Statistical model for standard seismicity and detection
of anomalies by residual analysis

Y. OGATA

Institute of Statistical Mathematics, Minami-Azabu 4-6-7, Minato-ku, Tokyo 106 (Japan)

Abstract

Ogata, Y., 1989. Statistical model for standard seismicity and detection of anomalies by residual analysis. In: F.Evison (Editor), Earthquake Hazard and Prediction: Time-Variable Earthquake Hazard. *Tectonophysics*, 169 (spec. sect.): 159–174.

A statistical point-process model is derived to describe the standard activity of earthquake occurrences by assuming that general seismicity is given by the superposition of aftershock sequences. The parameters are estimated by the maximum likelihood method. Using the estimated model, the *"residual point process"* of the data is defined and used to find the anomalies which are included in the data set but not captured in the considered model for the standard seismicity. For instance, seismic quiescences can be measured quantitatively by using the residual process. Some examples are provided to illustrate such analyses. Furthermore, a time series of the magnitudes on the residual point process is considered, to investigate its dependence either on the time or on the history of the seismicity. By assuming the exponential distribution at each time and modelling of the b-value, we can examine such dependences and estimate them. Two practical examples are shown.

Introduction

The quantitative study of seismic activity, based on the earthquake data set, seems to become more necessary at the present stage. Some empirical laws in conventional aftershock statistics are very useful in order to make a statistical point process model for the analysis of seismic activity in a geophysical region, since aftershocks occupy the largest portion of an earthquake catalog.

After a major earthquake we have many aftershocks. The number of aftershocks per unit time, $\lambda(t)$, decreases at a hyperbolic rate, as was first found by Omori (1894). The decay formula was given by:

$$\lambda(t) = \frac{K}{(t+c)} \tag{1}$$

where t is the lapse time from the main shock. In the 1960s and 1970s aftershock statistics were extensively investigated by Utsu (1969, 1970), who extended the formula in the following manner:

$$\lambda(t) = \frac{K}{(t+c)^p} \tag{2}$$

This is known as the modified Omori formula. Three parameters, K, c and p, have been conventionally determined from a graphical estimate of the log–log plot of the number versus lapse time from the main shock. According to the investigation of Utsu (1969), the p-values, which are considered to be a reflection of some geophysical effect, are known to vary from place to place.

Ogata (1983) proposed the maximum likelihood method to estimate the parameters of the modified Omori formula directly on the basis of origin times of aftershock occurrences. That is to say, assuming a non-stationary Poisson process with

decreasing occurrence rate (2) of the modified Omori formula, we have a log-likelihood function:

$$\log_e L(\theta) = \sum_{i=1}^{N} \log_e \lambda_\theta(t_i) - \int_S^T \lambda_\theta(t)\, dt \qquad (3)$$

where $\{t_i\}_{i=1,2,\ldots,N}$ are occurrence times of shocks in an observed time interval $[S,T]$. Taking the maximum of the function with respect to θ numerically, we can obtain the maximum likelihood estimate of $\theta = (K, c, p)$ simultaneously. Incidentally, this method provides the asymptotic error distribution of the estimate and, more importantly, provides the quantity for measuring the goodness-of-fit of the assumed models.

For how long do aftershocks continue?

The decreasing number of felt aftershocks of the Nobi earthquake ($M = 8.4$) of 1891, which was studied by Omori (1894) to derive his formula (1), are shown in Fig. 1. The solid circles in this figure are the data obtained by Omori for a time span of about ten years. The open circles are the data compiled by Utsu for another time span of 80 years. It is quite remarkable that the aftershock activity has been continuing according to the decay formula of Omori for almost one century.

Compared to the shocks felt at a certain observatory, there have not been very many shocks the magnitudes and hypocenters of which were
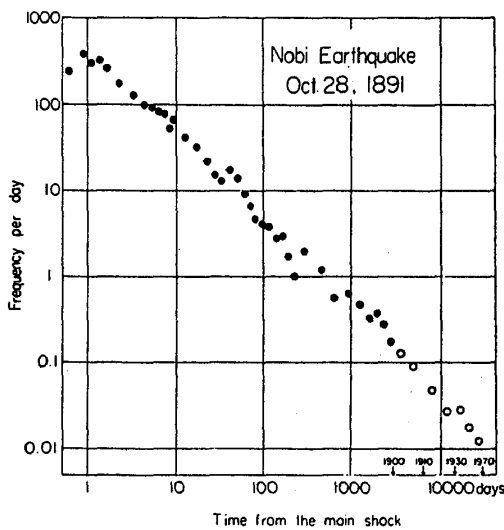

Fig. 1. Frequency of aftershocks of the Nobi earthquake felt at Gifu Local Meteorological Observatory (after Utsu, 1969).
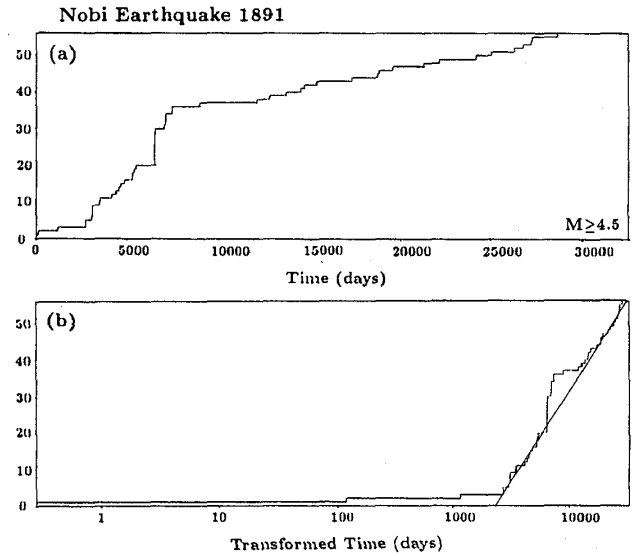

Fig. 2. Cumulative number of aftershocks of the Nobi earthquakes versus (a) original lapse time and (b) lapse time on a log scale, respectively, from the occurrence of the main shock: selected with a cut-off magnitude of 4.5 from the JMA catalog.

identified. For example, only 56 shocks with magnitude 4.5 and above in the aftershock area of the Nobi earthquake are obtained (see Fig. 2a) from the catalog compiled by the Japan Meteorological Agency (JMA): for convenience, the aftershock area of this earthquake is taken to be a circle with an 80 km radius. Nevertheless, we can see a similar feature to the above, supporting the original Omori formula (1), for a long time span. Figure 2b was plotted by taking the the lapse time from the main shock on a logarithmic scale. The cumulative number of shocks in Fig. 2b lie on a straight line up to the present (most shocks with $4.5 \leqslant M < 5.5$ are missing from the occurrence of the main shock up to about 2000 days in the JMA catalog, and the bulge around 9000 days includes other aftershocks). Note here that the logarithmic function is almost the integration of the original Omori decay function in such a way that;

$$\tau = K \log_e \frac{(t + c)}{c} = \int_0^t \frac{K}{(s + c)}\, ds \qquad (4)$$

Another example is the aftershock activity of the Tango earthquake ($M = 7.5$) of 1927. The JMA catalog has only 25 shocks (see Fig. 3a) in a circle of 50km radius with a cut-off magnitude of 4.5. However, the cumulative numbers in Fig. 3b in-
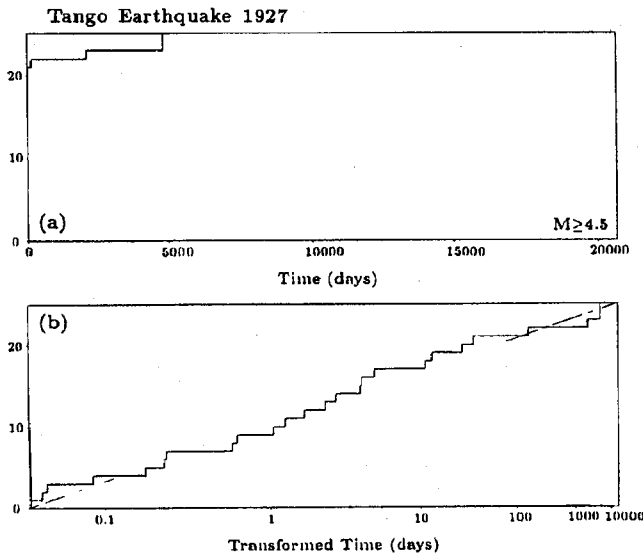
Tango Earthquake 1927

Fig. 3. Cumulative numbers of aftershocks of the Tango earthquake versus (a) original lapse time and (b) transformed lapse time in (5), respectively, from the occurrence of the main shock: selected with a cut-off magnitude of 4.5 from the JMA catalog.

crease linearly against the transformed time scale, which develops faster than the logarithmic scale. In fact, this transformation is obtained by the integration:

$$\tau = \int_0^t \frac{K}{(s+c)^p} \, ds \qquad (5)$$

of the estimated occurrence rate of the modified Omori formula, where $K = 3.71$, $c = 0.13$ and $p = 1.22$ for this data set.

