Chapter 8

Daily Variation in Earthquake Detection Capability: A Quantitative Evaluation

Takaki Iwata

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/54890

1. Introduction

Evaluating the capability for detection of an earthquake catalogue is the first step in statistical seismicity analysis. In recent years, many studies have focused on ways to assess the completeness magnitude (M_c), the lowest magnitude level at which all earthquakes are recorded and there are no missing earthquakes, in a global earthquake catalogue [1, 2], regional catalogues [3–6], and small-scale observations such as underground mines [7]. We cannot make full use of the available information in an earthquake catalogue if we overestimate M_c, and underestimation of M_c yields incorrect or biased results.

[8] is a representative example showing that an accurate choice of M_c is vital in a seismicity analysis. This study examined the global earthquake catalogue provided by the National Oceanic and Atmospheric Administration (NOAA), United States Department of Commerce, and reported that the number of earthquakes during nighttime was significantly higher than that during daytime. Some studies [9, 10] stated that the daily variation in seismic activity [8] had found was just a bias; [8] did not take into account the daily change in detection capability, and earthquakes with small magnitudes — for which records were incomplete — were used in the analysis.

Empirically, it is well known that the level of cultural noise has daytime/nighttime variations and that this causes a daily variation in the detection capability. Nonetheless, in many studies, the estimation of M_c or the evaluation of the detection capability is carried out for a catalogue ranging over several days, weeks, months or years without taking short-time variation into consideration; M_c is underestimated in daytime and overestimated in nighttime, compared with the true value of M_c.

For an appropriate and precise evaluation of the detection capability, it is vital to quantify this temporal change. In this chapter, we propose a statistical method for a quantitative evaluation of the daily variation in the detection capability and also present an example of

the method in real earthquake sequences. A short summary of this chapter has already been presented in a non-peer reviewed article [11], and this chapter focuses on showing the details of the approach and results that appeared in that article.

2. Data

The earthquake data used in this study were retrieved from the Japan Meteorological Agency (JMA) catalogue from January 2006 to December of 2010, with focal depths shallower than 30 km. The magnitude scale used in the JMA catalogue has been termed the "JMA magnitude", which is determined on the basis of velocity and displacement seismograms [12, 13].

The detection capability of earthquakes has regional variations, and in particular there is clear difference between inland and offshore regions. Because the main interest of this study is the temporal variation in the detection capability, to mitigate the influence of the regional differences, this study used the events occurring within the "Mainland" area defined by [14], which covers the inland and coastal regions of Japan (Figure 1). The total number of the events is 331,537.

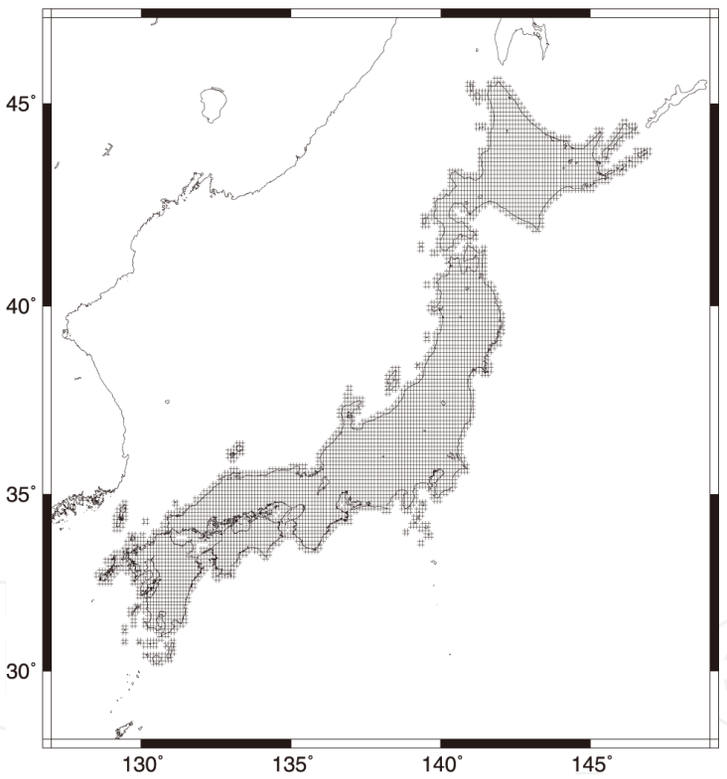

Figure 1. "Mainland" defined by [14], which covers the inland and coastal regions of Japan (after [14]).

Because the main object of the present study is to investigate the daily variation in earthquake detection capability, the data were divided into periods of one day, and the one-day sequences were stacked. The technical limitations of computer memory space and computing time make it difficult to handle hundreds of thousands events simultaneously in a Bayesian analysis, as described in the next section. Thus, the stacked sequences were constructed for each of the five years; five datasets were analyzed to examine the daily change in the detection capability.

http://dx.doi.org/10.5772/54890

3. Statistical method to evaluate the daily variation in detection capability of earthquakes

3.1. Evaluation of the detection capability by using a statistical model to represent an observed magnitude-frequency distribution

[15] proposed a probability density function $f(M)$ for an observed magnitude-frequency distribution of earthquakes over all magnitude range. The probability density function is assumed to be the product of two probability distributions.

The first comes from the Gutenberg-Richter (GR) law [16], which is equivalent to an exponential distribution:

$$w(M|\beta) = \beta \exp(-\beta M), \tag{1}$$

where the parameter β is related to the b-value of the GR-law and their relationship is described by $\beta = b \ln 10$.

The second is a detection rate function $q(M)$ showing the proportion of detected earthquakes to all earthquakes at magnitude M. Following the proposal by [17], the cumulative distribution function of a normal distribution is used as the detection probability function:

$$q(M|\mu,\sigma) = \int_{-\infty}^{M} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \tag{2}$$

We normalize the product of the two functions, and then we obtain the target probability density function as follows (Figure 2):

$$\begin{aligned} f(M|\mu,\beta,\sigma) &= \frac{w(M|\beta)q(M|\mu,\sigma)}{\int_{-\infty}^{\infty} w(M|\beta)q(M|\mu,\sigma)dM} \\ &= \exp(-\beta M)q(M|\mu,\sigma) \cdot \beta \exp\left[\beta\mu - \frac{\beta^2\sigma^2}{2}\right]. \end{aligned} \tag{3}$$

Of the three parameters (μ, β, and σ) in this statistical model, the parameter μ has the closest connection with the detection capability; μ indicates the magnitude at which the detection probability is 50 %. This means that the detection capability is better as the value of μ is smaller.

