Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes

YOSIHIKO OGATA*

This article discusses several classes of stochastic models for the origin times and magnitudes of earthquakes. The models are compared for a Japanese data set for the years 1885–1980 using likelihood methods. For the best model, a change of time scale is made to investigate the deviation of the data from the model. Conventional graphical methods associated with stationary Poisson processes can be used with the transformed time scale. For point processes, effective use of such *residual analysis* makes it possible to find features of the data set that are not captured in the model. Based on such analyses, the utility of seismic quiescence for the prediction of a major earthquake is investigated.

KEY WORDS: Akaike information criterion; Epidemic-type models; Conditional intensity; Likelihood; Marked point process; Seismic quiescence; Trigger models.

1. INTRODUCTION

It is widely accepted that some time after the occurrence of a major earthquake the aftershock activity dies off and background seismicity surpasses the aftershock activity. Prior to the next major earthquake, preseismic quiescence and then foreshocks are expected to appear in the focal region (Mogi 1968). It is not easy, however, to identify foreshocks before the occurrence of a main shock. In particular, foreshocks tend to be less numerous than aftershocks. Thus the seismic quiescence and related seismic gap have been studied by many seismologists for the purpose of earthquake prediction [e.g., see Kanamori (1981) and some other authors in the same volume].

Some have questioned the usefulness of seismic quiescence and suggested that it can be considered a mere result of the decaying activity of aftershocks from the last major earthquake (see Lomnitz 1982; Lomnitz and Nava 1983). The latter paper discusses the similarity between certain observed earthquake sequences and sequences featuring seismic quiescences generated from stochastic models that are based on simple statistical assumptions such as the Gutenberg–Richter law of magnitude frequency, Omori's decaying frequency law of aftershocks, and Båth's law.

To investigate this issue we have to investigate seismic quiescence by means of a quantitative comparison with background seismicity and aftershock activity. There have been some attempts to develop statistical techniques for the definition and detection of quiescence. For example, Habermann and Wyss (1984) developed a statistical test procedure to detect quiescences for seismic activity with data from which the effect of aftershocks is removed. Aftershocks, however, constitute the greatest proportion of shocks in an earthquake catalog, and there have been many detailed studies on aftershock sequences (e.g., see Utsu 1969, 1970, 1971, 1972). If these are effectively considered in the analysis, aftershocks can give us useful information for understanding the whole cycle of seismic activity (see Matsu'ura 1986; Ogata and Shimazaki 1984; Okada 1978).

This article selects a statistical model for the standard activity of earthquake series by comparing several possible models using likelihood methods. In the next section some possible stochastic models for seismic activity derived from empirical studies, mainly of aftershock statistics, are reviewed and discussed. In Section 3.1 an earthquake data set for which the models seem applicable is considered. Model comparisons, using this data set, are presented in Section 3.2.

Another feature of this article is the systematic use of a new kind of *residual* analysis for point process data, described in Section 3.3, which involves a change of time scale that depends on the conditional intensity rate of the estimated model. The model selected fits the data well enough that the residuals reveal certain events more clearly than the original data. This makes possible a precise quantitative definition of seismic quiescence in Section 4 as well as the detection of a stretch of outliers. Then it is demonstrated that seismic quiescence has some utility for statistical prediction of a major earthquake. The final section contains some additional discussion and some conclusions.

2. STATISTICAL MODELS FOR THE EARTHQUAKE PROCESS

2.1 Empirical Studies on Aftershock Statistics

The statistical properties of the occurrence of aftershocks have long been some of the main objects of seismological studies in connection with the processes of

* Yosihiko Ogata is Associate Professor, Institute of Statistical Mathematics (I.S.M.), Minami-Azabu 4-6-7, Minato-ku, Tokyo 106. This work is based on a seminar given at the Earthquake Research Institute (E.R.I.), University of Tokyo, October 1984, and this article was completed while the author was an invited research associate of the same institute in 1986–1987. The author is most grateful to Koichi Katsura of the I.S.M. for generous help, especially in drawing the necessary figures, and to Dr. Mark Berman of the Division of Mathematics and Statistics, Commonwealth Scientific and Industrial Research Organisation, Australia, for the stimulating discussion of the absolute goodness of fit of models relating to the present residual analysis. He also thanks Professors Shimazaki and Abe of the E.R.I., University of Tokyo, and Professor Oike of the Disaster Prevention Research Institute, Kyoto University, for their useful discussions and comments on seismicity, especially on quiescence. Professor Utsu of the E.R.I. kindly let the author know his identification of the main shocks, foreshocks, and aftershocks in his earthquake catalog. Finally, the author would like to thank the editors, anonymous referees, and Dr. David Findley of the U.S. Bureau of the Census, whose comments led to a significant revision of the article.

Journal of the American Statistical Association
March 1988, Vol. 83, No. 401, Applications

earthquake generation. Distributions of aftershocks in time, space, and magnitude are usually included in a general survey of destructive earthquakes. In this section some studies of aftershock statistics are reviewed briefly. Based on this review, some plausible point process models are subsequently developed to describe the standard activity of earthquake series.

The frequency of aftershocks per unit time interval (one day, one month, etc.) is well described by the modified Omori formula (Utsu 1961),

$$n(t) = K/(t + c)^p \qquad (K, c, p; \text{parameters}), \tag{1}$$

where t is the time since occurrence of the shock. K depends on the lower bound of the magnitude of aftershocks counted in $n(t)$, whereas p and c are known to be independent of this choice of lower bound. Utsu (1969, 1970) extensively investigated the aftershock sequences of this century's earthquakes in Japan. Plotting observed $n(t)$ versus time t on a log–log-scaled plane, he showed that the graph tends to a straight line whose slope is an estimate of p. With such plots, for example, Utsu (1969) discovered the remarkable fact that the aftershock activity of the Nobi earthquake of 1891 continued for at least 80 years with a continuously decreasing rate of occurrence.

It is known empirically that the magnitudes of the earthquakes follow roughly an exponential distribution.

$$\Pr\{\text{Mag} > M\} = e^{-\beta M}. \tag{2}$$

In the seismological literature this relationship, called Gutenberg–Richter's law of magnitude of frequency, is most often written in the form

$$\log_{10} F(M) = a - bM, \tag{3}$$

where $F(M)$ refers to the frequency of earthquakes with magnitudes not smaller than M [see Gutenberg and Richter (1944); see also Fig. 3 later in this article]. The constant b, called b value, is then related to β by the equation $\beta = b \cdot \log_e 10$ and commonly takes a value near unity.

Utsu (1971, pp. 420–427) studied theoretically the relation between the size of aftershock activity and the magnitude of the main shock, based on the aforementioned law of magnitude frequency (3) and some theoretical properties of order statistics. For a sequence of aftershocks triggered by a main shock with magnitude M, he found the relation

$$\log_{10} A = \alpha M + \text{constant}, \tag{4}$$

where A is the total number of aftershocks in the sequence and α is a constant. From this and the magnitude frequency law (2) for the main shocks, the frequency distribution of A takes the form

$$f(A) = \text{constant } A^{-q} \tag{5}$$

for some positive q. Empirical evidence in support of (5) was provided by plotting the logarithm of the cumulative frequency distribution of the number of aftershocks versus the magnitude of the corresponding main shock (Utsu 1971, fig. 133). For seismological readers particularly, I point out that (4) is similar in form to some widely used empirical formula like that given by Utsu and Seki (1955),

$$\log_{10} S = 1.02M - 4.0, \tag{6}$$

relating the aftershock area $S(km^2)$ and the magnitude M of the main shock for shallow earthquakes in Japan and its vicinity.

Combining (1) and (4), Utsu (1970, p. 229) gave the standard equation for the rate of occurrence of aftershocks having magnitude M_r and larger, t days after a large shallow earthquake of magnitude M_0:

$$n(t) = 10^{.85(M_0 - M_r) - 1.83}/(t + .3)^{1.3} \tag{7}$$

per day.

For constructing statistical models of earthquake occurrence, some authors (Lomnitz 1966; Lomnitz and Nava 1983; Utsu 1972) assumed that there was no detailed causality or interaction between the occurrence times of aftershocks, so the time sequence of aftershocks could be treated as a nonstationary Poisson process. Vere-Jones (1975) suggested that this assumption is mainly due to the analysis of Jeffreys (1938), in which daily numbers of aftershocks of the Tango earthquake of 1927 were tested by evaluating chi-squared statistics to establish the independence of the occurrence numbers. A similar result was obtained in Lomnitz and Hax (1966) by analyzing autocorrelations. I will argue, however, that these analyses do not really establish the independence of the precise times at which the aftershocks occur. Furthermore, the so-called secondary aftershocks, or the aftershocks of an aftershock (see Utsu 1970, pp. 214–216), are not explained by such models. According to Ogata (1983) the aftershocks of the Tango earthquake had at least two secondary sequences. In addition, a few analyses suggesting the interdependence of aftershocks can be obtained by using the theory of runs tests (Utsu 1970, pp. 217–229).

Although the independence of the occurrence of main shocks has been assumed in many models, some papers discuss the migration of large earthquakes and the causal relationships between seismic activity in different geophysical regions (Mogi 1968, 1973; Ogata and Katsura 1986; Utsu 1975). Many such reports have been given constantly at the meetings of the Seismological Society of Japan.

2.2 Trigger Models

The model discussed by Lomnitz and Nava (1983) is related to the trigger model suggested by Vere-Jones and Davies (1966), which Vere-Jones (1970) described as a special case of the Neyman–Scott clustering model. This model assumes a series of primary events (main shocks) distributed completely randomly in time. Each of these primary events may generate a secondary series of events (aftershocks). It is assumed that the conditional probability that an aftershock (of magnitude above a certain level) will occur in the small time interval $(t, t + dt)$, triggered by a main shock at time t_0, is equal to $\sigma_{t_0}(t)\, dt$

with

$$\sigma_{t_0}(t) = \xi \cdot f(t - t_0), \qquad t \geq t_0$$
$$= 0, \qquad t < t_0, \tag{8}$$

where $f(t)$ is a normalized function; that is, $\int_0^\infty f(t)\,dt = 1$ and ξ is the average number of secondary events produced by a primary event at t_0.

Vere-Jones and Davies (1966) suggested that the inverse power type of decay function

$$f(t) = (p - 1)c^{p-1}/(c + t)^p, \tag{9}$$

which arises from normalizing (1), gives a better fit than the negative exponential function. They assumed further that ξ itself is a random variable with finite mean and variance, probably as a technical requirement for the estimation of the parameters using the second-moment properties. In Lomnitz and Nava (1983), ξ is taken to be proportional to $M_m - M_r$, where M_m is the magnitude of the main shock and M_r is the cutoff magnitude, which together with the law of magnitude frequency implies that ξ has a negative exponential distribution. In this article, I will consider only a restricted form of trigger model, which will be described later.

2.3 Epidemic-Type Model

Another type of model appeared in applications to population genetics. Kendall (1949) introduced an age-dependent birth and death process such that for any individual of age x alive at time t, for the next interval $(t, t + dt)$ there are probabilities $g(x)\,dt$ of a birth and $h(x)\,dt$ of a death, independently for each individual. Hawkes (1971) considered the self-exciting process, which is a birth process [i.e., $h(x) = 0$] allowing immigration at a rate μ per unit time. He defined the process by means of the conditional intensity rate

$$\begin{aligned}\lambda(t) &= \mathrm{E}[dN(t) \mid \text{history of } N(s) \text{ at time } t]/dt \\ &= \lim_{\Delta \to 0} \Delta^{-1} \\ &\quad \times \Pr\{\text{event in } (t, t + \Delta) \mid \text{history of} \\ &\qquad N(s) \text{ at time } t\} \\ &= \mu + \sum_{t_i < t} g(t - t_i) = \mu + \int_0^t g(t - s)\,dN(s),\end{aligned} \tag{10}$$

where $N(t)$ is the cumulative number of events, $\{t_i\}$ in $(0, t]$. This process may also be viewed as a cluster process, different from the Neymann–Scott type, in which the process $N_c(t)$ of birth times of cluster centers [see Eq. (18)] is a Poisson process of rate μ formed by the arrival of immigrants. Associated with each event of $N_c(t)$ we have a cluster of subsidiary events formed by the births of all of the descendants of all generations of the immigrant (Hawkes and Oakes 1974). One of the differences between this model and the trigger model is that the latter includes only the first generation offspring whereas in the former all events have the possibility of possessing offspring.

Extending (10) to a multivariate point process, $\{t_i^m\}$, Hawkes (1971) also defined the mutually exciting model

$$\begin{aligned}\lambda_j(t) &= \mathrm{E}[dN_j(t) \mid \text{histories of } N_m(s) \\ &\qquad \text{for all } m \text{ at time } t]/dt \\ &= \mu_j + \sum_m \sum_{t_i^m < t} g_{jm}(t - t_i^m) \\ &= \mu_j + \sum_m \int_0^t g_{jm}(t - s)\,dN_m(s)\end{aligned} \tag{11}$$

for the discrete magnitude values of j and m. If we assume that the simplest forms $g_{jm}(t) = c(m)g_j(t)$ and consider the superposition $N(t) = \Sigma_m N_m(t)$ of the point process components, then the conditional intensity $\lambda(t) = \Sigma_j \lambda_j(t)$ is given by

$$\lambda(t) = \mu + \sum_{t_i < t} c(m_i) g(t - t_i), \tag{12}$$

where t_i is the occurrence time of the superposition $N(t)$, m_i is the corresponding magnitude of t_i, and $g(t) = \Sigma_j g_j(t)$. Further, $\mu = \Sigma \mu_m$ can be considered as a base rate that prevents the process from dying out. The model (12) coincides with the tagged Klondike-type model described in Lomnitz (1974) for earthquake series $\{(t_i, m_i)\}$ with $m_i \geq M_r$, where M_r is the cutoff magnitude. Lomnitz suggested the use of

$$g(t) = ae^{-\alpha t} \tag{13}$$

in view of Boltzman's theory of elastic aftereffect. Here I would also like to consider the model

$$g(t) = K/(t + c)^p, \tag{14}$$

which corresponds to (1). A technical extension of (13),

$$g(t) = \sum_{k=1}^{K} a_k t^{k-1} e^{-\alpha t}, \tag{15}$$

is proposed in Ogata and Akaike (1982) and also in Vere-Jones and Ozaki (1982). The pioneering application of the model (8) (including a trigger model as a special case) to earthquake data was carried out by Hawkes and Adamopoulos (1973), where a mixture of two exponentials is considered as an extension of (10) and a certain approximated log-likelihood is used to fit the model.

For the $c(m)$ in (12) I propose the use of

$$c(m) = e^{\beta(m - M_r)} \tag{16}$$

rather than $\beta(m - M_r)$ as given for ξ in (8) by Lomnitz and Nava (1983), since (16) seems to be consistent with (4)–(7) for the descendants of the first generation (primary aftershocks). The parameter β here measures the effect of magnitude in the production of descendants and is useful in characterizing the earthquake sequences quantitatively in relation to the classification into seismic types made by Mogi (1963) and Utsu (1970) (see also Ogata 1987). For example, earthquake swarms have small β values, and a value of β around 2 might be used as a standard for the aftershock activity in Japan and its vicinity [see Eqs. (6) and (7)].

