Next-day Earthquake Forecasts Generated by the ETAS Model and performance evaluation

Jiancang Zhuang

Institute of Statistical Mathematics, Tokyo

Outlines

1. **ETAS model in CSEP projects**

2. ***"Offline optimization, online forecasting"* Realization of space-time ETAS model in the CSEP projects.**

3. **Output examples of the ETAS model in the CSEP project.**

Space-Time Epidemic Type Aftershock Sequence (ETAS) model

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

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

- 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

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

Space-time ETAS model

- Conditional intensity

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

- Likelihood function

$$\ln L=\sum_{(t_i,x_i,y_i)\in[0,T]\times A}\ln\lambda(t_i,x_i,y_i)-\int_0^T\iint_A\lambda(t,x,y)dtdxdy$$

Space-time ETAS model

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

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

Contribution from background seismicity

$$\frac{\mu(x,y)}{\lambda(t,x,y)}$$

Space-time ETAS model

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

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

Contribution from background seismicity

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

Pr{event j is from background}

$$\varphi_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)}$$

Estimation problems

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

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

Contribution from background seismicity

Clustering part

Estimation problems

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

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

How to estimate time independent 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)

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

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)

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

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

Kernel functions, spline, tessellation, histogram, …

How to estimate background seismicity?

How to estimate clustering parameters?

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

Estimation problems

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

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

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

→ Kernel functions, spline, tessellation, histogram, …

How to estimate background seismicity?

→ ?

How to estimate clustering parameters?

→ Maximum likelihood estimate if background seismicity μ is known

Estimation problems

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

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

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

How to estimate background seismicity?

How to estimate clustering parameters?

Kernel function, spline, tessellation, histogram, …

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

Pr{event j is from background}

$$\varphi_j = \frac{\mu(x_j, y_j)}{\lambda(t_j, x_j, y_j)}$$

Estimation problems

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

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

Kernel functions

Kernel functions with each event weighted by φ_j

$$\hat{\mu}(x,y) = \frac{1}{T}\sum_j \varphi_j h(x-x_j, y-y_j; d)$$

$$\hat{\lambda}(x,y) = \frac{1}{T}\sum_i h(x-x_i, y-y_i; d)$$

Solution—estimating parameters and background rate simultaneously

Iterative 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 (forecast) 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 p.d.f.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 generated. Return with all the events in all generations

