Long-term earthquake forecasts based on the epidemic-type aftershock sequence (ETAS) model for short-term clustering

Jiancang Zhuang

Institute of Statistical Mathematics and Department of Statistical Science, the Graduate University for Advanced Studies, Tokyo, Japan

Abstract

Based on the ETAS (epidemic-type aftershock sequence) model, which is used for describing the features of short-term clustering of earthquake occurrence, this paper presents some theories and techniques related to evaluating the probability distribution of the maximum magnitude in a given space-time window, where the Gutenberg-Richter law for earthquake magnitude distribution cannot be directly applied. It is seen that the distribution of the maximum magnitude in a given space-time volume is determined in the long-term by the background seismicity rate and the magnitude distribution of the largest events in each earthquake cluster. The techniques introduced were applied to the seismicity in the Japan region in the period from 1926 to 2009. It was found that the regions most likely to have big earthquakes are along the Tohoku (northeastern Japan) Arc and the Kuril Arc, both with much higher probabilities than the offshore Nankai and Tokai regions.

Introduction

The Poisson model, or other renewal models, are often used to evaluate the long-term risk of large earthquakes in a particular region. However, clusters of aftershocks complicate the statistical analysis that evaluates the background seismic activity. The clustering features differ from place to place. To forecast the location of the large earthquakes, it is necessary to analyze the background seismicity, for which removal of temporal cluster members is considered to be of central importance. Both the background rate and the clustering structures are important to earthquake hazard estimation, because the background rate tells us the potential risk of the occurrence of an earthquake cluster, and the clustering structure determines the size of the earthquake cluster and the magnitude of the biggest event in a cluster.

In the literature of seismology, people traditionally use some window based methods[1-3] or link-based methods[4-6] to decluster the catalog or to identify earthquake clusters. The parameters in these declustering rules depend to a large degree on the intuitions and experience of researchers, sometimes even colored by the user's subjective impression of the declustered output. The space-time epidemic-type aftershock sequence (ETAS) model describes the features of earthquake clusters. It has become a standard model for testing hypotheses and a starting point for short-term earthquake forecasts.[7-12] Based on this, Zhuang *et al.*[13] proposed the stochastic declustering method to separate seismicity into a background component and a clustering component in a probability manner, following the initial ideas by Kagan and Knopoff.[14] This method can also produce stochastic versions of declustered catalogs similar to the output of the conventional earthquake declustering methods.

The aim of this article is to show how to use the background rate and the clustering structure to estimate the probability of future large earthquakes in a particular region. A brief description of the ETAS model is given and its use to separate the background rate from overall seismicity is discussed. Problems in estimating the background rate and the parameters of the clustering structure are examined. Formulas are then elaborated to evaluate the probabilities related to the magnitude of the largest descendant from an initial event with a given magnitude. Based on these probabilities, the probability density of the magnitude distribution of the largest earthquake in an arbitrary earthquake cluster is calculated. In order to illustrate this, we use the approach set out in the Japan Meteorological Agency (JMA) catalog to estimate earthquake risk in the Japan region.

Definition of the epidemic-type aftershock sequence model

In the ETAS model, there is no difference in triggering seismicity among foreshocks, mainshocks and aftershocks. Each event triggers its own offspring independently according to some probability rules. In brief, the time-varying seismicity rate (mathematically termed as conditional intensity) of this model takes the form of:

OPEN ACCESS

$$\lambda(t,x,y) = \mu(x,y) + \sum_{i:t_i<t} \kappa(m_i)\, g(t-t_i)\, f(x-x_i, y-y_i, m_i), \quad (1)$$

where
a) $\mu(x,y)$ represents the background seismicity rate; b) the term $\kappa(m_i)$ $g(t-t_i)\, f(x-x_i, y-y_i, m_i)$ is the contribution to seismicity rate by the ith event occurring previously, and

$$\kappa(m) = A\, e^{\alpha(m-m_c)}, \quad m \geq m_c, \quad (2)$$

is the mean number of direct offspring from an event sized m, m_c being the magnitude threshold, and

$$g(t) = \frac{p-1}{c}\left(1 + \frac{t}{c}\right)^{-p}, \quad t > 0, \quad (3)$$

and

$$f(x,y;m) = \frac{q-1}{\pi D\, e^{\gamma(m-m_c)}}\left(1 + \frac{x^2+y^2}{D\, e^{\gamma(m-m_c)}}\right)^{-q}. \quad (4)$$

represents the probability density functions for the occurrence times and locations of direct offspring, respectively. In equations (2), (3) and (4), A, α, c, p, D, q and γ are constant parameters.

