CMT data inversion using a Bayesian information criterion to estimate seismogenic stress fields

Toshiko Terakawa and Mitsuhiro Matsu’ura

Department of Earth and Planetary Science, The University of Tokyo, Bunkyo-ku, Tokyo 113–0033, Japan. E-mail: terakawa@eps.s.u-tokyo.ac.jp

Accepted 2007 October 10. Received 2007 October 10; in original form 2006 April 12

SUMMARY
We developed an inversion method to estimate the stress fields related to earthquake generation (seismogenic stress fields) from the centroid moment tensors (CMT) of seismic events by using Akaike’s Bayesian information criterion (ABIC). On the idea that the occurrence of an earthquake releases some part of the seismogenic stress field around its hypocentre, we define the CMT of a seismic event by a weighted volume integral of the true but unknown seismogenic stress field. Representing each component of the seismogenic stress field by the superposition of a finite number of 3-D basis functions (tri-cubic B-splines), we obtain a set of linear observation equations to be solved for the expansion coefficients (model parameters). We introduce prior constraint on the roughness of the seismogenic stress field and combine it with observed data to construct a Bayesian model with hierarchic, highly flexible structure controlled by hyper-parameters. The optimum values of the hyper-parameters are objectively determined from observed data by using ABIC. Given the optimum values of the hyper-parameters, we can obtain the best estimates of model parameters by using a maximum likelihood algorithm. We tested the validity of the inversion method through numerical experiments on two synthetic CMT data sets, assuming the distribution of fault orientations to be aligned with the maximum shear stress plane in one case and to be random in the other case. Then we applied the inversion method to actual CMT data in northeast Japan, and obtained the pattern of the seismogenic stress field consistent with geophysical and geological observations.

Key words: Inverse theory; Seismicity and tectonics; Dynamics: seismotectonics; Fractures and faults.

1 INTRODUCTION
The occurrence of earthquakes can be regarded as a stress release process in the Earth’s crust. This means that observed seismological data contain direct information about the internal stress state of the crust. Since the crust contains a number of pre-existing faults with various orientations and most earthquakes utilize them to release accumulated tectonic stress, we cannot correctly infer the stress state from a single seismic event (McKenzie 1969). Therefore, some statistical approach is necessary for unbiased inference of the stress state.

In the 1980s several stress inversion methods have been proposed to estimate the stress state of the Earth’s crust from the focal mechanism data (fault orientations and slip directions) of seismic events (Ellsworth & Zhonghui 1980; Gephart & Forsyth 1984; Michael 1984; Michael 1987). In these conventional methods, on the basic assumption that seismic slip occurs in the direction of the tangential component of the stress vector acting on a pre-existing fault (Wallace 1951; Bott 1959), the pattern of average deviatoric stress is determined for each partitioned block by minimizing the difference in direction between the predicted and observed slip vectors in the least-squares sense. However, as pointed out by Twiss & Unruh (1998), the conventional methods have an essential problem that the inversion of fault slip data provides direct constraint on the regional deformation rate but not the regional stress. This essential problem will be resolved by using centroid moment tensor (CMT) solutions of seismic events instead of focal mechanism solutions as data. The conventional methods have also some practical problems; one of which is that unavoidable ambiguity in choosing true fault planes would degrade the reliability of inversion results (Michael 1987; Hardebeck & Hauksson 2001), and another is that inverted stress patterns depend on the way of area partitioning unless we have sufficient data (Hardebeck & Hauksson 1999; Townend & Zoback 2001; Hardebeck & Michael 2004). The former, ambiguity in choosing true fault planes, is a technical problem, which can be resolved by using the slip shear stress component (SSSC) criterion (Angelier 2002). The latter, the model dependence of inverted stress patterns, is a universal and inherent problem in statistical inference, which cannot be resolved within the framework of the likelihood maximization principle.

