The M9 Tohoku–Oki Earthquake and Statistical Seismology
Yosihiko Ogata, Institute of Statistical Mathematics

1. Long-term Probability forecasting and repeating earthquakes
2. Prediction and diagnostic analysis of seismic activity
3. Discrimination of foreshocks and their operational forecasting

Long-term earthquake forecast

- The Earthquake Research Committee (ERC) of Japan opens the probabilities that earthquakes will occur in the future at main active faults and subduction-zones in Japan to the public.

Long-term evaluation of the subduction-zones

- >90%
- <7%

Long-term evaluation of the main active faults

- >90%
- 20%
- <50%
- 90%
The main focal area of major Earthquakes From Sanriku-Oki to Boso-Oki (ERC, 1999)

Miyagi-Ken-Oki Earthquake

<table>
<thead>
<tr>
<th>Date</th>
<th>Type of Earthquake</th>
<th>Magnitude</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1793.2月17日</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1835.1月20日</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1861.10.21日</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1897.2月20日</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906.11.8日</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978.6月12日</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Renewal process models

\[ \lambda(t \mid \text{History}) = \frac{f(t - t_{\text{last}})}{1 - F(t - t_{\text{last}})} \]

What is the BPT renewal process?

Brownian relaxation oscillator process

\[ S(t) = \lambda t + \sigma W(t) \]

- S(t): Stress accumulation
- W(t): Standard Brownian motion
- \( \lambda \): Loading rate
- \( \sigma \): Perturbation rate
- sf: Failure state
- s0: Ground state

Active faults

Fault 1
Fault 2
Fault 3
Fault 4
Fault n

Recurrence interval (Brownian passage time)
What is the BPT renewal process?

- **Brownian Passage Time (BPT) distribution**
  \[ f(x; \mu, \alpha) = \frac{\mu}{2\pi \alpha^2 x^3} \exp \left\{ -\frac{(x - \mu)^2}{2\mu \alpha^2 x} \right\} \]

- **Recurrence interval (Brownian passage time)**

- **Unknown parameters**
  - Mean: \( \mu = E[X_i] = \frac{s_f - s_0}{\alpha} \)
  - Coefficient of variation: \( \alpha = \sqrt{\frac{\text{Var}[X_i]}{E[X_i]}} = \frac{\sigma}{\sqrt{\lambda(s_f - s_0)}} \)

Forecasting method by ERC

- **Predictive distribution for interval time** \( X \) from the latest earthquake to the next earthquake

**Plug-in predictive density**
\[ \hat{f}(x) = f(x | \mu, \hat{\alpha}) \]

- Mean of past intervals \( X = (X_1 \cdots X_n) \)

- \( \hat{\mu} = \frac{\bar{X}}{T} = \frac{\sum_{i=1}^{n} X_i}{n} \)
- \( \hat{\alpha} = 0.24 \)

- MLE from 4 active faults where a lot of earthquakes turned out.

**Proposal**

- **Predictive distribution for interval time** \( X \) from the latest earthquake to the next earthquake

**Bayesian predictive density**
\[ \hat{f}(x) = \int \int f(x | \mu, \alpha) \pi(\mu, \alpha | T, X) \, d\mu \, d\alpha \]

**Estimation of prior distribution**

- **Fault 1**
- **Fault 2**
- **Fault 3**
- **Fault 4**
- **Fault n**

**ABIC and estimated prior distribution**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>ABIC</th>
<th>Gamma</th>
<th>Weibull</th>
<th>Exp</th>
<th>Uniform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lognormal</td>
<td>246.62</td>
<td>245.27</td>
<td>244.50</td>
<td>263.00</td>
<td>249.82</td>
</tr>
<tr>
<td>Gamma</td>
<td>245.78</td>
<td>244.44</td>
<td>243.67</td>
<td>262.66</td>
<td>249.94</td>
</tr>
<tr>
<td>Weibull</td>
<td>245.02</td>
<td>244.66</td>
<td>243.66</td>
<td>262.97</td>
<td>250.46</td>
</tr>
<tr>
<td>Exp</td>
<td>268.67</td>
<td>267.38</td>
<td>266.63</td>
<td>285.84</td>
<td>275.17</td>
</tr>
<tr>
<td>Uniform</td>
<td>251.86</td>
<td>250.48</td>
<td>249.69</td>
<td>268.92</td>
<td>257.63</td>
</tr>
</tbody>
</table>
### Stochastic model for repeating earthquakes

