Probabilistic Earthquake Forecasting: How to Live Better Between Complete Randomness and Complete Predictability

Jiancang Zhuang
Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan

August 25, 2024

Abstract

In earthquake forecasting, there is a significant gap between complete randomness and complete predictability. This presentation begins by discussing how to quantify predictability and outlining the current state of earthquake predictability from an information-theoretic perspective, together with the development line of of the ETAS model.

1 Introduction

Earthquake hazards continue to pose serious threats worldwide, as seen in recent devastating events like the February 2023 Turkey-Syria earthquake, which resulted in over 50,000 deaths and widespread destruction. Such disasters underscore the urgent need for effective earthquake prediction and preparedness. Earthquakes can trigger tsunamis, landslides, and aftershocks, compounding their destructive impact on communities, infrastructure, and economies. People hope that scientistic could predict these events before their occurrence so that we can take actions before hand to avoid our losses of lives and properties caused the earthquakes. However, it has been commonly accepted in ongoing seismological researches that predicting the exact time and location of earthquakes remains difficult.

2 Some history

The necessity of earthquake probability forecast arises from the viewpoints of both earthquake physics and statistical seismology. In 1892, John Milne, James Ewing, and Thomas Cray installed the first model seismometer in Japan, marking the start of modern seismology. Seismometers enable the detection of the occurrence of global earthquakes, such that we can calculate their occurrence time and hypocenter locations, and compile relatively complete earthquake catalogs. The primary applications of statistics in earthquake studies during this period were simple statistical techniques, such as linear regression and point estimates, among others, scattered among individual studies on different topics. The following subsections summarize the main findings during this stage.

The Gutenberg-Richter law for the magnitude-frequency relationship In 1944, Gutenberg and Richter published a formula describing the relationship between any magnitude, e.g., m, and the number of earthquakes in any given region and period (*Gutenberg and Richter*, 1944):

$$\log_{10} N(\geqslant m) = a - b\, m, \qquad \text{or,} \qquad N = 10^{a-b\,m}, \tag{1}$$

where $N(\geqslant m)$ is the number of earthquakes with a magnitude no less than m in the given region and period and b is the so-called Gutenberg-Richter b-value. In probability language, the magnitude follows an exponential distribution.

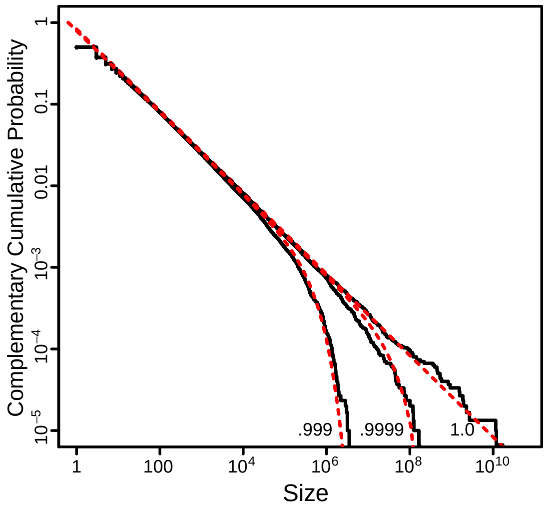

Figure 1: Complementary cumulative distributions for seismic moments of simulated earthquakes with ν=0.999, 0.9999, and 1 (black curves). Dashed curves represent the corresponding asymptotes of the Pareto (for ν=1) and the tapered Parto (for $\nu < 1$) distributions (after [*Zhuang et al.* (2016)] with modifications).

The Omori-Utsu formula for the aftershock frequency The Omori-Utsu formula describes the decay of the aftershock frequency with time after the mainshock, as an inverse power law in the form of

$$n(t) = K(t+c)^{-p}. \tag{2}$$

where t is the time from the occurrence of the mainshock, and K, c and p are constants (*Omori*, 1894; *Utsu*, 1957).

There are also many other empirical laws, such as the Båth law for the maximum magnitude of aftershocks and scaling laws related to earthquake magnitude.