In order to resolve the problem that inverted stress patterns depend on the way of area partitioning, Hardebeck & Michael (2006) proposed a kind of damped least squares method. They construct a family of parametric models with an adjustable damping
The Earth's crust contains a number of pre-existing faults with various orientations. Most earthquakes utilize the pre-existing faults to release accumulated tectonic stress, and so fault slip vectors are not necessarily in the plane of maximum shear stress (McKenzie 1969). Now we consider the situation that numerous pre-existing faults with random orientations are distributed in a region under given uniform stress. In this situation we calculated the $P$- and $T$-axes of possible seismic events for three representative cases; the incidence of events is (1) random for all the pre-existing faults, (2) proportional to the magnitude of shear stress acting on pre-existing faults (Tresca failure criterion) and (3) proportional to the ratio of shear stress to shear strength at pre-existing faults (Coulomb failure criterion). In the case (3) we define the shear strength of each pre-existing fault by the product of the standard value (0.6) of frictional coefficients and the normal stress acting on it. We plotted the $P$- and $T$-axes of seismic events for the cases of (1), (2) and (3) on stress diagrams in Figs 1(a), (b) and (c), respectively, with lower hemisphere equal-area projection. In either case, as can be seen from these diagrams, the $P$- and $T$-axes of seismic events tend to cluster about the maximum and minimum principal axes of the given uniform stress. This means that an ensemble average of the stress-release patterns of seismic events will provide a good estimate of the true tectonic stress pattern if the orientations of pre-existing faults are random and the frictional coefficients of them are the same. In the following subsections we show basic ideas for seismic source representation and Bayesian statistical inference to develop a robust inversion method to estimate tectonic stress fields from seismic data.

2.1 Seismic source representation

The source of an earthquake is brittle dynamic rupture in the Earth's crust, which occurs so as to release the elastic strain energy accumulated by tectonic loading. In the framework of linear elasticity we represent the inelastic process in the source region with its centre at $\mathbf{x} = \mathbf{x}_0$ by a moment tensor $M(\mathbf{x}_0, t) = [M_{\text{eq}}(\mathbf{x}_0, t)]$, which is a second-order symmetric tensor with the diagonal elements of force dipoles and the off-diagonal elements of force couples (Backus & Mulcahy 1976a, b). The moment tensor gives the same deformation field as the dynamic rupture brought about everywhere except the source region. The process in the source region itself cannot be described by the moment tensor, because the source region would be neither isotropic nor linearly elastic, as pointed out by Twiss & Unruh (1998) and Ben-Zion (2003).

Seismic wave fields $u_i(\mathbf{x}, t)$ are directly related to the moment tensor $M_{\text{eq}}(\mathbf{x}_0, \tau)$ by means of the spatial derivatives of Green's tensor $G_{ipq}(\mathbf{x}, t; \mathbf{x}_0, \tau)$ (e.g. Aki & Richards 1980; Kennett 1983):

$$u_i(\mathbf{x}, t) = \int_{-\infty}^{t} G_{ipq}(\mathbf{x}, t; \mathbf{x}_0, \tau) M_{\text{eq}}(\mathbf{x}_0, \tau) d\tau.$$  

(1)

In the problem of seismic wave propagation, since the Earth's crust and mantle behave like an isotropic linearly elastic body for short-period disturbance, we may use elastic Green's tensor. Then, the CMT solution $M_{\text{eq}} = M_{\text{eq}}(\mathbf{x}_0, t \to \infty)$ of a seismic event, which is

![Figure 1](https://academic.oup.com/gji/article-abstract/172/2/674/626152/fig1.png)
determined from the inversion analysis of seismic waveform data, can be regarded as a clearly defined quantity in the theory of linear elasticity. On the other hand, in order to obtain the focal mechanism solution (fault orientation and slip direction) of the seismic event, we need to assume the isotropic linear elasticity of the source region. If we use the focal mechanism solutions as data, the inverted stress field would be biased because of this inappropriate assumption.