Multiple aftershock sequences

The previous examples are aftershock sequences of *intraplate* earthquakes. Then how about the aftershock sequences of *interplate* earthquakes, such as those in subducting zones? The case of the Tokachi earthquake ($M = 8.2$) which occurred in 1952 off Southeast Hokkaido, is shown in Fig. 4a. Only 42 aftershocks are included in the JMA catalog within a circle of 100 km radius from the epicenter of the main shock, and over a time span up to April 1983. The cumulative numbers against the transformed time scale in Fig. 4b increase linearly for the time span of about 3000 days, but deviate upwards after that. This is interpreted as showing that this aftershock activity

then reached the level of the normal seismicity rate of the focal region.

Another case was discussed in Ogata and Shimazaki (1984), on the aftershocks of the 1965 Rat Islands earthquake ($M_w = 8.7$); see Fig. 5a. By fitting the simple modified Omori formula in (2) for the time span of the first 1000 days, we have in Fig. 5b cumulative numbers of shocks against the transformed time axis of the type (5). However, we see a bulge due to the largest aftershock, with $M_w = 7.6$, which occurred at $T_2 = 53.9$ days after the main shock, and this triggered the remarkable secondary aftershocks after that. Therefore the occurrence rate of the modified Omori formula (2) is extended to include secondary aftershocks (see Utsu, 1970) in such a way that:

$$\lambda(t) = \frac{K_1}{(t + c_1)^{p_1}} + H(t - T_2) \frac{K_2}{(t - T_2 + c_2)^{p_2}} \qquad (6)$$

where $H(t - T_2) = 0$ for $t < T_2$ and $= 1$ for $t > T_2$. Integrating this occurrence rate with the parameters of the maximum likelihood estimates, the transformation of the time $\tau = \int_0^t \lambda(s) \, ds$ is used in Fig. 5c, where the linear trend of the cumulative numbers continues up to about 2200 days after
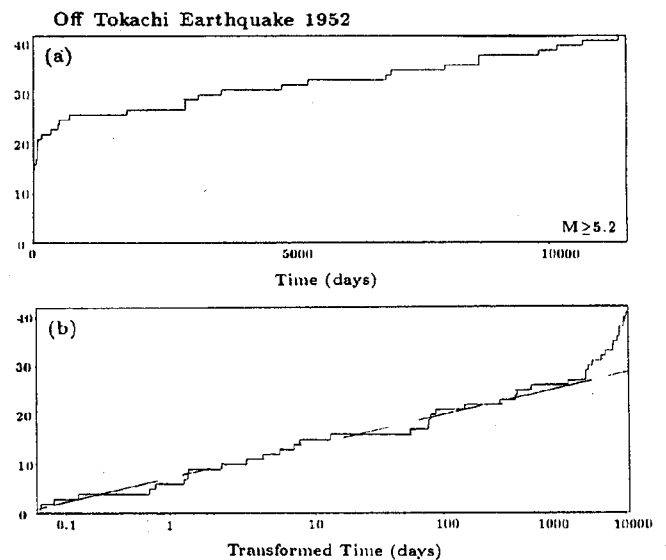
Off Tokachi Earthquake 1952

Fig. 4. Cumulative numbers of aftershocks of the Tokachi earthquake of 1952 versus (a) original lapse time and (b) transformed lapse time in (5), respectively, from the occurrence of the main shock: selected with a cut-off magnitude of 5.2 from the JMA catalog.
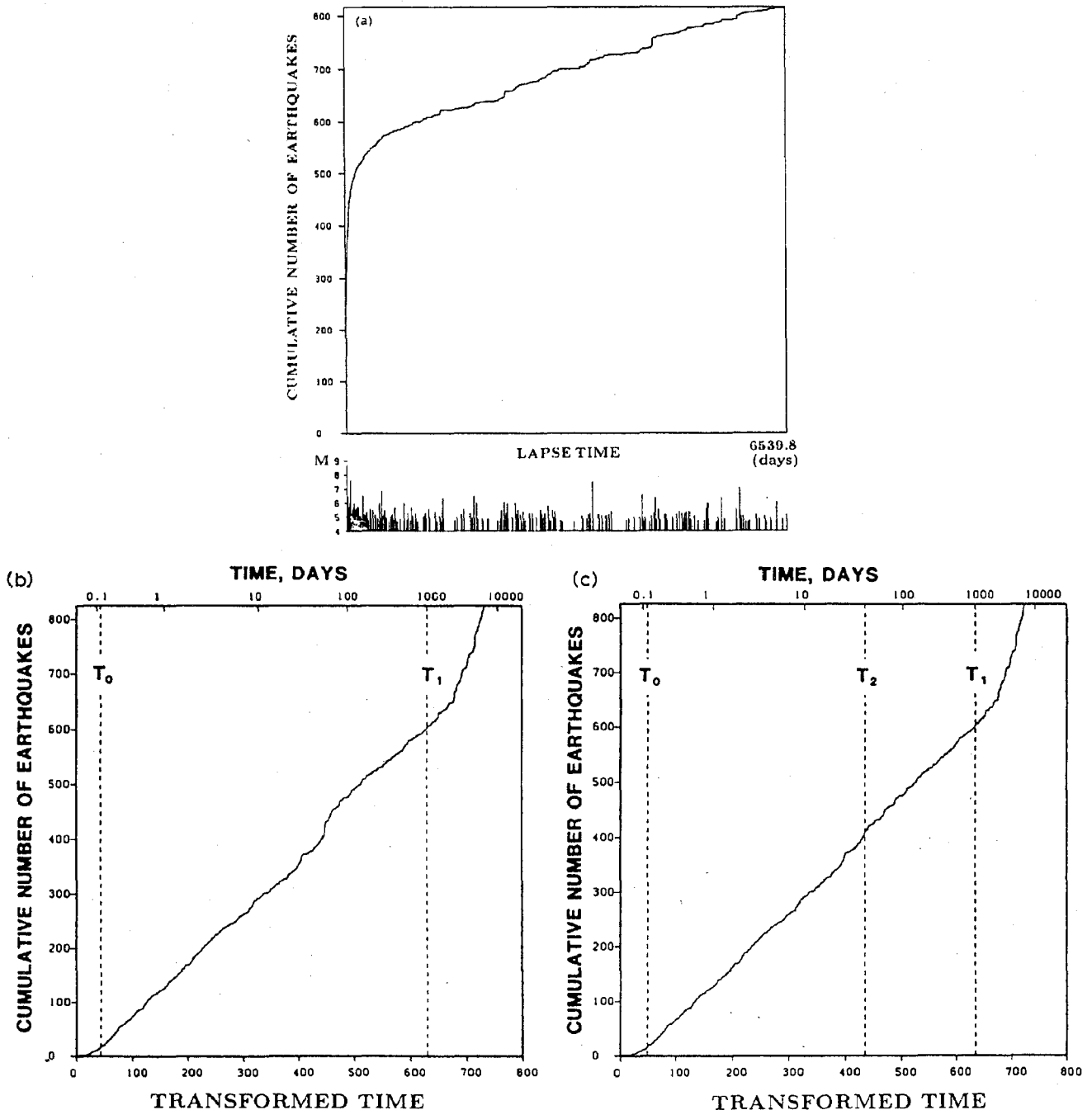
Fig. 5. Cumulative number of aftershocks of the Rat Island earthquakes of 1965 versus the following variety of lapse times. (a) Original lapse time. (b) Transformed lapse time in eqn. (5) with parameters $K = 85.088$ (shocks/day), $c = 0.204$ (days) and $p = 1.055$ which are obtained for the data in the time span between $T_0 = 3$ (hours) and $T_1 = 1000$ (days) after the main shock. (c) Transformed lapse time by the integral of eqn. (6) with parameters $K_1 = 82.284$ (shocks/day), $K_2 = 6.089$ (shocks/day), $c_1 = c_2 = 0.176$ (days), $p_1 = p_2 = 1.079$ and $T_2 = 53.9$ (days) which are obtained for the data in the time span between $T_0 = 3$ (hours) and $T_1 = 1000$ (days) from the main shock.

the main shock, and then begins the upward deviation which is similar to the case of the Tokachi aftershocks in Fig. 4b.

Of course, the occurrence rate for a multiple aftershock sequence is often necessary for the analysis of aftershocks (see Ogata, 1983), and this

is given by simply extending eqn. (6) to:

$$\lambda(t) = \sum_{m=1}^{M} H(t - T_m) \frac{K_m}{(t - T_m + c_m)^{p_m}}$$

$$= \sum_{T_m < t} \frac{K_m}{(t - T_m + c_m)^{p_m}} \qquad (7)$$

for the occurrence time T_m of the special aftershocks, where $T_1 = 0$, the occurrence time of the main shock, and the sum $\sum_{T_m < t}$ is taken for all m which satisfy the inequality $T_m < t$.

