Short-term earthquake forecast model using foreshock discrimination and aftershock decay

The Institute of Statistical Mathematics   Shunichi Nomura, Yosihiko Ogata

1. Summary

Foreshock discrimination is one of promising ways for short-time forecast of large mainshocks. We propose a statistical method to evaluate foreshock probabilities for growing earthquake clusters. By incorporating magnitude-frequency distribution of mainshocks, our method also provides occurrence probabilities of mainshocks of arbitrary target magnitudes. To validate our method, we divided the catalog into training and validation datasets and tested the predictive performance by cross-classified contingency table. Moreover, we evaluate the evolution of foreshock probabilities for the 2016 Kumamoto (M7.3) earthquake sequence.

2. Method

First we constructed earthquake clusters to define foreshocks, mainshocks and aftershocks. In this study, we analyzed the JMA catalog of M $\geq$ 4, in the region 128° E-148°E, 30° N-46° N, shallower than 100km depth and observed during period 1926-2016. We applied the 'Single-link' clustering (SLC) algorithm of Frohlich & Davis (1990) to the dataset. Specifically, earthquakes of less space–time distance than 0.3° (33.33 km or 30 days) are connected as belonging to a same cluster. The distance of SLC is given by $\sqrt{(\Delta d)^2 + (c\Delta t)^2}$ with $c = 0.01°$ /day, where Δd is the ordinary epicentre separation and Δt is the time difference. After clusters are composed, we define the largest event in a cluster as a mainshock and sub-events before and after the mainshock as foreshocks and aftershocks, respectively.

Ogata et al. (1996) revealed some statistics within a earthquake cluster are useful to discriminate foreshocks. Here we considered following feature statistics similar to those used in Ogata et al. (1996) and calculated them at each time that a cluster grows by adding a new event:

- N : Number of events in the cluster
- M_1, M_2 : 2 largest magnitudes of the cluster
- T : Time duration of the cluster (day)
- D : Mean pairwise epicentral distance in the cluster (km)
- (X, Y) : Mean longitude and latitude of epicenters (degree)

We evaluate probability that a mainshock of magnitude M_{main} occur after the present cluster with feature statistics $(N, M_1, M_2, T, D, X, Y)$ by following logistic regression:

$$\mathrm{logit}\{P(\mathrm{foreshock}|N, M_1, M_2, T, D, X, Y)\}$$
$$= f_1(N, M_1, M_2) + f_2(N, M_1, T) + f_3(N, M_1, D) + g(X, Y) + \varepsilon_{\mathrm{seq}}.$$

where logit $p = \log\{p/(1-p)\}$. f_1, f_2, f_3 are nonlinear 3-D B-spline functions and g is a thin-plate spline function. The last term $\varepsilon_{\mathrm{seq}}$ is random effects with respect to the composed clusters which represents other factors not included in our model.