In general, the CMT of a seismic event is defined by the volume integral of moment tensor density \( m_j(\xi) \) over the source region \( V_s \):

\[
M_j = \int_{V_s} m_j(\xi) dV. \tag{2}
\]

In the ideal case where displacement discontinuity \( [u_i(\eta)]^+ \) occurs across a fault surface \( \Sigma(\eta) \) embedded in a homogeneous elastic medium, on the basis of the elastodynamic representation theorem (e.g. Aki & Richards 1980), the moment tensor density can be written as

\[
m_j(\xi) = c_{ijkl} u_i(\eta)^+ n_l(\eta) \delta(\xi - \eta), \tag{3}
\]

where \( c_{ijkl} \) is the elastic tensor, \( n_l(\eta) \) is the unit normal vector of \( \Sigma \), and \( \delta(\xi - \eta) \) denotes the delta function localized at the fault surface \( \Sigma \). Substituting eq. (3) into eq. (2), we obtain

\[
M_j = \int_{\Sigma} c_{ijkl} [u_i(\eta)]^+ n_j(\eta) d\Sigma(\eta). \tag{4}
\]

For tangential displacement discontinuity (shear faulting) in an isotropic elastic medium, the above expression can be reduced to

\[
M_j = \mu \int_{\Sigma} \Delta u_i(\eta) [v_i(\eta)n_j(\eta) + v_j(\eta)n_i(\eta)] d\Sigma(\eta), \tag{5}
\]

where \( \mu \) is the rigidity of the medium, and \( \Delta u \) and \( \nu \) denote the magnitude and unit direction vector of fault slip, respectively. In the case of a planar fault with its area \( \Sigma \), normal vector \( n \), slip magnitude \( \Delta u \) and slip direction \( \nu \), defining the seismic moment as \( M_0 = \mu \Delta u \Sigma \), we obtain the familiar representation of CMT as

\[
M_j = M_0 [v_i n_j + v_j n_i]. \tag{6}
\]

Actually, the fault \( \Sigma \) is not a surface of discontinuity but a shear fracture zone with a finite thickness. Then, taking a surface \( S_f \) to consist of the upper and lower fault surfaces, \( \Sigma^+ \) and \( \Sigma^- \), joined at both ends of the fault, we can rewrite eq. (4) as

\[
M_j = \int_{S_f=S^+ \cup S^-} c_{ijkl} u_i(\eta)n_j(\eta)dS. \tag{7}
\]

We consider a region \( V \) bounded by the inner closed surface \( S_f \) and an outer closed surface \( S_o \) with a sufficiently large radius as shown in Fig. 2. Then, the integral over the surface of the source region in eq. (7) can be transformed into the integral over the surface of the region \( V \) surrounding the source region as

\[
M_j = -\int_{S=s_f \cup S_o} c_{ijkl} u_i(\eta)n_j(\eta)dS, \tag{8}
\]

because the surface integral of \( c_{ijkl} u_i n_j \) over the outer surface \( S_o \) with a sufficiently large radius tends to zero (Kostrov 1974). Here, it should be noted that the unit normal vector \( n(\xi) \) is taken to be outward the region \( V \). Applying Gauss’ divergence theorem to the right hand side of eq. (8), we finally obtain

\[
M_j = -\int_V \frac{\partial u_i(\xi)}{\partial \xi^j} dV = \int_V \Delta \tau_{ij}(\xi)dV, \tag{9}
\]

where \( \Delta \tau_{ij} \) represents static stress release associated with the seismic event. Since dynamic rupture growth is controlled by energy flow into the rupture zone from the surrounding region storing elastic strain energy (e.g. Aki & Richards 1980; Rice 1980; Hashimoto & Matsu’ura 2000, 2002; Fukuyama et al. 2002), the volume integral representation of CMT in eq. (9) is more essential than the surface integral representation in eq. (4).

