ANNALS OF GEOPHYSICS, 67, 4, S428, 2024; doi:10.4401/ag-9153
OPEN ACCESS

How the ETAS models were created, used, and evolved – Personal views and perspectives

Yosihiko Ogata*,1

(1) The Institute of Statistical Mathematics, 10-3 Midori-Cho, Tachikawa, Tokyo, Japan 190-8562

Article history: received September 19, 2024; accepted October 10, 2024

Abstract

Statistical seismology originated in Japan 130 years ago. Through the marriage with stochastic point processes in about half a century, we can now provide on-line estimation, real-time forecasting and direct diagnosis of seismic activity. This manuscript describes the origin and recent development of the ETAS models, their relationship to forecasting problems, and their diagnostic studies of the physical phenomena of seismic activity. I will take this opportunity to focus on my research experiences and views.

Keywords: AIC; ABIC; seismic activity; point process models; ETAS model: non-stationary ETAS model; space-time ETAS model; hierarchical space-time ETAS model; Bayesian optimization method; probability forecasts and their evaluation; Pre-discrimination of foreshocks. Penalized log likelihood; Maximum a posteriori (MAP) solution.

1. Introduction

I have always tried to demonstrate the usefulness of statistical methods by expressing many empirical laws and seismological hypotheses based on seismic statistics as statistical point process models. In this manuscript, I hope to convey the development of statistical seismology, which is closely related to the study of seismic activity, and its important role in earthquake prediction. I will follow the lead of the papers by Vere-Jones (2001, 2006) and would like to present my own research experience and views on how I have inherited the achievements of our predecessors. Therefore, it should be noted in advance that the references cited in this paper are mainly those of myself and my collaborators or are also those subjectively chosen. Some of the references cited in this paper are written in Japanese, including Report of the Coordinating Committee for the Earthquake Prediction (CCEP), but since artificially intelligent transformation software is making good progress these days, please refer to those reports as well if the readers are interested in. Readers might also be interested in an evaluation report of our research results up to about 2006 by an external committee (The Institute of Statistical Mathematics, 2006).

1.1 Statistical Point Processes

The late 1960s was a period of great exchange on point processes in the fields of probability theory, information theory and applications (see Lewis ed., 1972). Point processes have been studied extensively as a mathematical

representation of series of events and appositional marks (scalar quantities and vectors such as jumps or spatial locations) (e.g., Daley and Vere-Jones, 2003, 2008).

In this manuscript, we fix the lower earthquake magnitude threshold M_c, and consider all events with earthquakes above M_c. A new statistical point process model, other than the traditional Poisson processes and the renewal processes etc., the trigger model (Vere-Jones and Davies, 1966) was developed for seismicity applications, as a realization of the Neyman-Scott clustering process (Neyman and Scott, 1958). It consists of the primary events (mainshocks) following a stationary Poisson process, the secondary events (aftershocks) being induced only by the mainshocks, and the seismic activity being a superposition of the two. This is a probabilistic model that reflects the common knowledge of the seismological community at the time, but it was difficult to specify exactly which events in the data were mainshocks and which were aftershocks. Therefore, setting it up involves combinatorial complexity, and it was extremely difficult to compute them with the likelihood function.

Therefore, analyses of clustering point processes used inference based on the method of statistical moments (Vere-Jones and Davies, 1966; Vere-Jones, 1970). A student of Hawkes, Adamopoulos (1976), computed the trigger model and the Hawkes model (Hawkes, 1971) using the spectral Whittle likelihood of quadratic moments of the point-process (Bartlett, 1963), but the accuracy is poor unless the sample size is sufficiently large.

1.2 The Arrival of Point Processes in Japan

David Vere-Jones of New Zealand, a pioneer in working with seismic data using point process theory and models, visited Tokyo for several months in 1976 at the invitation of Hirotugu Akaike at the Institute of Statistical Mathematics (ISM), and he also visited Tokuji Utsu and other Japanese seismologists. The circumstances of this visit are described by Vere-Jones (2006) and by Ogata (2018). David gave us a series of special lectures on point processes at our institute. Immediately after listening to his lectures, we found an opportunity to develop new models using the central concept of point process prediction: namely, the "conditional intensity function",

$$\lambda(t|H_t) = P\{\text{an event occurs in } [t, t+\Delta)|H_t\}/\Delta + o(\Delta) \tag{1.1}$$

which is the derivative of the conditional probability of an event occurring in the future $(t, t+\Delta$, where $H_t = \{t_1, \cdots, t_n; t_i < t\}$ is the sequence of events times the history up to time t.

By modeling the conditional intensity function, we can consider forecasting for the times of future events. For example, with the conditional intensity function (1.1), we obtain the probability distribution $F\big(t - t_n|H_{t_n}\big)$ in a period from the time of the last event t_n to the time of the next event, and its density function $f\big(t - t_n|H_{t_n}\big)$. The following ratio

$$\lambda(t|H_t) = f\big(t - t_n|H_{t_n}\big)/\big\{1 - F\big(t - t_n|H_{t_n}\big)\big\}. \tag{1.2}$$

is called the Hazard equation. It has been shown that all general point processes are characterized by a conditional intensity function (Liptzel and Shryaev, 1978) by using the Hazard equation, with the exception of pathologically defined processes. Thus, solving Eq. (1.2) yields the distribution of time intervals between successive events corresponding to the given conditional intensity function.

Vere-Jones (1978) emphasized the importance of statistical models using conditional intensities of point processes and risk assessment around the metropolitan area, using a Japanese historical data around Kamakura area as an example, in a seminar given at the Earthquake Research Institute, the University of Tokyo, December 1976.

1.3 Likelihood and Simulation

In the early 1970s, the ISM time series research group (David called it the "Akaike school") began to work on the identification of the Hawkes self- and mutually excited point processes (Hawkes, 1971), analogous to the

auto-regressive (AR) model (Akaike, 1969) for point processes.

$$\lambda(t|H_t) = \mu + \sum_{t_i<t} g(t-t_i) = \mu + \int_0^t g(t-s)dN_s. \tag{1.3}$$

Incidentally, our interest in this model is described in an interview with Alan Hawkes (Hawkes and Chen, 2021): "Akaike was interested in this model because its conditional intensity function (1.3) is in the form of a linear regression of the time of occurrence of past events".

In fact, Akaike's interest was to estimate the parameter μ and the response function $g(\cdot)$ of the Hawkes point process using the maximum likelihood optimization method and use of the Akaike Information Criterion (AIC); (Akaike, 1973),

$$\text{AIC} = (-2)\ \text{max log likelihood} + 2 \times (\text{dimension of the fitted parameters}), \tag{1.4}$$

where the maximum likelihood estimate (MLE) of parametric model among competing models is adopted to predict the future data.

Fortunately, in the field of information theory, the likelihood function (Rubin, 1973) had just been derived theoretically. Thus, it was an immediate challenge for us to numerically obtain the MLE of μ and a parameterized function $g(\cdot)$.

On the advice of Hiro Akaike, Tohru Ozaki (1979) obtained the MLE of the coefficients numerically by maximizing the log-likelihood using an efficient quasi-Newtonian method such as the Davidn-Fletcher-Powell method, with $g(\cdot)$ as an exponential function and the sample data simulated by Eq. (1.2).

Inspired by this, I worked on three topics. The following two results were submitted together as a thesis. The other one was related to the modeling of seismic data, which is described in the next section.

(i) Large-sample theory of the MLE and likelihood ratio of general point processes characterized by conditional intensity functions, including the Hawkes-type models, as described above. In other words, the convergence of the MLE, the error distribution, and the chi-square distribution of the likelihood ratio statistic were proved using the ergodic property of the martingale and the central limit theorem (Ogata, 1978).

(ii) Proposed a new simulation method for point processes directly using the conditional intensity function. Ozaki's simulation method numerically solved the hazard equation, Eq. (1.2), iteratively until the convergence using Newton's method; but when the conditional intensity function becomes complex, such as in the multivariate case, the solution is not stable. Therefore, I focused on the thinning method of Lewis and Shedler (1979). This is a simulation method for non-stationary Poisson processes and is essentially the same as the conventional rejection sampling method, which generates a sample from an arbitrary density function. I generalized the thinning method and extended it to general multivariate (multichannel) point processes. The validity of the method was proved using martingale theory. Also, the simulations of point process models were performed to demonstrate its accuracy using likelihood ratio tests (Ogata, 1981). This method is now used as a standard method in many applications, including space-time point processes.

1.4 Statistical Causality Analysis

The third topic was my subsequent research on seismic activity up to the present. Instead of excluding aftershocks and clustering activity from seismic data (declustering), extensions of the Hawkes-type model

$$\lambda(t) = \mu + f(t) + C(t;T_0) + \sum_{t_i<t} g(t-t_i;M_i) + \sum_{u_j<t} h(t-u_j)\xi(u_j). \tag{1.5}$$

were used to analyze trends (e.g., long-term changes in data detection rates), periodicities (e.g., seasonal changes in precipitation), and causalities to determine the existence of induced effects from external data in other domains.

The external input data represent the times u_j and the labels $\xi(u_j)$, where the times $\{u_j\}$ can be evenly or irregularly spaced. Thus, the AIC was used to verify the seasonality of seismic activity in Japan and globally. Also, the causal relationship of seismic activity between shallow- and deep regions on plate boundaries (Ogata and Akaike, 1982; Ogata et al., 1982; Ogata and Katsura, 1986; Ogata, 1999). The self-excitation term $g(.\,;.)$ on the right side of Eq. (1.5) can be calculated efficiently in proportion to the number of earthquakes because a recursive structure can be used to calculate the likelihood function by using an exponential function multiplied by a polynomial, which is called a Laguerre-type function. The inverse power decay response function $g(.\,;.)$ appears in section 2.3.

For modeling the response function $h(.)$ of external force data, various specific applications have been considered; ground current anomalies (Zhuang et al., 2005), water injection volume into gas fields (Lei et al., 2013), volumetric strain meter data (Kumazawa et al., 2016), anomalies in ultra-low frequency magnetic data (Han et al., 2017), and slow slip moment ratios (Nishikawa, 2023). The causality model (1.5) can be used to investigate the statistical relationship between anomaly data and the precursory nature of large earthquakes.

2. Exploring Seismic Activity

2.1 Encounter with Statistical Seismology

In search of applied problems of point processes, I joined the Seismological Society of Japan since around 1980. I was particularly interested in the Seismic Activity Section meeting, where many reports on the analysis of microseismic activity were presented one after another. Statistical seismology, as a research field for describing and understanding earthquakes based on data, has been the main field of seismology in Japan since the Meiji era (1968-1912).

The main topics discussed in the field of statistical seismology since the synthesis paper by Aki (1956) include periodicity of earthquake occurrence, detailed study of aftershock activity, causal relationships with geophysical data such as pressure and tide level changes, epicentral movement, classification of seismic activity patterns, and comparison of different magnitude scales, b-values of G-R magnitude distributions. Many such studies on these have been conducted in Japan since the 1960s, and many of the results, especially on aftershock activity are still being referred to today. However, our main concern about the validity of the results of seismic activity analysis was the difficulty of statistical testing and reproducibility of hypotheses due to the clustering nature of earthquakes, such as aftershocks and cluster earthquakes.

On the other hand, statistical studies of aftershocks were ignored in the American seismological community, possibly because, unlike Japan, they do not have frequent large earthquakes. Richter did not appreciate aftershock studies and, contrary to Jefferys' view, regarded aftershocks as useless data garbage (private communication with Tokuji Utsu). They seemed to believe that the "mainshock" activity, after removing the garbage aftershocks, should be used to predict large earthquakes. This belief was carried over into statistical studies of the quiescence of seismicity, the seasonality of earthquake occurrence and synchronization with the Earth's tides, and the causality of anomalous phenomena from large earthquakes, in the second half of the 20th century. Therefore they performed various "de-clustering" methods (e.g., Matthews and Reasenberg, 1988; Wyss and Habermann, 1988; van Stiphout et al., 2012), which remove cluster members from the considered data in order to test the null hypothesis (i.e., stationary Poisson process) of the de-clustered data, and then make the scientific claim.

However, de-clustering methods vary and can lead to different results. This is not meaningful for the study of aftershock activity itself. In fact, long-term seismic activity is long-term memory dependent in time, with the autocorrelation of the inverse power decay in time, and fractal self-similarity in space (e.g., Mandelbrot and Wallis, 1969; Ogata and Abe, 1991; Ogata and Katsura, 1981; Guo and Ogata, 1997).

Even the global series of large earthquakes of magnitude 7 or greater (see Fig. 1) have a long memory history dependence (Ogata and Abe, 1991; Ogata, 2017b, 2021). However, the Perez and Scholz (1984) and Pacheco and Sykes (1992) made the completely erroneous assumption that such a series of the large earthquakes should be stationary Poisson processes and falsifying the magnitudes of the Abe catalog (Abe, 1981), which is based on the original Gutenberg's surface wave records. Similar concern are discussed by Di Giacomo and Storchak (2022) in comparison with the ISI-GEM catalog. Thus, I am concerned that the twentieth-century earthquake catalog of Engdahl and Villasenor (2002), which is based on the results of Perez and Scholz (1984) and Pacheco and Sykes (1992), is now widely used in earthquake engineering and used as the basic data for global earthquake damage history (Ogata, 2021).

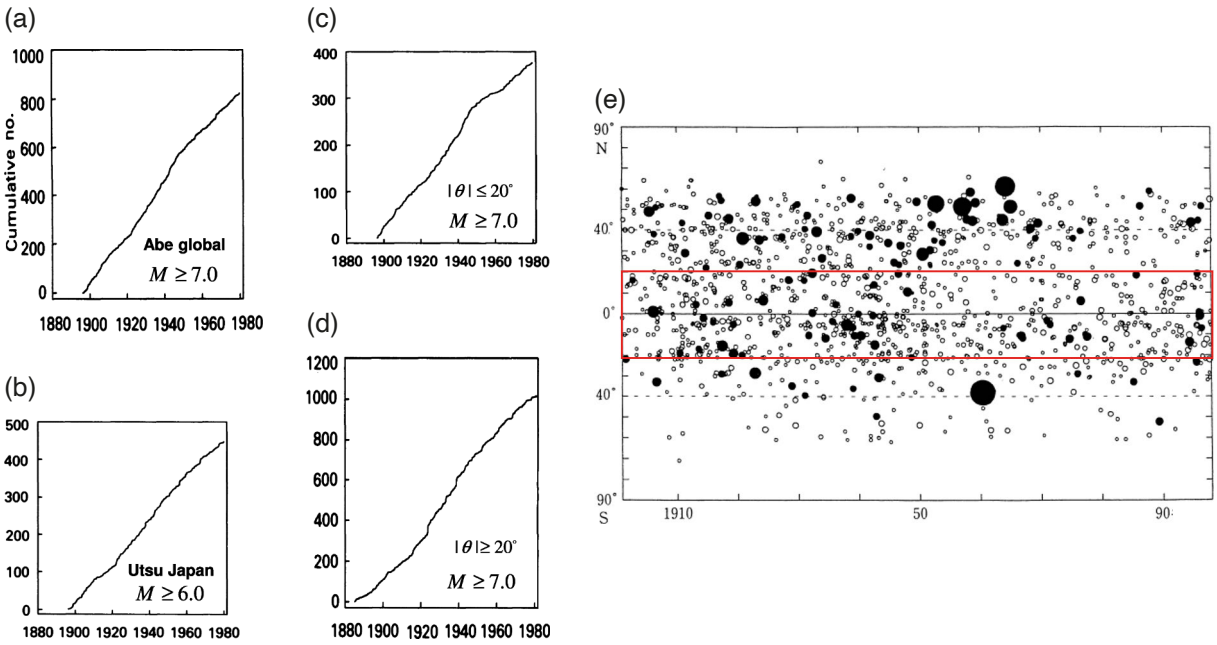


Figure 1. Cumulative number of shallow earthquakes (a) with $M_S \geq 7$ in the world for the period 1897-1980 (Abe, 1981); (b) with $M_{JMA} \geq 6$ in the regional area around Japan for 1885-1980 (Utsu, 1985); (c) and (d) $M_S \geq 7$ in the low-latitude region $\theta \leq 20°$ and high-latitude region $\theta \geq 20°$ of the world for the same period; (e) change over time in the epicentral longitude of the world's major earthquakes, black circles correspond to $M \geq 7.8$, white circles correspond to $7 \leq M \leq 7.8$, and circle size corresponds to M (Utsu, 1999, Fig. 12.1), which is mostly taken from to Abe (1981). The notable problems were that the cumulative curve in the whole world in (a) looks like a broken line, but the cumulative in the divided region in (b) and (d) are different. Those in (b) and (d) are similar despite different catalogs. Also, the space-time earthquake magnitude does not seem to be uniform. Thus, the same artificial magnitude bias of during the different periods over the world is difficult to hypothesize (see Ogata, 2021).

2.2 Development of Aftershock Studies in Japan

A large earthquake was followed by a series of small and medium-sized earthquakes, which later referred to as aftershocks. In Japan, the study of the large amount of aftershock data as important information has been instrumental in our understanding of the nature of seismic activity.

Aftershocks are extremely frequent immediately after the main shock and then gradually decrease over time, but it is difficult to return to the original level of activity even after a long period of time. Fusakichi Omori (Omori, 1894) was the first to discuss the quantitative relationship between aftershock frequency and the time course of aftershocks, noting that the attenuation of aftershock frequency from the 1891 Nobi (Mino-Owari) earthquake "did not fit well with an exponential function, as would be expected from the attenuation of physical phenomena, but it did fit well with a hyperbolic curve" over 10 years.

In 1957, Tokuji Utsu studied aftershocks in Japan and around the world and found that the decay rate of the aftershock frequency per unit time obeys

$$\nu(t) = K(t+c)^{-p}, \tag{2.1}$$

where t is the time elapsed since the main shock. A heuristic new method of Utsu was to plot the aftershock frequency $\nu(t)$ versus the elapsed time t on a double-logarithmic graph paper (see Fig. 2), and he discovered that the decay asymptotically rides on a straight line and obtained an estimate of the exponent p from the slope of the line. He then showed that the aftershocks of the Nobi earthquake continued to decay according to Eq. (2.1) for more than 100 ~ 130 years to the present (e.g., Utsu et al., 1995). Utsu called Eq. (2.1) the "modified Omori formula", but we refer to it as the "Omori-Utsu formula". The hyperbola mentioned by Omori refers to $p = 1$.

