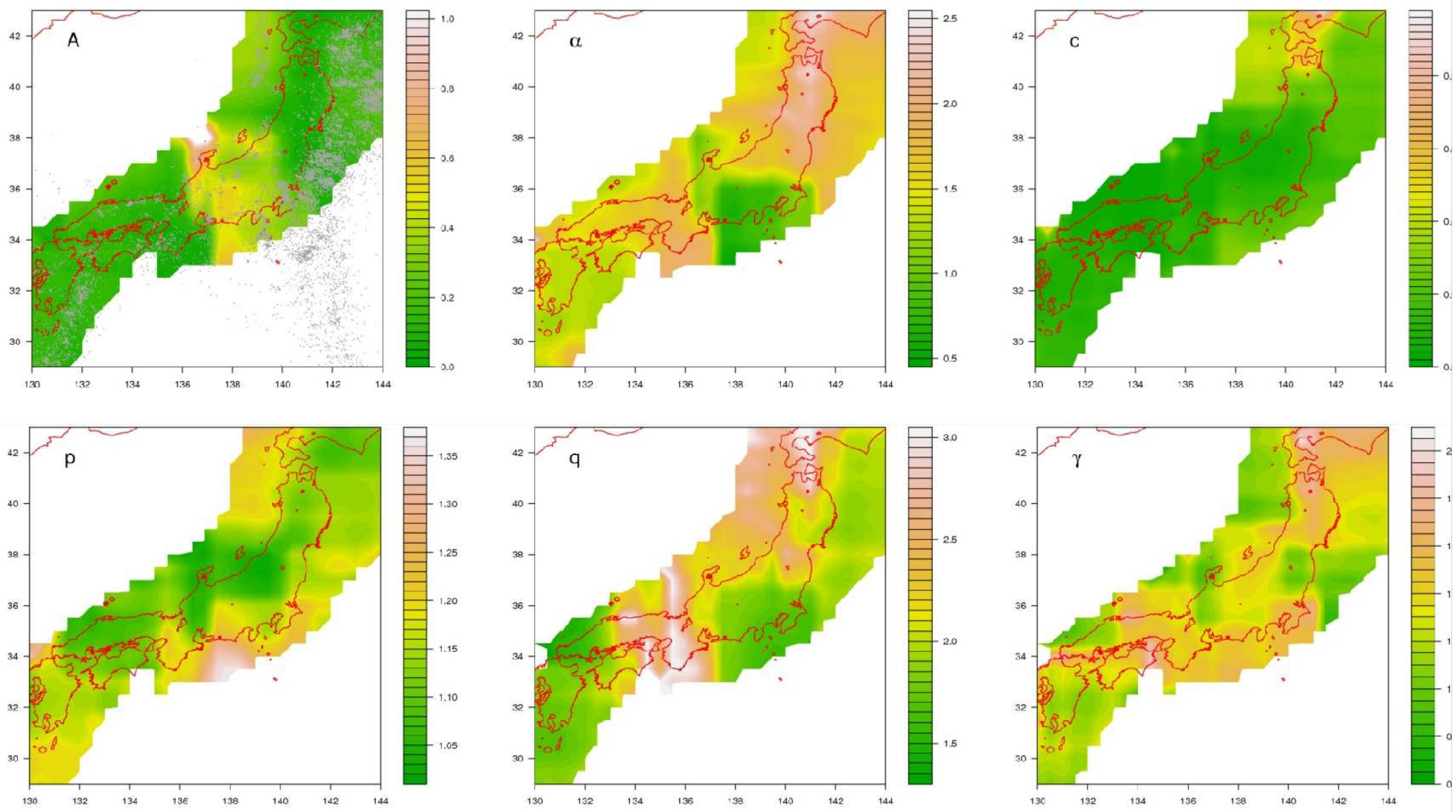


Data analysis

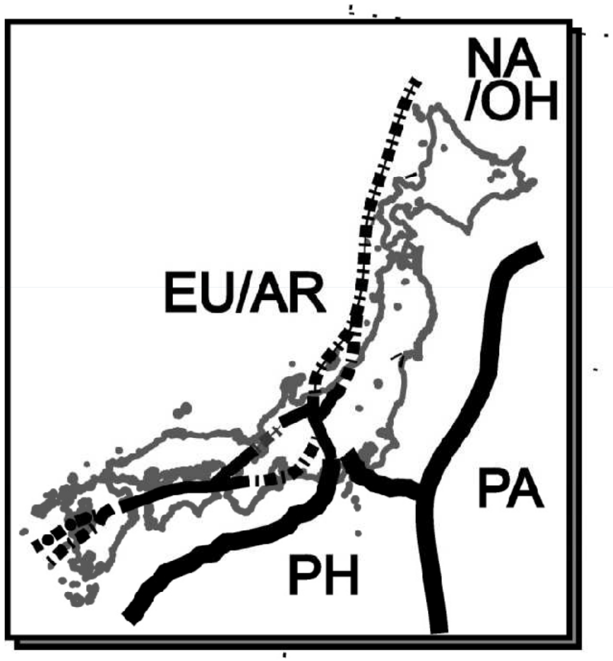
The data use for this analysis is the JMA (Japan Meteorological Agency) catalog.