### 2.2 Bayesian statistical inference

On the basis of the entropy maximization principle (Akaike 1977), Akaike (1980) has proposed a Bayesian information criterion (ABIC) for objective model selection in statistical inference. The introduction of ABIC enables us to freely construct a stochastic model for statistical inference by combining prior knowledge and prior information. The construction of Bayesian models is usually performed in the following way. First, on the basis of prior knowledge about a physical system (scientific theory), we determine the functional form of a stochastic model \( p(d|a; \sigma^2) \) that relates observed data \( d \) with model parameters \( a \). Here, \( \sigma^2 \) is an unknown scale factor of the covariance matrix of data errors. In the present case prior knowledge is represented by eq. (9) that relates observed CMT data with stress release associated with seismic events. Second, we represent prior information about the model parameters in the form of a probability density function \( p(a; \rho^2) \). Here, \( \rho^2 \) is an unknown scale factor controlling the reliability of prior information. In the present case we have prior information that the spatial distribution of stress release must be smooth in some degree (Yabuki & Matsu’ura 1992) because of the finiteness in shear strength of rocks. Such constraint can be regarded as the indirect prior information about model parameters, which regulates the structure of the stochastic model \( p(d|a; \sigma^2) \). Third, combining the prior information \( p(a; \rho^2) \) with the stochastic model \( p(d|a; \sigma^2) \) by Bayes’ rule, we construct a Bayesian model with hierarchic, highly flexible structure controlled by hyper-parameters, \( \sigma^2 \) and \( \rho^2 \):

\[
p(a; \sigma^2; \rho^2|d) = cp(d|a; \sigma^2)p(a; \rho^2). \tag{10}
\]

Here, it should be noted that the Bayesian model represents a family of parametric models with different structures.

For the family of parametric models ABIC gives an objective measure of the goodness of the hypothetic predictive distribution.
as an approximation to the true but unknown distribution. Therefore, we may use ABIC, which is defined by

$$\text{ABIC}(\sigma^2, \rho^2; d) = -2 \log \int p(a; \sigma^2, \rho^2; d)da + C,$$

(11)
to select the optimum values of hyper-parameters. Given the optimum values of hyper-parameters, which minimize ABIC, we can obtain the best estimates of model parameters by applying a maximum likelihood algorithm (Jackson & Matsu’ura 1985; Yabuki & Matsu’ura 1992).

3 Mathematical Formulation

On the basic ideas described in the previous section we mathematically formulate the inverse problem of estimating seismogenic stress fields from CMT data of seismic events. In the formulation we suppose most earthquakes utilize pre-existing faults to release accumulated tectonic stress and the pre-existing faults are randomly distributed with various orientations.

3.1 Observation equations

The occurrence of an earthquake releases some part of seismogenic stress in a finite region surrounding its hypocentre. Then, using a 3-D weighting function $p(x; x_s, M_0)$ with its centre at the hypocentre $x = x_s$ and its extent depending on the seismic moment $M_0$, we can relate the stress release distribution $\Delta \tau_{ij}(x)$ in eq. (9) with the seismogenic stress field $\tau_{ij}(x)$ as

$$\Delta \tau_{ij}(x) = p(x; x_s, M_0)\tau_{ij}(x).$$

(12)

Substituting the above relation into the volume integral representation of CMT in eq. (9), we obtain the equation that relates an observed CMT solution $M(x_s; M_0)$ with the true but unknown seismogenic stress field:

$$M_{ij}(x_s; M_0) = \int_V p(x; x_s, M_0)\tau_{ij}(x)dV + e_{ij}^{\text{data}}$$

(13)

for $i = 1, 2, 3; j = 1, 2, 3$.

Here, $e_{ij}^{\text{data}}$ are data errors of $M_{ij}$. As the weighting function $p(x; x_s, M_0)$ we take the 3-D Gaussian-type distribution with its peak at the hypocentre $x = x_s$ and the variance $L^2$ proportional to the two-thirds power of the seismic moment $M_0$:

$$p(x; x_s, M_0) = \frac{M_0}{(2\pi L^2)^{3/2}} \exp \left[-\frac{1}{2L^2}(x - x_s)^T(x - x_s)\right].$$

(14)

with

$$L = c M_0^{1/3}.$$  

(15)

The seismogenic moment $M_0$ scales with the cube of fault length (Aki 1972; Kanamori & Anderson 1975), and so eq. (15) means that the effective distance $L$ of stress release is proportional to fault length. In eq. (13), through the weighting function $p(x; x_s, M_0)$, we related a point CMT datum $M_{ij}(x_s; M_0)$ with the seismogenic stress field $\tau_{ij}(x)$ in an extent proportional to the seismic moment $M_0$. This corresponds to that the rupture growth of a seismic event is controlled by energy supply from its surrounding region with extent proportional to the seismic moment. On the other hand, the peak value of the weighting function does not depend on the seismic moment. This corresponds to the empirical law that the stress drop of earthquakes is independent of the scale (seismic moment) of them.