3 Unpredictability of earthquake

To forecast the occurrence of future disastrous earthquakes. Understanding the physical processes how earthquake ruptures are generated, accelerated, and stopped, is indispensable. In the 1990s, a series of papers by *Geller et al.* (1997) asserted that the occurrence of earthquakes cannot be precisely predicted. These papers led to a long argument on *Nature* (https://www.nature.com/nature/debates/earthquake/equake_frameset.ht Even though the view of earthquake forecasting or prediction have changed over past 20 years, these papers still influening the researches of earthquake forecasting studies. Here I use Vere-Jones' branching crack model to explain why earthquake occurrence cannot be deterministically predicted and discuss what are the potentially useful indices for evaluating the risk of future large earthquakes.

Vere-Jones's branching crack model Vere-Jones' branching crack model describes the earthquake rupture process at the micro scale [*Vere-Jones*, 1976, 1977; *Kagan*, 1982]. 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 dnew 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 **??**. 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 nth generation, $n = 0, 1, \cdots$, i.e.,

$$Y_{n+1} = \sum_{j=1}^{Y_n} X_j^{(n)}, \tag{3}$$

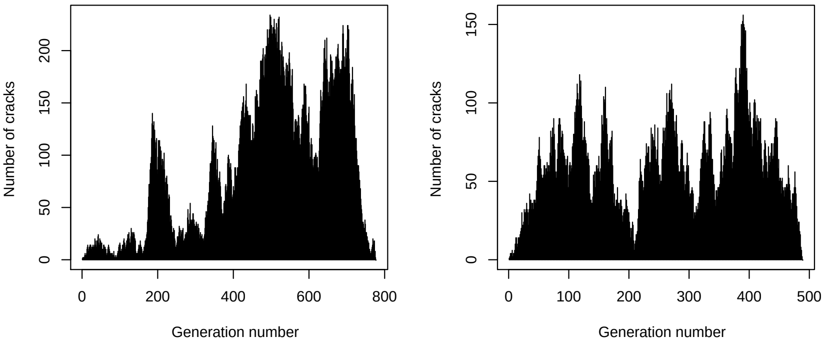

Figure 2: Simulated source time functions for two large events (c.f. *Zhuang et al.* (2016)) using a critical binary branching crack model ($p_0 = p_2 = 0.5$ and $p_1 = p_3 = p_4 = \cdots = 0$, p_i being the probability that an ancestor crack triggers i offspring).

where $\{X_j^{(n)} : j = 1, 2, \cdots, Y_n\}$ 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 ν and a finite variance σ^2. Parameter ν 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.

Such a simple model can be used to simulate most of the characteristics of earthquake sources. *Zhuang et al.* (2016) 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 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* (1982) 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 *Utsu and Seki* (1955), and the relationship between the mainshock magnitude and the number of aftershocks *Yamanaka and Shimazaki* (1990).

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.* (2016) simulated the source-time function with the branching crack model. Figure 2 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 rupture. This opinion supports *Rydelek and Horiuchi* (2006) against that of *Olson and Allen* (2005, 2006), 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.

Critical Zone 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.

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. *Zhuang et al.* (2021) discussed what possible approaches can be used to detect the critical zones.

Introducing point processes into statistical seismology In the 1970s with two important fact happened: the introduction of the point process model and the development of the theory of conditional intensity in the point process. First, the development of stochastic models for earthquake risks was a requirement for earthquake engineering. Building codes, with respect to the design of a building structure, required the consideration of the probability of the occurrence of the largest ground shaking event for a specific period into the future. In earthquake engineering, the stationary Poisson model (also referred to as the time-independent model) was often used to estimate the future earthquake hazard.

Considered as a classic reference, *Vere-Jones* (1970) proposed the point process to describe the process of earthquake occurrence times, focusing on the tools required to generate a functional and spectrum analysis. *Vere-Jones* (1973) introduced the use of the conditional intensity in statistical seismology, which is defined as the expectation of earthquake occurrence under the condition of previous knowledge on the earthquake process and/or external observations. They proposed the use of point processes to describe earthquake occurrence time series and developed tools to generate functional and spectral analyses.

The conditional intensity function is

$$\lambda(t)\mathrm{d}t = \Pr\{N[t, t+dt) > 0 \mid \text{Observation before } t\}. \tag{4}$$

defined as the expected probability of an earthquake occurring in the immediate future, conditional on the history of past seismic processes and external observational data. The advantage of using the conditional intensity function is that it is a natural concept for forecasting purposes, including estimation and simulation (*Ogata*, 1981).