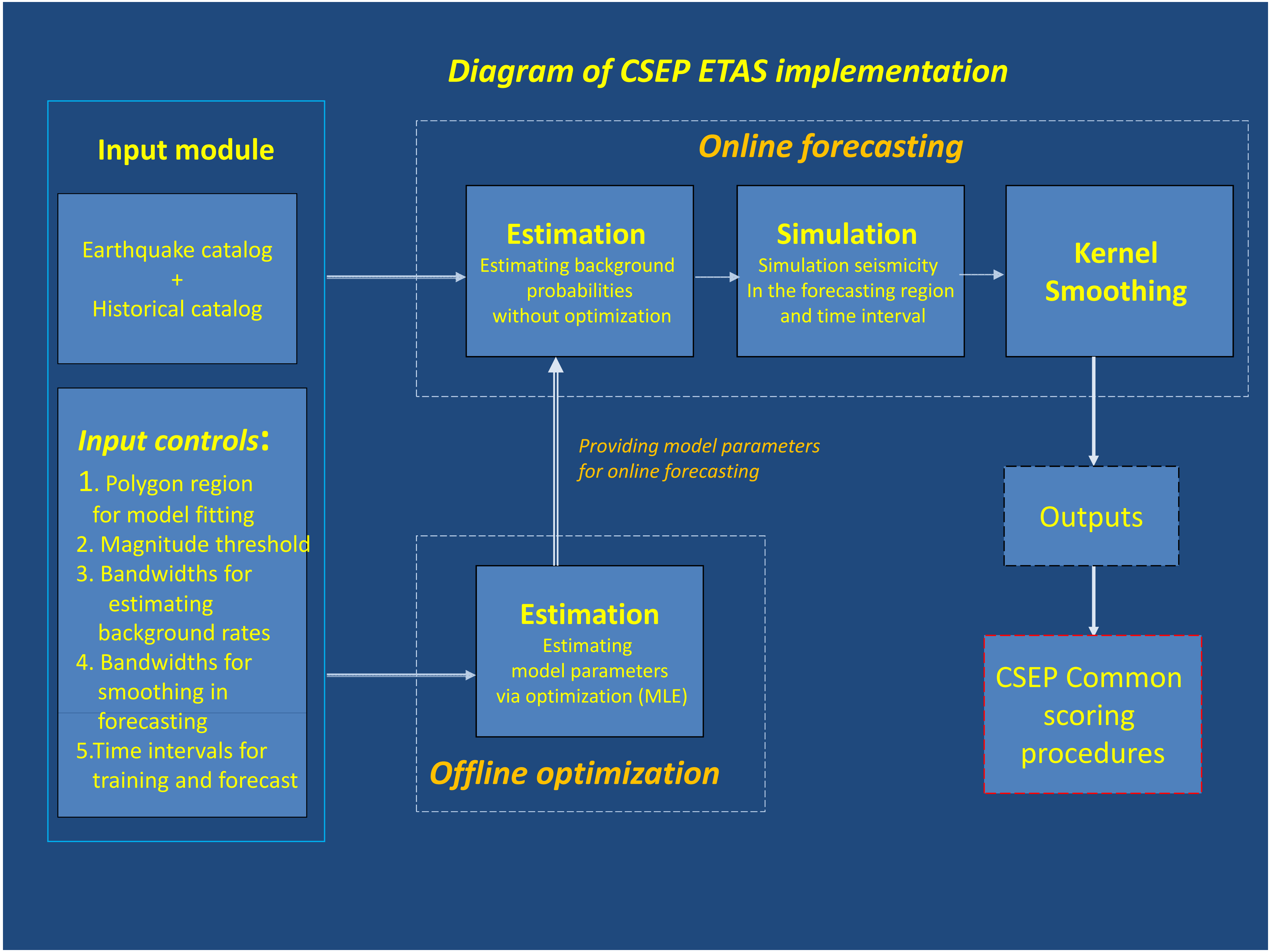
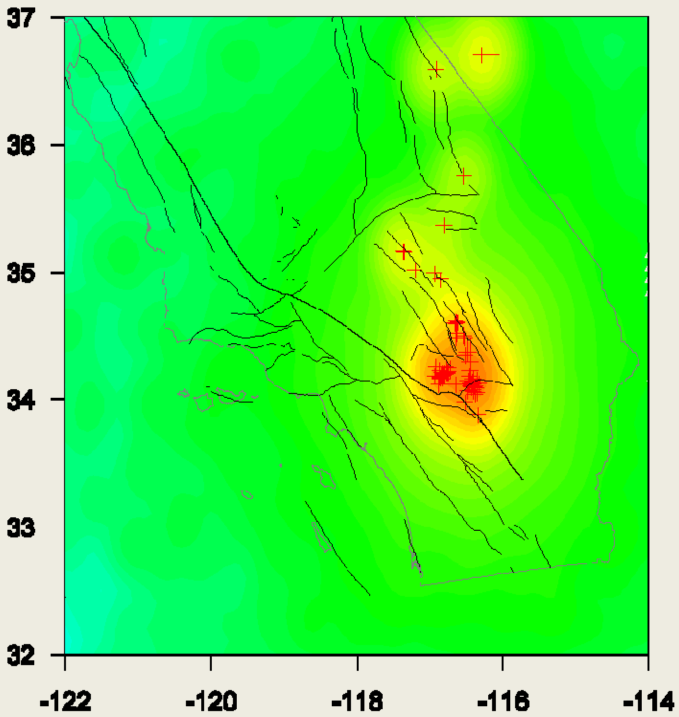
