Statistical features and cluster-based discrimination of foreshocks

Shunichi Nomura and Yosihiko Ogata
The Institute of Statistical Mathematics

Introduction

- Anomaly-type foreshock discrimination methods may provide high foreshock probabilities with optimized anomaly thresholds. However, they would miss some portion of foreshocks not satisfying the anomalies.
- Ogata et al. (1996) proposed a logistic regression model to evaluate foreshock probabilities for any seismic clusters. Furthermore, Ogata and Katsura (2012) validated its predictive performance by prospective tests during 1994-2011.
- We review the cluster-based foreshock discrimination by Ogata et al. (1996) and their successive works. Then, we introduce a modification of the model toward operational forecasting.

Cluster-based foreshock discrimination

Step 1. Seismicity Clustering

Step 2. Definition of Foreshocks

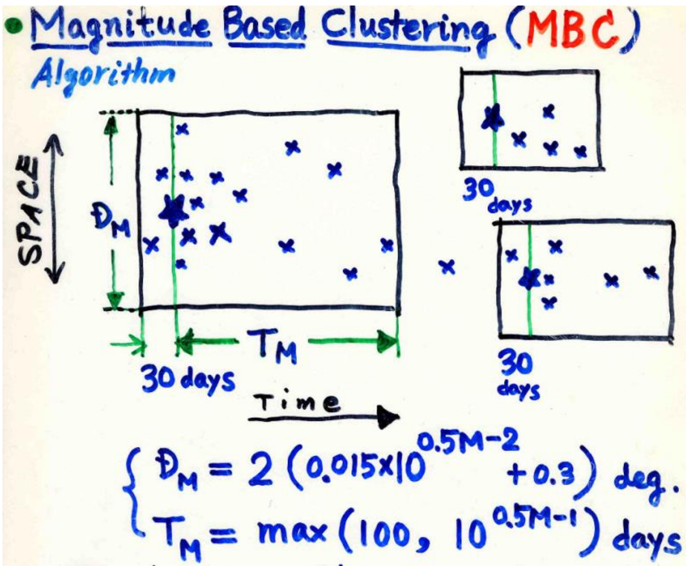
Step 3. Feature Extraction

Step 4. Evaluating Foreshock Probability

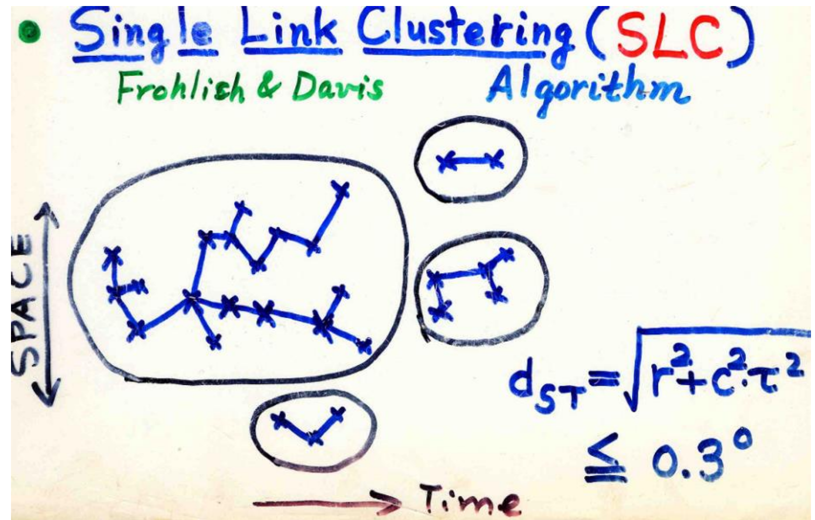
Step 5. Validation of Predictive Performance

Clustering algorithms

Magnitude-based
Utsu (1969,*GBHU*)



Single-Link
Frohlich & Davis (1990,*GJI*)



Reasenberg algorthm (Reasenberg, 1985,*JGR*)

Stochastic clustering (Zhuang et. al., 2002,*JASA;2004,JGR*)

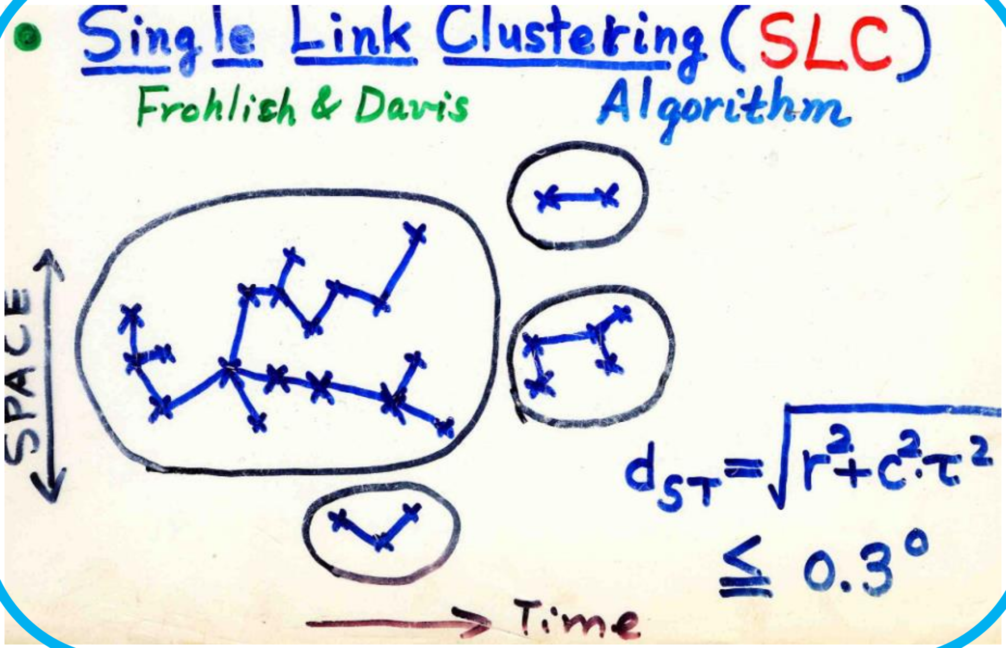
(1) How do we recognize that it is initial earthquake of a cluster?

(2) What is definition of foreshocks?

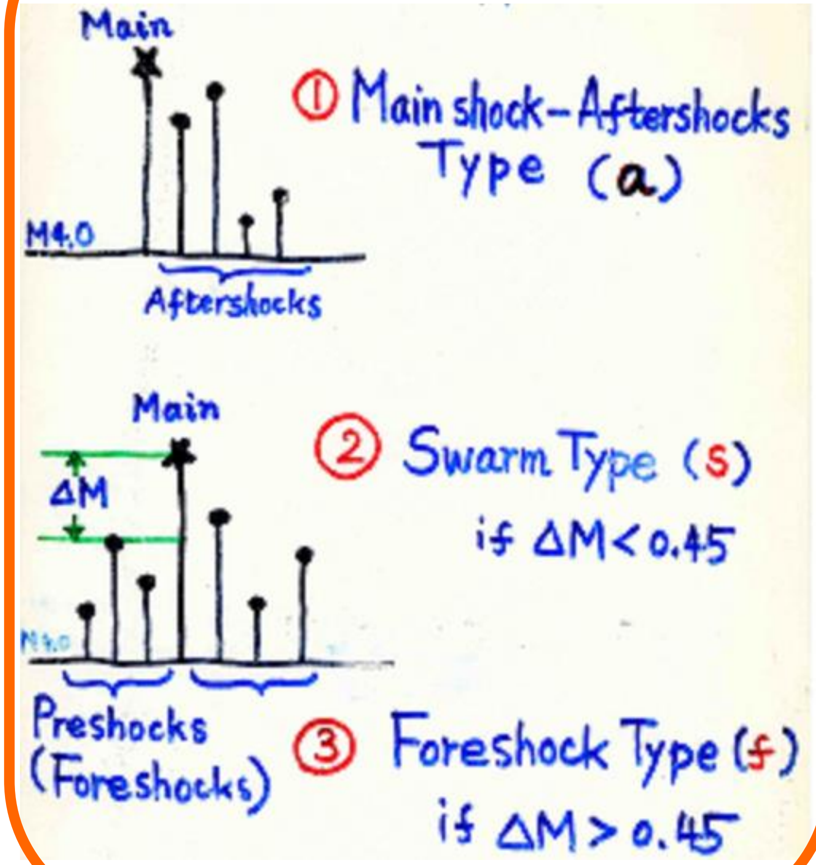
Earthquakes ($M \geqq 4.0$) in JMA catalog after 1926 are connected by the distance criterion