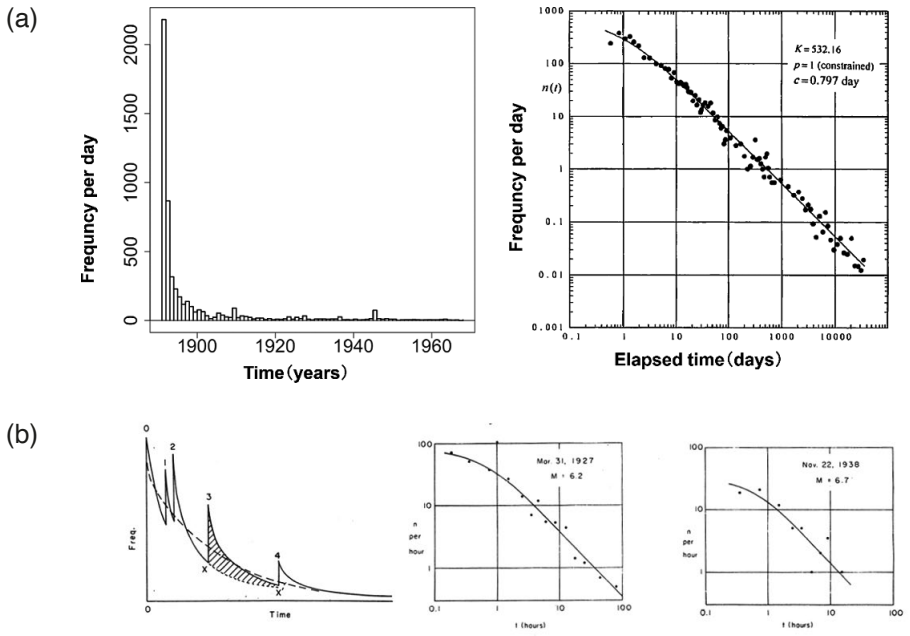


Figure 2. (a) The data for the first 10 years up to the end of the 19th century were compiled by Omori himself, and those for the remaining 95 years were compiled by Utsu (1985). Utsu used the double logarithmic plot for the aftershock study. The *p*-value estimate was obtained from the asymptotic slope of the plot, and the c-value is estimated from the degree of the curvature in the early part of the activity. This shows that the aftershock activity has continued for more than 120 years. (b) The lower panels show that the secondary aftershocks still obey the Omori-Utsu decay formula. These figures are taken from Utsu (1970). They were based on the inference by counting the number of aftershocks in time bins. Nowadays, the MLE method is used directly based on the occurrence times record (Ogata, 1983).

On the other hand, apart from the effect of missing aftershocks immediately after the mainshock, the interpretation of the coefficient *c* is still controversial. For example, in the case of a series with a complex rupture process (i.e., several triggered events are concentrated immediately after the main shock), the coefficient *c* may be related to the duration of the rupture and increases (Utsu, 1970; Utsu et al., 1995). I will discuss this again in section 4.2.

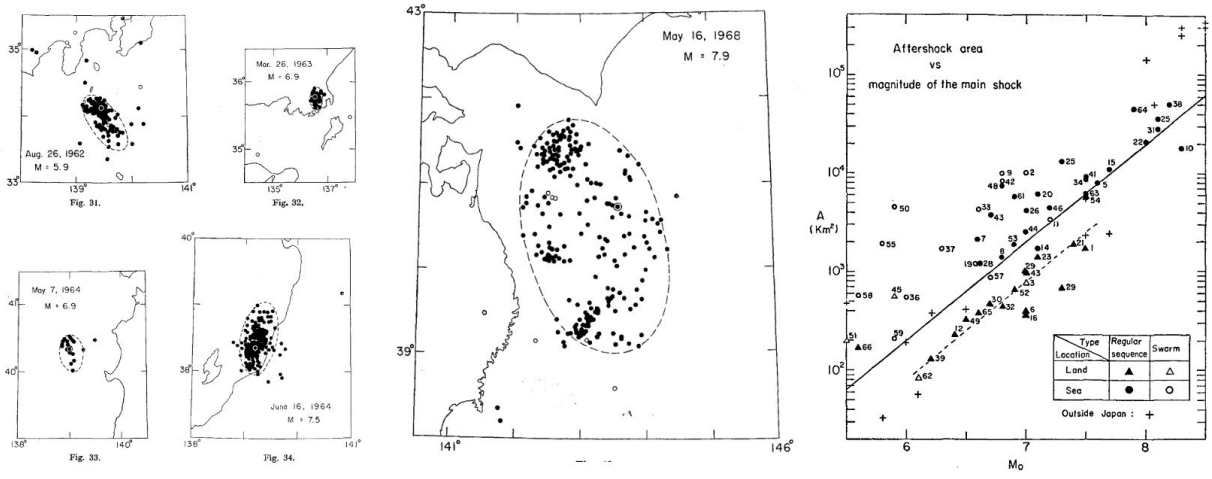


Figure 3. Five panels on the left: Ellipses contain 95% or more of the plotted aftershock epicenters within one month of the mainshocks falling within the ellipses. Right panel: Plot of aftershock area versus the mainshock magnitude. Open and closed circles and triangles represent large earthquakes in the region of Japan. The continuous black straight line is the linear regression indicating $\log Area = M - 3.7$. Note that the fitted dashed line for inland earthquakes, representing intra-plate earthquakes, is different from that for inter-plate earthquakes. These figures are taken from Utsu (1969).

Utsu (1970) also found secondary aftershocks (aftershocks of an aftershock, see Fig. 2) that increase dramatically after a large aftershock and then decay again, and found that such decay also follows the Omori-Utsu Eq. (2.1). This was a break with the conventional wisdom of the time that aftershocks were triggered only by mainshocks. In addition, Utsu and Seki (1954) and Utsu (1969) conducted a quantitative study of the relationship between the spatial extent of aftershocks and the magnitude of the main shock (Fig. 3); the empirical distribution of the magnitude differences between the main shocks and the largest aftershocks; and the statistical analysis of many other properties of aftershocks with respect to their various characteristics.

Ogata (1983) then considers Eq. (2.1) and its extension with secondary aftershocks as the (conditional) intensity function (1.1) of a non-stationary Poisson process and proposes to obtain the MLE of the parameters K, c, and p and the additional parameters together with their error estimates directly from the occurrence time records $\{t_i; i = 1, 2, \cdots\}$ above a certain size of M_c. This method also uses the AIC when model comparison is required (Ogata, 1983), when data immediately after the main shock are missing or when secondary aftershocks are included and so on. Today, when a moderate to large earthquake occurs, the predicted probability of aftershocks in time is calculated using Eq. (2.1) based on aftershock occurrence time data (e.g., Reasenberg and Jones, 1989; Hardebeck et al., 2024).

2.3 ETAS Model

In the mid-1980s, the Institute of Statistical Mathematics was reorganized from being under the direct control of the Ministry of Education into an inter-university research institute corporation. There, I organized a joint research seminar on "Mathematical Seismology" and a study group on complex systems and fractals with seismologists and conducted research on the characterization of seismic activity (Ogata and Katsura, 1991a; Guo and Ogata, 1997), including self-similarity and the maximum likelihood method for the fractal dimension (concentration of point clouds).

I sought a practical point-process model for predicting seismic activity directly from the source data, rather than merely interpreting them. Most of the clues are empirical aftershock laws and observations explored and established by the Utsu studies described in the previous section. The Epidemic-Type Aftershock Sequence (ETAS) model was developed to forecast earthquake occurrence rates above a certain threshold magnitude M_c

$$\lambda(t|H_t) = \mu + \sum_{t_i<t} \nu(t - t_i)\, e^{\alpha(M_i - M_c)}, \tag{2.2}$$

where $H_t = \{(t_i, M_i); t_i < t, M_i \geq M_c; i = 1,2, \cdots, N\}$ is the history of earthquake occurrence associated with the magnitude series. The response function $\nu(t)$ is the Omori-Utsu formula (2.1).

Depending on the magnitude size, larger earthquakes induce more aftershocks, and smaller earthquakes also induce a reasonable number of aftershocks. The conditional intensities at time t are the linear superpositions. The last exponential function in Eq. (2.2) reflects the spatial aftershock extent scaling of Utsu and Seki (1954), where the parameter α represents the magnitude efficiency, which when large results in a typical mainshock-aftershock type earthquake series, and when small results in the second type earthquake swarm in the sense of Utsu (1970a,b), i.e., a number of relatively large earthquakes occur with their aftershocks in a short period. The first type earthquake swarm will be discussed later in connection with the non-stationary ETAS model in section 2.5.

The parameter μ is called the background seismic activity rate, which is the intensity of earthquake occurrence specific to a region. However, some of the background seismic activity may include contributing aftershocks from older large earthquakes that have not fully decayed. The five parameters $\theta = (\mu, K, c, \alpha, p)$ of the ETAS model (Ogata, 1985, 1988, 1989) consist of the likelihood function

$$\log L(\theta) = \sum_{0<t_i<T} \log \lambda_\theta\left(t_i|H_{t_i}\right) - \int_0^T \lambda_\theta(t)dt. \tag{2.3}$$

Maximizing this function with respect to the parameters yields the MLE (see Fig. 4).

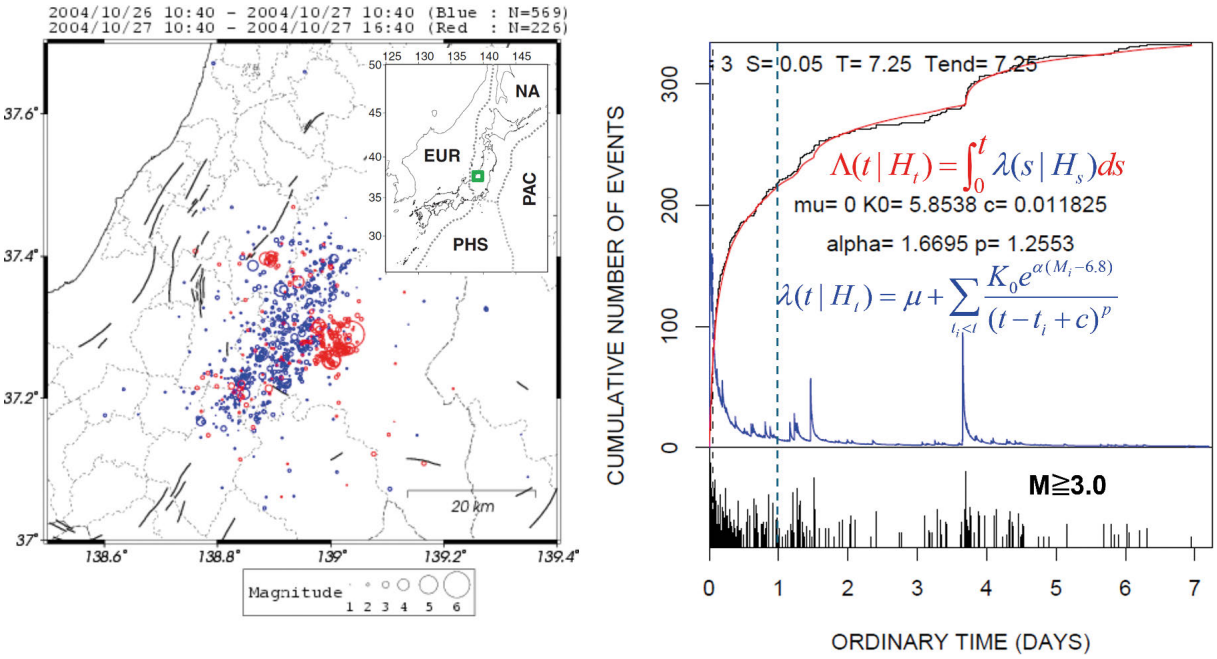


Figure 4. Left panel: Aftershock epicenter map of the 2004 M6.8 Niigata-Chuetsu earthquake. Right panel: Aftershock sequence up to about one week after the main shock. The blue spiky curve is the conditional intensity rate of the estimated ETAS model, which becomes a red increasing function when integrated over time from the start. It nearly overlaps with the empirical cumulative function (black curve) of the aftershocks over the entire period.

The ETAS model is often used to simulate standard seismic activity on a region-by-region basis; a faster simulation method than the general point process simulation method described in Section 1.3 can be implemented through the branching process interpretation of the ETAS model (Zhuang et al., 2004).

In addition, the log-likelihood calculation of the ETAS model requires $O(n^2)$ operations for n number of seismic data in the first term of (2.3), but Ogata et al. (1993) used a recursive structure (Ogata and Akaike, 1982) for numerical integration with a double exponential integral approximation formula, which is accurate and fast with $O(n)$ operations (Ogata, 2006a). A FORTRAN code and manual implementing this procedure are available from Ogata (2006c). A similar version is available in the R package from the Institute of Statistical Mathematics (2023).

2.4 Diagnosis using the ETAS model

Using the estimated ETAS model as a benchmark, we can find anomalous changes in actual seismic activity that cannot be explained by the model (Ogata, 1988). To see graphically whether seismic activity and aftershock activity are trending as forecasted, the expected cumulative number of earthquakes

$$\Lambda(t) = \int_0^t \lambda(s|H_s)ds \tag{2.4}$$

that will occur by time t is plotted. Then, instead of comparing the empirical cumulative step function $\{(t_i, i); i = 1,2,\cdots,N\}$with the theoretical expected curve $\{(t_i, \Lambda(t_i)); i = 1,2,\cdots,N\}$, the cumulative function of $\{(\Lambda(t_i), i); i = 1,2,\cdots,N)\}$ can show the clear discrepancy between the predicted cumulative curve and the empirical data (e.g., Fig. 5). In fact, if the occurrence data is generated by the ETAS model, the latter cumulative function will be on a straight line with a slope of 1.0.

A qualitative recognition of the changes in the seismicity pattern can be obtained by simply comparing the two-stage ETAS models before and after a change point. In fact, we first apply the ETAS model to the first part of the data before the change point, and then forecast the data after the change point using the same MLE parameters. The transformed data may then deviate significantly downward or upward, indicating relative quiescence or activation.

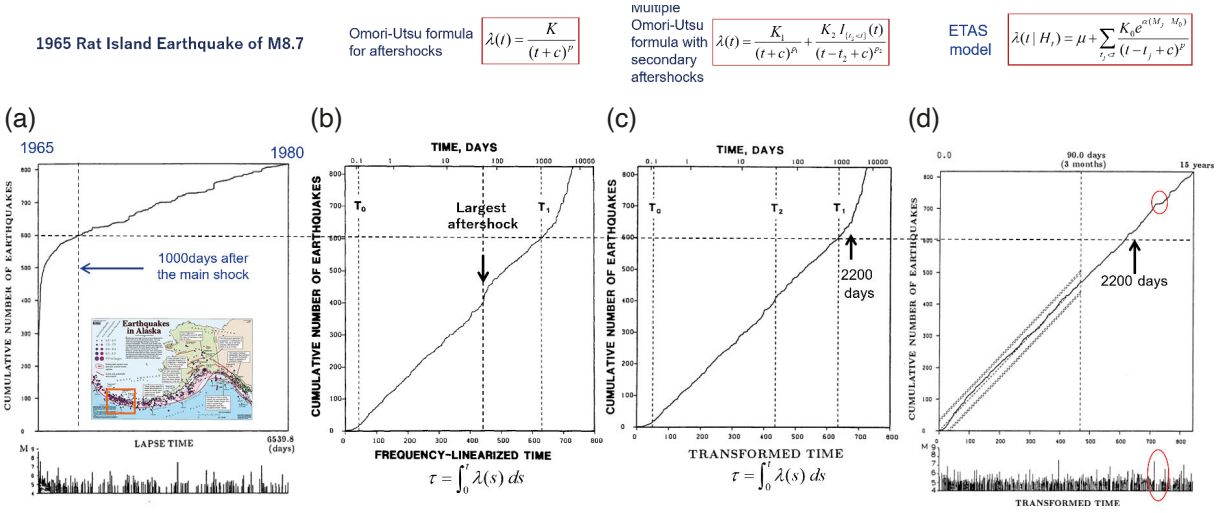


Figure 5. Example of diagnostic analysis. (a) Aftershock sequence from the great Rat Islands earthquake in the Aleutians. (b) If we fit an Omori-Utsu intensity rate for the first 1000 days and transform the time by the expected cumulative function, we get this cumulative plot. The deviation after the arrowed time is due to the largest aftershock, which is accompanied by additional aftershocks, called secondary aftershocks. (c) Thus, using the multiple Omori-Utsu model, we get a cumulative function that is almost on a straight line up to 2200 days, but then deviates upward. (d) If we fit the ETAS model, the graph is basically straight to the end, representing 15 years. The orange circled area is curious. If you look at the magnitude, it is a large earthquake, but there are almost no aftershocks. A nuclear test? At least, we confirmed that it was a real earthquake.

There are many anomaly findings by residual analysis as follows (Kumazawa et al., 2010, 2017, 2020; Ogata, 1992, 1999, 2001b, 2001, 2005a,b,c, 2006a, 2007a,b, 2011a, 2017a, 2017d; Ogata et al., 2003a; Ogata and Toda, 2010; Ogata and Tsuruoka, 2016).

In general, a diagnostic analysis is considered successful when the process of data generation is well understood and a residual series with anomalous variation of scientific significance can be extracted. To confirm its significance, we use AIC to compare models (Ogata, 1992, 1999; Kumazawa et al., 2010, 2013). For example, the model with ΔAIC smaller models than the other model has exp{−ΔAIC/2} times more likely to be accepted relative to the other model (Akaike, 1985; Ogata, 2017). After confirming the seismicity anomalies relative to the ETAS model, see Section 3.1 for our physical studies that caused the anomalies.

When multiple catalogs are available, as in the case of the 2004 M9 Sumatra earthquake, we can make a more careful study. Indeed, I found that the USGS catalog and the ISC catalog gave different results (Bansal and Ogata, 2013; Ogata, 2021), namely, the activation and quiescence of the seismicity relative to the respective ETAS models. In fact, I found that the ISC catalog contains a magnitude shift for several years before the M9.0 event. The difference in the compilation between the ISC and NEIC catalogs after 1996 is due to the acceptance/rejection of amplitude data from the International Data Center (IDC) of the Comprehensive Nuclear Test Ban Treaty Organization (CTBTO). See Bansal and Ogata (2013) and Ogata (2021) for more details. Thus, we ultimately inferred relative activation rather than relative quiescence.

For the physical interpretation of the relative quiescence or activation obtained by the diagnostic ETAS analysis, see Section 3.1.

2.5 Non-stationary ETAS Model

Transient swarms of seismic activity mixed with normal seismic activity may not be well represented by the stationary ETAS models. Llenos et al. (2009), Nishikawa et al. (2021), and Mitsui et al. (2021, 2022) examined the ETAS anomalies to discuss slow slip, fault creep and magma intrusion-related seismic activity by identifying abnormal seismic activity in the diagnostic analysis.