General seismic activity

Extending the occurrence rate in eqn. (7), we surmise that the seismic activity of a region is given by the superposition of the modified Omori hazard functions:

$$\lambda(t) = \mu + \sum_{t_i < t} \frac{K_i}{(t - t_i + c)^p} \qquad (8)$$

where each i refers to *any* shock at time t_i. Here μ is a constant occurrence rate that may correspond to the background seismic activity, the sum $\sum_{t_i < t}$ in eqn. (8) is taken for all i which satisfy the inequality $t_i < t$, and the constant K_i for each shock i is dependent on its magnitude M_i as well as the cut-off magnitude M_0 of the data set. Then, how do the constants $\{K_i\}$ depend on the corresponding magnitudes? In conclusion, the exponential function form:

$$K_i = K_0 \, e^{\beta(M_i - M_0)} \qquad (9)$$

has been chosen, where M_0 is a reference magnitude: for example, the cut-off magnitude is taken hereafter. The reason for this is briefly as follows. Utsu and Seki (1955) obtained the empirical formula:

$$\log_{10} S = 1.02 M - 4.0 \qquad (10)$$

for the relation between the aftershock area S and the magnitude M of the main shock (see Utsu (1971) for the data supporting this relationship). Relationship (10) suggests that the total number of aftershocks is roughly estimated as being proportional to the exponential function of a magnitude of the main shock, i.e.:

$$N \propto \exp\{\beta M\} \qquad (11)$$

Furthermore, assuming the Gutenberg–Richter law for magnitude frequency, the estimation (11) is supported by fig. 133 in Utsu (1971), which exhibits the power law decay of the cumulative frequency distribution of the number of aftershocks with $M \geqslant 6$ accompanying Japanese shallow earthquakes of $M \geqslant 6$.

Model (8) with (9) for the standard seismicity in terms of the occurrence rate of shocks is called the "*epidemic type model*". The parameter p is the same as the p-value of the decay rate of aftershocks. β measures the efficiency of a magnitude in generating its aftershocks, and this is useful in characterizing the seismicity of a focal region quantitatively: for example, swarm activities have smaller β-values than the standard (Ogata, 1987b), or possibly smaller p such that $p < 1$.

This model was fitted to the aftershocks of the Rat Island earthquake in the time span of the first 90 days to obtain estimates of the parameters. For the data, the NOAA hypocenter catalog was used with a cut-off magnitude of 4.7: see Ogata and Shimazaki (1984) for a detailed description of the data set. Here the body wave magnitudes were taken for most of the shocks, but moment magnitudes $M_w = 8.7$ and $M_w = 7.6$ were used for the main shock and the largest aftershock (at $T = 53.9$ days) respectively, and, furthermore, surface wave magnitudes were adopted if available in the NOAA catalog. Then, using the estimated hazard function of the form (8) with $\mu = 0.0$, $K_0 = 0.07212$, $c = 0.16682$, $\beta = 1.45702$ and $p = 1.34872$, the seismicity of the extended time span through 1982 is considered. We see in Fig. 6 that the cumulative numbers of shocks increase almost linearly, even on the extended part of the time span of 15 years, i.e. the upward deviation in Fig. 5c has disappeared. This shows that the present model provides a good fit to the seismicity of this area. Note here that we have used the transformation:

$$\tau_i = \int_0^{t_i} \lambda(s) \, ds$$

$$= \mu t_i + \int_0^{t_i} \sum_{t_j < t_i} \frac{K_0 \, e^{-\beta(M_j - M_0)}}{(s - t_j + c)^p} \, ds \qquad (12)$$

for each $i = 1, 2, \ldots$, where the sum $\sum_{t_j < t_i}$ is taken for all j which satisfy $t_j < t_i$.
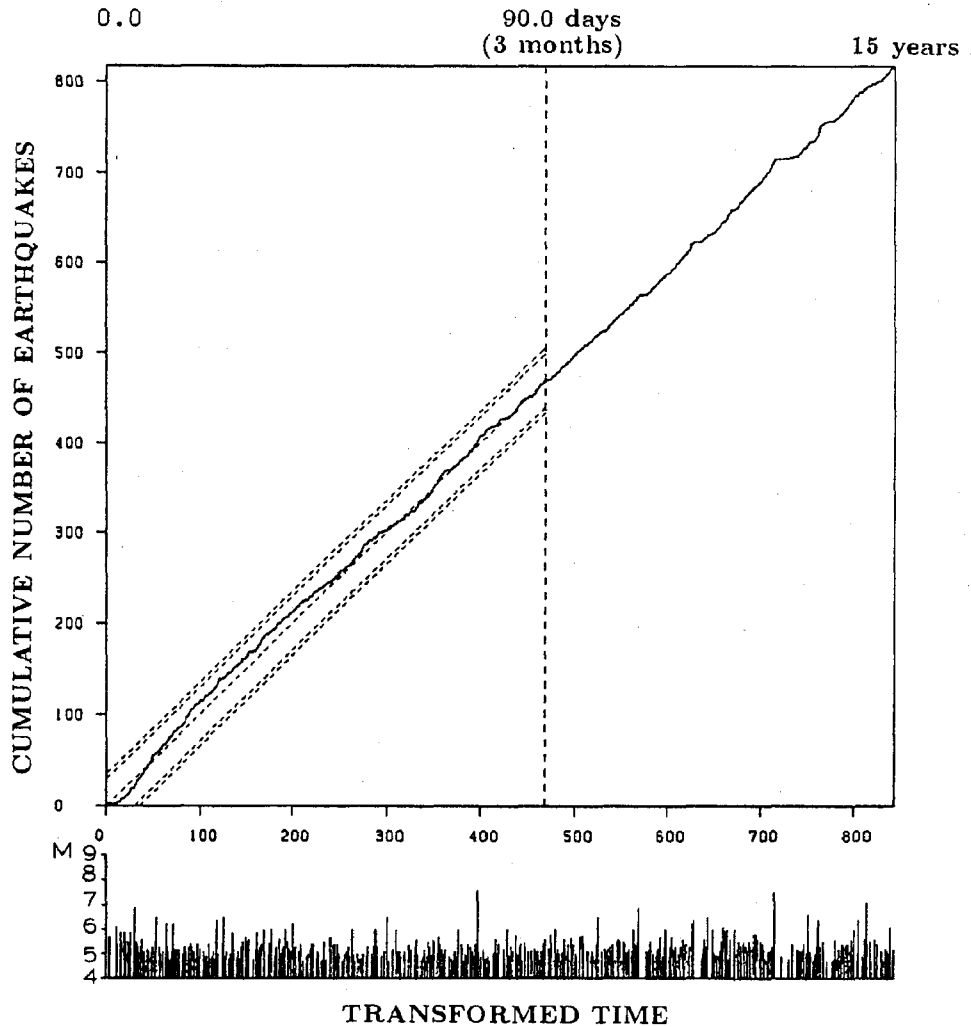
Fig. 6. Cumulative number of aftershocks of the Rat Island earthquake of 1963 versus the transformed lapse time in eqn. (12); for values of the parameters, see the text. The parameters are obtained for the time span between $T_0 = 0$ and $T_1 = 90$ (days), but the cumulative numbers are extrapolated for a further 15 years. The dotted lines indicate the average and two-sided 95% and 99% error bounds of the Kolmogorov–Smirnov statistic.

Let us consider general seismic activity in a fairly wide geophysical region. For the earthquakes within the polygonal area in Fig. 7, we obtain an origin time series of events, with their magnitudes, from the earthquake catalog of 96 years compiled by Utsu (1982), which is complete for magnitude 6 and above. This area is the most active in and around Japan (the cumulative number of events in this area is provided in Fig. 9a). By maximizing the log-likelihood, the estimated epidemic type model for the data is given by:

$$\lambda(t) = \mu + \sum_{t_i < t} \frac{K_0 \, e^{\beta(M_i - M_0)}}{(t - t_i + c)^p} \qquad (13)$$

