

# Critical zone of the branching crack model for earthquakes: Inherent randomness, earthquake predictability, and precursor modelling

Jiancang Zhuang<sup>1,2,3,a</sup>, Mitsuhiro Matsu'ura<sup>1</sup>, and Peng Han<sup>4</sup>

<sup>1</sup> Institute of Statistical Mathematics, Research Organization of Information and Systems, Tokyo 1908562, Japan

<sup>2</sup> Department of Statistical Science, the Graduate University for Advanced Studies, Tokyo, Japan

<sup>3</sup> London Mathematical Laboratory, London, UK

<sup>4</sup> Department of earth and space sciences, Southern University of Science and Technology, Shenzhen, P.R. China

Received 1 May 2018 / Accepted 9 December 2019

Published online 19 January 2021

**Abstract.** The branching crack model for earthquakes was developed by Vere-Jones and Kagan in the 1970s and the 1980s, respectively. With some simple and rational assumptions, its simulation results explain the Gutenberg-Richter magnitude-frequency relationship and the Omori-Utsu aftershock decay formula. By introducing the concept of the critical zone, this model can be connected with the asperity model, the barrier model, and the nucleation model through a parameter – criticality. Particularly, the size of the critical zone determines the maximum magnitude of potential earthquakes and the source of their anomalies. The key to earthquake forecasting is to determine whether the concerned area is in a critical state and how large the critical zone is. We discuss what kinds of anomalies are meaningful as candidates of earthquake precursors. Finally, we outline modelling strategies for earthquake precursors with low probability gains that are due to the inherent randomness of earthquake source processes.

## 1 Introduction

One of the most direct and most useful approaches for earthquake hazard mitigation is forecasting the probabilities of the occurrence of future disastrous earthquakes. Understanding the physical processes of how earthquake ruptures are generated, accelerated, and stopped, is indispensable for this purpose. In the 1990s, a series of papers by Geller et al. [9] asserted that the occurrence of earthquakes is not predictable. These papers led to a long argument on *Nature* ([https://www.nature.com/nature/debates/earthquake/equake\\_frameset.html](https://www.nature.com/nature/debates/earthquake/equake_frameset.html)). Even though the view of earthquake forecasting or prediction has changed over the past 20 years, these papers still influence earthquake forecasting studies. In this article, through simulations by using Vere-Jones' branching crack model, we explain why earthquake

<sup>a</sup> e-mail: [zhuangjc@ism.ac.jp](mailto:zhuangjc@ism.ac.jp), [zhuangjc@gmail.com](mailto:zhuangjc@gmail.com)

occurrence cannot be deterministically predicted and discuss what are the potentially useful indices for evaluating the risk of future large earthquakes.

## 2 Vere-Jones' branching crack model

### 2.1 Definition

Vere-Jones' branching crack model describes the earthquake rupture process at the micro scale [20,44,45]. This model does not assume any geometric shape for the earthquake rupture. In this model, the basic element of the earthquake rupture is the shear crack, which can be a small tangential slip on a small patch of the earthquake fault. Each crack independently triggers new cracks nearby on the fault according to some probability rules. In this way, the rupture process of an earthquake starts from a single crack and develops into an earthquake, as shown in Figure 1. This process is called the Galton-Watson process in mathematics: (1) The first generation with only one ancestor, i.e.,  $Y_0 = 1$ ; (2) the number of descendants in the  $(n + 1)$ th generation is the total number of direct offspring from each member of the  $n$ th generation,  $n = 0, 1, \dots$ , i.e.,

$$Y_{n+1} = \sum_{j=1}^{Y_n} X_j^{(n)}, \quad (1)$$

where  $\{X_j^{(n)} : j = 1, 2, \dots, Y_n\}$  is a set of independent copies of a nonnegative integer random variable  $X$ , representing the number of cracks triggered by a given crack (parent crack). We further assume that  $X$  has a finite expectation  $\nu$  and a finite variance  $\sigma^2$ . Parameter  $\nu$  plays an essential role in the modelling: when  $\nu < 1$ , the family tree extinguishes quickly; when  $\nu = 1$ , the family still extinguishes but slowly; when  $\nu > 1$ , there is some probability that the number of members in the family tree explodes. Thus, it is also called the criticality parameter [2].

### 2.2 Simulating earthquake moment

Zhuang et al. [64] simulated the Gutenberg-Richter magnitude-frequency relation, the source-time function, and the duration-moment relationship of earthquakes, under the following assumptions: (a) The energy released by each crack is the same; (b) the energies are released step-by-step, with each time step representing one generation of the branching process; (c) the energy of the earthquake is proportional to the total number of cracks; and (d) the rupture duration is proportional to the total number of generations. By a further assumption that there is a delay between each individual crack and its parent crack, Kagan [20] simulated the Omori-Utsu formula for the decay rate of aftershocks, the Utsu-Seki law for the relationship between the aftershock area and the mainshock magnitude [43], and the relationship between the mainshock magnitude and the number of aftershocks [49].

Zhuang et al. [64] showed

$$\Pr\{\text{moment} \geq M\} \sim \text{const} \times M^{-0.5} \exp(-M/M_r), \quad (2)$$

when the process is subcritical, where  $M_r = 2\sigma^2/(1 - \nu)^2$ . This is called the Kagan or tapered Pareto distribution [47], which is equivalent to a double-exponential distribution for the moment magnitude,

$$\Pr\{\text{magnitude} \geq m\} \sim \text{const} \times 10^{-0.75m - c10^{-dm}}, \quad (3)$$

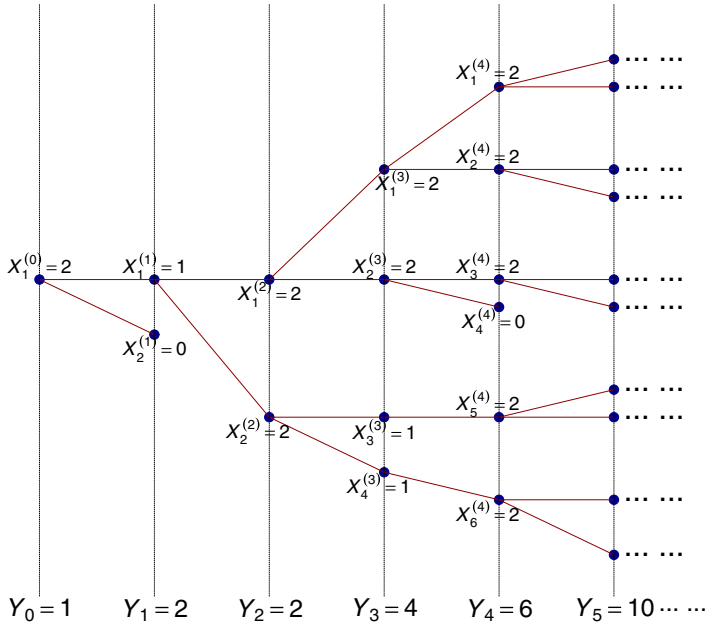


Fig. 1. A diagram illustrating the branching crack (Galton-Watson) process.

