Evaluating performance of earthquake prediction relative to a baseline model through the gambling scoring method

Jiancang Zhuang
Institute of Statistical Mathematics

Outlines

- Part I: How to use the background seismicity in long-term earthquake forecasting?
- Part II: How to use gambling scores without a reference model?
  Non-referenced gambling scores

Part I
Using Background Seismicity as a Reference Model

Spatiotemporal long-term earthquake probability models

- Poisson model (smoothed seismicity) – overpredict numbers
- Poisson model (smoothed declustered catalogs) – argument on declustering algorithms
- Background seismicity models – Obtained from stochastic declustering based on a clustering model, explicitly (space-time ETAS model) or implicitly (MISD).

Question: How to use background seismicity for long-term earthquake probability forecast

- ETAS model = Background + Triggered seismicity
- Is a model of "background seismicity" + "Gutenberg-Richter law" still valid?
- Answer: Yes, but only for background seismicity, not for the biggest event.

Assumptions of the model (A simpler version of the ETAS model)

- Background seismicity $\mu(x, y)$: function of spatial locations but stationary in time
- Each event of magnitude m triggers a cluster of mean size $\kappa(m) = Ae^{\alpha(m-m_c)}$
- The magnitude distribution of all the events are identically independently distributed according to the Gutenberg-Richter law

$$s(m) = 10^{-b(m-m_c)} b \ln 10, \quad m \geq m_c$$

- Temporal component is neglected, for the process is stationary
- Spatial component is neglected, because clusters concentrate at the location of the mainshock.

Magnitude distribution of the largest event in a cluster is not G-R law any longer, but

$$
\begin{aligned}
G(m) &= \Pr\{\text{the biggest event is less than } m\} \\
&= \Pr\{\text{all the events in the cluster is less than } m\} \\
&= \Pr\{\text{the initiate is less than } m \text{ and all events in each} \\
&\qquad \text{family of its direct children is less than } m\} \\
&= \int_{m_0}^{m} \Pr\{\text{all events in each child's family is less than } m \mid m_{initiate} = m^*\} \\
&\qquad\qquad \times s(m^*)\,\mathrm{d}m^* \\
&= \sum_{n=0}^{\infty} \int_{m_0}^{m} \Pr\{\text{all events in each child's family is less than } m \mid n \text{ children in total} \\
&\qquad m_{initiate} = m^*\} \Pr\{n \text{ children in total} \mid m_{initiate} = m^*\} \times s(m^*)\,\mathrm{d}m^* \\
&= \sum_{n=0}^{\infty} \int_{m_0}^{m} G^n(m) \frac{\kappa^n(m^*)}{n!} e^{-\kappa(m^*)} s(m^*)\mathrm{d}m^* \\
&= \int_{m_0}^{m} e^{G(m)\kappa(m^*)} e^{-\kappa(m^*)} s(m^*)\mathrm{d}m^* \\
&= \int_{m_0}^{m} e^{-\kappa(m^*)[1-G(m)]} s(m^*)\mathrm{d}m^*
\end{aligned}
$$

Zhuang & Ogata (PRE, 2006)

Magnitude distribution of the largest event

- Magnitude distribution of the largest event in a cluster is not G-R law any longer, but

$$G(m) = \int_{m_c}^{m} e^{-\kappa(m^*)[1-G(m)]} s(m^*)\mathrm{d}m^*$$

a functional equation but can be solved numerically. The complementary function (Zhuang and Ogata, 2006, Physical Review E)

$$F(m) = 1 - G(m) = 1 - \int_{m_c}^{m} e^{-\kappa(m^*)F(m)} s(m^*)\mathrm{d}m^*$$

