Properties of the probability distribution associated with the largest event in an earthquake cluster

Jiancang Zhuang and Yosihiko Ogata
Institute of Statistical Mathematics, 4-6-7 Minami Azabu, Minato-Ku, Tokyo 106-8659, Japan

Corresponding Author: Jiancang Zhuang, zhuangjc@ism.ac.jp

Abstract: The space-time epidemic-type aftershock sequence (ETAS) model is a stochastic branching process in which earthquake activity is classified into background and clustering components and each earthquake triggers other earthquakes independently according to certain rules. This paper gives the probability distributions associated with the largest event in a cluster and their limit properties for all three cases when the process is subcritical, critical and supercritical, and uses them to evaluate the probability of an earthquake to be a foreshocks, and to derive magnitude distributions of foreshocks and non-foreshock earthquakes. To verify these theoretical results, the JMA (Japan Meteorological Agency) earthquake catalog is analyzed, and we found no discrepancy between the analytic results and the inversion results by using the stochastic reconstruction method. We found that the proportion of foreshocks in background events is around 8%.

1. Introduction
The foreshock has been one of the most important topics in the research of seismicity pattern and earthquake prediction for decades. To understand foreshocks, we need a good modelling of earthquake clustering. In seismicity studies, the ETAS models are generally adopted (see e.g., Ogata, 1998). These models assume that the seismicity can be divided into a background component and a clustering component, and that each event, no matter it is from the background or it is directly triggered by another event, triggers its own offspring according to some general rules.

2. Definition of the ETAS model
In the ETAS model, the time varying seismicity rate (mathematically termed as conditional intensity) takes the form of

$$\lambda(t, x, y, m) = s(m) \lambda(t, x, y) = s(m) \left[ \mu(x, y) + \sum_{i:t_i<t} \xi(t, x, y; t_i, x_i, y_i, m_i) \right],$$

where $\mu(x, y)$ represents the background seismicity rate and $\xi(t, x, y; t_i, x_i, y_i, m_i)$ is the contribution to seismicity rate by the $i$th event occurring previously, in the form of

$$\xi(t, x, y; t_i, x_i, y_i, m_i) = k(m_i) g(t-t_i) f(x-x_i, y-y_i; m_i).$$

In (2),

$$k(m) = A e^{\alpha (m-m_c)}, \quad m \geq m_c,$$

is the mean number of direct offspring from an event sized $m$, and

$$g(t) = \frac{p-1}{c} \left(1 + \frac{t}{c}\right)^{-p}, \quad t > 0,$$

and

$$f(x, y; m) = \frac{q-1}{\pi D \exp[\gamma (m-m_c)]} \left(1 + \frac{x^2 + y^2}{D \exp[\gamma (m-m_c)]}\right)^{-q}.$$

are the p.d.fs for the occurrence times and locations of direct offspring, respectively.

3. Theoretical distributions associated with the largest magnitude of all the descendants from a given event
The probability of an event of magnitude $m$ to have no offspring greater than $m'$ can be derived from the model analytically (Zhuang and Ogata, 2004), i.e.,

$$\zeta(m, m') = \Pr\{\text{an event of magnitude } m \text{ has no offspring greater than } m'\}$$