where  $c$  and  $d$  are constants. When the process is critical, the seismic moment follows a Pareto distribution, i.e.,

$$\Pr\{\text{moment} \geq M\} \sim \text{const} \times M^{-0.5}, \tag{4}$$

with the corresponding magnitude distribution following the Gutenberg-Richter law, i.e.,

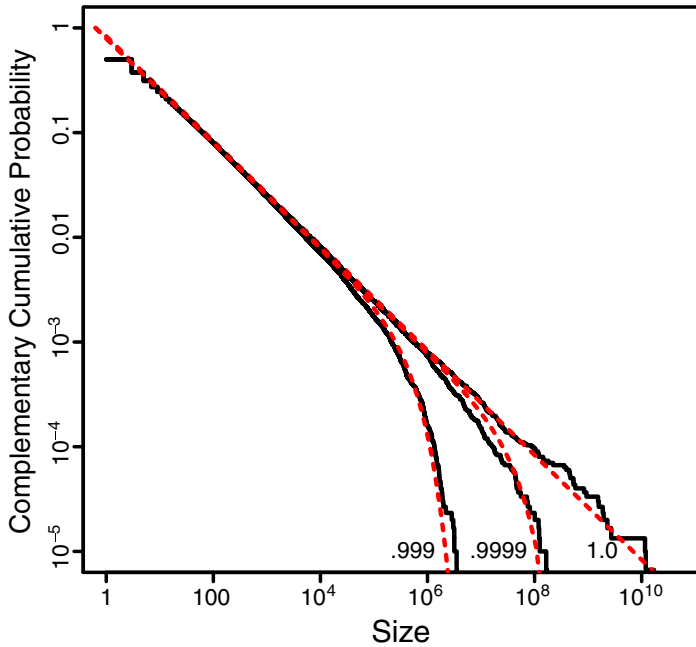
$$\Pr\{\text{magnitude} \geq m\} \sim \text{const} \times 10^{-0.75m}, \tag{5}$$

where 0.75 is the value of the so-called Gutenberg-Richter  $b$ -value. Figure 2 illustrates the simulation results and theoretical curves of seismic moments based on the branching crack model with different criticality parameters. Please note that the corner magnitude in the magnitude-frequency curves for subcritical cases is different from the corner magnitude given in [21,47], which arises from the limitation to earthquake magnitude caused by the distribution of pre-existing faults, with a scale of  $M6$  to 7.

### 2.3 Why we cannot know the magnitude of an earthquake before it stops?

Assuming that each branching generation is a time step and that the seismic moment released at each time step is the number of cracks, Zhuang et al. [64] simulated the source-time function with the branching crack model. Figure 3 gives two examples of simulations with larger numbers of cracks. Their patterns are quite similar to the source-time functions after smoothing, exhibiting single or multiple peaks and no fixed shapes. If the branching process does not stop at a certain time step, any number of cracks or peaks are possible to be produced in its continuation.

Such inherent randomness explains why the earthquake magnitude cannot be determined before the dynamic rupture process stops completely. That is to say, the magnitude of an earthquake cannot be predicted in a deterministic manner during its



**Fig. 2.** Complementary cumulative distributions for seismic moments of simulated earthquakes with  $\nu = 0.999$ ,  $0.9999$ , and  $1$  (solid curve). Dashed curves represent the corresponding asymptotes of the Pareto (for  $\nu = 1$ ) and the tapered Pareto (for  $\nu < 1$ ) distributions (after [64] with modifications).

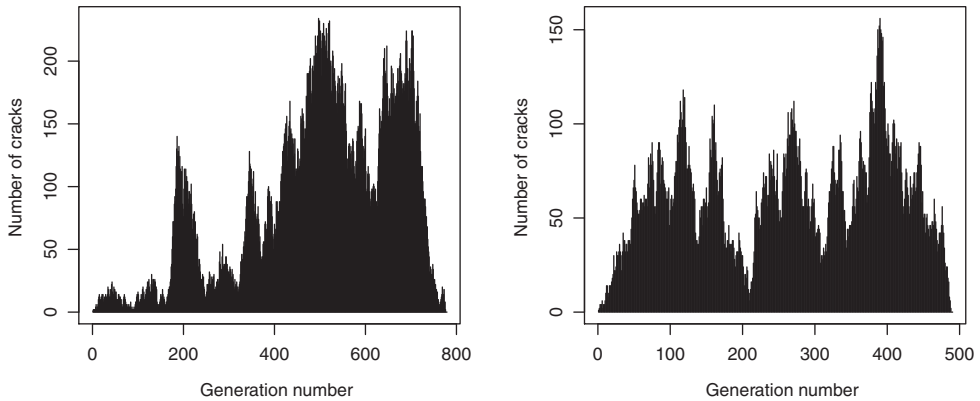
rupture. This opinion supports Rydelek and Horiuchi [39] against that of Olson and Allen [34,35], which declared that the earthquake magnitude could be determined completely by the first several  $P$ -phases.

The stopping time of a rupture process is closely related to the criticality. From the discussion in Section 2, we can see that, when the concerned area is subcritical, the rupture process can stop easily; that is, almost no large earthquakes can be generated there. On the other hand, large earthquakes can be easily generated when the concerned areas are in the critical state on account of the branching crack process.

### 3 Critical Zone

#### 3.1 Concept

One may have the permissive conclusion that the occurrence of earthquakes is completely unpredictable. We cannot even determine the magnitude of an earthquake during its rupture. How can we do it before the initiation of the rupture process, i.e., the occurrence of an earthquake? Such inherent randomness in the total number of branching cracks really does imply the impossibility of the deterministic prediction of earthquake occurrence time, location, and magnitude. However, such a conclusion has been drawn under a simple implicit assumption: The concerned area in which the earthquake occurs is infinitely large, and its state is homogeneously critical everywhere. An infinitely large seismogenic zone with homogeneous stresses everywhere is unrealistic. In the real world, seismogenic zones are finite in size and have inhomogeneous stresses in both time and space.



