Statistical features and cluster-based discrimination of foreshocks

Shunichi Nomura, Yosihiko Ogata

The Institute of Statistical Mathematics

1. Introduction

Foreshocks are promising clues for short-term forecasting of large mainshocks. Many studies have addressed certain features of foreshocks and predictability of mainshocks by foreshock detection. Ogata et al. (1995) composed and classified seismic clusters into foreshock, swarm, and mainshock-aftershock types. They discussed different trends in spatio-temporal distances and magnitude increments among those cluster types. Ogata et al. (1996) proposed a logistic regression model to evaluate the probabilities that seismic clusters will be the foreshock type or other clusters. This model can be applied to all earthquake clusters that may be foreshocks by defining the odds functions over all the feature spaces.

In this talk, we introduce a framework of cluster-based discrimination of foreshocks by Ogata et al. (1996). Also, we show the result of prospective experiment of foreshock discrimination in Ogata and Katsura (2012), which obtained as good predictive performance as that in the original paper. Furthermore, we discuss some future direction to improve the method by Ogata et al. (1996) in terms of other efficient features, time horizon, spatial ranges and mainshock magnitudes for forecast.

2. Methods

In this section, we introduce the procedure of probabilistic foreshock discrimination by Ogata et al. (1996). The Japan Meteorological Agency (JMA) catalog of $M \geq 4$ in the region 128–148°E, 30–46°N as observed from January 1, 1926 to October 31, 2017 at a depth shallower than 100 km is analyzed through this talk.

To define mainshocks and their sub-events, earthquake clusters should be constructed from that catalog. The seismic clusters were constructed by the so-called single-link clustering (SLC) algorithm of Frohlich and Davis (1990). Specifically, earthquake pairs whose spatio–temporal distances were less than 0.3° (33.33 km and 30 days) were linked as belonging to the same cluster. The spatio–temporal distance of SLC is defined by $\sqrt{(\Delta d)^2 + (c\Delta t)^2}$, where Δd is the epicentral distance in degrees, and Δt is the difference of occurrence times in a day. In addition, to separate clusters between earthquakes in the shallow crustal zones and deep plate subduction zones in Japan, we only linked earthquake pairs whose depth difference was less than 70 km. The spatio–temporal distance threshold of 0.3° and 30 days, was determined in Ogata et al. (1995) to be consistent with the clusters constructed by a window-based algorithm which depends on the magnitudes of the mainshocks.

When a seismic cluster ceases evolving after 30 days from the last event, its mainshock is defined by the largest event of the cluster. Then, in this talk, other earthquakes in the cluster are classified into preshocks before the mainshock and aftershocks after the mainshock. Ogata et al. (1996) defines a foreshock-type cluster by the clusters including preshocks whose largest magnitudes are less than the mainshock magnitude by 0.45 or more.

Ogata et al. (1995) revealed some feature statistics of earthquake clusters that are effective for foreshock discrimination. Ogata et al. (1995) considered time differences t_{ij}, epicenter separations r_{ij} and magnitude differences m_{ij} between the i-th and j-th events within a cluster. After standardizing those features t_{ij}, r_{ij} and m_{ij} into approximately uniformly distributed forms τ_{ij}, ρ_{ij} and γ_{ij}, Ogata et al. (1996) evaluated the probability $p_{c|n}$ that a cluster is foreshock-type by the following logistic model:

$$\text{logit}\, p_{c|n} = \text{logit}\, \mu(x_1,y_1) + \frac{1}{\#\{i<j\leq n\}}\sum_{i<j\leq n} (\mu_0 + f_1(\tau_{ij}) + f_2(\rho_{ij}) + f_3(\gamma_{ij}))$$

where logit $p = \log\{p/(1-p)\}$ is a logistic function. $\mu(x_1,y_1)$ indicates baseline probability dependent on the longitude and latitude of the first event in the cluster shown in Fig. 1. The factor $\#\{i < j \leq n\}$ is the number of pairs in the first n members in the cluster c. f_1, f_2 and f_3 are odds functions of foreshock-type from other types represented by third-order polynomial functions of τ_{ij}, ρ_{ij} and γ_{ij}, respectively.

3. Results

Ogata and Katsura (2012) validated the predictive performance of the model by Ogata et al. (1996) in their catalog during 1994-2011 after its publication. Fig. 2 indicates that relative frequencies of the actual foreshocks are basically larger than those of the other cluster types if the predicted foreshock probabilities. are larger than the average (7.9%) shown in a vertical line and vice versa. The evaluated probability varied in the range between 0 and 80% and most of the clusters which yield high probabilities were actually foreshock-type clusters.

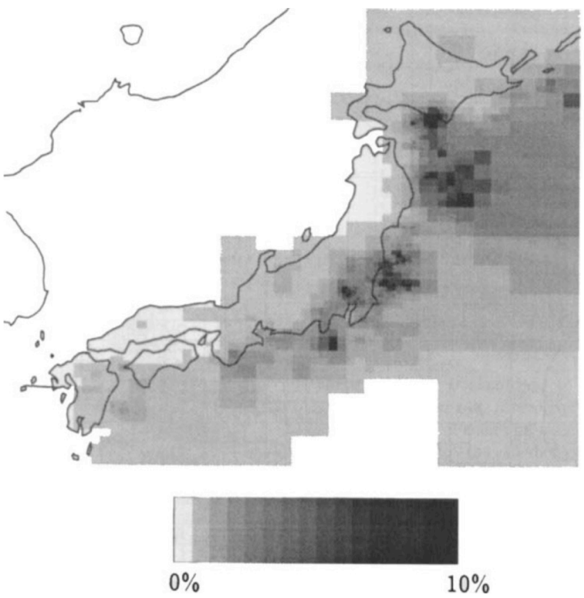

Fig. 1. Estimates of baseline probability function $\mu(x_1,y_1)$ of longitude x_1 and latitude y_1 of the first event in a cluster during the period 1926-1993 (Ogata et al., 1996).

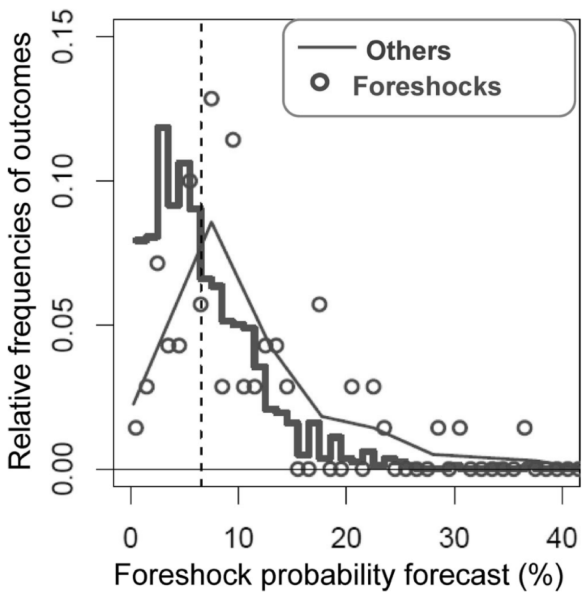

Fig. 2. Relative frequency of outcomes against foreshock probability forecasts for the multiplet clusters during the period 1994-2011 (Ogata and Katsura, 2012).