Testing annual earthquake predictions in China

Jiancang Zhuang  (Joint work with Changsheng Jiang)

Institute of Statistical Mathematics, Tokyo
Dep. Statist. Sci., the Graduate Univ. for Advanced Studies, Tokyo
The distribution of the alarmed regions and the precursor monitoring stations in the period of 1900 to 2003.
Evaluation of the Annual Reports of Consulting Meeting on Earthquake Tendency, China Earthquake Administration (Zhuang & Jiang, 2012, Tectonophysics)

From Reports of Annual Consulting Meeting on Earthquake Tendency for 1996.

Yellow areas: alarmed regions

Red dots: earthquakes occurred in the prediction period
Available scoring methods for earthquake predictions/forecasts

- R-score, or Hanssen-Kuiper skill score
- ROC curve or Molchan’s error diagram
- Information or entropy score
- RELM/CSEP tests: L-, N-, R-, M-, S-
- Residual based tests, ...
Why gambling score?

(1) To assign high weights for prediction of large earthquakes and to avoid the score being dominated by small events

(2) To deal with “no comment available” (NA) predictions

(3) To deal with irregular (non-gridded) predictions or forecasts

(4) To find out better parts of a model or a prediction algorithm even if its overall performance is worse than some well known models
Gambling score (1)

Risk of the predictor (probability to be failure):

Reference model:

\[ p_0 : \text{prob. that the prediction is correct} \]

predictor’s risk:

\[ 1 - p_0 \text{ (according to the reference model)} \]

Each time the forecaster make a prediction, he bets 1 point of his reputation. If he fails, he loses this point; if he wins, he should be rewarded fairly.
Gambling score (2)

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

Answer: 

\[ G = \frac{(1 - p_0)}{p_0} \]

where \( p_0 \) is the probability given by the reference model that the prediction is correct.

Return for each prediction:

<table>
<thead>
<tr>
<th>Prediction correct?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forecaster bets 1</td>
<td>( G )</td>
<td>-1</td>
</tr>
</tbody>
</table>
Gambling score (3)

*If the baseline model is true, expected return is zero.*

\[
E_\phi[R] = p_0 \times \left[ \frac{1 - p_0}{p_0} \right] + (1 - p_0) \times (-1)
\]

\[= 0.\]

**Choice of the baseline model:**

Should be the ones that are commonly accepted or the best one available, e.g., Poisson models or clustering models (ETAS or the Omori-Utsu formula) for aftershock activities.
How one can get positive gambling score?

\( q \) : The probability that the predictor bets.
\( p_0 \) : Probability that the prediction is true given by the reference model.
\( p^* \) : Probability that the prediction is true given by the true model.

The expected payoff is

\[
E[R] = E \left[ q \frac{1-p_0}{p_0} X + (1-X)q \right] - E[q]
\]

Since \( p_0 \) is usually a conditional expectation of \( p^* \) under some conditions, \( E[R]>0 \) requires \( \text{cov}(q, p^*) > \text{cov}(q, p_0) \).
Evaluation of the Annual Reports of Consulting Meeting on Earthquake Tendency, China Earthquake Administration (Zhuang & Jiang, 2012, Tectonophysics)

From Reports of Annual Consulting Meeting on Earthquake Tendency for 1996.

Yellow areas: alarmed regions

Red dots: earthquakes occurred in the prediction period
Analysis I: Partial Scores

Assumptions

1. Each alarm region is a positive prediction.
2. Non-alarmed regions are regarded as “No-Comment-Available” predictions.
Reference model

- Inhomogeneous Poisson model

Seismicity rate

\[ \lambda(x, y, m) = f(m \mid x, y)\lambda_0(x, y) \]

Magnitude dist.

\[ f(m \mid x, y) = b(x, y)10^{-b(x,y)(m-m_0)} \]

Probability at least 1 earthquake in Region \( S \) between magnitude \( m_1 \) and \( m_2 \) in unit time

\[ p_0(S, m_1, m_2) = 1 - \exp \left[ -T \int_{S} \int_{m_1}^{m_2} f(m \mid x, y)\lambda_0(x, y)dmdxdy \right] \]

Seismicity rate in a region \( S \)

\[ \int_{S} \lambda_0(x, y)dxdy = \frac{\# EQ \text{ in } (S \times T_0) \text{ and } \geq m_0}{\text{length of } T_0} \]