**Fig. 3.** Simulated source time functions for two large events (c.f. [64]) using a critical binary branching crack model ( $p_0 = p_2 = 0.5$  and  $p_1 = p_3 = p_4 = \dots = 0$ ,  $p_i$  being the probability that an ancestor crack triggers  $i$  offspring).

From the viewpoint of fracture mechanics, it is reasonable to link this criticality with the background stress level relative to the strength of the medium of the seismogenic zone. When the background tectonic stress increases to the crustal strength of the medium, the state of the seismogenic zone turns from subcritical to critical. Conversely, a critical zone turns into a subcritical one when the accumulated tectonic stress is released. In the critical zone, a rupture cannot easily stop. However, once this rupture extends beyond the critical zone and runs into a subcritical zone, the rupture stops quickly.

There are two types of subcritical zones surrounding a critical zone. In a Type I subcritical zone, the stress level is the same as the critical zone, but the material strength is higher. The Type II subcritical zone has the same strength as the critical zone, but its stress level is lower. Traditionally, two models have been proposed to describe the interaction between the earthquake rupture and the surrounding environment: the barrier model (e.g., [1,6]) and the asperity model (e.g., [23]), corresponding to Types I and II critical zones, respectively. In the barrier model, relatively weak patches are surrounded by strong patches, and an earthquake rupture is regarded as a stress-roughening process leaving unbroken patches. In the asperity model, relatively strong patches are surrounded by weak regions where stress is released by previous earthquakes or aseismic slip. Thus, Type I and II subcritical zones correspond to the barrier in the barrier model and the weak region in the asperity models, respectively.

The size of the critical zone restricts the spatial extension of potential cracking processes and limits the maximum magnitude of earthquakes that can occur in this area. Thus, the key points in evaluating the potential of a future disastrous earthquake are to determine whether the area of interest is in the critical state, and if so, how large the critical zone is. If we can determine the criticality status of the concerned area and evaluate the extent of such a critical zone through high resolution geophysical observations, we can then make more reliable forecasts (with high probability gain) regarding the occurrence time and magnitude of future large earthquakes. Due to the inherent randomness of the branching crack process, the forecast can only be made in probabilistic terms rather than in the form of affirmative predictions.

### 3.2 Determining the critical zone

In the following sections, we discuss what geophysical parameters are possibly useful for us to monitor the degree of criticality of an area.

**Corner magnitude,  $b$ -value, and activation of micro-seismicity.** From equation (1) and Figure 2, the most important parameter directly related to criticality is the corner magnitude  $M_r$ .  $M_r$  increases when the criticality of the branching cracking process increases and trends toward infinity when the process becomes critical. The value of  $M_r$  can be roughly estimated in the following way. First, we need to estimate at what scale the energy is released by each cracking. Roughly speaking, the smallest acoustic emission is comparable in energy to the fracture of a sharp pencil lead, which is about  $10^{-9}$  Joules (*Xinglin Lei, personal communication*). By using the relation  $M = \frac{2}{3} \log_{10} E - 3.2$  (with  $E$  in Joules), this is equivalent to an earthquake magnitude of  $-9.2$ . We simply take the standard deviation  $\sigma = 1$  and assume  $\nu = 0.99999$ . Substituting these two values into equation (3), we obtain a corner value  $M_r$  on the order of 10 Joules, which corresponds to an earthquake magnitude of  $-2.5$ . This magnitude is much less than the detection magnitude threshold of most monitoring networks. In real world, earthquakes have a different underlying process (mainly controlled by rock friction), and their minimum magnitude is generally estimated to be much larger than that. If the minimum released energy can be 1000 times higher than that of the fracture of pencil lead, the corner magnitude is about 0.5. Earthquakes of such a magnitude can be detected by our monitoring network but are difficult to record completely. This indicates that almost all seismogenic areas are in or near the critical state, with different sizes. This also explains why the observed  $b$  value is much larger than 0.75: The straight part in Figure 2 is not observed in practice as the corner magnitude is lower than the detection magnitude threshold. Kagan [22] hypothesised that the  $b$ -values obtained from seismic catalogues are, in fact, estimated based on events larger than the corner magnitudes or from a mixture of many tapered distributions with different corner magnitudes.

If the number of observed small earthquakes increases in an area, such an increase might be caused by two factors: (1) the increase in tectonic stress increases the occurrence rate of initiating cracks; (2) the increase in criticality of the branching cracking process increases the corner magnitude. When the criticality parameter increases to the critical value, the corner magnitude also shifts to larger values. In observation, we expect that the number of small events to increase. In addition, when the corner magnitude increases enough such that events at this magnitude are observable, the observed  $b$ -value decreases. This explains why, before some large earthquakes, low  $b$ -value anomalies (e.g., [29]) and accelerative moment release (AMR) phenomena (e.g., [3,4,18,38,41]) are evident.

**Load/unload response ratio (LURR):** Another meaningful parameter of criticality is the load/unload response ratio developed from the viewpoint of fracture mechanics (e.g., [50–53]). This parameter is based on the changes in the characteristics of solid materials when they evolve from the elastic phase to failure through the damage formation phase. Figure 4 illustrates a typical stress-strain curve for rock materials. Such curves are usually used to comprehensively describe the mechanical behaviour of rock materials under high-temperature and high-pressure conditions (see, e.g., [42]). When the rock is monotonically loaded, it evolves from an elastic phase to a damage formation phase and results in failure. In the elastic phase, the loading and unloading processes are reversible, that is, the loading and unloading moduli are identical. In contrast, in the damage formation phase, these moduli are different, and the loading

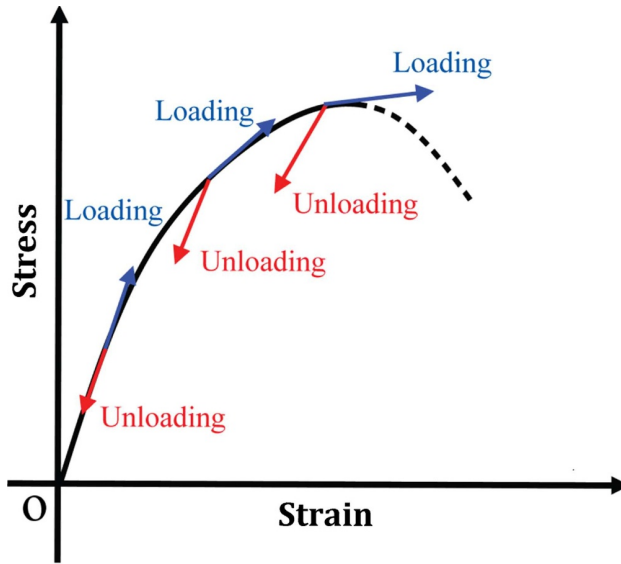


Fig. 4. An illustration of stress-strain relation of rock material (after [55] with modifications).