The values of the three parameters are estimated by the maximum likelihood method. The log-likelihood function ($\ln L$) is given by

$$\ln L(\beta,\mu,\sigma) = \sum_{i}^{N} \ln f(M_i|\beta,\mu,\sigma), \tag{4}$$

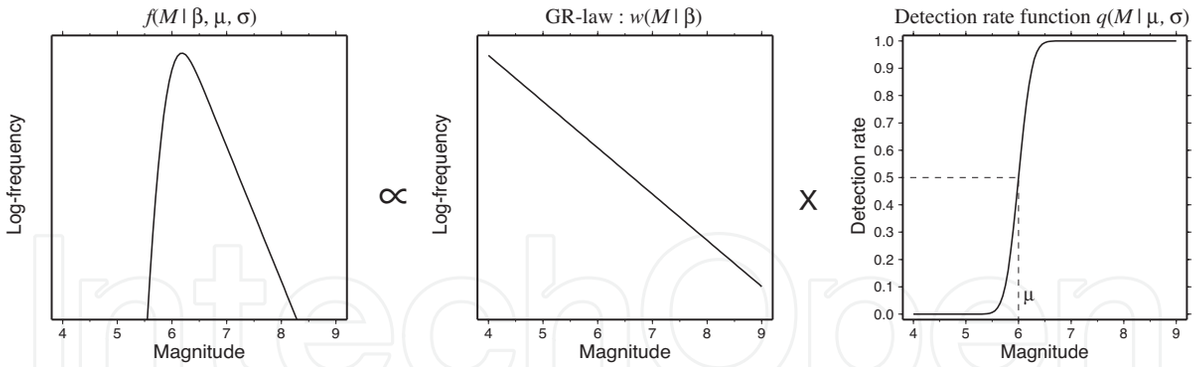

Figure 2. Schematic diagram showing the construction of the probability density function for an observed magnitude-frequency distribution of earthquakes over all magnitude range, as proposed by [15].

where N and M_i denote the number of the analyzed earthquakes and the magnitude of the i-th earthquake, respectively. The set of the values of the three parameters maximizing the log-likelihood function is the best estimate.

To observe the performance of this statistical model, we applied it to the magnitude-frequency distribution of the earthquakes taken from the JMA catalogue described in the previous section. As observed in Figure 3, the estimated curve of the statistical model fits well with the data, suggesting that this model is appropriate to describe the magnitude-frequency distribution on the JMA magnitude scale.

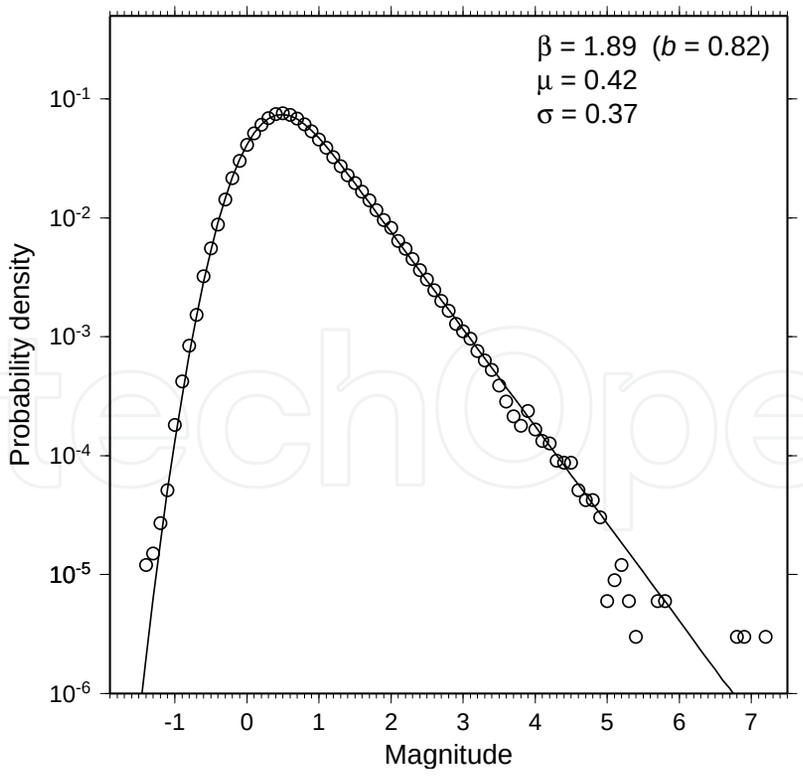

Figure 3. Observed (open circles) and estimated (solid line) probability density functions of the magnitudes of all the shallow (depth $\leq$ 30 km) earthquakes within the "Mainland" area [14], from January 2006 to December 2010. The best estimates of β (or b-value), μ, and σ are shown in the top right corner.

http://dx.doi.org/10.5772/54890

As a preliminary analysis to examine the daily variation, this model was applied to the datasets of nighttime (0.0–0.2 days) and daytime (0.4–0.6 days) earthquakes; we found a clear shift of the detection capability between the two datasets (Figure 4).

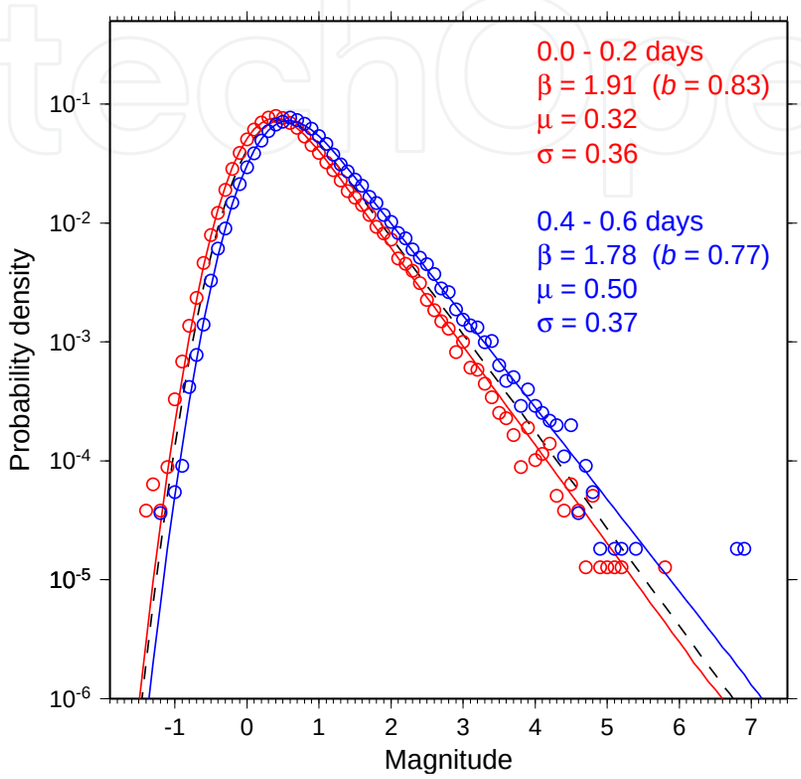

Figure 4. Observed (open circles) and estimated (solid line) probability density functions of the magnitude of nighttime (0.0–0.2 days, red) and daytime (0.4–0.6 days, blue) earthquakes. The dotted black line indicates the estimated probability density function for all the shallow earthquakes used in this study, as shown in Figure 3.