Equation (3) is the probability density form of the Omori-Utsu formula.[15-17] The form of equation (4) is based on the analysis in Zhuang *et al.*[18] and Ogata and Zhuang,[19] which are improved versions of those in Ogata.[20] Many other forms can also be found in papers published over last 20 years.[8,21]

Model estimation

To forecast seismicity by using the model specified by equation (1), we need to solve the following technical problems: i) estimating background seismicity rate; and ii) estimating the model parameters, (A, α, c, p, D, q, γ).

Maximum likelihood estimates

For an observation of the process, recorded as a list in the form of $\{(t_i, x_i, y_i, m_i): i=1, \cdots, N\}$, from a spatial region S and a time interval $[0,T]$, the likelihood has the standard form.[22]

$$\ln L = \sum_{i:(t_i,x_i,y_i)\in S\times[0,T]} \ln\lambda(t_i,x_i,y_i) - \int_0^T \iint_S \lambda(t,x,y)\, dx\, dy\, dt \quad (5)$$

If the background seismicity rate $\mu(x,y)$ is known, the model parameters, (A, α, c, p, D, q, γ), can be estimated through maximizing the likelihood function. In computation, the Davidon-Fletch-Powell method[23] is used to minimize $-\ln L$.

Estimating background seismicity

Many different approaches to assess background seismicity rate have also been used: i) using a rate proportional to total seismicity rate of all events or only of the big events in the catalog;[20,24] ii) using a declustering method to decluster the catalog and use the total rate in the declustered catalog as background rate;[20,25,26] iii) weighting each event by an estimated probability that it is a background event;[10,13] and iv) using the method introduced by Ogata,[27] which is a Bayesian smoothness prior on a tessellation grids to estimate the spatial variation of the background and the model parameter at the same time.

In this study, the third method is used because it is relatively simple and gives an unbiased estimate of the intensity function. First of all, once the conditional intensity function is estimated, the contribution of the spontaneous seismicity rate at the occurrence of the ith event,

$$\varphi_i = \frac{\mu(x_i,y_i)}{\lambda(t_i,x_i,y_i)}, \quad (6)$$

provides an estimate of the probability that the ith event is a background event, since, if we keep the ith event with probability φ_i for all the events in the process, we can obtain a process with the occurrence rate of $\mu(x, y)$.[28,29] Thus, the background intensity can be estimated by applying some smoothing techniques to the background catalog. Or, more efficiently, we can directly estimate the average by weighting all the events with their corresponding background probabilities. In this study, the variable kernel estimates[13] were chosen, *i.e.*

$$\hat{\mu}(x,y) = \frac{1}{T}\sum_i \varphi_i Z_{d_i}(x-x_i, y-y_i), \quad (7)$$

where i runs over all of the events in the whole process, T is the length of the time period of the process, and Z_d is the Gaussian Kernel function with a bandwidth d. The variable d_j represents the varying bandwidth calculated for each event j in the following way. Given a suitable integer n_p, find the smallest disk centered at the location of the jth event which includes at least n_p other earthquakes and whose radius is larger than some minimum value (*e.g.* a distance within 0.02 degrees, which is the order of the location error), δ, and let this minimum radius be d_j.

To find optimal values of n_p and δ, the above variable kernel estimates were applied to estimate the rate of a simulated inhomogeneous Poisson point process. Through cross-validation,[30] it is found that the optimal n_p is in $3 \sim 6$. The parameter δ does not influence the prediction too much, as long as the locations of points are not rounded to a certain precision. This parameter only becomes important when some points happen to overlap at one location caused by rounding the number.

Iterative algorithm

To estimate the model parameters and the background seismicity rate simultaneously,[10,13] the following iterative algorithm can be introduced.
A1. Given a fixed n_p and δ, (5 and 0.05 degree, equivalent to 5.56 km on the earth's surface, which is close to the location error of earthquakes), calculate the bandwidth h_j for each event (t_j, x_j, m_j), $j = 1, 2, \cdots, N$.
A2. Set $\ell \leftarrow 0$, and $u^{(0)}(x, y) \leftarrow 1$.
A3. Using the maximum likelihood procedure, fit the model with conditional intensity function

$$\lambda(t,x,y|\mathcal{H}_t) = \nu u^{(\ell)}(x,y) + \sum_{i:t_i<t} \kappa(m_i)g(t-t_i) \times f(x-x_i, y-y_i|m_i) \quad (8)$$

to the earthquake data, where κ, g and f are defined in equations (4), (5) and (6), and ν is the relaxing coefficient, which is introduced in order to fasten the convergency speed of this algorithm. The model parameters are (ν, A, α, c, p, D, q, γ).
A4. Calculate φ_j for $j = 1, 2, \cdots, N$ by using equation (6).
A5. Calculate $\mu(x, y)$ by using equation (7) and record it as $u^{(\ell+1)}(x, y)$.
A6. If $\max|u^{(\ell+1)}(x, y) - u^{(\ell)}(x, y)| > \varepsilon$, where ε is a given small positive number, then set $\ell \leftarrow \ell+1$ and go to step A3; otherwise, take $\nu u^{(\ell+1)}(x, y)$ as the background rate and also output φ_i.