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

where c is the constant to hold 1day = 1km

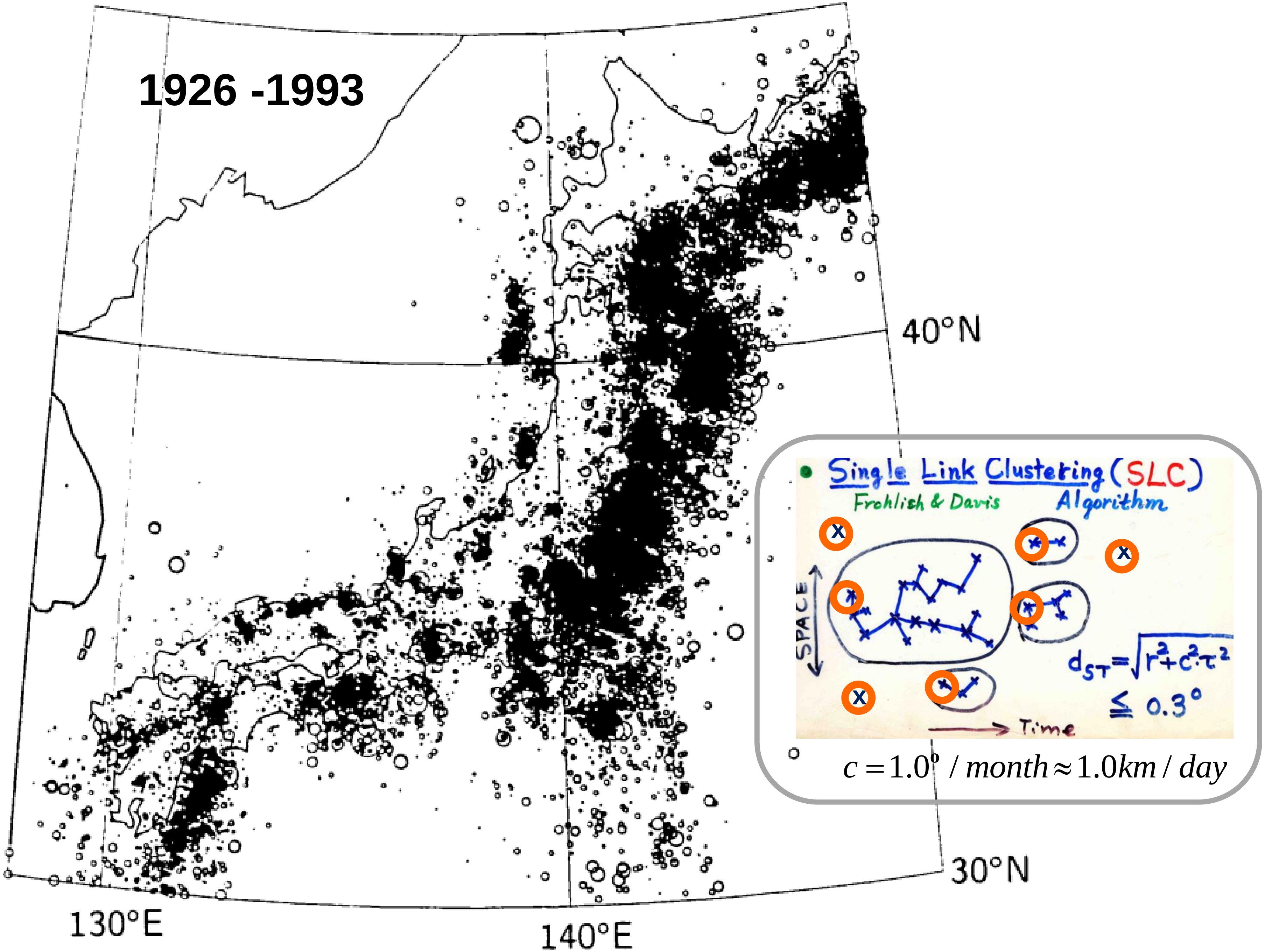


Cluster types

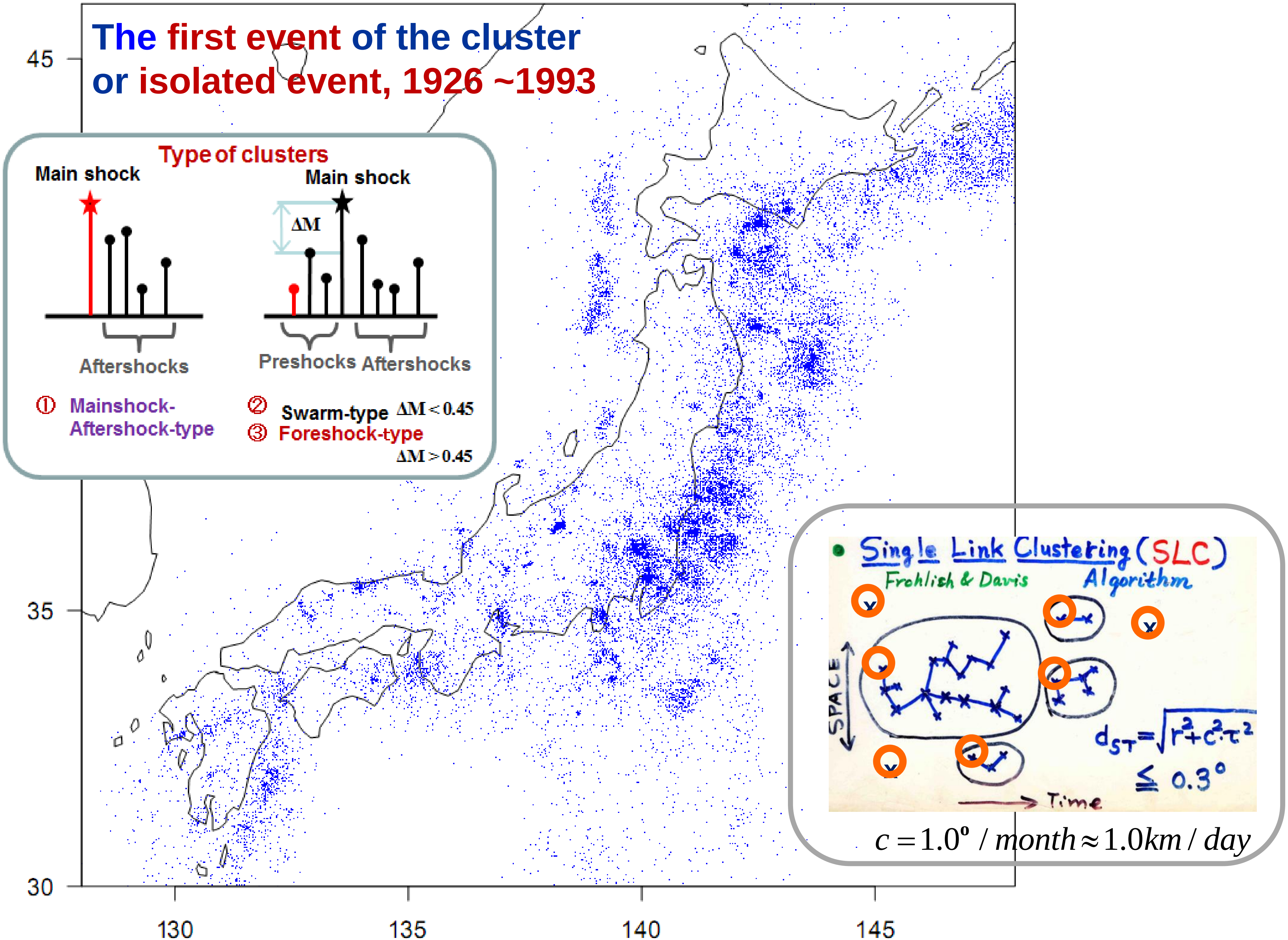


Numbers of foreshocks and swarms

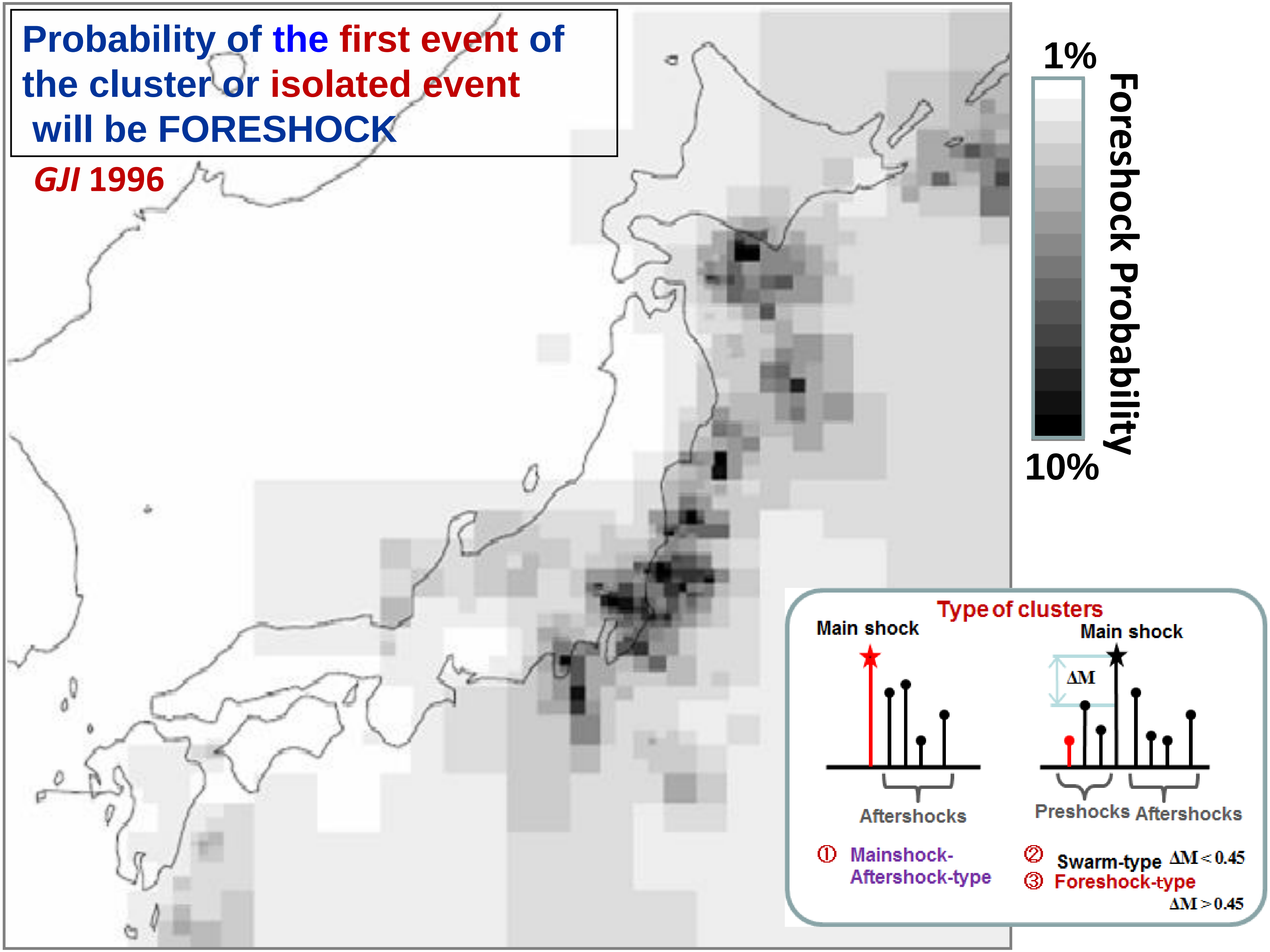
Single-link-clustering 1926–2009, M≧4 from the new JMA catalog								
Cluster	Foreshocks			Swarms			M.A.	All clusters
member#	#c	ratio(%)	s.e.(%)	#c	ratio(%)	s.e.(%)	#c	#c
≧1	1088	6.5	±0.2	--	--	--	14872	16784
≧2	156	5.6	±0.4	824	29.6	±0.9	1800	2780
≧3	78	7.6	±0.8	366	35.4	±1.5	589	1033
≧4	46	8.2	±1.2	213	38.2	±2.1	299	558
≧5	30	7.9	±1.4	145	38.1	±2.5	206	381
≧6	20	7.1	±1.5	108	38.2	±2.9	155	283
≧7	15	7.0	±1.7	77	36.0	±3.3	122	214
≧8	13	7.6	±2	59	34.7	±3.7	98	170
≧9	11	7.9	±2.3	46	33.1	±4	82	139
≧10	11	8.7	±2.5	44	34.9	±4.2	71	126



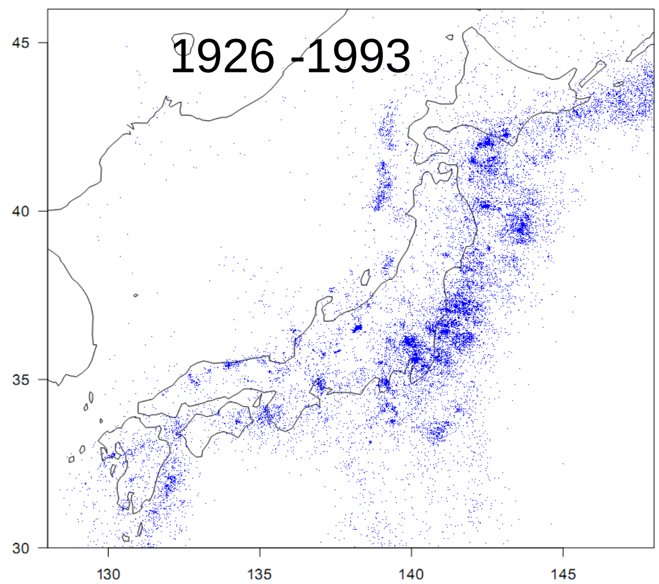
1926 -1993
40°N
30°N
130°E
140°E
Single Link Clustering (SLC)
Frohlish & Davis
Algorithm
SPACE
Time
$d_{ST}=\sqrt{r^2+c^2\cdot\tau^2}$
$\leq 0.3^\circ$
$c = 1.0^\circ / month \approx 1.0 km / day$



