

Conventional N-, L-, and R-tests of earthquake forecasting models without simulated catalogs

Masajiro Imoto¹, David A. Rhoades², Hiroyuki Fujiwara¹, and Naoko Yamamoto¹

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

²GNS Science, 1 Fairway Drive, Avalon, Lower Hutt 5010, P.O. Box 30-368, Lower Hutt, 5040 New Zealand

(Received January 12, 2010; Revised July 27, 2010; Accepted August 24, 2010; Online published March 4, 2011)

We propose a new procedure for testing the expected number (N-test), log likelihood (L-test), and log likelihood-ratio (R-test) of seismicity models. In these tests, scores obtained from observed earthquakes are compared with distributions of scores estimated from earthquakes expected from the models under test. We introduce a method to estimate the test score distributions analytically where uncertainties in magnitude and hypocentral parameters are involved. The analytical formulas used to estimate expected values and standard deviations of the test scores for earthquakes conforming to the test models were derived in earlier published studies, which allowed calculation of normal approximations by which to test score distributions. Using these two methods simultaneously, we can perform N-, L-, and R-tests for seismicity models without using any simulated catalogs. As a case study, the proposed procedure was applied to two seismicity models for Kanto, central Japan. To compare our procedure with the current one based on the Monte Carlo method, we randomly generated sets of 10,000 earthquake catalogs of two kinds: one set conforming to the model under test, and the other derived from the observed catalog allowing for uncertainties in magnitude and hypocentral parameters. The distributions of L-scores obtained from both sets are in good agreement with those obtained by the proposed procedure. This comparison suggests that the analytical approach presented here could be useful for conducting the N-, L-, and R-tests in a conventional way.

Key words: Seismicity models, probabilistic prediction, CSEP project, N-, L-, and R-scores, uncertainties in earthquake source parameters, simulated catalogs.

1. Introduction

With the development of statistical seismology, probabilistic predictions of earthquakes are now more common than has hitherto been possible. The probabilities of earthquake occurrence are usually estimated based on specific seismicity models. To provide a reliable probabilistic prediction, the model should pass well-defined tests. The Collaboratory for the Study of Earthquake Predictability (CSEP) Project (Jordan, 2006) has been organized to solve questions related to earthquake predictability and to develop an adequate experimental infrastructure to conduct scientific prediction experiments under rigorous conditions. The aim of the CSEP Project is to evaluate each proposed model with three statistical tests, i.e., the N-, L-, and R-tests (Kagan and Jackson, 1995). In these tests, observed scores are compared with the distributions of respective scores expected from a proposed model. When the observed scores fall within an acceptance range, the model is not rejected.

In calculating observed scores, effects due to uncertainties in the earthquake source parameters of location, depth, and magnitude are also taken into account in the CSEP procedures (Schorlemmer *et al.*, 2007). The underlying ratio-

nale for this step is that parameter uncertainties may cause earthquakes to be associated with bins differing from those originally assigned. A serious problem may arise if these uncertainties are ignored. For example, the involvement of a particular earthquake in the score can become largely a matter of chance if it is close to the boundaries of the test region or to the lower magnitude limit of the tests.

A large number of simulated catalogs are generated in the CSEP tests in order to obtain distributions of N-, L-, and R-scores expected from the proposed models (Schorlemmer *et al.*, 2007). The observed scores are estimated using simulated catalogs derived from an original one by allowing for uncertainties in earthquake source parameters. Generating a large number of catalogs consumes a great deal of computational time that is proportional to the number of proposed models and time and space segments necessary for forecasting.

In this paper, we propose a method for conducting these tests analytically without generating simulated catalogs. This method is based on the assumptions that the rate in each cell is far less than unity and that at most one earthquake occurs in one cell. For each test, we analytically derive two sets of the mean and variance of the respective score: that expected from a proposed model (Imoto, 2009; Imoto and Rhoades, 2010) and that of the probable score if uncertainties in hypocenter parameters are taken into account. The central limit theorem allows us to regard the

Copyright © The Society of Geomagnetism and Earth, Planetary and Space Sciences (SGEPSS); The Seismological Society of Japan; The Volcanological Society of Japan; The Geodetic Society of Japan; The Japanese Society for Planetary Sciences; TERRAPUB.

distribution of the scores as approximately normal, with the analytically obtained means and variances, if the number of earthquakes in the test area is large enough.

The proposed method is applied to two seismicity models for Kanto, central Japan. The first model (Hazmap model) is tentatively introduced as a candidate for testing and is a subset of a seismicity model used in estimating the recent seismic hazard maps for Japan. The second one is the EEPAS model (Rhoades and Evison, 2004, 2005), which was developed into a horizontal multi-layered model for seismicity in the Kanto region, central Japan (Rhoades and Evison, 2006; Imoto and Rhoades, 2010). The Hazmap model is configured to the tectonic setting in Kanto, where three plates (the Eurasian (North American, or Okhotsk) Plate, the Philippine Sea Plate, and the Pacific Plate) converge. Differences in configuration between the Hazmap and EEPAS models could be resolved by re-configuring the EEPAS model to be compatible with the Hazmap under a simple assumption. Thus, the Hazmap model could be compared with the EEPAS model in the R-test.

In order to compare our analytical approach with that involving the use of simulated catalogs, we have compared the distributions of L-scores by both methods. The results show that the means of L-scores for earthquakes conforming to a model under test and for those with parameter uncertainties taken into consideration are similar whether computed by our method or by the method involving simulated catalogs.

2. N-, L-, and R-Tests

In the N-, L-, and R-tests, scores obtained from the observed data are compared with a distribution of scores that could be expected assuming a given model to be correct. If the observed scores are not consistent with these distributions, the null hypothesis that the real earthquake sequences conform to the model may be rejected. If an observed score falls outside an acceptance range, the model should be rejected at the respective level of significance. In the following discussion of the three tests, we will introduce methods to estimate distributions of the test scores for a catalog with uncertain parameters. Using methods already presented in previous papers (Imoto, 2009; Imoto and Rhoades, 2010), these tests can be conducted without using simulated catalogs.

2.1 N-test

2.1.1 N-score expected from a proposed model The N-test checks the consistency of the observed number with the earthquake productivity of the proposed model. This test does not involve a (space, time, and magnitude) distribution of earthquakes. Assuming that the study space-time domain is divided into V_0 cells, the number of earthquakes in each cell follows the Poisson process. The Poisson rate in the i -th cell is noted as λ_i , i.e., the expected number of events, and its variance $\text{Var}[n_i(\Omega^1)] = \lambda_i^1$. The total expected number, $E[n(\Omega^1)]$, is given by

$$E[n(\Omega^1)] = \sum_{i=1}^{V_0} \lambda_i^1, \quad (1)$$

where the superscript 1 refers to the first model, and Ω^1 denotes data conforming to the first model. Earthquakes are assumed to be independent. The variance of the total number is given as follows.

$$\text{Var}[n(\Omega^1)] = \sum_{i=1}^{V_0} \text{Var}[n_i(\Omega^1)] = \sum_{i=1}^{V_0} \lambda_i^1 \quad (2)$$

If the number of earthquakes expected under the model remains above 10, its distribution may be well approximated by a normal distribution $g(n)$:

$$g(n) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{(n - \mu_1)^2}{2\sigma_1^2}\right\}, \quad (3)$$

where μ_1 and σ_1^2 are given by $E[n(\Omega^1)]$ and $\text{Var}[n(\Omega^1)]$, respectively. In any case, the Poisson distribution itself exactly defines the acceptance range for the test.

2.1.2 N-score from observed events with uncertain hypocentral parameters The observed number of events should be compared with the distribution of the N-score, where the score follows the Poisson distribution with the mean and the variance given in Eqs. (1) and (2). When we consider effects caused by errors in parameters, the observed number of events becomes uncertain. Some events occur in certain cells with certain probabilities, and others occur outside the magnitude and location limits of the study area. Here, we consider that the j -th event could be located in m_j different cells with probabilities of $P_{j,k}$ ($k = 1, 2, \dots, m_j$). The observed number of events, $n(\Omega^0)$, expected from catalogs with errors in parameters taken into account is given by

$$E[n(\Omega^0)] = \sum_{j=1}^{N_0} \sum_{k=1}^{m_j} P_{j,k}, \quad (4)$$

where Ω^0 denotes the actual earthquake occurrences in the study area, and N_0 is the number of events, the possible parameters of which belong to the study area. If it is certain that a specific event is inside the study area, the summation for the event in Eq. (4) becomes 1. However, for events near the border of the study area, the summation is not necessarily 1. The variance of $n(\Omega^0)$ is given as

$$\text{Var}[n(\Omega^0)] = \sum_{j=1}^{N_0} \left\{ \sum_{k=1}^{m_j} P_{j,k} \left(1 - \sum_{k=1}^{m_j} P_{j,k} \right) \right\}. \quad (5)$$

Under an assumption similar to that in the distribution of $g(n)$, the observed number of earthquakes may be well represented by a normal distribution, $f(n)$, with the mean and the variance given in Eqs. (4) and (5).