Theoretical distributions associated with the largest magnitude of all the descendants from a given event

The method for obtaining the background rate is based on the space-time ETAS model, *i.e.* the occurrence rate of the clusters. However, the hazard caused by earthquake clusters not only depends on their occurrence rate, but also on the magnitudes of large events in each cluster.

The largest magnitude in an arbitrary cluster and in a given space-time volume must be evaluated.

Basic equations

First, assume that the probability density function (p.d.f), $s(m)$, for all the magnitudes obeys the Gutenberg-Richter law (exponential distribution), *i.e.*

$$s(m) = \beta\, e^{-\beta(m-m_c)}, \quad m \geq m_c \tag{9}$$

where β is the constant linked with the Gutenberg-Richter's b-value by $\beta = b \ln 10$ and m_c is the magnitude threshold.

Note the number of children from an event of magnitude m is a Poisson random variable with a mean of $\kappa(m)$, *i.e.*

$$\Pr\{\text{an event sized } m \text{ has } n \text{ direct offspring}\} = \frac{[\kappa(m)]^n}{n!}\, e^{-\kappa(m)}. \tag{10}$$

The probability that an event of magnitude m has no offspring greater than m', which is the same as the probability that the largest descendant from an event of magnitude m is smaller than m', can be derived from the model analytically[31] in the following way:

$$
\begin{aligned}
\zeta(m,m') &= \Pr\{\text{an event sized } m \text{ has no offspring} > m'\} \\
&= \sum_{n=0}^{\infty} \Pr\{\text{each child is } \leq m' \text{ and has no offspring} \\
&\qquad\qquad > m' \mid m \text{ has } n \text{ direct offspring}\} \\
&\qquad\qquad \times \Pr\{m \text{ has } n \text{ children}\} \\
&= \sum_{n=0}^{\infty} \left[\int_{m_c}^{m'} s(u)\,\zeta(u,m')\,du \right]^n \frac{[\kappa(m)]^n}{n!}\, e^{-\kappa(m)} \\
&= \exp\left\{ -\kappa(m)\left[1 - \int_{m_c}^{m'} s(m^*)\,\zeta(u,m')\,du \right] \right\}.
\end{aligned}
\tag{11}
$$

It is evident that $\zeta(m, m')$ has the form

$$\zeta(m,m') = \exp[-\kappa(m)\, F(m')]. \tag{12}$$

where

$$F(m') = 1 - \int_{m_c}^{m'} s(m^*)\, \exp[-\kappa(m^*)\, F(m')]\, dm^* \tag{13}$$

represents the probability that the largest earthquake in an arbitrary cluster, including the initial event and all its descendants, is greater than m'. Thus, $F(m)$ gives the largest magnitude in an arbitrary cluster. In other words, the magnitude distribution of the largest event in a cluster is no longer Gutenberg-Richter law, but determined by functional equation (13). Considering that the background ground rate $\mu(x, y)$ gives the occurrence rate of earthquake clusters, a combination of $F(m)$ and the background rate provides information on the maximum magnitude of future earthquakes.

Asymptotical properties of $F(m)$

Because $F(m)$ is a function determined by $\kappa(m)$ and $s(m)$, it is influenced only by the parameters A, α and β in the model. Substituting equations (9) and (2) into equation (13) we obtain

$$
\begin{aligned}
F(m') &= 1 - \left[\Gamma_{-\frac{\beta}{\alpha}}\left(A\,F(m')\right) - \Gamma_{-\frac{\beta}{\alpha}}\left(A\,F(m')\,e^{\alpha(m'-m_c)}\right) \right] \\
&\quad \times \frac{\beta}{\alpha}\, [A\,F(m')]^{\frac{\beta}{\alpha}}
\end{aligned}
\tag{14}
$$

where $\Gamma_a(x) = \int_x^{\infty} e^{-u} u^a\, du$ is the complementary incomplete gamma function. An analysis with moment generating functions[32] also gave similar forms of the above equations. But they did not consider the case in which the process is supercritical. The properties of $F(m)$ are closely related to the criticality of the process.[31] According to the Appendix (see also Zhuang *et al.*[33]), the critical parameter ρ is determined by

$$\varrho = \int_{m_c}^{+\infty} \kappa(m)\, s(m)\, dm. \tag{15}$$

Substituting equations (2) and (9) into the above equation,

$$\varrho = \frac{A\beta}{\beta-\alpha}. \tag{16}$$

it is evident that $\beta > \alpha$ is required. According to the discussion in the appendix, F has the following properties:[31,34]