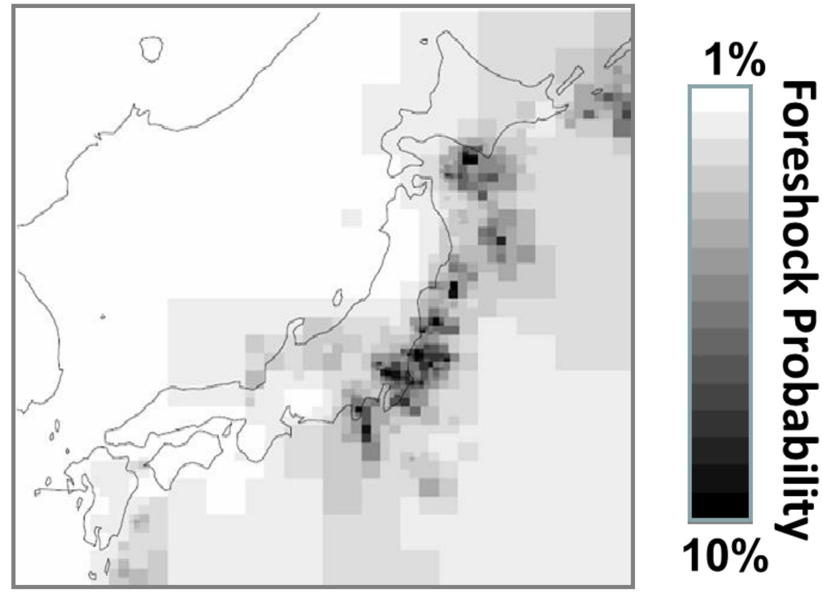
The first event of the cluster
or isolated event, 1926 ~1993
Type of clusters
Main shock
Main shock
ΔM
Aftershocks
Preshocks Aftershocks
① Mainshock-Aftershock-type
② Swarm-type ΔM < 0.45
③ Foreshock-type ΔM > 0.45
Single Link Clustering (SLC)
Frohlish & Davis Algorithm
SPACE
Time
$d_{ST} = \sqrt{r^2 + c^2 \tau^2} \leq 0.3^\circ$
$c = 1.0^\circ / month \approx 1.0 km / day$
45
35
30
130
135
140
145


Probability of the first event of the cluster or isolated event will be FORESHOCK
GJI 1996
1%
Foreshock Probability
10%
Type of clusters
Main shock
Main shock
ΔM
Aftershocks
Preshocks
Aftershocks
① Mainshock-Aftershock-type
② Swarm-type ΔM < 0.45
③ Foreshock-type ΔM > 0.45

Initial earthquakes of clusters or Isolated earthquakes



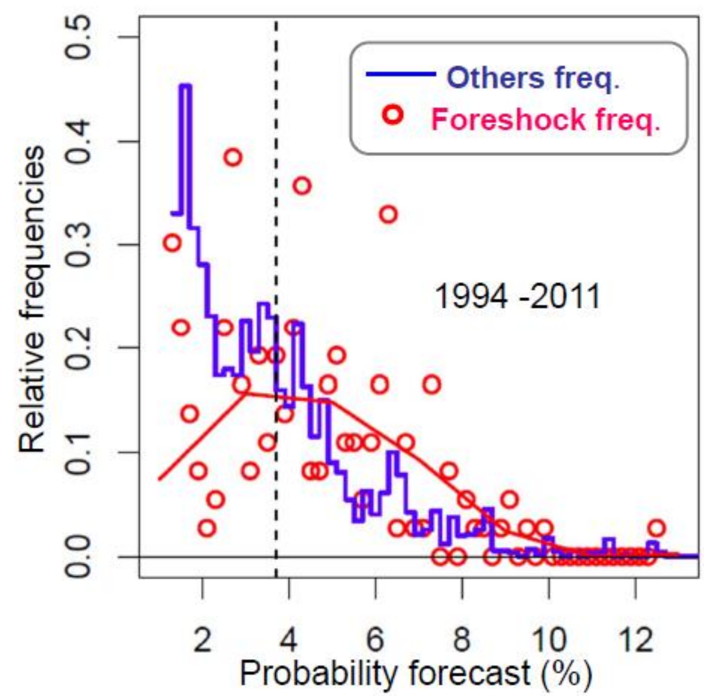
Probability of the initial earthquakes Being foreshock during 1926 – 1993



Forecasted results for 1994 – 2011Mar

Forecast	0-2.5%	2.5-5%	5%-	All
Foreshocks	33	84	65	182
Others	1572	1849	770	4191
All types	1605	1933	835	4373
Ratio(%)	2.1	4.3	7.8	4.2

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



Multiple earthquakes in a cluster

Measuring inter-events concentrations in a cluster and magnitude increments

$$\{(t_i^c, x_i^c, y_i^c, M_i^c) : i = 1,2,\ldots,n\}$$

Hypocenters in a cluster $c \in C$

① Origin-time differences

$$t_{ij}^c = t_j^c - t_i^c$$

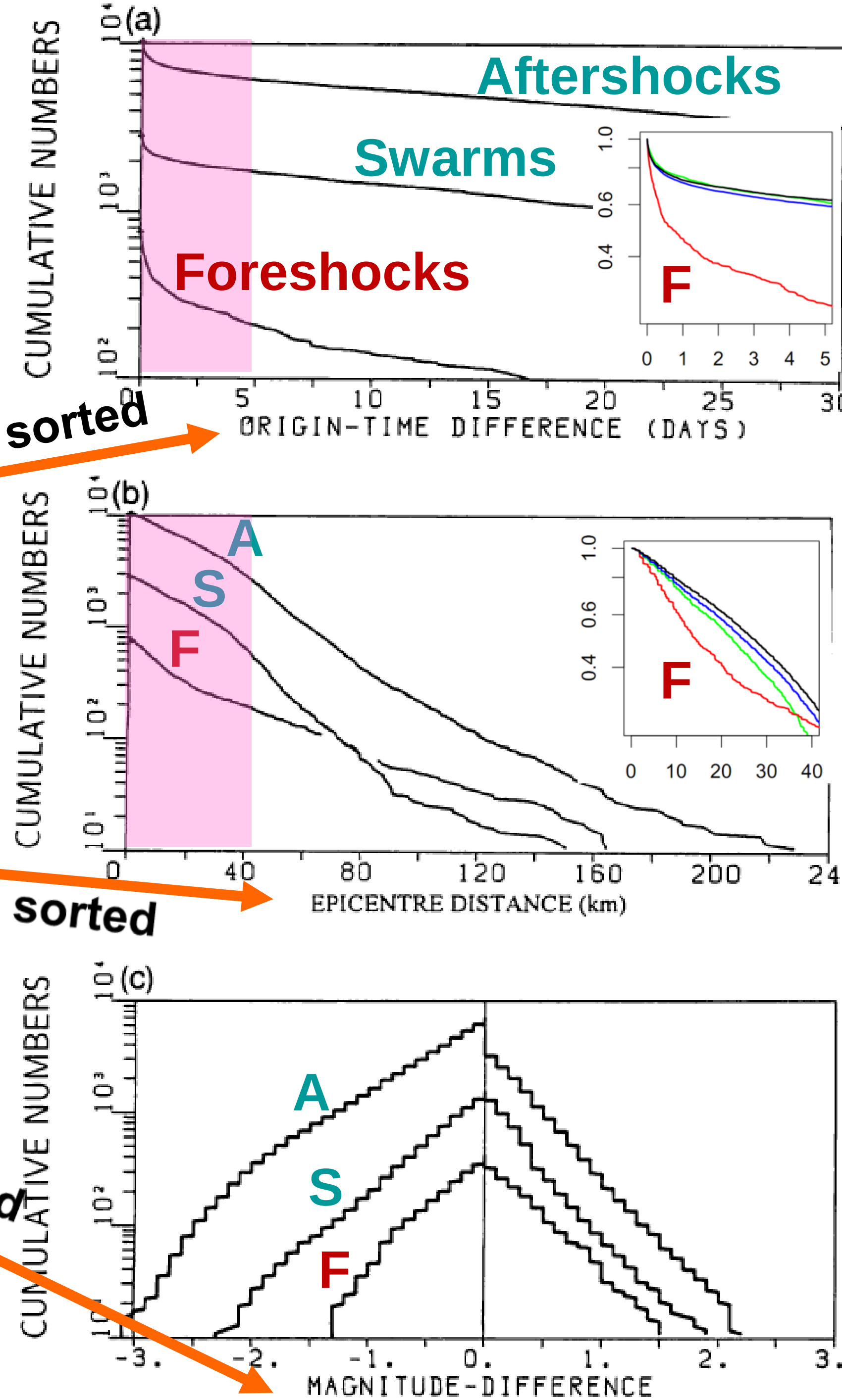
for any pairs i, j such that $i<j$

② Epicenter sparations

$$r_{ij}^c = \sqrt{(x_j^c - x_i^c)^2 + (y_j^c - y_i^c)^2} \quad (i<j)$$

③ Magnitude differences

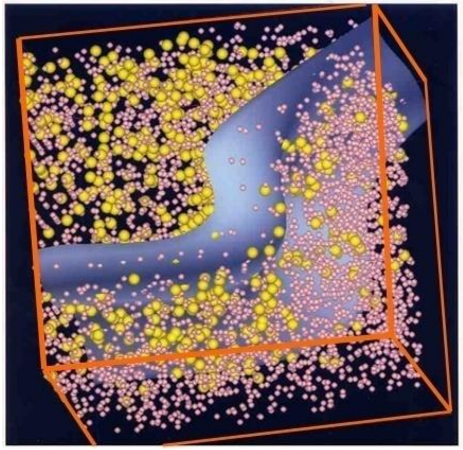
$$M_{ij}^c = M_j^c - M_i^c \quad (i<j)$$



Normalized time differnce, distance & magnitude difference in the unit cube

Time difference, Distance & Magnitude difference Normalization

$(t, r, g) \rightarrow (\tau, \rho, \gamma)$ in $[0,1]^3$ $\equiv$



Time Interval Transformation

$$\tau = \begin{cases} 0 & \text{for} \quad t \leq 0.01 \\ \log(100t)/\log(3000) & \text{for } 0.01 < t \leq 30 \\ 1 & \text{for} \quad 30 \leq t \end{cases}$$

Epicenter Separation Transformation

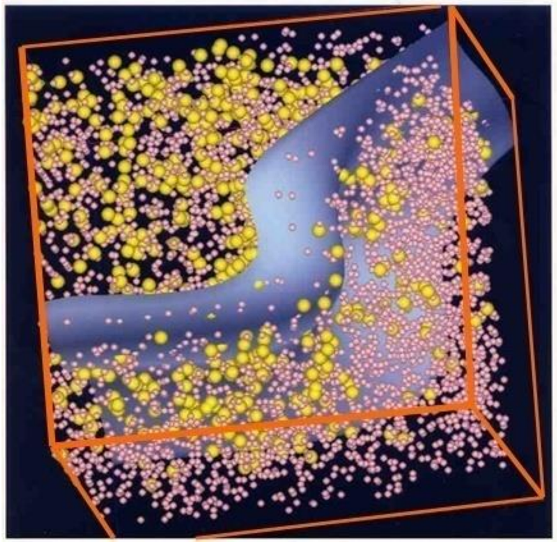
$$\rho = 1 - \exp\{-\min(r, 50)/20\}$$

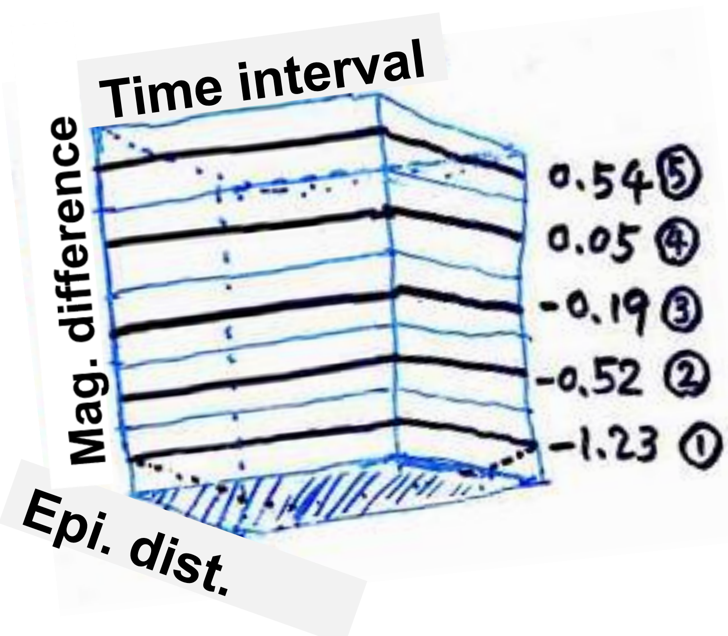
Magnitude Difference Transformation

$$\gamma = \begin{cases} (2/3)\exp\{g/\sigma_1\} & \text{for } g \leq 0 \\ (2/3) + (1/3)[1 - \exp\{-g/\sigma_2\}] & \text{for } g > 0 \end{cases}$$