The core idea of model development is bringing into existing stochastic models with more nonrandomness based on physical theory and observations. During this episode, for long-term earthquake hazard evaluations, renewal models were developed by modifying the inter-event time in the Poisson model to more general random distributions. The stress release model was developed by adding Reid's elastic rebound theory to the rate function of the Poisson model. For short-term earthquake forecasting, the Reasenberg and Jones model and the ETAS model were developed based on the Omori-Utsu formula.

Birth of the Hawks Process At about the same time, A. Hawkes, in a series of papers, gave the self- and mutually induced models in terms of conditional intensity functions and calculated the theoretical spectral functions of point processes (*Hawkes*, 1971a,b). In essence, a Hawkes process is a point process consisting of events generated by the induced effects of all past events under a background stationary Poisson process. Each event, whether it is a background event or an induced event, induces (excites) the occurrence of events in turn according to some probability rule. The model has conditional strength of the form

$$\lambda(t) = \mu + \sum_{i:t_i<t} g(t - t_i). \tag{5}$$

where, μ denotes the background rate, and $g(t)$ denotes the self-exciting effect. Specifically, it has been shown that the Hawkes process corresponds to the infection process (branching stochastic process) of an epidemic (*Kendall*, 1949; *Hawkes and Oakes*, 1974).

Statistical seismology before the ETAS model The ETAS model combines the Omori-Utsu formula used in aftershock analysis and the Hawkes model. Since the 1950s, the Omori-Utsu formula has been widely used in the analysis of aftershock activity. Utsu (1962) observed that not only the main shock but also large aftershocks may induce further aftershocks (secondary aftershocks). Such phenomena are shown in the following "multiple Omori-Uzu model".

$$\lambda(t) = K/(t-t_0)^p + \sum_{i=1}^{N_T} \frac{K_i\, H(t-t_i)}{(t-t_i+c_i)^{p_i}}, \tag{6}$$

where t_0 is the time of occurrence of the main shock, $t_i, i = 1, 2, \cdots, N_T$ is the time of occurrence of a significant aftershock, and,H is the Heaviside function. This observation is ground breaking because it overturns the conventional seismological hypothesis that the mainshock and aftershocks have different characteristics with respect to earthquake triggering. One difficulty in applying the multiple Omori-Utsu formula to general earthquake sequences is determining which earthquake triggers other events. The largest aftershocks are often accompanied by secondary aftershocks, but not always.

The ETAS model As explained by the multiple Omori-Uzu formula, aftershock activity often clearly includes secondary aftershocks. However, the distinction from primary aftershocks is generally not clear, and it is difficult to separate secondary aftershocks from seismic series data. In modeling earthquake occurrence, *Ogata* (1988) did not distinguish between triggering and triggered earthquakes, and supposed that each earthquake can trigger aftershocks, with the Omori-Utsu formula as the exciting response function and a positive exponential function of the magnitude as the productivity. This model was named as the Epidemic Type Aftershock Sequence (ETAS) model and expressed by a conditional function of

$$\lambda(t) = \mu + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i), \tag{7}$$

where $g(u)$ is the normalized form of the Omori-Utsuu formula $(p-1)/c(1+t/c)^{-p}$, and the weighting function $\kappa(m) = A \exp[\alpha(m-m_0)]$ is the expected number of earthquakes directly triggered by an earthquake of magnitude m. The magnitude of aftershocks need not be smaller than the triggering earthquake. An independent exponential distribution, i.e., the frequency of occurrence of the Gutenberg-Richter magnitude-frequency relationship is usually assumed for the purpose of theoretical discussion in simulation experiments of earthquake series. That is, the marked model takes the conditional intensity function of

$$\lambda(t, m) = s(m) \left[\mu + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i) \right], \tag{8}$$

where $s(m) = \beta e^{-beta(m-m_0)}$, $m \geq m_0$ is the probability density function form for the G-R law.