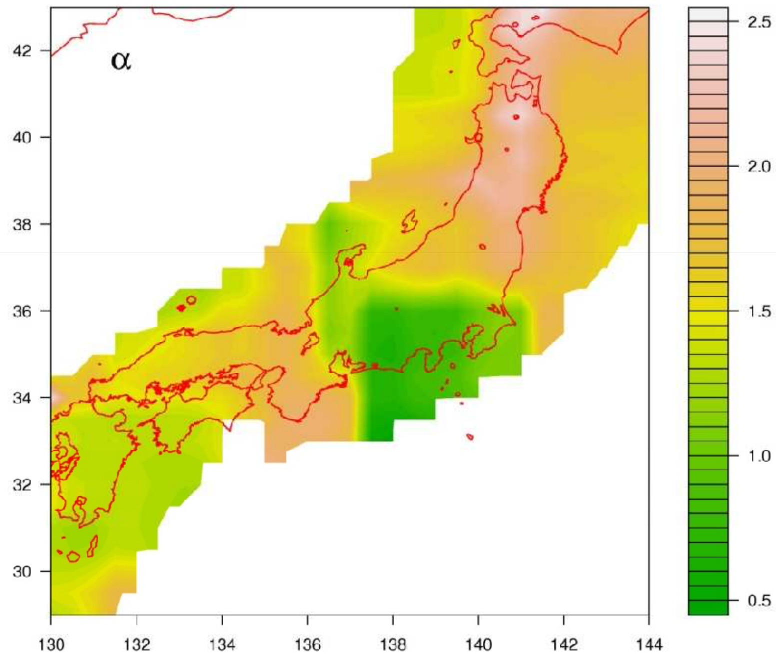


Spatial variation of parameters

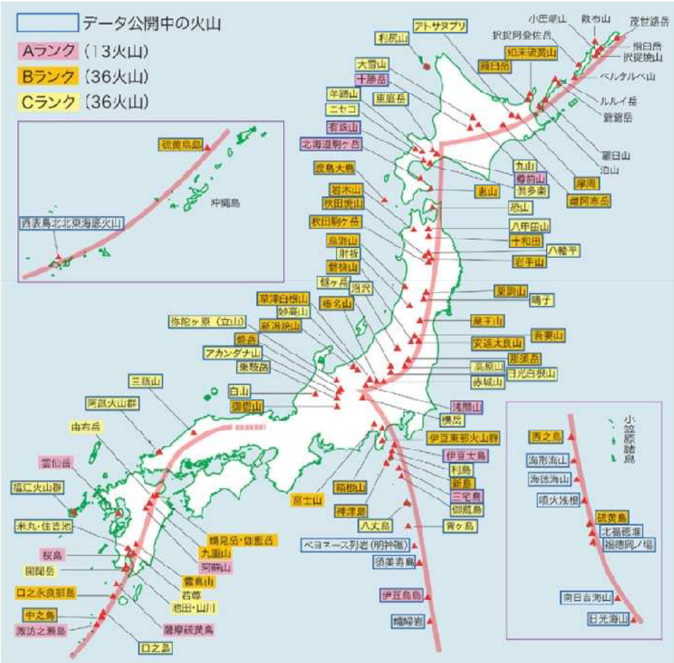


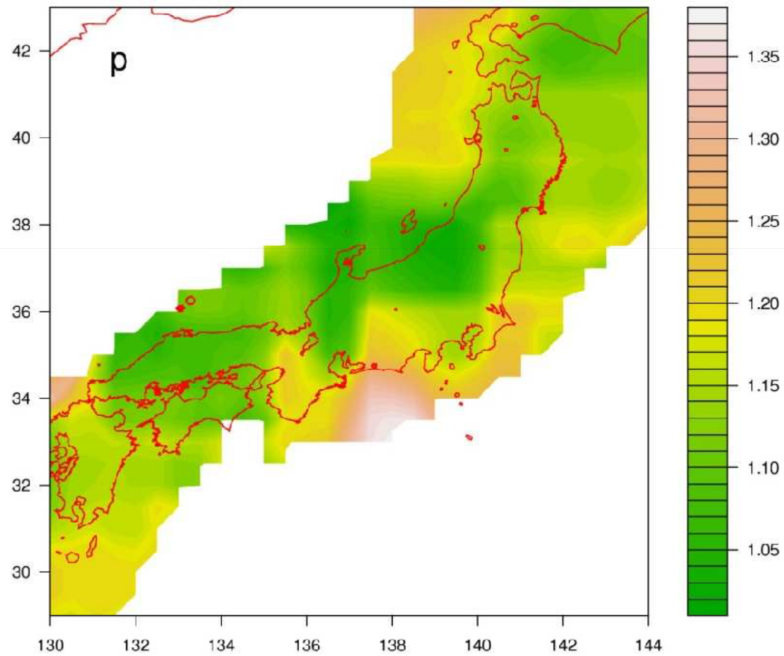
$$\kappa(m) = A\,\mathrm{e}^{\alpha(m-m_c)}$$





$$g(t) = \frac{p-1}{c}\left(1+\frac{t}{c}\right)^{-p}$$





Conclusion

1. By running online forecasting and offline optimization simultaneously, we can make quick forecasts and keep model parameters updating the observation history through heavy computation.

2. Results from fitting seiesmicity from the Japan region show that clustering characteristics of seismicity vary in space.

3. The local ETAS model can be used to handle spatial variations of model parameter by using MLLE to provide stable estimation of the ETAS parameters and forecasts for a small region with a small number of events.