

Performance of a seismicity model for earthquakes in Japan ($M \geq 5.0$) based on P -wave velocity anomalies

Masajiro Imoto

National Research Institute for Earth Science and Disaster Prevention, Tennodai 3-1, Tsukuba-shi, Ibaraki-ken 305-0006, Japan

(Received March 8, 2010; Revised May 13, 2010; Accepted June 18, 2010; Online published March 4, 2011)

We consider P -wave perturbations from a standard layered model for Japan, as a predictive parameter that may be useful for assessing regional seismogenesis. To assess the performance of a seismicity model with predictive parameters, we used the Kullback-Leibler statistic in terms of information gain per event (IGpe), which is the distance between two distributions of parameters, the background distribution (parameters over the entire space domain), and the conditional distribution (parameters at earthquake epicenters). We selected 198 epicenters of earthquakes with magnitudes ≥ 5.0 that occurred between 1961 and 2008 to estimate the conditional distribution. More than 3,000 points were selected at every point on a $0.1 \times 0.1^\circ$ grid for the background distribution. P -wave variations were considered at four different depths (10, 15, 20, and 25 km at each point) for both distributions. We compared the two distributions at each depth but found no significant difference in the average value of perturbations between them. As these distributions are well-approximated by normal distributions, IGpe can be estimated directly from the means and standard deviations of both distributions at each depth. We obtained an IGpe of ≤ 0.03 using a single parameter. However, when multiple parameters with correlations were considered, an IGpe of 0.3 was estimated, which means that the average probability across the 198 earthquakes is 1.35-fold higher than that of a Poisson process model.

Key words: Seismicity model, information gain, Kullback-Leibler quantity, P -wave velocity structure, predictive parameters, correlation, Japan.

1. Introduction

Seismicity models provide some of the most useful products in earthquake prediction research. The incorporation of various predictive parameters could result in better performing models. Utsu (1977, 1982) and many others (Rhoades and Evison, 1979; Aki, 1981; Hamada, 1983; Grandori *et al.*, 1988) have formulated expressions for earthquake probabilities based on precursory anomalies from a variety of measurements. Imoto (2006, 2007) proposed a method to build models based on multiple predictive parameters in which independence among parameters is not necessarily assumed as it has been in previous studies. His result implies that mutual correlations among predictive parameters for certain conditions could produce a better performance than expected for those cases in which the parameters are independent.

With the development of dense seismic networks and computational power, seismic wave velocity structures have been modeled to higher resolutions than has previously been possible. Seismogenesis must be closely related to the physical properties (e.g., pressure, temperature, and properties of geology) of focal areas. Of these, the P -wave velocity is more generally and systematically sampled than any other parameter. Many issues are involved in the rela-

tionship between seismic wave velocity perturbations and seismicities, and some of these have been discussed in only limited terms (i.e., AL-Shukri and Mitchell, 1988; Michael, 1988; Kaufmann and Long, 1996; Hauksson and Haase, 1997).

In the 1970s, a large body of literature was published on changes in seismic velocity before earthquakes. A change in P -wave velocity was interpreted using the dilatancy theory (hypothesis) that rocks underwent dilatation in the last stage before failure (Nur, 1972; Scholz *et al.*, 1973). To date, systematic and homogeneous measurements detecting such variations have not been obtained. Consequently, in the study reported here, variations in seismic velocity over time are not discussed.

After the Hi-net seismic network was established in Japan (Obara *et al.*, 2005), Matsubara *et al.* (2008) revealed fine structures of P - and S -wave velocities in Japan. Some of their remarkable findings are as follows. The high-velocity Pacific plate and Philippine Sea plate are clearly imaged to the depth of 150 km beneath northeastern and the southwestern Japan, respectively. High- V_p/V_s (P -wave velocity by S -wave velocity) zones are widely distributed beneath the volcanic front where seismic swarm activities, including moderately sized earthquakes, have often been observed. Non-volcanic tremors occur in the high- V_p/V_s zone at depths of 30–40 km beneath southwestern Japan where the oceanic crust of the Philippine Sea plate encounters the wedge mantle of the Eurasian plate.

Matsubara and Obara (2008) reported characteristic fea-

tures of the perturbations in zones beneath active faults, and this information may be incorporated into the construction of a seismicity model that performs better than any previous model. However, before building such a model, it is necessary to appropriately evaluate the information of each contributing parameter.

In the study reported here, we evaluate model performance in terms of information gain per event (IGpe; Daley and Vere-Jones, 2003; Imoto, 2004) in order to incorporate information on the *P*-wave velocity structure into current seismicity models. The results present a good example of cases in which correlations among predictive parameters increase predictive power more than those without correlations.

2. Method

The seismic hazard function is expressed as the expectation of the number of earthquakes in a space-time volume $d\mathbf{x}$ above some threshold magnitude (Daley and Vere-Jones, 2003). We consider both the unconditional and conditional probabilities of observing a potentially predictive parameter value of θ , which are represented by $g(\theta)d\theta$ (background density) and $f(\theta)d\theta$ (conditional density) and which are empirically determined with random samples of cells in the whole study volume and samples conditioned on occurrences of earthquakes in some cells. The hazard function at a space-time point (x), conditioned on a value of $\theta(x)$, is given by

$$h(\mathbf{x}|\theta)d\mathbf{x} = \frac{f(\theta(x))m_0}{g(\theta(x))V_0}d\mathbf{x}, \quad (1)$$

where m_0 is the number of earthquakes above the threshold, and V_0 is the space-time volume being studied (see Appendix).

Taking the Poisson model as the baseline, the IGpe (Daley and Vere-Jones, 2003; Imoto, 2004, 2007) for a large number of earthquakes is given by

$$\text{IGpe} = \int_R f(\theta) \ln \frac{f(\theta)}{g(\theta)} d\theta, \quad (2)$$

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 multivariate distributions, Imoto (2007) derived an analytical equation to estimate the IGpe value.

For the sake of convenience, the main results of the previous studies (Imoto, 2007) will be summarized below. For a single parameter θ_1 , the IGpe(θ_1) can be represented as

$$\text{IGpe}(\theta_1) = \frac{1}{2} \ln \frac{1}{\sigma_1^2} + \frac{\sigma_1^2 + \mu_1^2 - 1}{2}, \quad (3)$$

where μ_1 is the mean, and σ_1^2 is the variance of $f(\theta_1)$, and those of $g(\theta_1)$ are scaled to be 0 (mean) and 1 (variance).