From time model to space-time model Before the ETAS model was generalized to the space-time ETAS model, *Musmeci and Vere-Jones* (1992), *Kagan* (1991), and *Rathbun* (1993) gave different form of spatio-temporal models for modelling seismicity. Nowadays, the conditional intensity of the frequently used spatiotemporal ETAS model is adopted from *Ogata* (1998):

$$\lambda(t, x, y) = \mu(x, y) + \sum_{i:t_i<t} \kappa(m_i) g(t-t_i) f(x-x_i, y-y_i; m_i), \tag{9}$$

where (m) and $g(t)$ are the same as in (7), the spatial response kernel

$$f(x, y; m) = \frac{1}{\pi\sigma(m)} f_0 \left(\frac{x^2+y^2}{\sigma(m)} \right) \tag{10}$$

is the density function for the relative location of the triggered earthquakes from an event of magnitude m, with f_0 being a normal density function $f_0(\omega) = \frac{1}{2D^2} e^{-\frac{\omega}{2D^2}}$ or a scaled inversed power law $f_0(\omega) = \frac{q-1}{D^2}(1+\omega/D^2)^{-q}$ is considered, and the scaling function is $\sigma(m) = \kappa(m)$ or $\sigma(m) = [\kappa(m)]^{\gamma/\alpha}$.

Spatio-temporal ETAS models are widely used in seismic activity analysis (e.g., see, *Ogata*, 1998; *Ogata et al.*, 2003; *Zhuang et al.*, 2002, 2004; *Console et al.*, 2003; *Helmstetter et al.*, 2003; *Lombardi* *et al.*, 2010; *Guo et al.*, 2015a).

Nowadays, the ETAS model has been accepted as the standard model for describing seismic activity (see special issue of *Huang et al.* (2016)) and adopted as the main earthquake prediction models by research institutes and government agencies in major earthquake-prone countries (Schorlemmer et al., 2018). In particular, the United States Geological Survey (USGS) adopted the ETAS model as the UCERF3-ETAS model for short-term forecasts in the third all-California earthquake probability forecasting model (UCERF3) (*Field et al.*, 2017).

The usefulness of the ETAS model encourage researchers to extend this model to higher dimension and higher resolution of seismicity, including

- **Impact of mainshock rupture geometry** Aftershock clusters of larger mainshocks are usually not isotropic around the epicenter of the mainshock, but are rather aligned along the mainshock rupture. *Hainzl et al.* (2008) demonstrated via ETAS simulations that ignoring this fact can lead to a biased parameter estimation, particularly to an underestimation of the α parameter. To account for anisotropy, *Ogata* (1998) used an elliptical aftershock distribution. Recently, *Guo et al.* (2015b) developed a finite-source ETAS model that considers the influence that the rupture geometry of large earthquakes has on the aftershock locations.

- **Impact of hypocenter depths** *Guo et al.* (2015a) also developed a 3-D ETAS model to incorporate the hypocenter depth in the model formulation, where the spatial response is as follows:

$$f(x, y, z; m_i, z_i) = f(x, y; m_i)h(z, z_i) \tag{11}$$

  where $h(z, z_i)$ takes form of a rescaled Beta density over the depth range. *Zhuang et al.* (2019) showed that the 3-D and finite-source ETAS models performed better than the 2-D point source ETAS model when fitting them to the Italian catalog.

- **Location-dependent and time-dependent ETAS parameters** *Ogata* (2004) developed powerful Bayesian tools with penalized likelihoods to estimate the changes in the clustering characteristics in the form of spatial variation in the model parameters. *Zhuang* (2015) developed the weighted likelihood method to estimate the spatial variations in the space-time ETAS model, applying it to the seismicity of Japan.

- **Self-similar ETAS models** *Vere-Jones* (2005) developed a self-similar ETAS model to avoid problems caused by the cut-off magnitude in the ETAS model and enable fully self-similar features. Owing to the missing data problem for immediate aftershocks and small events, however, this model remains theoretically underdeveloped and has not yet been applied to actual seismicity data.