Solution of the functional equation

- Solving

$$F(m) = 1 - \int_{m_c}^{m} e^{-\kappa(m^*)F(m)} s(m^*)\mathrm{d}m^*$$

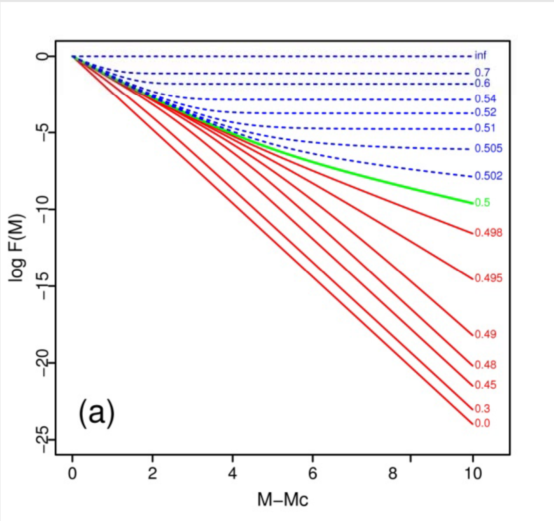
by iterations: for each *m*

$$F_{(k+1)}(m) = 1 - \int_{m_c}^{m} e^{-\kappa(m^*)F_{(k)}(m)} s(m^*)\mathrm{d}m^*$$

$$F_{(0)}(m) = 1$$

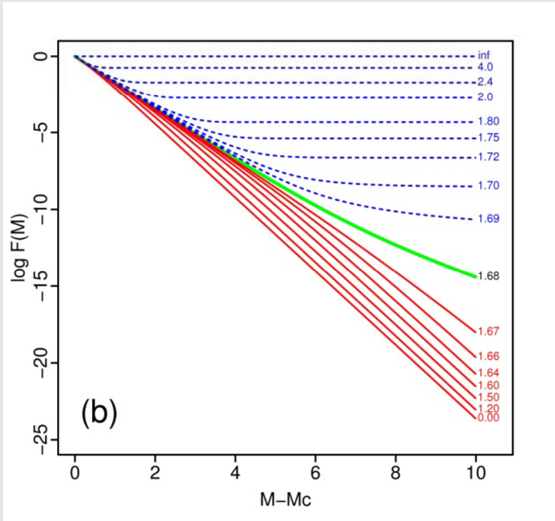
Properties of function $F(m)$

From *Zhuang and Ogata*, ***Physical Review E***, 2006



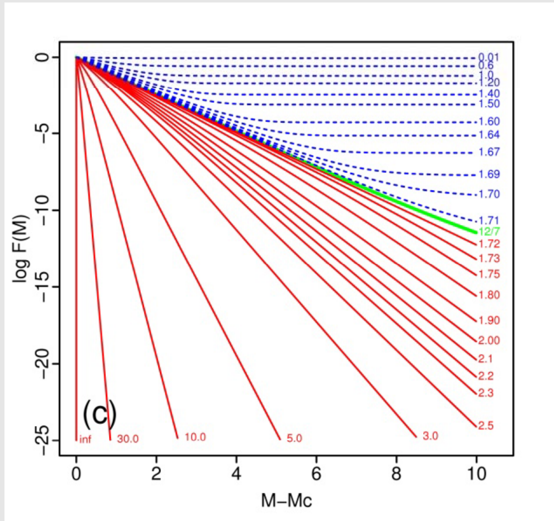
$\alpha = 1.2 \;\; \beta = 2.4$

A varies



$A = 0.3 \;\; \beta = 2.4$

α varies



$A = 0.3 \;\; \alpha = 1.2$

β varies

Distribution of the biggest magnitude in a space-time windows

Notations:

- Space-time window V
- Expected number of clusters in V,

$$\Lambda(V) = \iiint_V \mu(x, y)\mathrm{d}x\mathrm{d}y\mathrm{d}t$$

- Probability mass function for the number of clusters (background events) in V,

$$\Pr\{K = k\} = \frac{\Lambda^k(V)}{k!} e^{-\Lambda(V)}$$

- Magnitude p.d.f. $\quad s(m) = 10^{-b(m-m_c)} b \ln 10$

$$= \beta e^{-\beta(m-m_c)}, \qquad m \geq m_c$$

Distribution of the biggest magnitude in a space-time windows (cont.)

- Cumulative probability distribution function for the biggest magnitude in V

$$
\begin{aligned}
Q(V,m) \\
&= \Pr\{\text{events in all clusters are in } V < m\ \} \\
&= \sum_{n=0}^{\infty} \Pr\{\text{the biggest in each cluster is} < m \mid n \text{ cluster in total}\} \\
&\qquad \times \Pr\{n \text{ cluster in total}\} \\
&= \sum_{n=0}^{\infty} [G(m)]^n \frac{[\Lambda(V)]^n}{n!} \exp[-\Lambda(V)] \\
&= \exp[-\Lambda(V)[1-G(m)]] \\
&= \exp[-\Lambda(V)F(m)]
\end{aligned}
$$

Distribution of the biggest magnitude in a space-time windows (cont.)