In order to discretize eq. (13) we represent each component of the seismogenic stress field $\tau_{ij}(x)$ by the superposition of a finite number of 3-D basis functions $\Phi_n(x)$ as

$$\tau_{ij}(x) = \sum_{n=1}^{M} a_{ij}^n \Phi_n(x) + \Phi_0(x) \quad (i = 1, 2, 3; j = 1, 2, 3),$$

(16)

where $a_{ij}^n$ are the expansion coefficients (model parameters) to be determined from observed data, and $\Phi_0(x)$ denotes the remaining basis functions needed for making a complete set, which satisfies

$$\int_V \Phi_n(x)\Phi_0(x)dV = 0 \quad (m = 1, \ldots, M).$$

(17)

Substituting eq. (16) into eq. (13), we obtain a set of linear observation equations for $N$ seismic events as

$$M_{ij} = \sum_{n=1}^{M} F_{ij}^n a_{ij}^n + e_{ij}^{\text{data}, n}$$

(18)

with

$$F_{ij}^n = \int_V p(x; x_s^n, M_0^n)\Phi_n(x)dV,$$

(19)

and

$$e_{ij}^{\text{data}, n} = e_{ij}^{\text{model}, n} + e_{ij}^{\text{data}, n}.$$  

(20)

Here, it should be noted that the total errors $e_{ij}^{\text{model}}$ in the discretized observation equations (18) consists of the data errors $e_{ij}^{\text{data}, n}$ and the modelling errors $e_{ij}^{\text{model}, n}$, defined by

$$e_{ij}^{\text{model}, n} = \int_V p(x; x_s^n, M_0^n)\Phi_0(x)dV.$$  

(21)

The modelling errors come from deficiency in theoretical modelling (Jackson 1979); in the present case, incomplete representation of the true seismogenic stress field by a finite number of basis functions. Since the extent of the weighting function $p(x; x_s, M_0)$ for a seismic event is proportional to its seismic moment, the modelling errors are also proportional to the seismic moment. If we take a sufficiently large number of basis functions, the modelling errors become sufficiently small in comparison with data errors.

The CMT of seismic events is not a directly observable quantity but the quantity indirectly obtained from the inversion analysis of observed seismic waveform data. Therefore, the data errors of CMT also consist of observation errors and modelling errors. We may consider the observation errors to be independent of the scale of seismic events (seismic moment) but not the modelling errors. Larger events have generally more complicated fault geometry, rupture propagation, and slip distribution than smaller events. Nevertheless, we always use the simplest source model (CMT) with six degrees of freedom to represent the characteristics of seismic events with widely different scales. Then, the modelling errors of CMT data increase as the scale of seismic events (seismic moment) increases. From these considerations we may suppose that the total errors $e_{ij}^{\text{model}}$ of a CMT solution are proportional to the seismic moment $M_0^n$. This corresponds to that the magnitudes of seismic events have the estimation errors independent of them.
The seismic moment of the events that we use for inversion analysis takes various values in the range of several orders. In practical analysis, since the total errors of CMT data are proportional to the seismic moment, we normalize the both sides of the original observation equations (18) by the seismic moment. The normalized observation equations can be written in vector form as

\[ \mathbf{d} = \mathbf{F}a + \mathbf{e}, \]  

(23)

where \( \mathbf{d} \) is a \( 6N \times 1 \) normalized data vector, \( \mathbf{a} \) is a \( 6M \times 1 \) model parameter vector, \( \mathbf{F} \) is a \( 6N \times 6M \) normalized coefficient matrix, and \( \mathbf{e} \) is a \( 6N \times 1 \) normalized error vector. Assuming the normalized errors \( \mathbf{e} \) in eq. (23) to be Gaussian with zero mean and a covariance matrix \( \sigma^2 \mathbf{E} \), we obtain a stochastic model that relates the data \( \mathbf{d} \) with the model parameter \( \mathbf{a} \) as

\[ p(\mathbf{d} | \mathbf{a}; \sigma^2) = (2\pi \sigma^2)^{-3N/2} \| \mathbf{E} \|^{-1/2} \times \exp \left[ -\frac{1}{2\sigma^2} (\mathbf{d} - \mathbf{F}a)^T \mathbf{E}^{-1} (\mathbf{d} - \mathbf{F}a) \right], \]