2.4 Parameterization of the Models

To calculate log-likelihoods we need to describe the models in terms of the parameterized conditional intensity. Here a class of competitive models is summarized in such terms. First a unified form for the epidemic type of models is provided, as follows:

$$\lambda(t) = \mu + \sum_{t_i<t} g(t - t_i)e^{\beta(m_i - M_r)}$$
$$= \mu + \int_{M_r}^{\infty} \int_0^t g(t - s)e^{\beta(m - M_r)}\, dN(s, m), \quad (17)$$

where $\lambda(t)$ is the conditional intensity for the occurrence times of the point process, M_r is the cut off magnitude, and $dN(s, m)$ gives mass 1 to each shock (t_i, m_i).

Thus the parametric models (17) will be compared, with $g(t - s)$ defined by (14) or by (15) for integer values of κ [(13) is a special case in an interval]. In addition, we can examine the case in which $\beta = 0$ to determine whether the magnitude really affects the cluster size of the corresponding aftershocks; see also the last paragraph in Section 2.3.

Now I motivate the "restricted" trigger models, which will be considered subsequently. The estimation and comparison among the original trigger models was carried out by the methods using only second-moment properties, such as goodness of fit of some graphical statistics like the variance–time curve, the hazard function, the periodogram, and so on (see, e.g., Vere-Jones 1970; Vere-Jones and Davies 1966). The actual likelihood of such Neyman–Scott types of cluster models can only be calculated by an extraordinarily complicated computation allowing all possible choices of t_i for cluster centers (see, e.g., Baudin 1981) unless each shock (t_i, m_i) is identified in advance as a main shock or aftershock (foreshocks here are included in main shocks). The conventional approach to such identification is usually based on the observation of spatiotemporal plots. When the application of the trigger model is applied to such an identified data set, we call the model a *restricted trigger model*. That is to say, shocks are given by the decomposition

$$dN(t, m) = dN_c(t, m) + dN_a(t, m), \quad (18)$$

where N_c and N_a stand for the primary and secondary events, respectively. Assuming that the distribution of main shocks is stationary and random with rate μ, we have

$$\lambda_c(t) = \mu. \quad (19)$$

Corresponding to (8), the conditional intensity rate of an aftershock is given by

$$\lambda_a(t) = \int_{M_r}^{\infty} \int_0^t g(t - s)e^{\beta(m - M_r)}\, dN_c(s, m)$$
$$= \sum_{t_i^c<t} g(t - t_i^c)e^{\beta(m_i^c - M_r)}, \quad (20)$$

where M_r is the reference magnitude. Thus the intensity of the restricted trigger model is given by the combination of (19) and (20),

$$\lambda(t) = \lambda_c(t) + \lambda_a(t)$$
$$= \mu + \int_{M_r}^{\infty} \int_0^t g(t - s)e^{\beta(m - M_r)}\, dN_c(s, m)$$
$$= \mu + \sum_{t_i^c<t} g(t - t_i^c)e^{\beta(m_i^c - M_r)}. \quad (21)$$

Technically, (21) can be obtained from (17) by replacing the quantities $e^{\beta(m_i^a - M_r)}$ for aftershocks (t_i^a, M_i^a) by 0. I will compare (13) and (14) by using this restricted model and the epidemic model.

3. ANALYSIS

3.1 The Data and Their Features

Utsu (1982) compiled and published a complete catalog of earthquakes of magnitude 6 or more that occurred in Japan and its vicinity from 1885 through 1980. This catalog includes occurrence times, coordinates of epicenters, depths, and some other information. In addition, each shock has been classified as a main shock, a foreshock, or an aftershock. This classification was not published in Utsu (1982), but the author has kindly provided me with it privately. Thus it is also possible for us to get the log-likelihood values of the restricted trigger models (21), assuming that foreshocks are included in the category of main shocks.

I will consider shallow earthquakes of less than 100 kilometers depth with $M \geq 6.0$ in the polygonal region of the Off Tohoku district, with vertices at the points (42°N, 142°E), (39°N, 142°E), (38°N, 141°E), (35°N, 140.5°E), (35°N, 144°E), (39°N, 146°E), and (42°N, 146°E) (see Fig. 1), which is approximately between the lines of the Japan Trench and the aseismic front proposed by Yoshii (1975). This is a part of the northwestern Pacific seismic belt, where the Pacific plate is subducting beneath northeastern Japan in the Eurasian Plate, and is one of the most active seismic areas in Japan.

In this area, 483 shocks with magnitude 6 or more have occurred in the past 96 years, which means that about 5 such shocks occur on average each year. These data are listed in Table 1. Figure 2 shows the plots of the occurrence times of shocks with their magnitudes. The graphs of the cumulative number of shocks and cumulative square root of released energy are given in the same figure, where for the energy released by a shock with magnitude M I have used the relation $E = 10^{11.8+1.5M}$ (erg) according to Gutenberg and Richter (1954). From Figure 2 we may assume that the seismic activity here is stationary. An actual test of this assumption is not easy, especially when the process has long-range correlation (see Fig. 7). Nevertheless, stationary models are considered here in the prior belief that such geophysical activity for a long time span should be stationary. The magnitude distribution for this region is shown in Figure 3, which clearly supports the Gutenberg–

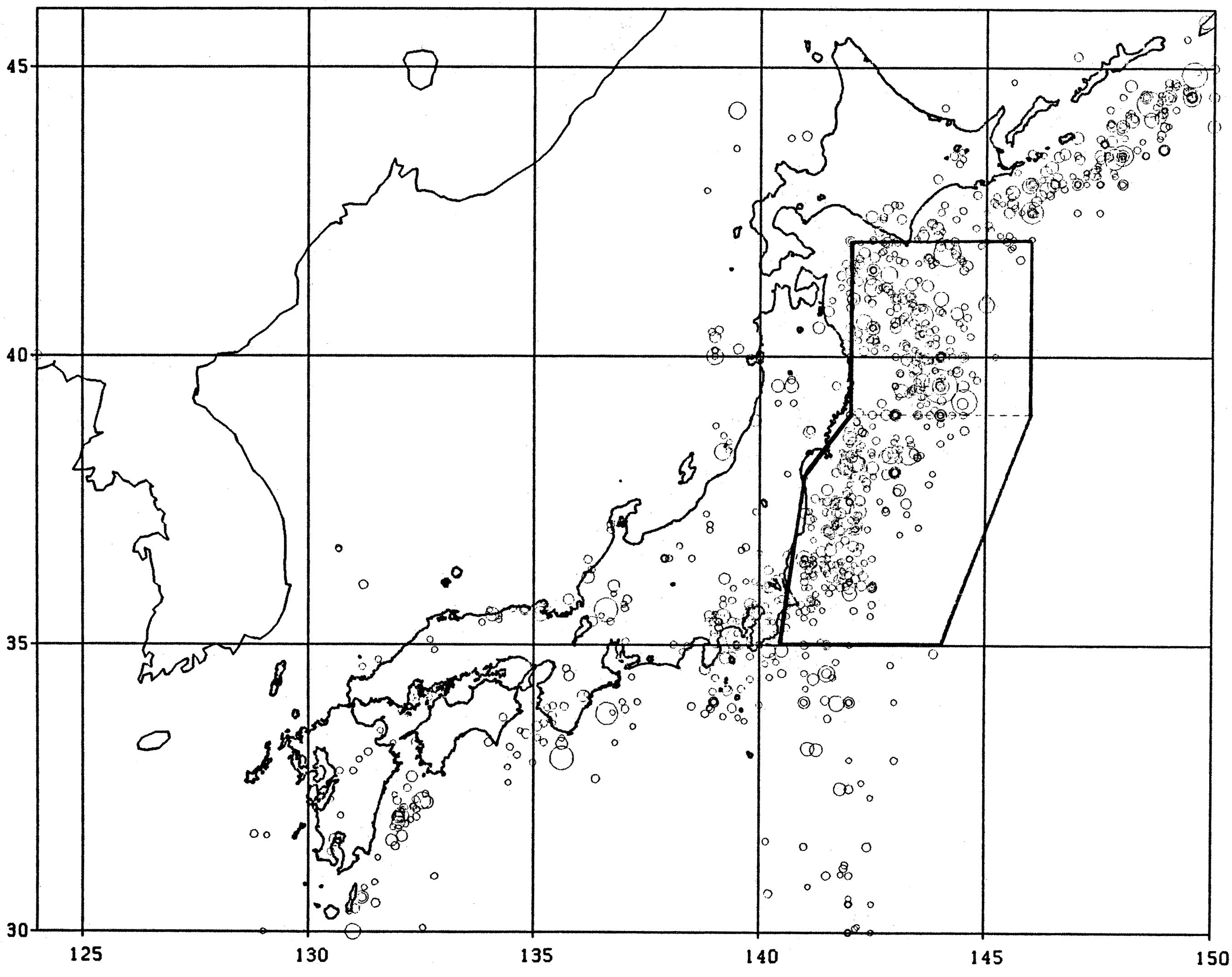

Figure 1. The Spatial Distribution of Large Earthquakes (M ≥ 6.0) 1895–1980 in Japan and Its Vicinity. The sizes of the open circles correspond to the sizes of the earthquakes. The data set for the polygonal region is considered.

Richter law of magnitude frequency [see (2) and (3), Sec. 2.1]. The fact that the graph of points associated with magnitudes near 6.0 lies on the straight line suggested by the Gutenberg–Richter law demonstrates to seismologists that the catalog includes essentially all of the events of this magnitude that occurred.

Using the intervals $X_1, X_2, \ldots$ between successive earthquakes, the log-survivor function, $\log P\{X > x\}$, is plotted in Figure 4. The estimated variance–time curve, $V(s) = \mathrm{var}\{N(t, t + s)\}$, is given in Figure 5. The empirical plot of the intensity of the so-called Palm distribution,

$$m_f(s) = \Pr\{\text{event in } (t + s, t + s + \Delta s) \mid \text{event at } t\}/\Delta s,$$

for some small Δs, is given in Figure 6. All of these plots suggest the strongly clustering nature of the occurrence times of shocks. In particular, Figure 6, with confidence error bounds assuming the Poisson model, suggests so-called long-range correlation. Indeed, if we plot an estimate of $| m_f(s) - \lambda_0 |$ (λ_0 is the mean intensity rate) on a log–log scale, we get Figure 7, which suggests inverse power decay of autocovariance at large time lags. This may support (14) rather than (13) and (15) for the time decay of the response function.

3.2 Comparison of Restricted Trigger Models and Epidemic Models

To compare the epidemic models with the restricted trigger models, I make use of likelihood analysis. Before writing the likelihood of the models, let us consider the assertion of Lomnitz (1982) and Lomnitz and Nava (1983) that the distribution of magnitude frequency is not affected by the occurrence times of shocks, especially just before large shocks. This means that the time series of magnitudes $\{m_i\}$ of earthquakes is independent of their occurrence times $\{t_i\}$. Based on this, the full likelihood of $\{(t_i, m_i)\}$ can be written in the following form:

$$\prod_{j=1}^{N} f_j(m_j \mid m^{(j-1)}; \eta) \prod_{j=1}^{N} g_j(t_j \mid m^{(j-1)}, t^{(j-1)}; \theta), \quad (22)$$