where $\mu = 0.00536$ (shocks/day), $K_0 = 0.08680$ (shocks/day), $c = 0.01959$ (days), $p = 1.0$ and $\beta = 1.61385$ (magnitude^{-1}) for the cut-off magnitude $M_0 = 6.0$. It is worthwhile here to note the following. According to the epidemic-type model it is claimed that every shock can have its aftershocks, more or less, with a stochastic frequency proportional to the quantity (9), and therefore deposes the postulate that only main shocks can stimulate the occurrence of aftershocks. This postulate is formulated by another similar model, called the *restricted trigger model*, to compare the goodness of fit of the models to the present data set, and it was found by Ogata (1988)
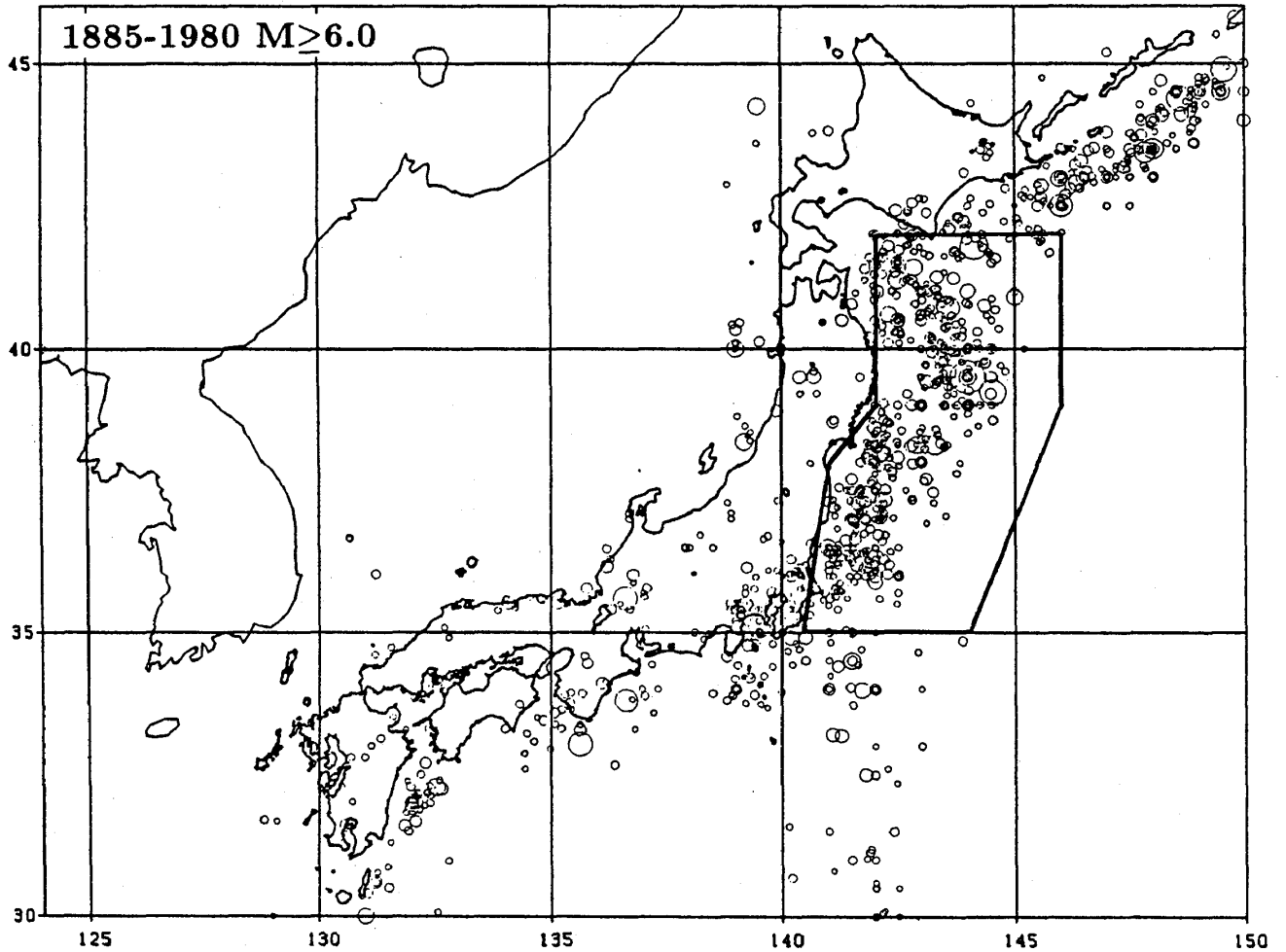
Fig. 7. The spatial distribution of large earthquakes for 1895–1980 in Japan and its vicinity, according to Utsu (1982). The data set for the polygonal region is considered in the text.

that the epidemic-type model is significantly better than the other one.

Using the transformation (12) for general seismicity, let us formulate the definition and implication of the *residual analysis* of seismicity. Suppose that the origin time data $\{t_i\}$ are generated by an occurrence rate $\lambda(t)$. Then consider the integral of the occurrence rate:

$$A_t = \int_0^t \lambda_\theta(s)\, ds \qquad (14)$$

which is an increasing function, as illustrated in Fig. 8. If we consider the time change using this increasing function such that $\tau_i = A_{t_i}$, then origin times $\{t_i\}$ are transformed one-to-one into $\{\tau_i\}$. It is known that the sequence $\{\tau_i\}$ is distributed according to the standard stationary Poisson process for the true parameter of the model. There-

fore, if the model provides a good fit to the seismicity, then the transformed data of the occurrence time, which we call the "*residual point process*", are approximated well by the standard stationary Poisson process. In other words, if we find a deviation of a characteristic property of the residual point process $\{\tau_i\}$ from that expected under the stationary Poisson process, this tells us about the existence of some features which are not captured by the estimated model. In this way, through the present modelling, we can improve our understanding of something that is included in the data set.

In Fig. 9a the original occurrence series in the polygonal area of Fig. 7 is shown, with magnitudes and cumulative numbers; and Fig. 9b is for the residual process transformed by using the
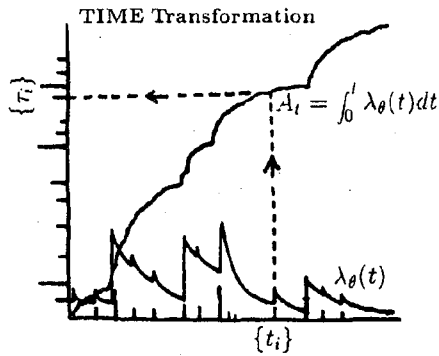
Fig. 8. A diagram of the time transformation of an original series of events into a residual point process.

estimated model (13). No remarkable clustering is seen from the cumulative of the residual process. There are many methods by which to test the residual process against the Poisson process (see Ogata, 1988), but here we are curious to know whether there are some locally unusual character-istics or outliers in the residual process.

To see such features, we count the number of points in the moving interval $[\tau - h, \tau]$. If the residual $\{\tau_i\}$ is the standard stationary Poisson process, then the counted number $N(\tau - h, \tau)$ is a Poisson random variable with mean h for any τ. Furthermore, the counted numbers, if h is not small, are almost distributed like a Gaussian pro-cess. We can use the well known Wilson–Hilferty-type transform which accelerates the approxima-tion to a Gaussian distribution (see Ogata, 1988). By setting the interval length to $h = 8$, for exam-ple, the time series of the counts with respect to τ is plotted at the bottom of Fig. 10. This seems to behave like a Gaussian process except in the part of the trajectory around the year 1938, which is significantly above the four-fold standard error. This is a kind of outlier due to the swarm-like shocks, which include five large shocks with mag-nitudes greater than $M = 7$. Abe (1977) subse-quently, and independently, reported the abnormality of this particular seismic activity.

Detection of seismic quiescence

It is frequently reported that before a major earthquake, seismic quiescence and then some-times foreshocks have occurred. On the other hand, some people have questioned the usefulness of seismic quiescence because it can be considered to be a mere effect of the decaying activity of aftershocks from the last major earthquake (see Lomnitz et al., 1982, 1983). Such a question can be disregarded here, because the time-transformed seismicity given in Fig. 10a is nothing but the residual process obtained by the modelling of such an effect.

Quiescence may correspond to the lowering of the trajectory of Fig. 10c. Comparing Fig. 10a with Fig. 10c, it is found that five large shocks with $M \geqslant 7.4$ took place within one year, after the trajectory in Fig. 10c had crossed above the level of -2. This combinatorial realization seems to be
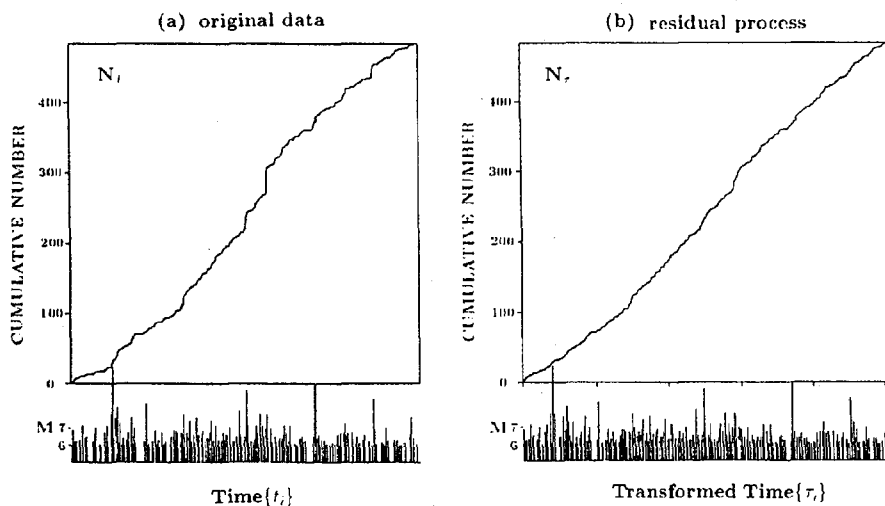