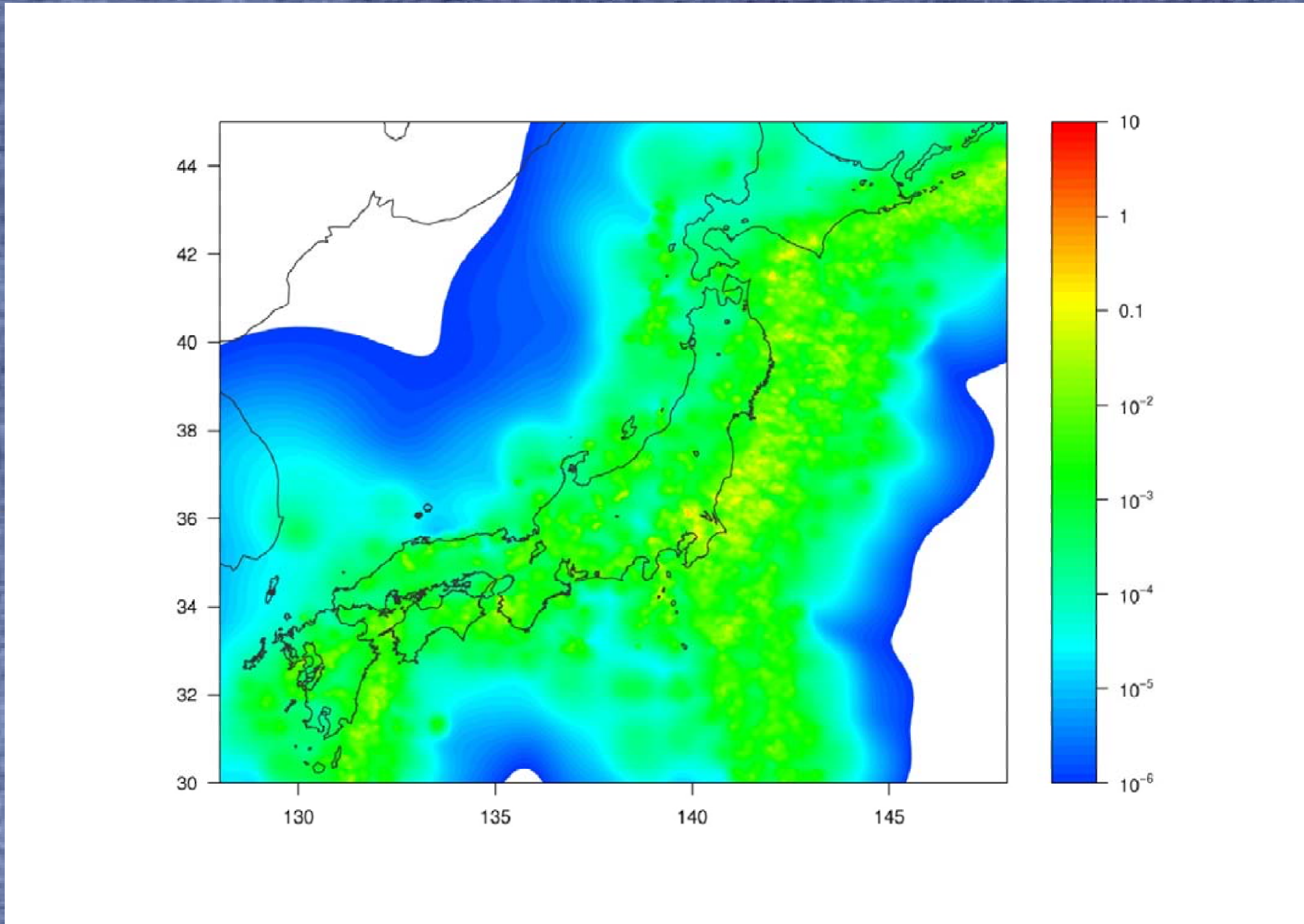
- Cumulative probability distribution function for the biggest magnitude in *V*