4 Where have we achieved in earthquake forecasting

The predictability of earthquake occurrences is illustrated in Figure 3. This figure situates predictability between complete randomness, represented by the Poisson process for time occurrences and the Gutenberg-Richter relationship for magnitude distribution, and complete determinism, where the target can be predicted with 100% precision. In recent years, the ETAS model has become a de facto standard model or null hypothesis for comparing with and testing other models and ideas (*Huang et al.*, 2016; *Zhuang et al.*, 2021; *Zhuang*, 2023), which in fact implies that the clustering is the largest predictable component in seismicity. In the RELM and CSEP projects, the highly scored models are almost among different versions of the ETAS models. Though there reports declaring non-seismicity precursors, their performance has not been fully validated.

5 Conclusion and future prospects

As a subject that aims to bridge the gap between physical and statistical models, statistical seismology has developed rapidly during the last several decades. A significant achievement is the formulation of conditional intensity models for quantifying time-varying seismicity rates. The conditional intensity is

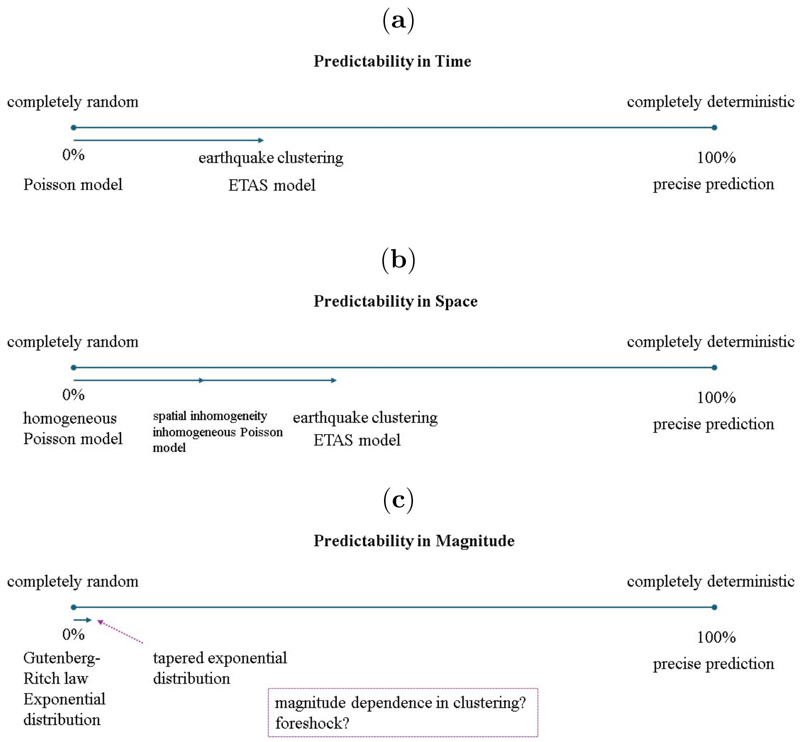

Figure 3: An illustration of earthquake predictability.

natural for forecasting and the evaluation of the forecasting performance can be completed in a measured and statistical manner using the probability gain framework. The ETAS model has especially become a de facto standard model, or null hypotheses, for comparison with other models and ideas. This suggests that improvements to our understanding of earthquake clustering can be quantified by developing new ideas into models and then comparing them to ETAS, or other models, using statistical hypothesis testing. Ultimately, the rigorous testing of forecast models is necessary to improve our ability to forecast seismic hazards.

Currently, owing to our inability to observe many of the fundamental processes of the system, as well as its inherent randomness, it is difficult to deterministically predict individual earthquakes. Therefore, statistical seismology, which places a larger focus on probabilistic forecasting, represents the best quantification method with respect to our state of knowledge. Furthermore, to provide more reliable earthquake forecast models, the challenge becomes the construction of models that can yield increased information gain with respect to a reference of the ETAS model.

With the rapid development of observation technologies, an increasing amount of observational data has been obtained. These new observations provide new theories and approaches to help us understand seismicity. Based on these new observations, seismologists can develop new methods to more efficiently analyze these data and new models to connect them to the earthquake process and tectonic environments, thus providing more knowledge on the earthquake occurrence process and and the ability to obtain reliable forecasts.

Acknowledgement

The authors thank David Vere-Jones for his long-term encouragement and generous support. This work was supported by supported by the Japan MEXT STAR-E project (JPJ010217) and Grants-in-Aid No. 19H04073 for Scientific Research from the Japan Society for the Promotion of Science (JSPS).