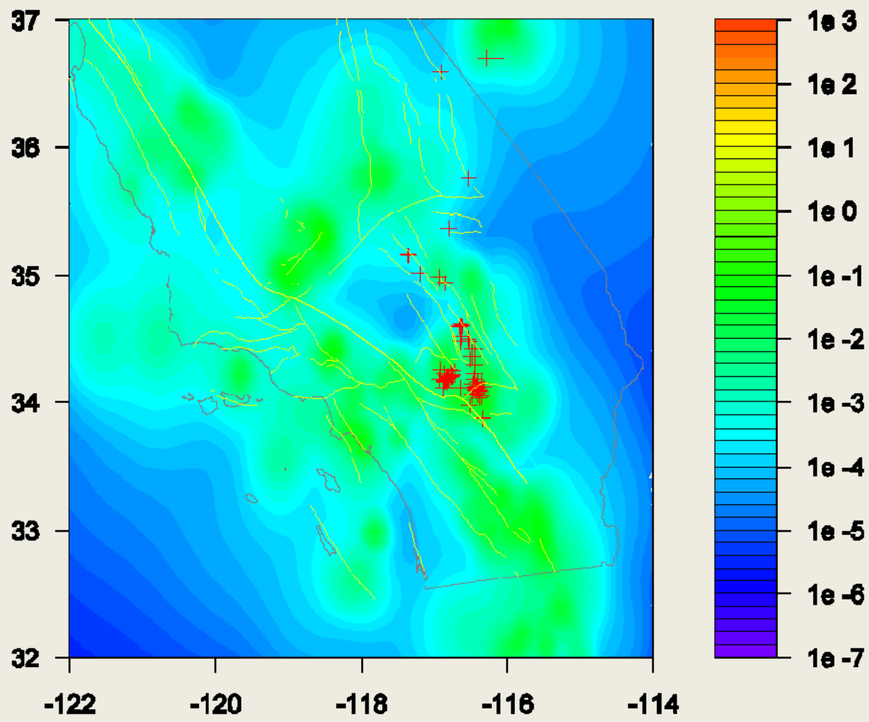