MLE of \( b \)

\[ \hat{b}(S) = \frac{\log 10}{m - m_0 + 0.05} \]
Yearly reputation returns of CEA predictions (blue dots) and their 5%, 50% and 95% percentiles under the assumptions of the reference Poisson models.
Confidence band by simulation

$$\sigma^{(\lambda)} \approx \frac{\hat{\lambda}}{\sqrt{n}} \quad \text{and} \quad \sigma^{(b)} \approx \frac{\hat{b}}{\sqrt{n}} \quad n: \# \text{ of eqs}$$

1. Use $\lambda_i$ and $b_i$ as means, and corresponding stand deviations, generate 2 normal r.v. $A_i$ and $B_i$.
2. Generate the process of earthquake occurrence in the training period and target period according to the Poisson model with rate $A_i$ and G-R $b$-value $B_i$.
3. Re-estimate $\lambda$ and $b$-value according to the generated process in the training period, and evaluate the reference probability and gambling gains.
4. Score the prediction according to the generated process in the target period: if a magnitude is generated within the predicted magnitude range, the score be the gambling gain, otherwise -1.
The cumulative probability function of the total reputation return for the CEA prediction under random conditions
Analysis II: Complete Scores

Assumptions

1. Each alarm region is a positive prediction.
2. Non-alarmed regions are regarded as negative predictions.
Extension to point process models

1. Consider a mark point process equipped with a conditional intensity (stochastic hazard function)

\[ \lambda(t, m) dt dm = \Pr\{N(t, t + dt) \times (m, m + dm) = 1 \text{ observation before } t\} \]

2. At \((t, m)\), the density of bet is \(b(t, m)\), and the payoff ratio for a success is

\[ [1 - \lambda_0(t, m) dt] / [\lambda_0(t, m) dt] = [\lambda_0(t, m) dt]^{-1} \]

\(\lambda_0(t, m)\): conditional intensity of the reference model

3. The total payoff on the interval \([0, T]\) is

\[ Q = \int_{0}^{T} \left[ \frac{b(t, m)}{\lambda_0(t, m)} N(dt dm) - b(t) dt dm \right] \]

with an expectation

\[ \mathbb{E}Q = \int_{0}^{T} b(t, m) \left( \frac{\lambda^*(t, m)}{\lambda_0(t, m)} - 1 \right) dt dm \]

\(\lambda^*(t, m)\): conditional intensity of the true model
Extension to point process models

1. Designing the bet function

\[ \beta(t, x, y, m) = \begin{cases} 1, & \text{if } (t, x, y, m) \text{ is in the alarm region} \\ -1, & \text{otherwise} \end{cases} \]

2. The return is

\[ R = \sum_{i \text{ in alarm region}} \frac{1}{\lambda_0(t_i, x_i, y_i, m_i)} - \sum_{i \text{ in non-alarm region}} \frac{1}{\lambda_0(t_i, x_i, y_i, m_i)} - \text{AREA (alarmed region)} + \text{AREA (non-alarmed regions)} \]

\( \lambda_0(t, m) \) : conditional intensity of the reference model
The distributions of the alarmed regions and the precursor monitoring stations in the period of 1900 to 2003.
Betting magnitude ranges (Ms)

\[ E: 5.0-8.5 \quad W: 5.5-9.0 \quad E: 5.0-8.0 \quad W: 5.5-8.5 \]

\[ R = \sum_{i \text{ in alarm region}} \frac{1}{\lambda_0(t_i, x_i, y_i, m_i)} - \sum_{i \text{ in non-alarm region}} \frac{1}{\lambda_0(t_i, x_i, y_i, m_i)} - \text{AREA (alarmed region)} + \text{AREA (non-alarmed regions)} \]
Figure 5: An illustration of the relation between the R scores for the CEA annual predictions and the inhomogeneous Poisson model:

(a) Predictions are made purely on knowledge of seismicity levels.
(b) Predictions made on precursory information, but not much more than information in seismicity levels.
(c) a method for improving CEA annual predictions: use information of seismicity levels.
To end this talk...

1. **Conclusion:** The performance of CEA annual predictions are evaluated. The partial score shows the inclusion of precursory information in those predictions, while the complete score shows the information of seismicity level is ignored.

2. **Suggestion:** Extending CSEP to an evaluation platform for predictions, not only for models.