and unloading processes are irreversible, indicating a destabilization of the material due to damage. Such a difference can be quantitatively described by the ratio of response rates in the loading and unloading periods,

$$Y = \frac{\Delta R_+ / \Delta P_+}{\Delta R_- / \Delta P_-}, \quad (6)$$

where  $\Delta P_+$  and  $\Delta P_-$  represent the increase and decrease in stress loaded to the material, respectively, and  $\Delta R_+$  and  $\Delta R_-$  are the corresponding changes in response.  $Y$  is called the load/unload response ratio (LURR) since it is the ratio of the loading response rate  $\Delta R_+ / \Delta P_+$  to the unloading response rate  $\Delta R_- / \Delta P_-$ . It is easy to see that  $Y = 1$  implies that the material is in the elastic phase,  $Y > 1$  corresponds to the damage formation phase, and  $Y \rightarrow \infty$  corresponds to the failure (fracture) stage. Yin, Yin & Yin, and Yin et al. [50–53] used the Coulomb failure stress induced by earth tides as the loading and unloading source on faults and the Benioff strains released by small earthquakes as the response to calculate the  $Y$ -value. The results show that the  $Y$ -value increases about the focal zone of future large earthquakes and the  $Y$ -anomaly areas have a positive correlation with the magnitude of potential large earthquakes, implying that the focal zone agrees with the critical zone.

**Criticality in earthquake cluster.** It is well known that observed seismicity is clustered. Such clustering behaviours can be explained by Monte Carlo simulations carried out by Kagan [20] and modelled by the Epidemic-Type Aftershock Sequence (ETAS) model (e.g., [30,31,33,59,60]). The purely temporal version of the ETAS model is specified by the conditional intensity (the expected seismicity rate given the condition that the seismicity in the past is known),

$$\lambda(t) = \frac{\mathbf{E}[N[t, t+dt] | \mathcal{H}_t]}{dt} = \mu + \sum_{i:t_i < t} \kappa(m_i)g(t-t_i), \quad (7)$$

for earthquakes with magnitudes greater than the given magnitude threshold  $m_c$ , where  $N[t, t + dt)$  is the number of earthquakes occurring in  $[t, t + dt)$ , and

$$\kappa(m) = A \exp[\alpha(m - m_c)] \quad \text{and} \quad g(t) = (p - 1)(1 + t/c)^{-p}/c \quad (8)$$

represent the Utsu-Seki law and the Omori-Utsu formula, respectively, in the form of the probability density function, with  $A$ ,  $\alpha$ ,  $c$ , and  $p$  as constants. The meaning of this model includes two aspects: (1) the background seismicity is a stationary Poisson process with rate  $\mu$ ; (2) each earthquake, irrespective of whether it is a background event or a triggered event, triggers its own offspring independently. Similar as in the branching crack model, there is also a criticality parameter in the ETAS model,

$$\rho_{\text{etas}} = \int_{m_c}^{\infty} \kappa(m) s(m) dm, \quad (9)$$

where  $s(m)$  is the probability density of the earthquake magnitude. When  $\rho_{\text{etas}} < 1$ , the process is stable and stationary; otherwise, the process is unstable and goes into a state of population explosion (see Zhuang et al. [63] for more discussion).

This model is used to describe the “standard” behaviour of earthquake clustering. Here, the word “standard” means that the seismicity process is stationary in both background and clustering behaviours. The background and the clustering seismicities can be separated from the entire catalogue of seismic events by using the stochastic declustering technique [10,24,59,60]. In many cases, neither the background nor the clustering seismicity is stationary. In particular, before some large earthquakes, the background seismicity rate increases, and the earthquake process becomes more clustered. Many such changes can be interpreted by the Coulomb failure stress changes on the earthquake fault (e.g., [28]). Currently there are many reports about the increase in microseismicity before large earthquakes. However, since the ETAS model fitting requires a complete catalogue of seismic events, such a quantitative analysis of microseismicity has not been able to be carried out.

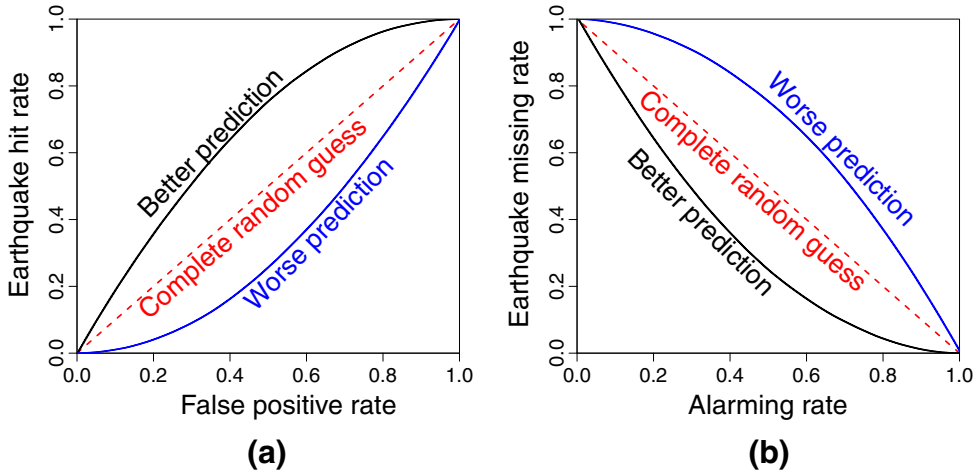
When the branching crack process trends toward the critical state, the increase of  $M_r$  can result in an increase in background seismicity and possibly a higher clustering effect (higher  $\rho_{\text{etas}}$ ). Thus, monitoring both  $\nu$  and  $\rho_{\text{etas}}$ , particularly, for microseismicity in local areas is useful to understand the critical states of seismicity and the degree of criticality of the crustal medium and the size of the critical zone. Such a clustering effect, related to small earthquakes, was discussed by Mignan [25].

**Changes in strain, gravity, and electromagnetic fields.** In addition to the above mentioned methods, there are many other possible ways to determine the extent of potential critical zones. For example, the crustal strain estimated from surface displacement data is closely related to the crustal stress change. Changes in the gravity field provide us with information about the vertical motion of the crustal medium (e.g., [5]). There are also many reports on the radiative signals related to the changes in geo-electromagnetic fields, coseismically (e.g., [16,17,36,37]) or preseismically (e.g., [7,11,12,14,61]). These observations may indicate changes in the underground stress field or the status of crustal materials, as discussed by Freund et al. [8].