Under the model, the probability α that the number of earthquakes is less than n is given by

$$\alpha = \int_0^n g(u) du = \Phi(n). \quad (6)$$

Taking into account the uncertainty in $n(\Omega^0)$ represented by the distribution $f(n)$, the expected value of the probability

$\bar{\alpha}$ that the number of earthquakes in the model is less than $n(\Omega^0)$ is given by

$$\bar{\alpha} = \int_0^{\infty} \Phi(n) \cdot f(n) dn. \quad (7)$$

If $\bar{\alpha}$ falls outside the acceptance range (for example, 0.05–0.95), the proposed model should be rejected.

2.2 L-test

2.2.1 L-score expected from a proposed model The L-test determines the consistency of the observed likelihood with that expected from sequences conforming to the model. This test measures the consistency of a forecast distribution with the observed one. The log likelihood for the i -th cell is given by

$$l_i(\Omega^1) = Y_i \text{Ln} \lambda_i^1 + (1 - Y_i) \text{Ln}(1 - \lambda_i^1), \quad (8)$$

where Y_i is 1 when an earthquake occurs and 0 otherwise, and Ln refers to the natural logarithm. It is assumed that λ_i is far less than unity and that at most one earthquake occurs in one cell. This condition might be satisfied by $\lambda_i^2 V_0$ less than 1 as a first order estimate. The expected value of the log likelihood, $E[l_i(\Omega^1)]$ is expressed by

$$E[l_i(\Omega^1)] = \lambda_i^1 \text{Ln} \lambda_i^1 + (1 - \lambda_i^1) \text{Ln}(1 - \lambda_i^1). \quad (9)$$

The variance is given by

$$\begin{aligned} \text{Var}[l_i(\Omega^1)] &= E \left[\left(\text{Ln} \frac{\lambda_i^1}{1 - \lambda_i^1} \right)^2 \cdot (Y_i^1 - \lambda_i^1)^2 \right] \\ &\approx \left(\text{Ln} \frac{\lambda_i^1}{1 - \lambda_i^1} \right)^2 \cdot \lambda_i^1 \end{aligned} \quad (10)$$

where the superscript $\hat{2}$ denotes the second power in Section 2.2 and 2.3 (Imoto, 2009; Imoto and Rhoades, 2010). When an earthquake in one cell is independent from that in another cell, the mean $E[l(\Omega^1)]$ and variance $\text{Var}[l(\Omega^1)]$ of the log likelihood for the whole domain are given by

$$E[l(\Omega^1)] = \sum_{i=1}^{V_0} \left[\lambda_i^1 \text{Ln} \frac{\lambda_i^1}{1 - \lambda_i^1} + \text{Ln}(1 - \lambda_i^1) \right] \quad (11)$$

and

$$\text{Var}[l(\Omega^1)] \approx \sum_{i=1}^{V_0} \left[\left\{ \text{Ln} \frac{\lambda_i^1}{1 - \lambda_i^1} \right\}^2 \lambda_i^1 \right]. \quad (12)$$

2.2.2 L-score from observed events with uncertain hypocentral parameters We first assume that there are no errors in the catalog. The log likelihood of one event occurring in the j -th cell (and no event in any other cell) is then given as

$$l_j(\Omega^0) = \text{Ln} \frac{\lambda_j^1}{1 - \lambda_j^1} + \sum_{i=1}^{V_0} \text{Ln}(1 - \lambda_i^1). \quad (13)$$

The difference in log likelihood between the case of one event in the j -th cell and no event, $\Delta l_j(\Omega^0)$, is given by

$$\Delta l_j(\Omega^0) = \text{Ln} \frac{\lambda_j^1}{1 - \lambda_j^1}. \quad (14)$$

If the event occurs in the j' -th cell, not in the j -th cell, the difference in likelihood is given by the same formula with j' in the place of j . If the location and/or magnitude parameters contain significant errors, an earthquake could be located in several different cells with certain probabilities.

If the j -th event is located in one of m_j cells with probabilities $P_{j,1}, P_{j,2}, \dots, P_{j,m_j}$, the expected value $E[\Delta l_j(\Omega^0)]$ of the difference in the log likelihood between the cases with and without the j -th earthquake is given by

$$\begin{aligned} E[\Delta l_j(\Omega^0)] &= \sum_{k=1}^{m_j} P_{j,k} \Delta l_{j,k}^1 \\ &= \sum_{k=1}^{m_j} P_{j,k} \left\{ \text{Ln} \frac{\lambda_{j,k}^1}{1 - \lambda_{j,k}^1} \right\} \end{aligned} \quad (15)$$

where the subscript $\lambda_{j,k}$ indicates the k -th possible cell for the j -th earthquake.

Summing up terms from all the events and the no-event terms, we have the following:

$$E[l^1(\Omega^0)] = \sum_{j=1}^{N_0} \sum_{k=1}^{m_j} P_{j,k} \left(\text{Ln} \frac{\lambda_{j,k}^1}{1 - \lambda_{j,k}^1} \right) + \sum_{i=1}^{V_0} \{ \text{Ln}(1 - \lambda_i^1) \}. \quad (16)$$

The variance of the log likelihood is given as

$$\begin{aligned} \text{Var}[l^1(\Omega^0)] &= \sum_{j=1}^{N_0} \left[\sum_{k=1}^{m_j} P_{j,k} \{ \Delta l_{j,k} - E[\Delta l_j(\Omega^0)] \}^2 \right. \\ &\quad \left. + \left(1 - \sum_{k=1}^{m_j} P_{j,k} \right) \cdot E[\Delta l_j(\Omega^0)]^2 \right] \\ &= \sum_{j=1}^{N_0} \left[\sum_{k=1}^{m_j} \left\{ P_{j,k} \left(\text{Ln} \frac{\lambda_{j,k}^1}{1 - \lambda_{j,k}^1} \right)^2 \right\} \right. \\ &\quad \left. - E(\Delta l_j(\Omega^0))^2 \right] \end{aligned} \quad (17)$$

In a similar way as for the N-test, the L-test can be performed based on the parameters given in Eqs. (11), (12), (16), and (17).

2.3 R-test

2.3.1 R-score expected from two models The R-test involves a comparison of the log likelihoods of two models in which each model in turn is regarded as the null hypothesis. The expected value of the log likelihood-ratio for the i -th cell, $E[R_i^{12}(\Omega^1)]$, is expressed by

$$E[R_i^{12}(\Omega^1)] = \lambda_i^1 \text{Ln} \left\{ \frac{\lambda_i^1 (1 - \lambda_i^2)}{\lambda_i^2 (1 - \lambda_i^1)} \right\} + \text{Ln} \left\{ \frac{(1 - \lambda_i^1)}{(1 - \lambda_i^2)} \right\}, \quad (18)$$

where superscripts 1 and 2 refer to the first and second models, respectively. Here it is assumed that earthquake sequences conform to the first model (Imoto, 2009; Imoto and Rhoades, 2010). The variance is given by

$$\begin{aligned} \text{Var}[R_i^{12}(\Omega^1)] &= E \left[\left\{ \text{Ln} \left\{ \frac{\lambda_i^1}{\lambda_i^2} \cdot \frac{(1 - \lambda_i^2)}{(1 - \lambda_i^1)} \right\} \right\}^2 \cdot (Y_i^1 - \lambda_i^1)^2 \right] \\ &\approx \left[\text{Ln} \left\{ \frac{\lambda_i^1}{\lambda_i^2} \cdot \frac{(1 - \lambda_i^2)}{(1 - \lambda_i^1)} \right\} \right]^2 \cdot \lambda_i^1. \end{aligned} \quad (19)$$

The mean $E[R^{12}]$ and variance $\text{Var}[R^{12}]$ of the log likelihood ratio for the whole domain are given by

$$E[R^{12}(\Omega^1)] = \sum_{i=1}^{V_0} \left[\lambda_i^1 \text{Ln} \left\{ \frac{\lambda_i^1 (1 - \lambda_i^2)}{\lambda_i^2 (1 - \lambda_i^1)} \right\} + \text{Ln} \left\{ \frac{1 - \lambda_i^1}{1 - \lambda_i^2} \right\} \right] \quad (20)$$

and

$$\text{Var}[R^{12}(\Omega^1)] \approx \sum_{i=1}^{V_0} \left[\text{Ln} \left\{ \frac{\lambda_i^1 (1 - \lambda_i^2)}{\lambda_i^2 (1 - \lambda_i^1)} \right\} \right]^2 \cdot \lambda_i^1. \quad (21)$$