$$Q(V,m) = \exp[-\Lambda(V)F(m)]$$

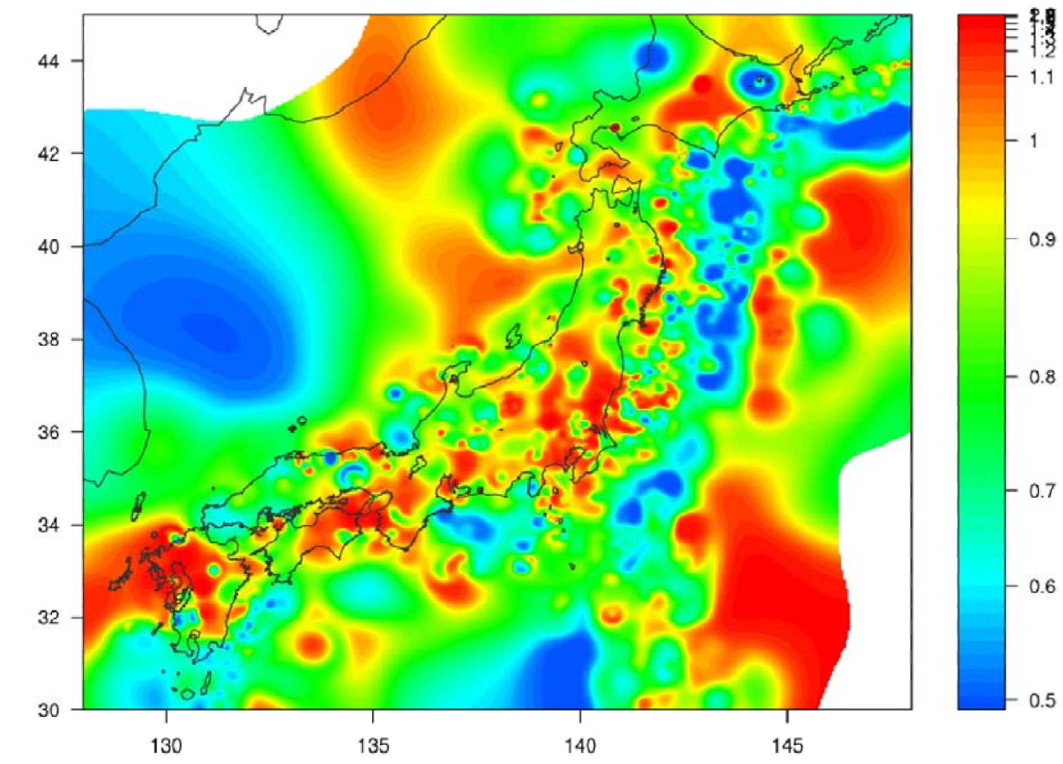
- However, it is still hard to evaluate the number of earthquake at a certain magnitude ranges, but only possible to evaluate the number of earthquake clusters covers this range.

Results: background ratesin the Japan region

(events/day/deg^2)