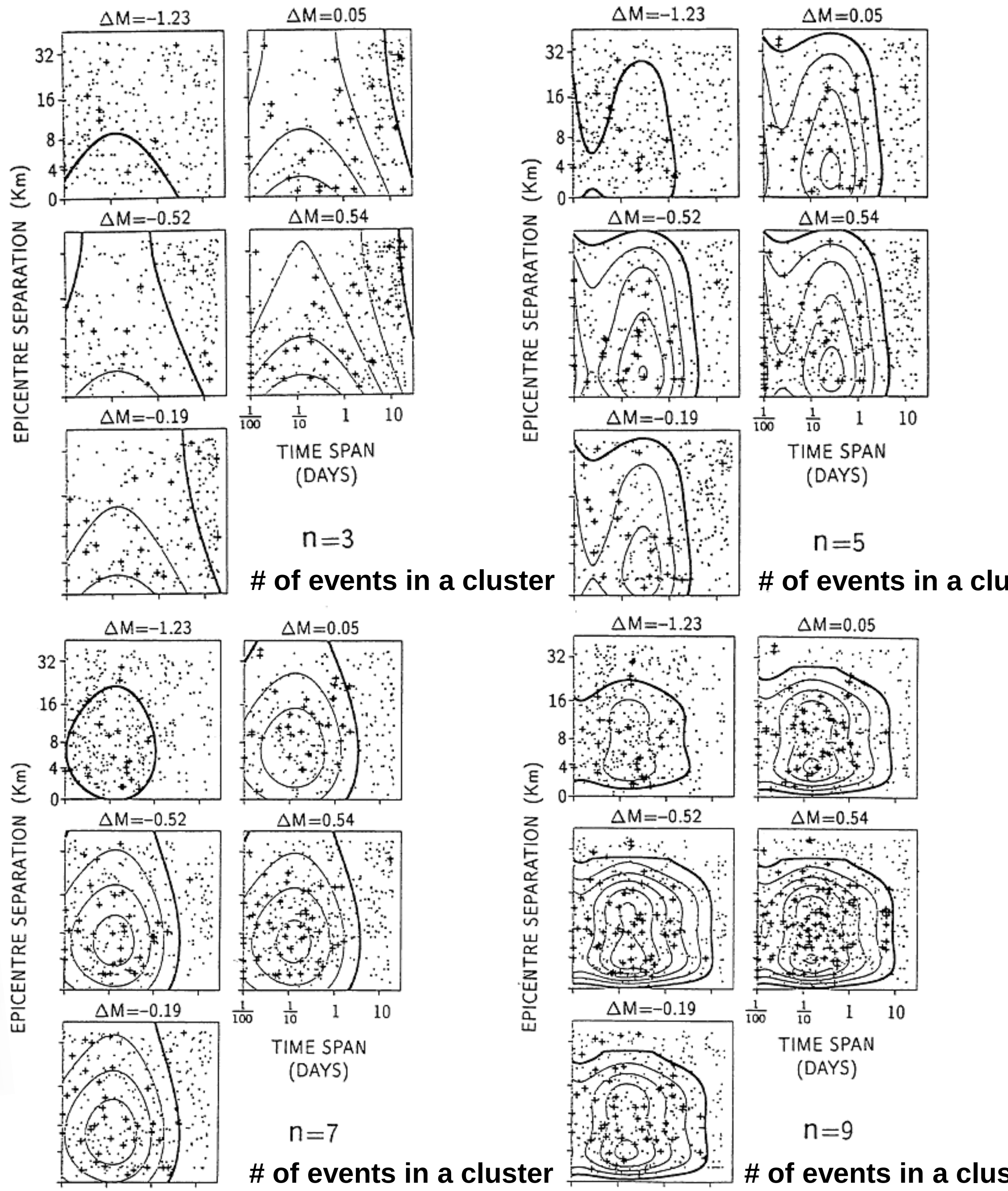
ただし $\sigma_1 = 6709, \sigma_2 = 0.4456$

Normalized time, distance & magnitude difference in the unit cube





Foreshock probab. on sliced planes of mag-difference



GJI 1996

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

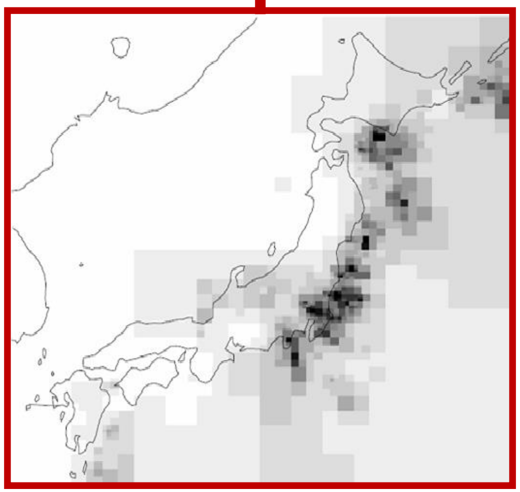
For plural earthquakes in a cluster, time differences t_{ij} (days), epicenter separation r_{ij} (km), magnitude difference g_{ij} are transformed into the unit cube

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

Probability p_c is calculated sequentially

$$\mathrm{Logit}(p_c) = \mathrm{Logit}\{\mu(x_1, y_1)\} + \frac{1}{\#\{i<j\}} \sum_{i<j} \left[a_1 + \sum_{k=1}^{3} b_k \gamma_{i,j}{}^k + \sum_{k=1}^{3} c_k \rho_{i,j}{}^k + \sum_{k=1}^{3} d_k \tau_{i,j}{}^k \right]$$

where $\mathrm{Logit}(p) = \log\left\{\frac{1-p}{p}\right\} = f \Leftrightarrow p = \frac{e^f}{1+e^f}$



Ogata, Utsu and Katsura, 1996, *GJI*)

k	a_k	b_k	c_k	d_k
1	8.018	–33.25	–1.490	–10.92
2		62.77	2.805	295.09
3		–37.66	–2.190	–1161.5

Hint of statistical modeling of probability prediction

Multi-element probability prediction formula

$$P(M \mid A, B, C, \cdots, S) = \frac{1}{1 + \left(\frac{1}{P_A} - 1\right)\left(\frac{1}{P_B} - 1\right)\left(\frac{1}{P_C} - 1\right)\cdots\left(\frac{1}{P_S} - 1\right) \Big/ \left(\frac{1}{P_0} - 1\right)^{N-1}}$$

$$logit\ P = logit\ P_A + logit\ P_B + logit\ P_C + \cdots + logit\ P_S - (N-1)\, logit\ P_0$$

where $$f = logit\ p \equiv \log\{p/(1-p)\} \Leftrightarrow p = \frac{e^f}{1-e^f}$$

Binomial likelihood

$$L(p) = \prod_{i=1}^{N} p_i^{\eta_i} (1-p_i)^{(1-\eta_i)}$$

$$\log L(p) = \sum_{i=1}^{N} \{\eta_i \log p_i + (1-\eta_i)\log(1-p_i)\}$$

Prospective foreshock forecast experiment

***GJI* 1996**

$$\mathrm{Logit}(p_c) = \mathrm{Logit}\{\mu(x_1, y_1)\} + \frac{1}{\#\{i<j\}} \sum_{i<j} \left[a_1 + \sum_{k=1}^{3} b_k \gamma_{i,j}{}^{k} + \sum_{k=1}^{3} c_k \rho_{i,j}{}^{k} + \sum_{k=1}^{3} d_k \tau_{i,j}{}^{k} \right]$$

***GJI* 2012**

Geophysical Journal International

Geophys. J. Int. (2012) **191,** 1237–1244 doi: 10.1111/j.1365-246X.2012.05645.x

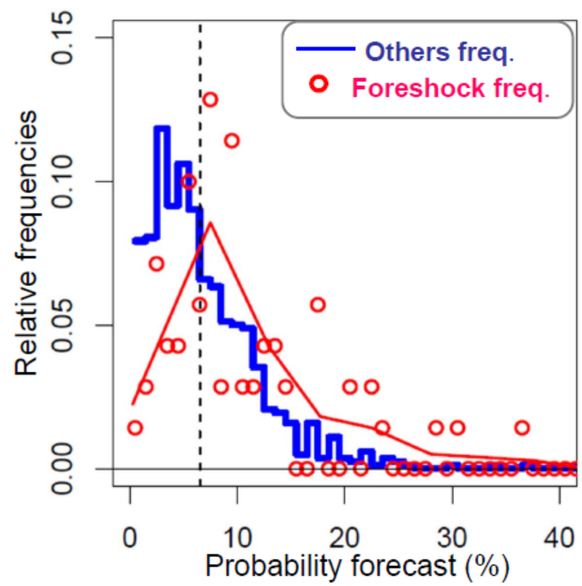
Prospective foreshock forecast experiment during the last 17 years

Yosihiko Ogata[1,2] and Koichi Katsura[1]

Forecast implementation during 1994 – April 2011

Forecast & performance

Forecast(%)	2.5	5%	10%	15%	|	
#Foreshocks	4	10	30	12	14 |	70
#Others	179	211	263	115	51 |	819
Totals	183	221	293	127	65 |	889
Relative Frequency	2.2	4.5	10.2	9.4	21.5 |	7.9



Forecasted sequence and evaluation （１９９４－２０１１Mar）

#	F?	#C	Pc	ENTRPY	CU~ENT	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
1	-	1	5.14%	-0.01537	-0.01537	5.14%									
2	-	2	10.06%	-0.06863	-0.08400	7.46%	12.66%								
3	-	1	18.58%	-0.16822	-0.25222	18.58%									
4	-	1	10.71%	-0.07592	-0.32814	10.71%									
5	-	1	0.15%	0.03586	-0.29228	0.15%									
6	-	1	1.70%	0.02028	-0.27200	1.70%									
7	-	4	9.50%	-0.06243	-0.33443	9.14%	11.17%	7.87%	9.82%						
8	-	1	6.03%	-0.02484	-0.35927	6.03%									
9	-	1	1.77%	0.01950	-0.33977	1.77%									
10	+	1	13.14%	1.27605	0.93628	13.14%									
.....															
875	+	80	9.2%	0.923	28.649	6.7%	27.8%	27.7%	20.1%	14.0%	14.2%	13.6%	11.6%	15.7%	11.9%
						10.1%	8.2%	10.1%	11.7%	10.9%	10.6%	11.5%	11.1%	9.9%	8.2%
						7.2%	6.8%	7.6%	7.3%	7.4%	6.7%	7.0%	7.0%	8.0%	8.5%
						8.6%	8.2%	8.0%	8.1%	8.4%	7.8%	7.3%	7.5%	7.8%	8.1%
						8.1%	7.8%	7.4%	7.7%	7.8%	7.6%	7.2%	7.2%	6.9%	6.8%
						6.7%	7.4%	8.0%	7.8%	7.6%	7.7%	8.3%	9.0%	8.7%	8.5%
						8.6%	8.3%	8.4%	8.2%	8.2%	8.0%	7.9%	7.9%	8.4%	8.4%
						8.6%	8.5%	8.6%	8.4%	8.2%	8.4%	8.3%	8.3%	8.1%	7.9%
.....															
880	-	11	2.44%	0.01266	31.60644	4.69%	4.77%	6.21%	3.42%	1.74%	1.24%	1.04%	0.90%	0.83%	0.97%
						1.03%									
881	-	16	2.11%	0.01604	31.62248	0.03%	0.25%	0.51%	0.83%	2.77%	2.21%	2.02%	3.19%	2.78%	2.50%
						2.43%	3.07%	2.92%	2.74%	2.84%	2.68%				
882	-	7	1.47%	0.02259	31.64507	0.06%	0.79%	1.70%	2.06%	1.90%	1.90%	1.88%			
883	-	1	4.51%	-0.00878	31.63629	4.51%									
884	-	1	3.84%	-0.00178	31.63451	3.84%									
885	+	7	5.04%	0.31698	31.95149	6.89%	7.42%	4.88%	3.98%	3.56%	4.05%	4.49%			
886	-	1	2.84%	0.00853	31.96002	2.84%									
887	-	1	7.00%	-0.03518	31.92483	7.00%									
888	-	1	7.65%	-0.04219	31.88264	7.65%									
889	-	1	7.83%	-0.04419	31.83845	7.83%									

