

Performance of a seismicity model based on three parameters for earthquakes ($M \geq 5.0$) in Kanto, central Japan

Masajiro Imoto

National Research Institute for Earth Science and Disaster Prevention, Ibaraki-ken, Japan

Abstract

We constructed a model of earthquakes ($M \geq 5.0$) in Kanto, central Japan, based on three parameters: the a and b values of the Gutenberg-Richter relation, and the v -parameter of changes in mean event size. In our method, two empirical probability densities for each parameter, those associated with target events (conditional density distributions) and those not associated with them (background density distributions), are defined and assumed to have a normal distribution. Therefore, three parameters are transformed by appropriate relations so that new parameters are normally distributed. The retrospective analysis in the learning period and the prospective test of testing period demonstrated that the proposed model performs better by about 0.1 units in terms of the information gain per event than the value summed up with those of the three parameters. The results are confirmed by a simulation with randomly selected model parameters.

Key words *seismicity model – information gain – Gutenberg-Richter relation – Kanto – Japan*

1. Introduction

A number of researchers (Utsu, 1977, 1982; Rhoades and Evison, 1979; Aki, 1981; Hamada, 1983; Grandori, *et al.*, 1988) have formulated expressions of earthquake probabilities based on precursory anomalies detected by multidisciplinary observation. Their formulas assume that different precursory phenomena occur independently of each other. In such a case, the probability expected from detecting

multiple precursory phenomena is given by the product of probability gains for respective observations and the probability estimated from secular seismicity.

Imoto (2006, 2007) extended their formulas to cases in which precursory anomalies are observed as continuous measurements. He further considered the effects originating from mutual correlations between two precursory anomalies, where he assumed two distributions for each precursory parameter: those associated with only space-time volumes in the vicinity of target events (conditional density distributions) and those associated with space-time volumes excluding target events within a short distance (background density distributions). Assuming normal distributions for the conditional and background densities from each discipline, he obtained analytical solutions for the general case in which mutual correlations exist among precursory anomalies in both distributions.

In this paper we attempt to build statistic models for earthquake probability in Kanto, central Japan, based on three parameters: the a -

Mailing address: Dr. Masajiro Imoto, National Research Institute for Earth Science and Disaster Prevention, Tennodai 3-1, Tsukuba-shi, Ibaraki-ken, 305-0006, Japan; tel: +81-29-863-7594; fax: +81-29-863-7876; e-mail: imoto@bosai.go.jp

and b - values of the Gutenberg-Richter relation and a parameter representing changes in mean event size. We evaluate model performance in terms of information gain per event ($IGpe$). We estimate $IGpe$ for the parameters and combinations of the parameters, which are compared with values estimated from data in both the learning period and a testing period.

2. Method

The hazard function is expressed as the expectation of the number of earthquakes in a space-time volume dx (Daley and Vere-Jones, 2003).

We consider unconditional and conditional probabilities of observing a parameter value of θ , which are represented by $g(\theta)d\theta$ (the «background» density) and $f(\theta)d\theta$ (the «conditional» density) and are empirically determined with random samples in the whole study volume and samples conditioned on occurrences of earthquakes.

The hazard function at a space-time point, x , conditioned on a value of θ , is given by

$$h(x|\theta)dx = \frac{f(\theta)m_0}{g(\theta)V_0}dx, \tag{2.1}$$

where m_0 is the number of target events and V_0 is the space-time volume for study.

Taking the Poisson model as the baseline, the information gain per event ($IGpe$: Daley and Vere-Jones, 2003; Imoto, 2004) for a large number of target events is given by

$$IGpe = \int_R f(\theta) \ln \frac{f(\theta)}{g(\theta)} d\theta, \tag{2.2}$$

where the integral is performed within the whole space of θ defined, R .

The above equation represents the fact that $IGpe$ is equivalent to the Kullback-Leibler quantity of information expressing the distance between two probability distributions. Assuming that $f(\theta)$ and $g(\theta)$ are normal distributions of multi-variables, Imoto (2007) derived an analytical form to estimate the $IGpe$ value.