Diagram of CSEP ETAS implementation
Input module
Earthquake catalog
+
Historical catalog
Input controls:
1. Polygon region for model fitting
2. Magnitude threshold
3. Bandwidths for estimating background rates
4. Bandwidths for smoothing in forecasting
5.Time intervals for training and forecast
Online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

Forecasting Landers aftershocks



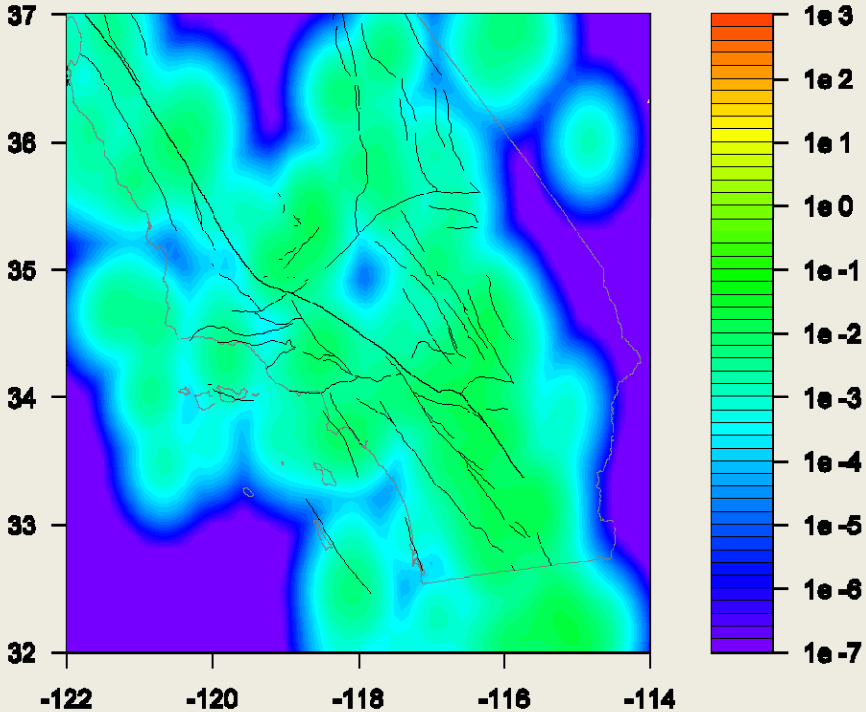