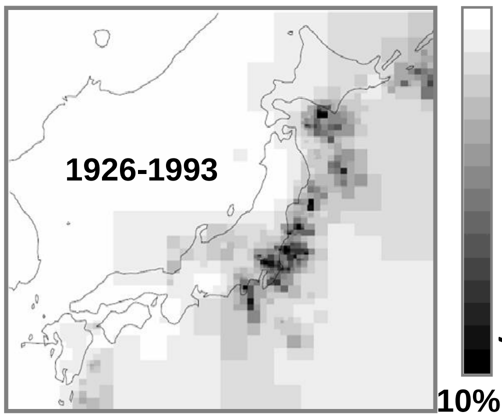
M7.3 Foreshock of 9 Mar 2011

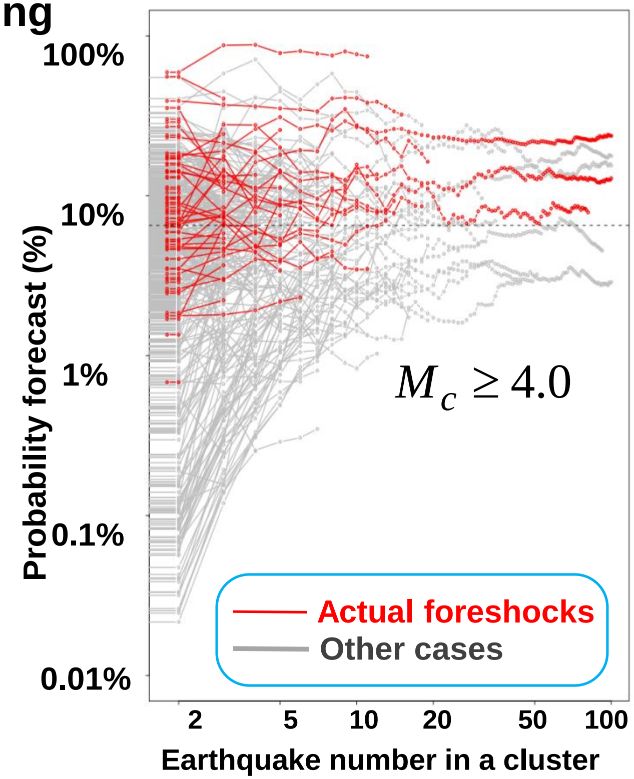
M9.0

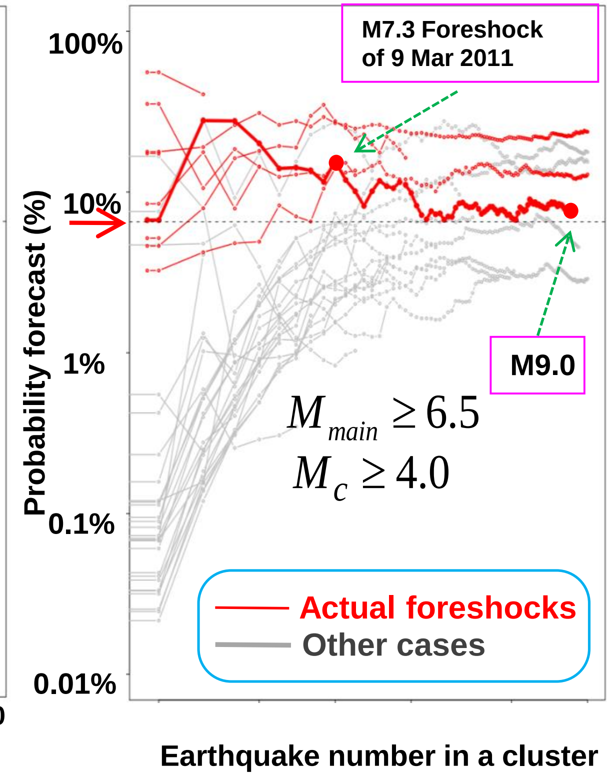
2*Entropy0 = 523.96; 2*Entropy = 460.29: 2*DEntropy = –63.68

Forecast implementation during 1994 – 2011 (March 11)

Probability of the first event of the cluster or isolated event will be FORESHOCK



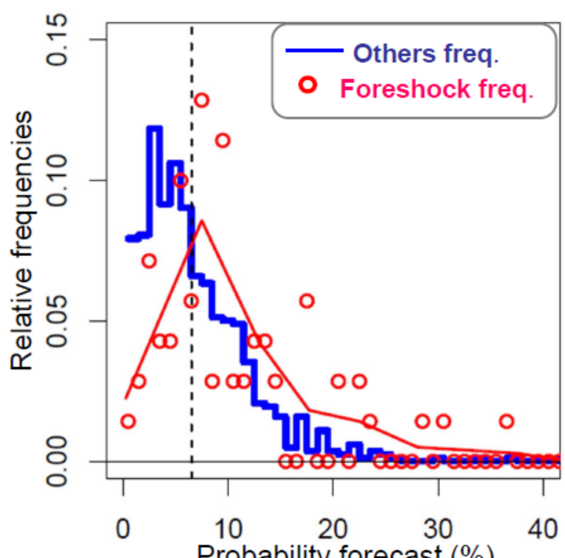




Forecasted results for 1994 – 2011Mar

Forecast	0-2.5%	2.5-5%	5%-	All
Foreshocks	33	84	65	182
Others	1572	1849	770	4191
All types	1605	1933	835	4373
Ratio(%)	2.1	4.3	7.8	4.2

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

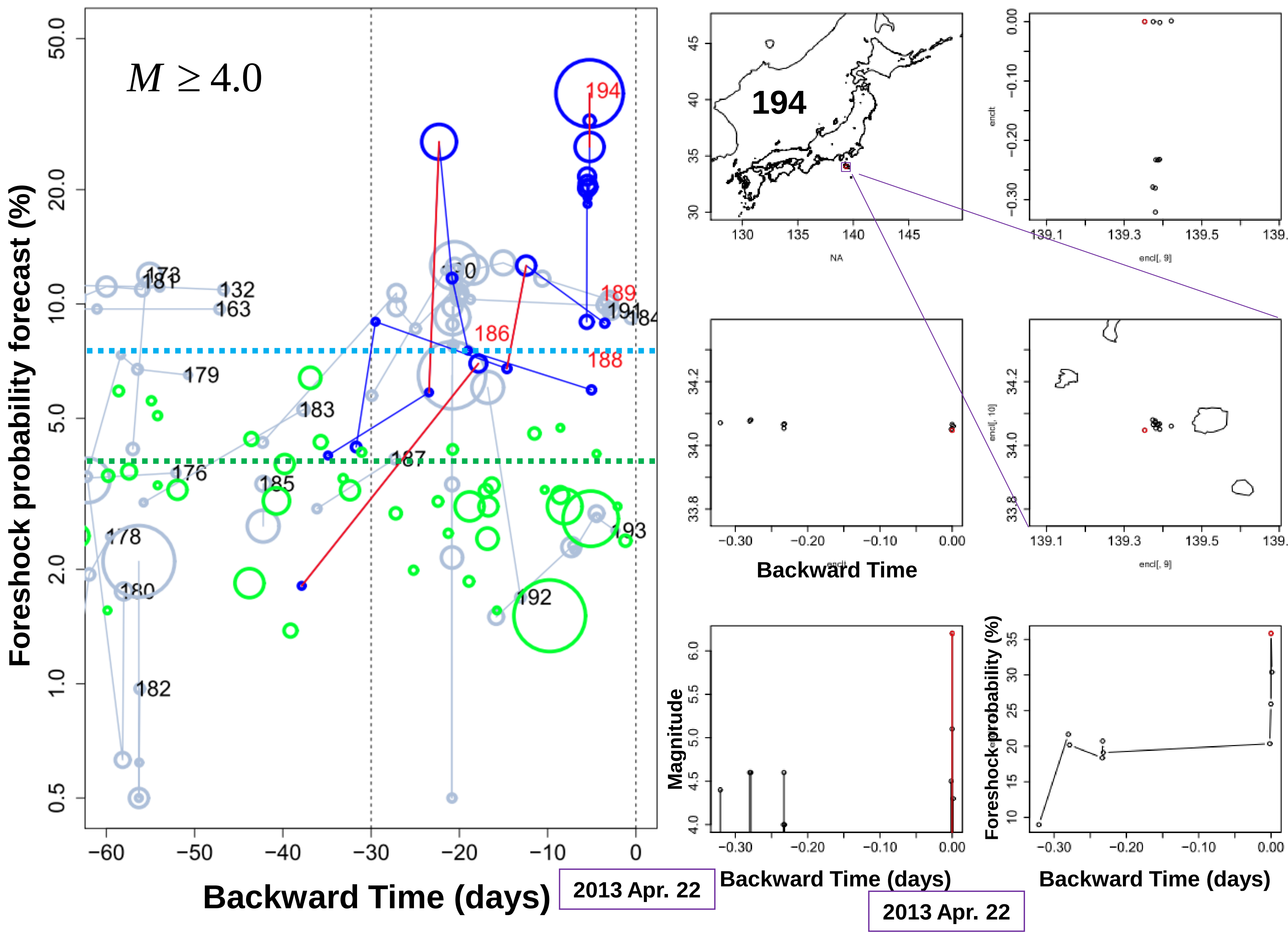


2*Entropy0 = 523.96 / 2*Entropy = 460.29
2*⊿Entropy = −63.68

Forecast & performance

Forecast(%)	2.5	5%	10%	15%		
#Foreshocks	4	10	30	12	14	70
#Others	179	211	263	115	51	819
Totals	183	221	293	127	65	889
Relative Frequency	2.2	4.5	10.2	9.4	21.5	7.9

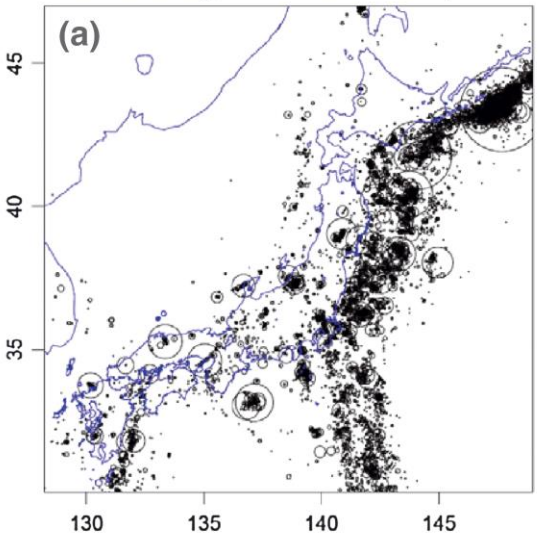
aic0 = 3178.62 aic1 = 3157.15
⊿aic = −21.47



$M \geq 4.0$
Foreshock probability forecast (%)
Backward Time (days)
194
189
186
188
2013 Apr. 22
194
Backward Time
Magnitude
Backward Time (days)
Foreshock probability (%)
Backward Time (days)
2013 Apr. 22

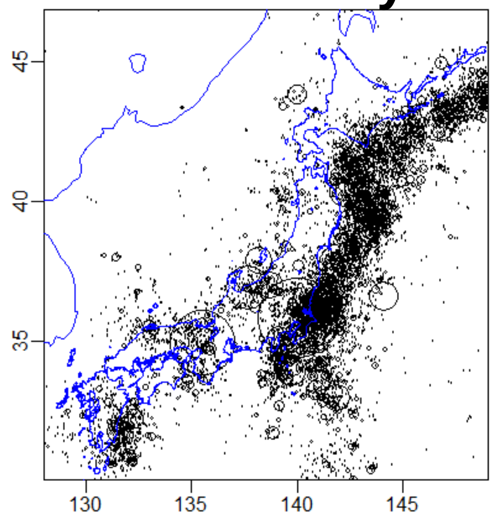
Forecasts in real and synthesized catalogs

JMA catalog 1994 – 2011.3.10, M ≥ 4

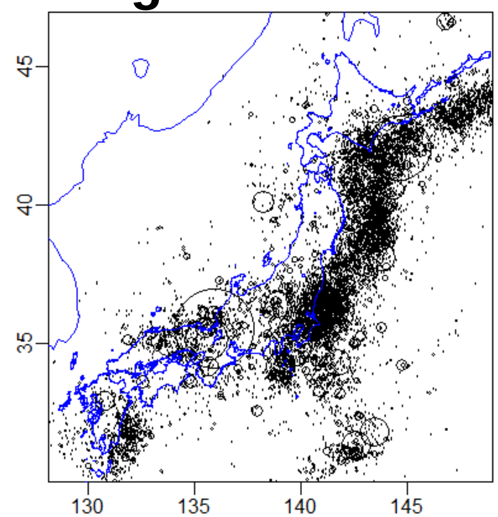


v.s.