Next, we consider n variables $\theta_1, \theta_2, \dots, \theta_n$ as 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 $f_i(\theta_i)$ for the conditional distribution and $g_i(\theta_i)$ for the background distribution. If variables $\theta_1, \theta_2, \dots, \theta_n$ are

mutually independent in both distributions and are normally distributed with the mean μ_i , and variance σ_i^2 for the conditional distribution and 0 and 1 for the background distribution, the IGpe can be represented as follows.

$$\begin{aligned} \text{IGpe}(\theta_1, \theta_2, \dots, \theta_n) = & \frac{1}{2} \ln \frac{1}{\sigma_1^2 \sigma_2^2 \dots \sigma_n^2} + \frac{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2}{2} \\ & + \frac{\mu_1^2 + \mu_2^2 + \dots + \mu_n^2}{2} - \frac{n}{2} \end{aligned} \quad (4)$$

We assume here that the correlation among the n variables $\theta_1, \theta_2, \dots, \theta_n$ only occurs in the conditional density distribution $f(\theta_1, \theta_2, \dots, \theta_n)$:

$$\begin{aligned} f(\theta_1, \theta_2, \dots, \theta_n) = & \frac{1}{\sqrt{2\pi^n} \sqrt{\det(\mathbf{C})}} \\ & \cdot \exp \left\{ -\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\mu})^t \mathbf{C}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}, \end{aligned} \quad (5)$$

where the superscript -1 refers to the inverse of a matrix, and the covariance matrix \mathbf{C} can be expressed as

$$[\mathbf{C}] = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ \rho_{ij}\sigma_i\sigma_j & & & \sigma_n^2 \end{pmatrix}, \quad (6)$$

where ρ_{ij} is the correlation coefficient between θ_i and θ_j .

By introducing an appropriate transformation of the $(\boldsymbol{\theta})$ coordinate system with an orthogonal matrix, the covariance matrix can be expressed as a diagonal matrix. At the same time, the vector $\boldsymbol{\mu}$ is transformed into $\boldsymbol{\mu}'$ with the same orthogonal matrix. Referring to the previous case, the IGpe is represented by

$$\begin{aligned} \text{IGpe}(\theta_1, \theta_2, \dots, \theta_n) = & \frac{1}{2} \ln \frac{1}{\lambda_1^2 \lambda_2^2 \dots \lambda_n^2} + \frac{\lambda_1^2 + \lambda_2^2 + \dots + \lambda_n^2}{2} \\ & + \frac{\mu_1'^2 + \mu_2'^2 + \dots + \mu_n'^2}{2} - \frac{n}{2} \\ = & \frac{1}{2} \ln \frac{1}{\det(\mathbf{C})} + \frac{\text{trace}(\mathbf{C})}{2} + \frac{|\boldsymbol{\mu}'|^2}{2} - \frac{n}{2}, \end{aligned} \quad (7)$$

where trace denotes the sum of the diagonal elements and is an invariant parameter for a unitary transformation, and λ_i^2 ($i = 1, 2, \dots, n$) represent the eigenvalues of \mathbf{C} . Comparing Eqs. (7) and (4), the first term in the right side of Eq. (7) exceeds that of Eq. (4) unless every ρ_{ij} is zero. The other three terms have the same values in both equations. Therefore, the IGpe for a conditional distribution of correlated variables always exceeds that with no correlation.

In general, some correlations among parameters may be observed in both distributions. The procedure from Eq. (5) to Eq. (7) could be applied after the covariance matrix for the background distribution is changed into the identity matrix by transformations of the coordinate system with an orthogonal matrix and a diagonal matrix.

Table 1. List of target earthquakes used for the conditional distribution.

YMD	HH	MM	Lat	Long	Depth	Mag	Gain	Ln (Gain)
1961/2/2	3	39	37.45	138.84	0	5.2	1.59	0.47
1961/3/16	7	16	32.02	130.69	0	5.3	1.66	0.51
1961/3/18	15	22	31.99	130.77	3.1	5.1	1.69	0.53
1961/5/7	21	14	35.05	134.51	23	5.9	1.85	0.62
1961/8/19	14	33	36.11	136.7	10	7	0.70	-0.36
1961/8/19	22	24	36.54	137.68	0	5	1.87	0.63
1962/3/13	15	7	37.07	139.13	3	5.1	2.44	0.89
1962/4/30	11	26	38.74	141.14	19	6.5	1.31	0.27
1962/12/10	6	16	39.8	140.87	16	5	0.66	-0.42
1963/1/28	13	5	43.59	144.71	10	5.3	1.52	0.42
1963/2/9	12	53	36.35	137.69	0	5.4	1.86	0.62
1963/3/27	6	34	35.81	135.79	13.1	6.9	1.57	0.45
1963/3/31	21	26	35.13	132.42	12	5.1	1.55	0.44
1963/4/21	22	2	35.26	137.61	4.1	5	1.41	0.35
1963/7/24	20	50	35.84	137.01	7	5	1.61	0.48
1963/7/30	17	27	33.86	135	8.1	5.2	0.78	-0.25
1963/8/11	16	37	38.72	141.15	8	5.2	1.17	0.15
1964/6/16	13	17	38.65	139.52	0	6.1	1.71	0.54
1964/6/16	15	52	38.02	139.01	15.1	5.6	0.57	-0.56
1964/11/3	20	9	34.74	138.84	0	5.3	0.93	-0.07
1965/1/13	17	40	38.72	141.13	0	5.2	1.19	0.18
1965/2/26	15	42	35.27	132.73	20	5.1	1.41	0.35
1965/4/20	8	41	34.88	138.3	20	6.1	3.02	1.11
1965/8/31	16	48	43.48	144.43	0	5.1	0.93	-0.07
1965/9/9	13	39	43.47	144.3	0	5.1	1.55	0.44
1965/11/23	2	57	36.52	138.23	0	5	2.07	0.73
1966/1/9	7	39	37.15	138.55	0	5.2	0.42	-0.86
1966/1/23	20	15	36.52	138.22	0	5.1	2.10	0.74
1966/4/5	17	51	36.58	138.32	0	5.4	2.09	0.74
1966/5/6	19	8	36.52	138.25	0	5	2.06	0.72
1966/5/26	7	49	35.35	136.5	20	5.1	1.75	0.56
1966/5/28	14	21	36.57	138.22	0	5.3	2.84	1.04
1966/6/12	9	43	36.53	138.32	0	5	2.04	0.71
1966/6/26	16	34	36.55	138.35	0	5	1.77	0.57
1966/8/3	3	48	36.47	138.2	0	5.3	2.02	0.7
1966/8/8	9	37	36.53	138.32	0	5.1	2.04	0.71
1966/8/28	13	9	36.47	138.13	0	5.3	1.43	0.36
1966/8/29	0	36	36.57	138.25	0	5.1	2.52	0.92
1966/9/14	10	14	36.57	138.25	0	5	2.52	0.92
1966/10/26	3	4	36.55	138.37	0	5.3	1.58	0.46
1966/11/12	21	1	33.07	130.27	20	5.5	1.54	0.43
1966/12/5	16	23	32.33	131.82	0	5	1.11	0.11
1967/1/16	12	32	36.48	138	0	5.2	2.32	0.84
1967/2/3	17	17	36.43	138.07	0	5.4	1.40	0.34
1967/3/2	3	39	36.5	138.3	0	5.1	1.87	0.62
1967/5/5	8	25	36.4	138.05	10	5.2	1.15	0.14
1967/9/14	19	38	36.43	138.15	10	5.1	1.58	0.46
1967/10/14	4	48	36.53	138.2	10	5.3	2.41	0.88
1967/11/4	23	30	43.48	144.27	20	6.5	1.53	0.42
1968/1/26	16	55	36.52	138.15	0	5.3	1.71	0.53
1968/2/21	10	44	32.02	130.72	0	6.1	1.78	0.57
1968/3/25	0	58	32.02	130.72	0	5.7	1.78	0.57
1968/3/30	4	4	34.17	135.17	0	5	1.98	0.68
1968/4/4	19	54	36.57	138.18	0	5.1	2.78	1.02
1968/6/15	11	14	37.15	138.67	10	5	0.66	-0.42
1968/8/18	16	12	35.22	135.38	0	5.6	1.58	0.46
1968/9/21	7	25	36.82	138.27	10	5.3	0.42	-0.88
1969/9/2	21	7	36.2	137.72	0	5	1.68	0.52
1969/9/9	14	15	35.78	137.07	0	6.6	3.03	1.11

Table 1. (continued).