Fig. 9. Series of magnitudes and cumulative numbers of the events in the area shown in Fig. 7. (a) Original data. (b) Residual point process.
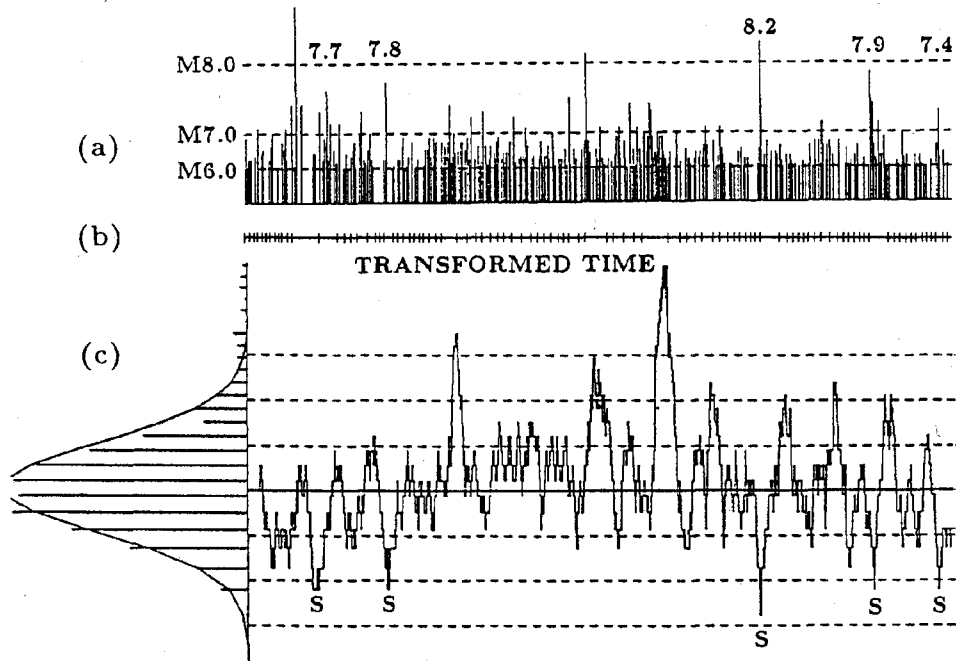
Fig. 10. (a) Magnitudes versus occurrence time of the residual process. (b) The transformed time location of 1 January for every year from 1885 to 1980. (c) Time series related to the number of points in the moving interval $[\tau - h, \tau]$ of the residual process: putting $h = 8$, each counted number is transformed into another variable by using the Wilson–Hilferty-type relation to accelerate the approximation to the standard normal random distribution $N(0, 1)$ at each time τ. Dotted lines are one-, two- and three-fold standard errors of the assumed normal distribution. The sideways histogram indicates the total length of the time span of τ of each Y-axis value, and the broken line corresponds to the theoretical distribution (approximately normal), assumed under the hypothesis that the residual process is a stationary Poisson process.
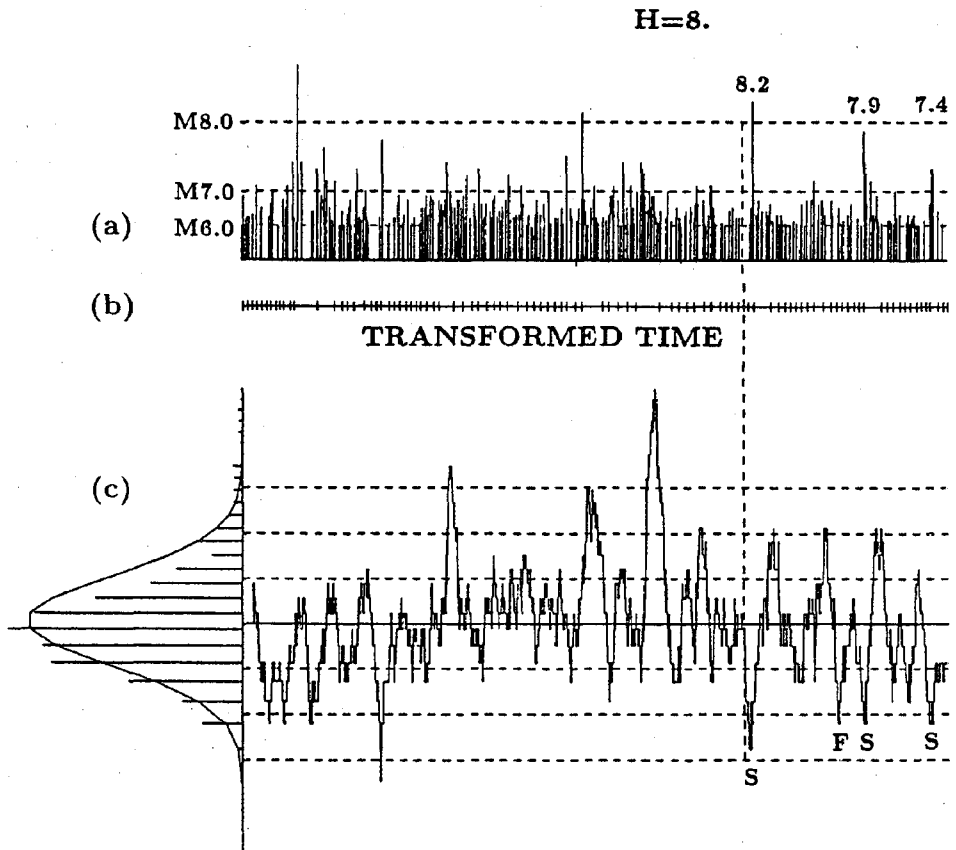

Fig. 11. (a) Magnitudes versus occurrence times of the residual process obtained by the estimated model (8) with (9) for the data in the first time span, 1885–1949. The vertical dotted line shows the end of 1949. (b, c) Similar to (b) and (c) in Fig. 10, but extrapolated for the last time span, 1950–1980.

a rare phenomenon (see Ogata, 1988, for assessments of the realization probability), and it is considered that this is related to the usefulness of seismic quiescence for predicting major earthquakes. Moreover, the stability of the residual process has been examined from the predictive viewpoint. The data were divided into two parts, 1885–1949 and 1950–1980, and the earlier 2/3 was used to fit the model. Then the residual process was extended to the other part (see Fig. 11) to show features similar to those of Fig. 10.

Now, returning to the case of seismic activity in a focal region, another two findings which are related to the present application must be mentioned. First, it was found that, even in a simple sequence of aftershocks of shorter time span, quiescence in our sense can exist, and that after such quiescence we expect to have a major aftershock which initiates secondary aftershocks (Matsu'ura, 1986). For example, let us analyse the aftershock sequence (Fig. 12) of the Western Nagano Prefecture earthquake ($M = 6.8$) which occurred in 1984 in the interior of Japan (see JMA (1986) for data). The epidemic-type model (8) with (9) is fitted to the series of shocks, with a cut-off magnitude of 2.9, for a time span of 150 days after
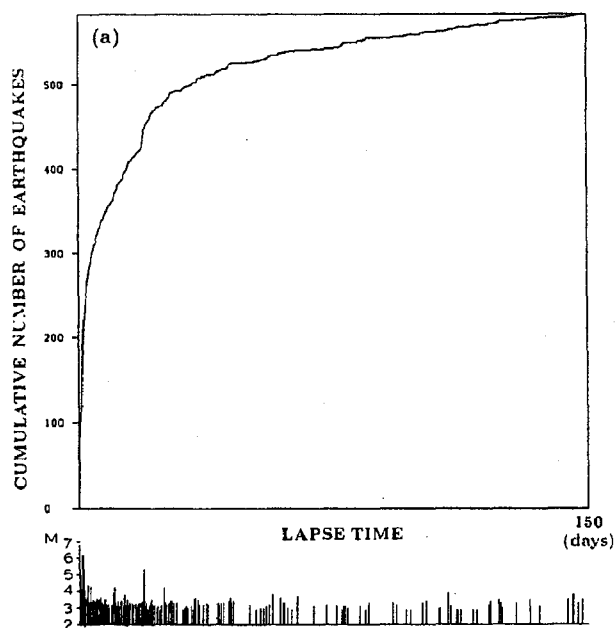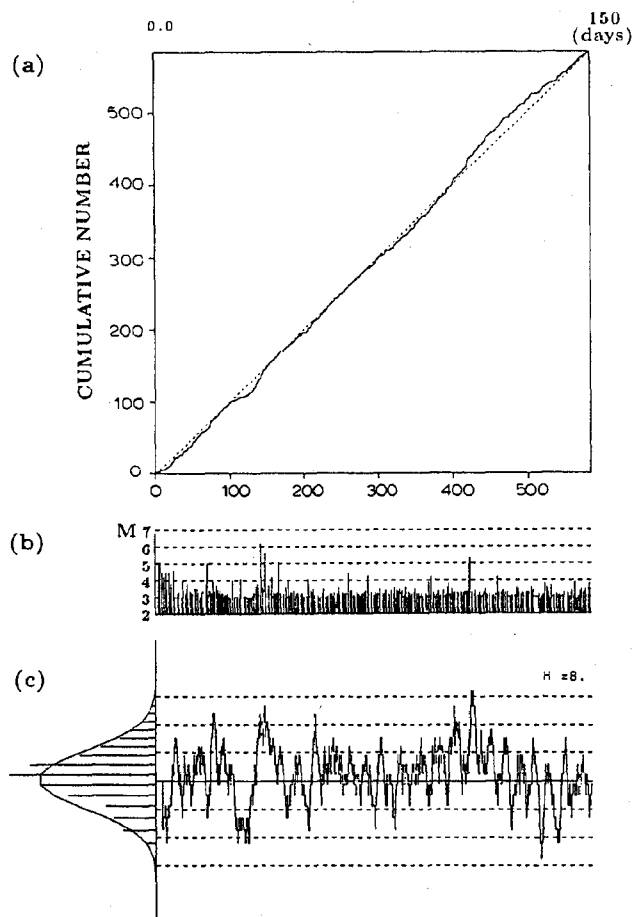