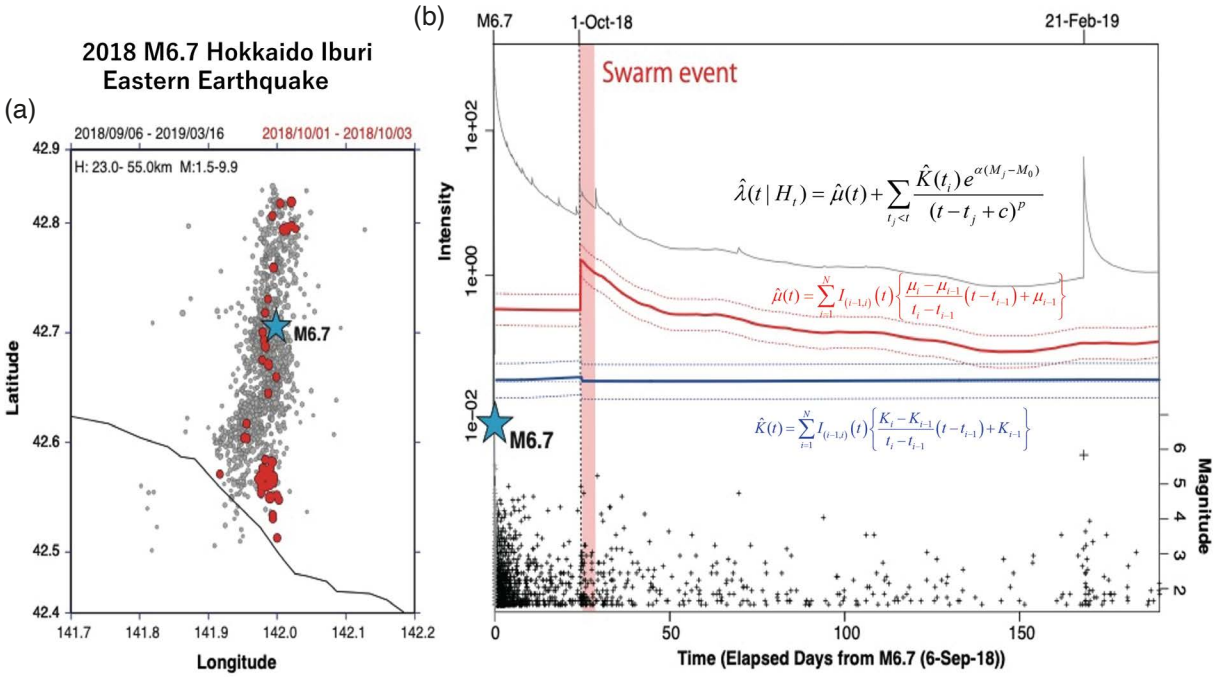


Figure 6. Aftershocks of the 2018 M6.7 Hokkaido Iburi-East Earthquake. (a) Epicenters of M ≥ 1.5 from M 6.7 on September 6, 2018 to March 16, 2019. The red disks show the epicenters for 4 days period from October 1, 2018, which is 24 days after the main shock. (b) The non-stationary ETAS model estimation for M ≥ 2.0 events. The red and blue curves show the background intensity and the aftershock productivity with 95% error, respectively. Events during the period of the pink band correspond to the red disk events in panel (a). Note that the background rate of the ETAS model increased sharply without being triggered by a significantly large aftershock.

It is difficult for the stationary ETAS model to fit transient earthquake swarm well. Therefore, Hainzl and Ogata (2005) and Kumazawa and Ogata (2013) used a non-stationary ETAS model of the form,

$$\lambda(t|H_t) = \mu(t) + \sum_{t_i<t} \frac{K(t_i)e^{\alpha(M_i-M_c)}}{(t-t_i+c)^p}, \tag{2.5}$$

where the background intensity rate $\mu(t)$ and the aftershock productivity $K(t)$ are time-dependent. Kumazawa and Ogata (2013, 2024a) assume that the functions are piecewise linear functions (line graphs) connecting the respective coordinates $(t_i, \mu_i)$and (t_i, K_i) where $\mu_i = \mu(t_i)$ and $K_i = K(t_i)$ for earthquake i (e.g., Fig. 6).

The $\mu(t)$ function could reflect stress changes due to, for example, large ruptures outside the region, slow slip around the adjacent region, or dynamic triggering by pore pressure intrusion within the faults. The function $K(t_i)$ represents the rupture chain effect (aftershock induction efficiency) of a past earthquake that occurred in the region at time t_i. The other parameters α, c, and p are assumed to be time-invariant and are determined according to the surrounding tectonic seismic activity.

To estimate the changes in the time functions involved in the non-stationary ETAS model (2.5), we need the number of the unknown coefficients to be twice the number of earthquake data. On the other hand, since the conditional intensity function (2.5) is linear with respect to the coefficients, the log-likelihood function (2.3) is convex and unimodal with respect to the coefficients (Ogata, 1978), which has the advantage that the inverse analysis can be numerically stable regardless of high dimensionality. Together with the smoothness constraints of the penalty function, the penalized log-likelihood function (Good and Gaskins, 1971) is as follows

$$\begin{aligned} & \log PL\left(w_\mu, w_K; \mu(t), K(t), c, \alpha, p\right) \\ = \quad & \log L\left(w_\mu, w_K; \mu(t), K(t), c, \alpha, p\right) - \sum_i w_\mu \left(\mu_i - \mu_{i-1}\right)^2/2 - \sum_i w_K (K_i - K_{i-1})^2/2 \end{aligned}$$

which has the advantage of having a unique maximum a posteriori (MAP) solution so that the inverse analysis can be robust regardless of high dimensionality. Kumazawa and Ogata (2013, 2014) proposed a method to invert the time variation of the model parameters using a Bayesian smoothing method to determine appropriate weights w_μ and w_K. The Akaike Bayesian Information Criterion (ABIC) (Akaike, 1980) is used to objectively estimate these weights (hyperparameters). In addition, the reproducibility of the data was verified by simulation experiments (Kumazawa and Ogata, 2013).

Predicting the termination period of active swarm activity from the background rates of non-stationary ETAS is particularly useful information for local people in the areas of geothermal or active water-induced seismicity.

2.6 Space-time ETAS models

The space-time ETAS model, an extension of the temporal ETAS model (2.2), includes the epicenter variable (xi, yi), predicts the spatial propagation of earthquake clusters in real time, showing where the risk is at the next moment and where and how it has been active in the past. In a general form,

$$\lambda_\theta(t,x,y|H_t) = \mu(x,y) + \sum_{\{i;t<t_i;M\geq M_c\}} \frac{K}{(t-t_i+c)^p} g(x,y;x_i,y_i,M_i)$$

predicts earthquake occurrence rates above a certain cutoff magnitude M_c, where $H_t = \{(t_i, x_i, y_i, M_i): t_i < t, M_i \geq M_c: i = 1,2,\cdots,N\}$ is the history of past earthquake occurrence associated with the magnitude series.

Various spatial factors $g(x,y;x_i,y_i,M_i)$ have been considered in this context. For example, Kagan (1991), Musmeci and Vere-Jones (1992), Ogata (1993), Rathbun (1993, 1996), Lombardi and Marzocchi (2010a,b) Falcone et al. (2010) and so on. Among them, different isotropic spatial factors have been considered such as Gaussian-type short-range decay with distance from the epicenters, inverse power decay with distance and their mixtures.

By the AIC comparison for goodness of fit, Ogata (1998, 2013, 2015) and Ogata and Zhuang (2006) show that the following mixed type is the best fit among the three types of spatial factors proposed by Ogata (1993) for seismicity in and around Japan. In addition, this is extended for possible anisotropic triggering spatial factors of the centroid coordinate as shown in Figs. 3 and 7, which is practically the best fit performance compared with the AIC. Finally, the following model is obtained

$$\lambda_\theta(t,x,y|H_t) = \mu\cdot v(x,y) + \sum_{\{i;t<t_i\}} \frac{K}{(t-t_i+c)^p}\left[\frac{(x-\bar{x}_i, y-\bar{y}_i)S_j^{-1}(x-\bar{x}_i, y-\bar{y}_i)^t}{e^{\alpha(M_i-M_0)}} + d\right]^{-q} \tag{2.6}$$

which is the best for practical application to the data sets in and around Japan. Here, before applying the maximum likelihood procedure to the space-time ETAS model (2.6), we need to prepare the following treatment. Namely, the two-dimensional normal distributions are applied to the epicenter data of all detected aftershocks immediately following a predetermined large earthquake (e.g., either within 1 day for estimation or within 1 hour for prediction; see Fig. 7). Specifically, AIC is used to select the best fit of the mean vector $(\bar{x}_j, \bar{y}_j)$ and the variance-covariance matrix S_j. This is done to predict the anisotropy of the aftershock pattern using the appropriate elliptical approximation and the centroid coordinate of the aftershock domain. For all other earthquakes smaller than the predetermined magnitude, we keep the original epicentral coordinate (x_i, y_i) and the unit matrix I for the matrix S_j. Finally, I obtain the MLE for seven parameters (μ, K, c, α, p, d, q). The FORTRAN codes to implement the present procedure and the English manual are available (Ogata et al., 2021).

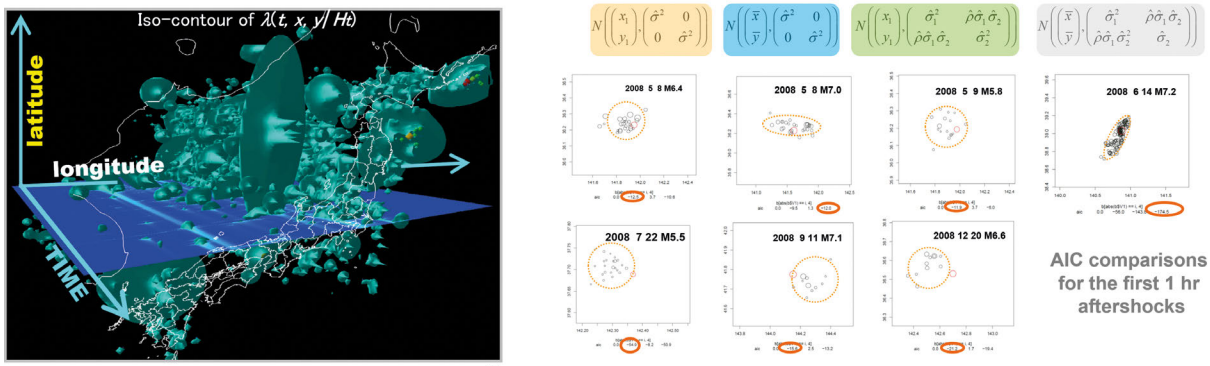
2.7 Hierarchical Space-Time ETAS Model

As more and more small earthquakes are recorded and databases become richer, the regional nature and complexity of earthquake occurrence patterns become more pronounced, and detailed statistical understanding and prediction become more difficult. To extract essential and reliable information, it is necessary to consider

$$\frac{\Pr\{an\ event\ in [t+\Delta t)\times[x+\Delta x)\times[y+\Delta y) \mid H_t\}}{\Delta t \Delta x \Delta y} \approx \lambda(t,x,y \mid H_t) \qquad \text{where} \quad H_t = \{(t_j, x_j, y_j, M_j);\ t_j < t\}$$

Space-Time ETAS model

$$\lambda_\theta(t,x,y \mid H_t) = \mu \cdot \nu(x,y) + \sum_{\{j:\, t_j<t\}} \frac{K}{(t-t_j+c)^p} \left\{ \frac{Q_j(x-\bar{x}_j, y-\bar{y}_j)}{e^{\alpha M_j}} + d \right\}^{-q} \qquad \text{where} \quad Q_j(x,y) = (x-\bar{x}_j, y-\bar{y}_j) S_j^{-1} \begin{pmatrix} x-\bar{x}_j \\ y-\bar{y}_j \end{pmatrix}$$



Log likelihood: $\log L(\theta) = \sum_{i=1}^{N} \log \lambda_\theta(t_i, x_i, y_i \mid H_{t_i}) - \int_0^T \iint_A \lambda_\theta(t,x,y \mid H_t)\, dx dy dt$

Figure 7. Space-Time ETAS Model. Left panel: An illustration of the iso-surface of the conditional intensity (occurrence rates) of the Space-Time ETAS Model. Given a history of earthquake occurrences prior to time t, the conditional intensity function is defined by the occurrence rate at time t and location (x, y). Right panels: The spatial triggering effect S_j is defined for some pre-specified large earthquake j. For the anisotropy of the aftershock regions, I estimate the variance-covariance matrix for the first short period after the mainshock using the AIC selection procedure. This model has 7 parameters to be estimated, and we maximize the log-likelihood function to obtain the MLE. See Ogata (1998, 2013, 2015) and Ogata and Zhuang (2006).

heterogeneous prediction models in time and space, and studies using these statistical models have become inevitable. Hierarchical Bayesian methods are needed to handle such large models with a large number of unknown coefficients of the parameters (2.6) for inverse problems and prediction.

The characteristic parameters (μ, K, α, p, q) of the space-time ETAS model (2.6) were represented by location-dependent functions to describe the spatial heterogeneity of seismic activity (Ogata et al., 2003). This model was referred to as the *full* Hierarchical Space-Time ETAS (*full* HIST-ETAS) model, which allows detailed regional characteristics to be obtained (Ogata et al., 2003; Ogata, 2004, 2011b,c). When only characteristic parameters (μ, K) are dependent, then it is called a *restricted* HIST-ETAS model.

The parameters are represented by piecewise linear functions (Delaunay surface functions) on a Delaunay triangular partition connecting the coordinates of the epicenters of adjacent earthquakes, and can represent highly accurate changes, especially in earthquake concentration areas (see Fig. 8).

Similarly, Ogata et al. (2019) constructed and estimated a hierarchical space-time ETAS model with a 3D Delaunay tetrahedron partition in a rectangular parallelepiped volume beneath the Tokyo metropolitan area to a depth of 100 km (see Fig. 9). Also, this model is added by the external effects in the 3D volume from the M9 Tohoku earthquake.

However, the total parameter coefficients of the HIST-ETAS model are several times (up to five times) larger than the data size. For stable optimal estimation of the model, the variability of the Delaunay surface function coefficients must be limited by imposing a smoothing penalty for large fluctuations, so the optimal prior distribution of the constraints among the parameter coefficients must be obtained to optimize the constraining weights (hyperparameters) using empirical Bayesian methods and the Akaike Bayes information criterion. Therefore, a penalized log-likelihood function (Good and Gaskins, 1971) is used to smooth the Delaunay surface functions instead of the line graphs (2.6). The coefficients of the Delaunay surface functions that maximize the penalized log-likelihood are then determined. This procedure is equivalent to obtaining the MAP distribution as the tomographic inversion but requires the use of empirical Bayesian methods and the Akaike Bayesian Information Criterion (ABIC) (Akaike,

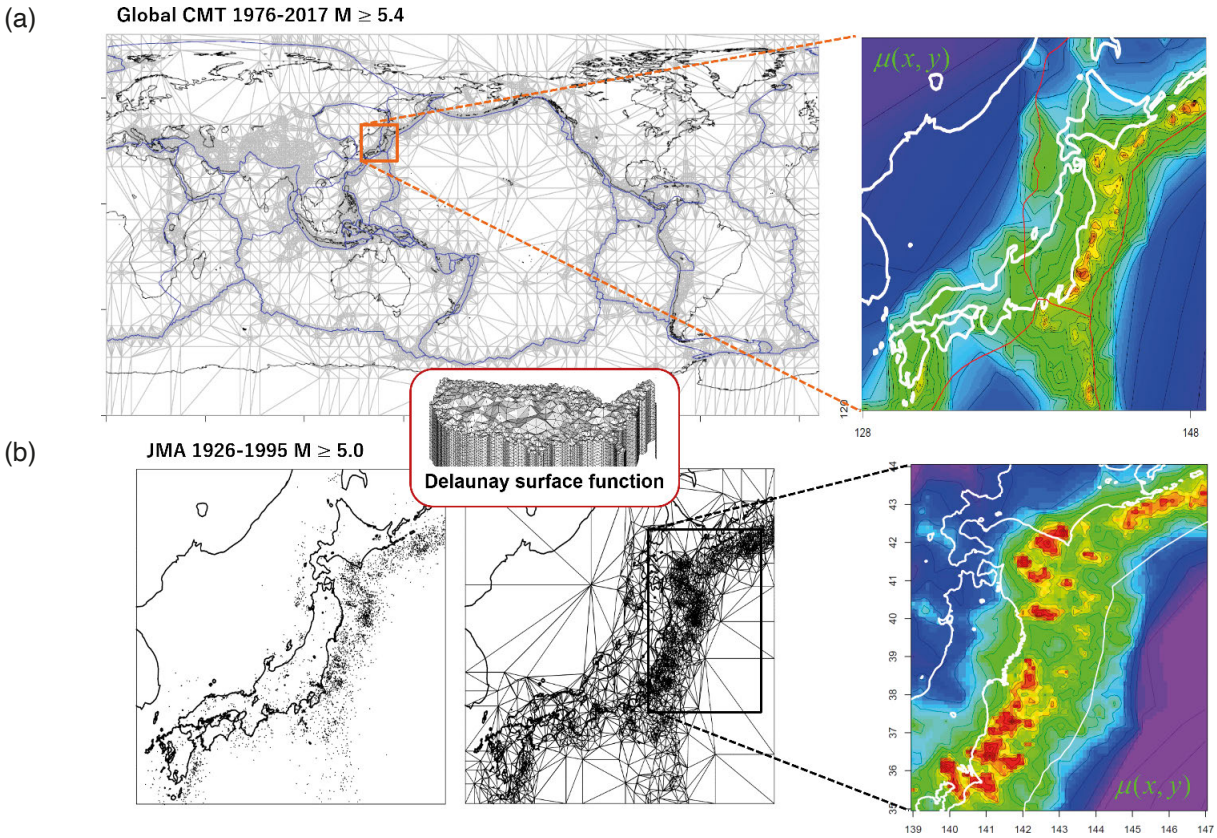


Figure 8. Delaunay triangulation mesh and optimal MAP solutions. (a) By constructing a Delaunay triangle tessellation mesh based on the location of the global earthquake (global CMT catalog $M_w \geq 5.4$) with some selected points, the HIST-ETAS model is optimized using ABIC. Restricting the optimal MAP inversion solution to the area around the Japanese archipelago and linearly complementing it, the image of $\mu(x,y)$ is shown in the upper right panel. (b) Similarly, an interpolated view of the HIST-ETAS optimal MAP solution can be obtained from the Delaunay triangle tessellation for the JMA data ($M_{JMA} \geq 5$) around the Japanese islands, but at a higher resolution.