Hereafter referring to Imoto (2007), the main results of previous works will be introduced for the sake of convenience in the later

application of the present study. For a single parameter θ_1 , the $IGpe$ value for θ_1 , $IGpe(\theta_1)$, can be represented as follows:

$$IGpe(\theta_1) = \frac{1}{2} \ln \frac{1}{\sigma_1^2} + \frac{\sigma_1^2 + \mu_1^2 - 1}{2}, \tag{2.3}$$

where, μ_1 and σ_1^2 are the mean and variance of $f(\theta_1)$ and those of $g(\theta_1)$ have been fixed at 0. and 1.

2.1. Correlations in neither the background nor the conditional distributions

Here, we consider n variables $\theta_1, \theta_2, \dots, \theta_n$ possessing joint density distributions $f(\theta_1, \theta_2, \dots, \theta_n)$ and $g(\theta_1, \theta_2, \dots, \theta_n)$ and their marginal distributions of θ_i are noted as $f_i(\theta_i)$ and $g_i(\theta_i)$ for the conditional and background distributions. If variables $\theta_1, \theta_2, \dots, \theta_n$ are mutually independent in both distributions and normally distributed, that is,

$$f(\theta_1, \theta_2, \dots, \theta_n) = \prod_{i=1}^n f_i(\theta_i) = \prod_{i=1}^n N(\mu_i, \sigma_i^2), \tag{2.4}$$

$$g(\theta_1, \theta_2, \dots, \theta_n) = \prod_{i=1}^n g_i(\theta_i) = \prod_{i=1}^n N(0, 1) \tag{2.5}$$

the following equation can be obtained, where $N(\mu_i, \sigma_i^2)$ refers to the normal distribution of the mean μ_i and variance σ_i^2 .

$$IGpe(\theta_1, \theta_2, \dots, \theta_n) = \frac{1}{2} \ln \frac{1}{\sigma_1^2 \sigma_2^2 \dots \sigma_n^2} + \frac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}{2} + \frac{\mu_1^2 + \mu_2^2 + \dots + \mu_n^2}{2} - \frac{n}{2} \tag{2.6}$$

2.2. Correlated conditional distribution

We assume here that the correlation among the n variables $\theta_1, \theta_2, \dots, \theta_n$ can be observed only in the conditional density distribution $f(\theta_1, \theta_2, \dots, \theta_n)$ and that the covariance matrix C can be expressed as

$$f(\theta_1, \theta_2, \dots, \theta_n) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det(C)}} \exp \left\{ -\frac{1}{2} (\theta - \mu)' C^{-1} (\theta - \mu) \right\} \tag{2.7}$$

and

$$[\mathbf{C}] = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ \rho_{ij}\sigma_i\sigma_j & & & \sigma_n^2 \end{pmatrix} \quad (2.8)$$

where the superscript -1 refers to the inverse of a matrix, ρ_{ij} the correlation coefficient between θ_i and θ_j .

By introducing an appropriate transformation of the coordinate system with an orthogonal matrix \mathbf{S} , the covariance matrix can be expressed in a diagonal matrix. At the same time, the vector μ is transformed into μ' with the same matrix.

Referring to the previous case, the *IGpe* value is represented by

$$\begin{aligned} IGpe(\theta_1, \theta_2, \dots, \theta_n) &= \frac{1}{2} \ln \frac{1}{\lambda_1^2 \lambda_2^2 \dots \lambda_n^2} + \\ &+ \frac{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_n^2}{2} + \frac{\mu_1^2 + \mu_2^2 + \dots + \mu_n^2}{2} - \frac{n}{2} = \quad (2.9) \\ &= \frac{1}{2} \ln \frac{1}{\det(\mathbf{C})} + \frac{\text{trace}(\mathbf{C})}{2} + \frac{|\mu|^2}{2} - \frac{n}{2} \end{aligned}$$

where *trace* denotes the sum of the diagonal elements and is an invariant parameter for a unitary transformation and λ_i^2 ($i=1,2,\dots,n$) eigen values.