References

Console, R., M. Murru, and A. M. Lombardi (2003), Refining earthquake clustering models, *Journal of Geophysical Research*, *108*(B10), 2468, doi:10.1029/2002JB002130.

Field, E. H., K. R. Milner, J. L. Hardebeck, M. T. Page, N. van der Elst, T. H. Jordan, A. J. Michael, B. E. Shaw, and M. J. Werner (2017), A Spatiotemporal Clustering Model for the Third Uniform California Earthquake Rupture Forecast (UCERF3]ETAS): Toward an Operational Earthquake Forecast, *Bulletin of the Seismological Society of America*, *107*(3), 1049–1081, doi:10.1785/0120160173.

Geller, R. G., D. D. Jackson, Y. Y. Kagan, and F. Mulargia (1997), Earthquakes cannot be predicted, *Science*, *275*, 1616–1617.

Guo, Y., J. Zhuang, and S. Zhou (2015a), A hypocentral version of the space–time ETAS model, *Geophysical Journal International*, *203*(1), 366, doi:10.1093/gji/ggv319.

Guo, Y., J. Zhuang, and S. Zhou (2015b), An improved space-time ETAS model for inverting the rupture geometry from seismicity triggering, *Journal of Geophysical Research: Solid Earth*, *120*(5), 3309–3323, doi:10.1002/2015JB011979, 2015JB011979.

Gutenberg, B., and C. F. Richter (1944), Frequency of earthquakes in California, *Bull. Seis. Soc. Am.*, *34*, 184–188.

Hainzl, S., A. Christophersen, and B. Enescu (2008), Impact of earthquake rupture extensions on parameter estimations of point-process models, *Bulletin of the Seismological Society of America*, *98*(4), 2066–2072, doi:10.1785/0120070256.

Hawkes, A. G. (1971a), Spectra of some self-exciting and mutually exciting point processes, *Biometrika*, *58*(1), 83–90, doi:10.1093/biomet/58.1.83.

Hawkes, A. G. (1971b), Point spectra of some mutually exciting point processes, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *33*(3), 438–443.

Hawkes, A. G., and D. Oakes (1974), A cluster process representation of a self-exciting process, *Journal of Applied Probability*, *11*(3), 493–503.

Helmstetter, A., G. Ouillon, and D. Sornette (2003), Are aftershocks of large californian earthquakes diffusing?, *Journal of Geophysical Research*, *108*(B10), 2483, doi:10.1029/2003JB002503.

Huang, Q., M. Gerstenberger, and J. Zhuang (2016), Current challenges in statistical seismology, *Pure and Applied Geophysics*, *173*(1), 1–3, doi:10.1007/s00024-015-1222-7.

Kagan, Y. (1991), Likelihood analysis of earthquake catalogues, *J. of Geophys. Res.*, *106, Ser. B7*, 135–148.

Kagan, Y. Y. (1982), Stochastic model of earthquake fault geometry, *Geophys. J. Roy. Astr. Soc.*, *71*, 659–691.

Kendall, D. G. (1949), Stochastic processes and population growth, *Journal of the Royal Statistical Society: Series B (Methodological)*, *11*(2), 230–264, doi:https://doi.org/10.1111/j.2517-6161.1949.tb00032.x.

Lombardi, A. M., M. Cocco, and W. Marzocchi (2010), On the increase of background seismicity rate during the 1997–1998 Umbria-Marche, central Italy, sequence: apparent variation or fluid-driven triggering?, *Bulletin of the Seismological Society of America*, *100*(3), 1138–1152, doi:10.1785/0120090077.

Musmeci, F., and D. Vere-Jones (1992), A space-time clustering model for historical earthquakes, *Annals of the Institute of Statistical Mathematics*, *44*, 1–11, 10.1007/BF00048666.

Ogata, Y. (1981), On Lewis' simulation method for point processes, *IEEE Transactions on Information Theory*, *IT-27*(1), 23–31.

Ogata, Y. (1988), Statistical models for earthquake occurrences and residual analysis for point processes, *Journal of the American Statistical Association*, *83*(401), 9–27, doi:10.1080/01621459.1988.10478560.

Ogata, Y. (1998), Space-time point-process models for earthquake occurrences, *Annals of the Institute of Statistical Mathematics*, *50*(2), 379–402, doi:10.1023/A:1003403601725.