(24)

where \( \sigma^2 \) is an unknown scale factor of the covariance matrix, and \( \| \mathbf{E} \| \) denotes the absolute value of the determinant of \( \mathbf{E} \). Since the total errors \( e_i \) of CMT data have been already normalized by the seismic moment \( M_s \), we suppose \( \mathbf{E} \) to be a \( 6N \times 6N \) unit matrix \( \mathbf{I} \) for simplicity. We can use the stochastic model \( p(\mathbf{d} | \mathbf{a}; \sigma^2) \) in eq. (24) as a basic device to extract information from observed data.

### 3.2 Bayesian modelling

In addition to observed data we usually have some prior constraint that regulates the structure of the parametric model in eq. (24). In the present case, from physical consideration, we impose prior constraint on the roughness of seismogenic stress fields \( \tau_{ij}(x) \). As a measure of the roughness of \( \tau_{ij}(x) \) we define the following quantity:

\[ r = \sum_{i=1}^{3} \sum_{j=1}^{3} \int \left[ \left( \frac{\partial \tau_{ij}}{\partial x_1} \right)^2 + \left( \frac{\partial \tau_{ij}}{\partial x_2} \right)^2 + \left( \frac{\partial \tau_{ij}}{\partial x_3} \right)^2 \right] dx. \]

(25)

where the volume integration is performed over the whole model region. Substituting eq. (16) into eq. (25), we obtain

\[ r = \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k=1}^{3} \sum_{l=1}^{3} a_{ij}^k R_{pq} a_{kl}^p \]

(26)

with

\[ R_{pq} = \int \left[ \frac{\partial \Phi_p(x)}{\partial x_1} \frac{\partial \Phi_q(x)}{\partial x_1} + \frac{\partial \Phi_p(x)}{\partial x_2} \frac{\partial \Phi_q(x)}{\partial x_2} + \frac{\partial \Phi_p(x)}{\partial x_3} \frac{\partial \Phi_q(x)}{\partial x_3} \right] dx. \]

(27)

or in vector form,

\[ r = \mathbf{a}^T \mathbf{G} \mathbf{a}, \]

(28)

with

\[ \mathbf{G} = [R_{ij} \mathbf{I}] \quad (I = 1, \ldots, M; J = 1, \ldots, M), \]

(29)

where \( \mathbf{I} \) is a \( 6 \times 6 \) unit matrix.

The roughness \( r \) defined in eq. (26) has a positive-definite quadratic form of the model parameters \( \mathbf{a} \). Then, we can express the prior constraint on the roughness of \( \tau_{ij}(x) \) in the form of a probability density function (pdf) with an unknown scale factor \( \rho^2 \) as

\[ p(\mathbf{a}; \rho^2) = (2\pi \rho^2)^{-N/2} \| \mathbf{A}_p \|^{1/2} \exp \left[ -\frac{1}{2\rho^2} \mathbf{a}^T \mathbf{G} \mathbf{a} \right], \]

(30)

where \( P \) is the rank of the matrix \( \mathbf{G} \), and \( \| \mathbf{A}_p \| \) denotes the absolute value of the product of non-zero eigenvalues of \( \mathbf{G} \). Combining the prior constraint \( p(\mathbf{a}; \rho^2) \) in eq. (30) with the stochastic model \( p(\mathbf{d} | \mathbf{a}; \sigma^2) \) in eq. (24) by Bayes’ rule, and introducing a new hyper-parameter \( \alpha^2 (= \sigma^2/\rho^2) \) instead of \( \rho^2 \), we obtain a hierarchical, highly flexible model controlled by the hyper-parameters \( \sigma^2 \) and \( \alpha^2 \):

\[ p(\mathbf{a}; \sigma^2, \alpha^2 | \mathbf{d}) = c(2\pi \sigma^2)^{-(6N+P)/2}(\alpha^2)^{P/2} \| \mathbf{E} \|^{-1/2} \| \mathbf{A}_p \|^{1/2} \times \exp \left[ -\frac{1}{2\sigma^2} s(\mathbf{a}) \right] \]