2.3. Correlations in both distributions

Now we consider a general case in which some correlations among parameters are observed in both distributions. It is assumed that the marginal distribution of each variable θ_i can be expressed in the form of a normal distribution and that correlations among variables are observed in the background distribution.

Let the correlation coefficient between θ_i and θ_j be γ_{ij} , thus the joint density distribution for the background distribution, $g(\theta_1, \theta_2, \dots, \theta_n)$, is given as follows:

$$g(\theta_1, \theta_2, \dots, \theta_n) = \frac{1}{\sqrt{2\pi}^n \sqrt{\det(\mathbf{B})}} \exp\left\{-\frac{1}{2}\theta' \mathbf{B}^{-1} \theta\right\} \quad (2.10)$$

where matrix \mathbf{B} is the covariance matrix (equivalent to the correlation matrix in this case), represented by

$$\mathbf{B} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ \gamma_{ij} & & & 1 \end{pmatrix}, \quad (2.11)$$

By operating successively appropriate transformations of the coordinate system with an orthogonal matrix and a diagonal matrix, the covariance matrix can be expressed in a unit matrix. The compatible transformations are applied to covariance matrix \mathbf{C} , and vector μ . After these transformations, the *IGpe* value can be estimated by eq. (2.9).

Once we have estimated the means and variances of the parameters together with the correlation matrices for both the conditional and background distributions, we can represent the hazard function of the combined model by substituting eqs. (2.7) and (2.10) into eq. (2.1). This function estimates the hazard rate at any point of interest conditioned on the parameter values observed at the respective point.

3. Seismicity model based on three parameters

As an application of the above formula, we attempt to build a seismicity model for moderate and larger earthquakes (target events $M \geq 5.0$) in Kanto, central Japan, based on three parameters: the *a*- and *b*- values in the Gutenberg-Richter relation (GR relation) and a parameter representing temporal change in mean event size, ν - value (Imoto 2003). The latter parameter is obtained from the difference between two mean event sizes, long-term and short-term mean. The long-term mean is defined as a simple mean size of earthquakes ($M \geq 2.0$) within 20km of the point over a period of 960 days prior to the assessment. The short-term mean is calculated in the same way as the long-term one except that an exponentially decaying weight with a time constant of 400 days is used. These space and time windows are reasonably select-

ed with careful consideration of characteristic features of seismic activity in Kanto, and optimization of respective parameters (Imoto, 2003). The ν -value could be considered a sort of b -value short-term variation since the mean event size is inversely proportional to the b -value. In this study, we use the hypocenter parameters for the period from 1980 to 2006 located by the Kanto-Tokai network operated by the National Research Institute for Earth Science and Disaster Prevention (NIED). Taking a balance between the duration of the catalogue and stable estimation, a longer time window, 3650 days is selected for the a - and b - values. The same spatial window and cutoff magnitude as that of the ν - value are used so as to simplify conditions. The cut-off magnitude is selected at 2.0 since the magnitude-frequency relation of the earthquakes in the present study volume exhibits a linear relationship between size and the cumulative number of earthquakes down to a magnitude of 2.0 (Imoto and Yamamoto, 2006).

We surveyed the a -, the b - and the ν - values within the study volume (200 x 200 x 90km³; Imoto and Yamamoto, 2006) from January 1990 to December 1999 for the region. In this specified space-time domain, clustering features of target earthquakes are not observed. The assessment was made at 2km-spaced points and at 10-day intervals. To ensure a reliable distribution, we selected only those points of assessment with both 100 or more earthquakes for 3650 days and 20 and more for 960 days.