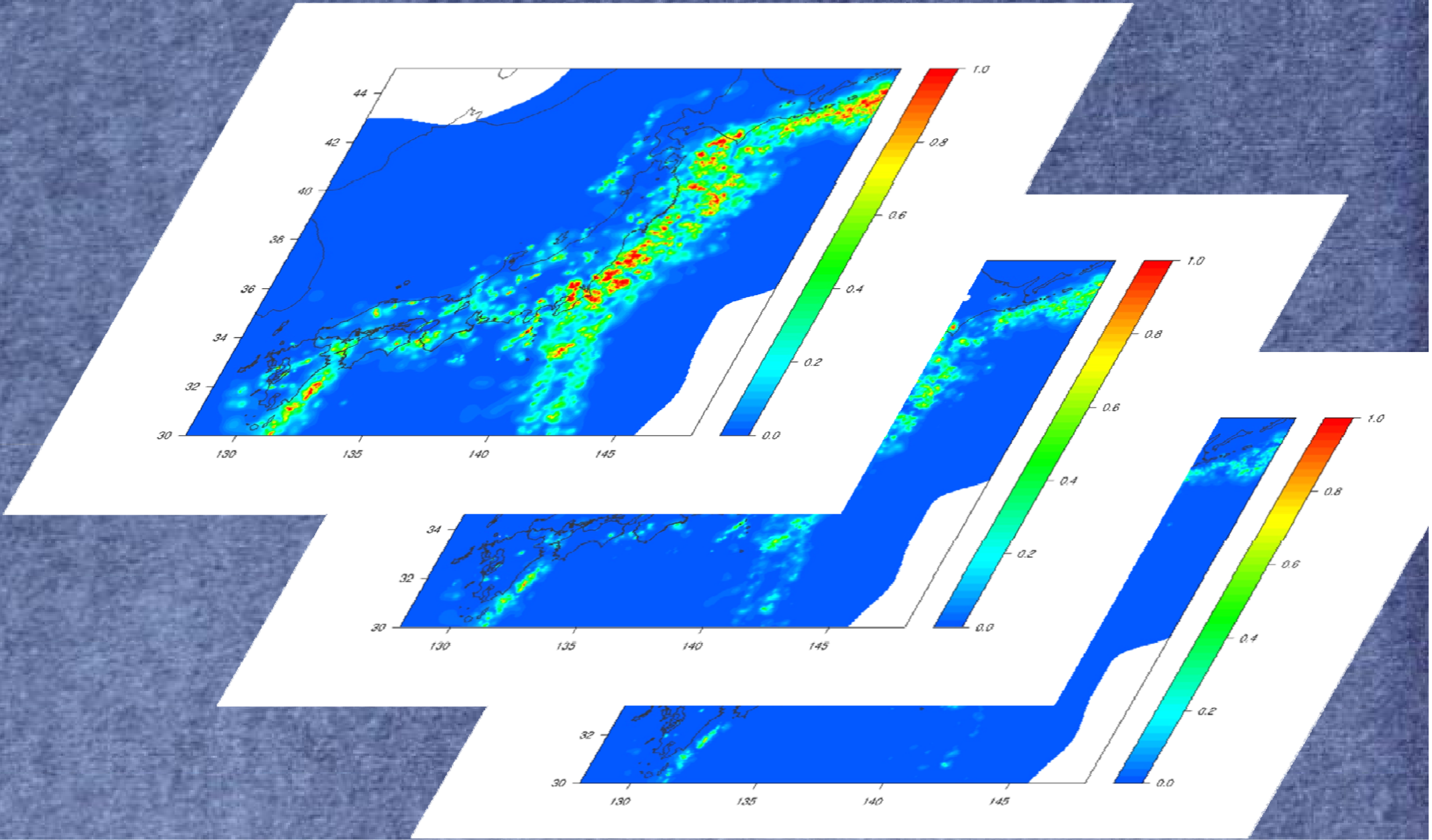


Results: b-value in the Japan region

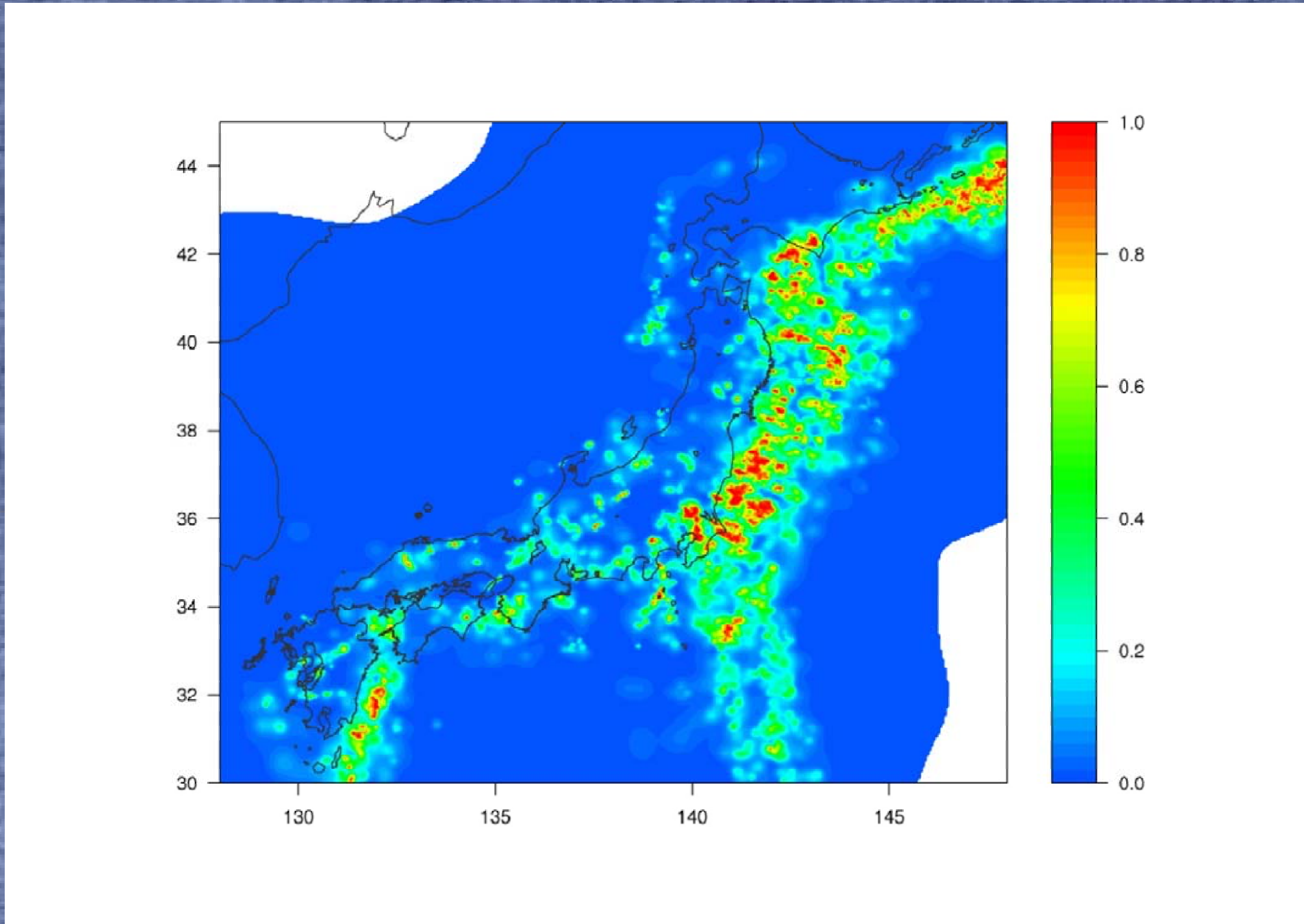


Results: Reference probability (M>=5, 6, 7)