i) subcritical case ($\varrho < 1$) where each family tree ultimately dies off and the whole process is stable and stationary. When m is large enough,

$$\lim_{m \to +\infty} \frac{F(m)}{s(m)} = \frac{1}{\beta(1-\varrho)}, \quad \text{or,} \quad F(m) \sim \frac{e^{-\beta m}}{1-\rho}. \tag{17}$$

ii) critical case ($\varrho = 1$), where each family tree dies off with a long tail and the population of the whole process in unit time experiences an unlimited increase. In this case, the asymptotic property of F branches according to the relative values of α and β.[34] When $\beta/2 < \alpha < \beta$,

$$F(m) \sim e^{-\alpha m}, \tag{18}$$

and when $\alpha \leq \beta/2$

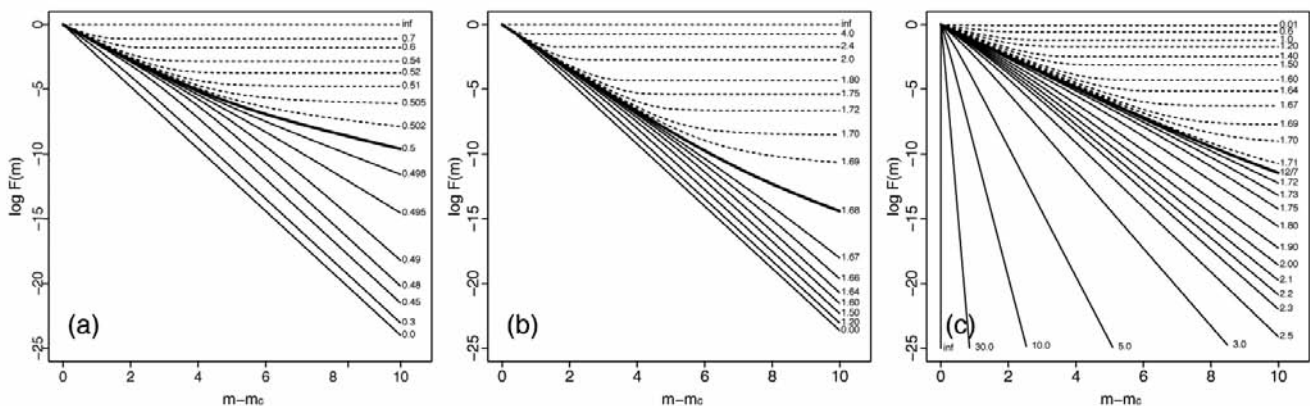$$F(m) \sim A^{-1}\sqrt{2}(1 - 2\alpha/\beta)^{1/2} e^{-\beta m/2}. \tag{19}$$

Figure 1. Influence of A, α and β to function $F(m)$. (A) Parameter A changes, but α=1.2 and β=2.4 are fixed. (B) Parameter α changes, but A=0.3 and β=2.4 are fixed. (C) Parameter β changes, but A=0.3 and α=1.2 are fixed. The thin solid, thick solid and dashed curves represent the subcritical, critical and supercritical regimes, respectively.

iii) supercritical case ($\varrho \geq 1$), where some of the family trees may never die off and the population of the whole process explodes, $F(m) \sim C$, where C is linked with the extinct probability $P_c(m)$ of the family starting from an event of magnitude m by $Pc(m) = e^{-C\kappa(m)}$.

Using the following iteration algorithm, $F(m)$ can be solved for each value of m

$$\begin{cases} F_{(0)}(m) = 1, \\ F_{(k)}(m) = 1 - \int_{m_c}^{m} e^{-\kappa(m^*) F_{(k-1)}(m)} s(m^*) \mathrm{d}m^*, \quad k = 1, 2, \cdots. \end{cases} \quad (20)$$

In Figure 1, first two of A, α and β are fixed and then the other is changed to see how $F(m)$ changes. As discussed above, the influence of all the three parameters, α, β and A, to $F(m)$ is through their influence to the criticality, *i.e.* the properties of $F(m)$ can be divided into three regimes: subcritical, critical and supercritical, where $F(m) \sim s(m)$, $F(m) \sim 1/\kappa(m)$ and $F(m) \sim C$, respectively.