2.3.2 R-score from observed events with uncertain hypocentral parameters If errors in earthquake source parameters are taken into account, the log likelihood ratio of model 1 to that of model 2, $E[R^{12}]$ is estimated from Eq. (16).

$$\begin{aligned} E[R^{12}(\Omega^0)] &= E[I^1(\Omega^0)] - E[I^2(\Omega^0)] \\ &= \sum_{j=1}^{N_0} \left[\sum_{k=1}^{m_j} \left\{ P_{j,k} \text{Ln} \frac{\lambda_{j,k}^1 (1 - \lambda_{j,k}^2)}{(1 - \lambda_{j,k}^1) \lambda_{j,k}^2} \right\} \right] \\ &\quad + \sum_{i=1}^{V_0} \left\{ \text{Ln} \frac{(1 - \lambda_i^1)}{(1 - \lambda_i^2)} \right\}. \end{aligned} \quad (22)$$

The variance of the log likelihood ratio is due only to the cells related to earthquakes and is given by

$$\begin{aligned} \text{Var}[R^{12}(\Omega^0)] &= \sum_{j=1}^{N_0} \left[\sum_{k=1}^{m_j} \left\{ P_{j,k} \left(\text{Ln} \frac{\lambda_{j,k}^1 (1 - \lambda_{j,k}^2)}{(1 - \lambda_{j,k}^1) \lambda_{j,k}^2} \right)^2 \right\} \right] \\ &\quad - E[\Delta R_j(\Omega^0)]^2 \end{aligned} \quad (23)$$

where

$$E[\Delta R_j(\Omega^0)] = \sum_{k=1}^{m_j} \left[P_{j,k} \cdot \text{Ln} \left(\frac{\lambda_{j,k}^1 (1 - \lambda_{j,k}^2)}{(1 - \lambda_{j,k}^1) \lambda_{j,k}^2} \right) \right]. \quad (24)$$

In a similar manner as for the above tests, the R-test can be performed based on the parameters given in Eqs. (20), (21), (22), and (23).

3. Models and Data

3.1 Seismic hazard map model

Probabilistic seismic hazard maps were prepared for Japan based on the long-term probability of earthquakes (Fujiwara, 2004; Fujiwara *et al.*, 2009). Future earthquakes in and around Japan are classified into several categories, such as along major and minor inland active faults, thrust faults along subduction zones, and others. The probabilistic seismic hazard estimates from every category are then merged into a total hazard.

In the present study, one candidate seismicity model tentatively considered is a subset of the long-term probability pertaining to the category of earthquakes without specified faults (Hazmap model). In this category, the long-term

probability of earthquakes is estimated based on a smoothed seismicity of past earthquakes. The target earthquakes in our tested time-space volume all belong to this category.

Figure 1(a) depicts the study area in Kanto, Japan, where three plates (the Eurasian (North American, or Okhotsk) Plate, the Philippine Sea Plate, and the Pacific Plate) converge. The Hazmap model configures the tectonic setting there using a three-layered structure with different depth ranges for different grid points every 0.1° . Figure 1(b) schematically represents a vertical section in the region. At the surface, the layer covers the depth range of 0–25 km, which includes earthquakes in the crust, most of which belong to the Eurasian Plate. The eastern border of this layer is indicated by the lightly colored dashed line in Fig. 1(a). Earthquakes in the subducted Philippine Sea Plate deeper than 25 km and shallower than the upper boundary of the subducted Pacific Plate are included in the second layer. This layer is limited to a region around Tokyo (indicated by solid lines in Fig. 1(a)). The third layer includes earthquakes originating from the subducted Pacific plate, which covers the region for study (indicated by the box in Fig. 1(a)). The boundary is set at 5 km above the interface between the Philippine Sea and Pacific Plates, since the inter-plate earthquakes occurring between them are mixed with earthquakes in the Pacific plates; a margin of 5 km is taken in order to account for fluctuations of inter-plate earthquakes. All of the parameters needed to draw these configurations were obtained from the Japan Seismic Hazard Information Station (J-Shis) site of the National Research Institute for Earth Science and Disaster Prevention (NIED, 2010).

3.2 EEPAS model

In the EEPAS model, the hazard function of earthquake occurrence, $h_{\text{eep}}(x)$, is defined for any time t , magnitude m , and location (x, y) where m exceeds a threshold magnitude m_c , and (x, y) is a point in the region of surveillance R . Each earthquake (t_i, m_i, x_i, y_i) contributes a transient increment $\lambda_i(t, m, x, y)$ to the future rate density in its vicinity, given by

$$\lambda_i(t, m, x, y) = f_{li}(t) g_{li}(m) h_{li}(x, y), \quad (25)$$

where f_{li} is the density of the probability distributions for time, g_{li} is that for magnitude, and h_{li} is that for location. These densities take the forms

$$g_{li}(m) = \frac{1}{\sigma_M \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{m - a_M - b_M m_i}{\sigma_M} \right)^2 \right], \quad (26)$$

$$\begin{aligned} f_{li}(t) &= \frac{H(t - t_i)}{(t - t_i) \ln(10) \sigma_T \sqrt{2\pi}} \\ &\quad \cdot \exp \left[-\frac{1}{2} \left(\frac{\log_{10}(t - t_i) - a_T - b_T m_i}{\sigma_T} \right)^2 \right], \end{aligned} \quad (27)$$

and

$$h_{li}(x, y) = \frac{1}{2\pi \sigma_A^2 10^{b_A M_i}} \exp \left[-\frac{(x - x_i)^2 + (y - y_i)^2}{2\sigma_A^2 10^{b_A M_i}} \right], \quad (28)$$