Brownian relaxation oscillator process (e.g. Matthews et al. 2002)

\[ dS_t = \lambda(t) dt + \sigma \ dW_t \quad (W_t \text{ standard B.M.}) \]

- **\( S_t \)**: shear stress
- **\( S_0 \)**: ground stress
- **\( \lambda(t) \)**: stress loading
- **\( \sigma dW_t \)**: stress perturbation

Interval lengths obey the BPT distribution:

\[
 f(x | \mu, \alpha) = \frac{\mu}{2\pi \alpha^2 x^3} \exp \left\{ -\frac{(x - \mu)^2}{2\mu \alpha^2 x} \right\} \\
 \text{mean: } \mu = (s_f - s_0) / \lambda \\
 \text{variance: } \alpha = \sigma / \sqrt{\lambda(s_f - s_0)}
\]

### Space-time model

- Non-stationary loading rates and dispersions

\[ dS_t = \lambda(x, y, t) \ dt + \sigma(x, y, t) \ dW_t \]

- We cannot assume the BPT for the interval distribution.

---

**Maximum Posterior estimates**

- **Mean interval**
  - \( \mu = E[X] = \frac{s_f - s_0}{\lambda} \)
  - \( \tilde{\mu} \)

- **Coefficient of variation**
  - \( \alpha = \frac{\sigma}{\sqrt{\lambda(s_f - s_0)}} \)
  - \( \tilde{\alpha} \)

**Renewal model (log Normal) for 163 repeating earthquake sequences**

Estimation 1993 - 2009

**Forecasted for 2010**

Result for 2010

Okada, Uchida, Aoki (2011)

S. Nomura (SOKENDAI D3)
Assume the following model

\[ dS_i = \lambda_0 \lambda(x, y, t) dt + \sigma_0 \sqrt{\lambda(x, y, t)} \, dW_i \]

- \( \lambda_0 \): Background loading rate
- \( \sigma_0 \): Background perturbation rate
- \( \lambda(x, y, t) \): space-time rates relative to the background loading rate \( \lambda_0 \)

Then we have the BPT intervals for each \((x, y)\) by time transformation \( t' = \int \lambda(x, y, t) \, dt \)

- \( \lambda(x, y, t) \) is regarded as changes of stress loading velocity, and parameterized by a 3D spline function with smoothness constraints.
- Objective Bayesian procedure (ABIC) is applied to attain the optimal constraints.

**Parkfield data (Nadeau)**

Micro-earthquakes (HRSN, Magnitude range \(-0.5 \sim 2\))

- period 1987/1/1 \sim 2004/9/28 (preceding 2004 Parkfield) with missing data during 1998/7/1 \sim 2001/7/26

Result

Loading rate change at each sub-region

Spatial change in time
Result on the dispersion coeff. $\alpha$
Dispersion coefficient $\alpha$ (proportional to the perturbation rate $\sigma$) is affected by the nearby earthquakes.

Prediction and diagnostic analysis of seismic activity

Geographical Survey Institute
Deformation of Japanese islands

Coseismic dislocations of Tohoku-Oki earthquake of M9
**Sudden Stress Decrease**

**FAILURE**

**STRESS**

**Quiescence**

**Secular stress increase**

**Sudden Stress Increase**

**Coulomb’s Stress Failure Stress**

\[ \Delta CFS = \Delta(\text{shear stress}) - (\text{friction coeff.}) \times \Delta(\text{normal stress}) \]

---

**ETAS model:** (Epidemic Type Aftershock Sequence model)

\[ \lambda(t) = \mu + \sum_{j: t_j < t} e^{\alpha(M_j - M_c)} v(t - t_j) \]

where

- \( t_j \) is occurrence time of \( j \)th event;
- \( M_j \) is magnitude of \( j \)th event;

\[ v(t) = \frac{K}{(t + c)^p} \]

Omori-Utsu formula for aftershock decay


Observed time interval: \([S, T]\)

Occurrence time data: \( t_1, t_2, \ldots, t_n \)

**Log-likelihood** function:

\[ \ln L(\theta; S, T) = \sum_{S < t_i < T} \ln \lambda(\theta(t_i)) - \int_S^T \lambda(\theta(t)) dt \]

parameters are (\( \mu, K, c, \alpha, p \)).