ETAS model



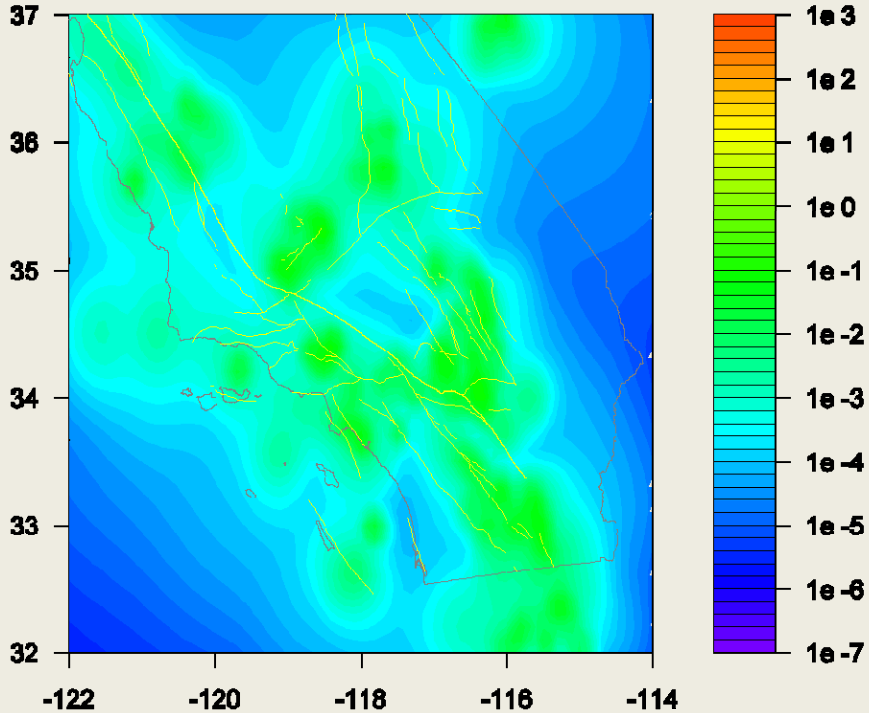
Poisson model

Unit: events/day/deg^2

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



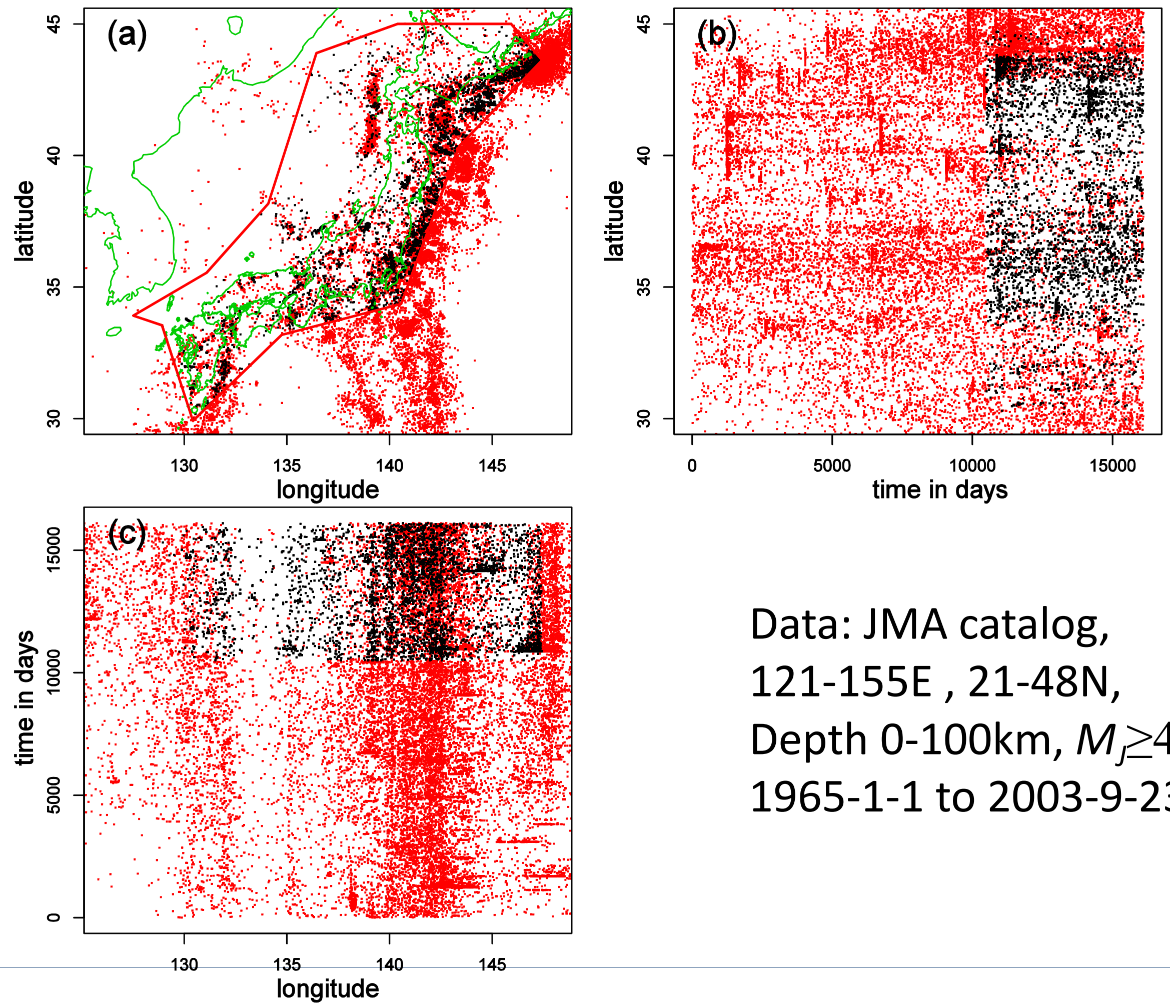
ETAS model



Poisson model

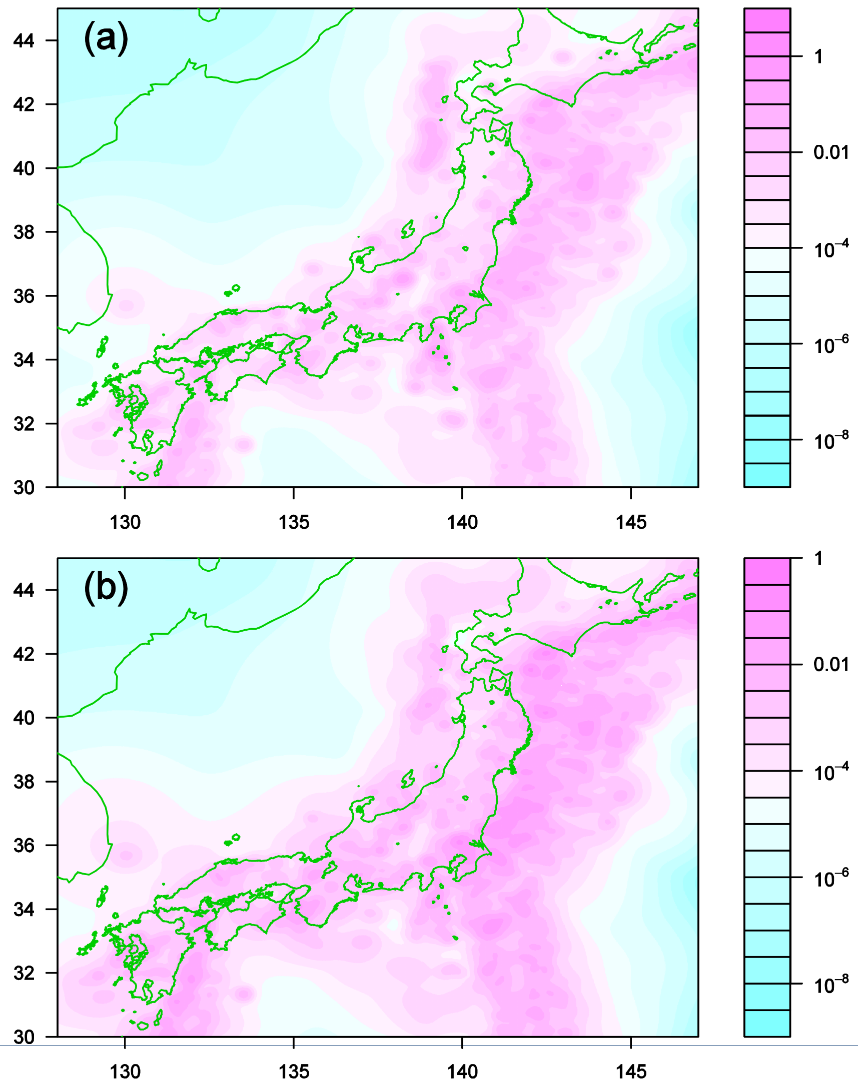
Unit: events/day/deg^2

Example 2: Retrospective forecast of the Tokachi-Oki Earthquake



Data: JMA catalog,
121-155E , 21-48N,
Depth 0-100km, $M_J \geq 4.0$,
1965-1-1 to 2003-9-23