## 4 Testing of earthquake precursors and probability gain

Observational data of any kind of anomaly must be converted into testable predictions before we determine whether they include precursory information about potential large earthquakes. A testable claim of a prediction or a forecast should include the





**Fig. 5.** Illustration of (a) the ROC curve and (b) the Molchan error diagram. Suppose that the entire forecasting space-time range is divided into  $n$  cells. We define the following numbers:  $a$  be the number of cells with successful predictions of occurrences;  $b$  be the number of cells with false alarms;  $c$  be the number of cells with successful prediction of non-occurrences; and  $d$  be the number of cells with occurrences failures to prediction. The numbers  $a$ ,  $b$ ,  $c$ , and  $d$  can be a function of the anomaly threshold. That is to say, whether an alarm is issued on a cell depends on the alarm level in this cell is above the threshold or not. The ROC curve is defined by the parameter curve of  $\frac{a}{a+d}$  against  $\frac{b}{b+c}$  for every possible alarm index threshold. The curve of the portion of failures to prediction  $\frac{d}{a+d}$  against the fraction of cells with positive prediction  $\frac{a+b}{a+b+c+d}$  is called the Molchan error diagram. A better alarm strategy than a completely random guess should have a ROC curve above the case of completely random guess and a Molchan error diagram below the completely random case.

following elements: a specific location or area, a specific span of time, and a specific magnitude range. Such testable earthquake predictions can be categorised into three classes: (1) binary (Yes/No) predictions, which give almost-sure statements about the occurrence of a future earthquake; (2) alarm-level type predictions, which give some indexing values showing the relative probabilities of earthquakes under consideration; and (3) probability forecasts, which are based on some probability model and provide the expectation of earthquakes occurring in a given space-time-magnitude window. A systematic scheme for verifying the precursory effect of anomalies in observations can be achieved through the ROC (Receiver Operating Characteristic) curve (e.g., [65]) and the Molchan error diagram (e.g., [26,27]). The ROC curve or Molchan error diagram provides comprehensive summaries of the performance of alarm-level type predictions. Here, we refer to Figure 5 for the usage of these two types of curves. Wang et al. [48] validated the precursory information in GPS displacement data in the Taupo region, the California region, and the Kanto region, Chen et al. [5] verified the explanatory effect of gravity anomalies for the occurrences of large earthquakes in western China, and Han et al. [14] checked the signals of short-term anomalies in geomagnetic observations, all using the Molchan error diagram. To deal with positive-only predictions, for example, the reverse tracing precursors [40], a gambling score is developed for such a purpose ([54,56,58]). Probability forecasts can be evaluated directly by probability gain or information gain (e.g., [46]).

The effectiveness of a precursor phenomenon can be measured by the probability gain, which is defined by

$$p_g = \frac{P(E|A)}{P(E)}, \quad (10)$$

where  $P(E|A)$  is the probability of earthquake occurrence conditioned on having observed the anomaly, and  $P(E)$  is the unconditional probability of an earthquake in the same time-space-magnitude range. This formula is from the Bayesian prediction formula for earthquake occurrence probabilities based on multiple precursors. Assuming that anomalous phenomena  $A_1, \dots, A_k$  are independent and the probability that  $A_i$  is followed by the expected earthquake is  $P(E|A_i)$ , then the probability of earthquake occurrence under the condition that multiple anomalies are observed is

$$\begin{aligned} P(E|A_1, A_2, \dots, A_n) &= \frac{P(E, A_1, A_2, \dots, A_n)}{P(A_1, A_2, \dots, A_n)} = P(E) \frac{P(A_1, A_2, \dots, A_n | E)}{P(A_1, A_2, \dots, A_n)} \\ &= P(E) \frac{P(A_1 | E) P(A_2 | E) \dots P(A_n | E)}{P(A_1) P(A_2) \dots P(A_n)} \\ &= P(E) \frac{P(E | A_1) P(E | A_2) \dots P(E | A_n)}{P(E) P(E) \dots P(E)} \\ &= P(E) \prod_{i=1}^K p_{g,i}, \end{aligned} \quad (11)$$

where  $p_{g,i} = P(E|A_i)/P(E)$  is the probability gain for anomaly  $A_i$ . Here, we used two assumptions of independence: the unconditional independence among anomalies yields

$$P(A_1, A_2, \dots, A_n) = P(A_1) P(A_2) \dots P(A_n)$$

and conditional independence among anomalies given occurrence of the earthquake ensures

$$P(A_1, A_2, \dots, A_n | E) = P(A_1|E) P(A_2|E) \dots P(A_n|E).$$

Such assumptions are usually used as the null hypothesis before any meaningful modelling is carried out. Though this assumption may not hold in practice, the above formula gives us a hint of a possible approach to improving earthquake forecasting performance, that is, combining different observations.

A common result from past studies is that the probability gain obtained by those precursors are quite low, usually from 2 to 4. With such low probability gains, the precursory effect cannot be determined by intuition but can only be verified by strict statistical tests. Comparing to the information provided by these anomalies, the earthquake clustering is the largest predictable component of the seismicity process, with a probability gain from 10 to  $10^3$  relative to the Poisson model (e.g., [15,19,57]).

Considering that the inherent randomness of the branching crack process disables us from predicting the final size of the earthquake before the entire branching crack process is completed, such precursory observations can only give us useful indications into whether the concerned area is in a critical state or not and how large the critical zone is. That is to say, there possibly exists a limit to the prediction powers of such precursors. The current probability gains estimated from the GPS displacement data by Wang et al. [48], gravity data by Han et al. [5], and geo-electromagnetic signals by Han et al. [14] all approximate a value of 3. Finding precursors with high probability gains as well as developing models and methods for combining different precursors are desired to obtain forecasts with higher probability gains. More physical mechanisms related to the earthquake preparation process should be incorporated in the model.

If pre-cursory signals expected from such models can be introduced, the resulting forecasts would gain in reliability.