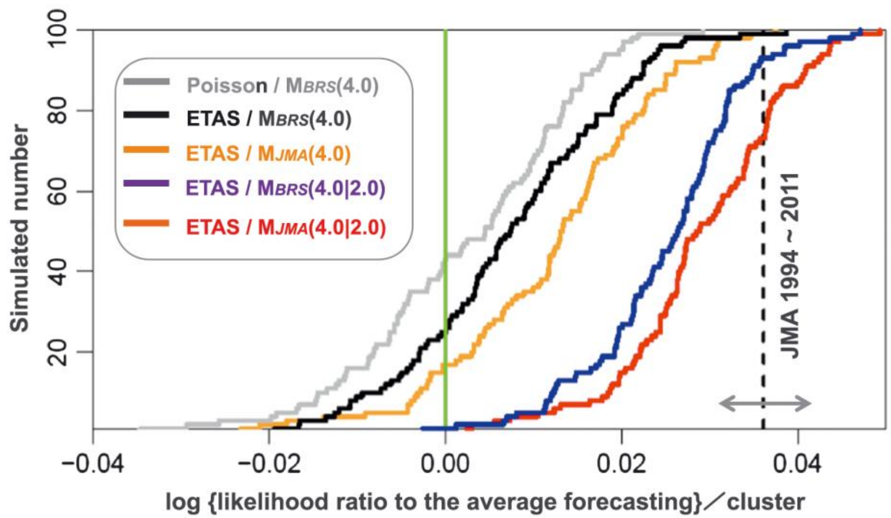
100 synthesized catalogs

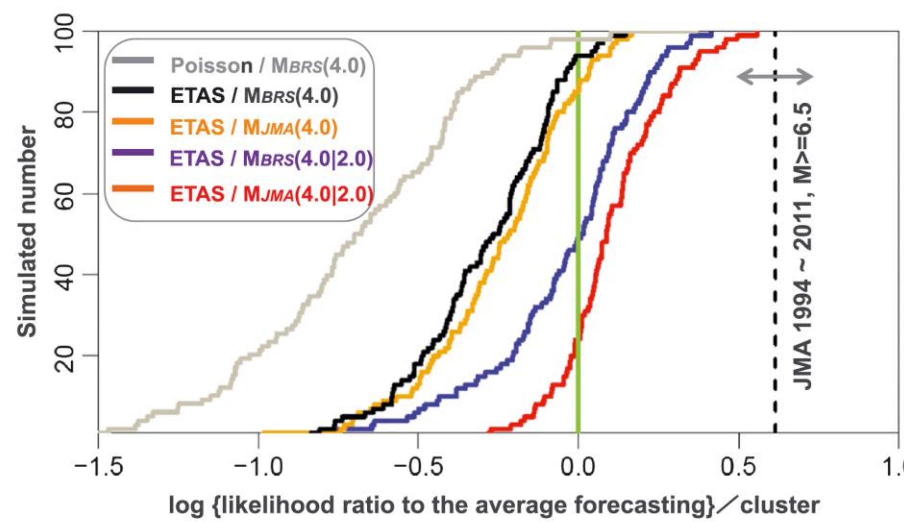


....



***GJR* 2014**

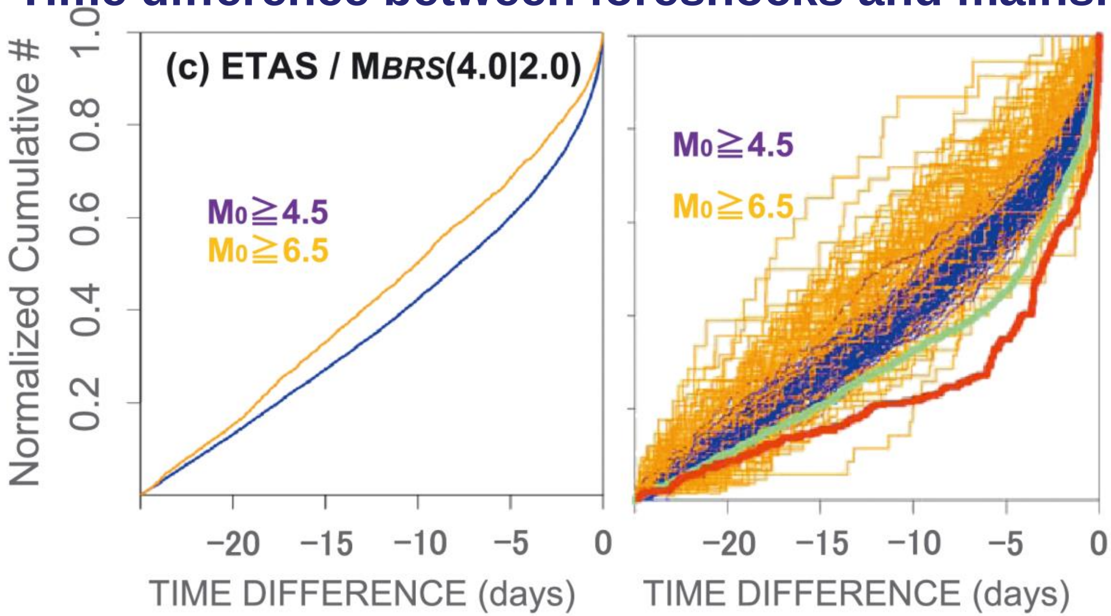




Foreshock features in real and synthesized catalogs

***GJR* 2014**

Time difference between foreshocks and mainshock



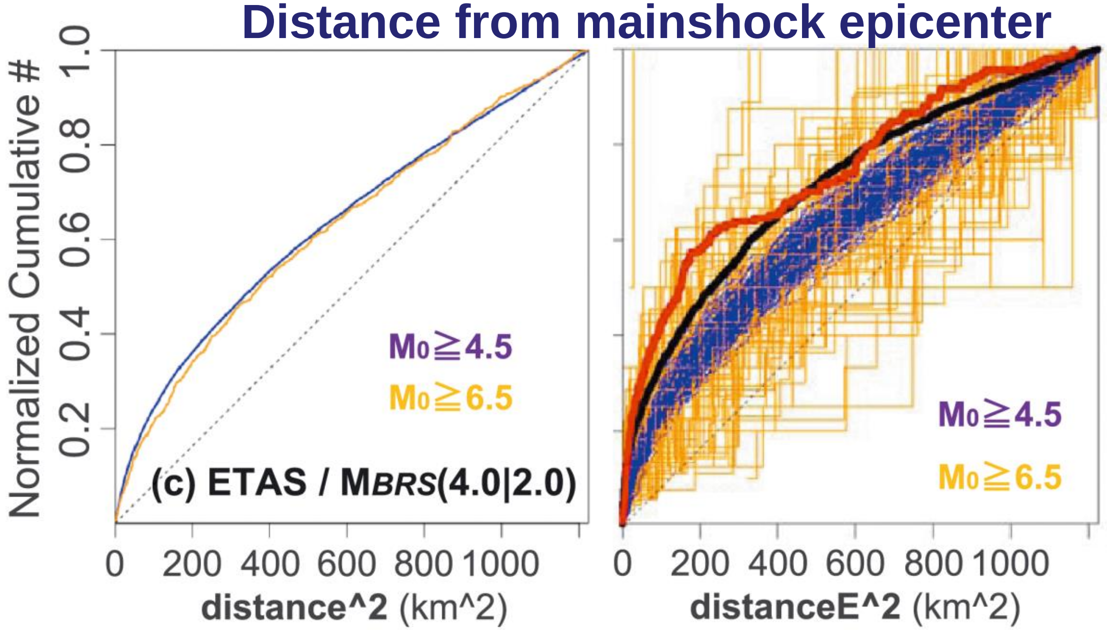
M0 ≥ 4.5 Real catalog

M0 ≥ 4.5 Synthesized

M0 ≥ 6.5 Real catalog

M0 ≥ 6.5 Synthesized

Distance from mainshock epicenter



M0 ≥ 4.5 Real catalog

M0 ≥ 4.5 Synthesized

M0 ≥ 6.5 Real catalog

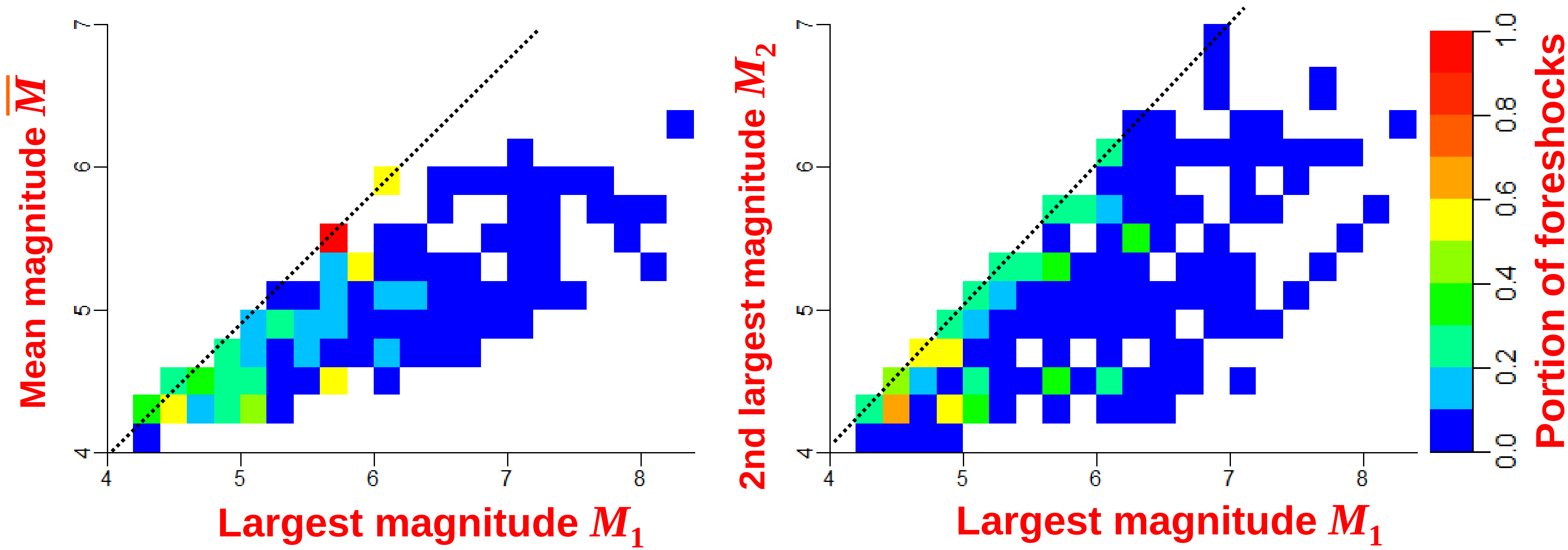
M0 ≥ 6.5 Synthesized

Exploring other features in foreshocks

- Statistical features of foreshocks used in Ogata et al. (1996) are based on difference in space, time and magnitude between every pair in clusters.
- Other features on magnitude histories can also be considered:
  1. Mean magnitude in the cluster : $\bar{M}$
     (equivalent to the maximum likelihood estimate of *b*-value)
  2. Magnitude gap between 2 largest foreshocks : $M_1 - M_2$
     (e.g. the 2016 Kumamoto earthquake (M7.3) is preceded by M6.5 and M6.4 foreshocks)

Exploring other features in foreshocks

- Other features on magnitude histories can also be considered:
  1. Mean magnitude in the cluster : $\bar{M}$
     (equivalent to the maximum likelihood estimate of *b*-value)
  2. Magnitude gap between 2 largest foreshocks: $M_1 - M_2$
     (e.g. the 2016 Kumamoto earthquake (M7.3) is preceded by M6.5 and M6.4 foreshocks)



Toward operational forecasting

- Toward operational forecasting, we modified certain aspects of the foreshock discrimination model proposed by Ogata et al. (1996). Specificlly, we set

  1. Forecasting period at 30 days from the last event,
  2. Targets of forecast at events whose magnitudes are over the largest foreshock magnitude,
  3. Exponential distribution for magnitudes of target events.