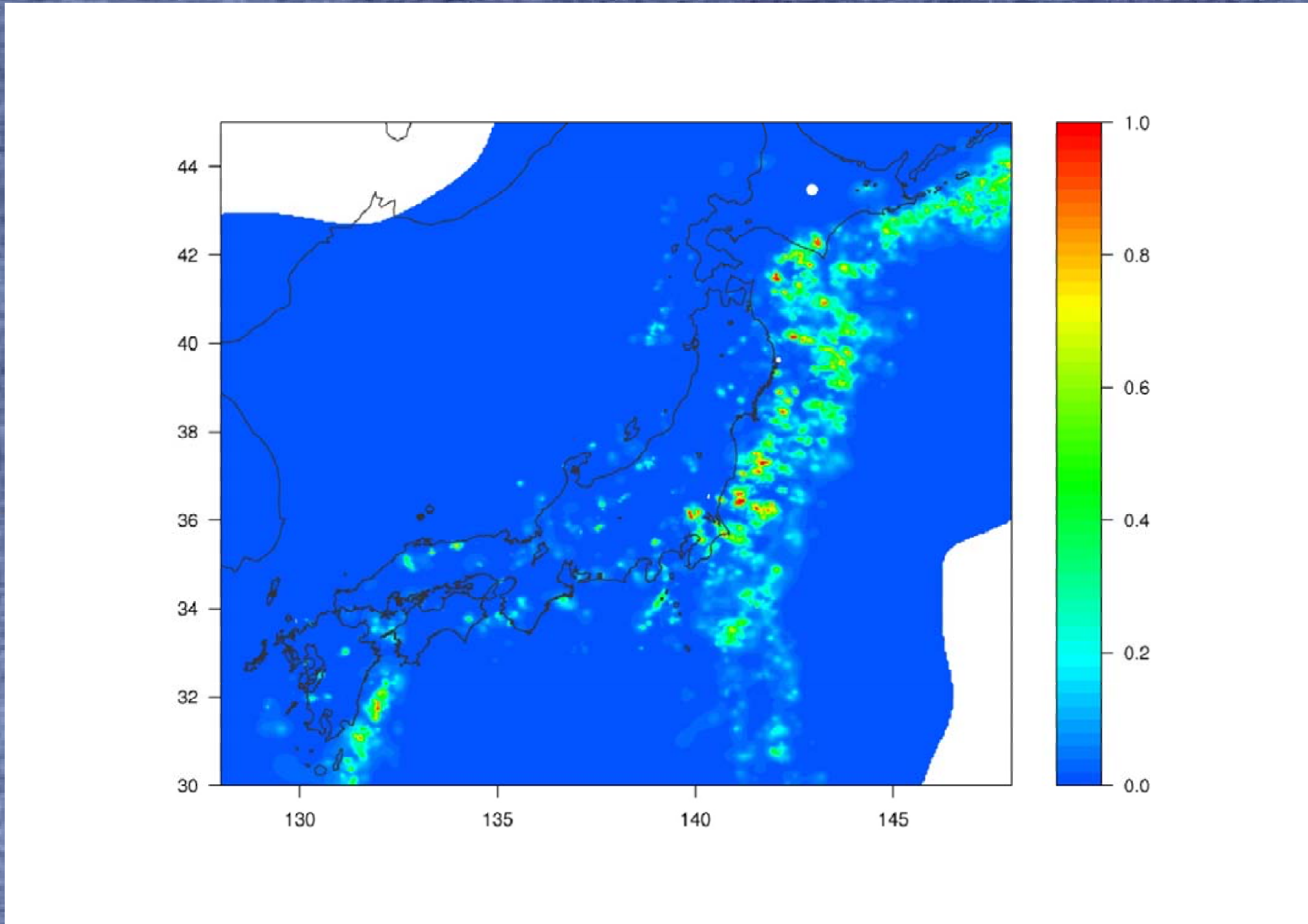
$A = 0.17, \qquad \alpha = 1.45$