3.2. Evaluation of the daily variation of the detection capability in a Bayesian framework

To quantify the daily variation of the detection capability, we estimated the temporal change in the three parameters μ, β, and σ that appeared in the model. The procedure of the estimation is similar to that used in [2, 18].

To represent the temporal variations in the parameters, we introduced a piecewise linear function or linear spline [19]. The nodal point of the spline were taken at each of the occurrence times of the earthquakes in a stacked sequence. For the sake of brevity, we defined $\theta_i^{(1)}$, $\theta_i^{(2)}$, and $\theta_i^{(3)}$ as the values of β, μ, and σ, respectively, at the occurrence time t_i of the i-th earthquake. Hence, the temporal variations of $\mu(t), \beta(t)$, and $\sigma(t)$ were represented by

$$\phi_1(t) = \beta(t) = \frac{\theta_{i+1}^{(1)} - \theta_i^{(1)}}{t_{i+1} - t_i}(t - t_i) + \theta_i^{(1)} \ ,$$
$$\phi_2(t) = \mu(t) = \frac{\theta_{i+1}^{(2)} - \theta_i^{(2)}}{t_{i+1} - t_i}(t - t_i) + \theta_i^{(2)} \ ,$$
$$\phi_3(t) = \sigma(t) = \frac{\theta_{i+1}^{(3)} - \theta_i^{(3)}}{t_{i+1} - t_i}(t - t_i) + \theta_i^{(3)} \quad \text{for } t_i \leq t < t_{i+1}. \tag{5}$$

The goal of the estimation is the optimization of the following parameters:

$$\begin{aligned} \boldsymbol{\theta} &= (\boldsymbol{\theta}^{(1)}; \boldsymbol{\theta}^{(2)}; \boldsymbol{\theta}^{(3)}) \\ &= (\theta_1^{(1)}, \ldots, \theta_N^{(1)}; \theta_1^{(2)}, \ldots, \theta_N^{(2)}; \theta_1^{(3)}, \ldots, \theta_N^{(3)}). \end{aligned} \tag{6}$$

The simultaneous optimization of such a large number ($= 3N$) of parameters is, however, an unstable process; we incorporated a smoothness constraint or roughness penalty on $\phi_i(t)(i = 1, 2, 3)$ to enhance the stability of the optimization.

Following the equation (4), the log-likelihood function in this case is given by

$$\ln L(\boldsymbol{\theta}) = \sum_i^N \ln f(M_i | \theta_i^{(1)}, \theta_i^{(2)}, \theta_i^{(3)}), \tag{7}$$

and the smoothness constraint is quantified by the following equation:

$$\Phi(\boldsymbol{\theta} | v_1, v_2, v_3) = \sum_{j=1}^{3} v_j \int_{T_s}^{T_e} \left[\frac{\partial}{\partial t} \phi_j(t) \right]^2 dt, \tag{8}$$

where v_j is the parameter controlling the trade-off between the goodness-of-fit of the statistical model to the data and the smoothness constraint, and $[T_s, T_e]$ denotes the domain of the analyzed time period.

In the analysis of an ordinary earthquake sequence such as [2, 18], the start and end points (T_s and T_e) are different but in the present case they are the same, midnight because the stacked one-day data are analyzed. Thus, the values of β, μ, and σ at T_s and T_e should connect smoothly. Considering this property and the equation (5) we rewrite the equation (8) as follows:

$$\Phi(\boldsymbol{\theta} | v_1, v_2, v_3) = \sum_{j=1}^{3} v_j \left[\sum_{i=1}^{N-1} \frac{(\theta_{i+1}^{(j)} - \mu_i^{(j)})^2}{t_{i+1} - t_i} + \frac{(\theta_1^{(j)} - \theta_N^{(j)})^2}{(t_1 - T_s) + (T_e - t_N)} \right] \tag{9}$$

Then, we introduced a penalized log-likelihood function [20, 21]

$$Q(\boldsymbol{\theta}|v_1, v_2, v_3) = \ln L(\boldsymbol{\theta}) - \Phi(\boldsymbol{\theta}|v_1, v_2, v_3), \tag{10}$$

and the maximization of $Q(\boldsymbol{\theta}|v_1, v_2, v_3)$ provides the best estimate of $\boldsymbol{\theta}$, and it depends on the values of $v_j (j = 1, 2, 3)$.

A Bayesian framework with the type II maximum likelihood approach [22] or the maximization of marginal likelihood [23] enables us to determine the values of v_j's objectively. We supposed that the prior distribution of $\boldsymbol{\theta}$ corresponding to the smoothness constraint is proportional to $\exp[-\Phi(\boldsymbol{\theta}|v_1, v_2, v_3)]$; the equation (9) and the consideration of the normalizing constant give us the prior distribution $\pi(\boldsymbol{\theta}|v_1, v_2, v_3)$ as follows:

$$\begin{aligned}
&\pi(\boldsymbol{\theta}|v_1, v_2, v_3) \\
&= \prod_{j=1}^{3} \prod_{i=1}^{N-1} \sqrt{\frac{v_j}{\pi(t_{i+1} - t_i)}} \exp\left[-\frac{v(\phi_{i+1}^{(j)} - \phi_i^{(j)})^2}{t_{i+1} - t_i} \right] \\
&\cdot \sqrt{\frac{v_j}{\pi(t_1 - T_s + T_e - t_N)}} \exp\left[-\frac{v(\phi_1^{(j)} - \phi_N^{(j)})^2}{t_1 - T_s + T_e - t_N} \right].
\end{aligned} \tag{11}$$

To find the marginal likelihood [24], we need to integrate the product of the prior distribution $\pi(\boldsymbol{\theta}|v_1, v_2, v_3)$ and the likelihood function $L(\boldsymbol{\theta})$ shown in the equation (7) over $\boldsymbol{\theta}$, but this integration is unachievable. This is because the integral of $\pi(\boldsymbol{\theta}|v_1, v_2, v_3)$ over $\boldsymbol{\theta}$ is infinite and is a so-called improper prior.