We classified the selected points into two categories, points of conditional distribution and points of background. A point of conditional distribution was defined as the grid point of assessment nearest to a target in space and immediately before it in time. This definition of a point belonging to the conditional distribution may be restrictive and other definitions could be explored. Although a study from this viewpoint would be important, it is beyond the scope of the present study. Only the present definition is considered in the study. The other grid points are classified as background. We thus obtained two distributions, the conditional distribution from 33 samples and the background distribution from about 6.6×10^7 , for each of the a -, b - and ν - values. Figures 1a, 1b and 1c illustrate

the empirical background (solid line) and conditional (gray line) distributions for a - b - and ν - values.

We have already proposed two types of seismicity model, one based on the ν - value (Imoto, 2003) and the other on the a - and b - values (Imoto, 2006). The hazard functions, $h_{ab}(x|a,b)$ and $h_\nu(x|\nu)$ are represented as

$$h_{ab}(x|a,b) = a_1 \cdot 10^{a-b_1(b-b_2)^2} \quad (3.1)$$

and

$$h_\nu(x|\nu) = c_0 \mathbf{B}(\nu(x)|c_1, c_2) \quad (3.2)$$

where \mathbf{B} denotes the Beta function, and a_1, b_1, b_2, c_0, c_1 and c_2 are model parameters that have been optimized by the maximum likelihood method in the point process analysis.

The hazard function of eq. (3.1) is not monotonic and takes its maximum at $b=b_1$. Equation (3.2) is similar to this. In transforming the b - and n - parameters into new ones with a normal distribution, we make the hazard function be monotonically increasing with a transformed parameter.

We assumed that the background distributions of the a -, b - and ν - values are respectively represented by an exponential function (dashed line in fig. 1a), a normal function (dashed line in fig. 1b), and a Beta function (dashed line in fig. 1c), which are fitted by the maximum likelihood method. Therefore, a - b - and ν - values are transformed into the a' , b' and ν' values as follows:

$$\int_2^a \frac{1}{a_0 - 2} \exp\left\{-\frac{1}{a_0 - 2}(a - 2)\right\} da = \int_{-\infty}^{a'} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \quad (3.3)$$

$$\int_{-\infty}^{b-} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(b-b_0)^2}{s^2}} db + \int_{b_+}^{\infty} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(b-b_0)^2}{s^2}} = \int_{-\infty}^{b'} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \quad (3.4)$$

where b_0 and s denote the mean and the standard deviation of the b -value in the background