where $t^{(j)} = (t_1, \ldots, t_j)$ and $m^{(j)} = (m_1, \ldots, m_j)$ and

Table 1. *Shallow Shocks (M ≥ 6.0) in OFF Tohoku Area for 1885–1980*

NO	YEAR	MO	DY	HR	MN	MAG	C	NO	YEAR	MO	DY	HR	MN	MAG	C	NO	YEAR	MO	DY	HR	MN	MAG	C
1	1885	2	9	2	0	6.0	0	84	1908	1	15	21	56	6.9	0	167	1923	5	31	14	55	6.2	1
2	1885	6	11	9	20	6.9	0	85	1908	1	18	1	5	6.0	0	168	1923	6	2	2	24	7.3	0
3	1885	7	29	5	30	6.0	0	86	1908	2	5	21	7	6.0	0	169	1923	6	2	5	14	7.1	2
4	1885	10	30	20	30	6.2	0	87	1908	6	27	23	21	6.1	0	170	1923	6	7	2	36	6.2	2
5	1885	12	7	13	2	6.3	0	88	1908	11	22	16	15	6.4	0	171	1923	9	2	18	49	6.3	2
6	1885	12	19	18	26	6.0	2	89	1909	9	17	4	39	6.8	0	172	1923	11	18	5	40	6.3	1
7	1886	4	13	5	44	6.3	0	90	1910	1	22	8	25	6.0	0	173	1923	12	27	23	39	6.4	0
8	1886	7	2	12	33	6.3	2	91	1910	5	9	18	53	6.0	1	174	1924	2	3	7	25	6.3	0
9	1887	5	29	0	50	6.4	0	92	1910	5	10	22	56	6.1	0	175	1924	5	31	21	2	6.3	1
10	1887	5	29	1	10	6.2	2	93	1910	5	12	12	22	6.0	2	176	1924	5	31	21	4	6.4	1
11	1888	2	5	0	50	7.1	0	94	1910	10	13	23	56	6.3	0	177	1924	8	6	23	22	6.3	0
12	1888	11	24	2	3	6.5	0	95	1912	1	4	4	4	6.1	0	178	1924	8	15	3	2	7.1	0
13	1889	3	31	6	42	6.6	0	96	1912	1	9	6	21	6.1	0	179	1924	8	15	8	27	6.7	2
14	1890	11	17	9	31	6.3	0	97	1912	6	8	13	41	6.6	0	180	1924	8	17	10	45	6.3	2
15	1891	4	7	9	49	6.7	0	98	1912	12	9	8	50	6.6	0	181	1924	8	17	11	10	6.6	2
16	1891	5	5	8	16	6.2	0	99	1913	2	20	17	58	6.9	0	182	1924	8	25	23	31	6.7	2
17	1891	7	21	20	19	7.0	0	100	1913	5	22	5	36	6.1	1	183	1925	2	7	2	11	6.0	0
18	1892	10	22	19	9	6.0	0	101	1913	5	29	19	14	6.4	0	184	1925	4	20	5	24	6.3	0
19	1894	2	25	4	18	6.8	0	102	1913	10	3	9	17	6.1	1	185	1925	6	2	14	18	6.4	0
20	1894	3	14	18	15	6.0	2	103	1913	10	11	18	10	6.9	0	186	1925	11	10	23	44	6.0	0
21	1894	8	29	19	55	6.6	0	104	1913	10	13	2	5	6.6	2	187	1926	4	7	4	33	6.3	0
22	1894	11	28	1	5	7.1	0	105	1914	2	7	15	50	6.8	0	188	1926	5	27	4	45	6.4	0
23	1894	12	1	18	37	6.3	0	106	1914	12	26	3	18	6.1	0	189	1926	9	5	0	37	6.8	0
24	1896	1	9	22	17	7.5	0	107	1915	3	9	0	29	6.8	0	190	1926	10	3	17	25	6.4	0
25	1896	1	10	5	52	6.0	2	108	1915	4	6	5	25	6.0	1	191	1926	10	19	9	29	6.2	0
26	1896	1	10	11	25	6.3	0	109	1915	4	6	14	32	6.2	0	192	1926	11	11	12	1	6.1	2
27	1896	2	23	19	42	6.1	2	110	1915	4	25	2	9	6.4	0	193	1927	1	18	6	58	6.4	0
28	1896	3	6	23	52	6.0	2	111	1915	5	28	2	26	6.0	2	194	1927	3	16	15	52	6.4	2
29	1896	4	11	23	0	6.0	2	112	1915	6	5	6	59	6.7	0	195	1927	7	30	23	18	6.4	0
30	1896	6	15	19	32	8.5	0	113	1915	7	9	7	21	6.4	0	196	1927	8	6	6	12	6.7	0
31	1896	6	16	4	16	7.5	2	114	1915	10	13	6	30	6.8	0	197	1927	9	30	16	38	6.3	0
32	1896	6	16	8	1	7.5	2	115	1915	10	14	4	43	6.2	2	198	1928	5	27	18	50	7.0	0
33	1896	7	29	17	44	6.1	2	116	1915	10	15	1	28	6.1	2	199	1928	5	29	0	35	6.7	2
34	1896	8	1	11	49	6.5	0	117	1915	10	15	3	40	6.3	2	200	1928	6	1	22	12	6.5	2
35	1896	9	5	23	7	6.5	2	118	1915	10	16	1	55	6.0	2	201	1928	6	2	7	6	6.0	2
36	1897	2	20	5	50	7.4	0	119	1915	10	17	0	21	6.1	2	202	1928	8	1	4	28	6.1	2
37	1897	2	20	8	47	7.0	2	120	1915	11	1	16	24	7.5	0	203	1929	3	15	10	57	6.0	2
38	1897	3	27	19	49	6.3	2	121	1915	11	1	16	50	6.7	2	204	1929	4	1	5	17	6.3	0
39	1897	5	23	21	22	6.9	2	122	1915	11	1	18	1	7.0	2	205	1929	4	16	9	53	6.3	0
40	1897	7	22	18	31	6.8	0	123	1915	11	2	0	43	6.2	2	206	1929	5	31	9	10	6.1	0
41	1897	7	29	22	45	6.0	2	124	1915	11	4	12	13	6.4	2	207	1929	6	27	1	49	6.1	0
42	1897	8	5	9	10	7.7	0	125	1915	11	18	13	4	7.0	2	208	1929	8	29	3	51	6.3	0
43	1897	8	6	8	48	6.3	2	126	1915	12	7	5	58	6.5	0	209	1930	5	1	9	58	6.6	0
44	1897	8	12	10	50	6.1	2	127	1916	3	18	9	58	6.6	0	210	1930	8	21	19	44	6.1	0
45	1897	8	16	16	50	7.2	2	128	1916	5	15	8	56	6.0	0	211	1931	3	9	12	48	7.6	0
46	1897	10	2	21	45	6.6	2	129	1916	7	17	3	16	6.8	0	212	1931	3	10	2	56	6.1	2
47	1897	12	4	9	18	6.2	2	130	1916	8	8	13	25	6.3	0	213	1931	6	23	15	14	6.3	1
48	1897	12	26	16	41	6.2	2	131	1916	8	21	23	33	6.2	2	214	1931	6	23	15	14	6.6	0
49	1898	4	23	8	37	7.2	0	132	1916	8	28	7	43	6.8	0	215	1931	8	18	14	40	6.0	2
50	1898	10	7	11	0	6.0	0	133	1916	11	24	13	4	6.6	0	216	1931	9	9	4	8	6.3	2
51	1898	12	16	1	47	6.0	2	134	1917	3	15	9	14	6.9	0	217	1932	6	22	9	36	6.2	0
52	1899	3	22	19	23	6.5	2	135	1917	4	21	12	53	6.3	2	218	1932	6	30	3	16	6.0	0
53	1899	8	3	18	52	6.0	0	136	1917	6	14	22	22	6.1	0	219	1932	7	10	16	45	6.1	0
54	1900	3	12	10	34	6.4	0	137	1917	7	29	23	32	7.3	0	220	1932	9	3	20	58	6.6	0
55	1900	8	5	13	21	6.6	0	138	1917	11	16	0	2	6.0	0	221	1932	9	5	12	8	6.1	2
56	1900	8	29	11	32	6.8	0	139	1917	12	6	20	39	6.3	0	222	1933	1	4	0	26	6.1	1
57	1900	9	24	12	32	6.0	2	140	1918	7	26	5	50	6.7	0	223	1933	1	7	13	6	6.8	1
58	1901	5	14	5	11	6.0	0	141	1918	9	13	18	8	6.1	0	224	1933	1	8	15	28	6.3	1
59	1901	6	15	18	34	7.0	2	142	1918	12	14	6	33	6.2	0	225	1933	3	3	2	30	8.1	0
60	1901	8	9	18	23	7.2	1	143	1919	5	3	9	52	7.4	0	226	1933	3	3	2	40	6.5	2
61	1901	8	10	3	33	7.4	0	144	1919	7	22	8	51	6.1	0	227	1933	3	3	3	25	6.6	2
62	1901	8	10	5	0	6.3	2	145	1919	8	4	3	8	6.7	0	228	1933	3	3	3	48	6.0	2
63	1901	8	11	20	31	6.0	2	146	1919	8	8	1	32	6.2	0	229	1933	3	3	5	42	6.8	2
64	1901	8	29	21	16	6.3	2	147	1919	9	12	23	54	6.1	2	230	1933	3	3	18	12	6.5	2
65	1901	9	30	19	44	6.2	2	148	1919	10	11	22	17	6.3	0	231	1933	3	3	18	37	6.2	2
66	1902	1	1	0	20	6.1	2	149	1919	12	20	9	28	6.3	0	232	1933	3	3	18	39	6.0	2
67	1902	1	31	10	42	6.6	1	150	1920	2	8	0	6	6.7	0	233	1933	3	4	0	2	6.0	2
68	1902	5	2	20	31	7.0	0	151	1920	9	17	0	8	6.5	0	234	1933	3	8	10	35	6.0	2
69	1902	7	1	17	19	6.3	2	152	1920	9	21	5	27	6.1	0	235	1933	4	2	7	41	6.0	2
70	1902	7	8	23	5	6.2	2	153	1920	11	9	2	38	6.3	0	236	1933	4	9	11	46	6.6	2
71	1902	8	7	18	22	6.1	2	154	1920	12	3	8	39	6.0	2	237	1933	4	9	11	57	6.0	2
72	1904	12	17	16	2	6.3	0	155	1920	12	20	5	11	6.8	0	238	1933	4	19	11	55	6.1	2
73	1904	12	24	11	46	6.0	0	156	1921	1	10	3	56	6.1	0	239	1933	4	23	16	13	6.3	2
74	1905	6	27	1	10	6.2	0	157	1921	3	3	12	2	6.9	0	240	1933	6	9	3	10	6.1	2
75	1905	7	7	1	21	7.8	0	158	1921	8	22	13	5	6.3	0	241	1933	6	13	6	8	6.1	0
76	1905	12	26	12	11	6.0	0	159	1922	1	23	7	5	6.5	0	242	1933	6	14	5	33	6.4	0
77	1906	2	4	15	24	6.0	2	160	1922	3	17	3	31	6.2	0	243	1933	6	19	6	37	7.1	0
78	1906	4	5	11	50	6.0	2	161	1922	4	27	18	15	6.0	2	244	1933	7	10	9	21	6.3	2
79	1906	4	9	2	37	6.1	0	162	1922	5	16	5	21	6.5	0	245	1933	7	21	8	14	6.2	2
80	1907	1	5	1	46	6.3	0	163	1922	6	3	13	56	6.1	2	246	1933	9	21	18	47	6.1	2
81	1907	4	23	9	57	6.0	0	164	1922	7	6	5	20	6.5	0	247	1934	4	7	4	9	6.3	0
82	1907	5	23	7	54	6.3	0	165	1922	12	9	7	33	6.8	0	248	1934	10	6	5	25	6.2	0
83	1907	12	2	22	53	6.7	0	166	1923	5	26	12	12	6.4	1	249	1935	2	20	5	10	6.0	0

Table 1 (continued)

NO	YEAR	MO	DY	HR	MN	MAG	C	NO	YEAR	MO	DY	HR	MN	MAG	C	NO	YEAR	MO	DY	HR	MN	MAG	C
250	1935	3	31	6	19	6.4	0	328	1943	3	14	21	43	6.3	1	406	1960	3	21	2	7	7.2	0
251	1935	7	19	9	50	6.9	0	329	1943	4	11	23	46	6.7	0	407	1960	3	21	9	34	6.0	2
252	1935	10	13	1	45	6.9	2	330	1943	4	13	4	43	6.2	2	408	1960	3	23	9	23	6.7	2
253	1935	10	13	2	0	6.5	2	331	1943	4	13	4	50	6.0	2	409	1960	3	23	10	7	6.1	2
254	1935	10	13	10	57	6.4	1	332	1943	6	13	14	11	7.1	0	410	1960	3	24	7	22	6.0	2
255	1935	10	18	9	11	7.1	0	333	1943	6	13	14	58	6.4	2	411	1960	6	16	0	36	6.2	1
256	1935	10	18	23	53	6.5	2	334	1943	6	13	17	36	6.4	2	412	1960	7	30	2	31	6.7	0
257	1935	10	19	6	52	6.3	2	335	1943	6	14	2	39	6.3	2	413	1960	8	13	16	11	6.2	0
258	1936	3	2	12	19	6.8	0	336	1943	6	15	1	22	6.1	2	414	1961	1	16	16	20	6.8	0
259	1936	3	11	5	36	6.3	2	337	1943	6	15	20	10	6.7	2	415	1961	1	16	20	19	6.4	2
260	1936	3	11	9	44	6.1	0	338	1944	2	1	14	16	6.8	0	416	1961	1	16	21	12	6.5	2
261	1936	6	3	11	15	6.0	0	339	1944	3	10	15	40	6.1	0	417	1961	1	16	23	3	6.1	2
262	1936	11	3	5	45	7.5	0	340	1944	6	6	20	48	6.0	0	418	1961	1	17	0	41	6.6	2
263	1936	11	14	9	58	6.0	0	341	1944	10	3	5	29	6.4	1	419	1961	2	23	13	16	6.4	0
264	1937	1	7	15	12	6.6	2	342	1945	2	10	13	57	7.1	0	420	1961	3	25	7	57	6.1	0
265	1937	1	20	9	3	6.0	0	343	1945	2	18	7	35	6.2	2	421	1962	4	12	9	52	6.8	0
266	1937	3	22	4	27	6.1	0	344	1945	2	18	19	8	6.6	2	422	1962	4	26	0	47	6.4	2
267	1937	7	27	4	56	7.1	2	345	1945	3	12	6	37	6.6	0	423	1963	5	8	19	22	6.1	0
268	1937	10	17	13	47	6.6	0	346	1945	4	10	10	22	6.4	0	424	1963	8	15	15	11	6.6	0
269	1937	12	10	22	28	6.1	0	347	1945	6	26	8	40	6.1	0	425	1964	1	10	13	50	6.1	0
270	1938	5	23	16	18	7.0	0	348	1946	5	10	7	27	6.1	0	426	1964	2	5	20	30	6.0	0
271	1938	9	22	3	52	6.5	0	349	1946	7	20	6	16	6.0	0	427	1964	4	16	10	4	6.0	0
272	1938	10	12	9	34	6.9	0	350	1946	8	3	22	6	6.1	0	428	1964	5	30	23	30	6.2	0
273	1938	10	29	22	8	6.4	0	351	1946	8	14	18	40	6.0	0	429	1965	3	17	1	46	6.4	0
274	1938	11	5	17	43	7.5	0	352	1947	1	3	12	57	6.0	0	430	1965	3	29	19	47	6.4	2
275	1938	11	5	19	50	7.3	2	353	1947	11	14	19	49	6.5	0	431	1965	6	13	16	6	6.0	0
276	1938	11	6	6	22	6.1	2	354	1948	3	15	20	24	6.0	0	432	1965	9	18	1	21	6.7	0
277	1938	11	6	17	53	7.4	2	355	1948	5	12	9	57	6.6	0	433	1965	9	23	7	8	6.2	2
278	1938	11	7	2	19	6.0	2	356	1948	5	12	10	21	6.1	2	434	1967	1	17	20	59	6.3	0
279	1938	11	7	6	4	6.2	2	357	1948	5	14	22	19	6.2	2	435	1967	11	19	21	6	6.0	0
280	1938	11	7	6	38	6.9	2	358	1948	9	23	9	52	6.0	0	436	1968	5	16	9	48	7.9	0
281	1938	11	7	9	48	6.2	2	359	1948	10	29	5	45	6.3	0	437	1968	5	16	10	4	6.2	2
282	1938	11	7	10	38	6.4	2	360	1949	5	22	6	40	6.3	0	438	1968	5	16	19	39	7.5	2
283	1938	11	7	10	45	6.2	2	361	1951	7	26	18	59	6.0	0	439	1968	5	17	1	13	6.1	2
284	1938	11	7	10	54	6.4	2	362	1951	7	29	8	4	6.1	0	440	1968	5	17	8	4	6.7	2
285	1938	11	7	11	27	6.0	2	363	1951	10	18	17	26	6.6	0	441	1968	5	23	4	29	6.3	2
286	1938	11	7	13	15	6.3	2	364	1952	3	4	10	22	8.2	0	442	1968	5	24	23	6	6.2	2
287	1938	11	8	4	33	6.5	2	365	1952	3	4	10	40	6.5	2	443	1968	6	12	22	41	7.2	2
288	1938	11	9	11	22	6.5	2	366	1952	3	5	4	56	6.4	2	444	1968	6	13	6	57	6.1	2
289	1938	11	9	18	15	6.3	2	367	1952	3	5	18	17	6.0	2	445	1968	6	17	20	52	6.4	2
290	1938	11	11	7	22	6.4	2	368	1952	3	10	2	3	6.8	2	446	1968	6	18	3	57	6.0	0
291	1938	11	14	7	31	7.0	2	369	1952	4	28	19	54	6.2	2	447	1968	6	22	10	12	6.1	2
292	1938	11	14	11	36	6.0	2	370	1952	5	14	9	36	6.1	2	448	1968	7	5	20	28	6.4	0
293	1938	11	16	20	8	6.6	2	371	1952	5	17	18	48	6.1	2	449	1968	7	12	9	44	6.4	2
294	1938	11	19	14	54	6.0	2	372	1952	5	20	3	32	6.5	2	450	1968	9	21	22	6	6.9	2
295	1938	11	22	10	14	6.9	2	373	1952	10	27	0	46	6.1	1	451	1968	10	8	5	49	6.2	2
296	1938	11	22	10	40	6.1	2	374	1952	10	27	0	53	6.4	1	452	1968	11	11	23	41	6.0	2
297	1938	11	22	12	24	6.0	2	375	1952	10	27	3	1	6.4	1	453	1968	11	14	3	41	6.0	2
298	1938	11	25	17	20	6.3	2	376	1952	10	27	4	19	6.5	0	454	1968	11	25	6	20	6.0	2
299	1938	11	29	22	39	6.4	2	377	1952	10	27	12	17	6.4	2	455	1969	3	17	0	54	6.1	0
300	1938	11	30	11	29	6.9	2	378	1952	10	28	15	30	6.3	2	456	1970	5	28	4	5	6.2	0
301	1938	12	1	0	16	6.1	2	379	1952	11	1	1	37	6.4	2	457	1970	5	28	7	35	6.0	2
302	1938	12	3	21	12	6.5	2	380	1953	1	19	13	57	6.0	2	458	1970	9	14	18	44	6.2	0
303	1938	12	7	22	4	6.4	0	381	1953	2	6	22	13	6.7	2	459	1970	12	7	5	20	6.1	0
304	1938	12	14	2	26	6.3	2	382	1953	4	4	14	52	6.2	0	460	1971	4	5	3	39	6.0	0
305	1938	12	19	6	45	6.0	2	383	1953	5	26	10	43	6.0	2	461	1971	8	2	16	24	7.0	0
306	1938	12	23	10	51	6.0	2	384	1953	12	7	23	11	6.4	0	462	1971	9	15	23	55	6.3	0
307	1939	1	24	13	1	6.1	2	385	1953	12	22	2	36	6.1	0	463	1971	9	24	10	9	6.1	2
308	1939	2	17	3	51	6.5	2	386	1954	4	5	8	14	6.1	2	464	1972	3	20	0	57	6.4	0
309	1939	8	22	9	6	6.3	2	387	1954	7	18	18	7	6.4	0	465	1973	9	5	22	3	6.1	0
310	1939	10	11	3	32	7.0	0	388	1954	9	12	16	43	6.2	0	466	1973	9	10	3	25	6.0	2
311	1939	10	11	3	51	6.4	2	389	1954	11	19	5	44	6.1	0	467	1973	11	19	22	1	6.4	0
312	1940	2	9	22	53	6.2	0	390	1955	5	1	18	55	6.1	0	468	1974	1	25	4	12	6.0	0
313	1940	11	14	19	33	6.2	0	391	1955	5	6	9	4	6.0	2	469	1974	3	3	13	50	6.1	0
314	1940	11	20	0	1	6.6	0	392	1956	2	10	9	2	6.0	0	470	1974	7	8	14	45	6.3	0
315	1941	2	9	13	16	6.1	0	393	1956	10	12	21	22	6.1	0	471	1974	10	10	15	48	6.2	1
316	1941	3	12	23	16	6.3	0	394	1956	11	21	16	33	6.2	0	472	1974	10	10	15	56	6.4	0
317	1941	3	13	6	37	6.2	2	395	1957	6	12	17	28	6.1	0	473	1974	10	12	15	14	6.2	2
318	1941	3	14	23	31	6.2	2	396	1958	2	16	15	4	6.1	0	474	1974	11	16	8	32	6.1	0
319	1941	3	19	11	45	6.3	2	397	1958	4	8	3	5	6.7	0	475	1975	5	4	18	31	6.0	0
320	1941	5	9	18	32	6.1	0	398	1958	4	8	3	30	6.2	2	476	1975	10	30	10	41	6.0	0
321	1941	5	9	18	38	6.0	2	399	1958	4	8	3	38	6.2	2	477	1976	11	8	17	19	6.2	0
322	1941	11	26	0	20	6.3	0	400	1958	4	11	9	58	6.3	2	478	1978	2	20	13	36	6.7	0
323	1942	2	21	16	7	6.5	0	401	1958	9	3	17	10	6.0	0	479	1978	4	7	8	29	6.1	0
324	1942	9	9	1	7	6.3	1	402	1959	1	22	14	10	6.8	0	480	1978	6	12	17	14	7.4	0
325	1942	9	21	14	53	6.2	0	403	1959	1	24	14	8	6.2	2	481	1978	6	14	20	34	6.3	2
326	1942	11	16	2	12	6.5	0	404	1959	10	26	16	35	6.8	0	482	1979	2	20	15	32	6.5	0
327	1943	3	14	20	59	6.1	1	405	1960	2	5	1	50	6.1	0	483	1980	1	13	0	57	6.1	0