greater than $m' \ | \ m \text{ has } n \text{ direct offspring}$$

$$= \exp \left\{ -\kappa(m) \left[ 1 - \int_{m_c}^{m'} s(m^*) \zeta(m^*, m') \, dm^* \right] \right\}. \quad (6)$$
It is evident that $\zeta(m, m')$ has the form of
\[
\zeta(m, m') = \exp[-\kappa(m) F(m')].
\] (7)
where
\[
F(m') = 1 - \int_{m_c}^{m'} s(m') \exp[-\kappa(m') F(m')] dm'
\]
\[
= 1 - \int_{m_c}^{m'} \beta e^{-\beta(m'-m_c)} \exp[-A e^{\alpha(m'-m_c)} F(m')] dm'
\]
\[
= 1 - \frac{\beta}{\alpha} [A F(m')]^{\frac{\beta}{\alpha}} \int_{A F(m')}^{1} e^{-\frac{\beta}{\alpha} - 1} e^{-u} du
\] (8)
represents the probability that the largest earthquake in an arbitrary cluster, including the initial event and all its descendants, to be greater than $m'$. Based on the analysis with generating functions, Saichev and Sornette (2005) give similar forms of the above equations. But they do not discuss the case when the process is supercritical. In this article, we are going to discuss the properties of $\zeta$ under all the three cases.

The extinct probability $P_c(m)$ can be derived as
\[
P_c(m) = \Pr\{\text{The family tree from an event of magnitude } m \text{ extinguishes}\}
\]
\[
= \exp\left[-\kappa(m) \left(1 - \int_{m_c}^{+\infty} s(m') P_c(m') dm'\right)\right].
\] (9)
Substitute $P_c(m) = \exp[-C \kappa(m)]$ into (9), we have
\[
C = 1 - \int_{m_c}^{+\infty} s(m') \exp[-C \kappa(m')] dm'.
\] (10)
Function $F(m)$ is closely related to $P_c(m)$ by
\[
\lim_{m \to +\infty} F(m) = C = -\frac{\log P_c(m)}{\kappa(m)}.
\] (11)
Thus, when the process is supercritical, $C > 0$ yields
\[
\lim_{m' \to +\infty} \zeta(m, m') = \lim_{m \to +\infty} e^{-\kappa(m) F(m')} = e^{-C \kappa(m)} = P_c(m').
\] (12)
For the subcritical case, which requires that $\beta > \alpha$ and $\varrho = \alpha/\beta - \alpha < 1$, it is easy to see that $\zeta(m, m') \to 1$ when $m' \to +\infty$ because $C = 0$. In this case,
\[
\lim_{m \to +\infty} \frac{F(m)}{s(m)} = \frac{1}{\beta(1 - \varrho)}.
\] (13)
Furthermore, in subcritical case, according to (7), for a fixed $m$, when $m'$ is large enough,
\[
\zeta(m, m') \approx \exp\left[-\frac{1}{\beta(1 - \varrho)} s(m') \kappa(m)\right] = \exp\left[-\frac{A'}{\beta(1 - \varrho)} e^{\alpha m - \beta m'}\right],
\] (14)
where $A' = Ae^{(\beta - \alpha)m_c}$. For the critical case,
\[
F(m) \approx e^{-\alpha (m - m_c)} \propto [\kappa(m)]^{-1},
\] (15)
and
\[
\zeta(m, m') \approx \exp\left[-A e^{\alpha (m' - m)}\right].
\] (16)
In conventional studies, foreshocks are defined as non-aftershock earthquakes that are followed by another larger earthquake occurring nearby. Here we define a foreshock by a background event that has at least one offspring, direct or indirect, with a larger magnitude. Here we refer to Zhuang et al (2004) on how to determine whether an earthquake is a background event or a triggered events or a foreshock by using the
stochastic declustering method. From (6), the probability that an event of magnitude $m$ is a foreshock is $H(m) = 1 - \zeta(m, m)$.

## 4. Data analysis and results

The dataset used in this study is the JMA catalogue in a range of longitude $121^\circ - 155^\circ E$, latitude $21^\circ - 48^\circ N$, depth 0~100 km, time 1926/January/1~1999/December/31 and magnitude $\geq M_J 4.2$. We choose a polygon as shown in Figure 1 as the target region and a time period of 10,000~26,814 days after 1926/Jan/1, and same depth and magnitude ranges as the whole data set for model fitting. The events outside of this study space-time range are used as complementary events for calculating the boundary effect.