1980) to fit the posterior function to the appropriate strength of the smoothness constraints in terms of the penalized log-likelihood.

We numerically obtain the MAP using the Laplace approximation (normal distribution approximation) of the posterior function at each iteration step and compute the ABIC that should be minimized by each set of constraint penalty weights. Details of our method can be found in Ogata et al. (2003) and Ogata (2004, 2011). A program and computational manual for estimation, short-term forecasting, and simulation implementation can be found in Ogata et al. (2021).

In addition, the following modeling can be used for more in-depth diagnostic analysis. Ogata et al. (2003) and Ogata (2004, 2011) propose that once the optimal MAP solution $\tilde{\lambda}(t,x,y|H_t)$ of the HIST-ETAS model is obtained, the 3D unknown function $\xi(t,x,y)$ in a piecewise linear function on the Delaunay tetrahedral partition (Delaunay volume function) is considered to represent a space-time anomaly from the estimated HIST-ETAS model, and then $\tilde{\lambda}(t,x,y|H_t)\xi(t,x,y)$ is re-fitted. Finding the optimal MAP solution $\tilde{\xi}(t,x,y)$ of the space-time Delaunay volume function by the ABIC smoothing shows the possibility of visualizing the space-time anomalous times and regions such as the quiescence or activation of seismic activity relative to the HIST-ETAS model. However, to clearly visualize such anomalies, the magnitude threshold of the original data set needs to be lowered to increase the information. To pursue this possibility, it is expected that the missing data correction, which will be discussed for the time case in Section 4.2, will be pursued appropriately for the space-time case.

In another way of space-time diagnostic analysis, the stochastic de-clustering method (Zhuang et al., 2002) together with the stochastic reconstruction method (Zhuang et al., 2004) can be applied to show either reproducibility or non-reproducibility of empirical statistical features obtained from the data.

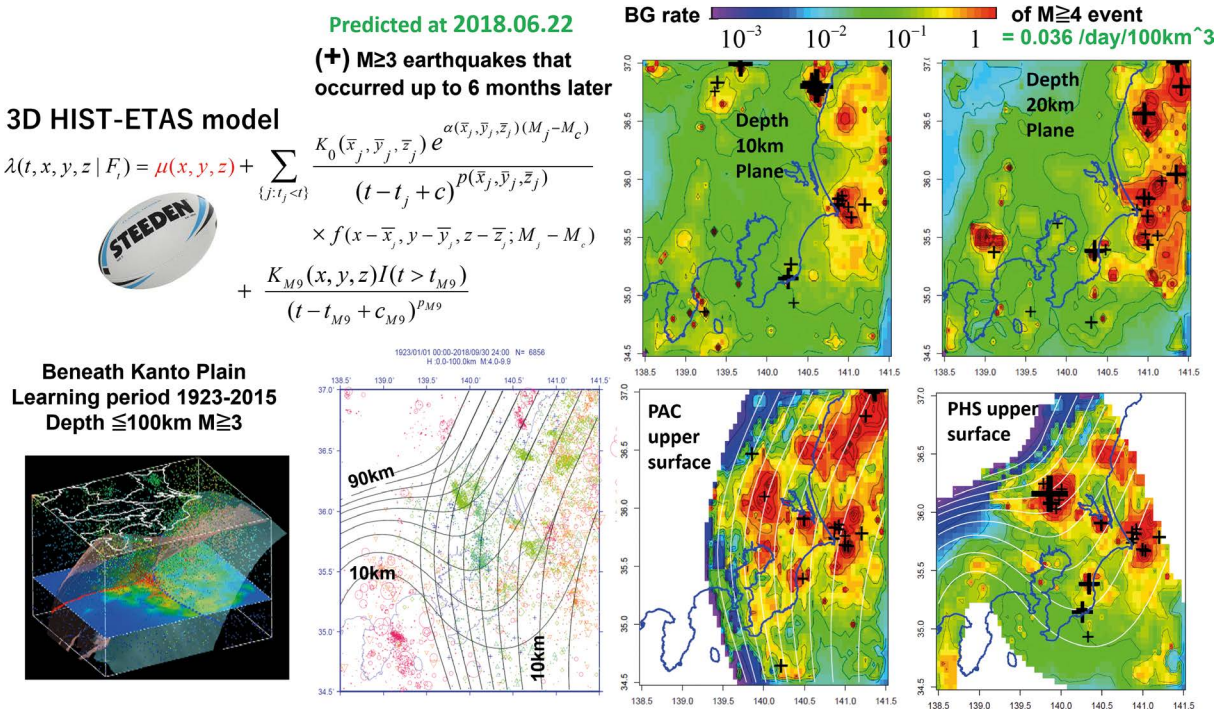


Figure 9. The HIST-ETAS and Omori-Utsu models are combined. The 3D models are constructed on Delaunay tetrohedron tessellation under the Kanto region. The effect of the M9 earthquake off the coast of Tohoku cannot be ignored. The 4 panels on the right show the predicted background intensities as of June 2018. The predicted 3D intensity rates are shown in 2D images intersecting the horizontal plane and the upper plate planes. Namely, these are images of the predicted intensities intersected by the 10 km and 20 km depth section planes, the Pacific plate plane, and the Philippine Sea plate plane. The plus signs indicate earthquakes that occurred in the vicinity of each surface during the following six months. The medium-term forecast proved to be quite successful.

2.8 Evaluating long-Term and Short-Term Forecasts

The estimated background seismic activity rate $\hat{\mu}(x,y)$ of the HIST-ETAS model in inland Japan is invariant over long-term intervals; it is independent of the learning target period, even when it includes or excludes the aftershock periods of the 2011 M9 Tohoku-oki earthquake. The high background rate retrospectively and apparently predicts the locations of large earthquakes with magnitude M6 or greated, including the locations of historically damaging earthquakes (Ogata, 2022). In addition, the background rate map appears to be consistent with the stress accumulation rates (see upper panels in Fig. 10) obtained from the GNSS data by Nishimura (2017) which also appears to correspond well with the major active fault alignments of M7 class or above. Thus, the lower panels show the 30-year probability forecast of an M ≥ 6 inland earthquake. The apparent different probability patterns between the background rates and the geodetic-based variation in northern Japan is still under consideration for the evaluation. But, for example, Ogata (2022) specifically confirmed that in inland Japan, the background intensity $\hat{\mu}_{full}(x,y)$ of the full HIST-ETAS model provides better long-term (i.e., about 30 years) forecast of the locations of large M ≥ 6.0 earthquakes than the estimated spatial 2D Poisson intensity $\lambda_{Poisson}(x,y)$, and further than the background rate $\hat{\mu}_{restricted}(x,y)$ of the restricted HIST-ETAS model.

Short-term forecasts $\hat{\lambda}_\theta(t,x,y|H_t)$ can be seen in the videos on YouTube by searching for "Yosihiko Ogata YouTube" https://www.youtube.com/channel/UCgfwpUeA-ztl5m5PsmnfOtA.

The performance score for short-term forecast is evaluated using the log-likelihood score (information gain score),

$$\log L(t_i,x_i,y_i;M_c) = \sum_{\{j:\, S<t_j<t_i\}} \log \hat{\lambda}\left(t_j,x_j,y_j|H_{t_j}\right) - \int_S^{t_i} \iint_{inland} \hat{\lambda}(u,x,y|H_u)dx\,dy\,du \tag{2.7}$$

of the considered HIST-ETAS models up to the last earthquake i. Here, for the MAP solution of the HIST-ETAS

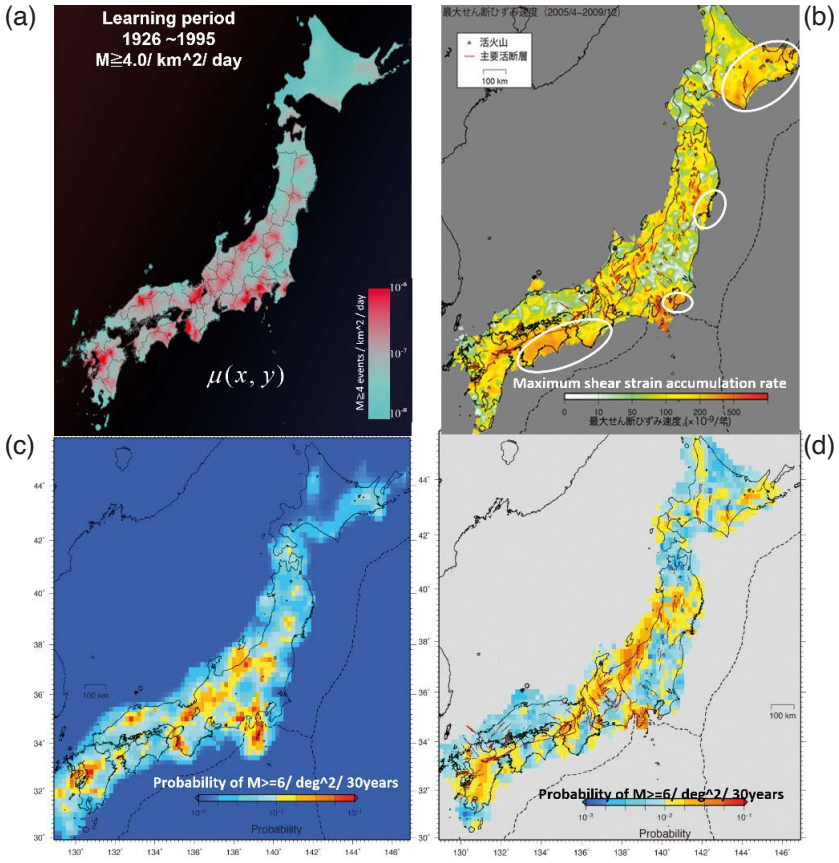


Figure 10. (a) The estimated background intensity $\mu(x,y)$ of the HIST-ETAS model in the interior per square kilometer per day of an earthquake with M ≥ 4. The extremes vary by a factor of 100. (b) The distribution of maximum shear strain rates for the period 2005-2009 at a high spatial resolution calculated by Nishimura (2017) using GNSS data. This shows a high correlation with $\mu(x,y)$ except for the circled areas where the strong stresses from plate subduction are included. (c) The expected 30 year probability of M ≥ 6 earthquakes per degree square, obtained by assuming the G-R law with $b = 0.9$, which is the MLE over the entire area of Japan. (d) The same 30-year probability recalculated from the corrected stress rates for the component pushed off the plate (Nishimura, 2024). Similarly, the background seismicity in California is given and discussed in Ogata (2017a).

models relative to the uniform Poisson intensity within the inland, which is based on the JMA inland data up to the year $S = 2018$, I have tested the performance of the short-term forecast for the incidence t_i in the period for the future 5 years from 2019. Also, the size of the difference between the increase and decrease of the compared score values can evaluate the details of the superiority or inferiority of the forecast results of certain earthquake occurrence times and locations. For predicted events, the performance of a model with a difference of $\log L - \log L_0$ in information scores is L/L_0 times more feasible than another model with a score of L_0 (Akaike, 1985; Ogata, 2017). I usually assume that the best fitted model is not always closest to the true with respect to the real data, but the other models may be useful in some periods and locations, so that I am reluctant to compare models by the average information gain score per event.

For example, Ogata (2024a) reports that the full HIST-ETAS model is generally superior for the recent period 2019-2023. However, this performs poorly for the recent period after the 2024 M7.6 Noto Peninsula earthquake. In terms of the spatial performance of the cumulative log-likelihood scores (2.7), the simple (i.e., non-Bayesian) space-time ETAS model performs better in regions with low seismic activity during the learning period up to the 2011 Tohoku earthquake, such as the Noto Peninsula region, eastern Fukushima Prefecture, northeastern Ibaraki Prefecture, and the eastern offshore area of Yamagata Prefecture. These correspond to the regions of very low background seismic activity. The HIST-ETAS models also do not perform well for volcanic swarms in the regions off the east coast of the Izu nor the Japanese Alps in central Japan.

On the other hand, in Collaboratory for the Study of Earthquake Predictability (CSEP; Schorlemmer, 2018), Nanjo et al. (2012) reported the prospective short-term forecasts performance of the models for the aftershocks of

the 2011 M9 Tohoku-oki earthquake from the learning period before the Tohoku-oki earthquake is poorly evaluated for all submitted different space-time ETAS models due to significantly high activations compared to the models after the M9 event (Ogata et al., 2013).

3. Diagnostic Analysis and Earthquake Physics

Since about the 1960s, there have been dramatic advances in the physics of earthquakes including the development of the theory of plate tectonics, which is the cause of stress accumulation in earthquake generation, and the theory of elastic repulsion, which states that the cause of earthquake shaking is the rapid discrepancy motion of fault planes. For each major earthquake, a fault motion model has been obtained by analyzing seismic waves. In this century, geodetic changes have been obtained from satellite observations, and the physical elucidation of the mode of earthquake occurrence and earthquake dynamics has progressed. Thus, seismic catalogs are now understood to represent the time and location of rupture onset, and the area and amount of slip of the ruptured fault. Data such as the direction of the fault plane and orthogonal plane, and the center of gravity of the fault have also been added. These data are used to study seismic activity.

3.1 Relative seismicity quiescence and relative activation

When seismic activity in an area appears to be quiescent, it may simply mean that aftershock activity from past earthquakes has ceased, or that aftershock activity from past earthquakes that should have been more active is inactive for some reason.

As a seismicity anomaly, Matsu'ura et al. (1995) prospectively reported a lower activity rate than the expected decay rate by the Omori-Utsu law for aftershocks of the 1995 M7.3 Kobe earthquake before its largest aftershock of M5.4. Similarly, for the aftershock activity of the 2005 M7.0 off the west coast of Fukuoka Prefecture, Ogata (2005a) reported a quiescence relative to the ETAS model before the maximum aftershock of M5.8 occurred.

The ETAS model allows us to account for such effects and to argue for a relative quiescence of the decaying aftershock activity relative to the expected level of activity (Matsu'ura, 1986; Ogata, 1988, 1989, 1992). In contrast, this is different from the quiescence for conventionally de-clustered events which is tested by the stationary Poisson hypothesis. Conventional de-clustering algorithms can lead to diverse and different results (e.g., Llenos et al., 2023), making it difficult to properly discuss data containing large aftershock activity from moderate to large earthquakes, and the accuracy of the results cannot be expected.

However, the relative quiescence is not always a precursor to a large earthquake or a moderate to large aftershock. I examined 76 aftershock sequences of past large earthquakes in Japan (Ogata, 2001b). See Ogata (2001); http://bemlar.ism.ac.jp/ogata/JGR01supplement/ for detailed records. Relative quiescence was observed in about 45% of the aftershock sequences, which is not particularly unusual. However, if the quiescence is prolonged for more than three months, the probability of an earthquake of the same or greater magnitude as the main shock occurring within 6 years is several times higher than for normal aftershocks in the vicinity of the aftershock area (e.g., within 200 km) in Japan (Ogata, 2001) or 3 ~ 4 times higher than probability gain induced by the space-time ETAS model (Ogata, 2017b).

One possible cause of relative quiescence is in the period when "slow slip" occurs within a nearby fault, causing stress shadows in the aftershock region. In such cases, there is a possibility that such slow slip may be a precursor to triggering a nearby fault rupture. We have been able to confirm several such slow slips by analyzing anomalous changes in GNSS geodetic data (e.g., Fig. 10) (e.g., Ogata, 2007a, 2010a, 2011a; Kumazawa et al., 2010), and I hope to analyze many cases so that we can monitor them in real time. Similarly, "relative activation" can be defined (Ogata, 2005, 2007, 2011a; Kumazawa and Ogata, 2013). In particular, Kumazawa et al. (2010) discussed the coexistence of a quiescent region and an activated region as precursors to the M9 Tohoku-oki earthquake. Specifically, following is found in many scenarios for retrospective interpretations of their physical significance that assume precursory slow slips on the down-dip extension of the focal fault solution of the forthcoming large earthquake, causing precursory changes in the patterns of the Coulomb Failure Stress (CFS) in a wide area around the focal event (Figs. 11 and 12), and also changes from the trend of time series of geodetic baseline distances between GNSS stations around the focal area (Figs. 12 and 13) (see details in Kumazawa et al., 2010, 2017, 2020; Ogata, 1992,

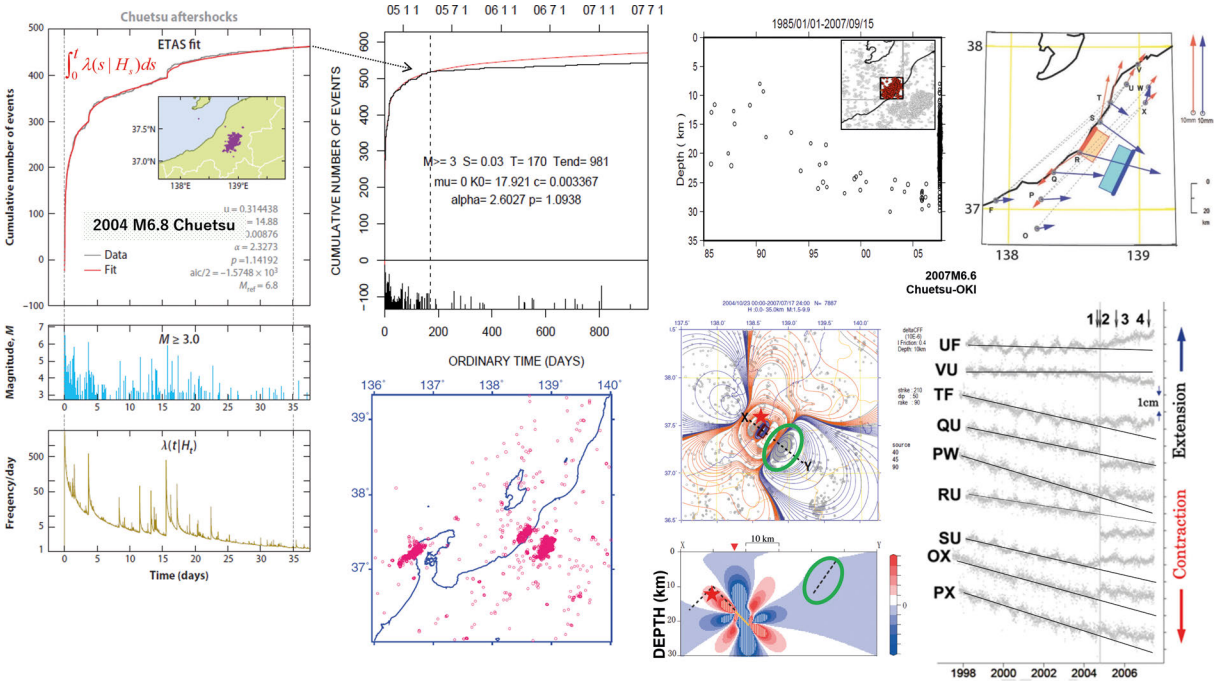


Figure 11. A physical scenario of the interaction between two neighboring earthquakes. On the upper left is the fitted ETAS for the aftershock series for the first week of the 2004 M6.8 Chuetsu earthquake. The red curve is the fitted cumulative ETAS curve. However, about six months later, the aftershocks lower than the expected rate during the period of two years, and the 2007 M6.8 Chuetsu-oki earthquake occurs nearby.