NOTE: NO, YEAR, MO, DY, HR, MN, and MAG, indicate earthquake number, year, month, day, hour, minute, and magnitude, respectively. In the column *C*, 0, 1, and 2 represent main shock, foreshock, and aftershock, respectively.

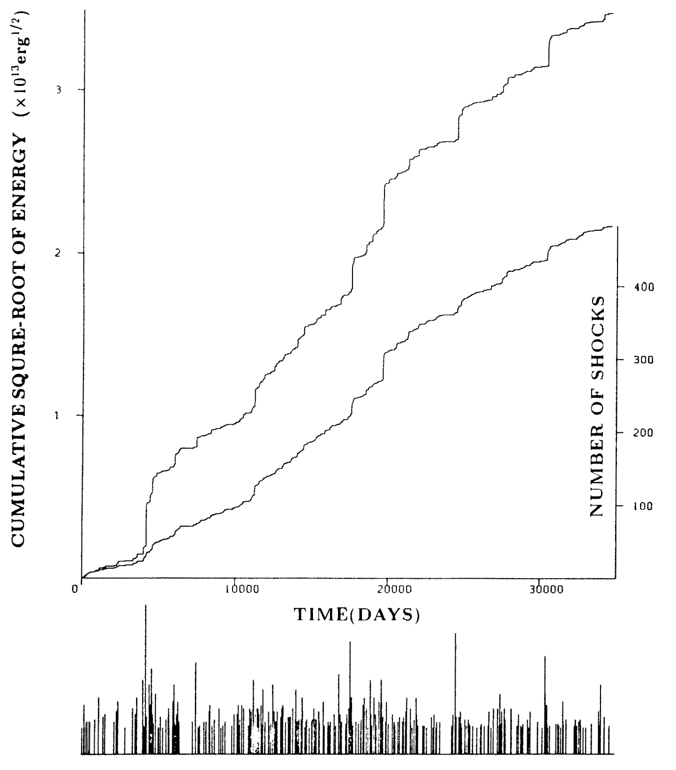

Figure 2. Cumulative Number of Shocks (lower solid line), Cumulative Square Root of the Released Energy (upper solid line), and Plot of the Magnitudes Versus the Occurrence Times of Shocks.

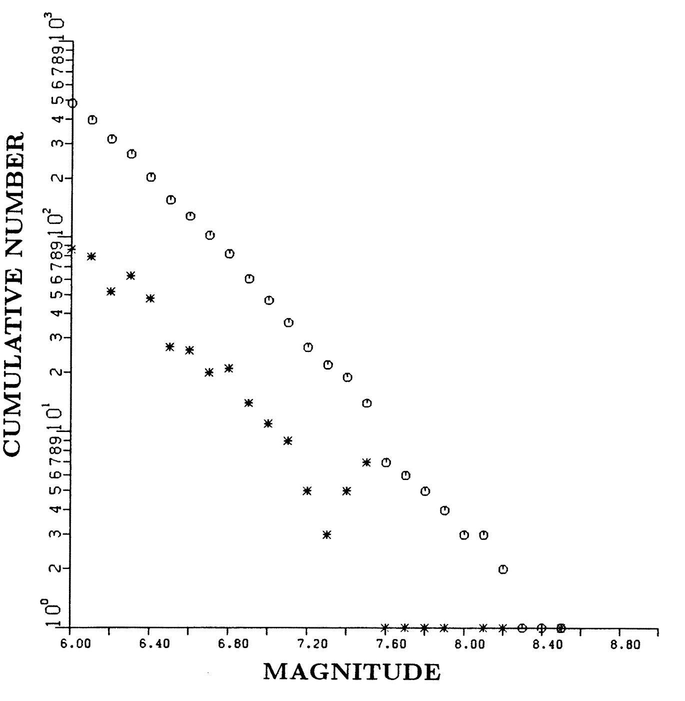

Figure 3. Density Distribution (*) and Cumulative Distribution (○) of Magnitudes for the Data Set.

the parameter vectors η and θ have no common components. Then maximization of the full likelihood is reduced to the maximization of each component. Although I will have to examine the first part of (22) afterwards in Section 4 to argue for the utility of seismic quiescence, here I am interested in the second part (the so-called conditional likelihood), the logarithm of which is written as

$$\log L(\theta) = \sum_{i=1}^{N} \log \lambda(t_i; \theta) - \int_0^T \lambda(t; \theta)\, dt, \quad (23)$$

where $\lambda(t; \theta)$ is the parameterized conditional intensity rate discussed in the previous section and $\{t_i\}$ is the set of occurrence times of earthquakes in an observed time interval $[0, T]$. Note here that the magnitude data $\{m_i\}$ of the shocks are also included in the function $\lambda(t; \theta)$, as in (17) and (21). We can expect the conditional likelihood to enjoy the standard large-sample theory under some regularity conditions such as stationarity and ergodicity of the joint series $\{(t_i, m_i)\}$.

The maximum of the log-likelihood and the estimates of the parameters in (23) can be obtained numerically by a standard nonlinear optimization technique such as that in Fletcher and Powell (1963). It is then possible to judge which of the models described in the preceding section provides the best fit to the earthquake process data. For this purpose, the Akaike information criterion (AIC) (Akaike 1974) is used as a measure for selecting the best among competing models for a fixed data set. This is defined by AIC = (-2)max(log-likelihood) + 2 (number of used parameters). The model with the smaller AIC shows the better fit to the data. It is sometimes useful to note that the log-likelihood ratio statistic takes the form

$$(-2)\log(L_0/L_1) = \mathrm{AIC}(H_0) - \mathrm{AIC}(H_1) + 2k, \quad (24)$$

where k denotes the difference between the numbers of parameters in H_0 and H_1. When the model H_1 contains the model H_0 as a restricted case, then under the null

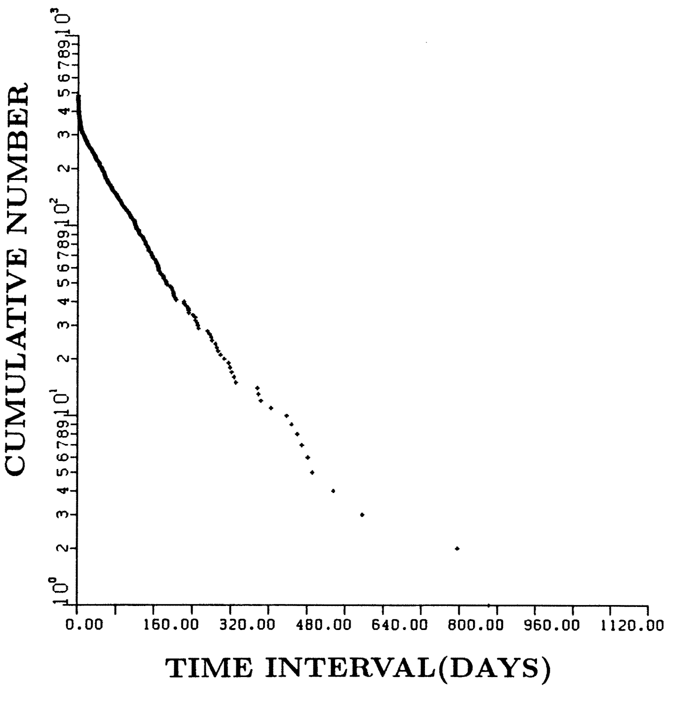

Figure 4. Empirical Log-Survivor Function of the Intervals Between Successive Earthquakes.

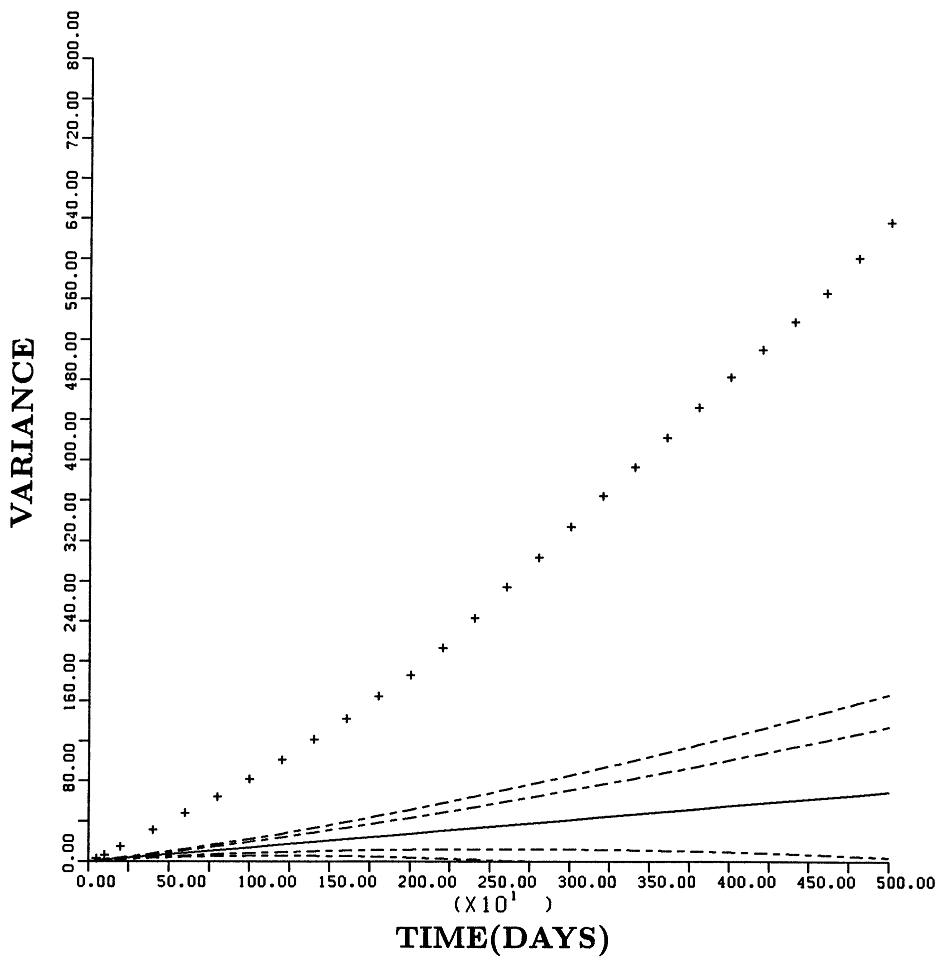

Figure 5. Estimated Variance–Time Curve (+ + +). The solid line is a theoretical one for the stationary Poisson case, and the dotted lines correspond to the 95.5% and 99.7% error bounds for individual estimates under the stationary Poisson process.

hypothesis H_0, the statistic $(-2)\log(L_0/L_1)$ is expected to have a chi-squared distribution with k df. The comparison of the minimum AIC procedure with the conventional likelihood ratio test is discussed in Akaike (1977, 1983). It is emphasized that we can further extend the comparison among nonnested models by using AIC, which was originally derived for such a purpose.