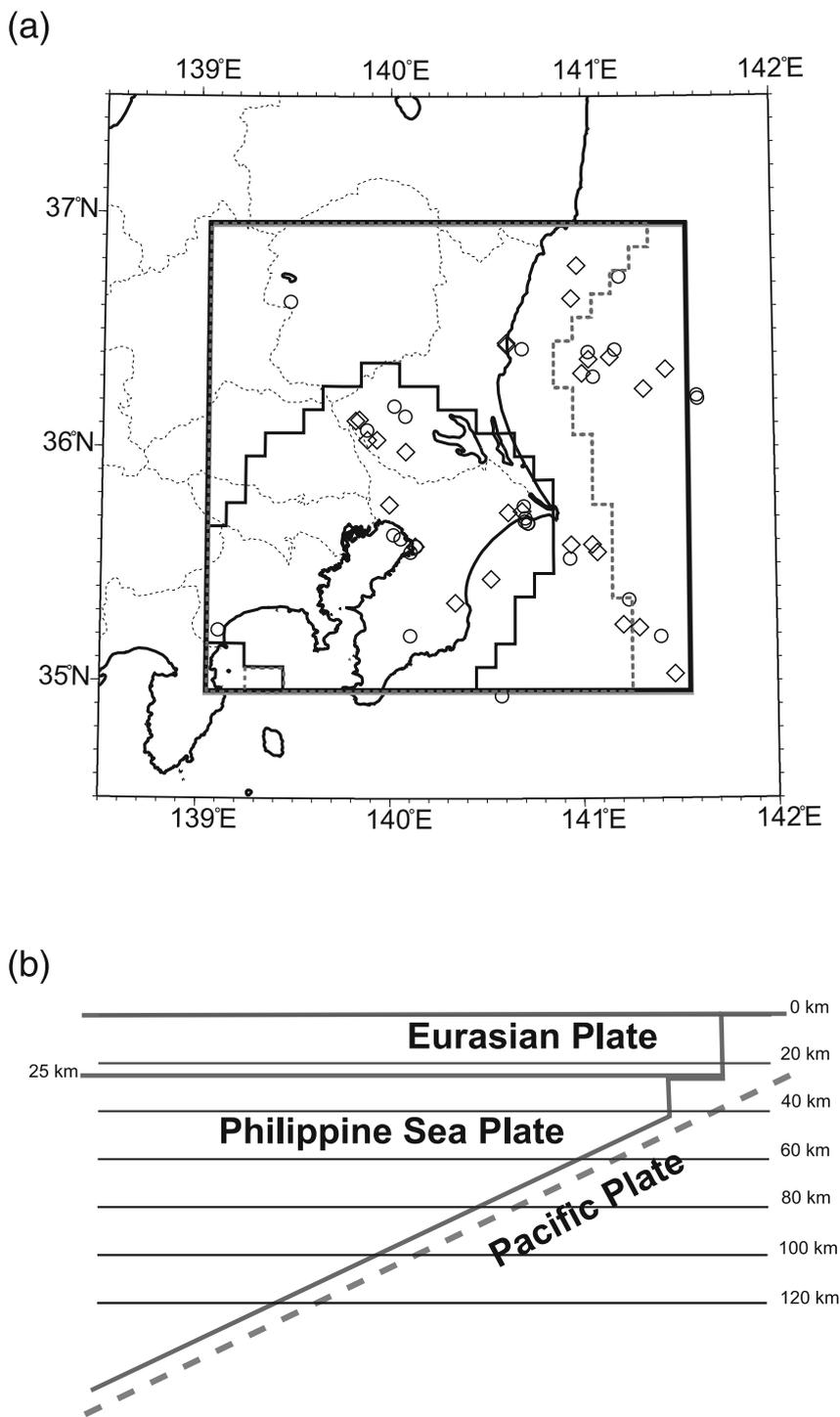


Fig. 1. (a) Map of target earthquakes with magnitude ≥ 5.0 from 1 January 2004 to 31 December 2008. Diamonds (circles) indicate the locations of target earthquakes with a probability greater (less) than 0.5 that the earthquake falls in any cells of the study volume. The study region is enclosed by a box. The lightly colored dashed line indicates an area covered by the top layer of the Hazmap model. The dark solid line indicates an area of the subducted Philippine sea plate as the second layer. The third layer covers the whole study area. (b) Schematic illustration of a vertical section of layered models for estimating hazard rates. The horizontally layered EEPAS model is reframed into the three-layered model labeling the Eurasian Plate, Philippine Sea Plate, and Pacific Plate.

where $a_M, b_M, \sigma_M, a_T, b_T, \sigma_T, \sigma_A,$ and b_A are parameters; otherwise, $H(s) = 1$ if $s > 0$ and 0.

The total rate density is obtained by summing over all past occurrences, including earthquakes outside R that could affect the rate density within R . More detail is given in previous studies (e.g., Evison and Rhoades, 2002, 2004; Rhoades and Evison, 2004). We do not distinguish the mag-

nitudes of target earthquakes. Therefore, the integral form of the rate density is used. A minor modification of the EEPAS model was made to study three-dimensional seismicity in the Kanto region, Japan (Rhoades and Evison, 2006), where the depth of 0–120 km is divided into six layers, and the EEPAS model is applied to each layer.

Differences in configuration between the Hazmap and

Table 1. List of target earthquakes. “Prob” in the last column indicates the probability that the earthquake falls in the study volume.

| | Date | LAT(deg) | LON(deg) | DEP(km) | Mag | Prob |
|----|----------|----------|----------|---------|-----|---------|
| 1 | 20040311 | 36.322 | 141.007 | 47.5 | 5.3 | 0.99908 |
| 2 | 20040404 | 36.390 | 141.153 | 49.0 | 5.8 | 1.00000 |
| 3 | 20040710 | 36.080 | 139.885 | 48.2 | 4.7 | 0.00621 |
| 4 | 20040821 | 35.037 | 141.480 | 38.4 | 5.4 | 0.97818 |
| 5 | 20041006 | 35.988 | 140.088 | 66.0 | 5.7 | 0.99992 |
| 6 | 20041017 | 36.255 | 141.328 | 49.0 | 5.7 | 0.99999 |
| 7 | 20050101 | 36.783 | 140.983 | 89.4 | 5.0 | 0.69145 |
| 8 | 20050208 | 36.140 | 140.085 | 67.0 | 4.8 | 0.06681 |
| 9 | 20050216 | 36.038 | 139.888 | 46.2 | 5.3 | 0.99977 |
| 10 | 20050411 | 35.727 | 140.620 | 51.5 | 6.1 | 0.99839 |
| 11 | 20050515 | 36.628 | 139.483 | 8.8 | 4.8 | 0.06681 |
| 12 | 20050515 | 35.755 | 140.702 | 50.8 | 4.7 | 0.00621 |
| 13 | 20050519 | 35.558 | 141.082 | 33.3 | 5.4 | 0.99984 |
| 14 | 20050530 | 35.690 | 140.707 | 51.2 | 4.7 | 0.00621 |
| 15 | 20050620 | 35.733 | 140.693 | 50.7 | 5.6 | 0.99964 |
| 16 | 20050711 | 36.413 | 141.040 | 47.8 | 4.7 | 0.00621 |
| 17 | 20050723 | 35.582 | 140.138 | 73.1 | 6.0 | 1.00000 |
| 18 | 20050728 | 36.125 | 139.845 | 51.1 | 5.0 | 0.69146 |
| 19 | 20050807 | 35.558 | 140.113 | 73.3 | 4.7 | 0.00621 |
| 20 | 20050808 | 36.338 | 141.445 | 46.0 | 5.6 | 0.99997 |
| 21 | 20050909 | 35.590 | 140.947 | 37.5 | 5.0 | 0.69146 |
| 22 | 20051016 | 36.038 | 139.937 | 47.1 | 5.1 | 0.93319 |
| 23 | 20051019 | 36.382 | 141.042 | 48.3 | 6.3 | 1.00000 |
| 24 | 20051116 | 36.308 | 141.065 | 47.8 | 4.8 | 0.06681 |
| 25 | 20051228 | 36.182 | 140.027 | 53.0 | 4.8 | 0.06681 |
| 26 | 20060201 | 35.760 | 140.003 | 101.0 | 5.1 | 0.93305 |
| 27 | 20060203 | 36.213 | 141.610 | 62.1 | 5.9 | 0.00587 |
| 28 | 20060520 | 35.202 | 140.112 | 73.8 | 4.7 | 0.00621 |
| 29 | 20060831 | 35.632 | 140.023 | 75.6 | 4.8 | 0.06681 |
| 30 | 20060907 | 35.590 | 141.057 | 38.2 | 5.1 | 0.93316 |
| 31 | 20070314 | 35.705 | 140.710 | 52.4 | 4.7 | 0.00621 |
| 32 | 20070620 | 35.235 | 141.297 | 24.9 | 5.0 | 0.69146 |
| 33 | 20070621 | 35.198 | 141.408 | 38.8 | 4.8 | 0.06681 |
| 34 | 20070816 | 35.443 | 140.530 | 30.8 | 5.3 | 0.99977 |
| 35 | 20070818 | 35.342 | 140.345 | 20.2 | 5.2 | 0.99379 |
| 36 | 20070921 | 35.248 | 141.215 | 35.0 | 5.1 | 0.93316 |
| 37 | 20071001 | 35.225 | 139.118 | 14.2 | 4.9 | 0.30854 |
| 38 | 20071007 | 35.683 | 140.725 | 44.5 | 4.8 | 0.06680 |
| 39 | 20071130 | 36.427 | 140.695 | 52.9 | 4.7 | 0.00621 |
| 40 | 20080116 | 36.423 | 141.180 | 47.0 | 4.7 | 0.00621 |
| 41 | 20080209 | 36.735 | 141.205 | 46.3 | 4.9 | 0.30813 |
| 42 | 20080308 | 36.452 | 140.612 | 57.0 | 5.2 | 0.99379 |
| 43 | 20080401 | 34.945 | 140.585 | 98.1 | 4.7 | 0.00141 |
| 44 | 20080404 | 36.120 | 139.827 | 53.5 | 5.0 | 0.69103 |
| 45 | 20080425 | 35.683 | 140.722 | 48.2 | 4.8 | 0.06679 |
| 46 | 20080508 | 36.227 | 141.607 | 50.6 | 7.0 | 0.01178 |
| 47 | 20080705 | 36.642 | 140.952 | 49.7 | 5.2 | 0.99379 |
| 48 | 20080822 | 36.442 | 140.615 | 55.9 | 5.2 | 0.99377 |
| 49 | 20080906 | 35.355 | 141.245 | 40.3 | 4.7 | 0.00621 |
| 50 | 20080921 | 35.615 | 140.062 | 70.8 | 4.8 | 0.06680 |
| 51 | 20081008 | 35.583 | 140.142 | 63.1 | 4.7 | 0.00621 |
| 52 | 20081231 | 35.532 | 140.940 | 42.7 | 4.7 | 0.00621 |

EEPAS models (Fig. 1(b)) prevent us from performing the R-test. However, one simple assumption could enable us to make the EEPAS model compatible with the Hazmap model. Assuming that the cell size of the EEPAS model is so small that a uniform Poisson process is maintained in each cell, we could divide a cell into multiple pieces of arbitrary size, the hazard rates of which are estimated to be proportional to their volume. In our case, only the depth range for the EEPAS model differs from that of the Hazmap. Therefore, we performed “divide” and “connect” upon cells

of the EEPAS model to make them compatible for the depth range only, since the horizontal grid spacing of $0.1 \times 0.1^\circ$ is common between the two models. The modified EEPAS model thus obtained is adapted in the present study for comparison with the Hazmap model. Hereafter, we refer to this modified EEPAS model as the EEPAS model.