The preceding earthquake activity in the Chuetsu-oki region supports the slow slip before the event, as it moves deeper (top 3rd column panel) and the contraction of the GNSS baseline distance along the coast slows down (top and bottom in the 4th column panel). We can then assume that this slow slip (green ellptical circles in the middle and bottom third column panels) promoted the main rupture and created a stress shadow in the aftershock area of the 2004 Chuetsu earthquake. As shown in the lower right four panels, the preceeding slow slip on a part of the Chuetsu earthquake fault not only promoted its main rupture, but also reduced the aftershock activity of the Chuetsu earthquake. The relative quiescence in the top panel of the second column is because the stress shadow covered the aftershock area of the 2004 Chuetsu earthquake and suppressed the aftershock activity.

1999, 2001b, 2001, 2005a,b,c, 2006a, 2007a,b, 2011a, 2017a, 2017d; Ogata et al., 2003a; Ogata and Toda, 2010; Ogata and Tsuruoka, 2016). In general, the transformed times of aftershocks may be uniform over the aftershock period, but they are often not locally uniform in either time or space as shown in Fig. 13. If we take the aftershock sequence and use the time transformation (2.4) using the Omori-Utsu formula to look at the spatiotemporal distribution within the aftershock swarm in a plot of latitude or longitude coordinates x_i versus $\{(\Lambda(t_i), x_i); i = 1,2,\cdots,N\}$, we can see that the transformed occurrence times may be uniform throughout the period, but each subregion may have different space-time patterns (Ogata, 2010, 2017a; Ogata and Toda, 2010; Ogata and Tsuruoka, 2016) as shown in Fig. 13. Ogata (2010) attempted to characterize these patterns using six scenarios derived from the rate- and state-dependent friction law of Dieterich (1994). Spatiotemporal local quiescence/activation patterns and local cluster drifts in the aftershock zone can be observed prior to a large aftershock. Such pattern anomalies suggest that local rupture, local slow slip, or fluid intrusion may be occurring within the aftershock zone.

From about 1990 to the present, as a member of the Coordinating Committee for Earthquake Prediction (CCEP), I have had many opportunities to analyze numerous case studies in and around Japan and have published many reports on them in the Report of the Coordinating Committee for Earthquake Prediction (CCEP; http://cais.gsi.go.jp/YOCHIREN/report.html). These have been summarized in Ogata (2007b, 2017d), and the details of the most of them have been published in international journals. Except for a few cases, the analysis reports are ex post rather than prospective. If the location of the main shock and its fault solution are known, it is possible to retrospectively interpret geodetic data and changes in seismic activity by assuming slow slip on the down-dip extension of the main fault, as shown in Figs. 11 and 12 (see also Ogata, 2007a, b, 2008, 2010a, b, 2011).

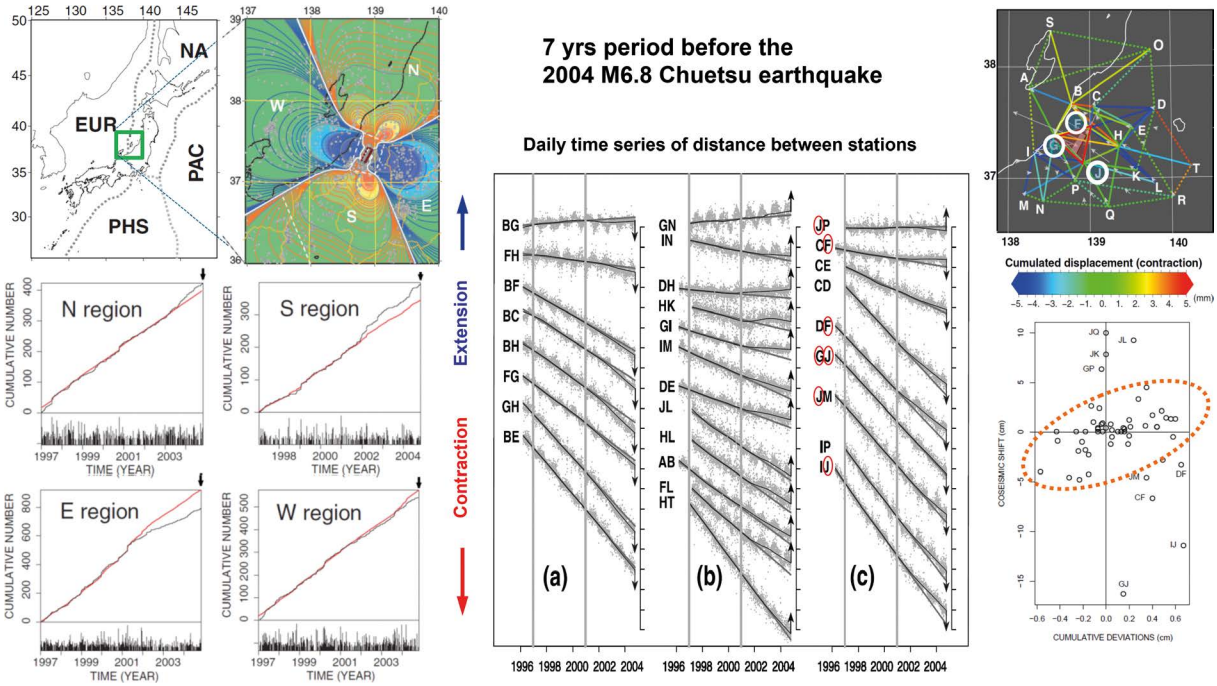


Figure 12. Precursor scenario for the 2004 M6.8 Chuetsu earthquake. Assuming that the Chuetsu earthquake was preceded by slow slip (the fault model is deep within the Chuetsu earthquake fault), we confirmed that the quiescence and activation of seismic activity in the 8 years preceding the Chuetsu earthquake correspond well to the positive and negative regions of the ΔCFS, when applied to the ETAS model (left diagrams). Although not shown here, local seismic activity was examined for each clustered sub-region containing a sufficient number of events for ETAS analysis, with harmonic results. The rectangle in the center is a model of the deep extension of the Chuetsu earthquake fault. This fault is assumed to slip slowly as in the similar mechanism solution of the main earthquake. The red and blue contours, warm and cold colors, represent positive and negative ΔCFS, respectively, for the major stress field in this region. When the crustal deformation (time series of baseline distance) is now considered using GPS data, it is consistent with slow slip except for the baselines including the three closest stations to the fault (the red and white circled characters in the panels). The predicted residual displacements of the baseline distance time series are positively correlated with the coseismic displacements of M6.8 except for the three most proximal stations.

Therefore, to prospectively link relative quiescence/activation and slow slip phenomena to earthquake forecasting in ETAS in the future, we need to explore the detection of seismicity anomalies and crustal deformation anomalies in advance and implement statistical models and methods that can efficiently detect them. However, in most cases, especially in the inland regions of Japan, slow slips are difficult to detect and reverse due to their small size and the many post-seismic movements caused by neighboring earthquakes.

As a solution to this problem, we can investigate how the stress changes by hypothetical slow slip fault models fit the earthquake occurrence patterns as well as the geodetic displacements of the GNSS stations. By constructing such a fault model and using the crustal deformation analysis method as shown in Figs. 11 and 12, the goodness of fit can be used as a constraint weight in the criteria for inverse analysis. Such forward inference has serious problems requiring a large amount of modeling effort as well as computational load, but it may be possible with the development of computer technology.

For medium- and short-term forecasting of a large interplate earthquake, episodic slow slip on a plate boundary may be a promising precursor (e.g., Nishikawa and Nishimura, 2023: Dascher-Cousineau and Burgmann, 2024). As discussed in Section 3.1, it must be discriminated from post-seismic or periodic slip. To empirically evaluate the probability gain (see Section 4.3) of a large earthquake occurrence, these slip types must be statistically classified based on the spatiotemporal pattern of seismic activity. See also the case for the relative quiescence discussed in Sections 3.1 and 4.3.

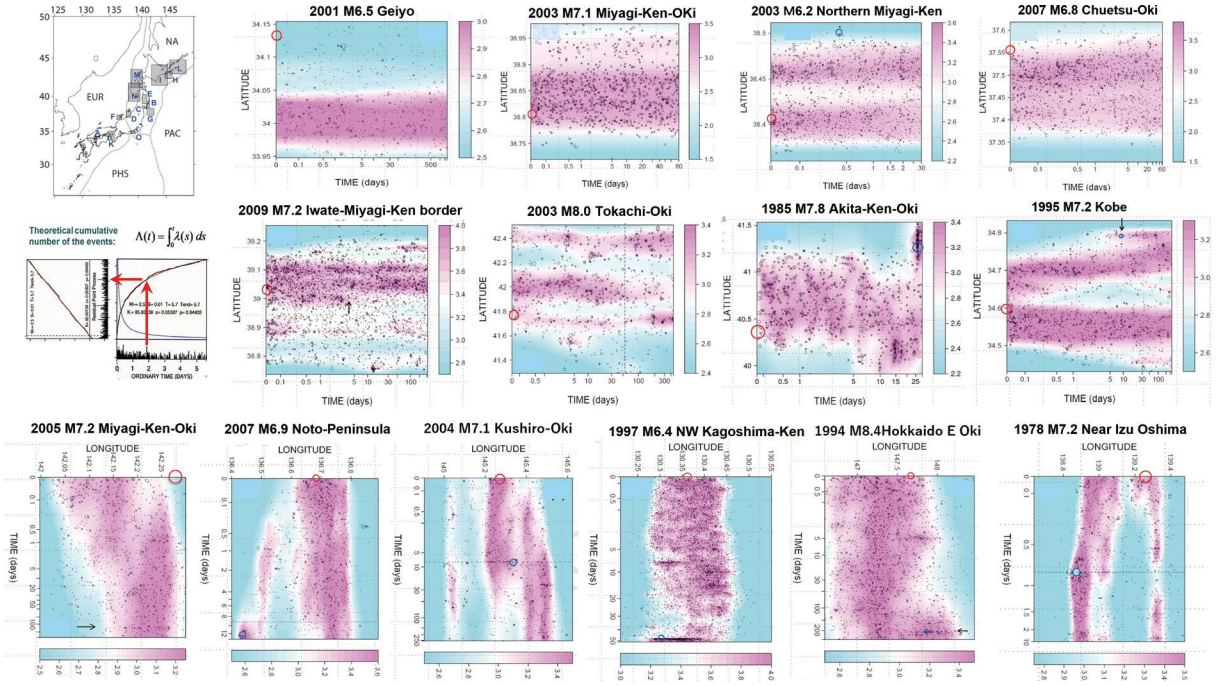


Figure 13. Examples of plots of homogeneous/heterogeneous space-time aftershock locations where latitude/longitude vs. time transformed by Eq. (2.4) for the Omori-Utsu formula, which are studied by Ogata (2010). The transformation by Eq. (2.4) is illustrated in the first column panel.

3.2 Earthquake Swarms and Non-stationary ETAS Model

Subsequent studies have extended the ETAS model to a wide range of seismic anomalies. Seismic activity can increase or decrease in response to changes in crustal stresses caused by seismic or non-seismic slip of surrounding faults, and it is also known that fluid intrusion into the fault gap can cause a swarm of seismic activity. Transient swarm earthquakes mixed in a normal ETAS system can be estimated indirectly by calculating their significance from the poor estimation results of the stationary ETAS model. For example, there are several reports on the relationship between earthquake swarms and slow slip at plate boundaries (e.g., Nishikawa and Ide, 2018; Nishikawa et al., 2021, Nishikawa et al., 2023).

These suggest that a non-stationary ETAS model in which the parameter $\mu(t)$ in Eq. (2.5) varies with time can be well fitted (Kumazawa et al., 2010, 2016, 2017; Kumazawa and Ogata, 2013, 2014). The time variation of the parameter $\mu(t)$ can provide a quantitative physical relationship between the weakening of a sticking fault due to the intrusion of fluids, such as magma or hydrothermal fluids, into the fault system and the rate of earthquake occurrence. For example, Kumazawa and Ogata (2013) applied non-stationary ETAS model to induced earthquakes in northern Japan immediately after the 2011 M9 Tohoku-oki earthquake and were able to provide supporting evidence that some of these inland earthquakes were caused by a reduction in fault strength due to fluid intrusion induced by the earthquake ground motions of the M9 Tohoku-oki earthquake. The reproducibility of the data was verified by simulation experiments.

Figure 14 illustrate frequent transient slow slips off the east coast of the Boso Peninsula and its northern part (off the east coast of Ibaraki Prefecture) have been reported in significant numbers by, tiltmeter anomaly records, GNSS geodetic recording anomalies, and highly frequent repeating (recurrent) earthquakes. Fitting the non-stationary ETAS model to these areas corresponds to sharp peaks in $\mu(t)$ values of background seismicity for most of the anomalies (Kumazawa and Ogata, 2024a).

In addition, the non-stationary ETAS model was used to investigate the predictability of earthquake swarms associated with magma intrusion in volcanic and geothermal regions (Fig. 14). Velocity changes in volumetric strainmeter records installed by the Japan Meteorological Agency (JMA) were reported to show a high correlation between the strain changes associated with magma ascent and the occurrence of earthquake swarms off the east coast of the Izu Peninsula. Kumazawa et al. (2016) found that the temporal changes in background seismic activity

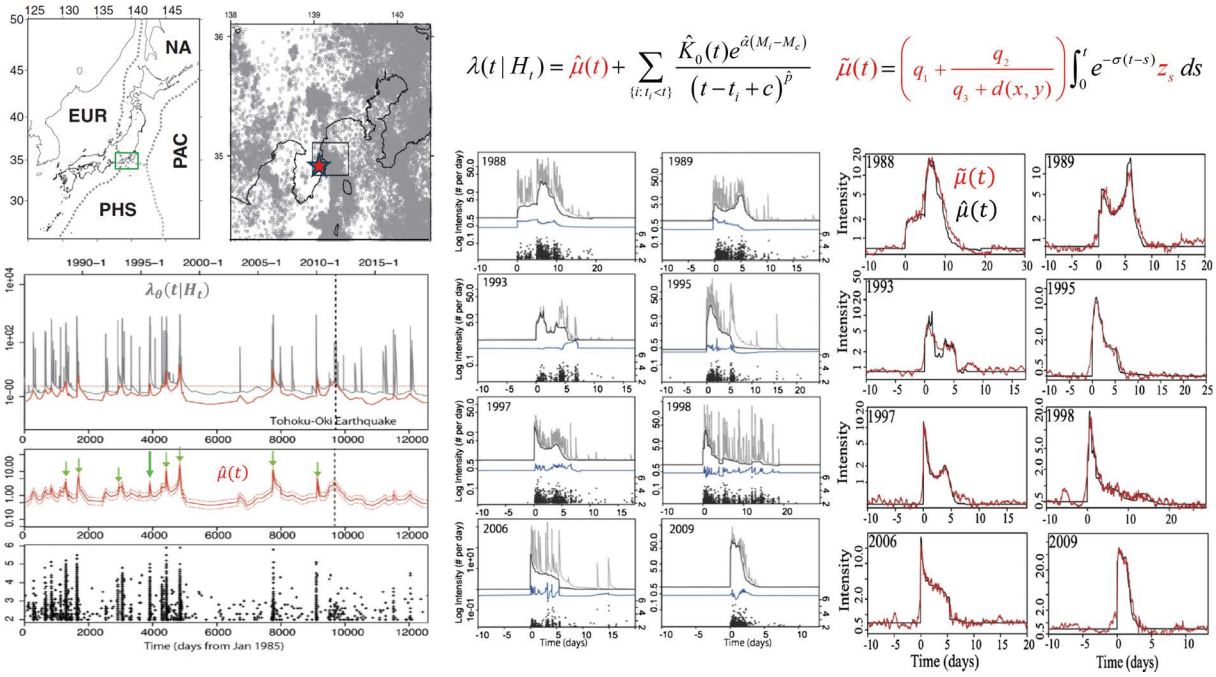


Figure 14. Offshore Izu Peninsula earthquake swarm. The three lower left panels show the MAP results of the non-stationary ETAS model for M ≥ 2.0 data from 1985 to 2019. The red curves in the lower left panels are the estimated background intensity of the non-stationary ETAS model with 95% confidence error, and the eight periods with arrows are considered swarm in Kumazawa et al. (2016). The central eight panels are magnified from those fitted by the non-stationary ETAS model and separately represent the groups indicated by the arrows, and the black curve is the MAP solution for the background intensity $\mu(t)$. The panels in the right eight panels are the $\mu(t)$ solution (black curve) and the prediction based on the rate of change of volumetric strain (red curve). The latter data are provided by the JMA, which continuously records the volumetric strain at the locations marked with an asterisk. The prediction function for the μ value calculated from the volumetric strain change z_s depends on the distance $d(x,y)$ from the strainmeter to the initial location of the swarm earthquake, and all other constant parameters being equal. All are in good agreement with the previously obtained non-stationary ETAS μ values, where the y-axis is the logarithmic scale.

in the non-stationary ETAS model were more sensitive to strain changes than the activity of the swarm itself. Here the volumetric strainmeter records are calibrated by removing the effects of, atmospheric pressure, ocean tides, groundwater level changes due to precipitation, and the coseismic changes.

Therefore, Kumazawa et al. (2016) propose a model that forecasts the conditional intensity rate of the extended ETAS model from the corrected volumetric strain records, and then links it to the forecasting the frequency of earthquake swarms. Thus, when the starting location of a swarm earthquake is given, we can accurately forecast the conditional intensity rate about half a day after the volumetric strainmeter change data (see Fig. 14). The results show the prospect of improving the forecasting performance of seismicity in seismic activity with external triggering factors by appropriately incorporating external data such as geodetic data. At the same time, it suggests the possibility of developing space-time inversion of crustal stress information from information such as seismic swarm activity and strainmeters.