Table 2 gives the maximum log-likelihoods and AIC values for the models considered. For the models where (15) is used, I considered values of κ up to 16, but here are listed only the results for $\kappa = 1$ and for the order $\hat{\kappa}$ with minimum AIC value. Table 2 clearly shows the following: (a) the effect of magnitude on cluster size is significant, (b) for $g(t)$, the modified Omori function (14) is significantly better than the Laguerre-type polynomials (15), and (c) the epidemic-type models (17) are signifi-

Table 2. Comparison of Models Fitted to the Data in Table 1

	Response function (15), including (13)				*Response function (11)*			
	Magnitude effect $\beta = 0$		*Magnitude effect $\beta > 0$*		*Magnitude effect $\beta = 0$*		*Magnitude effect $\beta > 0$*	
Model	$\kappa = 1^*$	$1 \leqq \kappa \leqq 16^*$	$\kappa = 1^*$	$1 \leqq \kappa \leqq 16^*$	$p = 1.0^*$	$p \neq 1.0^*$	$p = 1.0^*$	$p \neq 1.0^*$
				Trigger models				
Optimum orders	—	$x = 9$	—	$x = 7$	—	—	—	—
$-\log L(\theta)$	2,404.2	2,275.4	2,308.6	2,271.0	2,348.5	2,347.3	2,251.3	2,249.4
Number of parameters	3	11	4	10	3	4	4	5
AIC	4,814.4	4,572.8	4,625.3	4,562.0	4,703.0	4,702.6	4,510.5	4,508.8
				Epidemic-type models				
Optimum orders	—	$x = 11$	—	$x = 11$	—	—	—	—
$-\log L(\theta)$	2,288.4	2,236.1	2,248.0	2,205.1	2,226.4	2,226.4	2,185.2	2,185.0
Number of parameters	3	13	4	14	3	4	4	5
AIC	4,582.8	4,498.3	4,504.0	4,438.2	4,458.8	4,460.7	4,378.4	4,380.0

* Restrictions

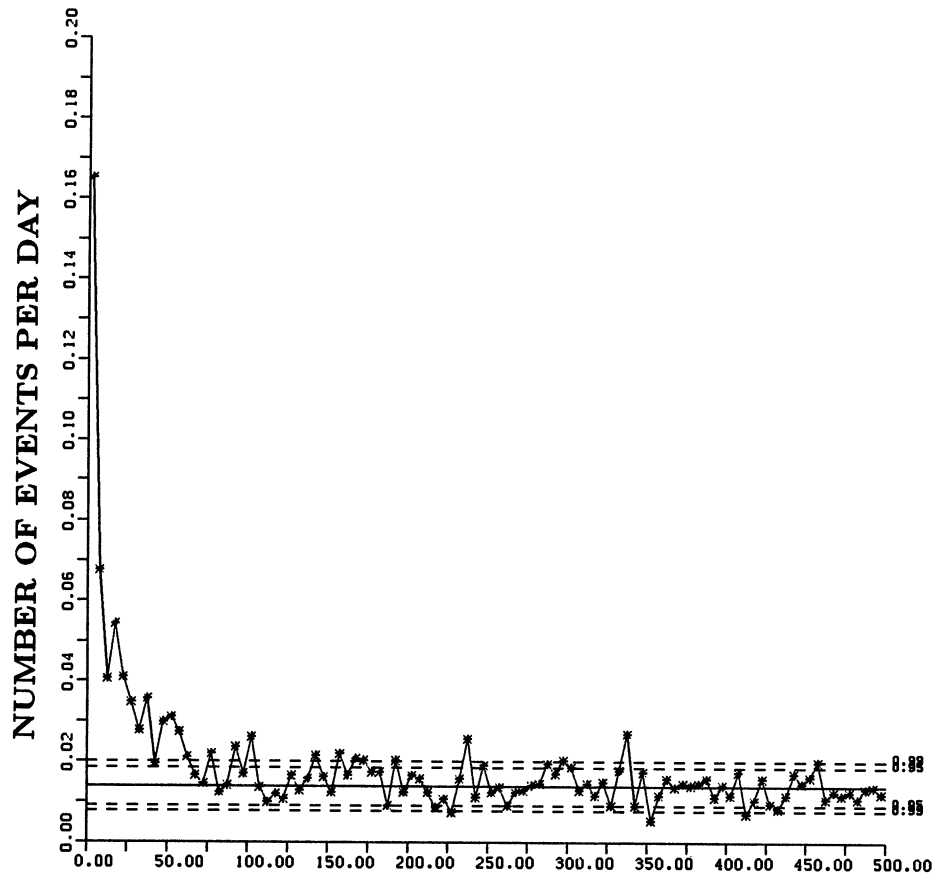

Figure 6. Histogram for Estimating $m_f(s)$, the Intensity Rate Under the Palm Probability. The dotted lines indicate 95.5% and 99.7% error bounds for individual estimates assuming the stationary Poisson process.

cantly better than the restricted trigger models (21). Especially, the model (17) with (14) is significantly better than any of the other models. Estimates of the parameters of the best model are listed in Table 3; here the reference magnitude was taken to be $M_r = 6.0$. The plot of the estimated conditional intensity rate is shown in Figure 8.

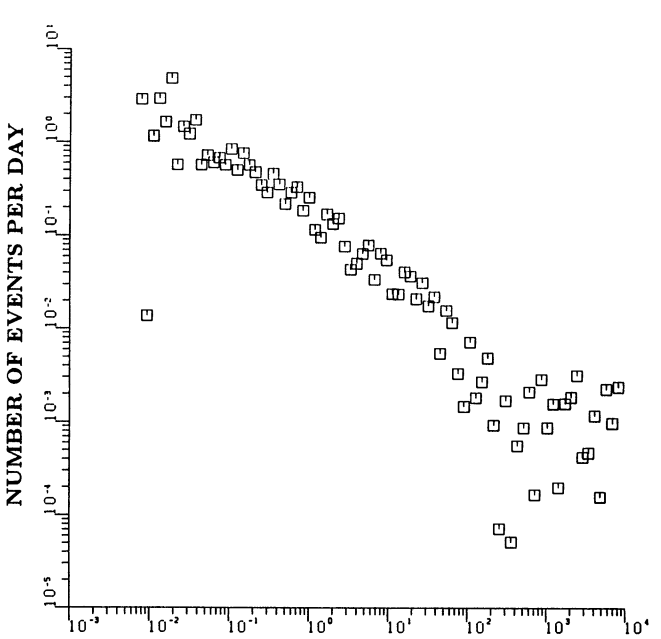

Figure 7. Plot of $| m_f(s) - \lambda_0 |$ Versus s on a Log–Log Scale.

3.3 Residual Analysis of Point Process Data

AIC is useful for the comparison of competing models. Having obtained the best model among those considered, however, there remains the possibility of the existence of a still better model. So, we usually check whether the major features of the given data can be reproduced by the estimated models [see Ogata (1981) for efficient and systematic simulation of point processes using the conditional intensity function]. If any one of the important features is not reproduced, we must consider further models whose AIC values can be compared with those of the previous best model. Thus it may be useful to develop graphical techniques for amplifying the features of the data that deviate from the model, if any.

Suppose that the point process data $\{t_i\}$ are generated

Table 3. Estimated Parameters of the Best Model (M_r = 6.0)

μ	K	c	p	β
.00536 (shocks/day)	.017284 (shocks/day)	.01959 (days)	1.0	1.61385

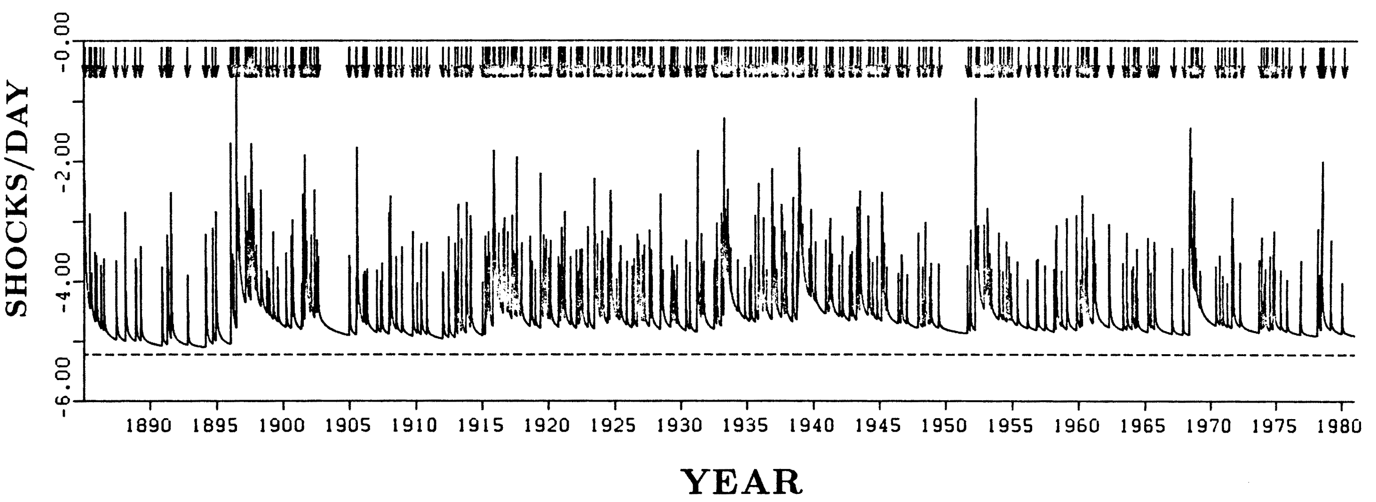

Figure 8. Estimated Conditional Intensity Rate of (17) With Parameter Values in Table 3. The rate is plotted in the logarithmic scale. The arrows downwards indicate the occurrence times of shocks in the Off Tohoku area.

by the conditional intensity $\lambda(t)$. Consider the integral of the conditional intensity

$$\Lambda(t) = \int_0^t \lambda(t)\, dt, \tag{25}$$

which is a monotonically increasing function because $\lambda(t)$ is nonnegative. If we consider the random time change $\tau = \Lambda(t)$ from t to τ, then $\{t_i\}$ is transformed one-to-one into $\{\tau_i\}$. It is well known that $\{\tau_i\}$ has the distribution of a stationary Poisson process of intensity 1 (see, e.g., Papangelou 1972). Therefore, if the estimated conditional intensity $\lambda(t; \hat{\theta})$ is a good approximation to the true $\lambda(t)$, then the transformed data $\{\tau_i\}$ are expected to behave like a stationary Poisson process. In other words, a deviation from a property of $\{\tau_i\}$ from that expected of a stationary Poisson process implies the existence of a corresponding feature of the data $\{t_i\}$ that is not captured by the model $\lambda(t; \hat{\theta})$. The intensity $\lambda(t; \hat{\theta})$ represents a model for prediction, whereas the transformed data $\{\tau_i\}$ may be regarded as "noise," or "residuals" in a wide sense, of the point process data $\{t_i\}$. This sequence $\{\tau_i\}$ will be called the *residual process* and is a further example of generalized residuals similar to those discussed in Cox and Snell (1968).

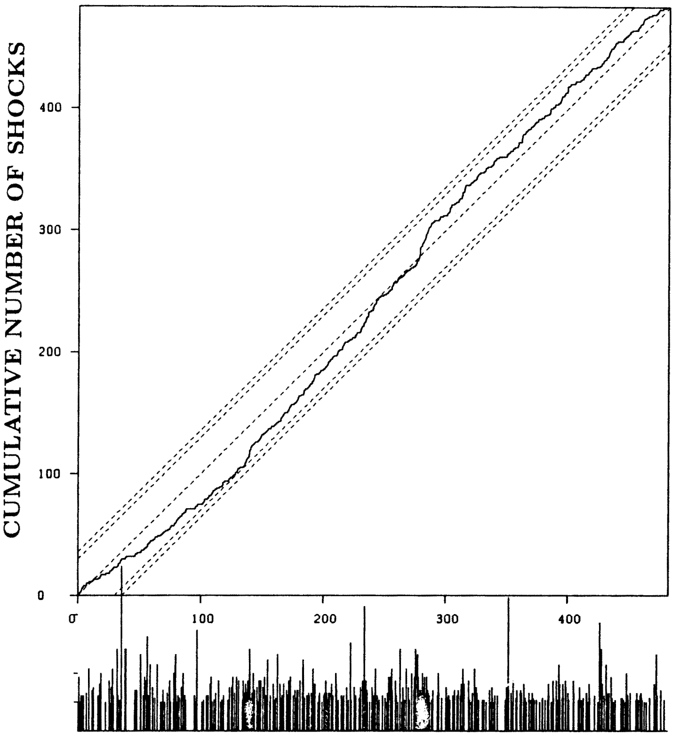

Figure 9. Cumulative Numbers of the Residual Process and Magnitudes Versus the Transformed Time τ_i. The dotted lines indicate the average and two-sided 95% and 99% error bounds of the Kolmogorov–Smirnov statistic.

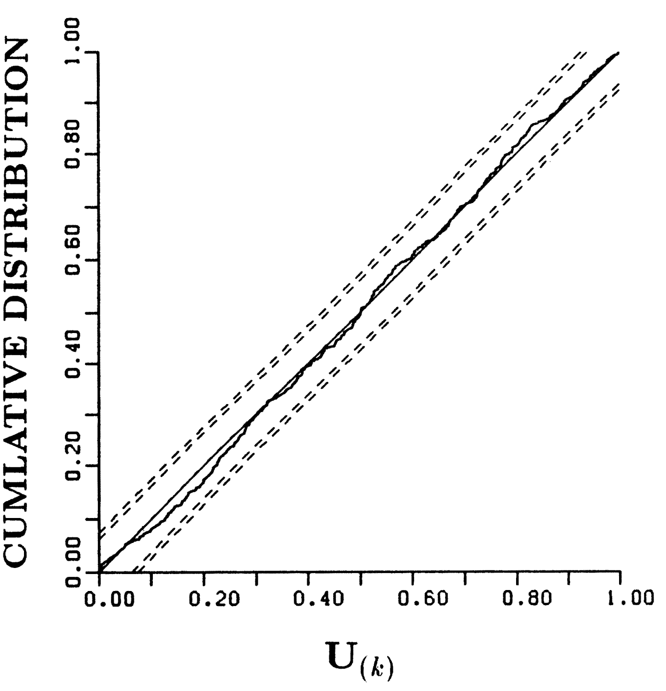

Figure 10. Empirical Distribution of U_k. The dotted lines indicate the 95% and 99% error bounds of the Kolmogorov–Smirnov statistic assuming the uniform distribution.

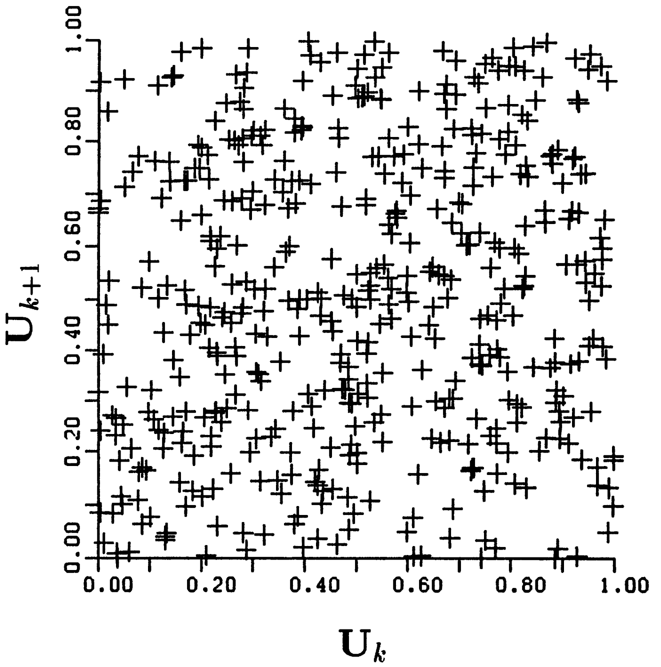

Figure 11. U_k Versus U_{k+1} Plot.