3.3 Data

We use the catalog of earthquakes from 2004 to 2008 located by the Japan Meteorological Agency (JMA, 2009). The Hazmap model estimates hazard rates of earthquakes

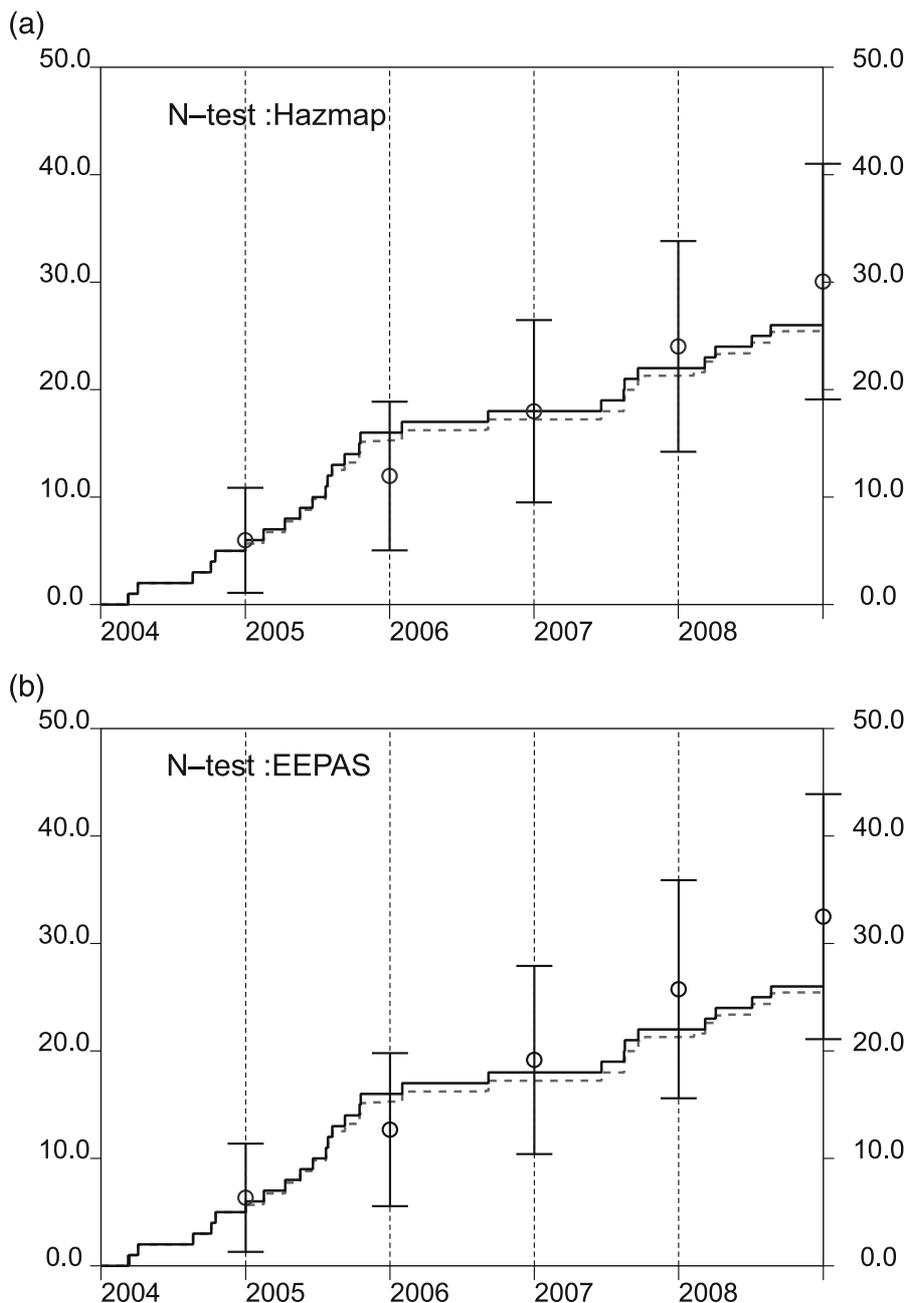


Fig. 2. N-test for January 2004 to December 2008. (a) Cumulative numbers of target earthquakes are compared with those expected at the end of each year from sequences conforming to the Hazmap model. The ranges of one standard deviation are indicated with error bars. The gray broken line (solid line) plots the cumulative numbers of earthquakes ($M \geq 5.0$) for cases with (without) errors in magnitude. (b) EEPAS model tested.

with magnitudes ≥ 5.0 . Epicenters of earthquakes ($M \geq 4.7$) used in the study are plotted in Fig. 1(a), where a solid box indicates the area examined. Here, we selected earthquakes of magnitude ≥ 4.7 , assuming a standard deviation of 0.1 in magnitude determination and deducing that earthquakes with magnitude > 4.7 have non-negligible probabilities (in the order of $\geq 1\%$) that they become > 5.0 . Table 1 lists the times of occurrence, locations, and magnitudes of earthquakes used in the present study. The last column indicates the probability that the earthquakes occur in cells of the model. Earthquakes with a probability < 0.5 are indicated by a circle in Fig. 1(a). Uncertainties in location are determined from the standard deviations of parameters registered in the JMA catalog.

4. Results

Figure 2(a) and 2(b) compares the cumulative numbers of target earthquakes from 1 January 2004 to 31 December 2008 with those expected at the end of each year based on sequences conforming to the Hazmap and EEPAS models. The uncertainty ranges of two standard deviations are indicated with error bars. The solid (gray, broken) lines plot the cumulative numbers of earthquakes ($M \geq 5.0$) for the cases without (with) errors in magnitude. These figures indicate that each model is self-consistent from the viewpoint of earthquake productivity, with the cumulative numbers falling within the uncertainty ranges in four of the five comparisons made.