In addition, we quantitatively investigated the seismicity diversity in the central Kyushu region, including the foreshocks and aftershocks of the 2016 M6.4 and M7.3 Kumamoto earthquakes by applying various point process models, including the ETAS and non-stationary ETAS models, from the seismic activity in the aftershock area before and after the earthquake. A detailed analysis of the foreshock and aftershock series was also performed for the 2018 Hokkaido Eastern Iburi earthquake, the study of which clearly showed that a portion of the aftershock activity was of swarming type and explicitly showed the portion of partial fluid intrusion (Fig. 6), (Kumazawa et al., 2019).

Seismic activity in the northern Noto Peninsula began in late 2020, followed by the M6.5 earthquake in May 2023, and then the M7.6 earthquake on New Year's Day 2024. The mechanism of this series of seismic activity until just before the M7.6 event was analyzed by Kumazawa and Ogata (2024b) using the non-stationary ETAS model and

various spatiotemporal statistical models. In the aftershock series of the M6.5 event in 2023, it was estimated that the region immediately before the M5.9 maximum aftershock showed a quiescence, and the background seismicity increased immediately after the mainshock, then decreased, and finally increased again after a short time (Ogata and Kumazawa, 2024a, b). This suggests that the fluid that was injected into the ground by the seismic motion immediately after the M6.5 earthquake was processed and the fluid pressure increased again.

A similar method was applied to a series of 2010-2014 earthquake swarms in southern Italy (Petrillo et al., 2024).

4. Magnitude Series Forecasts

The simplest statistical model is exponential distribution with respect to magnitude, namely, the Gutenberg-Richter law (G-R law) (Gutenberg and Richter, 1956), and they showed that the coefficient b value is different in different parts of the world. This value is also known to be different in different parts of Japan (Ogata, 2011b their Figs. 3 and 4; 2011c their Fig. 3) and can vary with time even in the same location. In particular, the temporal variation of b-values is thought to reflect changes in the stress field in the region based on the rock fracture experiments and is expected to provide a useful information for earthquake forecasting. For example, Suyehiro (1966) showed that b-values can differ significantly between foreshocks and aftershocks. Traditionally, the coefficient b has been estimated as the slope of a line using least squares or other methods based on a logarithmic plot of magnitude versus cumulative number of aftershocks. Utsu (1965) proposed the estimator by the method of moments, which Aki (1965) soon showed it being the MLE and gave a theoretical estimation error.

According to Utsu (1971), the number of papers describing the b-value at that time was not less than 250. At present, more than 50 years have passed since then, and it is assumed that a considerable number (possibly more than ten thousand) of papers have been written in retrospect. This means that it is difficult to discuss on b-value changes except on a case-by-case basis in space and time when applying the simple b-value estimate (i.e., reciprocal mean magnitude) to extremely heterogeneous geology and complex stress changes. I find it much difficult to make predictions using b-values and I describe some b-value related issues in subsection 4.1.

4.1 Discrimination of Foreshocks

The ETAS model (2.2) is the probability of an earthquake above a certain magnitude threshold. However, the magnitude distributions associated with the ETAS model are not forecasted, where the magnitude sequence is independently and identically distributed according to the G-R law for a given b-value. Namely, it follows an exponential distribution above a certain lower threshold magnitude M_c.

Thus, the simplest model for forecasting aftershock rates, both in terms of occurrence times and magnitudes is $g(t, M) = K(t + c)^{-p} \times 10^{-b(M-M_c)}$ that of Reasenberg and Jones (1989). Here the ETAS model (2.2) can be considered instead of the Omori-Utsu formula (e.g. Omi et al., 2014).

In practice, the UCERF3-ETAS method (Field et al., 2017; US Geological Survey, 2022), which implements medium- and long-term forecasts, uses characteristic magnitude distributions of intrinsic earthquakes for earthquake fault segments throughout California, and an upper bound for the space-time ETAS model. For the magnitude sequences, the space-time ETAS model is combined with a tapering G-R distribution. However, the magnitude sequence remains independently distributed.

In fact, the magnitude time series may in some cases depend on historical data such as earthquake occurrence patterns. The key clue is the probabilistic discrimination of "foreshocks". When a new earthquake cluster starts in a region, I want to prospectively predict the probability that it will be of the "foreshock" type in real time. Ogata et al. (1995, 1996) separate earthquake clusters and link all those closer than a certain space-time distance to the earthquake catalog of M ≥ 4.0 (single-link method).

The largest earthquake in each cluster is called the main shock, and if the magnitude difference between the main shock and the largest foreshock in the cluster is greater than 0.45, the foreshock is called a foreshock type. In other words, we estimate the probability that an earthquake that is 0.45 magnitude larger than the largest earthquake in the cluster so far will occur within one month. The magnitude difference should be as large as common sense dictates, but since this would reduce the proportion of foreshock-type clusters in the total cluster, we chose 0.45 for a meaningful statistical discussion (Ogata et al., 1995).

Namely, Ogata et al. (1995, 1996) defined earthquake clusters and proposed a formula for evaluating the probability of a significantly larger earthquake occurring in a month based on the statistical identification of the space-time-magnitude configuration patterns of previous earthquakes within the cluster.

In the 17 years period since we proposed such a formula, Ogata and Katsura (2012) validated the forecasting results. The performance was much better than the constant probability of a foreshock based on the log-likelihood ratio criterion. In particular, the results of the discrimination results were quite clear when restricted to clusters of earthquakes with large main shocks (e.g., $M_{main} \geq 6.5$). That is, when compared to the empirical constant probability foreshock forecast using the log-likelihood ratio criterion (equivalent to the information gain score), the performance was far superior.

Subsequently, Nomura and Ogata (2020, 2023) created a model to evaluate the probability of each magnitude larger than the largest earthquake to date in a cluster. They then calculated and tested online the probability of the larger earthquakes for ongoing seismic activity throughout Japan and also the world, by an assumed magnitude and 30-day imminence, based on recent forecasting experiments using the JMA catalog and USGS global catalogs.

Thus, the probabilistic identification of clusters that may be foreshock-type can be a valid forecast (Ogata, 2017b; Ogata et al., 2018). The superiority of the forecast is verified by the log-likelihood ratio criterion (information gain score) compared to the forecast results of the standard G-R distribution model (5.1) (Ogata et al., 2018; Nomura and Ogata, 2020, 2023).

From a different perspective to support the magnitude dependence, Ogata and Katsura (2014) created the first type of synthetic catalog by sequentially inputting the magnitude time series from the JMA catalog into the ETAS model and using the time of occurrence and epicenter coordinates, while the second type was created by bootstrapping the magnitude sequences from the JMA catalog and inputting them into the ETAS model. The second type of synthetic catalog was created to ensure independence and input into the ETAS model. Each synthetic catalog satisfies the properties of the spatio-temporal characteristics of the foreshock statistics (Helmstetter et al., 2003), but the data assuming independence of the magnitude series show significantly worse probability gain of foreshocks than the original JMA data.

Ogata (2018) investigated the information gain score of three types of spatio-temporal conditional probability models in forecasting the magnitude of future earthquakes relative to an independently distributed G-R model with a constant *b*-value. The first one is a spatially smoother *b*-value model; the second is a model in which the reciprocal of the *b*-value (i.e., the denominator of the MLE) forecasts by auto-regression of nearby magnitude sequences; and the third is a model that forecasts magnitude based on foreshock identification methods within earthquake clusters.

The forecasting performance of the three models shows that, except for the last model, the information gain score did not exceed the basic G-R law in the Japan area. These experiments may indicate that the temporal and spatial smoothing models for *b*-values are not suitable for forecasting the magnitude sequence of the new rupture, suggesting that it can only be captured by short-term local changes.

4.2 Data Heterogeneity and the ETAS Models

The JMA source catalog of the Japan region for the period since 1923, a database of earthquake occurrence patterns that has been greatly expanded over the long period of time, is spatiotemporally heterogeneous with the development of observing systems. We model the non-uniformity of the source catalog and consider a detection function for detected data (Ogata and Katsura, 1993) for long-term and for aftershock sequences (Ogata and Katsura, 2006). The detection distribution density function for earthquakes of magnitude M is proposed to be modeled by an exponential function (the G-R law) multiplied by a cumulative normal distribution (error function) with mean μ and standard deviation σ (see left plots in Fig. 15). Here, μ represents the magnitude value at which 50% of the earthquakes are detected, and σ represents the range of magnitudes at which earthquakes are partially observed. In my experience, σ reflects the size of the area of aftershock activity and the spatiotemporal heterogeneity of the stations that recorded the seismic waves.

Model-fitted results will be biased if the spatial and temporal heterogeneity of earthquake detection rates and the lack of data immediately following a relatively large-magnitude earthquake (mainshock) are not considered. Thus, the number of usable data is limited by the magnitude limit to ensure uniformity, which has been solved by setting an appropriate lower magnitude limit or avoiding problematic time spans.

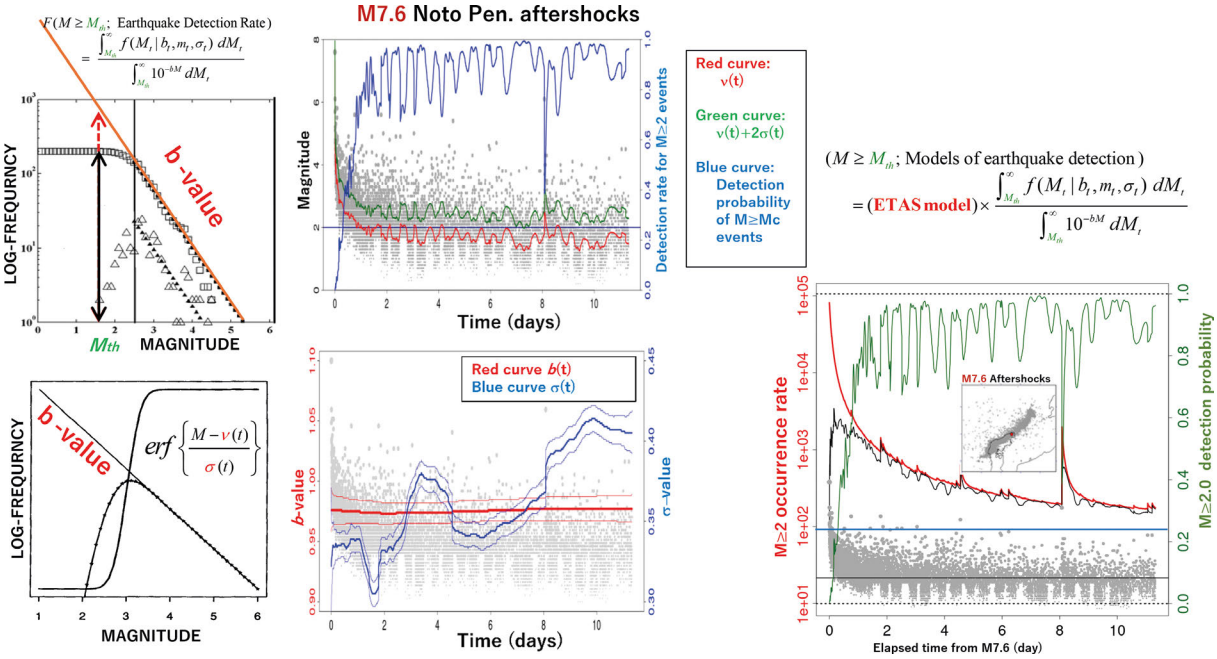


Figure 15. Time evolution of the aftershock detection rate immediately after the 2024 M7.6 Noto Peninsula earthquake and ETAS model estimation [Ogata and Kumazawa, 2024]. The left panels illustrate the detection rate models. Upper middle panel: red and green curves shows the 50% detection rate μ values and μ+2σ values, respectively, and blue curve is the detection rate of aftershocks with M ≥ 2.0. Lower middle panel: Red and blue curves with errors show the change in b and σ values. Right panel: Green curve, detection rate for M ≥ 2.0 aftershocks; black and red curves show the apparent and the underlying ETAS intensity function for M ≥ 2.0 aftershocks. Gray dots in all panels show magnitude-time of all detected earthquakes. All curves are ABIC optimal MAP solutions.

However, Ogata and Katsura (2006) proposed a solution to the data deficiency by incorporating the simultaneous estimation model of seismic detection rate and time variation of b-value of Ogata and Katsura (1993) into the Omori-Utsu model. To characterize the non-stationarity of the magnitude time series of detected earthquakes, we assume parameters $b(t)$, $\mu(t)$ and $\sigma(t)$ that depend on the elapsed time t. The Bayesian model is considered and applied for these smoothing estimates by using the ABIC. See Omi et al. (2013, 2014), Ogata and Kumazawa (2024a,b) and others, for the Reasenberg-Jones-type real-time aftershock forecasting methods that account for missing measurements immediately after the main shock.

A conventionally used approach is to set the time dependence of the lower magnitude limit M_c to be detectable with a probability (97.5%) greater than, say, $\mu + 2\sigma$, for example, and then fit the ETAS model (e.g., Helmstetter et al., 2006; Page et al., 2016). However, even in this upper range of magnitudes, earthquakes are missed. Therefore, parameter estimation requires considering the detection rate as a function of earthquake magnitude, as in Ogata and Katsura (2006) and Omi et al. (2013).

The calibrated ETAS model is the underlying ETAS model multiplied by the detection rate of aftershocks with M ≥ Mc. Thus, we can obtain the MLE of the underlying ETAS model by maximizing the log-likelihood function of the calibrated ETAS model using the inhomogeneous aftershock data immediately after the main shock (Fig. 15).

We can also apply the calibrated non-stationary ETAS model and compare its goodness of fit with the stationary ETAS model using the ABIC minimization procedure (Ogata, 2024a,b; Ogata and Kumazawa, 2024a,b).

Aftershocks can occur in two ways: statically, due to the change in shear stress, and dynamically, due to the reduction in fault strength caused by pore fluid pressure (Nakagomi et al., 2021; Hardebeck and Harris, 2022, Hardebeck et al., 2024). We want to investigate whether a statistical model can quantitatively categorize them based on the temporal distribution of aftershocks. The key clue is to estimate the temporally varying background intensity of aftershock activity, and the non-stationary ETAS model seems to be effective in this respect (Ogata and Kumazawa, 2024a,b). A more detailed analysis would be possible if the problem of missing aftershocks could be overcome. We would like to deepen our understanding of the process of aftershock generation in conjunction with

the magnitude dependence of the spatiotemporal distribution of aftershocks, since a certain pattern of temporal variation of b-values is also observed in aftershock sequences.

4.3 Multi-Element Probability Predictions

In Section 2.8, it was explained that reasonable long-term forecasts of large earthquakes can be obtained from the background rate $\mu(x,y)$ of the HIST-ETAS model. In addition to this, the probability gains of the multi-element probability forecast equation are expected to increase when information is available on the probability gains associated with, for example, foreshock identification (short-term forecast) and to relative quiescence (medium-term forecast). The multi-element probability forecast equation (Utsu, 1979; Aki, 1981; Ogata, 2013, 2017a,b) is expected to increase significantly in practical use. In retrospective studies, Utsu (1979) calculated the multi-element probability forecast for the 1978 M7.6 Izu-Oshima-Kinkai earthquake, and Cao and Aki (1983) also calculated the probability for four large M7 to M8 earthquakes in China. Although the 1975 M7.3 Haichen event is considered a successful forecast, while the 1976 M7.8 Tangshan event is considered a failure, Cao and Aki (1983) show that both have the probability of about 10% per day of the immediately preceding the day of the event. Both cases have a very high probability per day given the expected magnitude size from medium and longer term forecasts. There seem to be deep political reasons why it is taught that China was lucky at Haicheng and that a harsh, unpredictable reality struck back at Tangshan (Mearns and Sornette, 2021).

Recently, I also calculated the retrospective results of the multi-element probability forecasts for the 2016 M7.3 Kumamoto earthquake (Ogata, 2017a,b). These studies made me realize the importance of developing model-based residuals to increase the probability gain (Ogata, 2013, 2017a,b).

The concept of multi-element prediction has been found theoretically in statistical seismology studies. However, multi-element probability prediction has not been operational for 40 years. This is mainly because the seismological community's desired goal for "earthquake prediction" in the past has been to enthusiastically discover anomalous phenomena with a high probability gain of causing earthquakes, and such prediction is very limited due to low alarm rate, resulting in an overwhelming number of large earthquakes occurring unexpectedly (low alarm rate). Therefore, we need to systematically collect as many anomalies as possible, even if they have low probability gains (low success rate).

5. Final Remarks

Hiro Akaike once said, "Statistical models are like the telescopes and microscopes of data analysis." It is well known that since ancient times, the driving force behind scientific discovery has been methodological innovation, such as the discovery and improvement of telescopes and microscopes. The proof of a scientific hypothesis is determined by the results of its predictions. Although the subjects of statistical science are truly complex systems, appropriate statistical models not only fulfill predictions to a certain extent, but also reveal new scientific facts, also called "anomalies" or "unexpected" that were not considered when the model was created, from real data. Statistical models and the graphical and pictorial representations based on them can serve as a scientific method to clearly show what is barely visible and what is not, leading to new insights.

I have tried to demonstrate the usefulness of statistical methods by expressing a number of empirical laws based on seismic statistics and physical hypotheses of seismology as statistical point process models. For example, the ETAS model and its extensions, which were constructed based on empirical aftershock activity to forecast not only reasonable short-term rates from a database of earthquakes, offers the possibility of being used as a benchmark to detect subtle anomalies in seismic activity. We hope that we have conveyed the development of statistical seismology, which is closely related to the study of seismic activity, and its importance in forecasting.

Acknowledgments. The author is supported by the "Seismological Research Innovation Promotion Program (STAR-E)" (Grant No. JPJ010217) of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT). Comments and suggestions by Dr. Ilaria Spassiani and an anonymous reviewer have been helpful in clarifying the present paper.

References

Abe, K. (1981). Magnitudes of large shallow earthquakes from 1904 to 1980.

Adamopoulos, L. (1976). Cluster models for earthquakes: Regional comparisons., Math. Geol., 8, 463-475.

Akaike, H. (1969). Fitting autoregressive models for prediction, Ann. Inst. Statist. Math., 21, 243-247.

Akaike, H. (1974). A new look at statistical model identification, IEEE Trans. Automat. Contrl.,19, 6, 716-723, doi:10.1109/TAC.1974.1100705.

Akaike, H. (1980). Likelihood and the Bayes procedure, in Bayesian Statistic, J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds. University Press, 143-166, (discussion 185-203); also in 'Selected Papers of Hirotugu Akaike' edited by E. Parzen,K Tanabe and G. Kitagawa, Springer-Verlag, Tokyo.