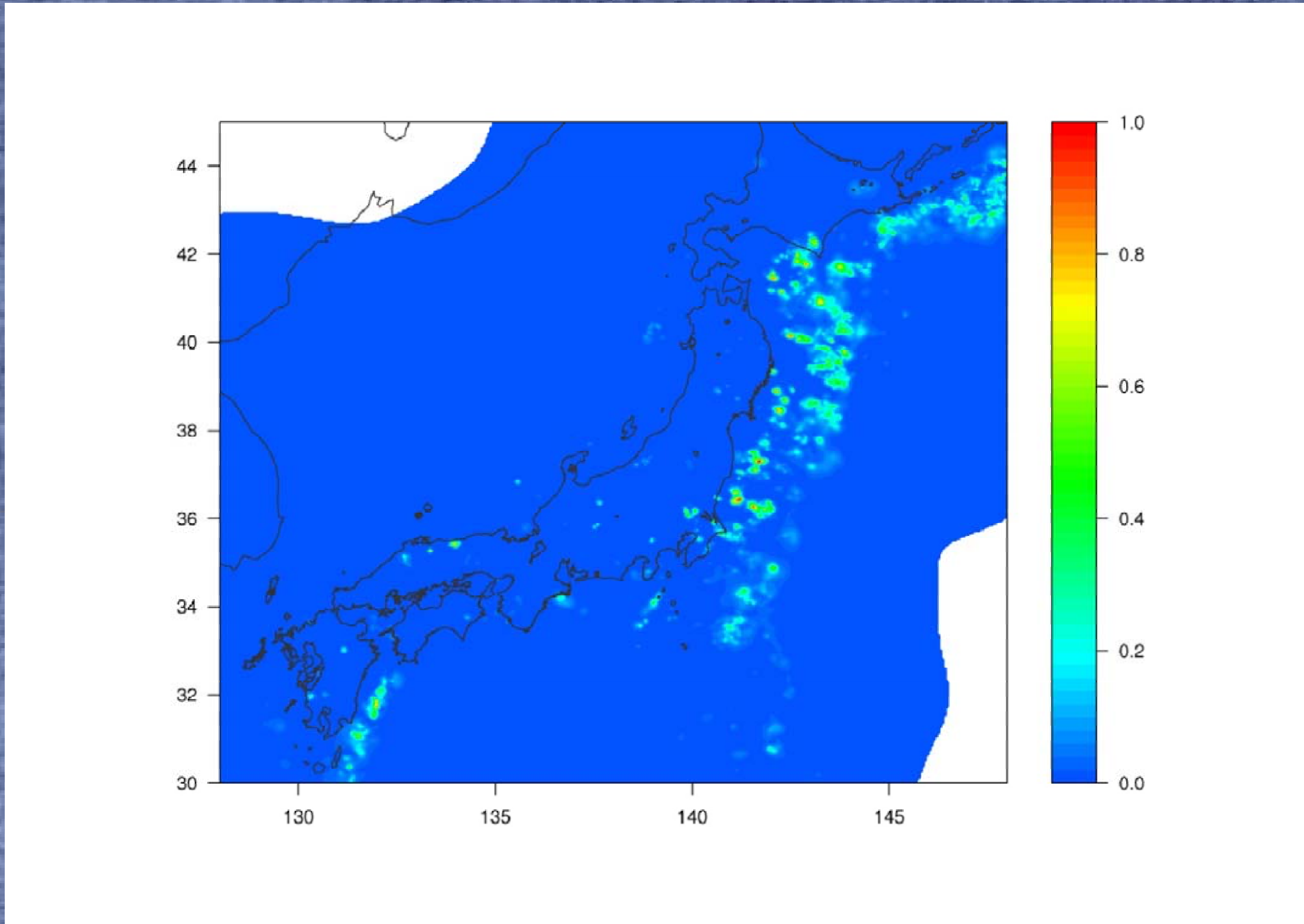
Results: Reference probability Pr{#(M>=5) in 1deg × 1deg × 1yr grid}