To deal with this computational problem we isolated $\theta_N^{(j)} (j = 1, 2, 3)$ from $\boldsymbol{\theta}$, because the integral of $\pi(\boldsymbol{\theta}|v_1, v_2, v_3)$ over

$$\begin{aligned}
\boldsymbol{\theta}_{-N} &= (\boldsymbol{\theta}_{-N}^{(1)}; \boldsymbol{\theta}_{-N}^{(2)}; \boldsymbol{\theta}_{-N}^{(3)}) \\
&= (\theta_1^{(1)}, \ldots, \theta_{N-1}^{(1)}; \theta_1^{(2)}, \ldots, \theta_{N-1}^{(2)}; \theta_1^{(3)}, \ldots, \theta_{N-1}^{(3)}).
\end{aligned} \tag{12}$$

is finite; we rewrote the original prior distribution $\pi(\boldsymbol{\theta}|v_1, v_2, v_3)$ as $\pi_{-N}(\boldsymbol{\theta}_N|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ which is a proper prior with respect to $\boldsymbol{\theta}_{-N}$.

Then, to obtain the marginal likelihood $\mathcal{L}$, we integrated out the product of the prior distribution $\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ and the likelihood function $L(\boldsymbol{\theta})$ over $\boldsymbol{\theta}_{-N}$:

$$\mathcal{L}(v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) = \int_{\Theta} L(\boldsymbol{\theta}) \pi_{-N}(\boldsymbol{\theta}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) d\boldsymbol{\theta}_{-N}, \tag{13}$$

where Θ denotes the parameter space of $\boldsymbol{\theta}_{-N}$.

We intended to find the set of the values of $v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}$, and $\theta_N^{(3)}$ which maximizes the marginal likelihood, because such a set is the best estimate of the six parameters [22, 23].

In this hierarchical Bayesian scheme, the six parameters are often called as hyperparemeters which govern the prior distribution.

The maximization of the marginal likelihood $\mathcal{L}(v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ was achieved in the following way. In the first step, we intend to maximize the integrand $\ln L(\boldsymbol{\theta})\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ in the equation (13) with respect to $\boldsymbol{\theta}_{-N}$, which is equivalent to the maximization of the penalized log-likelihood function $Q(\boldsymbol{\theta}|v_1, v_2, v_3)$ appeared in the equation (10).

The logarithm of the integrand in the equation (13) is approximated by a quadratic form at the initial value of $\boldsymbol{\theta}_{-N} = \boldsymbol{\theta}_{-N0}$ (and $\boldsymbol{\theta} = \boldsymbol{\theta}_0$):

$$\begin{aligned}
&\ln L(\boldsymbol{\theta})\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) \\
\approx\ &\ln L(\boldsymbol{\theta}_0)\pi_{-N}(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) \\
&+\boldsymbol{g}(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) \cdot (\boldsymbol{\theta}_{-N} - \boldsymbol{\theta}_{-N0}) \\
&-\frac{1}{2}(\boldsymbol{\theta}_{-N} - \boldsymbol{\theta}_{-N0})H(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})(\boldsymbol{\theta}_{-N} - \boldsymbol{\theta}_{-N0})^T,
\end{aligned} \tag{14}$$

where $\boldsymbol{g}(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ and $H(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ are the gradient vector and the negative of the Hessian matrix (second partial derivatives) of the integrand at $\boldsymbol{\theta}_{-N} = \boldsymbol{\theta}_{-N0}$, respectively. The symbol T denotes the transpose of a vector (or a matrix).

If the choice of the initial value $\boldsymbol{\theta}_{-N0}$ is appropriate and is not far from $\boldsymbol{\theta}_{-N} = \hat{\boldsymbol{\theta}}_{-N}$ which maximizes the integrand, the quadratic form is an upward convex and $H(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ is a positive definite. Hence, Cholesky decomposition [25] enables us to factorize the matrix $H(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ into the following form:

$$H(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) = AA^T, \tag{15}$$

where A is a lower triangular matrix.

The quadratic form is maximized at $\boldsymbol{\theta}_{-N} = \boldsymbol{\theta}'_{-N0}$ which satisfies

$$A^T(\boldsymbol{\theta}_{-N} - \boldsymbol{\theta}_{-N0}) = A^{-1}\boldsymbol{g}(\boldsymbol{\theta}_{-N0}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) \tag{16}$$

through the Newton method. We replace $\boldsymbol{\theta}_{-N0}$ by $\boldsymbol{\theta}'_{-N0}$ and iterate the procedure until $\boldsymbol{\theta}_{-N}$ converges to find $\boldsymbol{\theta}_{-N} = \hat{\boldsymbol{\theta}}_{-N}$.

At $\boldsymbol{\theta}_{-N} = \hat{\boldsymbol{\theta}}_{-N}$ where the integrand $\ln L(\boldsymbol{\theta})\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ is maximized, the second term of the equation (14) vanishes because $\boldsymbol{g}(\hat{\boldsymbol{\theta}}_{-N}|v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ is a zero vector. Hence, around $\boldsymbol{\theta}_{-N} = \hat{\boldsymbol{\theta}}_{-N}$, the integrand is approximated by

http://dx.doi.org/10.5772/54890

$$\begin{aligned}
&\ln L(\boldsymbol{\theta})\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)}) \\
&\approx \ln L(\hat{\boldsymbol{\theta}}_0)\pi_{-N}(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)}) \\
&\quad -\frac{1}{2}(\boldsymbol{\theta}_{-N}-\hat{\boldsymbol{\theta}}_{-N})H(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})(\boldsymbol{\theta}_{-N}-\hat{\boldsymbol{\theta}}_{-N})^T
\end{aligned} \tag{17}$$

from the equation (14).

Then, in the second step, the integral in the equation (13) is computed by the Laplace approximation [26]. Consequently, the log marginal likelihood $\ln \mathcal{L}(v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})$ is given by

$$\begin{aligned}
&\ln \mathcal{L}(v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)}) \\
&\approx \ln L(\hat{\boldsymbol{\theta}}_{-N})\pi(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)}) \\
&\quad -\frac{1}{2}\ln \det H(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)}) + \frac{n}{2}\ln 2\pi,
\end{aligned} \tag{18}$$

where n is the number of parameters included in $\hat{\boldsymbol{\theta}}_{-N}$ and $n = 3(N-1)$ in this case.

By repeating the two steps, we attempt to find the values of the six hyperparameters which maximize the value of $\ln \mathcal{L}(v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})$. Eventually, we can find the optima of the hyperparameters and $\boldsymbol{\theta}_{-N}$.