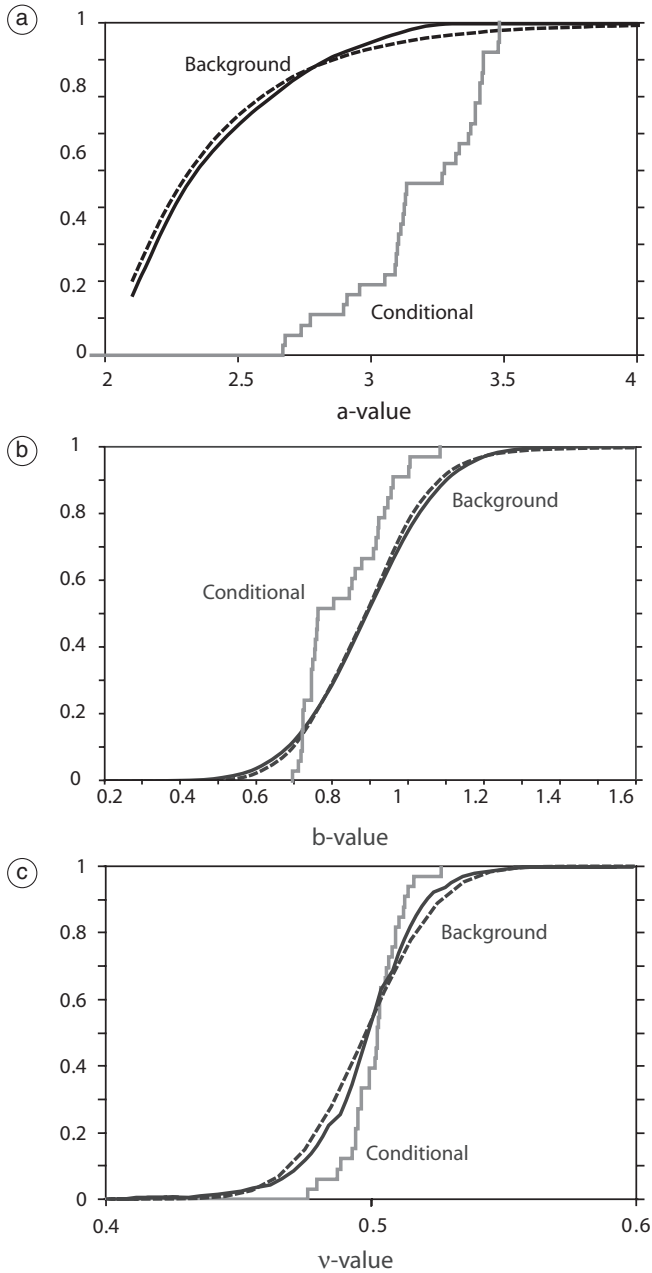


Fig. 1a-c. Cumulative background and conditional distributions for the parameters. (a) Empirical background distributions for the a -value are plotted as a black solid curve, and conditional distributions as a gray step line. An exponential distribution is fitted to the background distribution (dashed curve). (b) Same as fig. 1a but for the b -value. A normal distribution is fitted to the background distribution (dashed curve). (c) Same as fig. 1a but for changes in mean event size. Beta function is fitted to the background distribution (dashed curve).

distribution, and b_+ and b_- are related to the b -value by the relation:

$$(b_- - b_2)^2 = (b_+ - b_2)^2 = (b - b_2)^2, b_+ \geq b_-, \quad (3.5)$$

$$\int_0^{v_-} \mathbf{B}(y | c_3, c_4) dy + \int_{v_+}^1 \mathbf{B}(y | c_3, c_4) dy = \int_{-\infty}^v \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad (3.6)$$

where c_3 and c_4 are the maximum likelihood estimates of the model parameters for the B function fitted to the background distribution of the v - value. The upper limit of the integral of the first term on the left side in eq. (3.6), v_- and the lower limit of the second term, v_+ satisfy the relations as follow:

$$\begin{aligned} \mathbf{B}(v_- | c_1, c_2) &= \mathbf{B}(v_+ | c_1, c_2) = \\ &= \mathbf{B}(v | c_1, c_2), v_+ \geq v_-, \end{aligned} \quad (3.7)$$

Figures 2a, 2b and 2c illustrate the background distribution in black and the conditional distribution in gray for a' , b' and v' . Normal distributions fitted to the conditional and background distributions are indicated with dashed lines in gray and black. These figures indicate that two distributions in each set are more or less appropriately approximated with normal distributions. The parameters of these normal

distributions are summarized in table I. The standard deviations of these estimates for the conditional distributions are given below in parentheses since those of the background distributions are much smaller. This could be justified as follows. The sample size of the background distributions is by far larger than that of the conditional distributions. The transformations have been performed so the background distributions follow a normal distribution with mean 0 and variance 1. Only relative values of the conditional distribution to the background distribution are involved in calculating *IGpe*. Weak correlations among the three parameters are observed for the background distributions. Tables IIa and IIb summarize the correlation matrices in the background and conditional distributions. In the latter table, the standard deviations of the correlation coefficients in the conditional distributions, which are estimated from Fisher transformation (Imoto, 2007), are given in parentheses. It can be concluded that at least the correlation between a' - and v' - values is significant. This suggests that the formula of Aki and others is not applicable to the present case.

4. Results and discussion

The last column in table I indicates *IGpe* for each parameter, where both distributions are as-

Table I. Terms of normal distributions for each parameter and its *IGpe* value. The standard deviations for the estimates in the conditional distributions are given in parentheses.