Example 2: Retrospective forecast of the Tokachi-Oki Earthquake

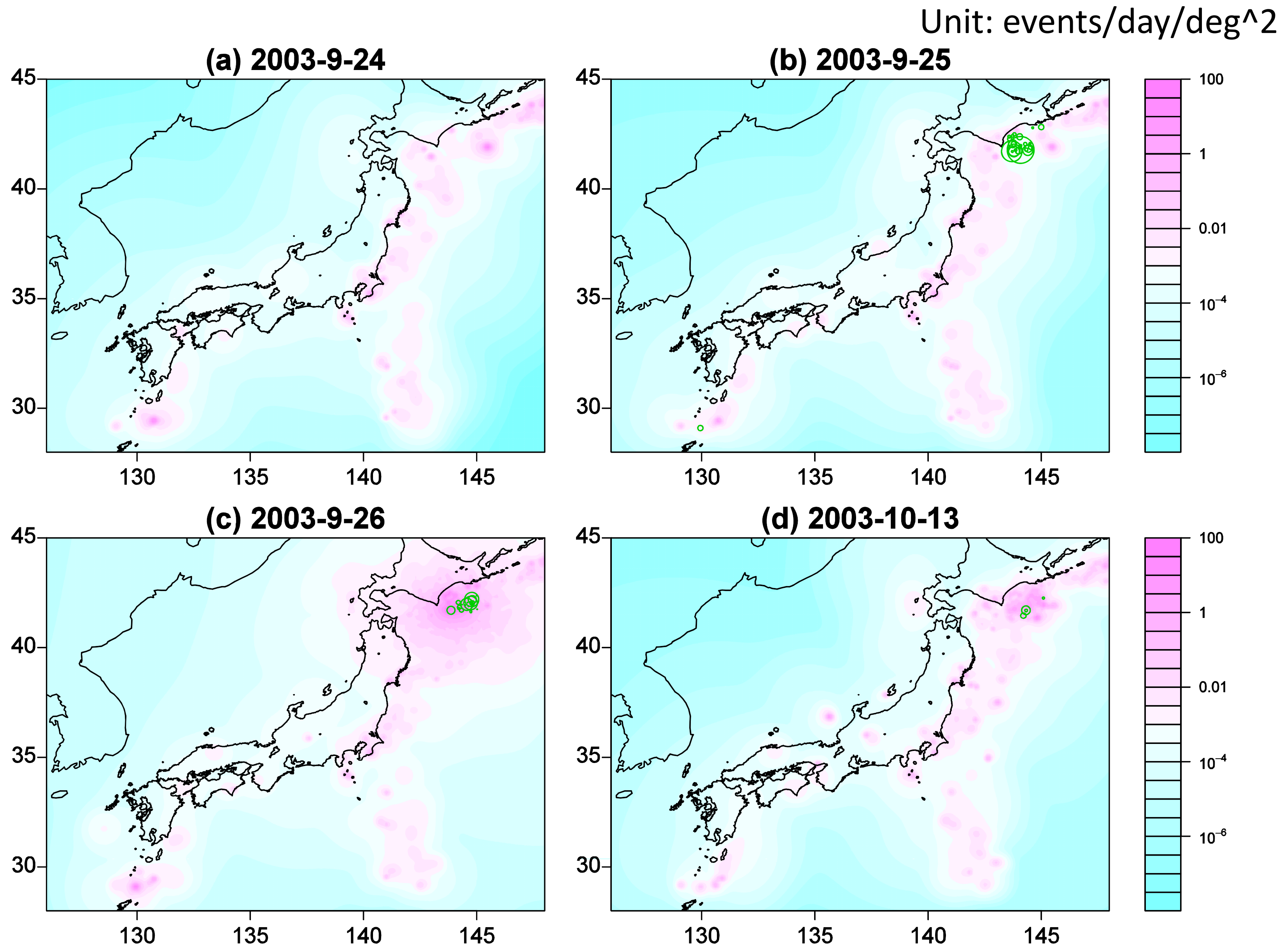


Time independent total seismicity rate

Background seismicity rate

Unit: events/day/deg^2

Example 2: Retrospective forecast of the Tokachi-Oki Earthquake



Typical scoring methods (1)

(1) Contingency-table based test for **Deterministic: Yes/No prediction** :

	Eq obs.	No eq. obse	Row sum
Predict eqs	a	b	a+b
Predict no eq	d	c	c+d
Column sum	a+d	b+c	N=a+b+c+d

Hanssen-Kuiper skill score, or R-score, namely

$$R_{HK} = \frac{a}{a+d} - \frac{b}{b+c}$$

Scoring alarming-level type of predictions

Molchan's error diagram (Molchan 1990, 1991, 1997, 2003; ROC curve):

usually for alarming-level based forecasts

a: number of successful forecasts of occurrence;

b: number of false alarms;

c: number of successful forecasts of non-occurrence;

d: number of failures to predict.

$$\nu = \frac{d}{a+d} \sim \tau = \frac{a+b}{a+b+c+d}$$

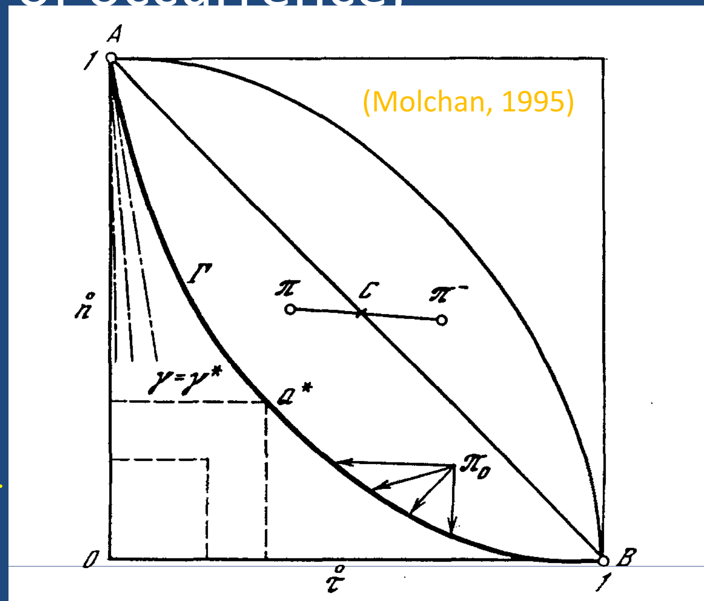


Figure 1

Gambling score (1)

Each time the forecaster make a prediction, he bets 1 point of his reputations. If he fails, he looses this point; if he wins, he should be rewarded fairly.

Question: How to fairly reward the forecaster for a success?

Gambling score (2)

Question: How to reward the forecaster for a success fairly?