Probabilities of a future large earthquake

It should be remembered that $F(m)$ is the survivor probability function of the magnitude of the biggest event in a cluster, including the initial event and all its descendants. Given a space-time window V and a magnitude m, the probability that the largest event in V is greater than m can be evaluated in the following way:

$$\begin{aligned} p_0(m, V) &= \Pr\{\text{the largest event in } V \text{ is greater than } m\} \\ &= \Pr\{\text{at least 1 event greater than } m \text{ occurs in } V\} \\ &= 1 - \Pr\{\text{all biggest events in each cluster are less than } m \text{ in } V\} \\ &= 1 - \sum_{n=0}^{\infty} \Pr\{n \text{ clusters occur in total}\} \\ &\qquad \times \Pr\{\text{all events in } n \text{ clusters} < m \text{ in } V \mid n \text{ clusters in total}\} \\ &= 1 - \sum_{n=0}^{\infty} [1 - F(m)]^n \frac{\Lambda^n(V)}{n!} e^{-\Lambda(V)} \\ &= 1 - e^{-\Lambda(V) F(m)}. \end{aligned} \quad (21)$$

where $\Lambda(V) = \iiint_V \mu(x,y) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t$ is the mean number of clusters occurring in V.

Here it is assumed that all events of a cluster occur at the same location and at the same time, which may not be true. However, for long-term risk evaluation, such a simple approximation is reasonable and also reduces the calculation tasks.

It should be noted that the above equation only holds true when the process is stationary, *i.e.* subcritical. In the case of supercritical or critical processes, the population of the process explodes to infinity. Except for the trivial case of $\Lambda(V) = 0$, the largest event in V can theoretically exceed any magnitude.

Results

To illustrate the above procedures, the JMA (Japan Meteorological Agency) earthquake catalog was used to evaluate the example of the long-term probabilities of large earthquakes in and around the Japan region. First, data were selected from the JMA catalog in the ranges of longitude 128°-148°E, latitude 30°-45°N, depth 0-100 km, over the period from 1 January 1926 to 31 December 2009, and magnitude $M_J \geq 4.0$. There are 36,144 events in this dataset. Figure 2 shows the locations of the epicenters and occurrence times of events in the dataset.

Before fitting the space-time ETAS model to the catalog, some preprocessing of data is necessary. For the earthquake catalog as in this study, which covers records of a time period of 38 years, completeness and homogeneity are usually problems for statistical analysis. If seismicity in some regions or the whole region has an increasing trend, the fitting
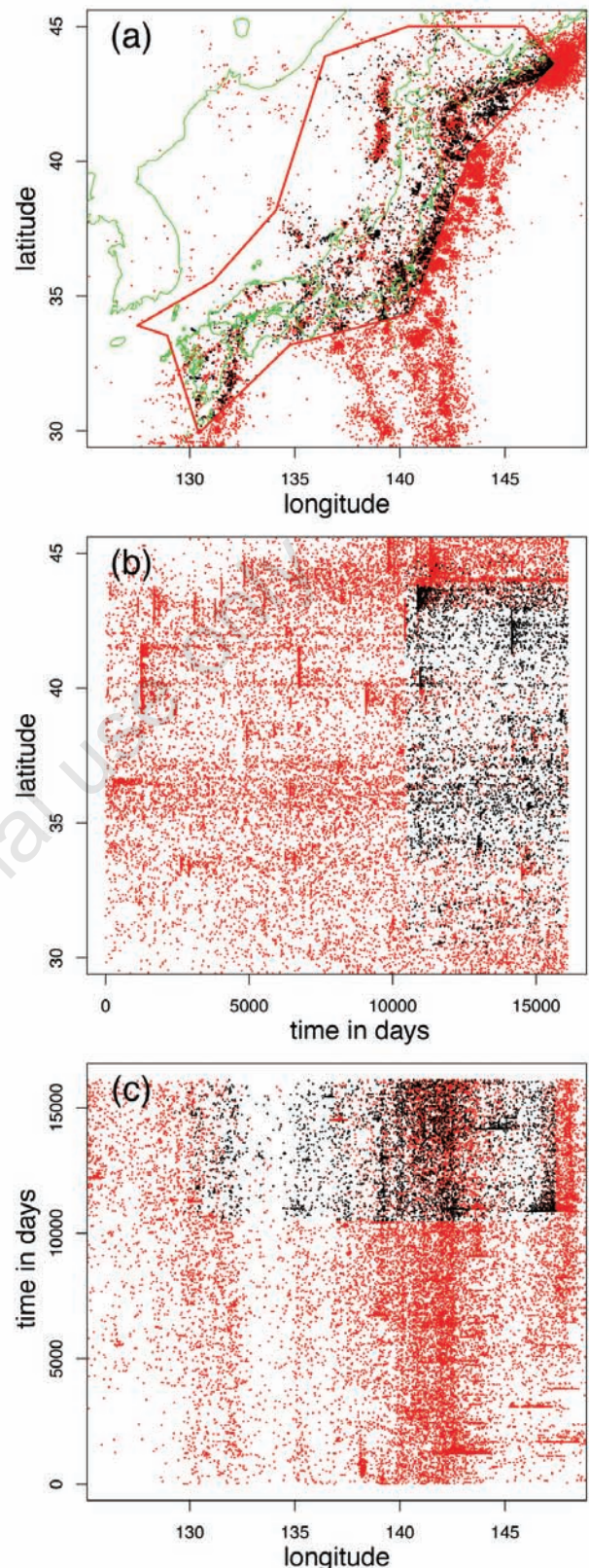