Maximize the function \( \Rightarrow \hat{\theta} = (\hat{\mu}, \hat{K}, \hat{c}, \hat{\alpha}, \hat{p}) \) M.L.E.

Computer codes available. Please visit WEB home page of our project team:

Theoretical cumulative number of the events:

$$\Lambda(t) = \int_0^t \lambda(s) \, ds$$

$$\tau_i = \Lambda(t_i)$$

Triggered activities

See the following slides

Northern Nagano Pref. Sakae Vil.

福島県いわき市付近の誘発地震活動

M3 ≥ M3.0 ≥ M3.5

Stress Rate Decrease

Failure

Stress

Secular stress increase

Quiescence

M ≥ 4.0

M ≥ 3.5

M ≥ 3.0
Foreshocks led by the M7.3 event

Foreshocks led by the M7.3 event

Aftershocks

Space vs transformed-time plot of the foreshock activity till M9 mega event
Aftershocks of Tohoku-Oki Earthquake of M9 (cont.)

Theoretical cumulative number of the events: \( \Lambda(t) = \int_{0}^{t} \lambda(s) \, ds \)

2010 Chile Earthquake (Mw8.8) Aftershocks

2004 Sumatra Earthquake (Mw9.0) Aftershocks
New Zealand Darfield Earthquake Aftershocks and Christchurch Earthquake

Extensive Quiescence before the M9 earthquake

Space-Time ETAS model (Ogata, 1998, AISM)

$$\lambda(t, x, y) = \mu + \sum_{j: t_j < t} \frac{K}{(t - t_j + c)^{\alpha}} \left\{ \frac{Q_j(x, y)}{e^{sM_j}} + d \right\}$$

Heterogeneity in Space (Ogata, 2004, JGR and 2011, EPS)

$$\mu \Rightarrow \mu(x, y);$$
$$K \Rightarrow K(x_j, y_j);$$
$$\alpha \Rightarrow \alpha(x_j, y_j);$$
$$p \Rightarrow p(x_j, y_j);$$
$$q \Rightarrow q(x_j, y_j)$$

Delaunay-based function: an illustration
Location Dependent Space-Time ETAS model and occurrence data \{\{(x_i, y_i, M_i)\}_{i=1, \ldots, n}\} in \([0, T] \times A\) are given. Then Log Likelihood is

$$
\log L(\theta) = \log L(p_1, p_2, \ldots, p_n) \equiv \sum_{i=1}^{n} \log \lambda_i(x_i, y_i) - \int \int \lambda(x, y) dx dy,
$$

where \(\theta = (\omega_1, \omega_2, \ldots, \omega_n)\).

Penalized Log Likelihood

$$
Q(\theta; w_p, w_k, \omega_1, \omega_2, \ldots, \omega_n) = \log L(\theta) - \text{penalty}(\theta; w_p, w_k, \omega_1, \omega_2, \ldots, \omega_n)
$$

where the penalty is

$$
\int \int dxdy \left[ w_1(u + \tilde{u})^2 + w_2(K_1 + K_2)^2 + w_3(u^2 + \tilde{u}^2 + \omega_1^2 + \omega_2^2 + \omega_3^2) + w_4(u^2 + \tilde{u}^2) + w_5(u^2 + \tilde{u}^2) \right]
$$

Flatness constraints

Choose \(\rho\) that maximize \(\Lambda(\rho)\)
or minimize

$$
ABIC = \text{log} \max \Lambda(\rho) + 2 \times \text{dim}(\rho)
$$

Akaike Bayesian information criterion (Akaike, 1980)

Hierarchical Space-time ETAS model (HIST-ETAS)

$$
\lambda(t, x, y | H_j) = \mu(x, y) \sum_{(i, j) \in \mathcal{C}} \frac{K(x, y)}{(t - t_j + \epsilon)^{\alpha}} \left( \frac{S(x, y) S(x - x_j, y - y_j) e^{-\alpha(x - x_j, y - y_j)^2}}{\lambda(t_i, x_j, y_j | H_{t_i})} \right)^{\alpha(x, y)}
$$