## 5 Point-process modelling of explanatory effect of anomalies

There are three kinds of approaches to investigating the effect of anomalies, “data-driven”, “theory-driven”, and “experiment-driven” approaches. In the “data-driven” approach, we search for the correlation in occurrence times and locations between anomalies and earthquakes, and then attempt to explain the physical mechanisms of such correlations. The shortcoming of this approach is that there are many correlated datasets or sequences but without any causal relation, especially when dealing with big data. In the “theory-driven” approach, we start from some physical model, predict what phenomena we can observe, and then correct the model if the predicted phenomena cannot be observed or are different in quantity. The shortcoming of this approach is that the earthquake source is too complicated to be theoretically well modelled. Computer simulations also belong to the “theory-driven” approach. In the “experiment-driven” approach, we reproduce the high-pressure and high-temperature environment of deep underground, and then study the failure process of the crustal materials in the laboratory. This approach is limited by the difficulties in creating the same high-temperature and high-pressure environment and in understanding the actual composition of underground materials. In this section, we focus on the “data-driven” type statistical modelling of the correlation between earthquake occurrence and anomalies.

Statistical modelling of earthquake precursors can be dated back to the 1980s. Ogata and Katsura [32] used the LINLIN model (Hawkes’ self-exciting and mutually exciting processes) to evaluate the precursory information included in an external process. The time-varying event occurrence rate of this model is expressed as

$$\lambda(t) = \mu + \sum_{i: t_i < t} h_1(t - t_i) + \sum_{j: s_j < t} h_2(t - s_j), \quad (12)$$

where  $\{t_i : i = 1, 2, \dots\}$  represents the time series of the observed earthquakes, and  $\{s_j : j = 1, 3, \dots\}$  represents the time series of the observed anomalies. The excitation components take the forms of

$$h_1(t) = e^{-\alpha t} \sum_{l=1}^K a_l t^l, \quad \text{and} \quad h_2(t) = e^{-\beta t} \sum_{l=1}^L b_l t^l, \quad (13)$$

where  $K$  and  $L$  are two positive integers,  $a_l$  ( $l = 1, 2, \dots, K$ ),  $b_l$  ( $l = 1, 2, \dots, L$ ),  $\alpha$ , and  $\beta$  are parameters to be estimated. Zhuang et al. [61,62] used this model and a similar discrete version to model the explanatory effect of the observed data from the ultra-low frequency components of underground electric signals at 4 stations in the vicinity of Beijing in forecasting the occurrence of events with  $M \geq 4$  within a 300 km circle centred on Beijing, and showed that the explanatory effect is highly significant. The probability gain is about 2 – 10, similar to the values obtained by Wang et al. [48], Chen et al. [5], and Han et al. [14].

On the other hand, in probability earthquake forecasting, the ETAS model gives quite a high probability gain (e.g., [15,57]), up to  $10^2 \sim 10^3$  locally, which is much higher than the obtained probability gain for non-seismicity-based precursors. This suggests that the most effective way to model the explanatory effect of precursors is to construct a statistical model as an extension of the ETAS model, since the

clustering effect is the biggest predictable component in seismicity. Han et al. [13] tried the following model

$$\lambda(t) = \mu + \sum_{i:t_i < t} \kappa(m_i) g(t - t_i) + \sum_{j:s_j < t} h(t - s_j), \quad (14)$$

where  $\kappa(m)$  and  $g(t)$  are the same as those defined in equation (7) and

$$h(t) = B \exp \left[ -\frac{(t - T)^2}{2D^2} \right], \quad (15)$$

with  $B$ ,  $T$ , and  $D$  as constant parameters. In the above equation,  $B$  is the predictive power of the precursor,  $T$  represents the time lag between the anomaly and the expected earthquake, and  $D$  is the length of the alarmed time window. The unknown parameters of such a model can be estimated through the standard maximum likelihood procedure.

Using the above model, Han et al. [13] analysed the correlation between the occurrence times of the anomalies in magnetic records at the Kakioka station and the  $M \geq 4.0$  earthquakes within 100 km from the station (Fig. 6a). They found that the magnetic signals at this station exhibit a weak explanatory effect for the occurrence of nearby earthquakes, more significant for large events, up to an average probability gain of 1.40 for events of magnitude 5.0+ against the ETAS model. Figure 6b illustrates both self-excitation (high spikes) and the explanatory effects of anomalies (low bumps).

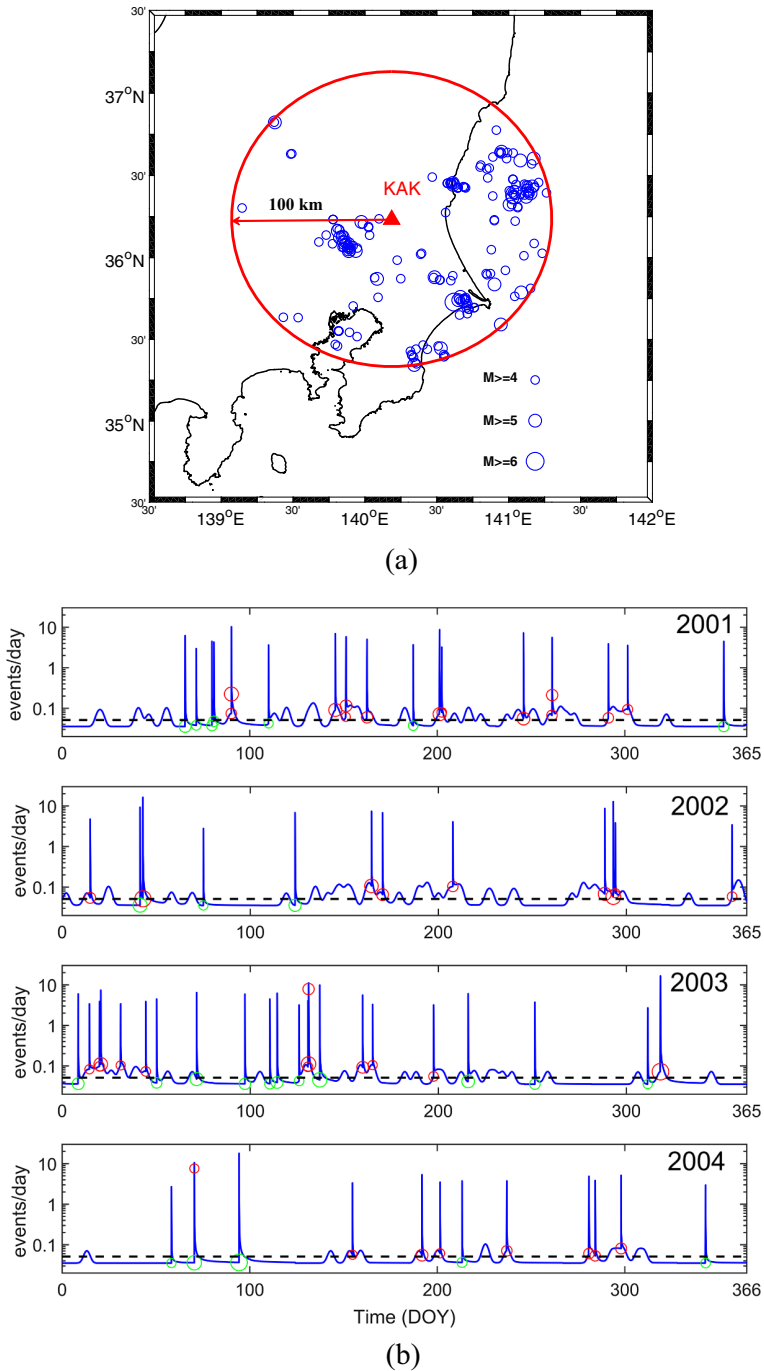
More long-term systematic observations of geophysical and geochemical data and systematic modelling studies are needed to acquire models that can issue forecasts with higher probability gains. In Zhuang et al. [62], with a combination of electric signals from different stations, the probability gains can be increased from 3 to 4.