	Background		Conditional		IGpe
	Av	Std	Av	Std	
a'	0	1	0.964 (0.180)	1.063 (0.136)	0.47
b'	0	1	0.669 (0.172)	0.986 (0.128)	0.22
v'	0	1	0.283 (0.115)	0.649 (0.088)	0.18

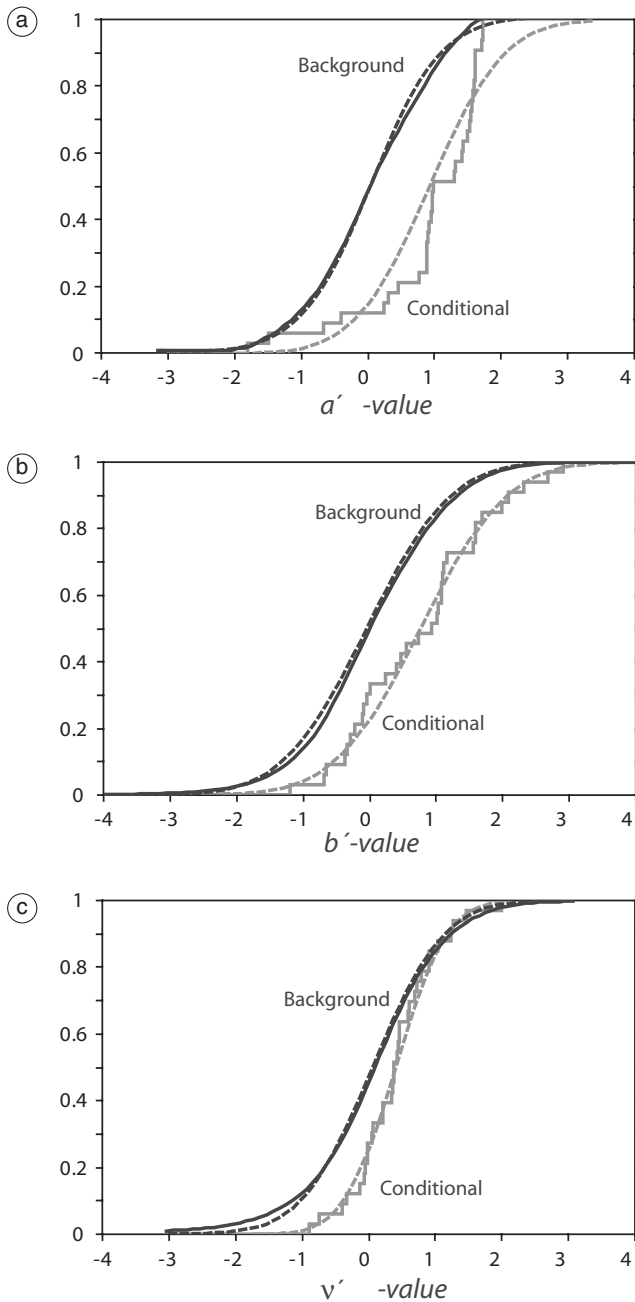


Fig. 2a-c. Cumulative background and conditional distributions for the transformed parameters. (a). Empirical background distribution for the transformed a -value is plotted as a black solid curve, and the conditional distribution as a gray step line. Their fitted normal distributions are plotted as dashed lines in black and gray. (b) Same as fig. 1a but for the b -value. (c) Same as fig. 1a but for changes in mean event size.