(31)

with

\[ s(\mathbf{a}) = (\mathbf{d} - \mathbf{F}a)^T \mathbf{E}^{-1} (\mathbf{d} - \mathbf{F}a) + \alpha^2 \mathbf{a}^T \mathbf{G} \mathbf{a}. \]

(32)

Here, it should be noted that the hyper-parameter \( \alpha^2 \) controls the relative weight of the observed CMT data to the prior constraint. Then, our problem is to find the values of \( \mathbf{a}, \sigma^2 \) and \( \alpha^2 \) that maximize the posterior pdf in eq. (31) for given data \( \mathbf{d} \).

### 3.3 Inversion algorithm

For certain fixed values of the hyper-parameters \( \sigma^2 \) and \( \alpha^2 \), Jackson & Matsu’ura (1985) have obtained the solution \( \mathbf{a}^* \) that maximizes the posterior pdf in eq. (31) and its covariance matrix \( \mathbf{C}(\mathbf{a}^*) \) as

\[ \mathbf{a}^* = (\mathbf{F}^T \mathbf{E}^{-1} \mathbf{F} + \alpha^2 \mathbf{G})^{-1} \mathbf{F}^T \mathbf{E}^{-1} \mathbf{d}, \]

(33)

\[ \mathbf{C}(\mathbf{a}^*) = \alpha^2 (\mathbf{F}^T \mathbf{E}^{-1} \mathbf{F} + \alpha^2 \mathbf{G})^{-1}. \]

(34)

In actual problems, the hyper-parameters \( \sigma^2 \) and \( \alpha^2 \) are free. To determine the optimum values of \( \sigma^2 \) and \( \alpha^2 \), as stated in the Section 2.2, we can use Akaike’s Bayesian information criterion (ABIC). The definition of ABIC in the present context is given by

\[ \text{ABIC}(\sigma^2, \alpha^2 | \mathbf{d}) = -2 \log \int p(\mathbf{a}; \sigma^2, \alpha^2 | \mathbf{d}) d\mathbf{a} + C, \]

(35)

where \( p(\mathbf{a}; \sigma^2, \alpha^2 | \mathbf{d}) \) is the posterior pdf in eq. (31). The integration of the posterior pdf over the whole parameter space has been carried out by Yabuki & Matsu’ura (1992) as

\[ \int p(\mathbf{a}; \sigma^2, \alpha^2 | \mathbf{d}) d\mathbf{a} = c(2\pi \sigma^2)^{-(6N+P-6M)/2} \| \mathbf{E} \|^{-1/2} \| \mathbf{A}_p \|^{1/2} \times \| \mathbf{F}^T \mathbf{E}^{-1} \mathbf{F} + \alpha^2 \mathbf{G} \|^{-1/2} \exp \left[ -\frac{1}{2\sigma^2} s(\mathbf{a}^*) \right]. \]

(36)

The minimum of ABIC is realized by maximizing the above quantity (marginal likelihood). From the necessary condition for the minimum of ABIC we obtain

\[ \sigma^2 = s(\mathbf{a}^*)/(6N + P - 6M). \]

(37)

Substitution of the above relation into Eq. (36) yields

\[ \text{ABIC}(\sigma^2) = (6N + P - 6M) \log s(\mathbf{a}^*) - P \log \sigma^2 + \log \| \mathbf{F}^T \mathbf{E}^{-1} \mathbf{F} + \alpha^2 \mathbf{G} \| + C^*, \]

(38)

where \( C^* \) is a constant term independent of \( \sigma^2 \). The value of \( \sigma^2 \) that minimizes ABIC can be found by iterative numerical search. Once the optimum value \( \sigma^2 \) of the hyper-parameter \( \sigma^2 \) has been found, the best estimates \( \hat{\mathbf{a}} \) of the model parameters \( \mathbf{a} \), the optimum value \( \hat{\sigma}^2 \) of the hyper-parameter \( \sigma^2 \) and the corresponding covariance matrix \( \mathbf{C}(\hat{\mathbf{a}}) \) are obtained from eqs. (33), (37) and (34), respectively, by replacing \( \alpha^2 \) with \( \hat{\alpha}^2 \).
4 VALIDITY TESTS USING SYNTHETIC CMT DATA