Figure 3(a) and 3(b) compares the likelihoods of the

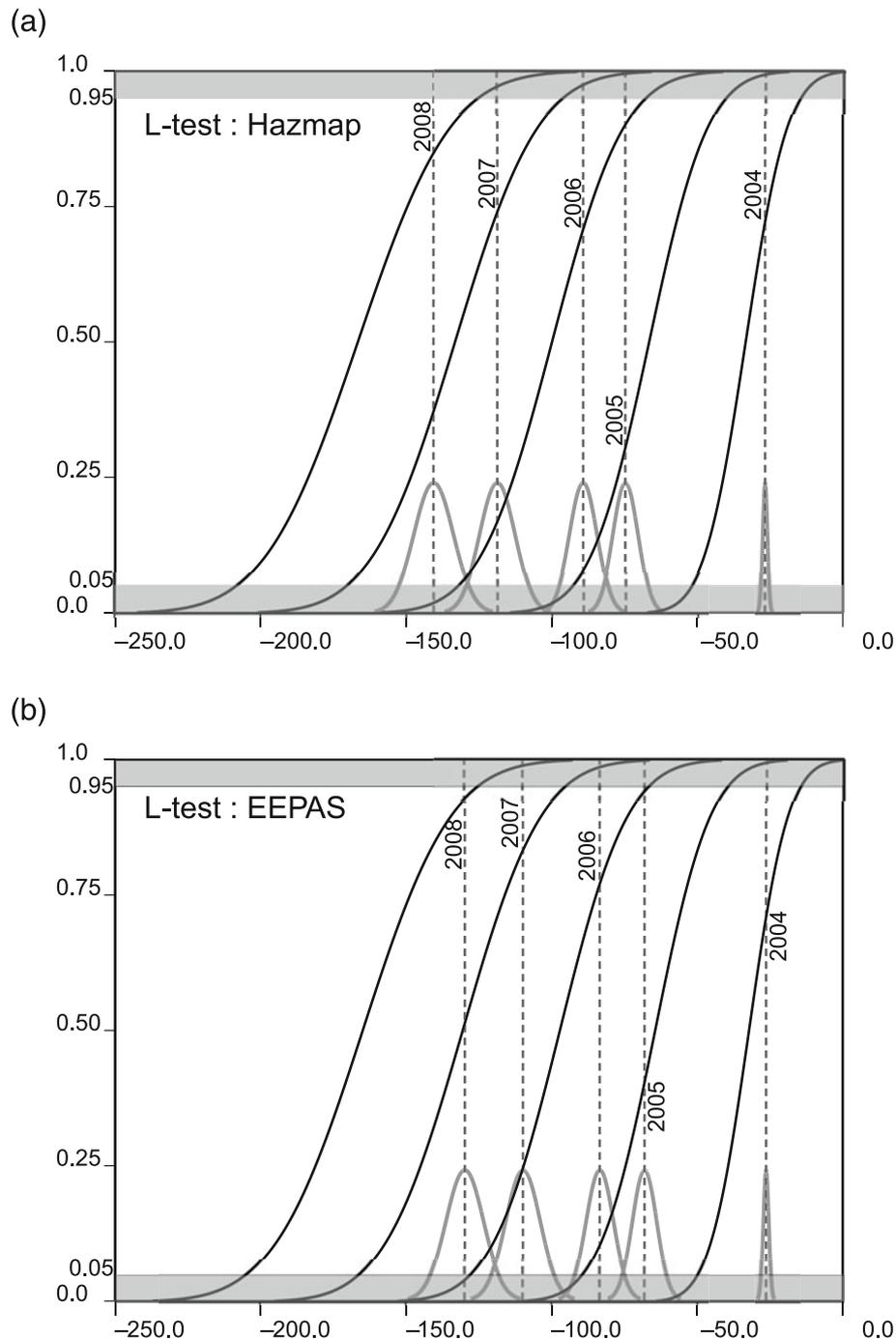


Fig. 3. L-test for January 2004 to December 2008. (a) Observed likelihoods of the Hazmap model are compared with the distributions of likelihoods calculated for sequences conforming to the Hazmap model. Five sets of charts depict likelihoods obtained for the year 2000 to the end of each respective year. The vertical broken line indicates the likelihood estimated with the observed earthquakes. The distribution of likelihood expected from the model is indicated with a solid line. The upper and lower 5% ranges of expected values are indicated with shading. Each year is labeled at the intersection of two lines. The ranges of the observed likelihoods due to uncertainties in the earthquake source parameters are given in normal density function form (arbitrary units) at the bottom. (b) The same as Fig. 3(a) but for the EEPAS model.

Hazmap and EEPAS models calculated from the observations with the distributions of those calculated for sequences conforming to the proposed models. Five sets of charts present the likelihoods obtained for the year 2004 to the end of the respective year being compared. The vertical broken line indicates the likelihood estimated with Eq. (16). The distribution of likelihoods expected from Eqs. (11) and (12) is indicated with a solid line. The upper and lower 5% rejection regions of expected values are indicated with shading.

Each year is indicated at the intersection of two lines. The ranges of the observed likelihood expected from Eqs. (16) and (17) resulting from uncertainties in the hypocentral parameters are given in normal density function form (arbitrary units) at the bottom. These figures demonstrate that each observed likelihood is within the acceptance range of expected values and suggest that neither model is rejected by the L-test.

Figure 4(a) compares the difference in the log-likelihood

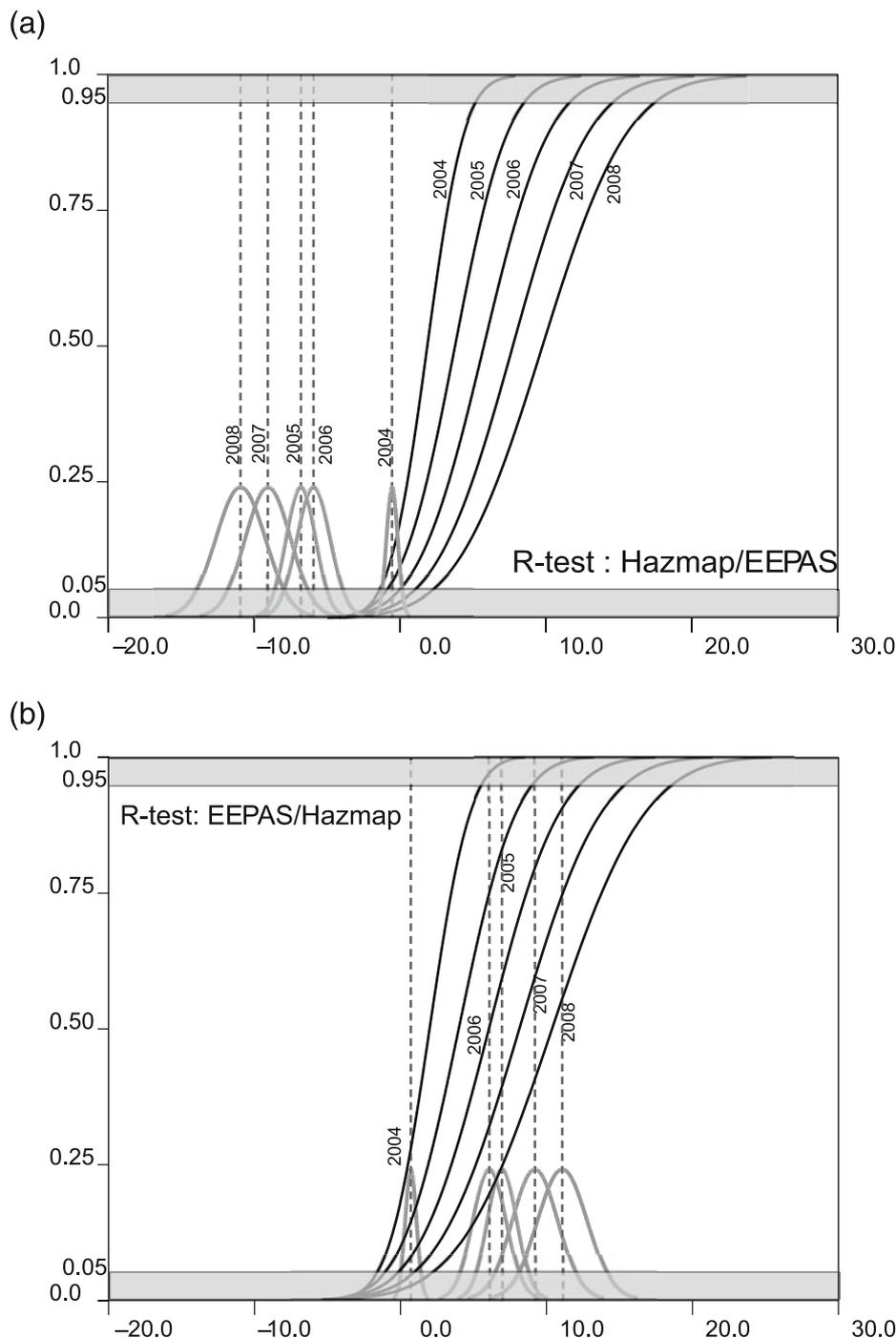


Fig. 4. (a) R-test of the Hazmap model versus the EEPAS model, where a sequence of target earthquakes conforming to the Hazmap model is assumed. See Fig. 3 for the notations. (b) R-test of the EEPAS model versus the Hazmap model, where a sequence of target earthquakes conforming to the EEPAS model is assumed.

(R-score) between the Hazmap and EEPAS models over the testing period with that expected from the Hazmap model. Vertical broken lines indicate the R-scores of the observed sequence of earthquakes at the end of respective years, and solid lines indicate the expected distributions from sequences conforming to the Hazmap model. The upper and lower 5% rejection ranges of expected values are indicated with shading. The intersections of broken lines and solid curves primarily remain below the 5% level except for the year 2004. The null hypothesis that earthquake occurrences conform to the Hazmap model can be rejected with more

than 95% confidence, in favor of the EEPAS model. This conclusion stands even if uncertainties of the parameters are taken into account.

Figure 4(b) compares the observed R-score with that expected from distributions under the EEPAS model. The intersections of broken lines and solid curves remain in the acceptance range (between 5% and 95%). This would be maintained even if we consider the extreme value of R-score that is estimated as side lobes of the normal density function at the bottom. Based on this result, the null hypothesis that earthquake occurrences conform to the EEPAS model