Table IIa,b. Correlation matrix (a) Observed in background distribution (b) Observed in conditional distribution. The standard deviations are given in parentheses.

	a'	b'	ν'
a'	1.000	0.049	0.108
b'	0.049	1.000	0.079
ν'	0.108	0.079	1.000

	a'	b'	ν'
a'	1.000	-0.215 (0.225)	-0.402 (0.201)
b'	-0.215 (0.225)	1.000	-0.022 (0.236)
ν'	-0.402 (0.201)	-0.022 (0.236)	1.000

sumed to be normally distributed with the parameters in the respective columns. If the three parameters are not correlated in the two distributions, the resultant *IGpe* value, which is the case of Aki (1981) and others, is given by the summation of each *IGpe* value in the second to the last column. In the actual case, some correlation is observed in both distributions, as indi-

cated in table IIa (background distribution) and table IIb (conditional distribution). Following the procedure given in the previous sections, we can calculate an *IGpe* value expected from a model combining three parameters. The first row of table III lists the *IGpe* values thus obtained.

In order to confirm the validity of *IGpe* estimates, we estimate the probability gain of each target event directly from the hazard function given in eq. (2.1) for each parameter. We can obtain the hazard function of the combined model by substituting eqs. (2.7) and (2.10) into eq. (2.1) with the correlation matrices. Hereafter, the model with the parameters fixed is referred to as the combined model. Using this hazard function, the retrospective analysis has been performed for data from 1990 to 1999 (learning period) and the forward test done for that from 2000 to 2006 (testing period). The results obtained are listed in the second and third rows in Table III. Comparing each *IGpe* value in the second row to the corresponding value in the first row, we can conclude that the parameter estimation and the model construction are consistent. Results obtained from data independent of model construction (testing period) indicate some gaps between expectation and observation. A gain of 0.51 units and a loss of 0.09 units in *IGpe* value from expectation are observed for a'- and b'- values. These gaps result in gains of 0.48 units for the simple sum-

Table III. Summary of the *IGpe* values. Those expected from the formula are listed in the first row. Those calculated as average values of probability gains for each target event are listed in the second row for the learning period and third row for the testing period. Those obtained by simulation are listed in the last row, where the standard deviations of these estimates are given in parentheses. The sum in the third to last column indicates the simple summation of the three *IGpe* values. The *IGpe* value of the combined model is given in the second to last column. The difference between these two values gained by an effect of correlations is given in the last column.

	Events	a'	b'	ν'	Sum_up	Combined	Diff.
Estimation		0.47	0.22	0.18	0.88	0.98	0.10
Learning Period	33	0.43	0.22	0.17	0.82	0.98	0.16
Testing Period	37	0.98	0.13	0.25	1.36	1.41	0.05
Simulation		0.49 (0.17)	0.25 (0.12)	0.21 (0.10)	0.95 (0.23)	1.16 (0.30)	0.21 (0.20)

mation and 0.43 units for our proposed combined model. Even in this case, it is notable that the combined model performs better than simple summation.

To confirm the above results, we estimated each $IGpe$ value and its standard deviation using 1000 sets of randomly generated samples. In generating a set of samples, we only consider the variations of parameter values in the conditional distributions, which are given in tables II and IIb. For each set of samples, we calculate values corresponding to each $IGpe$ in table III, in just the same way as that of the first row. The means and standard deviations are estimated from the 1000 acquired sets of $IGpe$ values and listed. The values in the first row are more or less similar to the respective values in the last row. A relatively large gap is observed in the combined case, the second to the last column, but it is still in the range of the standard deviation. The difference in the $IGpe$ value in the last column suggests that the estimate by the simulation is no smaller than its standard deviation, and means that the $IGpe$ value produced by the combined model is probably larger than that estimated from the formula by Aki and others. It should be noted again that the application of the latter formula is no longer appropriate in the present case with a significant correlation in the conditional distribution between two parameters. Accordingly, it can be concluded that the combined model proposed in the present paper is more appropriate than that formulated by Aki and others, from the viewpoints of its performance and assumption of independency.