Akaike, H. (1985). Prediction and entropy, in A Celebration of Statistics, A. C. Atkinson and E. Fienberg (Editors), Springer-Verlag, New York, 1-24, doi:10.1007/978-1-4613-8560-8_1.

Aki, K. (1956). Some Problems in Statistical Seismology, Zisin ii 8, 4, 205-228, doi:10.4294/zisin1948.8.4_205.

Aki, K. (1965). Maximum likelihood estimate of b in the formula log N = a - bM and its confidence limits, Bull. Earthq. Res. Inst., Univ. Tokyo 43, 237-239.

Aki, K. (1981). A probabilistic synthesis of precursory phenomena, Earthquake Prediction (eds. D. W. Simpson and P. G. Richards), Maurice Ewing Series, 4, 566-574, AGU, Washington, DC, 566-574.

Bansal, A. R. and Y. Ogata (2013). A non-stationary epidemic type aftershock sequence model for seismicity prior to the December 26, 2004 M9.1 Sumatra-Andaman Islands mega-earthquake, J. Gophys. Res. 118, 616-629, doi:10.1002/jgrb.50068.

Bartlett, M. S. (1963). The spectral analysis of point processes, J. R. Statist. Soc. B., 25, 264-296.

Cao, T. and K. Aki (1983). Assigning probability gain for precursors of four large Chinese earthquakes, J. Geophys. Res. 88, B3, 2185-2190. doi:10.1029/JB088iB03p02185.

Daley, D. J. and D. Vere-Jones (1988). An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods, Springer Nature.

Daley, D. J. and D. Vere-Jones (2008). An Introduction to the Theory of Point Processes, Volume II: General Theory and Structure, Second Edition, Springer Nature.

Dascher-Cousineau, K. and R. Burgmann (2024). Global subduction slow slip events and associated earthquakes, Sci. Adv., doi:10.1126/sciadv.ado2191.

Dieterich, J. (1994). A constitutive law for the rate of earthquake production and its application to earthquake clustering. J. Geophys. Res., 99, 2601-2618.

Di Giacomo, D. and D. A. Storchak (2022). One hundred plus years of recomputed surface wave magnitude of shallow global earthquakes, Earth Sys. Sci. Data, 14,2, 393-409, doi:10.5194/essd-14-393-2022.

Engdahl, E. R. and A. Villasenor (2002). Global Seismicity: 1900-1999. In: Lee, H. K., H. Kanamori, P. C. Jennings et al., Eds. Handbook of Earthquake and Engineering Seismology, Part A, Academic Press, Amsterdam. https://www.sciencedirect.com/science/article/abs/pii/S0074614202802443?via%3Dihub.

Falcone, G., R. Console and M. Murru (2010). Short-term and long-term earthquake occurrence models for Italy: ETES, ERS and LTST, Ann. Geophys., 53, 3, 41-50.

Field, E. H., K. R. Milner, J. L. Hardebeck, M. Page et al. (2017). A Spatiotemporal clustering model for the Third Uniform California Earthquake Rupture Forecast (UCERF3-ETAS): Toward an Operational Earthquake Forecast, Bull. Seism. Soc. Am., 107, 3, 1049-1081, doi:10.1785/0120160173.

Good, I. J. and R. A. Gaskins (1971). Nonparametric roughness penalties for probability densities, Biometrika, 58, 255-277.

Guo, Z. and Y. Ogata (1997). Statistical relations between the parameters of aftershocks in time, space and magnitude, J. Geophys. Res., 102, B2, 2857-2873.

Gutenberg, B. and C. F. Richter (1944). Frequency of earthquakes in California, Bull. Seism. Soc. Am., 34, 185-188.

Gutenberg, B. and C. F. Richter (1954). Seismicity of the Earth and Associated Phenomena, 2nd edn., Prinston University Press, Princeton, NJ.

Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. J. R. Statist. Soc., B 33, 438-43.

Hawkes, A. G. and J. Chen (2021). A Personal History of Hawkes Processes. Proceedings of the Institute of Statistical Mathematics 69 (2), 133-143 (English version). https://www.ism.ac.jp/editsec/toukei/pdf/69-2-123.pdf.

Hainzl, S. and Y. Ogata (2005). Detecting fluid signals in seismicity data through statistical earthquake modeling, J. Geophys. Res. 110, B05S07, doi:10.1029/2004JB003247.

Hardebeck, J. L. and R. A. Harris (2022). Earthquakes in the shadows: why aftershocks occur at surprising locations. The Seismic Record, 2, 3, 207-216, doi:10.1785/0320220023.

Hardebeck, J. L., A. L. Llenos, A. J. Michael, M. T. Page et al. (2024). Aftershock Forecasting, Ann. Rev. Earth Planet. Sci., 52, doi:10.1146/annurev-earth-040522-102129.

Han, P., K. Hattori, J. Zhuang, C. H. Chen et al. (2017). Evaluation of ULF seismo-magnetic phenomena in Kakioka, Japan by using Molchan's error diagram, Geophys. J. Int., 208, 1, 482-490, doi:10.1093/gji/ggw404.

Helmstetter, A., D. Sornette and J. R. Grasso (2003), Mainshocks are aftershocks of conditional foreshocks: How do foreshock statistical properties emerge from aftershock laws, J. Geophys. Res., 108, B1, 2046, doi:10.1029/2002JB001991.

Helmstetter, A., Y. Y. Kagan and D. D. Jackson (2006). Comparison of short-term and time-independent earthquake forecast models for southern California, Bull. Seism. Soc. Am., doi.10.1785/0120050067.

Jordan, T. H., Y.-T. Chen, P. Gasparini, R. Madariaga, I. Main et al. (2011). Operational earthquake forecasting: State of knowledge and guidelines for implementation, Final Report of the International Commission on Earthquake Forecasting for Civil Protection, Ann. Geophys., 54, 4, 315-391, doi:10.4401/ag-5350.

Kagan, Y. Y. (1991). Likelihoos analysisof earthquake catalogs, J. Geophys. Res., 106, 438-443.

Kumazawa, T., Y. Ogata and S. Toda (2010). Precursory seismic anomalies a nd transient crustal deformation prior to the 2008 Mw = 6.9 Iwate-Miyagi Nairiku, Japan, earthquake, J. Geophys. Res., 115, B10312, doi:10.1029/2010JB007567.

Kumazawa, T. and Y. Ogata (2013). Quantitative description of induced seismic activity before and after the 2011 Tohoku-Oki Earthquake by non-stationary ETAS models, J. Geophys. Res., 118, 12, 6165-6182, doi:10.1002/2013JB010259.

Kumazawa, T. and Y. Ogata (2014). Non-stationary ETAS models for nonstandard earthquakes, Ann. Applied Stat., 8, 3, 1825-1852, doi:10.1214/14-AOAS759, http://projecteuclid.org/euclid.aoas/1414091236.

Kumazawa, T., Y. Ogata, K. Kimura, K. Maeda and A. Kobayashi (2016). Background rates of swarm earthquakes that are synchronized with volumetric strain changes, Earth Planet. Sci. Lett., 442, 51-60, doi:10.1016/j.epsl.2016.02.049.

Kumazawa, T., Y. Ogata and H. Tsuruoka (2017). Measuring seismicity diversity and anomalies by point process models: Case studies before and after the 2016 Kumamoto Earthquakes in Kyushu, Japan, Earth, Planets and Space, 69, 169, doi:10.1186/s40623-017-0756-6.

Kumazawa, T., Y. Ogata and H. Tsuruoka (2019). Characteristics of seismic activity before and after the 2018 M6.7 Hokkaido Eastern Iburi earthquake, Earth Planets and Space, 71, 130, doi:10.1186/s40623-019-1102-y.

Kumazawa, T., Y. Ogata and S. Toda (2020). Wide-area seismicity anomalies before the 2011 Tohoku-Oki earthquake, Geophys. J. Int., 223, 2, 1304-1312, doi:10.1093/gji/ggaa356.

Kumazawa, T. and Y. Ogata (2024a). Non-stationary ETAS model: How it works for externally forcing. Seismol. Res. Lett. (Special issue on Statistical Seismology), doi:10.1785/0220240166.

Kumazawa, T. and Y. Ogata (2024b). Spatial and temporal variations of the 3-year earthquake swarm activities leading up to the M7.6 Noto Peninsula earthquake Research Memorandum (Technical report) 1228. Institute of Statistical Mathematics, https://www.ism.ac.jp/editsec/(Submitted to Earth Planet Space).

Kumazawa, T and Y. Ogata (2024c). Analysis of aftershock activity of the 2024 Noto Peninsula M7.6 Earthquake, Report of CCEP 112. https://cais.gsi.go.jp/YOCHIREN/activity/243/image243/034-036.pdf.

Lei, X., S. Ma, W. Chen, C. Pang et al. (2013). A detailed view of the injection-induced seismicity in a natural gas reservoir in Zigong, southwestern Sichuan Basin, China, doi:10.1002/jgrb.50310.

Lewis, P. A. W. ed. (1972). Stochastic Point Processes: statistical analysis, theory and applications, Wiley, N.Y.

Lewis, P. A. W. and G. S. Shedler (1979). Simulation of nonhomogeneous Poisson processes by thinning, Naval Res. Logis. Quart., 26, 3, 403-413, https://onlinelibrary.wiley.com/doi/10.1002/nav.3800260304.

Liptzel, R. S. and A. N. Shiryaev (1978). Statistics of Random Processes, Part II, Springer, Berlin.

Mandelbrot, B. B. and J. R. Wallis (1969). Some long-run properties of geophysical records, Water Resor. Res., 5, 321-340.

Llenos, A. L., J. J. McGuire and Y. Ogata (2009). Modeling seismic swarms triggered by aseismic transients., Earth Planet Sci. Lett., 281, 1-2, 59-69, doi:10.1016/j.epsl.2009.02.011.

Llenos, A. L., A. J. Michael, A. M. Allison, M. Shumway et al. (2023). Forecasting the Long-Term Spatial Distribution of Earthquakes for the 2023 U.S. National Seismic Hazard Model Using Gridded Seismicity, Bull. Seism. Soc. Am., 114, 4, doi:10.1785/0120230220.

Lombardi, A. M. and W. Marzocchi (2010a). The ETAS model for daily forecasting of Italian seismicity in the CSEP experiment, Ann. Geophys., 53, 155-164.

Lombardi, A. M. and W. Marzocchi (2010b). A double-branching model applied to long-term forecasting of Italian seismicity (ML ≥ 5.0) within the CSEP project, Ann. Geophys., 53, 31-39.

Matsu'ura, R. S. (1986). Precursory quiescence and recovery of aftershock activities before some large aftershocks, Bull. Earthq. Res. Inst. Univ. Tokyo, 61, 1-65, https://repository.dl.itc.u-tokyo.ac.jp/records/32875.

Matsu'ura, R., N. Hirata and S. Urabe (1995). Quasi-real-time watch of the aftershock activity change of Hyogoken-Nanbu Earthquake-Prediction of Jan. 25 23 h 16 m M4.7 aftershock. Prediction of Jan. 25 23 h 16 m M4.7 aftershock, Report of CCEP 54 (7-15) https://cais.gsi.go.jp/YOCHIREN/report/kaihou54/07-15.pdf.

Matthews, M. V. and P. A. Reasenberg (1988). Statistical methods for investigating quiescence and other temporal seismicity patterns, Pure Appl. Geophys., 126, 357-372, doi:10.1007/BF00879003.

Mearns, E. and D. Sornette (2021). The Great 1976 Tangshan Earthquake: Learning from the 1966-1976 Chinese Prediction Program. Cambridge Scholar Publishing, Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK (2021), ISBN: 1-5275-7164-5 ISBN13: 978-1-5275-7164-8, 235, https://www.cambridgescholars.com/product/978-1-5275-7164-8/.

Mitsui, Y., H. Muramatsu and Y. Tanaka (2021). Slow deformation events between large intraslab earthquakes at the Tonga Trench, Sci. Rep., 11, 257, https://www.nature.com/articles/s41598-020-80728-w#Sec2.

Mitsui, Y., K. Uehara, I. Kosugi and K. Matsuo (2022). Along-strike distribution of seismicity and large slow slip correlated with gravity at the Nankai Trough, Earth and Planet. Sci. Lett., 598, 117824-117824, doi:10.1016/j.epsl.2022.117824.

Mismeci, F. and D. Vere-Jones (1992). A space-time clustering model for historicalearthquakes, Ann Inst. Statist. Math., 44, 1-11, https://www.ism.ac.jp/editsec/aism/pdf/044_1_0001.pdf.

Nakagomi, K., T. Terakawa, S. Matsumoto and S. Horikawa (2021). Stress and pore fluid pressure control of seismicity rate changes following the 2016 Kumamoto, Earth Planets and Space, 73, 11, doi:10.1186/s40623-020-01329-5.

Nanjo, K. Z., H. Tsuruoka, S. Yokoi, Y. Ogata et al. (2012). Predictability study on the aftershock sequence following the 2011 off the Pacific coast of Tohoku, Japan, earthquake: First results, Geophys. J. Int., 191, 2, 653-658, doi:10.1111/j.1365-246X.2012.05626.x.

Neyman, J. and E. Scott (1958). Statistical approach to problems of cosmology, J. R. Statist. Soc., Series B, 20, 1-43.

Nishikawa, T. and S. Ide (2018), Recurring slow slip events and earthquake nucleation in the source region of the M7 Ibaraki-Oki Earthquakes revealed by earthquake swarm and foreshock activity, J. Geophys. Res., doi:10.1029/2018JB015642. https://drive.google.com/file/d/16Ah30 kGNuq3za038b2r_I-LNLSOvkGu5/view.

Nishikawa, T., T. Nishimura and Y. Okada (2021). Earthquake Swarm Detection along the Hikurangi Trench, New Zealand: Insights into the Relationship between Seismicity and Slow Slip Events, J. Geophys. Res., 126, e2020JB020618, doi:10.1029/2020JB020618. doi:10.1029/2020JB020618.

Nishikawa, T. and T. Nishimura (2023). Development of an Epidemic-Type Aftershock-Sequence Model Explicitly Incorporating the Seismicity-Triggering Effects of Slow Slip Events, J. Geophys. Res., doi:10.1029/2023JB02645784).

Nishikawa, T., S. Ide and T. Nishimura (2023). A review on slow earthquakes in the Japan Trench, Progr. Earth Planet. Sci., 10, 1, doi:10.1186/s40645-022-00528-w.

Nishimura, T. (2017). Strain concentrated zones in the Japanese Islands clarified from GNSS data and its relation with active fault and inland earthquakes, Active Fault Research, 46, 33-39, doi:10.11462/afr.2017.46_33.

Nomura, S. and Y. Ogata (2020). Foreshock discrimination models for various prediction schemes and evaluation of their predictive performance. Report of RCCP 103 (12-8), 361-366. https://cais.gsi.go.jp/YOCHIREN/report/kaihou103/12_08.pdf.

Nomura, S. and Y. Ogata (2023). Cluster-based foreshock discrimination model with flexible time horizon and mainshock magnitudes, Progr. Earth Planet. Sci., 10, 20, doi:10.1186/s40645-023-00548-0.

Ogata, Y. (1978). The asymptotic behavior of maximum likelihood estimators for stationary point processes, Ann. Inst. Statist. Math., 30, 1, 243-61, https://www.ism.ac.jp/editsec/aism/pdf/030_2_0243.pdf.

Ogata, Y. (1981). On Lewis' simulation method for point processes, IEEE Trans. Info. Theory, IT-27, 23-31, doi:10.1109/TIT.1981.1056305.

Ogata, Y. (1983). Estimation of the parameters in the modified Omori formula for aftershock frequencies by the maximum likelihood procedure, J. Phys. Earth, 31, 115-124, https://www.jstage.jst.go.jp/article/jpe1952/31/2/31_2_115/_article/-char/en.

Ogata, Y. (1985). Statistical models for earthquake occurrences and residual analysis for point processes, Research Memorandum (Technical report), 288, Inst. Math., Tokyo. https://www.ism.ac.jp/editsec/resmemo-j.html.

Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes, J. Amer. Statist. Soc., 83 401, 9-27, doi:10.1080/01621459.1988.10478560.

Ogata, Y. (1989). Statistical model for standard seismicity and detection of anomalies by residual analysis, Techtonophysics, 169, 159-174, doi:10.1016/0040-1951(89)90191-1.

Ogata, Y. (1992). Detection of precursory relative quiescence before great earthquakes through a statistical model, J. Geophys. Res., 97, B13, 19845-19871, doi:10.1029/92JB00708.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences, Ann. Inst. Statist.l Mat, 50, 2, 379-402, https://www.ism.ac.jp/editsec/aism/pdf/050_2_0379.pdf.

Ogata, Y. (1999). Seismicity analyses through point-process modelling: A review, in Seismicity Patterns, Their Statistical Significance and Physical Meaning M. Wyss, K. Shimazaki and A. Ito eds. Birkhauser Verlag, Basel, Pure Appl. Geophys., 155, 2-4, 471-507, doi:10.1007/s000240050275.

Ogata, Y. (2001a). Exploratory analysis of earthquake clusters by likelihood-based trigger models, Festscrift Volume for Professor Vere-Jones, J. Appl. Probab., 38A, 202-212, ISSN: 0021-9002.

Ogata, Y. (2001b). Increased probability of large earthquakes near aftershock regions with relative quiescence, J. Geophys. Res., 106, B5, 8729-8744, doi:10.1029/2000JB900400.

Ogata, Y. (2004). Space-time model for regional seismicity and detection of crustal stress changes, J. Geophys. Res. 109, B3, B03308, doi:10.1029/2003JB002621.

Ogata, Y. (2005a). Relative quiescence reported before the occurrence of the largest aftershock (M5.8) in the aftershocks of the 2005 earthquake of M7.0 at the western Fukuoka, Kyushu, and possible scenarios of precursory slips considered for the stress-shadow covering the aftershock area. Report of CCEP 74 (10-19) 529-535 https://cais.gsi.go.jp/YOCHIREN/report/kaihou74/10-19.pdf.

Ogata, Y. (2005b). Detection of anomalous seismicity as a stress change sensor, J. Geophys. Res., 110, B5, B05S06, doi:10.1029/2004JB003245.

Ogata, Y. (2005c). Synchronous seismicity changes in and around northern Japan preceding the 2003 Tokachi-oki earthquake of M8.0, J. Geophys. doi:10.1029/2004JB003323.

Ogata, Y. (2006a). Statistical Analysis of Seismicity – updated version (SASeis2006), ISM Computer Science Monographs, No. 33, The Institute of Statistical Mathematics, Tokyo, Japan. https://www.ism.ac.jp/editsec/csm/index.html.