Background rate

$$
\mu(x, y) = \sum_{(i, j) \in \mathcal{C}} \frac{K_0(x, y)}{(t - t_j + \epsilon)^{\alpha}} \left( \frac{S(x, y) S(x - x_j, y - y_j) e^{-\alpha(x - x_j, y - y_j)^2}}{\lambda(t_i, x_j, y_j | H_{t_i})} \right)^{\alpha(x, y)}
$$

Stochastic declustering

(Zhuang, Ogata & Vere-Jones., 2004 JASA and 2005, JGR)

Accept earthquake \(i\) as a background event with probability:

$$
\frac{\mu(x_i, y_i)}{\lambda(t_i, x_i, y_i | H_{t_i})}
$$
M ≥ 5.0

Original data

De-clustered data

T. Kumazawa (SOKEN-DAI D3)

1926 ~ 2011 Mar. M≥5.0

Transformed time by ETAS model
Okada, Uchida, Aoki (2011)
Renewal model (log Normal) for 163 repeating earthquake sequences
Estimation 1993 - 2009

Forecasted for 2010
Result for 2010

Operational Probability
Forecasting of Foreshocks and Evaluations:
15 years periods till the Tohoku-Oki Event
Earthquakes \( (M \geq 4.0) \) in JMA catalog during 1994 ~ Mar. 2011 are connected by the distance criterion

\[
d_{ST} = \sqrt{\Delta^2_{space} + (c\Delta_{time})^2} \leq 0.3^\circ \text{ (or 33.33km)}
\]

where \( c \) is the constant so as to hold 1 day = 1 km

(1) How do we recognize that it is initial earthquake of a cluster?

(2) What is definition of foreshocks?

Cluster types

<table>
<thead>
<tr>
<th>Cluster types</th>
<th>Foreshocks</th>
<th>Swarms</th>
<th>M.A.</th>
<th>All clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>member #</td>
<td>#c</td>
<td>ratio(%)</td>
<td>s.e.(%)</td>
<td>#c</td>
</tr>
<tr>
<td>1</td>
<td>467</td>
<td>3.7 ±0.2</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>2</td>
<td>125</td>
<td>6.5 ±0.6</td>
<td>584</td>
<td>30.5 ±1.1</td>
</tr>
<tr>
<td>3</td>
<td>57</td>
<td>8.0 ±1</td>
<td>271</td>
<td>37.9 ±1.8</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
<td>8.7 ±1.5</td>
<td>153</td>
<td>40.5 ±2.5</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
<td>7.4 ±1.7</td>
<td>93</td>
<td>38.4 ±3.1</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>6.8 ±1.9</td>
<td>63</td>
<td>35.6 ±3.6</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>7.9 ±2.4</td>
<td>44</td>
<td>34.6 ±4.2</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>7.8 ±2.6</td>
<td>31</td>
<td>30.1 ±4.5</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>9.2 ±3.1</td>
<td>29</td>
<td>33.3 ±5.1</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>10.3 ±3.4</td>
<td>25</td>
<td>32.1 ±5.3</td>
</tr>
</tbody>
</table>

Initial earthquakes of clusters or Isolated earthquakes

<table>
<thead>
<tr>
<th>Forecasted results for 1994 – Mar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecast</td>
</tr>
<tr>
<td>Foreshocks</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>All types</td>
</tr>
<tr>
<td>Ratio(%)</td>
</tr>
</tbody>
</table>

Diff. entropy = -22.7
Diff. AIC = -40.0 (cross-table)

Measuring inter-events concentrations in a cluster and magnitude increments

\[
(t_i^c, x_i^c, y_i^c, M_i^c) : \text{Hypocenters in a cluster } \mathcal{C} \in \mathcal{C}
\]

1. Origin-time differences
   \[ t_{ij}^c = t_i^c - t_j^c \]
   for any pairs \( i \neq j \) such that \( i < j \)

2. Epicenter operations
   \[ r_{ij}^c = \sqrt{(x_i^c - x_j^c)^2 + (y_i^c - y_j^c)^2} \]

3. Magnitude differences
   \[ M_{ij}^c = M_i^c - M_j^c \quad (i < j) \]
Time difference, Distance & Magnitude difference Normalization

\[(t, r, g) \rightarrow (\tau, \rho, \gamma) \text{ in } [0,1]^3 \]