Figure 2. (A) Epicenter map of the shallow earthquakes ($M_J \geq 4.0$ and depth ≤100, 1 January 1926 to 22 September 2003) from the JMA catalog. (B) A space-time plot of latitudes against occurrence times, and (C) a space-time plot of occurrence times against longitudes. The polygon in (A) represents the target region used in the likelihood computation, and the black dots and red dots represent the target events and complementary events, respectively.
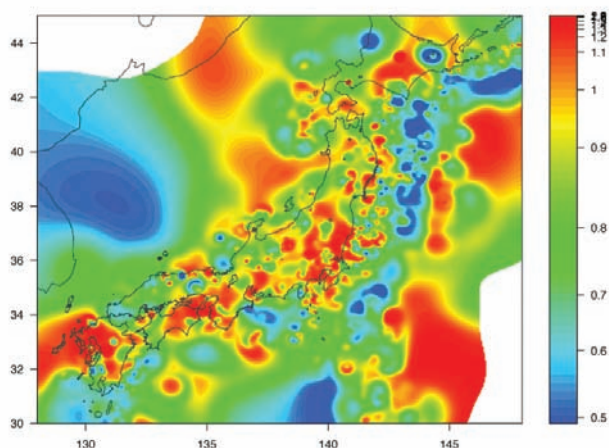
Figure 3. Background seismicity rate (occurrence rate of earthquake clusters) [events/(day·deg²)] of earthquakes with magnitude $M_J \geq 4.0$, obtained based on the space-time ETAS model.
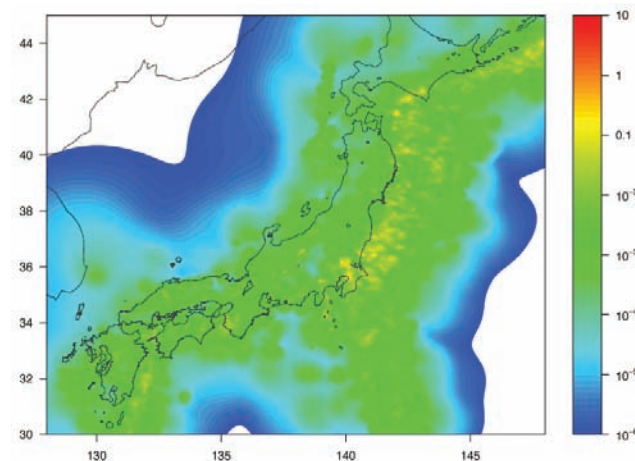

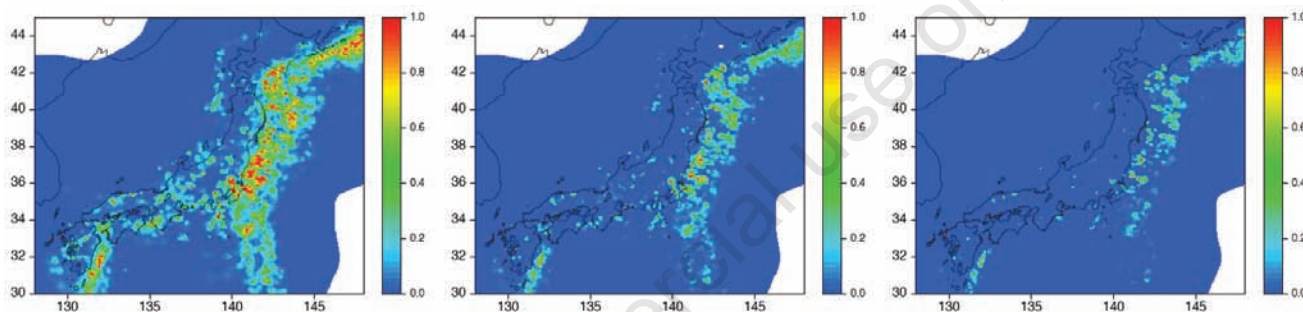**Figure 4. Spatial variations of b-values estimated by variable kernel estimates.**

Figure 5. Spatial variations of the occurrence probabilities that at least one earthquake in a 1°×1°× l-year space-time volume is greater than M_J 5.0 (left panel), M_J 6.0 (middle panel) and M_J 7.0 (right panel).

results often do not converge or they converge to some unreasonable values. Thus, a space-time range was chosen, in which the seismicity seems to be relatively and visually complete and homogeneous above 4.0, as the target space-time range of data to fit the model (Figure 2A and Table 1 in Zhuang[35]). The events outside this target space-time range are used as complementary events for calculating the conditional occurrence rate in order to illuminate the boundary effect (see also Zhuang[35]).