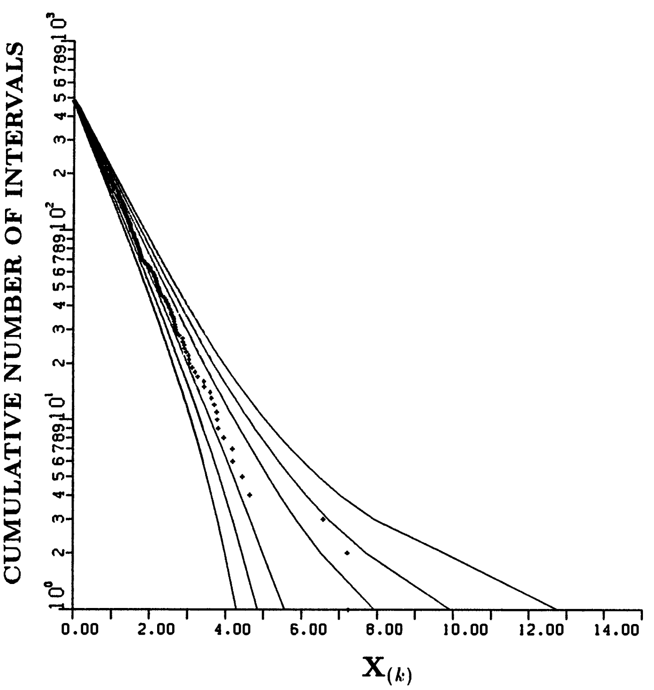

Figure 12. Empirical Log-Survivor Function of the Intervals Y_k of the Residual Process. The real lines indicate the percentiles that correspond to one-, two-, and three-fold standard errors of the normal distribution.

Any conventional graphic tests for complete randomness can be useful for residual analysis. Figure 9 shows the plot of the cumulative number of points $\{\tau_i\}$ versus transformed time $\tau = \Lambda(t)$. The dotted lines display the two-sided 95% and 99% error bounds of the Kolmogorov–Smirnov statistics, assuming the uniform empirical distribution. Ogata and Shimazaki (1984) applied this sort of residual analysis of trend to an aftershock sequence to reveal the time at which the background seismicity surpasses the aftershock activity. Another possible test was described by Berman (1983), who discussed the residual analysis of the transformed interarrival times properties; that is to say, he considered whether

$$Y_k = \tau_k - \tau_{k-1} = \Lambda(t_k) - \Lambda(t_{k-1}), \quad k = 1, \ldots, N, \tag{26}$$

are iid exponential random variables with unit mean, and hence whether the statistics $U_k = 1 - \exp(-Y_k)$ are iid uniform random variables on [0, 1). The empirical distri-

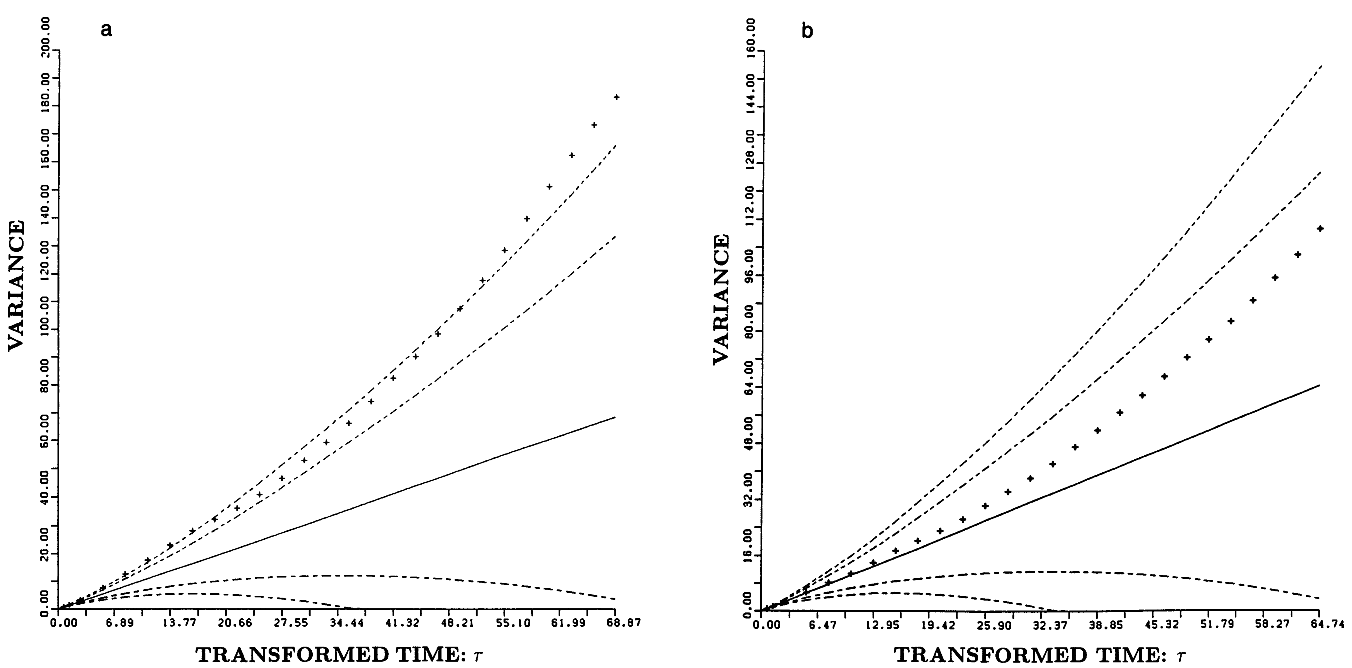

Figure 13. (a) Estimated Variance–Time Curve for the Residual Process. The solid line is the theoretically expected variance–time curve for the stationary Poisson process. The dotted lines indicate for individual estimates 95.5% and 99.7% error bounds for individual estimates. (b) Estimated variance–time curve for the residual process after removing the Shioya-Oki swarms in 1938.

bution of U_k for our data set is plotted in Figure 10 with the same error bounds as in Figure 9. Berman also suggested a plot to test for independence of the intervals. Assuming that any serial correlation present is likely to show up in neighboring intervals, he plotted U_k against U_{k+1}. The plot for our data set is given in Figure 11. The plots for $\{\tau_k\}$ corresponding to Figures 4, 5, and 6 are given in Figures 12, 13a, and 14, respectively. Here the lines for two-sided error bounds for individual estimates correspond to 99.7%, 95.5%, and 68.3% (only for Fig. 12). The variance–time curve $V(\tau)$ in Figure 13a still indicates some clustering features, despite the fact that Figures 9–12 and 14 seem to support the hypothesis that the noise process $\{\tau_i\}$ is a stationary Poisson process.

Thus I suspect that there may be some local unusual characteristics in $\{\tau_i\}$ invalidating the stationary Poisson assumption. To find such features, consider the number of points $\Delta N = N(\tau - h, \tau)$ in the interval $(\tau - h, \tau)$. If the residual process $\{\tau_i\}$ is stationary Poisson, then $\Delta N = N(\tau - h, \tau)$ is a Poisson random variable with mean h for each τ. Here, if we use the transformation suggested in Shimizu and Yuasa (1984),

$$\xi = \xi(\Delta N, h) = \frac{33\Delta N + 29 - h - (32\Delta N + 31)[h/(\Delta N + 1)]^{1/4}}{9(\Delta N + 1)^{1/2}}, \tag{27}$$

then ξ is well approximated by a normal random variable with mean 0 and variance 1. This transformation was obtained by an idea similar to the derivation of the well-known Wilson–Hilferty transform.

Setting $h = 8$, for example, the time series of ξ as a function of τ is plotted in Figure 15(3a). It behaves like a Gaussian process, except for the part of the trajectory around 1938. This is a stretch of *outliers* associated with a swarm of large earthquakes, including magnitudes 7.4 (May 23), 7.7 (Nov. 5), 7.8 (Nov. 5), 7.7 (Nov. 6), and 7.1 (Nov. 6) at Shioya-Oki (Off Fukushima Prefecture) in the southern part of the region shown in Figure 1. After finding these outliers, I learned that Abe (1977) had already reported and described the unusual features of these shocks in detail. That is to say, it is very rare that a swarm includes a series of shocks with such large magnitudes or such a large energy release. In addition, he pointed out the fact that there had been no such major earthquake in the focal region for at least 800 years, whereas the frequency of large earthquakes, $M \geq 7.5$, in other parts of the Off Tohoku area was about one per 100 years. Finally, I considered another variance–time curve $V(\tau)$ in Figure 13b, obtained by removing the swarm from the residuals $\{\tau_i\}$: the corresponding numbers i of the removed data are shown from 274 through 302 in Table 1. This variance–time curve suggests that the remaining part of the residuals satisfies the hypothesis of stationary Poisson process. Therefore, I conclude that the seismic activity (in the sense

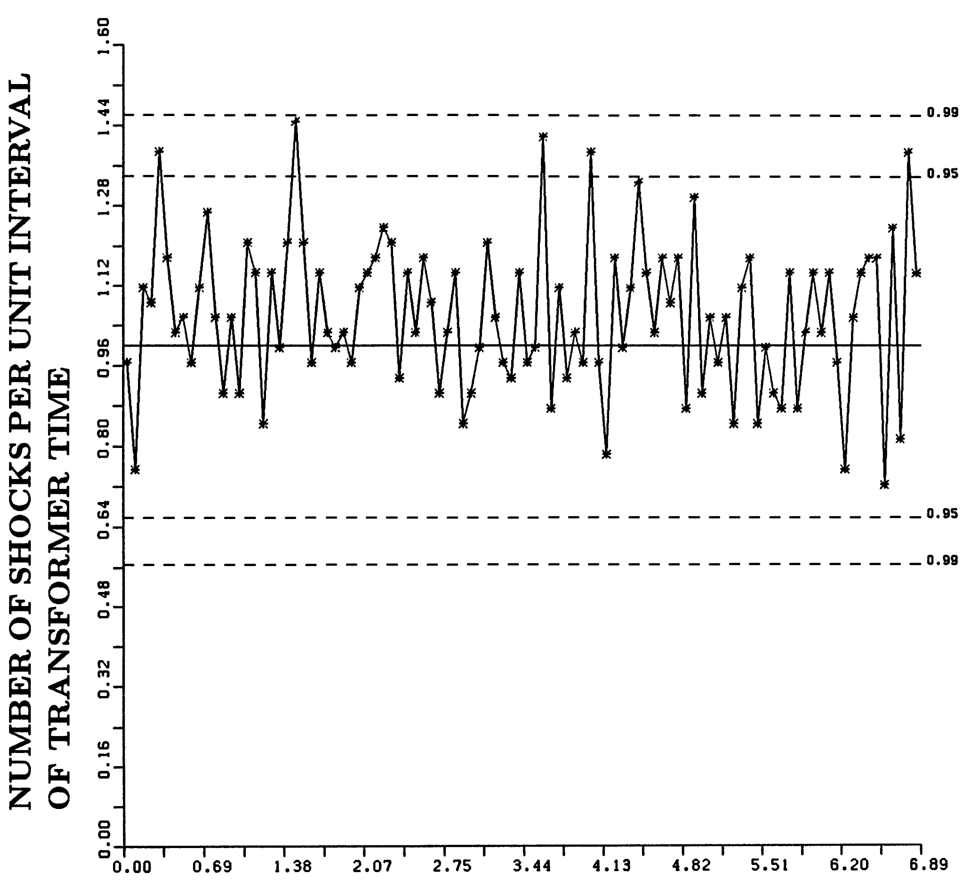

Figure 14. Histogram for Estimating $M_t(\tau)$ of the Residual Process. The dotted lines indicate 95.5% and 99.7% error bounds for individual estimates assuming the stationary Poisson process.

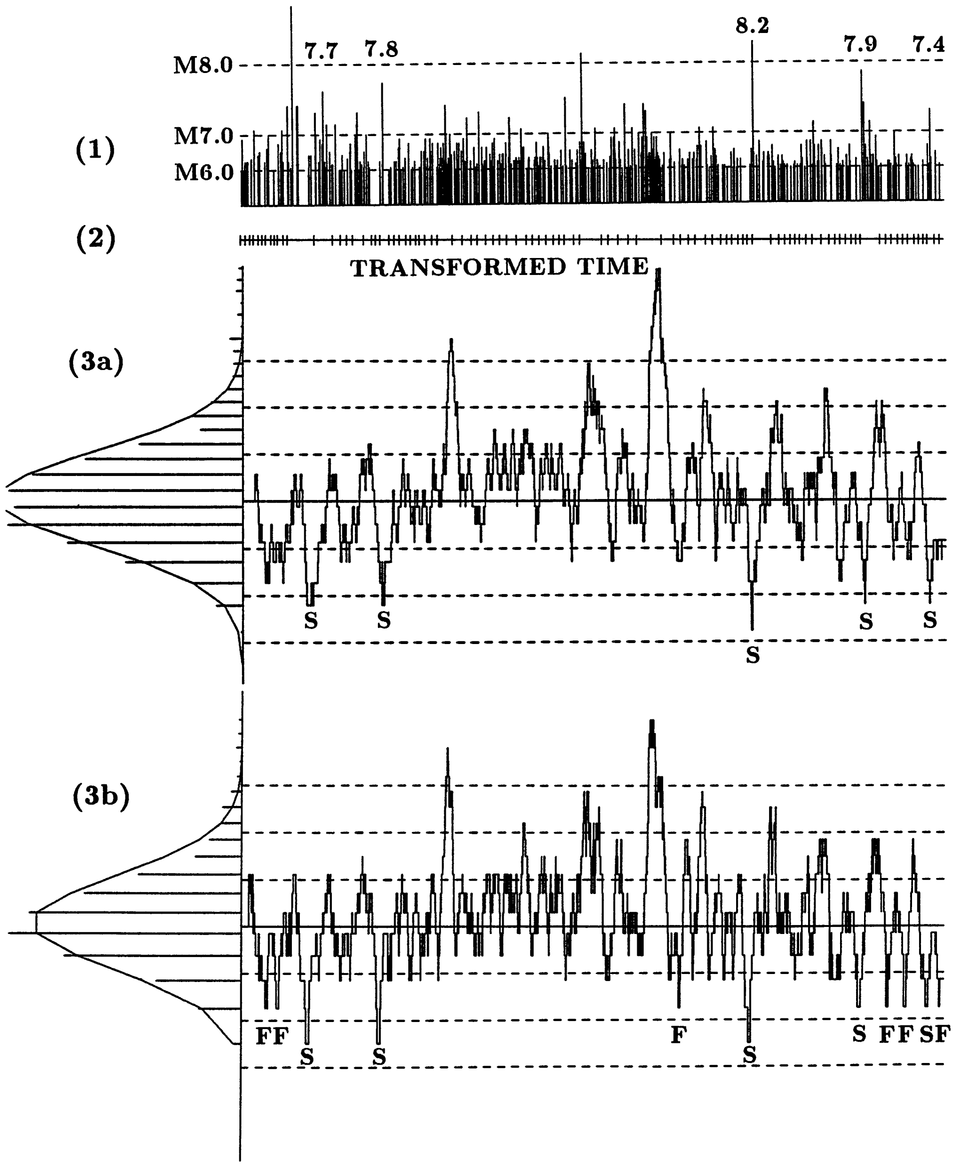

Figure 15. (1) The magnitudes versus occurrence times of the residual process. (2) The transformed times of January 1 of every year from 1885 to 1980. (3a) Time series of the number of points ΔN_τ in the interval (τ-h, τ) of the residual process. Putting h = 8, ΔN_τ is transformed into ξ_τ by the relation (27). The dotted lines are one-, two-, and three-fold standard errors of the approximated normal distribution. The sideways histograms indicate the total lengths of the time span of τ with the same values of ΔN_τ, and the broken line corresponds to the theoretical distribution (approximate normal) assuming that the residual process is the stationary Poisson process. (3b) Time series of ξ_τ for h = 5.

of frequency of shocks) of the remaining part seems to be well described by the selected model, with estimated parameters as given in Table 3.