The model parameters estimated by maximizing the likelihood function are $\hat{A} = 0.1692$, $\hat{c} = 0.0189 \, \text{day}$, $\hat{\alpha} = 1.5014$, $\hat{\beta} = 1.1181$, $\hat{D}^2 = 0.0008640$, $\hat{q} = 1.9107$, $\hat{\gamma} = 1.0761$ and $\hat{\beta} = 1.9585$.

There are two ways to evaluate $\zeta(m, m')$: one is based on solving (7) with the parameters from the fitting results; the other is based on stochastic reconstruction. The theoretical $\zeta(m, m')$ and the reconstruction $\hat{\zeta}(m, m')$ for the JMA are shown in Figure 2. In calculating the theoretical one in the left panel, we use the empirical magnitude distribution for the events falling in the study space-time range as $s(m)$. The overall impression of these two images is their similarity to each other, even though the theoretical one is more smoother than the reconstructed one.

After we plot $H(m)$ and $\hat{H}(m)$ in Figure 3, we can see that, even if there is no big discrepancy between these two functions, the probability for an event to produce at least 1 larger descendant is about 15%. Such a high probability may be explained in the following way. According to Zhuang et al. (2004), the characteristics of the background ground events are different between the background events and their descendants. For example, the magnitude distributions of background and triggered events are different, denoted here by $s_b(m)$ and $s_c(m)$. Also a triggered event triggers more direct offspring in average than a background event of the same magnitude. Here we denote the productivity from a background event and a triggered event by
Figure 3: Foreshock probabilities of foreshocks in the JMA catalogue. Circles, triangles and crosses represent the reconstructed $\tilde{H}(m)$, $\tilde{H}_c(m)$ and $\tilde{H}_b(m)$ for all the events, background events and triggered events, respectively. The black, red and orange solid lines are theoretical results for all the events, background events and triggered events, respectively.

For a triggered event,

$$\zeta_c(m, m') = \exp \left\{ -\kappa_c(m) \left[ 1 - \int_{m_c}^{m'} s_c(m^*) \zeta_c(m^*, m') \, dm^* \right] \right\}.$$  \hspace{1cm} (17)

For a background event of magnitude $m$,

$$\zeta_b(m, m') = \exp \left\{ -\kappa_b(m) \left[ 1 - \int_{m_c}^{m'} s_c(m^*) \zeta_c(m^*, m') \, dm^* \right] \right\}.$$  \hspace{1cm} (18)

Using the stochastic reconstruction techniques introduced by Zhuang et al (2004), we reconstruct $\kappa_b$, $\kappa_c$ and $s_c$, and then use them to calculate $\tilde{H}_b$ and $\tilde{H}_c$, as shown in Figure 3. As we can see, three theoretical curves are still close to the reconstructed ones after the above corrections. Basically, the probability that a background earthquake of magnitude about 4.2 to 5 triggers a larger earthquake is around 7%.

5. Conclusions

In this paper, based on the space-time ETAS model, we obtained the key equation for the probability for the magnitude of the largest descendant from a given ancestor, as $\zeta(m, m')$ in (6) and their asymptotic properties. We define foreshocks naturally by the background events that have at least one larger descendant. The probability that a background event is a foreshock can easily obtained.

We also analyze the JMA catalog to verify our theories. To obtained $\zeta(m, m')$, two methods are used: one is directly computed from the integral equations, (7) and (8), and the other is by using stochastic reconstruction. Because of the different behaviors in triggering offspring between background events and triggered events, we evaluate the corresponding $\zeta$-function for both background events and triggered events. For a background event of a magnitude from 4.2 to 6.0, the probability that it is a foreshock is about 8%, which is very close to the results obtained by using a conventional foreshock definition and a conventional declustering method.

6. Acknowledgement

Helpful discussions with Annemarie Christophsen, Kazuyoshi Nanjo, Martha Savage, Euan Smith, Didier Sornette are David Vere-Jones are gratefully acknowledged.

7. Main references