Fig. 13. Residual point process of the aftershock occurrences of the Western Nagano Prefecture earthquake: (a) cumulative numbers versus the transformed time; (b) magnitudes versus the transformed occurrence times; and (c) time series of the number of points on a moving interval with $h = 8$ (see the caption to Fig. 10 for a more detailed explanation of the graph).

the main shock. Estimated parameters are $\mu = 0.0$ (shock/day), $K_0 = 0.03187$ (shocks/day), $\beta = 1.68555$ (magnitude^{-1}), $c = 0.03783$ (days) and $p = 1.13758$. Figure 13, thus obtained, exhibits the cumulative number of points on the residual process, the time series of magnitudes on the residual process, and the number of points on the residual process in a moving interval. It is clearly seen from this figure that the quiescence appeared just before the occurrence of the largest aftershock.

Second, since there was remarkable swarm activity for about 7 years before the main shock, it may be useful to make a similar analysis to the above. From 1978 up to the occurrence time of the main shock, the JMA catalog has 208 events with

Fig. 12. Series of events and their cumulative numbers for the aftershocks of the Western Nagano Prefecture earthquake ($M = 6.8$) of 1984.
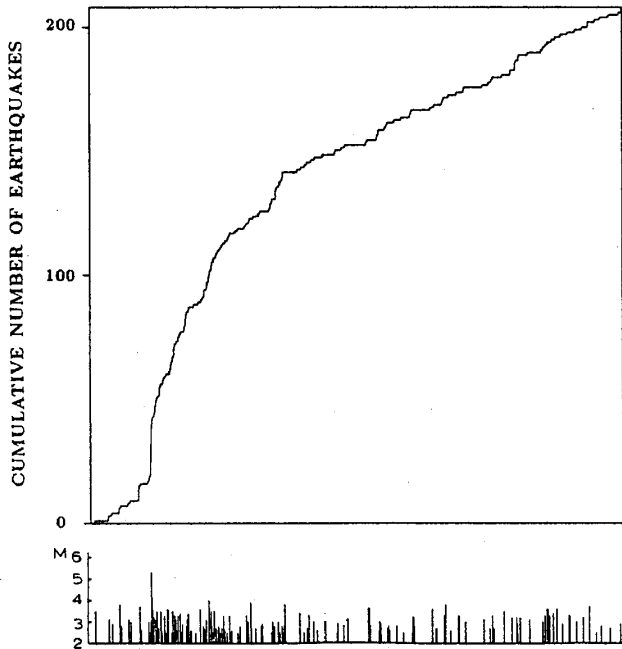
Fig. 14. Cumulative numbers and magnitudes of the swarm activity before the Nagano Prefecture earthquake versus lapse time from 1978 to the occurrence time of the main shock: events are selected with a cut-off magnitude of 2.5 from the JMA catalog.

magnitudes not smaller than 2.5: with this cut-off the data set seems to be complete for this particular time span (see Fig. 1.2.4 in JMA, 1986). The cumulative number and the magnitudes versus the occurrence time for the swarm activity are shown in Fig. 14. The epidemic-type model (8) with (9) and $M_0 = 2.5$ is fitted to the data, and the estimated parameters are $\mu = 0.02145$ (shocks/day), $K_0 = 0.007841$ (shocks/day), $\beta = 2.16391$ (magnitude^{-1}), $c = 0.01704$ (days) and $p = 0.84956$. Note the significant differences of the estimates of β and p here from those for the aftershocks of the Western Nagano earthquake given above. Incidentally, it is statistically confirmed that the seismicity from 1978 until February 1985, with a cut-off magnitude of 2.9 throughout, drastically changed before and after the main shock: here we have used the AIC (Akaike Information Criterion; Akaike, 1973) which is defined by:

$$AIC = (-2) \max_\theta \log_e L(\theta) + 2 \dim \theta$$

where the dimension of θ (dim θ) means the number of adjusted parameters. A model with a smaller AIC value means a better fit to the data.

In this particular case, two models are compared: the first hypothesis is that the seismicity did not change, so that the single epidemic model is fitted throughout; and for the second hypothesis, on the other hand, two epidemic-type models with different parameters are needed to describe the seismicity change. For such a problem of change point detection, the following comparison of the AIC's is useful (see Kitagawa and Akaike, 1978, for example); that is, AIC(swarms) + AIC(aftershocks) = 770.0 + (−2484.8), while AIC(swarms + aftershocks) = (−1705.7), where the aftershocks include the main shock. The former is considerably smaller, which supports the change of seismicity type.
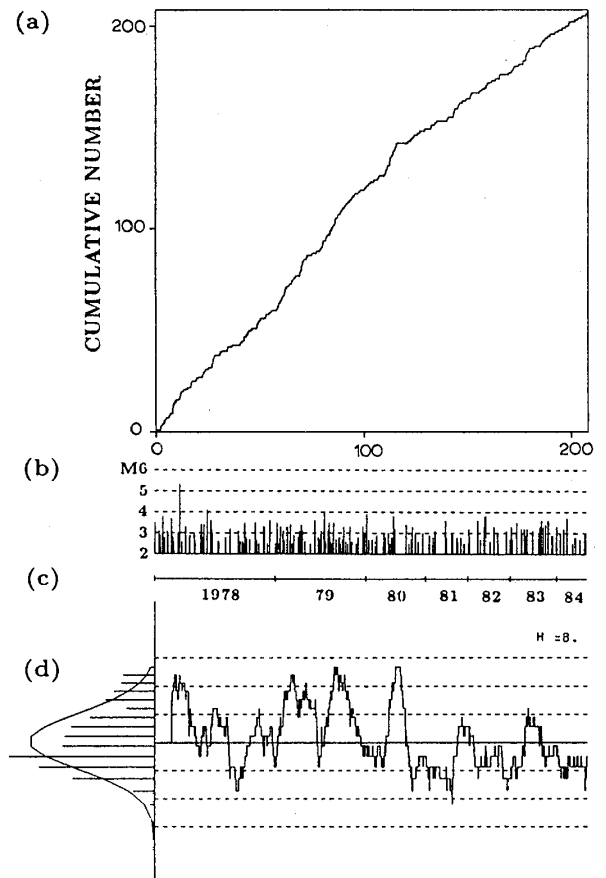

Fig. 15. Residual point process of the swarm occurrences before the Western Nagano Prefecture earthquake, where the parameters used are in the text. (a, b) Cumulative numbers and magnitudes versus the transformed time of occurrences, respectively. (c) Location of the transformed time of 1 January for every year from 1978 to 1984. (d) Time series of the number of points of the moving interval with $h = 8$ (see the caption to Fig. 10 for a more detailed explanation of the graph).

Now, using the estimated parameters for the swarms, $\mu = 0.02145$ (shocks/day), $K_0 = 0.00784$ (shocks/day), $c = 0.01704$ (days), $\beta = 2.16391$ (magnitude^{-1}) and $p = 0.84956$ for the cut-off magnitude $M_0 = 2.5$, Fig. 15 is obtained for the residual process of the swarms. Very clear quiescence in our sense is shown from about four years before the main shock, in both the cumulative and the counted events in the interval $[\tau, \tau + 8]$. To examine the stability of the residual process from the predictive viewpoint, as was done in Fig. 11, the first two years' data is used to fit the model, and then the residual process is extended to the other part. The parameters used here are $\mu = 0.03949$ (shocks/day), $K_0 = 0.01010$ (shocks/day), $c = 0.02544$ (days), $\beta = 2.03080$ (magnitude^{-1}) and $p = 0.82178$. Thus we obtain Fig. 16, from which we see a clearer quiescence.
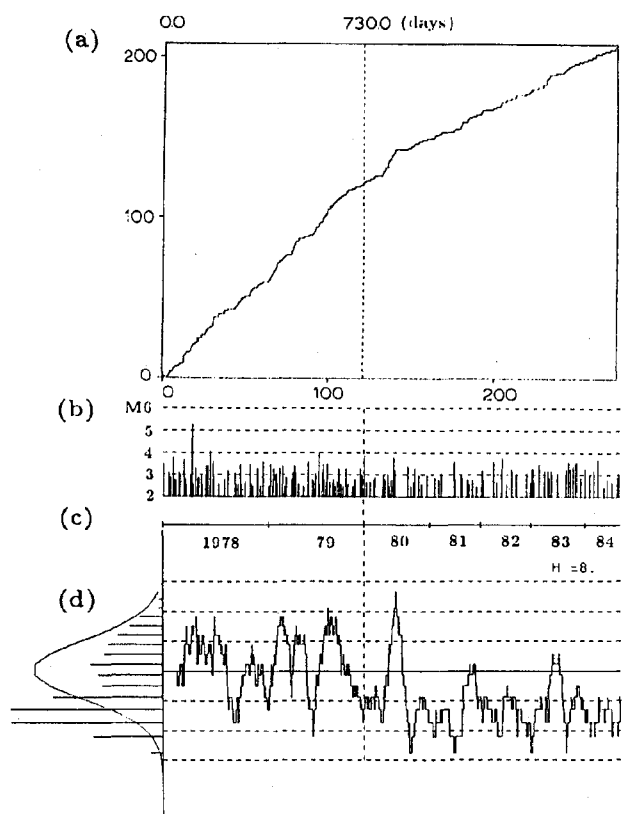

Fig. 16. (a, b) Cumulative number and magnitudes, respectively, versus occurrence time of the residual process of the swarms obtained by the epidemic-type model with the estimated parameters in the text for the data of the first time span, 1978–1979. The vertical dotted line shows the end of 1979. (c, d) Similar to (c) and (d) in Fig. 14, respectively, but extrapolated for the last time span from the beginning of 1980 to the occurrence time of the main shock.