Nevertheless, while the analysis of the correlation between the mechanical and seismic input data, facilitated by a probability model, does provide an improved understanding of the earthquake preparation process, its power is naturally limited by and depends on the used model. More physical mechanisms related to the earthquake preparation process should be incorporated in the model.

## 6 Discussion and conclusions

From the above discussion, the complexity of the source process of an earthquake can be produced with a simple mechanical model, and the randomness of the final size (magnitude) of an earthquake is inherent. Therefore, it is difficult, or perhaps impossible, to predict earthquakes in a deterministic manner or with almost-sure certainty, according to our current understanding. However, in our opinion, more reliable earthquake forecasts with significantly higher probability gain can be obtained, in which the key is to determine the size of the critical zone and to monitor when the system enters the critical state, through observations of microseismicity and other geophysical/geochemical variables. Such a conclusion is based on the following facts.

- (a) The biggest difference between the real world and the theoretical model of critical branching crack process is that the earth's crust is not in a critical state everywhere and all the time, but only some areas are in the critical state for some periods. The size of the critical zone controls the possible maximum magnitude of future earthquakes.
- (b) The critical zone can be detected by its stress field and other phenomena. Among many proposed earthquake precursors, we believe that the acceleration of microseismicity, the  $b$ -value, the LURR, and the criticality parameter in the ETAS



**Fig. 6.** (a) Location of the Kakioka Station and epicentre locations of nearby earthquakes of  $M4.0+$  from 2001 to 2010. (b) Conditional intensity (dark blue curve) from fitting model (14) to the recorded geomagnetic anomalies at the Kakioka station and the  $M4.0+$  earthquakes within 100 km from it, for the years of 2001 to 2004. In (b), the average level of seismicity is marked by a horizontal dashed line. The circles mark the values of the conditional intensity at the occurrence times of each earthquakes. The red and green circles represent, respectively, the earthquakes occurred at times that the conditional intensity is higher than the average level and at times that the conditional intensity is lower than the average level.

model, are good indicators of the current critical level. Other observations, such as GPS displacement, gravity field changes, and electromagnetic field changes are shown to be useful, even though the probability gain is lower than anticipated.

- (c) Due to the inherent randomness of the cracking process, these precursory indicators have an upper limit of probability gain in forecasting, which may be a bit far from satisfactory. High performance forecasts can be and possibly can only be made based on multidisciplinary precursors. From our discussion in Section 4, we can obtain a much higher probability gain by combining different observations. However, locating a place with long-term geophysical or geochemical observations is not trivial. Such observations need to be dense in space so that the area of the critical zone can be determined. These observations also need to be conducted for an extended period of time so that their performance, individual or combined, can be evaluated. For example, there are 859 seismic stations in Japan, thus, it is worthwhile observing magnetic or other non-seismic signals at these stations as well.
- (d) As the clustering effect is the largest predictable component of the seismicity, the modelling of the explanatory effect of the anomalies for earthquakes should be constructed from the ETAS model, as done in Han et al. [13] with external excitations. In this way, the ETAS model is automatically used as the benchmark for testing new forecasting models through standard model selection procedures, such as the Akaike Information Criterion (AIC).

In summary, the authors are still optimistic about earthquake forecasting. The main task is to develop monitoring technologies that can help us detect effective precursory anomalies and determine the size of the critical zone and the critical status of areas of interest. Moreover, the development of statistical inference and modelling methods for multidisciplinary precursors is also indispensable.

This study is supported by the JSPS KAKENHI Grant (19H04073) and China Earthquake Science Experiment Project, CEA (2019CSES0105, 2019CSES0106). Peng Han is also partially supported by the National Natural Science Foundation of China (41790465). The authors thank Dr. Friedemann Freund and an anonymous reviewer for their constructive comments.

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Publisher's Note** The EPJ Publishers remain neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

1. K. Aki, *J. Geophys. Res.: Solid Earth* **89**, 5867 (1984)
2. K.B. Athreya, P.E. Ney, *Branching Processes*, (Springer Verlag, 1972)
3. D.D. Bowman, C.G. Ouillon, C.G. Sammis, A. Sornette, D. Sornette, *J. Geophys. Res. Solid Earth* **103**, 24359 (1998)
4. C.G. Bufe, D.J. Varnes, *J. Geophys. Res. Solid Earth* **98**, 9871 (1993)
5. S. Chen, C. Jiang, J. Zhuang, *Pure Appl. Geophys.* **173**, 305 (2016)
6. S. Das, K. Aki, *J. Geophys. Res. Solid Earth* **82**, 5648 (1977)
7. A. Du, Q. Huang, S. Yang, *Geophys. Res. Lett.* **29**, 94–1 (2012)
8. F. Freund, G. Ouillon, J. Scoville, D. Sornette, submitted to *Europhysics* [arXiv:1711.01780] (2020)