5. Conclusions

This paper introduces a way to combine multi-disciplinary observations into one hazard function and demonstrates its superior performance to that expected from the well known formula of Aki and others. The combined model was confirmed using actual data taken from both the learning period and the subsequent testing period. The quantitative matching between the $IGpe$ values predicted by our formula and those observed for the learning period is reasonable and demonstrates only the logical

consistency in our derivations. However, this is not the case for the data taken from the testing period. The combined model performs only 0.05 units better than Aki's formula, but by 0.43 units better than that of the learning period. This gap is primarily because the $IGpe$ value of the a' parameter for the testing period is 0.51 units greater than that observed for the learning period. This corresponds to an increase of about 1.6 times the probability gain. This change could be interpreted by an increase of this factor in the number of sample earthquakes since such increases raise the hazard rate by the same factor through eq. (3.1) with the baseline (the Poisson rate) resuming the same value as before. The average number of earthquakes conditioned on a target event in the testing period exceeds that observed in the learning period by a factor of 1.5. This evidence is consistent with the fact that the seismicity rate of target events in the testing period, 37 events in 7 years, is 1.5 times larger than that in the learning period, 33 events in 10 years.

In summary, we attempted to build up a seismicity model based on three parameters, the a - and b - values of the Gutenberg-Richter magnitude frequency relation and a parameter of changes in mean event size. Applying the formula derived by Imoto (2007) to the data observed by the Kanto-Tokai network (NIED), we estimated that the $IGpe$ value of the model becomes about 0.1 units greater than that of the simple summation formulated by Aki and others. The model performs consistently with this estimation both in the learning period and the testing period.

Acknowledgments

This manuscript was greatly improved by the comments of two anonymous reviewers and editor Massimo Cocco.

REFERENCES

- AKI, K. (1981): A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction*, edited by D.W. SIMPSON and P.G. RICHARDS, 566-574, Agu.
DALEY, D. J. and D. VERE-JONES (2003): An introduction to the

- theory of point processes, vol. 1, *Elementary theory and methods*, Second edition, (Springer, New York), pp. 469.
- GRANDORI, G., E. GUAGENTI and F. PEROTTI (1988): Alarm systems based on a pair of short-term earthquake precursors, *Bull. Seism. Soc. Am.*, **78**, 1538-1549.
- HAMADA, K. (1983): A probability model for earthquake prediction, *Earthquake Prediction Res.*, **2**, 227-234.
- IMOTO, M. (2003): A testable model of earthquake probability based on changes in mean event size, *J. Geophys. Res.*, **108**, ESE 7.1-12 No. B2, 2082, doi:10.1029/2002JB001774.
- IMOTO, M. (2004): Probability gains expected for renewal process models, *Earth Planets Space*, **56**, 563-571.
- IMOTO, M. (2006): Statistical models based on the Gutenberg-Richter a and b values for estimating probabilities of moderate earthquakes in Kanto, Japan, in *Proceedings of The 4th International Workshop on Statistical Seismology, January 9-13, 2006*, ISM Report on Research and Education, ISM, Tokyo, Japan, **23**, 116-119.
- IMOTO, M. (2006): Earthquake probability based on multidisciplinary observations with correlations, *Earth Planets Space*, **57**, 1447-1454.
- IMOTO, M. (2007): Information gain of a model based on multidisciplinary observations with correlations, *J. Geophys. Res.*, **112**, B05306, doi: 10.1029/2006JB004662.
- IMOTO, M. and N. YAMAMOTO (2006): Verification test of the mean event size model for moderate earthquakes in the Kanto region, central Japan, *Tectonophysics*, **417**, 131-140.
- RHOADES, D. and F. EIVISON (1979): Long-range earthquake forecasting based on a single predictor, *Geophys. J. R. astr. Soc.*, **59**, 43-56.
- UTSU, T. (1977): Probabilities in earthquake prediction, *Zisin II*, **30**, 179-185, (in Japanese).
- UTSU, T. (1982): Probabilities in earthquake prediction (the second paper), *Bull. Earthq. Res. Inst.*, **57**, 499-524, (in Japanese).

(received July 6, 2007;
accepted July 2, 2008)