Further examples are given by Matsu'ura (1986) who, analysing the residuals of simpler models of modified Omori formulae (2), (6) and (7), extensively investigated 11 aftershock sequences in Japan for rather short time spans and found 18 time intervals of quiescence together with a recovery of activity before 14 major aftershocks which are not smaller than the main shock by more than 1.2 units of magnitude.

Time evolution of *b*-values

Finally, let us discuss the time evolution of *b*-values for magnitude distributions of seismic activity. A first question is that of whether or not the magnitudes $\{M_i\}$ as random variables are independently and identically distributed. As an example, the residual process of the Western Nagano Prefecture aftershocks is again considered here. Figure 17a shows plotted the time evolution of the maximum likelihood estimates of *b*-values, $\{\hat{b}_i\}$, and their error bars, where $\hat{b}_i = N_i/\sum_{k=1}^{N}(M_i - M_0)$, for the *i*th divided time interval, where $M_0 = 2.85$ is the cut-off magnitude. Figure 17b shows the optimum smoothed time evolution of *b*-values and one-fold error bounds which were automatically obtained by an objective method using the spline function, assuming that the change of *b*-value is smooth (Ogata and Katsura, 1988). It is strongly suspected that the low *b*-value at the beginning is due to the absence of smaller events, because this is immediately after the occurrence of the main shock. On the other hand, the trough around the largest aftershock seems very interesting. A natural question arises as to whether the lowering of the *b*-value started before the occurrence of the largest aftershock or whether this is just an artificial effect of the smoothing method.

This question is related to the other statistical question of whether or not the magnitude series depends on the history of either the magnitudes or the occurrence pattern of the residual process. To investigate this, let us consider a flexible class of statistical model. A time-dependent family of exponential distributions:

$$F(x \mid H_t) = \text{Prob}\{M_t - M_0 < x \mid H_t\}$$
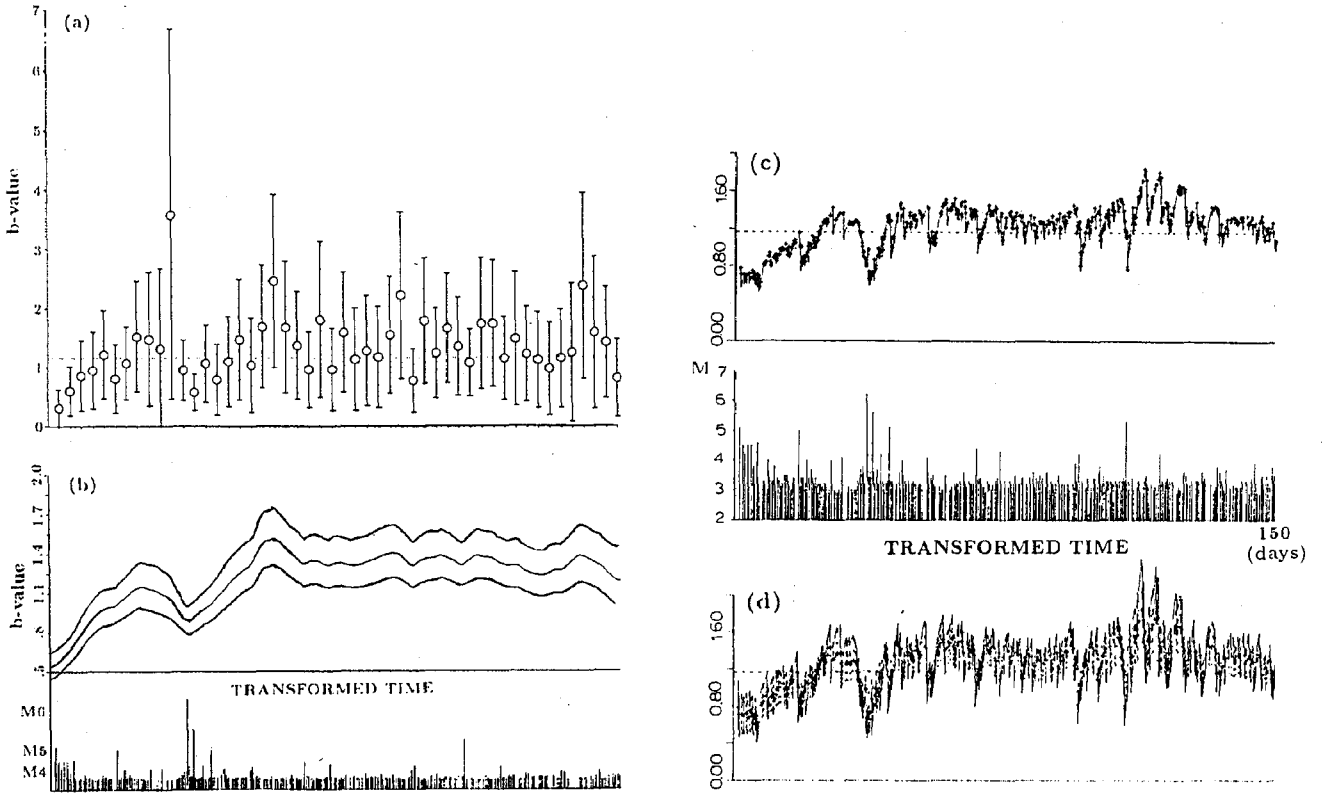$$= 1 - 10^{-b(t)x} = 1 - e^{-x/\sigma(t)} \qquad (15)$$

Fig. 17. Time evolution of b-values in the residual process of the aftershocks. (a) Histogram of the maximum likelihood estimates and their error bars. (b) Spline b-value function estimated by the objective Bayesian method and its one-fold error bounds. (c) History-dependent estimate of b-values $b(t) = \log_{10} e / \sigma(t)$ by using the model (16) with (17) (see text for the estimated parameters). (d) One- and two-fold error bounds of (c).

conditional on the history H_t of the seismicity, is constructed in such a way that $b(t)$ and $\sigma(t)$ are dependent on H_t and are described by the following relation:

$$\sigma(t) = \mu + \sum_{\tau_i < \tau} g(\tau - \tau_i) + \sum_{\tau_i < \tau} h(\tau - \tau_i) M_i \quad (16)$$

where $g(\cdot)$ and $h(\cdot)$ are response functions of an event and are parameterized by the following type polynomials:

$$g(x) = \sum_{j=1}^{J} a_j x^{j-1} e^{-cx}$$

and:

$$h(x) = \sum_{k=1}^{K} b_k x^{k-1} e^{-dx} \quad (17)$$

From (15) the likelihood function is written as:

$$L(\theta) = \prod_{i=1}^{N} \left\{ \frac{1}{\sigma(t_i)} e^{-(M_i - M_0)/\sigma(t_i)} \right\} \quad (18)$$

where $M_0 = 2.85$ for the present data set. Maximizing this function with respect to the parameters $\theta = (\mu, c, d, a_1, \ldots, a_J, b_1, \ldots, b_K)$ for fixed orders J and K of the polynomials in eqn. (17), we obtain the maximum likelihood estimates. To find the optimal orders, we use AIC defined in the previous section. The pair of orders (J, K) with the smallest AIC value indicates the best fit among the considered pairs. The AIC's for the present data are listed in Table 1.

TABLE 1

List of AIC's

J			(h)		
	K: 0	1	2	3	4
0	17.2	5.3	7.1	6.5	7.3
(g) 1	20.1	−5.1	−3.3	−2.5	
2	21.4	−3.1	−1.3	−0.5	
3	22.7	−1.5	−0.4	1.1	

The minimum AIC is attained by $(J, K) = (1, 1)$, which shows that the magnitude distribution, for this particular data set, is dependent on the past occurrence times of shocks as well as their magnitudes, and that both the response functions $g(\cdot)$ and $h(\cdot)$ should be taken as simple exponential functions. The maximum likelihood parameters of the estimated model (16) with (17) are $\mu = 0.50924$, $c = 0.021872$ (day^{-1}), $d = 0.25859$ (day^{-1}), $a_1 = -0.0056282$ and $b_1 = 0.073947$.