YMD	HH	MM	Lat	Long	Depth	Mag	Gain	Ln (Gain)
1970/2/27	19	13	36.97	137.65	0	5.1	0.90	-0.11
1970/4/9	1	43	36.43	138.1	0	5	1.11	0.1
1970/10/16	14	26	39.2	140.75	0	6.2	2.14	0.76
1971/2/26	4	27	37.13	138.35	0	5.5	2.00	0.69
1971/7/23	7	7	35.55	138.97	10	5.3	1.00	0
1972/4/14	4	29	34.9	132.93	10	5.2	0.64	-0.45
1972/8/20	19	9	38.6	139.95	20	5.3	1.69	0.52
1972/8/31	16	54	35.28	135.62	10	5.1	1.65	0.5
1972/8/31	17	7	35.88	136.77	10	6	2.45	0.9
1972/9/6	20	42	32.75	130.43	10	5.2	1.13	0.12
1973/2/25	19	9	34.73	132.42	0	5	1.01	0.01
1973/9/21	11	21	35.1	134.52	10	5.1	2.53	0.93
1973/10/27	13	44	35.2	133.27	10	5.1	2.99	1.1
1974/5/9	8	33	34.57	138.8	10	6.9	0.81	-0.21
1975/1/23	23	19	33	131.13	0	6.1	1.83	0.6
1975/4/21	2	35	33.13	131.33	0	6.4	1.89	0.64
1976/6/16	7	36	35.5	139	20	5.5	2.04	0.72
1976/8/18	2	18	34.78	138.95	0	5.4	0.33	-1.11
1977/5/2	1	23	35.15	132.7	10	5.6	1.17	0.16
1977/6/28	11	46	32.9	130.72	10	5.3	1.54	0.43
1978/1/15	7	31	34.83	138.88	20	5.8	0.06	-2.73
1978/6/4	5	3	35.08	132.7	0	6.1	1.07	0.07
1978/10/7	5	44	35.78	137.5	0	5.4	5.29	1.67
1978/11/23	10	43	34.77	139.02	0	5.1	1.19	0.17
1978/12/3	22	15	34.88	139.18	20	5.5	2.31	0.84
1979/3/17	12	26	31.95	130.57	20	5	1.41	0.35
1979/8/17	15	0	38.12	139.25	20	5	0.90	-0.11
1979/10/16	7	45	35.28	135.88	10	5	2.33	0.85
1979/12/28	23	54	34.92	134.37	20	5	1.31	0.27
1980/1/25	20	11	38.58	141.7	0	5.1	1.38	0.32
1980/6/29	16	20	34.92	139.23	10	6.7	1.52	0.42
1980/7/31	10	13	38.7	141.2	0	5	1.09	0.09
1981/1/30	6	19	32.43	129.85	0	5.3	1.18	0.17
1981/5/19	0	31	37.08	137.7	0	5	1.39	0.33
1982/1/8	5	37	40.02	140.48	0	5.2	1.51	0.41
1983/3/6	6	32	35.69	136.02	8	5	1.03	0.03
1983/8/8	12	47	35.52	139.02	22	6	1.63	0.49
1983/10/16	19	39	37.14	137.97	15	5.3	0.45	-0.79
1983/10/31	1	51	35.42	133.92	15	6.2	1.07	0.06
1984/2/14	1	53	35.59	139.1	25	5.4	1.25	0.22
1984/5/30	9	39	34.96	134.59	17	5.6	1.10	0.09
1984/6/25	6	29	34.76	132.58	12	5.3	1.29	0.26
1984/8/6	17	30	32.76	130.18	7	5.7	1.92	0.65
1984/9/14	8	48	35.83	137.56	2	6.8	1.23	0.2
1984/10/3	9	12	35.83	137.62	5	5.4	0.07	-2.69
1985/2/26	19	53	35.84	137.58	7.1	5.2	0.42	-0.88
1985/3/28	16	13	38.88	140.73	5.1	5.3	1.12	0.12
1985/7/2	13	20	35.38	133.61	13.1	5.1	1.60	0.47
1985/10/3	20	57	35.18	135.86	8	5.3	0.99	-0.01
1985/11/27	9	1	35.62	135.75	11	5.2	1.43	0.36
1986/3/7	3	25	36.03	137.5	4	5.1	2.83	1.04
1986/5/26	11	59	40.08	141.2	10	5	1.24	0.22
1986/6/27	20	18	39.04	140.95	11.1	5	1.39	0.33
1986/7/28	9	43	32.47	130.48	13	5.1	2.31	0.84
1986/12/30	9	38	36.64	137.92	3	5.9	1.03	0.03
1987/5/9	12	54	34.15	135.4	8	5.6	2.08	0.73
1987/5/28	6	3	35	135.53	16.1	5	2.02	0.7
1987/11/18	0	57	34.24	131.46	8	5.4	1.27	0.24
1988/7/31	8	40	34.97	139.21	5	5.2	1.00	0

Table 1. (continued).