The same estimates as previously described[35] were used in this study, i.e. $\hat{\nu}$ =0.383 (events/day), $\hat{A}$ =0.232 (events), $\hat{c}$ =0.00578 day, $\hat{\alpha}$ =1.41 (m⁻¹), $\hat{p}$ =1.08, $\bar{d}$ =1.01×10⁻⁵ (deg²), $\hat{q}$ =1.59 and $\hat{\gamma}$ =1.38 (m⁻¹), which is also similar to the results of Zhuang *et al.*[9,10] The estimated background seismicity rates are shown in Figure 3.

The spatial variations of the β-value (b-value) in the magnitude distribution are evaluated by using the MLE (maximum likelihood estimate) method: at the locations of each 0.1°×0.1° grid,

$$\hat{\beta}(x,y) = \frac{1}{\bar{m}(x,y;200) - m_c + 0.05},$$

where $\bar{m}(x, y; k)$ represents the mean magnitude of the k closest earthquakes to (x, y), and 0.05 is the correction of the estimation error caused by rounding the magnitude to 1 digit in the catalog. The results are shown in Figure 4.

Substituting $\mu(x, y)$, $\beta(x, y)$, and the values of and $\hat{\alpha}$ into equation (21), the probabilities that at least one event occurs in a 1°×1°×1-year space-time window with a magnitude no less than M_J 5, M_J 6 and M_J 7 are evaluated and plotted in Figure 5. From Figure 5 a series of conclusions can be drawn. i) The regions with highest probabilities are the Tohoku (northeastern Japan) Arc and the Kuril Arc, and the northern end of the Ryukyu Arc, which is east to the Kyushu Island. ii) Comparing these locations to the Tokai and Nankai regions, which are both offshore south of the Honshu and the Shikoku islands, such probabilities are relatively lower than the Tohoku (northeastern Japan) Arc and the Kuril Arc regions, and the northern end of the Ryukyu Arc. iii) Along the middle ridge of Honshu island, which makes up the major part of the volcanic front in the Japan region, occurrence probabilities in unit space-time volume are much lower than in the coastal and offshore regions. iv) The coastal and offshore regions of the sea around Japan have similar earthquake probabilities to the Tokai and Nankai regions.

Conclusions

The ETAS model gives a natural description of earthquake clusters that can be used as a starting point for studies of seismicity patterns.

Based on the stochastic declustering method, we can objectively decompose the seismicity into the background and clustering components. The background component can be used to estimate long-term earthquake hazards as declustered catalogs (declustered by window-based or link-based declustering methods). However, only using a simple combination of the background rate and the Gutenberg-Richter law for the magnitude distribution might underestimate such risks, since the maximum magnitude in an earthquake cluster no longer obeys the Gutenberg-Richter law, but a different distribution as described above by equation (14). Instead, more viable forecasts can be made by using the equation (21) to correct the bias. For data analysis, the seismicity in and around the Japan region was used to show how to apply the procedures and techniques introduced in this study. The results show that the Tohoku Arc and Kuril Arc are the regions at the greatest potential risk of having large earthquakes in the future and that this risk is much higher than in the Tokai and Nankai regions where scientists and the public at large expect one or more earthquakes of M7.5 or over.

The data analysis in this study is based on the hypothesis that future seismicity follows similar patterns seen in the past. This may not be true. For example, some areas currently in a quiescent period may be re-activated in the future. Certainly, any practicable models for long-term earthquake forecast should consider these possible large-scale changes in seismicity. However, before considering more *advanced* conditions, a solid reference model should be built to provide a null hypothesis by which to test whether these more complicated models offer any advantages.