9. R.G. Geller, D.D. Jackson, Y.Y. Kagan, F. Mulargia, *Science* **275**, 1616 (1997)
10. S. Hainzl, Y. Ogata, *J. Geophys. Res. Solid Earth* **110**, B05S07 (2005)
11. P. Han, K. Hattori, Q. Huang, T. Hirano, Y. Ishiguro, C. Yoshino, F. Febriani, *Nat. Hazards Earth Syst. Sci.* **11**, 965 (2011)
12. P. Han, K. Hattori, M. Hirokawa, J. Zhuang, C.H. Chen, F. Febriani, H. Yamaguchi, C. Yoshino, J. Liu, S. Yoshida, *J. Geophys. Res. Space Phys.* **119**, 4998 (2014)
13. P. Han, J. Zhuang, K. Hattori, Y. Ogata, *AGU Fall Meeting Abstracts* (2016) S23E-03
14. P. Han, K. Hattori, J. Zhuang, C.-H. Chen, J.-Y. Liu, S. Yoshida, *Geophys. J. Int.* **208**, 482 (2017)
15. A. Helmstetter, Y.Y. Kagan, D.D. Jackson, *Bull. Seismol. Soc. Am.* **96**, 90 (2006)
16. Q. Huang, *Proc. Jpn Acad. Ser. B* **78**, 173 (2002)
17. Q. Huang, H. Ren, D. Zhang, C.Y., *Proc. Jpn Acad. Ser. B* **91**, 17(2015)
18. S.C. Jaumé, L.R. Sykes, *Pure Appl. Geophys.* **155**, 279 (1999)
19. T. Jordan, Y.-T. Chen, P. Gasparini, R. Madariaga, I. Main, W. Marzocchi, G. Papadopoulos, G. Sobolev, K. Yamaoka, J. Zschau, *Ann. Geophys.* **54**, 316 (2011)
20. Y.Y. Kagan, *Geophys. J. R. Astr. Soc.* **71**, 659 (1982)
21. Y.Y. Kagan, *Geophys. J. Int.* **148**, 520 (2002)
22. Y.Y. Kagan, *Tectonophysics* **490**, 103 (2010)
23. T. Lay, H. Kanamori, An asperity model of large earthquake sequences, in *Earthquake Prediction, an International Review, Maurice Ewing Series, IV*, edited by D. Simpson, P. Richards (AGU, Washington, D.C., 1981), pp. 579–592
24. A.M. Lombardi, M. Cocco, W. Marzocchi, *Bull. Seismol. Soc. Am.* **100**, 1138 (2010)
25. A. Mignan, *Sci. Rep.* **4**, 4099 (2014)
26. G.M. Molchan, *Tectonophysics* **193**, 267 (1991)
27. G.M. Molchan, Y.Y. Kagan, *J. Geophys. Res.* **97**, 4823 (1992)
28. S. Nandan, G. Ouillon, J. Woessner, D. Sornette, S. Wiemer, *J. Geophys. Res. Solid Earth* **121**, 1890 (2016)
29. K.Z. Nanjo, N. Hirata, K. Obara, K. Kasahara, *Geophys. Res. Lett.* **39**, L20304 (2012)
30. Y. Ogata, *J. Am. Stat. Assoc.* **83**, 9 (1988)
31. Y. Ogata, *Ann. Inst. Stat. Math.* **50**, 379 (1998)
32. Y. Ogata, K. Katsura, *J. Appl. Probab.* **23**, 291 (2004)
33. Y. Ogata, J. Zhuang, *Tectonophysics* **413**, 13 (2006)
34. E.L. Olson, R.M. Allen, *Nature* **438**, 212 (2005)
35. E.L. Olson, R.M. Allen, *Nature* **442**, E6 (2006)
36. H. Ren, Q. Huang, X. Chen, *Geophys. J. Int.* **204**, 147 (2015)
37. H. Ren, Q. Huang, X. Chen, *Geophys. J. Int.* **206**, 205 (2016)
38. R. Robinson, *Geophys. J. Int.* **140**, 568 (2000)
39. P. Rydelek, S. Horiuchi, *Nature* **442**, E5 (2006)
40. P. Shebalin, V. Keilis-Borok, A. Gabriellov, I. Zaliapin, D. Turcotte, *Tectonophysics* **413**, 63 (2006)
41. D. Sornette, C.G. Sammis, *J. Phys. I Fr.* **5**, 607 (1995)
42. A. Sygała, M. Bukowska, T. Janoszek, *J. Sustainable Min.* **12**, 45 (2013)
43. T. Utsu, A. Seki, *Zisin* **7**, 233 (1955)
44. D. Vere-Jones, *Pure Appl. Geophys.* **114**, 711 (1976)
45. D. Vere-Jones, *J. Int. Assoc. Math. Geol.* **9**, 455 (1977)
46. D. Vere-Jones, *Comput. Seismol.* **30**, 248 (1998)
47. D. Vere-Jones, R. Robinson, W. Yang, *Geophys. J. Int.* **144**, 517 (2001)
48. T. Wang, J. Zhuang, T. Kato, M. Bebbington, *Geophys. Res. Lett.* **40**, 2631 (2013)
49. Y. Yamanaka, K. Shimazaki, *J. Phys. Earth* **38**, 305 (1990)
50. X. Yin, C. Yin, *Sci. Chin. (Ser. B)* **5**, 512 (1991) (in Chin)
51. X. Yin, X. Chen, Z. Song, C. Yin, *Pure Appl. Geophys.* **145**, 701 (1995)
52. X. Yin, L. Zhang, H. Zhang, C. Yin, Y. Wang, Y.X. Zhang, K. Peng, H. Wang, Z. Song, H. Yu, J. Zhuang, *Pure Appl. Geophys.* **163**, 2317 (2006)
53. X.-C. Yin, *Earthquake Res. Chin.* **3**, 1 (1987) (in Chin)
54. D.J. Zecher, J. Zhuang, *Geophys. J. Int.* **182**, 1319 (2010)
55. L.-P. Zhang, J. Zhuang, *Tectonophysics* **509**, 191 (2011)

56. J. Zhuang, *Geophys. J. Int.* **181**, 382 (2010)
57. J. Zhuang, *Earth Planet Space* **63**, 207 (2011)
58. J. Zhuang, C. Jiang, *Tectonophysics* **524/525**, 155 (2012)
59. J. Zhuang,, Y. Ogata, D. Vere-Jones, *J. Am. Stat. Assoc.* **97**, 369 (2002)
60. J. Zhuang, C.-P. Chang, Y. Ogata, Y.-I. Chen, *J. Geophys. Res.* **110**, B05S13 (2005)
61. J. Zhuang, D. Vere-Jones, H. Guan, Y. Ogata, L. Ma *Pure Appl. Geophys.* **162**, 1367 (2005)
62. J. Zhuang, Y. Ogata, D. Vere-Jones, L. Ma, H. Guan, in *Imaging, modeling and assimilation in seismology*, edited by Y. Li (De Gruyter, Germany, 2013), Vol. II, pp. 351–377
63. J. Zhuang, M.J. Werner, D.S. Harte, *Phys. Rev. E* **88**, 062109 (2013b)
64. J. Zhuang, D. Wang, M. Matsu'ura, *Bull. Seismol. Soc. Am.* **106**, 1832 (2016)
65. M.H. Zweig, G. Campbell, *Clin. Chem.* **39**, 561 (1993)