Answer:

$$G = (1 - p_0) / p_0$$

p_0 : prob. given by the reference model that the prediction is correct

Return for each prediction

Earthquake occurrence	Yes	No
Forecaster predicts Yes	G_{yes}	-1
Forecaster predicts No	-1	G_{no}
Forecaster predicts Yes with prob. *p*	$G_{yes}p$-(1-p)	(1-p)G_{no}-p

scoring methods for probability forecasts

Information score (likelihood ratio, Vere-Jones 1998):

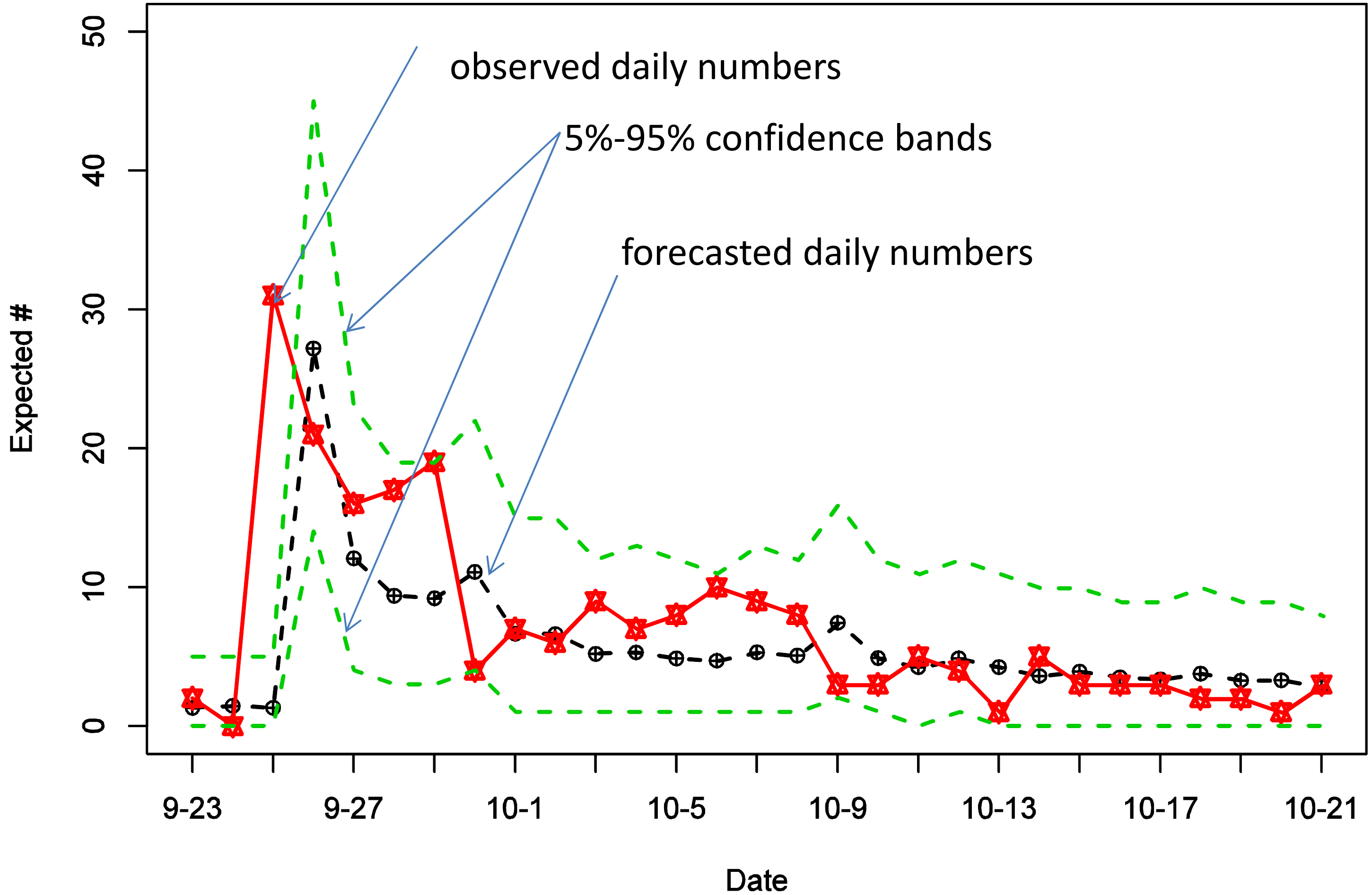
$$I=\sum_i \left[X_i \log \frac{p_i}{p_i^{(0)}} + (1-X_i) \log \frac{1-p_i}{1-p_i^{(0)}} \right]$$

p_i:prob. of earthquake occurrence given by model;

$p_i^{(0)}$: prob. given by baseline model;