Results: Reference probability Pr{#(M>=6) in 1deg×1deg×1yr grid}



Results: Reference probability Pr{#(M>=7) in 1deg × 1deg × 1yr grid}



Future researches

1. Incorporating spatially varying A and alpha values – HIST ETAS model
2. Model testing against other models of smoothed seismicity

Part II
Gambling scoring method

Gambling score

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

Publications of applications of the gambling score

Zhuang J. (2010) *Gambling scores for earthquake predictions and forecasts* . **Geophysical Journal International**. 181: 382-390

Zechar J. D. and Zhuang J., (2010) , *Risk and return: evaluating Reverse Tracing of Precursors earthquake predictions.* **Geophysical Journal International**, 182, 1319-1326.

G. Molchan and L. Romashkova (2010) *Gambling scores in earthquake prediction analysis*, accepted by **Geophysical Journal International**.

Zhuang and Jiang (2010), *Scoring annual earthquake prediction in China*, submitted to Tectonophysics.

Gambling score without a reference model

- Binary occurrence of earthquakes: 1 for grids with earthquake, and 0 for non-earthquake grid
- $X_1, X_2, \cdots, X_N$
- Modeler *i* bets
-

$p_{1,i}, p_{2,i}, \cdots, p_{N,i}$ at least 1 earthquake occurs

$1-p_{1,i}, 1-p_{2,i}, \cdots, 1-p_{N,i}$ no earthquake occurs

Gambling score without a reference model

- Reward principle: the winners take and divide all.
- Reward to Modeller *i*

$$\frac{p_i}{\sum_j p_j}, \text{ if 1 or more event occur;}$$

$$\frac{1-p_i}{\sum_j (1-p_j)}, \text{ if no event occurs;}$$

The model that gets the highest reward is the best.

Gambling score without a reference model

- Numbers of earthquakes on each grids:
- $X_1, X_2, \cdots, X_N$
- Modeler *i* bets
-

$p^{(j)}_{1,i}, p^{(j)}_{2,i}, \cdots, p^{(j)}_{N,i}$ for the probablities that j earthquakes occur, $j = 0, 1, 2, \cdots$

Gambling score without a reference model

- Reward principle: the winners take and divide all.
- Reward to Modeller i for the jth grid

$$\frac{p_{j,i}^{(X_j)}}{\sum_k p_{j,k}^{(X_j)}}, \text{ if } X_j \text{ events occur in } j\text{th grid.}$$

- The model that gets the highest reward is the best.

Gambling score without a reference model

- occurrence of earthquakes:

  $\{(t_i, x_i, m_i): i=1,2,\ldots\}$

-

- Modeler *i* gives a conditional intensity

  $$\lambda_i(t, x, m)$$

Gambling score without a reference model

- Reward principle: the winners take and divide all.
- Reward to Modeller i for the jth grid

$$\frac{L_i(Grid_j)}{\sum_k L_k(Grid_j)},$$

$L_i(Grid_j)$: likelihood of Model i on Grid j.

- The model that gets the highest reward is the best.