Cluster-based Foreshock Discrimination

Step 1. Seismicity Clustering

Step 2. Definition of Foreshocks

Step 3. Feature Extraction

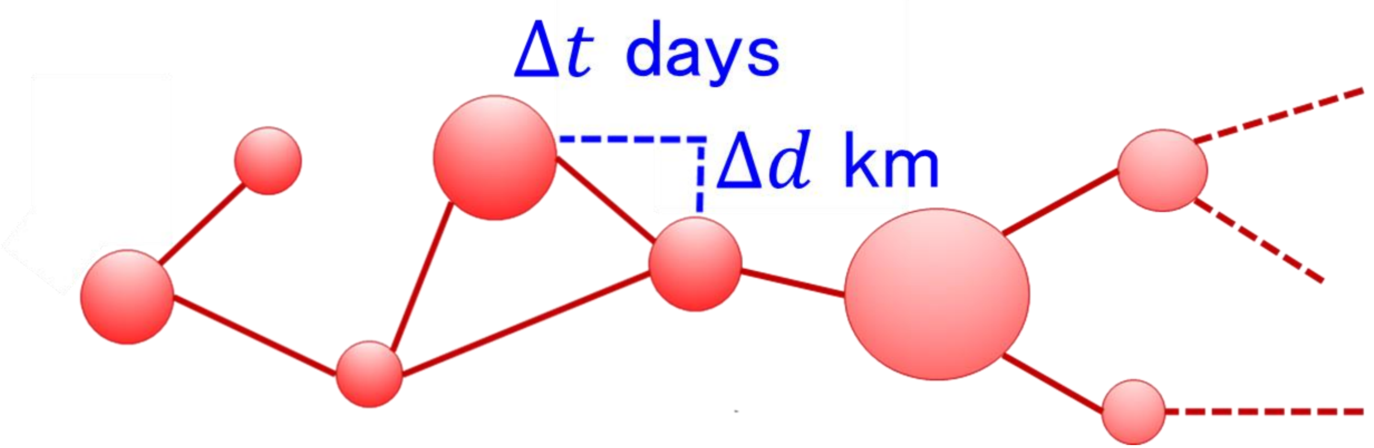
Step 4. Evaluating Foreshock Probability

Step 5. Validation of Predictive Performance

Step 1. Seismicity Clustering

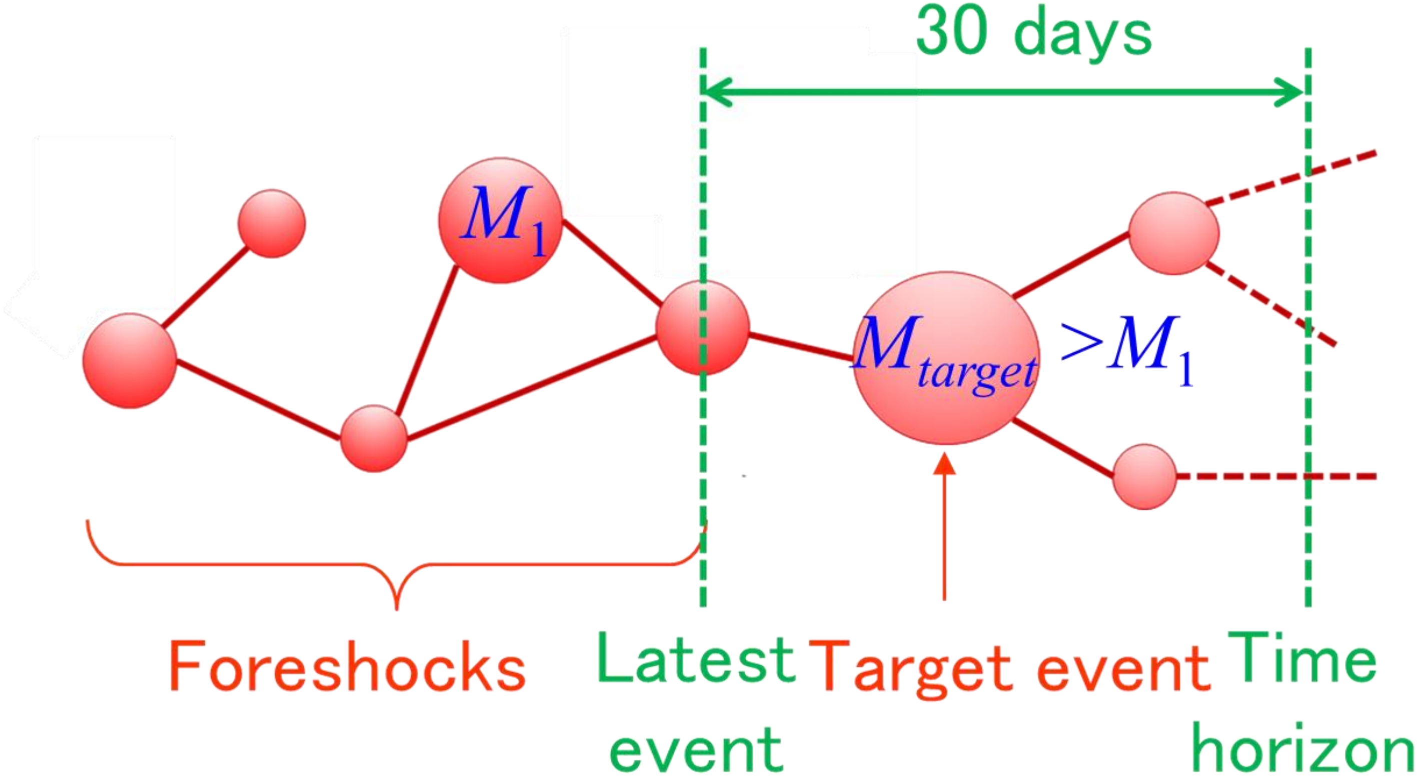
- We analyzed the JMA catalog of $M \geq 4$ in the region 128–148°E, 30–46°N observed from January 1, 1926 to October 31, 2017 at depth shallower than 100 km.
- First, we construct seismic clusters from that catalog by single-link method by Frohlich & Davis (1990) which connects pairs of earthquakes within certain space-time distance.

$$\sqrt{(\Delta d)^2 + (c\Delta t)^2} \leq 33.33\text{km} \quad (c = 1.11 \text{ km/day})$$



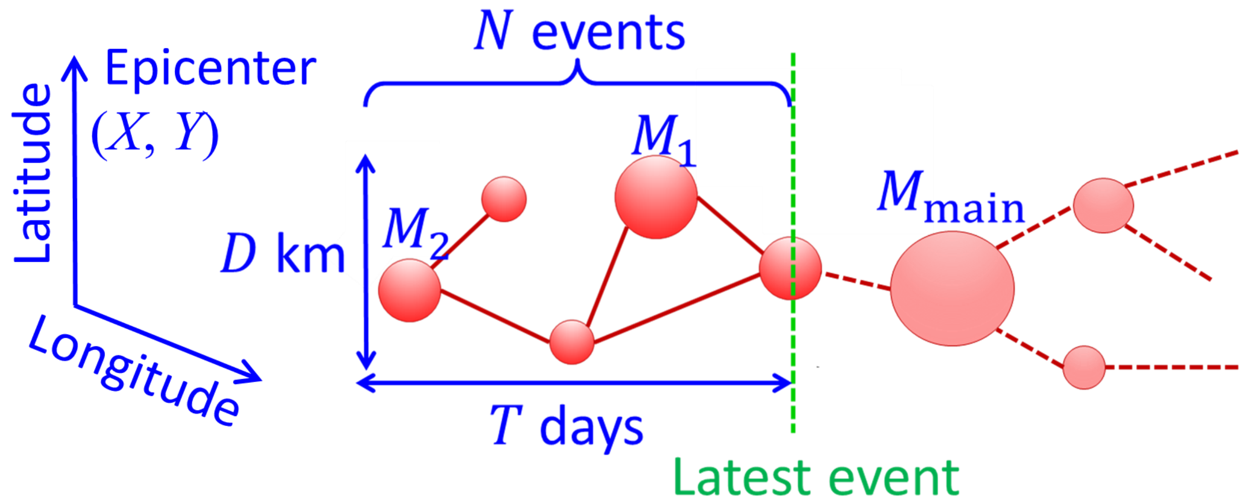
Step 2. Definition of Targets and Foreshocks

- For each evolving seismic cluster, we define the target event of forecast by the larger earthquake than the largest event up to the present in each evolving cluster.
- We define foreshocks by the evolving cluster that a target event occurs within 30 days after the latest event.