With a sufficiently large n, the quadratic and Laplace approximations are good [27] and $\ln L(\boldsymbol{\theta})\pi_{-N}(\boldsymbol{\theta}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})$ is approximately a multidimensional Gaussian distribution. In this case, the inverse of the negative of the Hessian matrix $H(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})^{-1}$ gives the estimation errors of the parameters [15]; the $[i+j(N-1)]$-th diagonal component of $H(\hat{\boldsymbol{\theta}}_{-N}|v_1,v_2,v_3,\theta_N^{(1)},\theta_N^{(2)},\theta_N^{(3)})^{-1}$ is the standard error of $\theta_i^{(j)}$.

3.3. Model comparison

So far, we assumed that each of the parameters β, μ, and σ has daily variations. For more appropriate statistical modelling, however, it is necessary to examine the significance of the temporal variations of the three parameters.

To apply the model where at least one of the three parameters β, μ, and σ is not assumed to have the daily variation to data, we fix v_j('s) at 0 where j takes the value(s) corresponding to the parameter(s) without the temporal variation. Then the procedure described in section 3.2 is conducted, but we exclude $\boldsymbol{\theta}^{(j)}$ in the equation (6) and consider this exclusion in the procedure following the equation.

The possbile combination of the allowance or non-allowance of the temporal variation in β, μ, and σ yields eight cases in total. We introduce Akaike's Bayesian Information Criterion ABIC [23] in the comparison of the goodness-of-fit of the eight cases to data:

$$\begin{aligned} \text{ABIC} = &-2(\text{maximum ln } \mathcal{L}(v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)}) \\ &+2(\text{number of non-fixed hyperparameters}), \end{aligned} \tag{19}$$

where $\ln \mathcal{L}(v_1, v_2, v_3, \theta_N^{(1)}, \theta_N^{(2)}, \theta_N^{(3)})$ is the marginal likelihood given by the equation (18). The model with smaller ABIC value is considered to have a better fit to data.

4. Results of the analysis of the JMA catalogue

On the basis of the comparison of the ABIC values for the eight models, we found that not only μ but also β and/or σ have the significant daily changes in the four stacked sequences except in the sequence for 2009.

Figure 5 shows the estimated daily variation of β (or b-value) found in the sequence for 2007. In this sequence the value of β drastically decreases at around 0.4 days (i.e., approximately 9:30 a.m.). On 27 March 2007 and 16 July 2007, there were two major earthquakes, the Noto Hanto earthquake with $M_w = 6.9$ and the Chuuetsu-oki earthquake with $M_w = 6.6$. Their occurrence times are coincidentally close to 10 a.m. (9:41 a.m. and 10:13 a.m. for the Noto Hanto and Chuuetsu-oki earthquakes, respectively), and these two major earthquakes are followed by active aftershock sequences (see Figure 5b). Thus, it is suspicious that the drastic decrease of β is the daily variation of our interest in this study and the two aftershock sequences would cause the decrease of β. To confirm this point the aftershock sequences following the Noto Hanto and Chuuetsu-oki earthquakes were excluded, and then the same analysis was applied; the model where only the daily variation of μ is allowed was chosen as the best model.

Year	Excluded sequence and area
2006	Earthquake swarm in Izu Peninsula (139.00-139.45°E, 34.65-35.15°N)
2007	The Noto Hanto earthquake and its aftershock sequence (136.60-137.00°E, 37.00-37.50°N) The Chuuetsu-oki earthquake and its aftershock sequence (138.30-138.80°E, 37.30-37.65°N)
2008	The Iwate-Miyagi Nairiku earthquake and its aftershock sequence (140.50-141.10°E, 38.70-39.30°N)
2010	Earthquake sequence in Nakadoori, Fukushima Prefecture (139.95-140.10°E, 37.20-37.35°N)

Table 1. Active sequences which may affect the examination of the daily variation in β, μ, and σ. The corresponding areas are shown in parentheses.

In similar to this case, from the sequences for 2006, 2008, and 2010, the active sequences listed in Table 1 were excluded and we re-analyzed them after the exclusion. Consequently, the case with only the daily variation of μ was considered as the best case for these three

http://dx.doi.org/10.5772/54890

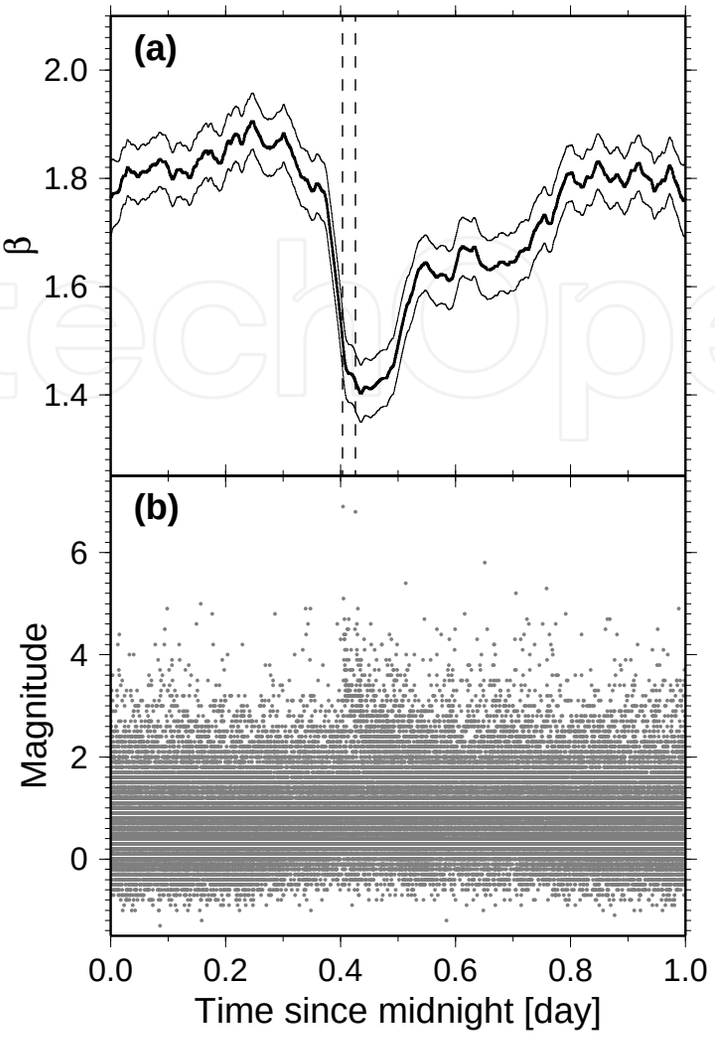

Figure 5. (a) Estimated daily variation of β for the stacked sequence for 2007 (bold line), and its two standard error bands (thin lines), as a function of the elapsed time since midnight. The two dotted lines indicate the occurrence times of the Noto Hanto and Chuuetsu-oki earthquakes. (b) Magnitude-time plot of the stacked sequence for 2007.

years, the same as that for 2007 (Table 2). As previously mentioned, in the original sequence for 2009, the case where only the daily variation of μ is allowed has been chosen as the best one. This is probably because no significant seismic activity occurred in that year.