YMD	HH	MM	Lat	Long	Depth	Mag	Gain	Ln (Gain)
1988/9/29	17	23	35.92	139.19	15	5.1	0.25	-1.39
1989/7/9	11	9	34.99	139.11	3	5.5	1.70	0.53
1989/10/27	7	41	35.26	133.37	13	5.3	3.99	1.38
1989/11/2	4	57	35.26	133.37	14.1	5.5	4.01	1.39
1990/1/11	20	10	35.11	135.98	11	5	1.34	0.29
1990/2/20	15	53	34.76	139.23	5.1	6.5	1.66	0.51
1990/8/5	16	13	35.21	139.09	13.1	5.3	1.61	0.47
1990/9/29	7	57	34.99	134.29	11	5.4	1.56	0.44
1990/11/11	4	5	44.28	142.13	0	5	1.50	0.41
1990/11/21	10	44	35.28	133.34	13.1	5.1	2.82	1.04
1990/11/23	19	33	35.27	133.36	14	5.2	3.61	1.28
1990/12/1	20	23	35.28	133.34	12	5.1	2.70	0.99
1990/12/7	18	38	37.21	138.56	14.1	5.4	0.61	-0.49
1991/8/28	10	29	35.32	133.19	13	5.9	0.74	-0.3
1991/10/28	10	9	33.92	131.16	18	6	1.17	0.16
1992/5/20	12	29	36.12	135.95	9.1	5	1.63	0.49
1993/4/2	19	5	37.59	137.45	5	5	1.66	0.51
1993/4/23	5	18	35.81	137.5	7.1	5.1	3.87	1.35
1993/5/7	4	57	37.62	137.4	20	5	1.57	0.45
1993/7/20	0	10	36.35	137.65	4	5	1.78	0.58
1994/2/13	2	6	32.08	130.49	5	5.7	1.28	0.25
1994/5/8	17	2	34.08	135.12	10.1	5	0.75	-0.29
1994/9/7	12	54	32.01	131.15	19	5.3	4.67	1.54
1994/12/18	20	7	37.29	139.89	6	5.5	2.13	0.75
1995/1/17	5	46	34.6	135.04	16.1	7.3	1.12	0.11
1995/1/17	7	38	34.79	135.44	12	5.4	1.47	0.39
1995/1/25	23	15	34.79	135.3	14.8	5.1	0.78	-0.25
1995/2/18	21	37	34.44	134.81	15.9	5	1.67	0.51
1995/3/17	0	8	35.74	137.56	10.4	5.3	1.29	0.26
1995/4/1	12	49	37.89	139.25	16.2	5.6	0.48	-0.74
1995/10/1	20	48	34.95	139.15	0	5	1.50	0.41
1996/2/7	10	33	35.94	136.62	12.1	5.3	1.23	0.21
1996/3/6	23	35	35.47	138.95	19.6	5.5	2.67	0.98
1996/8/11	3	12	38.91	140.63	8.6	6.1	0.63	-0.47
1997/3/4	12	51	34.96	139.17	2.6	5.9	1.39	0.33
1997/3/26	17	31	31.97	130.36	11.9	6.6	1.30	0.26
1997/4/3	4	33	31.97	130.32	14.8	5.7	1.54	0.43
1997/5/13	14	38	31.95	130.3	9.2	6.4	1.40	0.33
1997/5/24	2	50	34.5	137.5	23.1	6	0.45	-0.8
1997/6/25	18	50	34.44	131.66	8.3	6.6	1.62	0.48
1997/9/4	5	15	35.26	133.38	8.9	5.5	4.01	1.39
1998/2/21	9	55	37.27	138.79	19.1	5.2	1.53	0.43
1998/4/22	20	32	35.17	136.56	7.8	5.5	1.42	0.35
1998/4/26	7	37	34.96	139.17	6.1	5	1.28	0.25
1998/5/3	11	9	34.96	139.18	4.7	5.9	1.32	0.27
1998/7/1	2	22	36.62	137.91	8.9	5	1.20	0.19
1998/8/3	20	9	37.21	139.99	7.6	5.2	0.38	-0.97
1998/8/12	15	13	36.24	137.63	2.8	5	2.52	0.93
1998/8/16	3	31	36.33	137.62	3.2	5.6	1.83	0.6
1998/9/3	16	58	39.81	140.9	7.9	6.2	0.98	-0.02
1998/9/15	16	24	38.28	140.76	13.2	5.2	0.80	-0.22
1999/2/26	14	18	39.15	139.84	20.6	5.3	0.30	-1.21
1999/3/16	16	43	35.28	135.93	11.6	5.2	2.05	0.72
1999/11/7	3	34	36.06	135.79	15	5	1.44	0.37
2000/6/8	9	32	32.69	130.76	10.3	5	1.61	0.47
2000/10/6	13	30	35.27	133.35	9	7.3	3.18	1.16
2001/1/4	13	18	36.96	138.77	11.2	5.3	2.35	0.85
2001/1/12	8	0	35.47	134.49	10.6	5.6	1.32	0.28
2001/3/31	6	9	36.82	139.38	4.7	5.2	2.27	0.82

Table 1. (continued).

YMD	HH	MM	Lat	Long	Depth	Mag	Gain	Ln (Gain)
2001/8/25	22	21	35.15	135.66	8.2	5.4	2.70	0.99
2002/9/16	10	10	35.37	133.74	9.6	5.5	1.82	0.6
2003/7/26	7	13	38.4	141.17	11.9	6.4	0.00	-6.24
2004/10/23	17	56	37.29	138.87	13.1	6.8	1.25	0.22
2004/10/27	10	40	37.29	139.03	11.6	6.1	3.23	1.17
2004/11/1	4	35	37.21	138.9	8.5	5	2.22	0.8
2004/11/4	8	57	37.43	138.91	18	5.2	1.14	0.13
2004/11/8	11	15	37.4	139.03	0	5.9	1.04	0.04
2004/12/14	14	56	44.08	141.7	8.6	6.1	4.01	1.39
2004/12/28	18	30	37.32	138.98	8	5	1.56	0.44
2005/3/20	10	53	33.74	130.18	9.2	7	2.03	0.71
2005/4/10	20	34	33.67	130.28	4.7	5	2.41	0.88
2005/4/20	6	11	33.68	130.29	13.5	5.8	2.74	1.01
2005/5/2	1	23	33.67	130.32	11.4	5	2.99	1.1
2005/6/20	13	3	37.23	138.59	14.5	5	0.60	-0.51
2005/8/21	11	29	37.3	138.71	16.7	5	0.91	-0.09
2006/4/21	2	50	34.94	139.2	7.1	5.8	1.63	0.49
2007/4/15	12	19	34.79	136.41	16	5.4	0.63	-0.45
2007/7/16	10	13	37.56	138.61	16.8	6.8	1.25	0.22
2007/8/18	16	55	35.34	140.35	20.2	5.2	0.13	-2
2008/6/14	8	43	39.03	140.88	7.8	7.2	2.36	0.86

Once we have estimated the means and variances of the parameters together with the correlation matrices for both the conditional and the background distributions, we can represent them by $f(\theta)$ and $g(\theta)$ and thus calculate the hazard function of Eq. (1). This function estimates the hazard rate at any point of interest conditioned on the parameter values observed at that point.

3. Data

We consider a seismicity model for earthquakes $M \geq 5.0$ in Japan based on *P*-wave velocity perturbation data. We use the hypocenter parameters for 1961–2008 determined by the Japan Meteorological Agency (JMA). This period is selected to balance both the number of earthquakes and the accuracy of estimated locations. In terms of the complex tectonic setting in and around Japan, we restricted ourselves to earthquakes shallower than 30 km.

Matsubara *et al.* (2008) constructed three-dimensional *P*- and *S*-wave velocity models beneath all of Japan at depths of 0–40 km, with a 0.2° grid spacing in the horizontal direction and a 5- to 10-km spacing in the vertical direction. They also constructed a velocity model down to the depth of 400 km with less densely spaced grids. In general, velocity variations from a standard velocity model are estimated since *P*- and *S*-wave velocities strongly depend on depth. Therefore, comparing variations at the same depth may be useful for obtaining characteristic seismogenic features of a focal area.