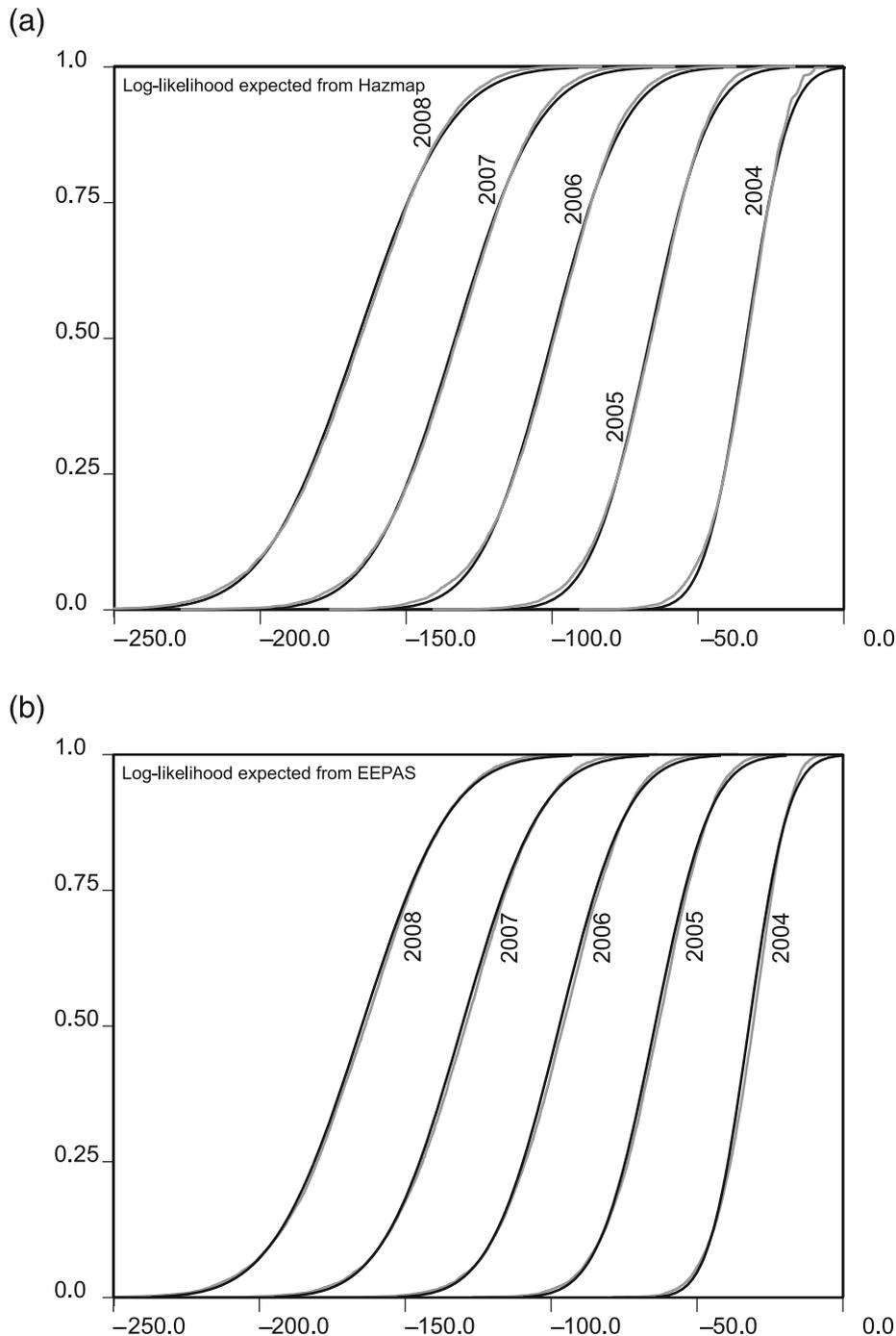


Fig. 5. (a) Comparisons of log-likelihood distributions between those obtained with the proposed formulas (dark lines) and those from 10,000 simulated catalogs (light lines). Both cases conform to the Hazmap model. (b) Same as for 5(a) but for the EEPAS model.

rather than the Hazmap model cannot be rejected.

5. Discussion

In this paper, we have derived the means and variances of the distributions for N-, L-, and R-scores in the cases of (1) seismicity compatible with the proposed model and (2) catalogs with errors in the observed parameters. With these means and variances, N-, L-, and R-tests could be performed for the proposed models in a simple way if distributions associated with the tests are well approximated by normal distributions with means and variances. This approximation is basically guaranteed by the central-limit

theorem if the number of earthquakes is sufficiently large. For the first case in Section 2.1.1, the Poisson distribution is approximated with a normal distribution. This approximation is likely accomplished with a total expected number of earthquakes exceeding 10.

The next case, presented in Section 2.1.2, is that in which binomial distributions are approximated by a normal distribution, and it should be carefully considered. One of the various rules that may be used to decide whether a sample size is large enough for this approximation is that both the expected value and the value of the sample size minus the expected value must be greater than 5. If we consider a case

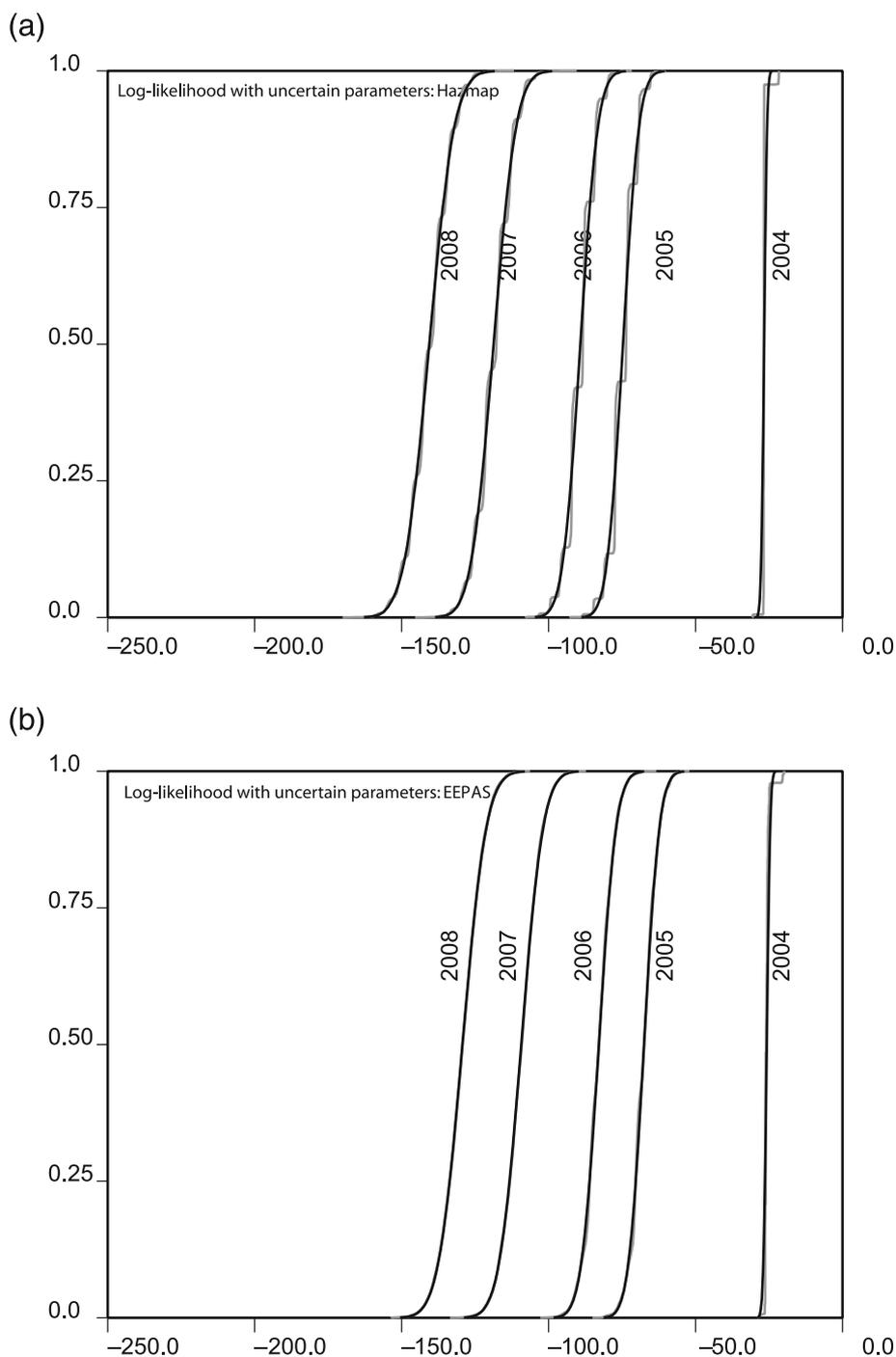


Fig. 6. (a) Comparisons of log-likelihood distributions for the catalog with uncertain parameters. Likelihood distributions are estimated with the proposed formulas (dark lines) and obtained from 10,000 simulated catalogs (light lines). The Hazmap model is used in both cases. (b) Same as 6(a) but for the EEPAS model.

of only errors in location, only a small proportion of earthquakes contributes to the variance of $n(\Omega^0)$ in Eq. (5), i.e., only earthquakes near the border of the test area contribute to the variance since only such events are origins of changes in $n(\Omega^0)$.

Therefore, the above rule for the approximation may not be satisfied. However, if we consider the case of errors both in locations and magnitude, the rule must be satisfied in most cases. For example, provided that we observe ten earthquakes with magnitudes exceeding the cutoff M_c in the test area and the standard deviation of the magnitude

determination is 0.2, we expect about 21 earthquakes in the magnitude range between $M_c - 0.4$ and $M_c + 0.4$, where the b -value of the Gutenberg-Richter relation is assumed to be 1. We could expect seven earthquakes larger than M_c and 14 less than this, which may satisfy the above rule.