4. SEISMIC QUIESCENCE

In contrast to the assertion of Lomnitz and Nava (1983) that seismic quiescence is merely a consequence of the decaying activity of aftershocks, there are some reports discussing the presence of quiescence with quantitative observation of earthquake frequencies. Okada (1978) found some seismicity lowering prior to major earthquakes; that is, the cumulative curve of felt earthquake frequency per month seems to bend lower than the frequency predicted by the combination of the constant rate background seismicity and the series of aftershock-type activity. I am interested here in solving the point at issue by using residual analysis. First of all, from Figure 12 we cannot find a statistically exceptionally long interval $Y_i = \tau_i - \tau_{i-1}$ under the assumption of a stationary Poisson process with intensity rate 1. In addition, we cannot find significant lowering of the number $\Delta N = N(\tau - 8, \tau)$ in Figure 15(3a); there are five times when the trajectory of $\xi = \xi(\Delta N, 8)$ defined in (27) crosses the second-lowest error bound (i.e., $\xi = -2$) from above, but this does not seem unusual for the stationary Poisson process with sample size 483. Do these observations support the aforementioned assertion of Lomnitz and Nava (1983)?

If we look at Figure 15(3a) in comparison with the plot of the transformed data $\{(\tau_i, M_i)\}$ we can see that the five large shocks with $M \geq 7.4$ (i.e., 7.7, 7.8, 8.2, 7.9, and 7.4) took place within a year after ξ_τ had crossed the level $\xi = -2$ from below. It seems unlikely that such coincidence is wholly due to the choice of h and ξ. For example, a similar plot was made for $h = 5$ in Figure 15(3b), where we see that the three large shocks with $M \geq 7.7$ (i.e., 7.7, 7.8, and 8.2) occurred within a year after ξ_τ had crossed the level $\xi = -2$ from below and that five shocks with $M \geq 7.4$ (the same as before) occurred within a year after the level crossing of $\xi = -1.5$, which took place 11 times. Are these combinatorial realizations rare phenomena or not? This relates to the question of usefulness of seismic quiescence for predicting the major earthquakes. I tentatively *define* the *quiescence* as the time span of τ or t when the trajectory of ξ_τ is lower than a certain level, such as -2.0 or -1.5, like the above.

Before assessing the probability of the aforementioned events, we have to make assumptions about the stochastic law governing the magnitudes of shocks. From Figure 3 we may assume that the marginal distribution of magnitudes throughout the 98 years in the area is negative exponential. But what should we assume about the joint distribution of magnitudes and occurrence times? Recall here the independence assumptions used to obtain the full likelihood in (22). As mentioned, in studying aftershock sequences, Lomnitz (1966) put forward a hypothesis of "magnitude stability," that is to say, stationarity of the magnitudes and their independence from the occurrence times; it was shown in his paper that the local averages of *aftershock* magnitudes fluctuated slightly about a mean value. This hypothesis was implicitly supported by Utsu (1962) in his detailed trend analysis of three Alaskan aftershock sequences.

Now can this hypothesis extend to the general sequences of the mixture of *main shock* and *aftershocks* in Figure 2? This question relates to the explanation of so-called Båth's law. Båth's law claims the different mean magnitudes between the group of main shocks and group of aftershocks, and Lomnitz and Nava (1983) interpreted this by the different b-values [see Eq. (3)] between the two groups. On the other hand, Vere-Jones (1966, 1975) carefully checked such differences using Utsu's table (Utsu 1961) and concluded that Båth's law can be adequately explained by the simple assumption that the magnitudes in an earthquake sequence form a random sample independently selected from a distribution having the exponential form. In addition, the experiment by Utsu (1971, pp. 425–426) suggests that the Gutenberg–Richter law of magnitude frequency with the same b-value throughout the whole seismic sequence can imply significantly different b-values between the group of main shocks and the group of aftershocks.

It was concluded in Section 3 that the residual time process is a stationary Poisson process. Now a joint model for both magnitude and time will be discussed. This discussion is based on the assumption that one b-value applies to the full shock sequence. If we assume that independence of the magnitudes M_i from the past occurrence times of shocks and their magnitudes, $\{(t_j, M_j); t_j < t_i\}$ as well as stationarity, then the residual process $\{(\tau_i, M_i)\}$ is approximately a stationary *compound Poisson process*. Thus we can now assess the probability of the realized combination of the seismic quiescences and the succeeding occurrence of a major earthquake within a year. Since the number of shocks with $M \geqslant 7.7$ is six in 96 years, 6/96 (per year) may be given for the rate of such shocks. Similarly, 19/96 and 47/96 per year are the rates of shocks with $M \geqslant 7.4$ and $M \geqslant 7.0$, respectively. I observed the numbers of the previously defined quiescences in Figure 15 and also counted the major shocks of the specified magnitudes that occurred within a year after the end of the quiescences. Using all of these observations the probabilities of the realized combinations were assessed. These are listed in the row for (i) of Table 4. It is seen that all of the assessed probabilities are very small.

On the other hand, I generated the five sequences (ii)–(vi) of the point process [see Fig. 16 and Ogata (1981) for the simulation method], using the estimated model (15) with the same magnitudes as the original data given in chronological order, and then calculated the probabilities of the realizations of the same combinations as before for each realized residual process. Rows (ii)–(vi) of Table 4 show the observed combinations and the assessments of their probabilities, which exhibit the ordinary probabilities as expected. The implication of the observed small probability events is that the magnitude distribution is *dependent* on the past occurrence times of shocks. In other words, seismic quiescence as defined previously can provide useful information for predicting a major earthquake.

Finally, the stability of the residual process was examined from the predictive viewpoint. The data were divided into two parts, and the earlier part (1885–1949) was used to fit the selected model. Estimated values of parameters, for comparison with Table 3, were $\mu = .00662$, $K = .01769$, $c = .01928$, $p = 1.0$, and $\beta = 1.51912$ with the

Table 4. Probability Assessments of Realized Combinations

Case outcome	$M \geqq 7.4$ ($\xi < -2$), S:F	$M \geqq 7.7$ ($\xi < -2$), S:F	$M \geqq 7.4$ ($\xi < -1.5$), S:F
These data (i)	5:0 (.00030)	3:0 (.00024)	5:6 (.0060)
Simulated (ii)	0:2 (.64)	—	2:8 (.34)
Simulated (iii)	2:2 (.18)	0:1 (.94)	4:13 (.25)
Simulated (iv)	0:3 (.52)	—	3:11 (.31)
Simulated (v)	2:3 (.26)	1:4 (.28)	3:13 (.42)
Simulated (vi)	1:6 (.79)	0:2 (.88)	1:16 (.87)

NOTE: *S* indicates the number of cases in which major earthquakes with magnitudes of the given inequalities took place within a year after the end of the defined quiescences in the text. *F* indicates the number of cases in which the major shock did not take place. The assessed probabilities for the outcomes (*S:F*) are shown in parentheses.

(i)

(ii)

(iii)

(iv)

(v)

(vi)

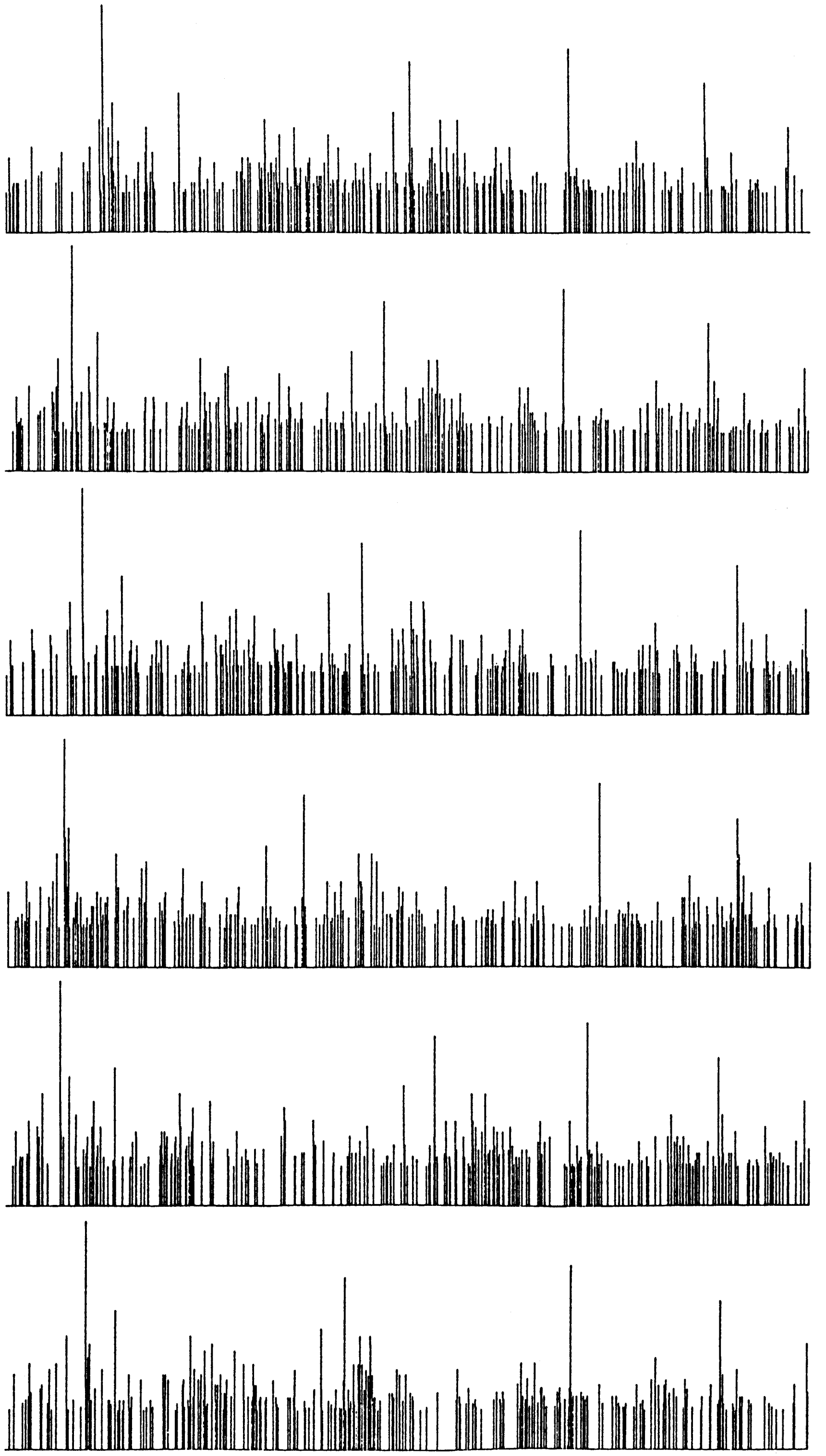

Figure 16. The Magnitudes Versus Occurrence Times Plot. (i) Real data; (ii)–(vi) simulated data. The order of the magnitudes are the same throughout (i)–(vi).

reference magnitude $M_r = 6.0$. Then Figure 17, corresponding to Figure 15(3a), is obtained for 1885–1980, in which similar features are seen. In 1951, especially, remarkable quiescence is observed before the large shock ($M8.2$) in 1952. In the period 1950–1980 there are four cases in which $\xi = \xi(\Delta N, 8)$ crosses the level $\xi = -2$ from below, within one year after which three large shocks with $M > 7.4$ occurred.

5. CONCLUSION AND SOME REMARKS

We have seen that the epidemic-type model that includes the effect of magnitude gave a better fit to the data than any of the restricted trigger models. This model is defined in terms of the conditional intensity rate, or seismic risk function of time, based on the following simple assumptions: (a) The background seismic activity is gen-

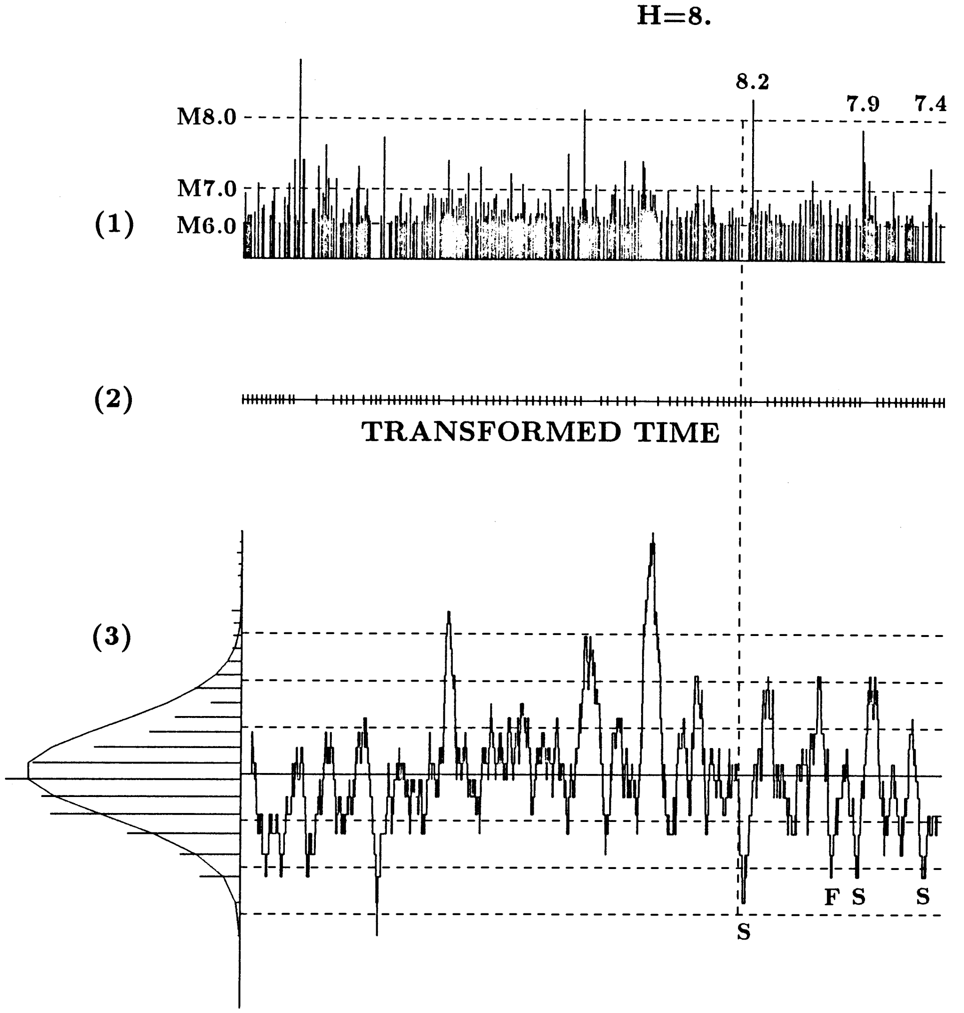

Figure 17. (1) Magnitudes Versus Occurrence Times of the Residual Process Obtained by the Selected Model for the First Part of the Data (1885–1949). The vertical dotted line shows the end of 1949. (2) and (3) are similar to (2) and (3a) in Figure 15 but are extrapolated for the last part (1950–1980).

erated by a stationary Poisson process with a constant hazard rate. (b) Each shock has a risk of stimulating aftershocks proportional to $e^{\beta M}$, where M is the magnitude of the shock. (c) The hazard rate of aftershocks decreases with time according to the modified Omori law, $K/(t + c)^p$.