Figure 6 depicts the daily variations of μ estimated for the sequences obtained in each of the five years with the exclusion of the significant activities listed in Table 1 (except 2009). There exist minor differences in these five profiles of μ, but they follow the almost the same pattern. First, during the time period between 0.0 and 0.2 days (i.e., between 0 a.m. and 5 a.m.) μ takes the smallest value and the detection capability is the best. Then the value of μ gradually increases as the time elapses, and it has a local peak around 0.4 days. Following the local peak μ shows a transient decrease corresponding to the recovery of the detection capability during lunchtime. At around 0.6 days (between 2 p.m. and 3 p.m.) it has a local peak again, and then it becomes smaller as the time gets closer to midnight. The differences of the maximum and minimum values of μ are 0.21–0.24.

Models			ΔABIC				
temporal variation			Year				
β	μ	σ	2006	2007	2008	2009	2010
			1160.9	1125.1	1092.6	1405.3	1448.9
√			266.7	320.3	236.5	495.7	383.5
	√		0.0	0.0	0.0	0.0	0.0
		√	389.8	352.2	336.1	450.91	546.6
√	√		2.0	1.8	2.0	2.0	2.0
√		√	127.4	139.4	103.8	183.93	208.6
	√	√	2.0	2.0	2.0	1.1	1.1
√	√	√	3.9	4.0	4.0	1.4	1.5

Table 2. Model comparison through ABIC values. The differences in ABIC values, ΔABIC, from cases where only the daily variation in μ is allowed are shown. The mark √ indicates the allowance of the daily variation in β, μ, and σ.

5. Discussions

To some extent, we can figure out the temporal pattern of the detection capability through the magnitude-time plots in Figure 6. In these plots there are some absences of small earthquakes with $M = -0.5$ or below during daytime. These plots also show a short recovery of the detection capability around noon. [28] has found a similar pattern of the detection capability in California through the hourly distribution of the number of reported earthquakes in the Advanced National Seismic Systems (ANSS) catalogue. Note that, however, both magnitude-time plots and hourly distributions provides us with only a qualitative pattern of the detection capability. For an understanding of the precise characteristics of an earthquake catalogue it is necessary to evaluate quantitatively the temporal change of the detection capability, and the statistical approach shown in this chapter is effective for such a purpose.

Of the three parameters β (or b-value), μ, and σ in the statistical model to describe an observed magnitude-frequency distribution, only μ shows the significant daily variation whereas β and σ do not show any variation. The temporal changes of b-value have been reported in some studies [15, 29, 30], and laboratory experiments [31–33], depth dependence [34–36], and faulting-style dependence of the b-value [37, 38] suggest a relation between stress state and b-value. We could consider air temperature and atmospheric pressure as physical factors which may cause the diurnal cycle of the stress changes in the Earth's crust. However, a constant value of β implies that the diurnal stress changes following the daily change of the temperature and/or pressure are insufficient to affect the magnitude-frequency distribution of earthquakes.

It is difficult to arrive at a plausible interpretation for a constant value of σ, because the factors controlling the value of σ still remain unknown. It has been suggested that σ is related to the spatial distribution of seismograph stations [2]. Following this suggestion the result of the unchanged σ is reasonable, because the spatial distribution of seismograph stations does not have daily variation.

As briefly described in section 4, the differences between the maximum and minimum values of μ are less than 0.25. In Figure 6, the values of μ in the case where we do not consider the daily variation of μ are indicated by the dotted lines; the difference between the maximum value of μ with the daily variation and that without the variation is approximately 0.15.

http://dx.doi.org/10.5772/54890

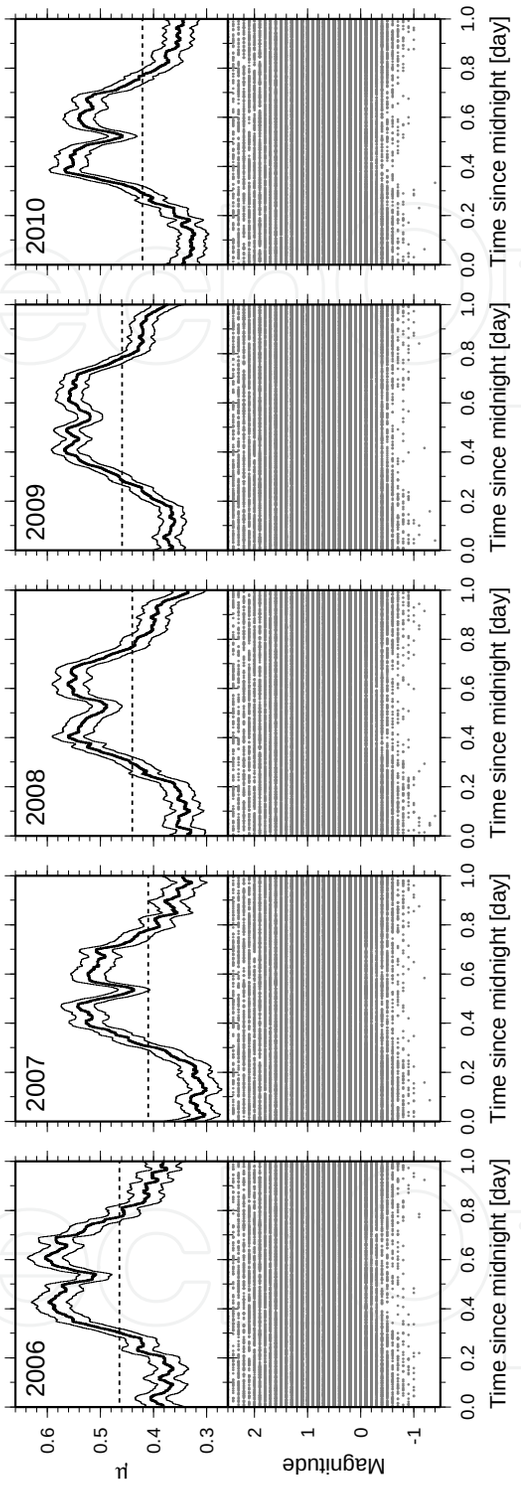

Figure 6. (Top) Estimated daily variations of μ in the stacked sequences of each of the five years (bold line) with the exclusion of the significant activities listed in Table 1, and their two standard error bands (thin lines), as a function of the elapsed time since midnight. The horizontal dotted lines shows the best estimate of μ in the case where we do not consider the temporal variation of μ. (Bottom) Magnitude-time plot of the stacked sequences for 2007. Note that only earthquakes with $M \leq 2.5$ are plotted to clarify the absences of small earthquakes during daytime but earthquakes with $M > 2.5$ are used in the analysis.