Step 3. Feature extraction

- Extract the following features from an evolving cluster and update them by every occurrence of an new event.
  - Number of earthquakes : $N(\geq 2)$
  - Largest magnitude : M_1
  - Difference between two largest magnitudes: $\Delta M = M_1 - M_2$
  - Duration time : T (days)
  - Mean pairwise distance : D (km)
  - Central location in longitude and latitude : X, Y (deg)



Step 4. Evaluating Foreshock Probability

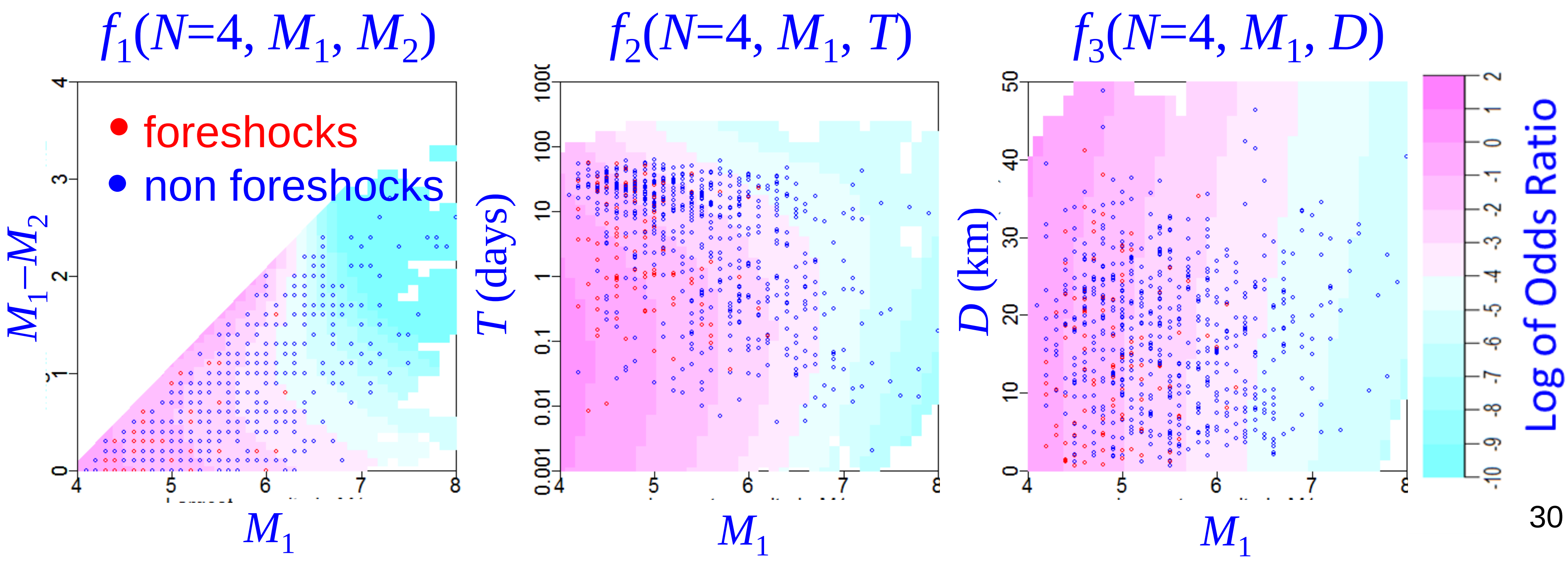
- Estimate the following non-linear logistic regression model from training dataset observed during 1926-1999.

$$\text{logit}\, P(\text{foreshock} \mid N, M_1, \Delta M, T, D, X, Y)$$
$$= g(X, Y) + f_1(N, M_1, \Delta M) + f_2(N, M_1, T) + f_3(N, M_1, D) + \alpha_s$$

  - $P(\text{foreshock} \mid N, M_1, \Delta M, T, D, X, Y)$: Foreshock probability
  - $\text{logit}\, p = \log\{p/(1-p)\}$: Logit function $(0 < p < 1)$
  - g: Location-dependent foreshock probability
  - $f_1,\ f_2,\ f_3$: log of odds ratio between foreshock and others (represented by cubic B-spline functions)
  - α_s: Random effect for the s-th cluster

Step 4. Estimated Log of Odds Ratio Functions

- As the largest magnitude M_1 gets smaller, the odds ratio gets higher because M_1 is the cut-off magnitude of target events.
- Within the same M_1, The odds ratio is higher as the magnitude difference of the two largest events $\Delta M = M_1 - M_2$ is smaller and the time duration T is shorter.

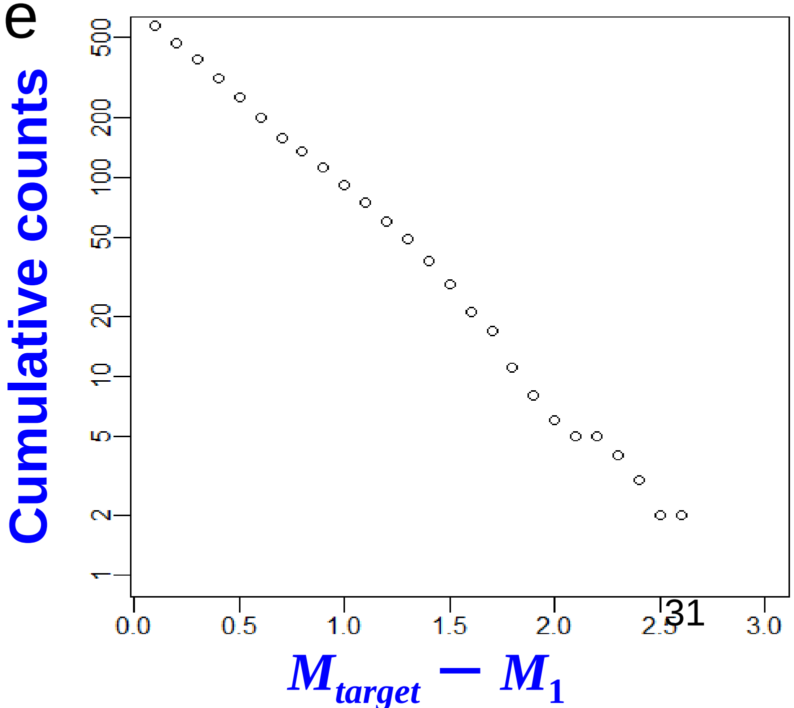


Step 4. Magnitude of Target Event

- Given a foreshock cluster, we also predict magnitude of the target event by exponential distribution over the largest foreshock magnitude:

$$p(M_{target} > M_1 + m \mid \text{foreshock}, M_1) = 10^{-0.8m}. \quad m = 0.1, 0.2, \ldots$$

  - M_{target} : Magnitudes of the target event
  - M_1 : Largest foreshock magnitude



Step 5. Validation of Predictive Performance

- We applied the model estimated from training dataset observed during 1926-1999 to validation dataset observed during 2000-2017.
- The evaluated foreshock probabilities are consistent with the 95% confidence interval of portion of foreshocks in the validation dataset.

Evaluated foreshock probability		0-10%	10-20%	20-30%	30-40%	40-50%	50-60%	60-70%	70-80%	Total
N=2	All clusters	469	489	220	59	14	2	0	0	1,253
	Foreshock clusters	27	67	45	17	4	2	0	0	162
	Portion of foreshocks &	6%	14%	20%	29%	29%	100%	N/A	N/A	13%
	95% confidence interval	4-8%	11-17%	15-26%	18-42%	8-58%	16-100%	N/A	N/A	11-15%
N=5	All clusters	129	57	27	13	9	3	2	2	242
	Foreshock clusters	8	12	4	3	3	1	1	2	34
	Portion of foreshocks &	6%	21%	15%	23%	33%	33%	50%	100%	14%
	95% confidence interval	3-12%	11-34%	4-34%	5-54%	7-70%	8-91%	1-99%	16-100%	10-19%
N=10	All clusters	68	7	3	5	4	2	2	0	91
	Foreshock clusters	5	3	1	2	3	1	0	0	15
	Portion of foreshocks &	7%	43%	33%	40%	75%	50%	0%	N/A	16%
	95% confidence interval	2-16%	10-82%	8-91%	5-85%	19-99%	1-99%	0-84%	N/A	10-26%

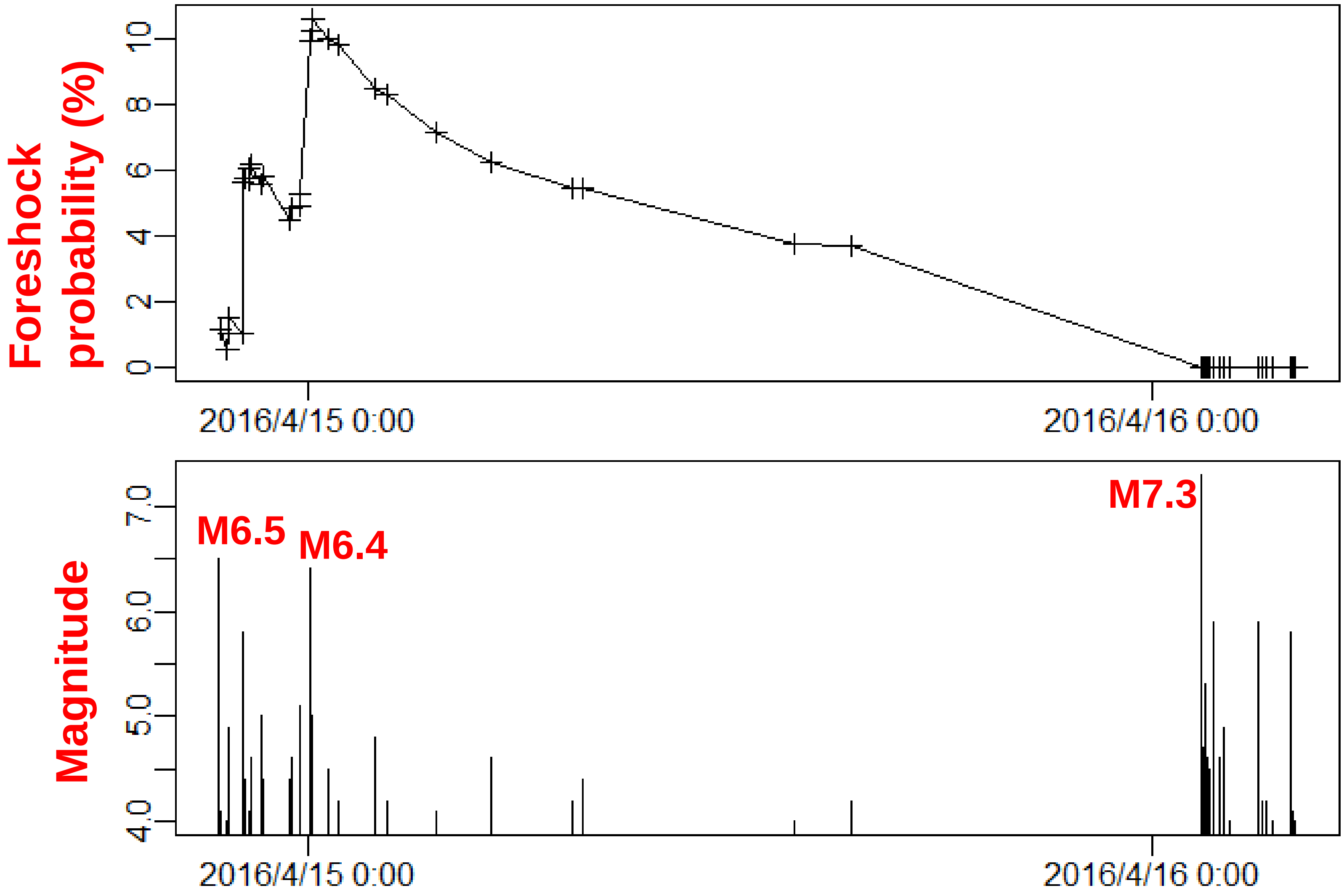
Step 5. Validation of Predictive Performance for M6+

- When we limit magnitudes of target events to M6+, the evaluated foreshock probabilities get lower but are still valid for prediction.

Evaluated foreshock probability for target events of M6+		0-5%	5-10%	10-15%	15-20%	Total
N=2	All clusters	1,229	22	2	0	1,253
	Foreshock clusters	18	0	1	0	19
	Portion of foreshocks &	1%	0%	50%	N/A	2%
	95% confidence interval	1-2%	0-15%	13-99%	N/A	1-10%
N=5	All clusters	205	30	6	1	242
	Foreshock clusters	6	5	0	0	11
	Portion of foreshocks &	3%	17%	0%	0%	5%
	95% confidence interval	1-6%	6-35%	0-46%	0-98%	2-8%
N=10	All clusters	81	6	3	1	91
	Foreshock clusters	6	1	1	1	9
	Portion of foreshocks &	7%	17%	33%	100%	10%
	95% confidence interval	3-15%	0-64%	1-91%	3-100%	5-18%

Evaluation for Foreshock Sequences

- 2016 Kumamoto (M7.3)



Summary

- We proposed cluster-based foreshock discrimination models using information of magnitudes, space, and time in seismic clusters.
- The actual portion of foreshocks was consistent with the evaluated foreshock probabilities in retrospective and prospective tests.
- We will try to develop an ensemble model of our forecasts with aftershock forecasts given by e.g. ETAS models and submit that model to the CSEP Japan Testing Center.

Thank you