In our case, we presume the standard deviation of the magnitude determination to be 0.1, and earthquakes of magnitude ≥ 4.7 are considered in our test. In total, 51 events are used, among which 24 events are registered with magnitude < 5.0 . The expected number of events for the five years is 25.4 and is given as the summation of the probabil-

Table 2. Summaries of means and standard deviations of L-score. (a) Those obtained from earthquakes conforming to the Hazmap model (top) and the EEPAS model (bottom) by 10,000 simulation catalogs (left: simulation) and the proposed formulas (right: analytical sol.). (b) Those obtained from earthquakes perturbed with uncertainties in hypocenter parameters.

| (a) | | | | | |
|--------|----------------|------|-----------------|------|-----|
| Hazmap | Simulation | | Analytical Sol. | | Std |
| | E[λ] | Std | E[λ] | Std | |
| 2004 | -33.4 | 11.4 | -33.3 | 11.3 | |
| 2005 | -66.6 | 16.0 | -66.5 | 16.0 | |
| 2006 | -99.9 | 19.5 | -99.9 | 19.5 | |
| 2007 | -133.0 | 22.5 | -133.2 | 22.6 | |
| 2008 | -166.3 | 25.1 | -166.6 | 25.2 | |
| EEPAS | | | | | |
| 2004 | -32.0 | 10.5 | -32.1 | 10.6 | |
| 2005 | -64.2 | 15.0 | -64.4 | 15.1 | |
| 2006 | -96.9 | 18.4 | -97.1 | 18.5 | |
| 2007 | -130.3 | 21.4 | -130.5 | 21.4 | |
| 2008 | -164.7 | 24.0 | -164.9 | 24.1 | |

| (b) | | | | | |
|--------|----------------|-----|-----------------|-----|-----|
| Hazmap | Simulation | | Analytical Sol. | | Std |
| | E[λ] | Std | E[λ] | Std | |
| 2004 | -26.5 | 0.8 | -26.5 | 0.8 | |
| 2005 | -74.6 | 4.3 | -74.5 | 4.3 | |
| 2006 | -89.1 | 4.7 | -89.0 | 4.7 | |
| 2007 | -118.5 | 6.1 | -118.6 | 6.1 | |
| 2008 | -140.3 | 6.8 | -140.4 | 6.8 | |
| EEPAS | | | | | |
| 2004 | -25.9 | 0.9 | -25.9 | 0.9 | |
| 2005 | -67.7 | 4.2 | -67.7 | 4.2 | |
| 2006 | -83.1 | 4.7 | -83.0 | 4.7 | |
| 2007 | -109.4 | 5.9 | -109.4 | 5.9 | |
| 2008 | -129.4 | 6.5 | -129.4 | 6.4 | |

ities in the last column of Table 1. Accordingly, the sample size minus the expected value becomes 25.6, which is much greater than 5.0 and satisfies the above rule.

Imoto (2009) presented an example in which a distribution of L-scores for a proposed model is well approximated by a normal distribution, in that the expected number of earthquakes is ten. Figure 5(a) (5(b)) compares the likelihood distributions expected from the Hazmap (EEPAS) model between those derived by our formula (dark lines) and those obtained from 10,000 simulated catalogs (light lines), which correspond to the method by Schorlemmer *et al.* (2007). Our Gaussian approximations fit those obtained from simulation fairly well, primarily between the upper and lower 10 percentiles. The approximations of longer periods are better than those of shorter periods (i.e., the approximation improves as the number of earthquakes increases).

Figure 6(a) compares the likelihood distributions of the Hazmap model observed for the catalog with uncertain parameters with those derived by our formula (dark line indicates a cumulative form of that in Fig. 3) and those obtained from 10,000 perturbed sequences from the original catalogs (light line). The perturbed sequences simulate the method of Schorlemmer *et al.* (2007). Figure 6(b) presents the results for the EEPAS model. It is quite difficult to distinguish between the two lines except in a few segments. Therefore, we can conclude that our method estimates the effects due to uncertainties of parameters fairly well.

Table 2(a) summarizes the means and the standard deviations of the distributions in Fig. 5(a) and 5(b). The differences in the means between simulated catalogs and analytical solutions are mostly <0.2 , and differences in the standard deviations are <0.1 . Table 2(b) summarizes the means and the standard deviations of the distributions in Fig. 6(a) and 6(b). The differences in both the means and the standard deviations are at most 0.1. These results suggest that the L-test conducted using our method should lead to similar results to that using simulated catalogs (Schorlemmer *et al.*, 2007).

In summary, we present here methods by which to perform N-, L-, and R-tests without generating random catalogs conforming to the test models or catalogs modified from the observed catalog with uncertainties in the parameters. We applied the proposed method to the Hazmap and EEPAS models in Kanto, central Japan. The N- and L-tests for the years 2004–2008 confirm the self-consistency of both models. The values for the R-test for the last 5 years suggest that the EEPAS model is superior to the Hazmap model over this period. This 5-year comparison is in no way a meaningful test of the Hazmap model itself, which is a long-term model designed for a time period of 30 years. Rather, it suggests that the time-varying EEPAS model contains information about earthquake occurrence on a 5-year timescale that is not contained in longer term estimates of seismicity.

Our comparison of L-scores derived analytically with

those derived from simulated catalogs indicates no significant difference between the two sets of scores, thus implying that the proposed method is an alternative to the current one. However, caution is warranted because the analytical method is only reliable when the assumptions on which it is based are satisfied.

Acknowledgments. This manuscript was greatly improved by the comments of an anonymous reviewer and K. Nanjo.

References

- Evison, F. F. and D. A. Rhoades, Precursory scale increase and long-term seismogenesis in California and northern Mexico, *Ann. Geophys.*, **454**, 479–495, 2002.
- Evison, F. F. and D. A. Rhoades, Demarcation and scaling of long-term seismogenesis, *Pure Appl. Geophys.*, **161**, 21–45, 2004.
- Fujiwara, H., National seismic hazard mapping project of Japan, *Proceedings of the 5th U.S.-Japan Natural Resources Meeting, U.S. Geological Survey, Open-File Report 2005-1131*, 14, 2004.
- Fujiwara, H., S. Kawai, S. Aoi, N. Morikawa, S. Senna, N. Kudo, M. Ooi, K. Hao, K. Wakamatsu, Y. Ishikawa, T. Okumura, T. Ishii, S. Matsushima, Y. Hayakawa, N. Toyama, and A. Narita, A study on “National seismic hazard maps for Japan”, *Technical Note National Research Institute for Earth Science and Disaster Prevention*, No336, 2009 (in Japanese).
- Imoto, M., Comments on the N-, L- and R-tests for seismicity models, *Zisin 2*, **61**, 207–209, 2009 (in Japanese).
- Imoto, M. and D. A. Rhoades, Seismicity models of moderate earthquakes in Kanto, Japan, utilizing multiple predictive parameters, *Pure Appl. Geophys.*, **167**, 831–843, 2010.
- JMA, *The Seismological and Volcanological Bulletin of Japan*, Japan Meteorological Agency, ISSN 1349-8320, 2009.
- Jordan, T. H., Earthquake predictability, brick by brick, *Seismol. Res. Lett.*, **77**, 3–6, 2006.
- Kagan, Y. Y. and D. D. Jackson, New seismic gap hypothesis: five year after, *J. Geophys. Res.*, **100**, 3943–3960, 1995.
- NIED, J-SHIS at http://www.wold.j-shis.bosai.go.jp/j-shis/index_en.html (as of July 1, 2010).
- Rhoades, D. A. and F. F. Evison, Long-range earthquake forecasting with every earthquake a precursor according to scale, *Pure Appl. Geophys.*, **161**, 47–72, 2004.
- Rhoades, D. A. and F. F. Evison, Test of the EEPAS forecasting model on the Japan earthquake catalogue, *Pure Appl. Geophys.*, **162**, 1271–1290, 2005.
- Rhoades, D. A. and F. F. Evison, The EEPAS forecasting model and the probability of moderate-to-large earthquakes in central Japan, *Tectonophysics*, **417**, 119–130, 2006.
- Schorlemmer, D., M. C. Gerstenberger, S. Wiemer, D. D. Jackson, and D. A. Rhoades, Earthquake likelihood model testing, *Seismol. Res. Lett.*, **78**, 17–29, 2007.

M. Imoto (e-mail: imoto@bosai.go.jp), D. A. Rhoades, H. Fujiwara, and N. Yamamoto