Ogata, Y. (2004), Space-time model for regional seismicity and detection of crustal stress changes, *Journal of Geophysical Research*, *109*(B3), B03,308, doi:10.1029/2003JB002621.

Ogata, Y., L. M. Jones, and S. Toda (2003), When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California, *Journal of Geophysical Research*, *108*(B62318), doi:doi:10.1029/2002JB002009.

Olson, E. L., and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature*, *438*, 212–215, doi:10.1038/nature04214.

Olson, E. L., and R. M. Allen (2006), Is earthquake rupture deterministic? (Reply), *Nature*, *442*, 6, doi:10.1038/nature04964.

Omori, F. (1894), On the aftershocks of earthquakes, *Journal of the College of Science, Imperial University of Tokyo*, *7*, 111–200.

Rathbun, S. L. (1993), Modeling marked spatio-temporal point patterns, *Bulletin of the International Statistical Institute*, *55, Book 2*, 379–396.

Rydelek, P., and S. Horiuchi (2006), Earth science: Is earthquake rupture deterministic?, *Nature*, *442*, E5–E6, doi:10.1038/nature04963.

Utsu, T. (1957), Magnitude of earthquakes and occurrence of their aftershocks, *Zisin (J. Seismol. Soc. Jap.)*, *10*, 35–45 (in Japanese).

Utsu, T., and A. Seki (1955), A relation between the area of aftershock region and the energy of the mainshock, *Zisin*, *7*, 233–240.

Vere-Jones, D. (1970), Stochastic models for earthquake occurrence, *J. Roy. Stat. Soc. Series B (Methodological)*, *32*(1), 1–62 (with discussion).

Vere-Jones, D. (1973), The statistical estimation of earthquake risk, *New Zealand Statistician*, *8*, 7–16.

Vere-Jones, D. (1976), A branching model for crack propagation, *Pure and Applied Geophysics*, *114*(4), 711–725, doi:10.1007/BF00875663.

Vere-Jones, D. (1977), Statistical theories of crack propagation, *Journal of the International Association for Mathematical Geology*, *9*(5), 455–481, doi:10.1007/BF02100959.

Vere-Jones, D. (2005), A class of self-similar random measure, *Advances in Applied Probability*, *37*(4), pp. 908–914.

Yamanaka, Y., and K. Shimazaki (1990), Scaling relationship between the number of aftershocks and the size of the main shock, *Journal of Physics of the Earth*, *38*(4), 305–324, doi:10.4294/jpe1952.38.305.

Zhuang, J. (2015), Weighted likelihood estimators for point processes, *Spatial Statistics*, *14*(B), 166–178, doi:http://dx.doi.org/10.1016/j.spasta.2015.07.009.

Zhuang, J. (2023), Statistical seismology, in *Encyclopeida of Earth Sciences Series: Encyclopedia of Mathematical Geosciences*, edited by B. S. Daya Saga, M. J. Cheng, Q., and F. Agterberg, Springer Nature Switzerland AG, Cham, doi:10.1007/978-3-030-26050-7_34-1.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002), Stochastic declustering of space-time earthquake occurrences, *Journal of the American Statistical Association*, *97*(3), 369–380.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2004), Analyzing earthquake clustering features by using stochastic reconstruction, *Journal of Geophysical Research*, *109*(3), B05,301, doi:10.1029/2003JB002879.

Zhuang, J., D. Wang, and M. Matsu'ura (2016), Features of earthquake source process simulated by Vere-Jones' branching crack model, *Bulletin of the Seismological Society of America*, *106*(4), 1832, doi:10.1785/0120150337.

Zhuang, J., M. Murru, G. Falcone, and Y. Guo (2019), An extensive study of clustering features of seismicity in Italy from 2005 to 2016, *Geophysical Journal International*, *216*(1), 302–318, doi:10.1093/gji/ggy428.

Zhuang, J., M. Matsufura, and P. Han (2021), Critical zone of the branching crack model for earthquakes: Inherent randomness, earthquake predictability, and precursor modelling, *European Physics Journal Special Topics*, *230*, 409–424, doi:10.1140/epjst/e2020-000272-7.