References

1. Utsu T. Aftershock and earthquake statistics (I): Some parameters which characterize an aftershock sequence and their interrelations. J Fac Sci Hokkaido Univ Ser VII 1969;3:129-95.
2. Gardner J, Knopoff L. Is the sequence of earthquakes in southern California, with aftershock removed, Poissonian? Bull Seismol Soc Am 1974;64:1363-7.
3. Keilis-Borok VI, Kossobokov VI. Time of increased probability for the great earthquakes of the world. Comput Seismol 1986;19:45-58.
4. Reasenberg P. Second-order moment of central California seismicity, 1969-1982. J Geophys Res 1985;90:5479-95.
5. Frohlich C, Davis SD. Single-link cluster analysis as a method to evaluate spatial and temporal properties of earthquake catalogs. Geophys J Int 1990;100:19-32.
6. Davis SD, Frohlich C. Single-link cluster analysis, synthetic earthquake catalogs, and aftershock identification. Geophys J Int 1991;104:289-306.
7. Console R, Murru M. A simple and testable model for earthquake clustering. J Geophys Res 2001;106:8699-711.
8. Helmstetter A, Ouillon G, Sornette D. Are aftershocks of large Californian earthquakes diffusing? J Geophys Res 2003;108:2483.
9. Zhuang J, Christophosen A, Savage MK, et al. Differences between spontaneous and triggered earthquakes: their influences on foreshock probabilities. J Geophys Res 2008;113:B11302.
10. Zhuang J, Ogata Y, Vere-Jones D. Analyzing earthquake clustering features by using stochastic reconstruction. J Geophys Res 2004; B5:B05301.
11. Hainzl S, Ogata Y. Detecting fluid signals in seismicity data through statistical earthquake modeling. J Geophys Res 2005;110:B05S07.
12. Lombardi AM, Cocco M, Marzocchi W. On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, central Italy, sequence: apparent variation or fluid-driven triggering? Bull Seismol Soc Am 2010;100:1138-52.
13. Zhuang J, Ogata Y, Vere-Jones D. Stochastic declustering of space-time earthquake occurrences. J Am Stat Assoc 2002;97:369-80.
14. Kagan Y, Knopoff L. Statistical search for non-random features of the seismicity of strong earthquakes. Phys Earth Planet Inter 1976;12:291-318.
15. Omori F. On the aftershocks of earthquakes. J Coll Sci Imperial Univ Tokyo 1894;7:111-200.
16. Utsu T. A method for determining the value of b in a formula logn = a − bM showing the magnitude-frequency relation for earthquakes. Geophys Bull Hokkaido Univ 1965;13:99-103.
17. Utsu T, Ogata Y, Matsu'ura RS. The centenary of the Omori formula for a decay law of aftershock activity. J Phys Earth 1995;43:1-33.
18. Zhuang J, Chang CP, Ogata Y, Chen YI. A study on the background and clustering seismicity in the Taiwan region by using a point process model. J Geophys Res 2005;110:B05S13.
19. Ogata Y, Zhuang J. Space–time ETAS models and an improved extension. Tectonophysics 2006;413:13-23.
20. Ogata Y. Space-time point-process models for earthquake occurrences. Ann Inst Stat Math 1998;50:379-402.
21. Console R, Murru M, Lombardi AM. Refining earthquake clustering models. J Geophys Res 2003;108:2468.
22. Daley DD, Vere-Jones D. An introduction to theory of point processes – Vol 1: Elementary theory and methods (2nd ed). New York: Springer; 2003.
23. Fletcher R, Powell MJD. A rapidly convergent descent method for minimization. Computer J 1963;6:163-8.
24. Musmeci F, Vere-Jones D. A variable-grid algorithm for smoothing clustered data. Biometrics 1986;42:483-94.
25. Helmstetter A, Kagan YY, Jackson DD. Comparison of short-term and time-independent earthquake forecast models for southern California. Bull Seismol Soc Am 2006;96:90-106.
26. Werner MJ, Helmstetter A, Jackson DD, Kagan YY. High-resolution long-term and short-term earthquake forecasts for California. Bull Seism Soc Am 2011;101:1630-48.
27. Ogata Y. Space-time model for regional seismicity and detection of crustal stress changes. J Geophys Res 2004;109:B03308.
28. Ogata Y. On Lewis' simulation method for point processes. IEEE Transact Informat Theory 1981;IT-27:23-31.
29. Zhuang J. Second-order residual analysis of spatiotemporal point processes and applications in model evaluation. J Royal Stat Soc Series B (Stat Methodol) 2006;68:635-53.
30. Picard R, Cook D. Cross-validation of regression models. J Am Stat Assoc 1984;79:575-83.
31. Zhuang J, Ogata Y. Properties of the probability distribution associated with the largest event in an earthquake cluster and their implications to foreshocks. Phys Rev E 2006;73:046134.
32. Saichev A, Sornette D. Distribution of the largest aftershocks in branching models of triggered seismicity: theory of the universal Båth law. Phys Rev E 2005;71:056127.
33. Zhuang J, Harte D, Hainzl S, et al. Basic models for seismicity. Commun Online Resource Stat Seism Analysis 2012; ThemeV. Available from: url:http://www.corssa.org/
34. Vere-Jones D, Zhuang J. Distribution of the largest event in the critical epidemic-type aftershock-sequence model. Phys Rev E 2008; 78:047102.
35. Zhuang J. Next-day earthquake forecasts by using the ETAS model. Earth Planet Space 2011;63:207-16.