This suggests that, if we determine M_c without taking into consideration the daily variation of the detection capability, $M_c - 0.2$ would be appropriate as the completeness magnitude to avoid any unexpected biases in our seismicity analysis.

In the present study, we use the smoothness constraint, as shown in the equation (9), which reflects the cyclic property of data. Suppose that we use a different smoothness constraint

$$\Phi(\boldsymbol{\theta}|v_1, v_2, v_3) = \sum_{j=1}^{3} v_j \left[\sum_{i=1}^{N-1} \frac{(\theta_{i+1}^{(j)} - \mu_i^{(j)})^2}{t_{i+1} - t_i} \right], \tag{20}$$

where we do not impose a smoothness constraint between the start and end points on the values of β, μ, and σ. If we analyze a sufficiently large dataset such as the five sequences used in this study, the introduction of the constraint without considering the cyclic property is not problematic. This is because a sufficiently large sample size reasonably reduce random fluctuations or estimation errors, and therefore the estimated values of the parameters at the start and end points connects smoothly without the smoothness constraint. In actual, the analyses with the constraints given by the equations (20) and (9) gives almost identical temporal profiles of μ for the five sequences.

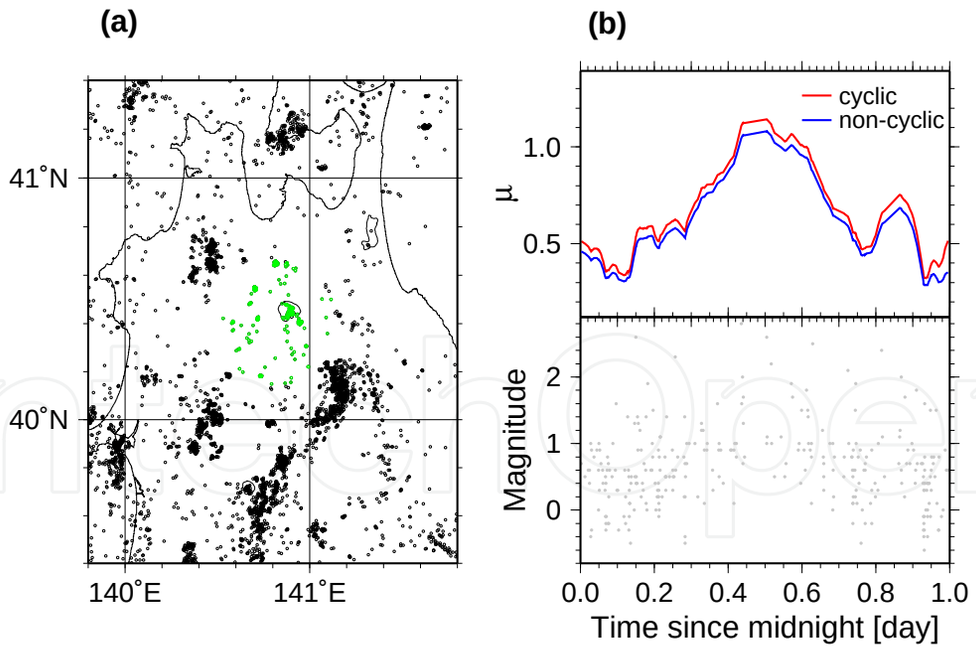

Figure 7. (a) Map showing the 300 earthquakes (green circles) used in the demonstrative analysis. (b) (top) Estimated daily variations of μ using the constraint following the equation (9) (red) and those following the equation (20). (bottom) Magnitude-time plot of the one-day stacked sequence of the 300 earthquakes.

http://dx.doi.org/10.5772/54890

For a small dataset, however, it is important to use the appropriate smoothness constraint. As a demonstration of the importance, a dataset containing 300 earthquakes distributed locally in a northern part of Japan (Figure 7a) was analyzed using the two constraints. The estimated daily variation of μ is shown in Figure 7b; we can observe systematic deviation between the two cases. The ABIC value with the smoothness constraint of the equation (9) is smaller than that corresponding to the equation (20), and the difference between the two ABIC values is 2.5, which is statistically significant. This demonstration implies that an analysis with an unreasonable constraint or prior distribution, which does not involve the essential characteristic of data and is statistically invalid, may leads us to an incorrect assessment of detection capability.

6. Concluding remarks

This chapter provided a statistical technique to make a quantitative evaluation of the daily variation of earthquake detection capability. As an example of its application to actual data, the datasets taken from the recent JMA catalogue were analyzed, and a guideline for the choice of completeness magnitude in the JMA catalogue was shown.

It should be noted that the results shown in this chapter were derived from an analysis where the earthquake sequence over an wide area was analyzed simultaneously. The daily variation of cultural noise, however, must have location dependency. Thus, in a manner similar to the analysis shown in Figure 7, we should examine regional earthquake sequences comprehensively to deal with an earthquake catalogue in a more sophisticated manner.

In this chapter we considered only the daily variation of detection capability, but human activities also have a weekly periodicity [28]. Additionally, we often observe seasonal (annual) variations in seismic noise level [39, 40], which are mainly caused by meteorological factors. Thus, the development of an appropriate method to handle such multiple periodic variations of earthquake detection capability is a necessary challenge to determine the completeness magnitude more accurately.

Acknowledgments

This study was partially supported by the Grants-in-Aid 23240039 for Scientific Research (A) by the Ministry of Education, Culture, Sports, Science and Technology, Japan, and by the ISM Cooperative Research Program (2011-ISM·CRP-2007). To obtain the results shown in this chapter, the author used the supercomputer system of the Institute of Statistical Mathematics, Japan. Figures were generated using the GMT software [41].

Author details

Takaki Iwata

The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan

References

[1] Woessner J, Wiemer S. Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. Bulletin of the Seismological Society of America 2005;95(2):684-698.

[2] Iwata T. Low detection capability of global earthquakes after the occurrence of large earthquakes: investigation of the Harvard CMT catalogue. Geophysical Journal International 2008;174(3):849-856.

[3] Schorlemmer D, Woessner J. Probability of detecting an earthquake. Bulletin of the Seismological Society of America 2008;98(5):2103-2117.

[4] Schorlemmer D, Mele F, Marzocchi W. A completeness analysis of the National Seismic Network of Italy. Journal of Geophysical Research-Solid Earth 2010;115B4:B04308;doi10.1029/2008JB006097.