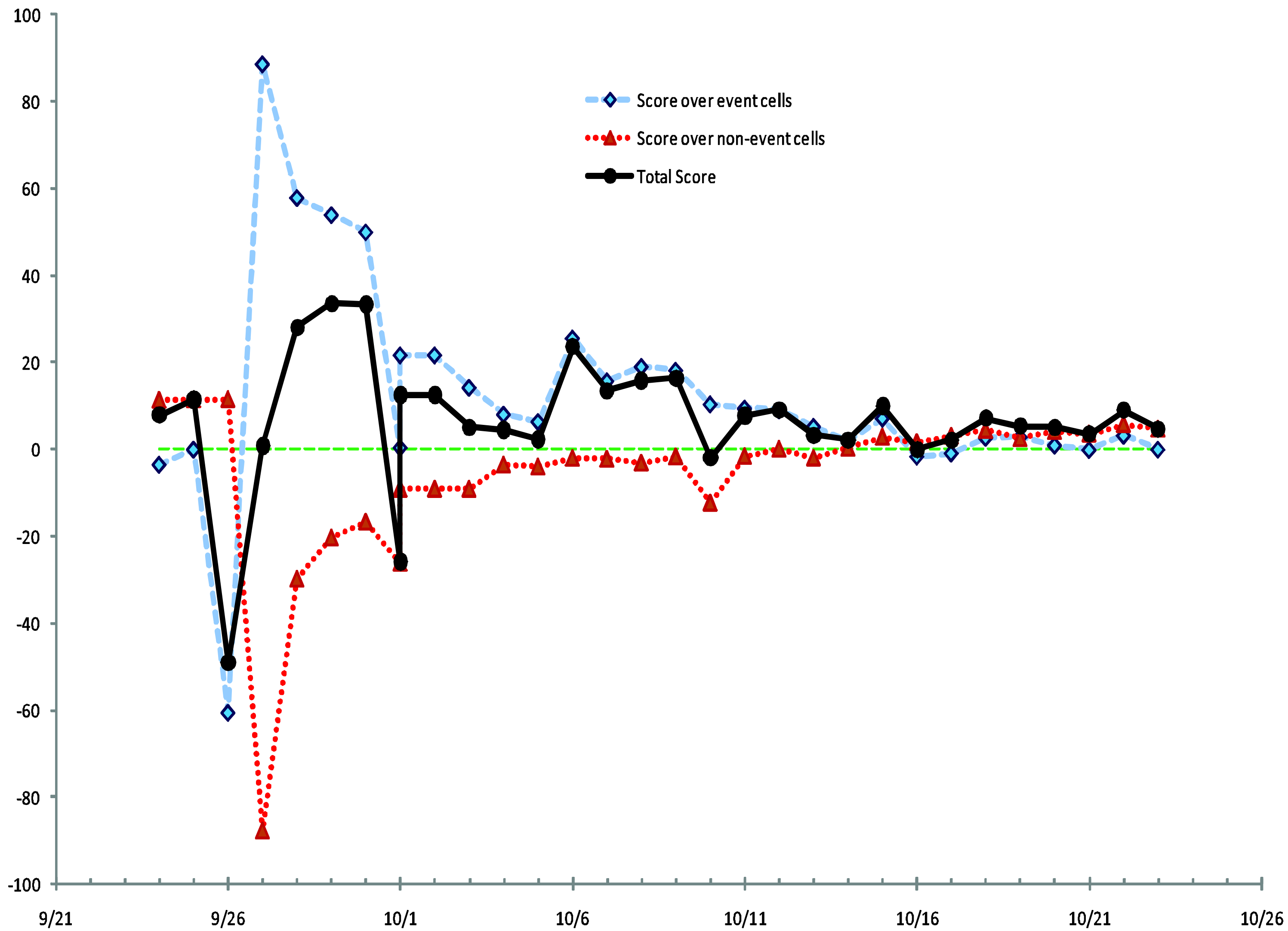
X_i: =0,if no event occurs;=1 otherwise.

Example 2: Retrospective forecast of the Tokachi-Oki Earthquake



Temporal variation of forecasted daily numbers of earthquake during 2003-09-24 to 2003-10-22.

Example 2: Retrospective forecast of the Tokachi-Oki Earthquake



Daily information gains against the Poisson model

Conclusions

- The space-time ETAS model has been implemented as an “off-line optimization and online forecasting” scheme in the Japan and SCEC CSEP projects. It consists of four components: (1) off-line optimization; (2) a simulation procedure; (3) a smoothing procedure.
- Using the ETAS model, I have made retrospective experiments on 1-day forecasts of earthquake probabilities in the Japan region before and after the Tokachi-Oki earthquake in September 2003, in the format of contour images.
- These forecasts were test against the reference model, the Poisson process which is stationary in time but spatially inhomogeneous. As expected, the forecasts based on the ETAS model catch the temporal and spatial features of the aftershock sequence, and the ETAS model performs better than the Poisson model.

This presentation is partially based on *Zhuang, 2011, EPS*

Thank you.