Accordingly, we consider a two-dimensional seismicity model in which hazard rates at horizontal spacing grids are defined. *P*-wave velocity differences at four different depths (10, 15, 20, and 25 km) for each point are used. We consider that a set of these four parameters plays important roles as predictive parameters. More than 3,000 points with reliable velocity anomalies are selected at every point

of a $0.1 \times 0.1^\circ$ grid for the background distributions, which mostly cover inland parts of Japan, with the exception of Hokkaido Island. To estimate the conditional distributions, we select 198 epicenters of earthquakes (Table 1) with magnitudes ≥ 5.0 that occurred between 1961 and 2008.

4. Information Gain per Event

Figure 1 illustrates the empirical background distributions (dark solid line) for the four parameters and the normal functions fitted to them (light dashed line). Each background distribution is generally well-approximated by a normal function. In the same way, Fig. 2 illustrates the conditional background distributions (dark solid line) and the normal functions fitted to them (light dashed line). For the conditional distributions at 20 and 25 km, the normal approximation is not a close fit.

The chi square-test for goodness-of-fit was performed within the framework of the null hypothesis that *P*-wave velocity differences at each depth possess a normal distribution. The hypothesis for samples at either 10 or 15 km is accepted at the 10% level of significance. The hypothesis for samples at either 20 or at 25 km is accepted at the 1% level of significance, which may appear higher than usual but is assessed to be adequate for fitting with a function of two parameters.

The parameters of these normal distributions are summarized in Table 2. The last column of the table indicates the IGpe for each predictive parameter, calculated using Eq. (3), where both distributions are assumed to be normally distributed. It is obvious that no large differences exist between conditional and background distributions. If we use the predictive parameter separately, an IGpe of 0.03 at most is expected for the parameter measured at a depth of 25 km.

However, correlations among the four parameters are ob-

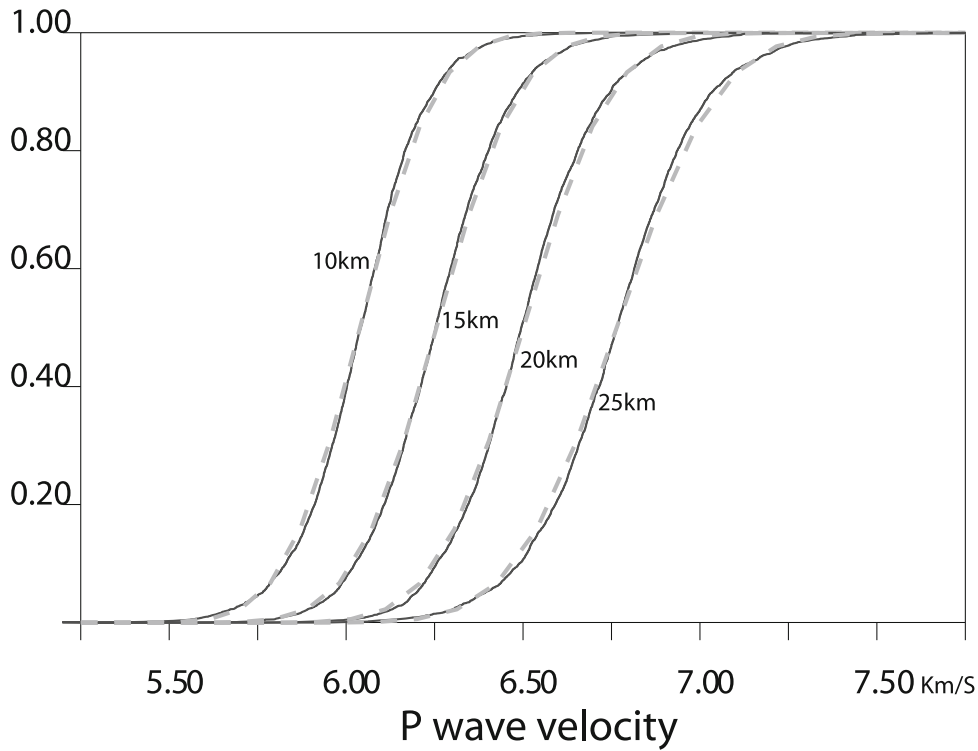


Fig. 1. Cumulative background distributions for the four parameters. Empirical background distributions (dark line) and normal distributions (light dashed line) fitted to the background distributions.

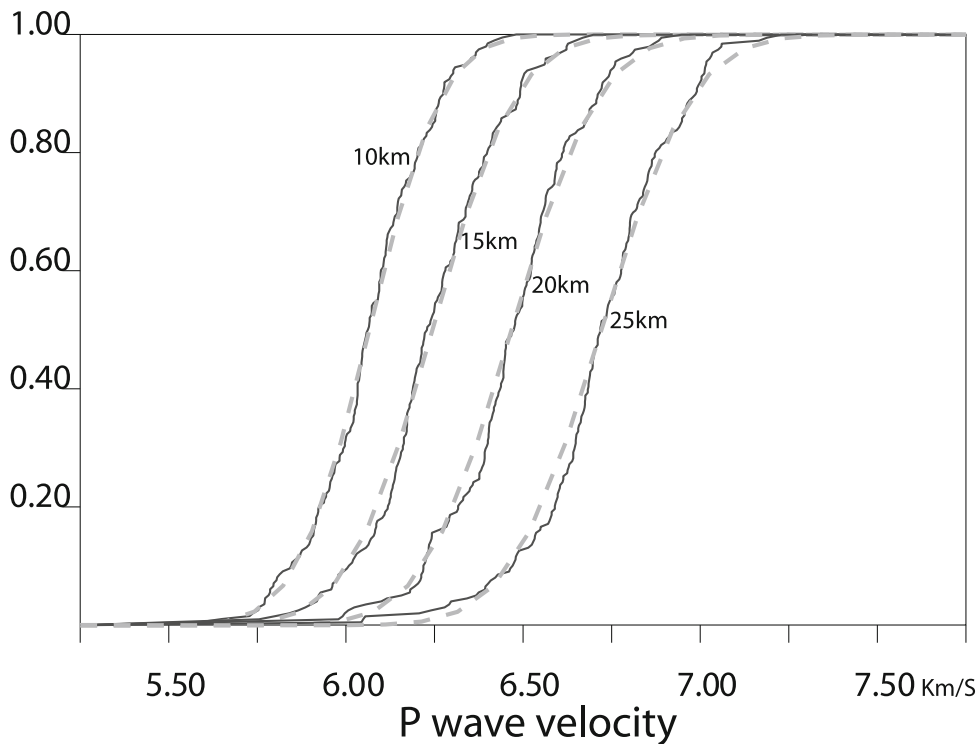


Fig. 2. Cumulative conditional distributions for the four parameters. Empirical conditional distributions (dark line) and normal distributions (light dashed line) fitted to the conditional distributions.

served for both distributions. Table 3 summarizes the correlation matrices in the background (lower left) and conditional (upper right) distributions. The coefficient in the conditional distribution always exceeds the corresponding one in the background distribution. Specifically, for the corre-

lation between parameters at depths of 10 and 25 km, the coefficient of 0.498 in the conditional distribution is larger than that of 0.103 in the background distribution. These features of the correlation matrices suggest a better predictive power with correlated parameters than that expected from a

Table 2. Terms of normal distributions for each parameter and its IGpe value.