Ogata, Y. (2006a). Monitoring of anomaly in the aftershock sequence of the 2005 earthquake of M7.0 off the coast of western Fukuoka, Japan, by the ETAS model, Geophys. Res. Lett. 33, 1, L01303, doi:10.1029/2005GL024405.

Ogata, Y. (2006b). Seismicity anomaly scenario prior to the major recurrent earthquakes off the east coast of Miyagi Prefecture, northern Japan, Tectonophysics, 424, 291-306, doi:10.1016/j.tecto.2006.03.038.

Ogata, Y. (2006c). Statistical Analysis of Seismicity – updated version (SASeis2006), ISM Computer Science Monographs, No.33, The Institute of Statistical Mathematics, Tokyo, Japan. https://www.ism.ac.jp/editsec/csm/.

Ogata, Y. (2007a). Seismicity and geodetic anomalies in a wide area preceding the Niigata-Ken-Chuetsu earthquake of 23 October 2004, central Japan, J. Geophys. Res. B10301, doi:10.1029/2006JB004697.

Ogata, Y. (2007b). The 40-year history of the Coordinating Committee for Earthquake Prediction (Looking back over the past 10 years) (4). The Institute of Statistical Mathematics, Research Organization of Information and Systems, Inter-University Research Institute. 85-88. https://cais.gsi.go.jp/YOCHIREN/history/40_history.pdf.

Ogata, Y. (2010a). Anomalies of seismic activity and transient crustal deformations preceding the 2005 M7.0 earthquake west of Fukuoka, Pure Appl. Geophys., 167 (8-9), doi:10.1007/s00024-010-0096-y. 8-9), doi:10.1007/s00024-010-0096-y.

Ogata, Y. (2010). Space-time heterogeneity in aftershock activity, Geophys. J. Int., 181 (3), 1575-1592, doi:10.1111/j.1365-246X.2010.04542.x.

Ogata, Y. (2011a). Pre-seismic anomalies in seismicity and crustal deformation: case studies of the 2007 Noto Hanto earthquake of M6.9 and the 2007 Chuetsu-oki earthquake of M6.8 after the 2004 Chuetsu earthquake of M6.8, Geophys. J. Int., 186(1), 331-348, doi:10.1111/j.1365-246X.2011.05033.x. Erratum, Geophys. 188(3), 1454, doi:10.1111/j.1365-246X.2011.05306.x, http://onlinelibrary.wiley.com/doi/10.1111/j.1365-246X.2011.05306.x/abstract.

Ogata, Y. (2011b). Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity, Earth, Planets and Space, 63 (3), 217-229, doi:10.5047/eps.2010.09.001.

Ogata, Y. (2011c). Earthquake forecast in Japanese Islands based on location-dependent b-values and space-time ETAS model. PCCEP 83 (1-2), 5-8. https://cais.gsi.go.jp/YOCHIREN/report/index83.e.html.

Ogata, Y. (2013). A prospect of earthquake prediction research. Statistical Science, 28, 521-541. MR3161586.

Ogata, Y. (2015). Stochastic prediction of earthquakes – A strategy for the research. Proceedings of the Institute of Statistical Mathematics 63(1) 3-27. https://www.ism.ac.jp/editsec/toukei/pdf/63-1-003.pdf.

Ogata, Y. (2017a). Statistics of Earthquake Activity: Models and Methods for Earthquake Predictability Studies, Annual Review of Earth and Planetary Sciences, 45, 497-527, doi:10.1146/annurev-earth-063016-015918.

Ogata, Y. (2017b). On spontaneous seismicity rate in Japan inland. report of CCEP 97 9-12,. https://cais.gsi.go.jp/YOCHIREN/report/kaihou97/01_03.pdf.

Ogata, Y. (2017d). The 50-year history of the Coordinating Committee for Earthquake Prediction (Looking back over the past 10 years) (4). The Institute of Statistical Mathematics, Research Organization of Information and Systems, Inter-University Research Institute. 67-76. https://cais.gsi.go.jp/YOCHIREN/history/CCEP50th/CCEP50th_1_3_1_(12).pdf.

Ogata, Y. (2018). Forecasting of a large earthquake: An outlook of the research, Seismological Res. Lett., 88, 4, 1117-1126, doi:10.1785/0220170006.

Ogata, Y. (2021). Visualizing heterogeneities of earthquake hypocenter catalogs: modeling, analysis, and compensation, Progress in Earth and Planet. Sci., 8, 8, doi:10.1186/s40645-020-00401-8.

Ogata, Y. (2022). Prediction and validation of short-to-long-term earthquake probabilities in inland Japan using the hierarchical space-time ETAS and space-time Poisson process models, Earth Planets and Space, 74, 110, doi:10.1186/s40623-022-01669-4.

Ogata, Y. (2024a). Experiments in prediction of crustal activity (2) – Short-term probability prediction and evaluation of inland earthquakes (2019 – 2024.01.08). Report of CCEP 112(10-4). http://cais.gsi.go.jp/YOCHIREN/report/kaihou112/10_4.pdf.

Ogata, Y. (2024b). Estimating real seismicity rates from nonhomogeneously recorded earthquake catalogs. Seismol. Res. Lett., Special Issue on Statistical Seismology, 1-13, doi:10.1785/0220240153, https://pubs.geoscienceworld.org/ssa/srl/article/doi/10.1785/0220240153/648984/Estimating-Real-Seismicity-Rates-from.

Ogata, Y. and H. Akaike (1982). On linear intensity models for mixed doubly stochastic Poisson and self-exciting point processes. J. R. Statist. Soc. B, 44, 102-107, https://www.jstor.org/stable/2984715.

Ogata, Y., H. Akaike and K. Katsura (1982). The application of linear intensity models to the investigation of causal relations between a point process and another stochastic process, Ann. Inst. Statist. Math., 34, B, 373-387, https://www.ism.ac.jp/editsec/aism/pdf/034_2_0373.pdf.

Ogata, Y. and K. Shimazaki (1984). Transition from aftershock to normal activity: the 1965 Rat Islands earthquake aftershock sequence, Bull. Seism. Soc. Am., 74 (5), 1757-1765, doi:10.1785/BSSA0740051757.

Ogata, Y. and K. Katsura (1986). Point-Process Models with Linearly Parametrized Intensity for Application to Earthquake Data, Journal of Applied Probability, 23, Essays in Time Series and Allied Processes, 291-310.

Ogata, Y. and K. Abe (1991). Some statistical features of the long-term variation of the global and regional seismic activity, International Statistical Reviews, 59, 139-161, https://www.jstor.org/stable/pdf/1403440.pdf.

Ogata, Y. and K. Katsura (1991). Maximum likelihood estimates of the fractal dimension for random spatial patterns, Biometrika, 78, 3, 463-474.

Ogata, Y. and K. Katsura (1993). Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake, Geophys. J. Int., 113, 727-738.

Ogata, Y., T. Utsu and K. Katsura (1995). Statistical features of foreshocks in comparison with other earthquake clusters, Geophys. J. Int., 121, 233-254.

Ogata, Y., L. M. Jones and S. Toda (2003a). When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California, J. Geophys. Res., 108, B6, 2318, doi:10.1029/2002JB002009.

Ogata, Y., K. Katsura and M. Tanemura (2003). Modelling heterogeneous space-time occurrences of earthquakes and its residual analysis, Applied Statistics (J. R. Statist. Soc., series C), 52, 4, 499-509.

Ogata, Y. and K. Katsura (2006). Immediate and updated forecasting of aftershock hazard, Geophys. Res. Lett. 33, 10, L10305, doi:10.1029/2006GL025888.

Ogata, Y. and J. Zhuang (2006). Space-time ETAS models and an improved extension, Tectonophysics, 413, 1-2, 13-23, doi:10.1016/j.tecto.2005.10.016.

Ogata, Y. and S. Toda (2010). Bridging great earthquake doublets through silent slip: On- and offfault aftershocks of the 2006 Kuril Island subduction earthquake toggled by a slow slip on the outer rise normal fault the 2007 great earthquake, J. Geophys. Res. 115 (B06318), doi:10.1029/2009JB006777.

Ogata, Y. and K. Katsura (2012). Prospective foreshock forecast experiment during the last 17 years, Geophys. J. Int., 191(3), 1237-1244, doi:10.1111/j.1365-246X.2012.05645.x.

Ogata, Y., K. Katsura, G. Falcone, K. Z. Nanjo, and J. Zhuang (2013). Comprehensive and topical evaluations of earthquake forecasts in terms of number, time, space, and magnitude, Bull. time, space, and magnitude, Bull. Seismol. Soc. Am. 103(3), 1692-1708, doi:10.1785/0120120063.

Ogata, Y. and K. Katsura (2014). Comparing foreshock characteristics and foreshock forecasting in observed and simulated earthquake catalogs, J. Geophys. Res., 119, 11, 8457-8477, doi:10.1002/2014JB011250.

Ogata, Y. and H. Tsuruoka (2016). Statistical monitoring of aftershock sequences: a case study of the 2015 Mw7.8 Gorkha, Nepal, earthquake, Earth Planets Space, 68, 44, doi:10.1186/s40623-016-0410-8.

Ogata, Y., K. Katsura, H. Tsuruoka and N. Hirata (2018). Exploring magnitude forecasting of the next earthquake, Seismol. Res. Lett., 89, 4, 1298-1304, doi:10.1785/0220180034.

Ogata, Y., K. Katsura, H. Tsuruoka and N. Hirata (2019). High-resolution 3D earthquake forecasting beneath the greater Tokyo area, Earth Planets Space, 71, 113, doi:10.1186/s 40623-019-1086-7.

Ogata, Y., K. Katsura, M. Tanemura, D. Harte and J. Zhuang (2021). Hierarchical Space-Time Point-Process Models (HIST-PPM): Software Documentation https://www.ism.ac.jp/editsec/csm/pdf/csm_035.pdf and Supplementary material for codes. Computer Science Monographs 35, The Institute of Statistical Mathematics.

Ogata, Y. and T. Kumazawa (2024a). Analysis of Aftershock Activity of the 2024 Noto Peninsula M7.6 Earthquake112, 11-12, Report of CCEP, http://cais.gsi.go.jp/YOCHIREN/report/kaihou112/11_12.pdf.

Ogata, Y. and T. Kumazawa (2024b). Inverse Analysis of Seismic Activity Rate Changes with Severe Missing Data – Comparison of aftershock activity Patterns immediately following the 2023 M6.5 and 2024 M7.6 Noto Peninsula Earthquakes, Earth, Planets and Space, 76, 112, doi:10.1186/s40623-024-02057-w.

Omi, T., Y. Ogata, Y. Hirata and K. Aihara (2013). Forecasting large aftershocks within one day after the main shock, Sci. Rep., 3, 2218, doi:1038/srep02218. 10.1038/srep02218.

Omi, T., Y. Ogata, Y. Hirata and K. Aihara (2014). Estimating the ETAS model from an early aftershock sequence, Geophy. Res. Lett. 41, 850-857, doi:10.1002/2013GL058958.

Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes, Ann. Inst. Statist. Math., 31, 145-155. MR0541960, https://www.ism.ac.jp/editsec/aism/pdf/031_1_0145.pdf.

Omori, F. (1894). On after-shocks of earthquakes, J. Coll. Sci. Imp. Univ. Tokyo 7, 111-200.

Pacheco, J. F. and L. R. Sykes (1992). Seismic moment catalog of large shallow earthquakes 1900 to 1989, Bull. Seismol. Soc. Am., 82, 3, 1306-1349, doi:10.1785/BSSA0820031306.

Perez, O. J. and C. H. Scholz (1984). Heterogeneities of the instrumental seismicity catalog (1904-1980) for strong shallow earthquakes, Bull. Seism. Soc. Am., 74, 2, 669-686, doi:10.1785/BSSA0740020669.

Petrillo, G., T. Kumazawa, F. Napolitano, P. Capuano et al. (2024). Fluids-Triggered Swarm Sequence Supported by a Non-stationary Epidemic-Like Description of Seismicity, Seismol. Res. Lett., doi:10.1785/0220240056.

Reasenberg, P. A. and L. M. Jones (1989). Earthquake hazards after a mainshock in Califomia, Science, 243 1173-1176.

Rubin, I. (1972). Regular point processes and their detection, IEEE Trans. Info. Theory, 18, 547-557, doi:10.1109/TIT.1972.1054897.

Rathbun, S. L. (1993). Modeling marked spatio-temporal point patterns, Bull. Int. Statistical Institute, 5, 2, 379-396.

Rathbun, S. L. (1996). Asymptotic proprerties of the maximum likelihood estimator for spatio-temporal point processes, J. Stat. Plan. Inference 51, 55-76. https://bemlar.ism.ac.jp/zhuang/Refs/Refs/rathbun1996mle.pdf.

Schorlemmer, D, M. J. Werner, W. Marzocchi, T. H. Jordan et al. (2018). The Collaborative for the Study of Earthquake Predictability: Achievements and Priorities, Seismol. Res. Lett., 89, 4, 305-1313, doi:10.1785/0220180053.

Suyehiro, S. (1966). Difference between aftershocks and foreshocks in the relationship of magnitude to frequency of occurrence for the great Chilean earthquake of 1960, Bull. Seismol. Soc. Amer., 56, 185-200.

The Institute of Statistical Mathematics (2006). Report of the External Review Committee on the Statistical Seismology Research Project, https://www.ism.ac.jp/evaluation/index_j/toukei_jisingaku2006.pdf.

The Institute of Statistical Mathematics (2023). SAPP Statistical Analysis of Point Processes with R, Jun 9, 2023 (Ver. 1.0.9-1), https://jasp.ism.ac.jp/ism/SAPP/index_e.html.

U.S. Geological Survey (2022). UCERF3--ETAS, http://wgcep.org/UCERF3-ETAS.html.

Utsu, T. (1965). A Method for Determining the Value of b in a Formula log n = a – bM showing the Magnitude-Frequency Relation for Earthquakes. Geophys. Bull. Hokkaido Univ., 13, 99-103, doi:10.14943/gbhu.13.99, https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/13887/1/13_p99-103.pdf.

Utsu, T. (1970a). Problems related to temporal distribution of earthquakes (Part 3) I. Temporal characteristics of aftershocks, foreshocks, and clusters, Geophys. Bull. Hokkaido Univ., 23, 49-71 doi:10.14943/gbhu.23.49, https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/13973/1/23_p49-71.pdf.

Utsu, T. (1969). Aftershocks and earthquake statistics (I): Some parameters which characterize an af-tershock sequence and their interaction, J. Faculty Sci., Hokkaido University, Series VII, Geophysics, 3, 129-195, https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/8683/1/3%283%29_p129-195.pdf.

Utsu, T. (1970b). Aftershocks and earthquake statistics (II): Further investigation of aftershocks and other earthquakes sequence based on a new classification of earthquake sequences, J. Faculty Sci., Hokkaido University, Series VII, Geophysics, 3, 379-441, https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/8684/1/3%284%29_p197-266.pdf.

Utsu, T. (1979). Calculation of the Probability of Success of an Earthquake Prediction (In the Case of the Izu-Oshima-Kinkai Earthquake of 1978), Report of CCEP, 21, 7-4, 164-166, https://cais.gsi.go.jp/YOCHIREN/report/kaihou21/07_04.pdf.

Utsu, T. (1985). Catalog of large earthquakes in the region of Japan from 1885 through 1980: correction and supplement. Bull. Earthq. Res. Inst., Univ. Tokyo 60:639-642 in Japanese with English summary and figure captions, https://progearthplanetsci.springeropen.com/articles/10.1186/s40645-020-00401-8.

Utsu, T. (1999). Seismicity Studies: A Comprehensive Review, 876, University of Tokyo Press, ISBN4-13060728-6 printed in Japan.

Utsu, T. and A. Seki (1954). A Relation between the area of the after-shock region and the energy of the main-shock. Zisin ii 7: 233-240, https://www.jstage.jst.go.jp/article/zisin1948/7/4/7_4_233/_article/-char/ja/.

Utsu, T., Y. Ogata and R. S. Matsu'ura (1995). The centenary of the Omori formula for a decay law of the aftershock activity, Journal of Physics of the Earth, 43, 1-33. https://www.jstage.jst.go.jp/article/jpe1952/43/1/43_1_1/_article/-char/en.

Van Stiphout, T., J. Zhuang and D. Marsan (2012). Seismicity declustering, Community Online Resource for Statistical Seismicity Analysis, doi:10. 5078/corssa-52382934, Available at http://www.corssa.org/export/sites/corssa/.galleries/articles-pdf/vanStiphout_et_al.pdf.

Vere-Jones, D. (1970). Stochastic models for earthquake occurrences (with discussion), J. R. Statist. Soc.: Series B (Methodological), 32, 1, 1-45, doi:10.1111/j.2517-6161.1970.tb00814.x.

Vere-Jones, D. (1978). Earthquake prediction – A statistician's view, J. Phys. Earth. 26, 129-146, https://www.jstage.jst.go.jp/article/jpe1952/26/2/26_2_129/_pdf.

Vere-Jones, D. (2001). The marriage of statistics and seismology, J. Appl. Probab., 38A, 1-5, in Probability, Statistics and Seismology, Applied Probability Trust, https://www.jstor.org/stable/3215861.

Vere-Jones, D. (2006). The development of statistical seismology: A personal experience, Tectonophysics, 413, 1-2, 5-12, doi:10.1016/j.tecto.2005.10.005.

Vere-Jones, D. and R. B. Davies (1966). A statistical survey of earthquakes in the main seismic region of New Zealand, Part 2, Time series analyses, New Zealand J. Geology and Geophys. 9, 251-284, doi:10.1080/00288306.1966.10422815.

Wyss, M. and R. E. Habermann (1988). Precursory seismic quiescence, Pure Appl. Geophys., 126, 2-4, 319-332, doi:10.1007/BF00879001.

Zhuang, J., Y. Ogata and D. Vere-Jones (2002). Stochastic declustering of space-time earthquake occurrences, J. Am. Stat. Assoc., 97, 369-380, doi:10.1198/016214502760046925.

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

Zhuang , J., D. Vere-Jones, H. Guan, Y. Ogata and L. Ma (2005). Preliminary analysis of observations on the ultra-low frequency electric field in a region around Beijing (with data), Pure Appl. Geophys., 162, 1367-1396, doi:10.1007/s00024-004-2674-3.

***CORRESPONDING AUTHOR: Yosihiko OGATA**,
The Institute of Statistical Mathematics, 10-3 Midori-Cho, Tachikawa, Tokyo, Japan 190-8562
e-mail: ogata@ism.ac.jp