A new method of residual analysis for a modeled point process was developed. This is based on the change of time scale (25) using the estimated conditional intensity (seismic risk function of time). I investigated whether the transformed time-scale version of the data is distributed as a compound Poisson process. The residual analysis for seismicity in the Off Tohoku area indicated that the Poisson model in the changed time scale seems to be acceptable, except during a time period with untypical characteristics. The magnitude distribution, however, is *not* independent of the history of occurrences, especially just after seismic quiescences. In other words, seismic quiescence can be useful for predicting a coming major earthquake.

There were two major earthquakes of the $M8$ class, that is, Off Sanriku 1896 ($M8.5$) and 1933 ($M8.1$), prior to which I did not find quiescence in Figure 15. Both of these earthquakes have unusual characteristics compared with the other major shocks. The 1896 earthquake was a "tsunami earthquake" (Kanamori 1972), which is characterized by an unusually large amount of low-frequency waves. Because of this, the estimated instrumental magnitude is only 6.8 (Utsu 1979), although Abe (1979) gave a magnitude 8.6 estimated from the size of the tsunami. The 1933 Sanriku ($M8.1$) earthquake has been characterized as having the "normal-fault" type of mechanism; that is to say, it was probably caused by gravitational pull exerted by the sinking lithosphere (Kanamori 1971), whereas most of the other major shocks in the east off the Tohoku area are caused by compressive forces. Excepting these, we are left with the striking conclusion that the mechanism underlying very large shocks can affect beforehand the seismicity of a broad area.

I also made a plot of ξ, similar to that in Figure 15 for only the main shocks identified by Utsu (see the data in Table 1). If we assume the stationary Poisson model for these shocks in the original time scale, we obtain results similar to those of the previous section. It is generally a very hard task, however, to carry out such identification objectively by using the available earthquake catalog (see the last two paragraphs in Sec. 2.1).

Finally, many seismologists believe that there is a *spatial* feature of the earthquake process, known as the *seismic gap*, which is related to the seismic quiescence. This feature, if identified early, could be used not only for the prediction of the location of a large shock in space, but in magnitude as well as in time: A large gap can be expected to be associated with a strong earthquake (see, e.g., Ohtake, Matumoto, and Latham 1977). In the present data set, I have so far not succeeded in identifying seismic gaps. This might be due to the relatively small number of shocks compared with the size of the area. Even if we have a more extensive data set, I believe that such identification would be difficult without suitable modeling of the space–time properties. For example, Motoya (1984) investigated the seismic activity of the whole Hokkaido area and its vicinity, using the catalog of micro-earthquakes (1977–1984) compiled by Hokkaido University. He divided the area into 42 regions and constructed a monthly time series of the percentage of the numbers of regions whose micro-earthquake activity was lower that month than their historical rates. In spite of his success in finding four clear quiescences in the time series just before large shocks, he could not identify the seismic gaps. The Hokkaido area is very active seismically, so the data set is complex. In addition, Motoya's methods are less sophisticated than mine. Using data for all of Japan and analyzing aftershocks instead of main shocks with the techniques of this article, Matsu'ura (1986) recently found 18 quiescences in 11 aftershock sequences before unusually large aftershocks not smaller by more than 1.2 units of magnitude than the main shock. Some space–time plots of hypocenters in her paper show that the recovered activity after the quiescence tends to cluster near the hypocenter of the forthcoming large aftershock. Of course, determining the seismic gap for a large aftershock is easier than for a main shock, because knowledge of the hypocenter of the main shock identifies the general area in which large aftershock might occur. In addition, we may need some additional prior information, such as stress release or accumulation in each focal region, to identify the gaps for the main shocks over a wide area.

Another task for the future is the modeling of the mutual relationship between the occurrence times of the earthquake series and the corresponding magnitudes. It was indicated that the full likelihood for the relation is not as written in (22). The proper model involves a kind of feedback system that is not easily modeled.

Some Fortran programs for the analysis of point processes, including graphical residual analysis, are available in the program package TIMSAC-84 (Akaike et al. 1984, part 2). The other programs used in this article are now being prepared for publication by Ogata and Katsura.

[*Received November 1985. Revised May 1987.*]

REFERENCES

Abe, K. (1977), "Tectonic Implications of the Large Shioya–Oki Earthquakes of 1938," *Tectonophysics,* 41, 269–289.

——— (1979), "Size of Great Earthquakes of 1837–1974 Inferred From Tsunami Data," *Journal of Geophysical Research,* 84, 1561–1565.

Akaike, H. (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control,* AC-19, 716–723.

——— (1977), "On Entropy Maximization Principle," in *Application of Statistics,* ed. P. R. Krishnaiah, Amsterdam: North-Holland, pp. 27–41.

——— (1983), "Information Measure and Model Selection," *Bulletin of the International Statistical Institute,* 50, Book 1, 277–291.

Akaike, H., Ozaki, T., Ishiguro, M., Ogata, Y., Kitagawa, G., Tamura, Y. H., Arahata, E., Katsura, K., and Tamura, Y. (1984), *Time Series Analysis and Control Program Package,* TIMSAC-84, Tokyo: The Institute of Statistical Mathematics.

Baudin, M. (1981), "Likelihood and Nearest Neighbor Distance Properties of Multidimensional Poisson Cluster Processes," *Journal of Applied Probability,* 18, 879–888.

Berman, M. (1983), Comment on "Likelihood Analysis of Point Processes and Its Applications to Seismological Data," by Y. Ogata, *Bulletin of the International Statistical Institute,* 50, Book 3, 412–418.

Cox, D. R., and Snell, E. J. (1968), "A General Definition of Residuals," *Journal of the Royal Statistical Society,* Ser. B, 30, 248–275.

Fletcher, R., and Powell, M. J. D. (1963), "A Rapidly Convergent Method for Minimization," *Computer Journal,* 6, 163–168.

Gutenberg, B., and Richter, C. F. (1944), "Frequency of Earthquakes in California," *Bulletin of the Seismological Society of America,* 34, 185–188.

——— (1954), *The Seismicity of the Earth,* Princeton, NJ: Princeton University Press.

Habermann, R. E., and Wyss, M. (1984), "Background Seismicity Rates and Precursory Quiescence: Imperial Valley, California," *Bulletin of the Seismological Society of America,* 74, 1743–1756.

Hawkes, A. G. (1971), "Point Spectra of Some Mutually Exciting Point Processes," *Journal of the Royal Statistical Society,* Ser. B, 33, 438–443.

Hawkes, A. G., and Adamopoulos, L. (1973), "Cluster Models for Earthquakes—Regional Comparisons," *Bulletin of the International Statistical Institute,* 45, Book 3, 454–461.

Hawkes, A. G., and Oakes, D. A. (1974), "A Cluster Process Representation of a Self-Exciting Process," *Journal of Applied Probability,* 11, 493–503.

Jeffreys, H. (1938), "Aftershocks and Periodicity in Earthquakes," *Gerland's Beitrage für Geophysik,* 53, 111–139.

Kanamori, H. (1971), "Seismological Evidence for a Lithospheric Normal Faulting—The Sanriku Earthquake of 1933," *Physics of the Earth and Planetary Interiors,* 4, 289–300.

——— (1972), "Mechanism of Tsunami Earthquakes," *Physics of the Earth and Planetary Interiors,* 6, 346–359.

——— (1981), "The Nature of Seismicity Patterns Before Large Earthquakes," in *Earthquake Prediction* (Maurice Ewing Series, 4), eds. D. W. Simpson and P. G. Richards, Washington, DC: American Geophysical Union, pp. 1–19.

Kendall, D. G. (1949), "Stochastic Processes and Population Growth," *Journal of the Royal Statistical Society,* Ser. B, 11, 230–264.

Lomnitz, C. (1966), "Magnitude Stability in Earthquake Sequences," *Bulletin of the Seismological Society of America,* 56, 247–249.

——— (1974), *Global Tectonics and Earthquake Risk,* London: Elsevier.

——— (1982), "What Is a Gap?," *Bulletin of the Seismological Society of America,* 72, 1411–1413.

Lomnitz, C., and Hax, A. (1966), "Clustering in Aftershock Sequences," *American Geophysical Union Geophysics Monograph,* 10, 502–508.

Lomnitz, C., and Nava, F. A. (1983), "The Predictive Value of Seismic Gaps," *Bulletin of the Seismological Society of America,* 73, 1815–1824.

Matsu'ura, R. S. (1986), "Precursory Quiescence and Recovery of Aftershock Activities Before Some Large Aftershocks," *Bulletin of the Earthquake Research Institute,* 61, 1–65.

Mogi, K. (1963), "Some Discussions on Aftershocks, Foreshocks, and Earthquake Swarms—The Fracture of a Semi-infinite Body Caused by an Inner Stress Origin and Its Relation to the Earthquake Phenomena," *Bulletin of the Earthquake Research Institute,* 41, 615–658.

——— (1968), "Source Locations of Elastic Shocks in the Fracturing Process in Rocks (1)," *Bulletin of the Earthquake Research Institute,* 46, 1103–1125.

——— (1973), "Relationship Between Shallow and Deep Seismicity in the Western Pacific Region," *Techtonophysics,* 17, 1–22.

Motoya, Y. (1984), "Study of Seismicity in and Around Hokkaido With Relation to Earthquake Prediction," unpublished D.Sc. thesis, Hokkaido University, Dept. of Geophysics.

Ogata, Y. (1981), "On Lewis' Simulation Method for Point Processes," *IEEE Transactions on Information Theory,* IT-27, 23–31.

——— (1983), "Estimation of the Parameters in the Modified Omori Formula for Aftershock Frequencies by the Maximum Likelihood Procedure," *Journal of Physics of the Earth,* 31, 115–124.
——— (1987), "Long Range Memory Property of Earthquake Occurrences and Statistical Models for Standard Seismic Activity" (in Japanese), in *Suri Zisin Gaku* (*Mathematical Seismology*) *II,* Cooperative Research Report 3, Tokyo: The Institute of Statistical Mathematics, pp. 115–124.
Ogata, Y., and Akaike, H. (1982), "On Linear Intensity Models for Mixed Doubly Stochastic Poisson and Self-Exciting Point Processes," *Journal of the Royal Statistical Society,* Ser. B, 44, 102–107.
Ogata, Y., and Katsura, K. (1986), "Point-Process Models With Linearly Parametrized Intensity for Application to Earthquake Data," in *Essays in Time Series and Allied Processes: Papers in Honour of E. J. Hannan,* eds. J. Gani and M. B. Priestley, *Journal of Applied Probability,* 23A, 291–310.
Ogata, Y., and Shimazaki, K. (1984), "Transition From Aftershock to Normal Activity: The 1965 Rat Islands Earthquake Aftershock Sequence," *Bulletin of the Seismological Society of America,* 74, 1757–1765.
Ohtake, M., Matumoto, T., and Latham, G. V. (1977), "Seismicity Gap Near Oaxaca, Southern Mexico as a Probably Precursor to a large Earthquake," *Pageoph,* 115, 375–385.
Okada, H. (1978), "Premonitory Seismicity Lowering Triggered by a Remote Large Earthquake," *Geophysical Bulletin of Hokkaido University,* 37, 43–51.
Papangelou, F. (1972), "Integrability of Expected Increments of Point Processes and Related Random Change of Scale," *Transactions of the American Mathematical Society,* 165, 483–506.
Shimizu, R., and Yuasa, M. (1984), "Normal Approximation for Asymmetric Distributions," *Proceedings of the Institute of Statistical Mathematics,* 32, 141–158.
Utsu, T. (1961), "A Statistical Study on the Occurrence of Aftershocks," *Geophysical Magazine,* 30, 521–605.
——— (1962), "On the Nature of Three Alaskan Aftershock Sequences of 1957 and 1958," *Bulletin of the Seismological Society of America,* 52, 279–297.
——— (1969), "Aftershocks and Earthquake Statistics (I)—Some Parameters Which Characterize an Aftershock Sequence and Their Interaction," *Journal of the Faculty of Science,* Hokkaido University, Ser. VII (Geophysics), 3, 129–195.
——— (1970), "Aftershocks and Earthquake Statistics (II)—Further Investigation of Aftershocks and Other Earthquake Sequences Based on a New Classification of Earthquake Sequences," *Journal of the Faculty of Science,* Hokkaido University, Ser. VII (Geophysics), 3, 197–266.
——— (1971), "Aftershocks and Earthquake Statistics (III)—Analyses of the Distribution of Earthquakes in Magnitude, Time and Space With Special Consideration to Clustering Characteristics of Earthquake Occurrence (1)," *Journal of the Faculty of Science,* Hokkaido University, Ser. VII (Geophysics), 3, 379–441.
——— (1972), "Aftershocks and Earthquake Statistics (IV)—Analyses of the Distribution of Earthquakes in Magnitude, Time, and Space With Special Consideration to Clustering Characteristics of Earthquake Occurrence (2)," *Journal of the Faculty of Science,* Hokkaido University, Ser. VII (Geophysics), 4, 1–42.
——— (1975), "Correlation Between Shallow Earthquake in Kwanto Region and Intermediate Earthquake in Hida Region, Central Japan," *Zisin* (*Journal of the Seismological Society of Japan*), 2nd ser., 28, 303–311.
——— (1979), "Seismicity of Japan From 1885 Through 1925: A New Catalog of Earthquakes of $M \geqq 6$ Felt in Japan and Smaller Earthquakes Which Caused Damage in Japan" (in Japanese), *Bulletin of the Earthquake Research Institute,* 54, 254–308.
——— (1982), "Catalog of Large Earthquakes in the Region of Japan From 1885 Through 1980," *Bulletin of the Earthquake Research Institute,* 57, 401–463.
Utsu, T., and Seki, A. (1955), "Relation Between the Area of Aftershock Region and the Energy of the Main Shock," *Zisin* (*Journal of the Seismological Society of Japan*), 2nd ser., ii, 7, 233–240.
Vere-Jones, D. (1966), "A Note on the Statistical Interpretation of Båth's Law," *Bulletin of the Seismological Society of America,* 59, 1535–1541.
——— (1970), "Stochastic Models for Earthquake Occurrence" (with discussion), *Journal of the Royal Statistical Society,* Ser. B, 32, 1–62.
——— (1975), "Stochastic Models for Earthquake Sequences," *Geophysical Journal of the Royal Astronomical Society,* 42, 811–826.
Vere-Jones, D., and Davies, R. B. (1966), "A Statistical Survey of Earthquakes in the Main Seismic Region of New Zealand. Part 2. Time Series Analyses," *New Zealand Journal of Geology and Geophysics,* 9, 251–284.
Vere-Jones, D., and Ozaki, T. (1982), "Some Examples of Statistical Estimation Applied to Earthquake Data," *Annals of the Institute of Statistical Mathematics,* 34, 189–207.
Yoshii, T. (1975), "Proposal of the 'Aseismic Front,' " *Zisin* (*Journal of the Seismological Society of Japan*), 2nd ser., 28, 365–367.