	Background		Conditional		IGpe
	Av	Std	Av	Std	
10 km	6.039	0.169	6.066	0.163	0.014
15 km	6.254	0.183	6.243	0.187	0.002
20 km	6.500	0.195	6.467	0.196	0.015
25 km	6.764	0.225	6.720	0.204	0.028

Table 3. Correlation matrices. Lower left: Observed in background distribution. Upper right: Observed in conditional distribution.

	10 km	15 km	20 km	25 km
10 km	—	0.877	0.679	0.498
15 km	0.829	—	0.887	0.681
20 km	0.451	0.817	—	0.901
25 km	0.103	0.437	0.808	—

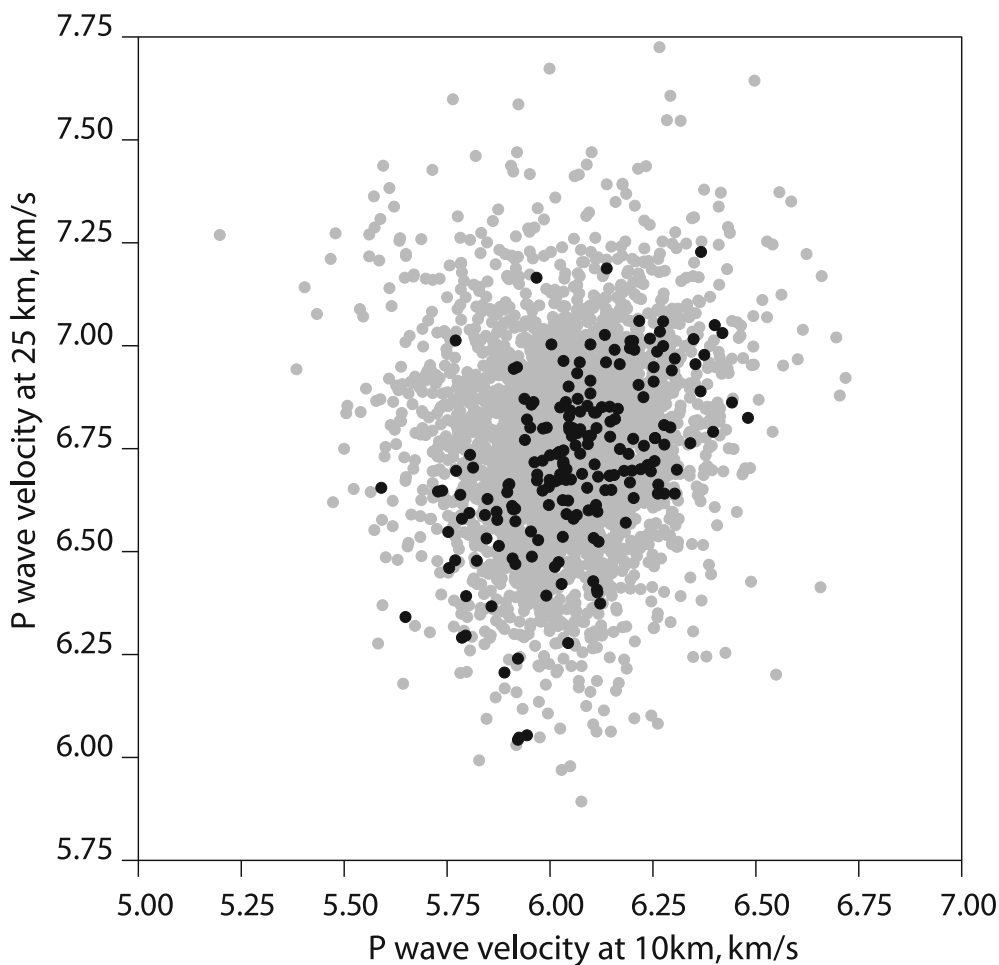


Fig. 3. Joint distributions of the predictive parameters at depths of 10 and 25 km. The ordinate represents *P*-wave velocity at a depth of 25 km, and the abscissa represents *P*-wave velocity at a depth of 10 km. Dark symbols indicate plots of the conditional distributions, and light symbols indicate those of background distributions.

single parameter, as indicated by Eq. (7). Given the values in Tables 2 and 3, we can construct the background and conditional densities. The hazard function is obtained by multiplying $f(\theta)/g(\theta)$ by an average rate (Poisson rate). Using the formula developed by Imoto (2007), we can estimate an IGpe of 0.30 for the seismicity model of this hazard function. This value is equivalent to a probability gain of 1.35 across all target earthquakes.

Figure 3 plots joint distributions between the parameters at 10 and 25 km for both cases. The plot of the conditional distribution (dark circles) is more concentrated than that of the background distribution (light circles). The plot of the conditional distribution is more or less located in a lower right part of the background distribution, which corresponds to the evidence showing that velocity becomes higher beneath an epicenter than the average velocity at a depth of

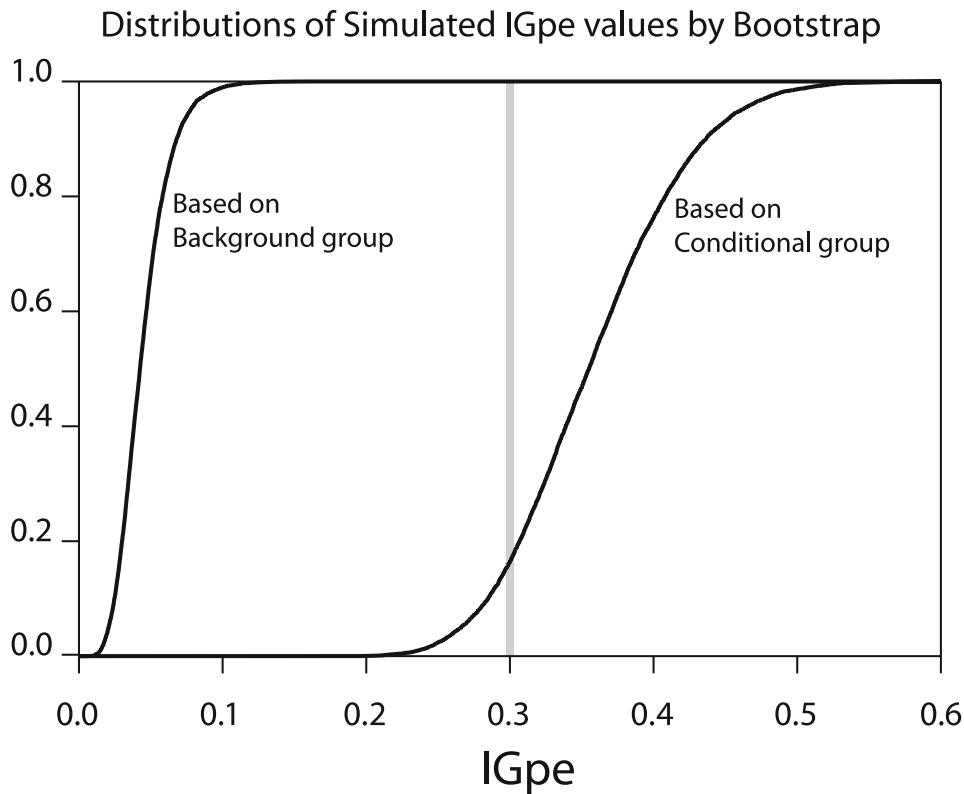


Fig. 4. Cumulative distributions of IGpe values simulated by a bootstrap method. The left curve with mean 0.04 and standard deviation 0.02 denotes IGpe values obtained based on background distributions, and the right curve with mean 0.36 and standard deviation 0.06 denotes IGpe values based on conditional distributions. The vertical line in the right curve indicates the IGpe estimated in the actual case.