Using this model, we have assessed b-value evolution $b(t) = \log_{10} e/\sigma(t)$ in Figs. 17c and d for the same aftershock sequence of the Western Nagano Prefecture earthquake. The approximate one- and two-fold error bounds are also obtained, since the inverse matrix of the Hessian $(-\partial^2 \log L(\theta)/\partial\theta_j \, \partial\theta_k)_{j,k=1,\ldots,5}$ of the log-likelihood provides the covariance matrix of the error distribution of the estimate, and since $\sigma(t)$ in eqn. (16) with eqn. (17) is linear with respect to the parameters other than c and d (see Ogata and Katsura, 1988, for example).

It is remarkable that b-values in the quiescent period before the largest aftershock are high, which means that smaller shocks are predominant. This may be consistent with the fact that a quiescence is usually found by a suitable selection of the cut-off magnitude. Furthermore, b-values diminished just before the largest aftershock: this is also consistent with the fact that the b-value of foreshocks is often low, and therefore these may be called foreshocks of the largest aftershock. Also, it is statistically natural that larger shocks are more likely to occur under a smaller b-value: note that $b(t)$ or $\sigma(t)$, in the sense of eqn. (15) and eqn. (16), is the predictive quantity. Another large shock of $M = 5.3$ took place in the latter part in Fig. 17c, and the b-value decreased just before this occurrence.

If the magnitude frequencies are assumed to depend only on the current time t, then $b(t)$ or $\sigma(t)$ can be parameterized by the splines, polynomials or Fourier expansions to find the optimal order. For example, to analyse the same swarm activity before the Nagano Prefecture earthquake, $\sigma(t)$ is parameterized by the standard polynomials, and AIC's with the likelihood functions (18) are computed up to the 10th order. The minimum
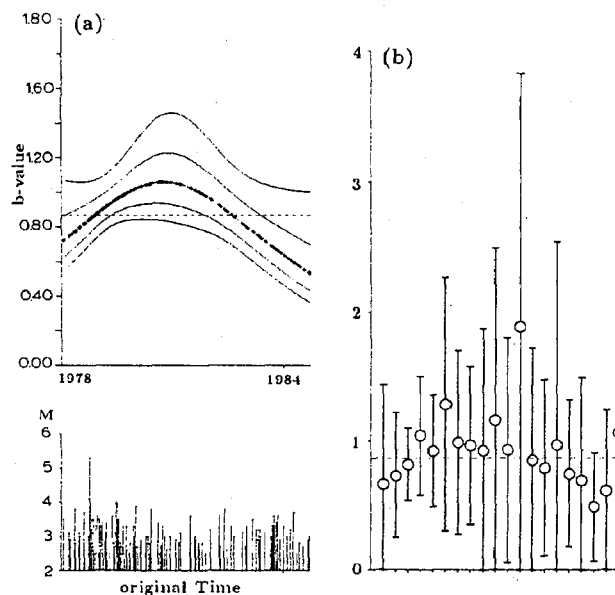

Fig. 18. The real time evolution of the b-values of swarms. (a) b-value function $b(t) = (\log_{10} e)/\sigma(t)$ with the polynomial (19) and the one- and two-fold error bounds. (b) Histogram of the maximum likelihood estimates and their error bars.

AIC is attained at the second order polynomial with the following estimated coefficients:

$$\sigma(t) = 0.66402 - 1.1000(t/T) + 1.1722(t/T)^2 \tag{19}$$

where t (days) is the lapse time from the beginning of 1978 and T is the total length of the observed interval. The b-value evolution $b(t) = \log_{10} e/\sigma(t)$ is then obtained as shown in Fig. 18, together with the assessed one- and two-fold error bounds, using the Hessian matrix of the likelihood function. It is thus seen that the b-values decreased significantly until the occurrence of the main shock. For the interest of readers, the conventional b-value estimates, similar to Fig. 17a, are also included in Fig. 18. Some details relevant to this section are described in Ogata (1987a).

Conclusions

It seems that the decay rate of aftershock sequences of the two intraplate earthquakes in Japan has followed the modified Omori formula (2) for very long time spans up to the present. In more active seismic areas, such as interplate regions, a complex model is necessary to describe the occurrence rate of earthquakes: extending the decay

rate (7) for multiple aftershock sequences, the epidemic-type model (8) with (9) is proposed for the general seismic activity. This model is based on the following simple assumptions: (a) the background seismic activity is given by a stationary Poisson process with a constant occurrence rate μ; (b) every shock has a risk of stimulating aftershocks proportionate to $\exp\{\beta M\}$, where M is the magnitude of the shock; (c) the occurrence rate of aftershocks decreases with elapsing time according to the modified Omori law $K_0/(t + c)^p$.

Another innovative method of the present paper is the systematic use of *residual analysis* of the point process data, the time scale of which is changed on the basis of the integration of the occurrence rate of the estimated model. Using the residual analysis effectively, we can find some interesting characteristic features of the data set which were not included in the modelling (a)–(c). For example, one of the residual analyses for data off the Pacific Coast of northeast Japan indicates the good fit of the model except for a particular time period. However, the magnitude distribution does not seen to be independent of the history of the occurrences, and is closely related to the existence of seismic quiescences just before some of the great shocks. Similar quiescences were seen in the swarms before the Western Nagano Prefecture earthquake, and even in its aftershock sequence before the major aftershocks. Therefore it may well be that quiescence represents the time span when the seismic rate is relatively but significantly lower than the expected rate of the standard seismicity, no matter how active or calm it is.

Assuming an exponential distribution for magnitude frequency at each time, the time evolution of the b-value is detected: this may be described by the choice of optimal statistical model for b-values of the original or the residual point processes; i.e. sometimes a time change may be necessary to investigate the causal relation to the history of the occurrence pattern of shocks.

Finally, the maximum likelihood method is efficient and sensitive to work out, so that the selection of a good quality data set, especially with good estimation of the magnitudes, seems to be crucial in fitting the models which have been introduced.

References

Abe, K., 1977. Tectonic implications of the large Shioya–Oki earthquakes of 1938. Tectonophysics, 41: 269–289.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Control. AC-19: 716–723.

JMA, 1986. The Report on the Nagano Prefecture earthquake, 1984. Tech. Rep., Japan Meteological Agency, 107.

Kitagawa, G. and Akaike, H., 1978. A procedure for the modelling of non-stationary time series. Ann. Inst. Stat. Math., 30: 351–363.

Lomnitz, C., 1982. What is a gap? Bull. Seismol. Soc. Am., 72: 1411–1413.

Lomnitz, C. and Nava, F.A., 1983. The predictive value of seismic gaps. Bull. Seismol. Soc. Am., 73: 1815–1824.

Matsu'ura, R.S., 1986. Precursory quiescence and recovery of aftershock activities before some large aftershocks. Bull. Earthquake Res. Inst., Univ. Tokyo, 61: 1–65.

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.

Ogata, Y., 1987a. Statistical models for the analysis of b-values (in Japanese). In: K. Oike (Editor), Statistical Methods for the Detection of Anomalies. Cooperative Research Report, 4. Institute of Statistical Mathematics, Tokyo, pp. 30–44.

Ogata, Y., 1987b. Long range memory property of earthquake occurrences and statistical models for standard seismic activity (in Japanese). In: M. Saito (Editor), Suri Zisin Gaku (Mathematical Seismology) II. Cooperative Research Report, 3. The Institute of Statistical Mathematics, Tokyo, pp. 115–124.

Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. J. Am. Stat. Assoc., 83 (no. 401, Applications): 9–27.

Ogata, Y. and Katsura, K., 1988. Likelihood analysis of inhomogeneity for marked point patterns. Ann. Inst. Stat. Math., 40: 29–39.

Ogata, Y. and Shimazaki, K., 1984. Transition from aftershock to normal activity: the 1965 Rat Islands earthquake aftershock sequence. Bull. Seismol. Soc. Am., 74: 1757–1765.

Omori, F., 1894. On the aftershocks of earthquake. J. Coll. Sci. Imp. Univ. Tokyo, 7: 111–200.

Utsu, T., 1969. Aftershocks and earthquake statistics (I): some parameters which characterize an aftershock sequence and their interaction. J. Fac. Sci., Hokkaido Univ., Ser. VII (Geophys.), 3: 129–195.

Utsu, T., 1970. Aftershocks and earthquake statistics (II): further investigation of aftershocks and other earthquakes sequence based on a new classification of earthquake se-quences. J. Fac. Sci., Hokkaido Univ., Ser. VII (Geophys.), 3: 379–441.

Utsu, T., 1982. Catalog of large earthquakes in the region of Japan from 1885 through 1980. Bull. Earthquake Res. Inst., Univ. Tokyo, 57: 401–463.

Utsu, T. and Seki, A., 1955. Relation between the area of aftershock region and the energy of the main shock. Zisin (J. Seismol. Soc. Jpn.), 2nd Ser., ii, 7: 233–240.