Time Interval Transformation
\[\tau = \begin{cases} 
0 & \text{for } 0 \leq t < 0.01 \\
\log(100t)/\log(3000) & \text{for } 0.01 \leq t \leq 0.30 \\
1 & \text{for } 0.30 \leq t
\end{cases} \]

Epicenter Separation Transformation
\[\rho = 1 - \exp\{-\min(r, 50)/20\}\]

Magnitude Difference Transformation
\[\gamma = \begin{cases} 
(2/3)\exp\{g/\sigma_g\} & \text{for } g \leq 0 \\
(2/3) + (1/3)[1 - \exp\{-g/\sigma_g\}] & \text{for } g > 0
\end{cases} \]

\[\text{ただし } \sigma_g = 6709, \sigma_2 = 0.4456\]

Algorithm of foreshock probability calculations in case of plural earthquakes in a cluster

For plural earthquakes in a cluster, time differences \(t_{ij}\) (days), epicenter separation \(r_{ij}\) (km), magnitude difference \(g_{ij}\) are transformed into the unit cube

\[(t_{ij}, r_{ij}, g_{ij}) \rightarrow (\tau_{ij}, \rho_{ij}, \gamma_{ij}) \in [0,1]^3\]

Probability \(p_i\) is calculated sequentially

\[
\logit(p_i) = \logit(\mu(x_i, y_i)) + \frac{1}{\#(i < j)} \sum_{k \neq i} a_k \rho_{ij}^k + \sum_{k \neq i} b_k \gamma_{ij}^k + \sum_{k \neq i} c_k \tau_{ij}^k
\]

Here \(\mu(x, y)\) indicates probability of initial earthquake at location \((x, y)\), and the 2nd term calculates arithmetic mean of polynomials of the normalised space-time magnitude variables for all pairs of earthquakes \((i < j)\) in a cluster, where the coefficients \(a_k, b_k, c_k, d_k\) are as follows.

\[
\begin{align*}
\{a_1, a_2, a_3, a_4\} &= \{1\logit(\cdot), \logit(\cdot, \cdot)\} \\
b_k &= \{k\} \\
c_k &= \{k\} \\
d_k &= \{k\}
\end{align*}
\]

<p>| | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>8.018</td>
<td>-33.25</td>
<td>1.40</td>
<td>10.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a2</td>
<td>62.77</td>
<td>2.81</td>
<td>295.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a3</td>
<td>-37.66</td>
<td>-2.19</td>
<td>-1161.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Probability forecast (%)

<table>
<thead>
<tr>
<th>Earthquake number in a cluster</th>
<th>Mc</th>
<th>Mc</th>
<th>M7.3 Foreshock of 9 Mar 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.0</td>
<td>4.0</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>0.1%</td>
<td>0.1%</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td>0.01%</td>
<td>0.01%</td>
<td>0.01%</td>
</tr>
<tr>
<td></td>
<td>1.0%</td>
<td>1.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Earthquake number in a cluster

\(2 \text{Mc} \geq 4.0\)  \(M_{\text{max}} \geq 6.5\)

Earthquake number in a cluster

\(2 \text{Mc} \geq 4.0\)  \(M_{\text{max}} \geq 6.5\)

Others freq

Foreshocks freq

Forecast & performance
1994-2011

\(2^\ast \text{Entropy0} = 523.96; 2^\ast \text{Entropy} = 460.29\)

\(2^\ast \text{Entropy} = -63.88\)

\(2^\ast \text{Entropy} = -43.68\)
Thank you very much for your attention.

Software available for the ETAS fitting, diagnostic analysis and its manual. Please visit WEB home page of our project team:
http://www.ism.ac.jp/~ogata/Ssg/ssgE.html

Very soon, computer programs are available for the Hierarchical Space-Time ETAS (HIST-ETAS) by a Bayesian procedure (ABIC).
前震の確率は本震が起きてからどのように変わったか？

青線が本震直前以降の確率変化。左図が一個の前震の場合で右図が複数個の前震（赤線）の場合。

群れの先頭（孤立地震を含む）が前震である確率の地域性

群れの先頭（孤立地震を含む）が前震である確率の地域性
Thank you very much for your attention.

Software available.
Please visit WEB home page of our project team:
http://www.ism.ac.jp/~ogata/Ssg/ssgE.html