10 km but becomes lower than the average velocity at a depth of 25 km (Table 2). A moderate correlation between two parameters is observed in the plot of the conditional distribution, whereas no clear correlation is observed in that of the background distribution (Table 3). In the present case, the contributions from every depth sum up to a negligible IGpe value of 0.06, but the correlations among parameters could lead to a useful IGpe value of 0.3.

5. Discussion and Conclusions

To confirm the above IGpe estimates, we have performed a simulation study to obtain the distribution of IGpe values by a bootstrap method. In generating a set of samples, we consider only the variations of parameter values in the conditional distributions (Tables 2 (right group) and 3 (upper right)). Two different data sources, the conditional distribution and the background distribution, are adopted. In both simulations, we randomly select 198 samples from the distribution, which are assigned as a simulated conditional distribution. After calculating means, variances, and correlation coefficients among parameters, we are able to estimate IGpe. By iterating 10,000 sets, we are also able to find the distribution of the IGpe value, which gives an average IGpe value and the standard deviation. When we select samples from the conditional distribution, we obtain an average of 0.36 and a standard deviation of 0.06 (right line in Fig. 4). The observed IGpe thus falls within one standard deviation of the mean. In contrast, when we use the background distribution, we obtain an average of 0.04 and a standard deviation of 0.02 (left line in Fig. 4). These results suggest

that the observed IGpe is not obtained by chance from the background distribution of the parameters.

Matsubara and Obara (2008) studied the relationship between the seismic velocity structure and the active tectonic faults in the Japan Islands. They first estimated velocity variations at depths of 5, 10, 15, and 20 km and then they compared the values beneath the fault zones with the nationwide averages. They found that velocity becomes higher than the average velocity in the shallow part beneath the fault zones but becomes lower than the average velocity in the deeper part. Based on this finding, they suggested that seismic velocity anomaly could contribute to the detection of blind active faults. Although their finding has not been examined quantitatively, it implies that a *P*-wave velocity model could contribute to the assessments of the seismogenesis of shallow earthquakes of moderate and large magnitude.

Taking into account the close relationship between large earthquakes and active faults, we focus on epicenters of earthquakes with a magnitude ≥ 5.0 at a shallow depth as the conditional group. It may be possible to adopt fault zones as a conditional group, but epicenters of earthquakes are more exactly defined and more easily selected than fault zones. Even with these simple selections, we are able to construct a seismicity model that could possibly assess the seismogenesis of shallow earthquakes. It may be possible to propose more effective models after various predictive parameters have been examined. However, how such models would perform remains to be seen.

Figure 5 shows the probability gains at every point of a

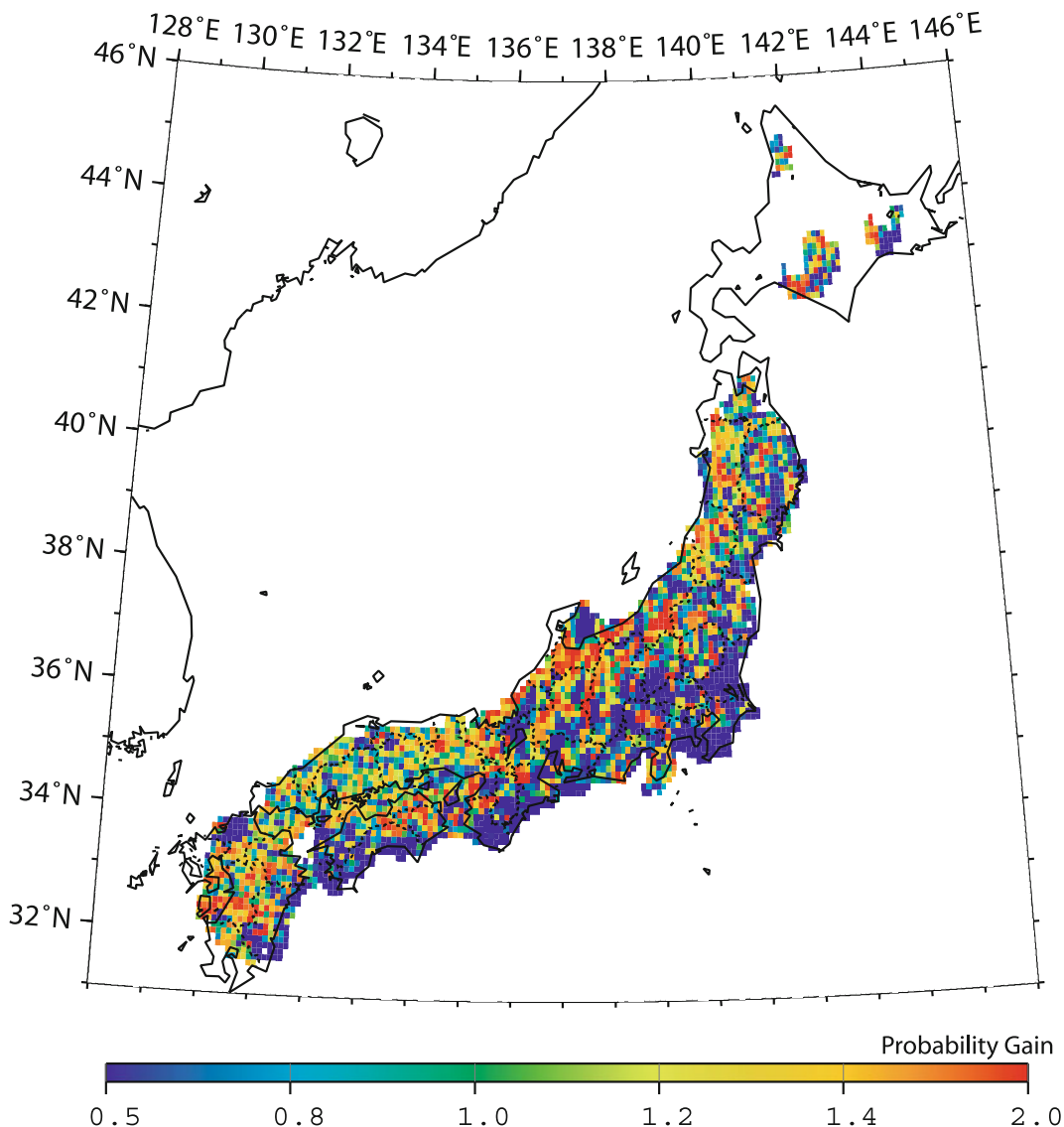


Fig. 5. A map of probability gains ($f(\theta)/g(\theta)$ values) at every point of a 0.1×0.1 grid.