[5] Nanjo KZ, Ishibe T, Tsuruoka H, Schorlemmer D, Ishigaki Y, Hirata N. Analysis of the completeness magnitude and seismic network coverage of Japan. Bulletin of the Seismological Society of America 2010;100(6):3261-3268.

[6] Mignan A, Werner MJ, Wiemer S, Chen CC, Wu YM. Bayesian Estimation of the Spatially Varying Completeness Magnitude of Earthquake Catalogs. Bulletin of the Seismological Society of America 2011;101(3):1371-1385.

[7] Plenkers K, Schorlemmer D, Kwiatek G. On the Probability of Detecting Picoseismicity. Bulletin of the Seismological Society of America 2011;101(6):2579-2591.

[8] Shimshoni.M. Evidence for Higher Seismic Activity During Night. Geophysical Journal of the Royal Astronomical Society 1971;24(1):97-99.

[9] Flinn EA, Blandford.RR, Mack H. Evidence for Higher Seismic Activity During Night. Geophysical Journal of the Royal Astronomical Society 1972;28(3):307-309.

[10] Knopoff L, Gardner JK. Higher seismic activity during local night on the raw worldwide earthquake catalogue. Geophysical Journal of the Royal Astronomical Society 1972;283:311-313.

[11] Iwata, T. Quantitative analysis of the daily variation of earthquake detection capability. 2012;Chikyu Monthly;34(9):504-508 (in Japanese).

[12] Funasaki, J., Earthquake Prediction Information Division, Seismological and Volcanological Department, Japan Meteorological Agency. Revision of the JMA Velocity Magnitude. Quarterly Journal of Seismology 2004;67(1-4):11-20(in Japanese with English abstract and figure captions).

[13] Katsumata, A. (2004) Revision of the JMA Displacement Magnitude. Quarterly Journal of Seismology 2004;67(1-4):1-10(in Japanese with English abstract and figure captions).

[14] Nanjo KZ, Tsuruoka H, Hirata N, Jordan TH. Overview of the first earthquake forecast testing experiment in Japan. Earth Planets and Space 2011;63(3):159-169.

[15] Ogata Y, Katsura K. Analysis of Temporal and Spatial Heterogeneity of Magnitude Frequency-Distribution Inferred from Earthquake Catalogs. Geophysical Journal International 1993;113(3):727-738.

[16] Gutenberg B, Richter CF. Frequency of earthquakes in California. Bulletin of the Seismological Society of America 1944;34(4):185-188.

[17] Ringdal F. On the estimation of seismic detection thresholds. Bulletin of the Seismological Society of America 1975;656:1631–42.

[18] Iwata, T. Revisiting the global detection capability of earthquakes during the period immediately after a large earthquake: considering the influence of intermediate-depth and deep earthquakes. Research in Geophysics 2012;2(1):24–28.

[19] Powell MJD. Approximation theory and methods. New York: Cambridge University Press; 1981.

[20] Good IJ, Gaskins RA. Nonparametric Roughness Penalties for Probability Densities. Biometrika 1971;58(2):255-277.

[21] Eggermon PPB, LaRiccia VN. Maximum Penalized Likelihood Estimation, vol I: Density Estimation. New York: Springer; 2001.

[22] Good IJ. The estimation of probabilities. Cambridge: The MIT Press; 1965.

[23] Akaike H. Likelihood and Bayes Procedure. In:Bernardo JE. et al. (eds.) Bayesian Statistics. Valencia: University Press;1980. p143-166.

[24] Kass RE, Raftery AE. Bayes factors. Journal of the American Statistical Association 1995;90(430):773-795.

[25] Gill, P. E., Murray, W., Wright, MH. Numerical linear algebra and optimization, vol. I. Redwood City: Addison-Wesley; 1991.

[26] Tierney L, Kadane JB. Accurate Approximations for Posterior Moments and Marginal Densities. Journal of the American Statistical Association 1986;81(393):82-86.

[27] Konishi, S., Kitagawa, G. Information Criteria and Statistical Modeling Springer, New York; 2008.

[28] Atef AH, Liu KH, Gao SS. Apparent Weekly and Daily Earthquake Periodicities in the Western United States. Bulletin of the Seismological Society of America 2009;99(4):2273-2279.

[29] Iwata T, Young RP. Tidal stress/strain and the b-values of acoustic emissions at the Underground Research Laboratory, Canada. Pure and Applied Geophysics 2005;162(6-7):1291-1308.

[30] Nanjo KZ, Hirata, N, Obara, K., Kasahara, K. Decade-scale decrease in *b* value prior to the *M*9-class 2011 Tohoku and 2004 Sumatra quakes. Geophysical Research Letters 2012;39(20):L20304;doi:10.1029/2012GL052997.

[31] Scholz, C. The frequency-magnitude relation of microfracturing in rock and its relation to earthquakes 1968;58(1):399-415.

[32] Amitrano D. Brittle-ductile transition and associated seismicity: Experimental and numerical studies and relationship with the *b* value. Journal of Geophysical Research-Solid Earth 2003;108(B1):2044;doi:10.1029/2001JB000680.

[33] Lei XL. How do asperities fracture? An experimental study of unbroken asperities. Earth and Planetary Science Letters 2003;213(3-4):347-359.

[34] Mori J, Abercrombie RE. Depth dependence of earthquake frequency-magnitude distributions in California: Implications for rupture initiation. Journal of Geophysical Research-Solid Earth 1997;102(B7):15081-15090.

[35] Gerstenberger M, Wiemer S, Giardini D. A systematic test of the hypothesis that the b value varies with depth in California. Geophysical Research Letters 2001;28(1):57-60.

[36] Amorèse, D., Grasso JR, Rydelek PA. On varying *b*-values with depth: results from computer-intensive tests for Southern California. Geophysical Journal International 2010;180(1):347-360.

[37] Schorlemmer D, Wiemer S, Wyss M. Variations in earthquake-size distribution across different stress regimes. Nature 2005;437(7058):539-542.

[38] Narteau C, Byrdina S, Shebalin P, Schorlemmer D. Common dependence on stress for the two fundamental laws of statistical seismology. Nature 2009;462(7273):642-645.

[39] Fyen J. Diurnal and Seasonal-Variations in the Microseismic Noise-Level Observed at the Noress Array. Physics of the Earth and Planetary Interiors 1990;63(3-4):252-268.

[40] Hillers G, Ben-Zion Y. Seasonal variations of observed noise amplitudes at 2-18 Hz in southern California. Geophysical Journal International 2011;184(2):860-868.

[41] Wessel P., Smith, WHF. New, improved version of the Generic mapping Tools released. EOS 1998;79:579.