$0.1 \times 0.1^\circ$ grid. In general, a probability gain is defined by a ratio of the hazard function (Eq. (1)) to the Poisson rate (m_0/V_0), which becomes equal to the $f(\theta)/g(\theta)$ value in the present case. Although the map indicates some parts of high probability gain up to 2.0, an average over the values at 198 epicenters (Table 1) becomes 1.35. Imoto and Rhoades (2010) combined two models, namely, the Every Earthquake a Precursor According to Scale model (EEPAS, Rhoades and Evison, 2006) and a three-parameter model (Imoto, 2008), into a better performance model in which the hazard rate of the EEPAS model is treated as a surrogate precursor. In a similar way, we can combine the present parameters and an appropriate seismicity model into a better performance model. A study focusing on this point will be conducted in the future.

In summary, we have attempted to assess the performance of a seismicity model for shallow earthquakes in Japan based on a *P*-wave velocity model. Applying the formula derived by Imoto (2007) to the *P*-wave velocity data, we assessed that IGpe of the model is 0.3 units, after incorporating the correlations among the parameters. The

bootstrap method suggests that this IGpe value could not be obtained by chance from the background distribution.

Acknowledgments. The author would like to express his thanks to M. Matsubara for providing a convenient tool to use their *P*-wave velocity data. He also thanks Euan Smith and Masashi Kawamura for their detailed reviews of this article and their valuable comments.

Appendix A.

The hazard function is expressed as the expectation of the number of earthquakes in a space-time volume $d\mathbf{x}$. For small $d\mathbf{x}$, the expectation becomes equal to the probability of an earthquake occurring in the volume, denoted as $P(E|\theta)$. By Bayes' theorem,

$$h(\mathbf{x}|\theta)d\mathbf{x} = P(E|\theta) = \frac{P(\theta|E) \cdot P(E)}{P(\theta)}, \quad (\text{A.1})$$

where $P(\theta)$ and $P(\theta|E)$ are unconditional and conditional probabilities of observing θ . We consider that $P(\theta)$ and $P(\theta|E)$ are represented by $g(\theta)d\theta$ and $f(\theta)d\theta$, respectively. The probability of an earthquake, $P(E)$, is given by

$m_0/V_0 d\mathbf{x}$. Thus

$$h(\mathbf{x}|\theta)d\mathbf{x} = \frac{f(\theta)m_0}{g(\theta)V_0}d\mathbf{x}, \quad (\text{A.2})$$

is obtained.

References

- Aki, K., A probabilistic synthesis of precursory phenomena, in *Earthquake Prediction*, edited by D. W. Simpson and P. G. Richards, 566–574, AGU, 1981.
- AL-Shukri, H. J. and B. J. Mitchell, Reduced seismic velocities in the source zone of New Madrid earthquakes, *Bull. Seismol. Soc. Am.*, **78**, 1491–1509, 1988.
- Daley, D. J. and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, vol 1, Elementary theory and methods, Second edition, 469 pp, Springer, New York, 2003.
- Grandori, G., E. Guagenti, and F. Perotti, Alarm systems based on a pair of short-term earthquake precursors, *Bull. Seismol. Soc. Am.*, **78**, 1538–1549, 1988.
- Hamada, K., A probability model for earthquake prediction, *Earthq. Predict. Res.*, **2**, 227–234, 1983.
- Hauksson, E. and J. S. Haase, Three-dimensional Vp and Vp/Vs velocity models of the Los Angeles basin and central Transverse Ranges, California, *J. Geophys. Res.*, **102**, 5423–5453, 1997.
- Imoto, M., Probability gains expected for renewal process models, *Earth Planets Space*, **56**, 563–571, 2004.
- Imoto, M., Earthquake probability based on multidisciplinary observations with correlations, *Earth Planets Space*, **57**, 1447–1454, 2006.
- Imoto, M., Information gain of a model based on multidisciplinary observations with correlations, *J. Geophys. Res.*, **112**, B05306, doi:10.1029/2006JB004662, 2007.
- Imoto, M., Performance of a seismicity model based on three parameters for earthquakes ($M \geq 5.0$) in Kanto, central Japan, *Ann. Geophys.*, **51**(4), 727–736, 2008.
- Imoto, M. and D. Rhoades, Seismicity models of moderate earthquakes in Kanto, Japan utilizing multiple predictive parameters, *Pure Appl. Geophys.*, **167**, 831–843, doi 10.1007/s00024-010-0066-4, 2010.
- Kaufmann, R. D. and L. T. Long, Velocity structure and seismicity of southeastern Tennessee, *J. Geophys. Res.*, **101**, 8531–8542, 1996.
- Matsubara, M. and K. Obara, Relationship between the seismic velocity structure and the active tectonic faults in the crust beneath Japan islands, *Eos Trans. AGU*, **89**(53), Fall Meet. Suppl., Abstract T53C-1954, 2008.
- Matsubara, M., K. Obara, and K. Kasahara, Three-dimensional P- and S-wave velocity structures beneath the Japan Islands obtained by high-density seismic stations by seismic tomography, *Tectonophysics*, **454**, 86–103, 2008.
- Michael, A. J., Effects of three-dimensional velocity structure on the seismicity of the 1984 Morgan Hill, California, aftershock sequence, *Bull. Seismol. Soc. Am.*, **78**, 1199–1221, 1988.
- Nur, A., Dilatancy, pore fluids and premonitory variations of ts/tp travel times, *Bull. Seismol. Soc. Am.*, **62**, 1217–1222, 1972.
- Obara, K., K. Kasahara, S. Hori, and Y. Okada, A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention, *Rev. Sci. Instrum.*, **76**, 021301, 2005.
- Rhoades, D. and F. Evison, Long-range earthquake forecasting based on a single predictor, *Geophys. J. R. Astron. Soc.*, **59**, 43–56, 1979.
- Rhoades, D. and F. Evison, The EEPAS forecasting model and the probability of moderate-to-large earthquakes in central Japan, *Tectonophysics*, **417**, 119–130, 2006.
- Scholz, C., L. R. Sykes, and Y. P. Aggarwal, Earthquake prediction: a physical basis, *Science*, **181**, 803–810, 1973.
- Utsu, T., Probabilities in earthquake prediction, *Zisin II*, **30**, 179–185, 1977 (in Japanese).
- Utsu, T., Probabilities in earthquake prediction (the second paper), *Bull. Earthq. Res. Inst.*, **57**, 499–524, 1982 (in Japanese).

M. Imoto (e-mail: imoto@bosai.go.jp)