In this section we test the validity of the CMT data inversion method through numerical experiments. The Earth’s crust contains a number of pre-existing faults with various sizes and orientations. The pre-existing faults are much weaker than the surrounding rocks, and so most earthquakes utilize the pre-existing faults to release accumulated tectonic stress. We suppose that seismic slip occurs in the direction of the tangential component of the stress vector acting on a pre-existing fault. This means that the slip vector is not necessarily in the plane of maximum shear stress there (McKenzie 1969). In validity tests we consider two extreme cases. In the first case (Case 1) we assume that every pre-existing fault is aligned with the plane of maximum shear stress there. In the second case (Case 2) we assume that the orientations of pre-existing faults are completely random. The actual distribution of fault orientations will be between these two extreme cases.

In both cases, for simplicity, we consider a 1-D true stress field $\tau_{ij}(x)$ on a segment of the x-axis ($-165 \text{ km} \leq x \leq 165 \text{ km}$) as shown in Fig. 3. Here, each component of the true stress field contains three modes with different characteristic length. In the 1-D problem the CMT of a seismic event with its hypocentre $x$ and seismic moment $M_0$ is represented by a weighted line integral of $\tau_{ij}(x)$ over the x-axis:

$$M_i(x; M_0) = \int_{-\infty}^{\infty} p(x; x_i, M_0) \tau_{ij}(x) dx \quad (i = 1, 2, 3; j = 1, 2, 3),$$

(39)

where $p(x; x_i, M_0)$ is the 1-D marginal weighting function defined by

$$p(x; x_i, M_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x; x_i, M_0) dy dz$$

$$= \frac{M_0}{\sqrt{2\pi L}} \exp \left[ -\frac{1}{2L^2}(x - x_i)^2 \right].$$

(40)

Next, for a seismic event with its hypocentre $x^*$ and seismic moment $M_0^*$, we calculate the elements of CMT in the following way. We consider a local coordinate system for each seismic event, defined by three orthonormal vectors; a unit normal vector $n(x^*)$ specifying fault orientation, a unit direction vector $\nu(x^*)$ specifying fault slip, and the outer product of them $n(x^*) \times \nu(x^*)$. In the local coordinate system, denoting $\nu$, $n \times \nu$ and $n$ as the $x$, $y$ and $z$ axes, respectively, we can represent the CMT of any seismic event as

$$M^* = M_0^* \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$  

(41)

If the three orthonormal vectors (actually the normal vector $n$ and the slip direction vector $\nu$) are known, we can transform the local-coordinate representation of CMT in eq. (41) to the global-coordinate representation as

$$M(x^*; M_0^*) = R^T(x^*) M^* R(x^*).$$

(42)

Here, $R$ is the $3 \times 3$ rotation matrix composed of the three orthonormal vectors.

Figure 3. The true seismogenic stress field used for numerical experiments. Every stress component is normalized by the maximum shear stress.

Figure 4. The distribution of seismic events to generate synthetic CMT data. (a) The hypocentre distribution of seismic events on the x-axis. (b) The cumulative number $N$ of seismic events versus logarithmic seismic moment $\log M_0$. The $N - \log M_0$ distribution of seismic events obeys a power law in the range of $13.6 \leq \log M_0 \leq 16.6$.

4.1 Synthetic data

To make data sets for validity tests, first, we determine the hypocentres $x_0^*$ and the seismic moments $M_0^*$ of 1400 events by using uniform random numbers and random numbers that obey Gutenberg-Richter’s magnitude-frequency distribution, respectively. In Figs 4(a) and (b) we show the hypocentre distribution and the log-magnitude seismic moment–frequency distribution of generated events, respectively.

In the local coordinate system, denoting $\nu$, $n \times \nu$ and $n$ as the $x$, $y$ and $z$ axes, respectively, we can represent the CMT of any seismic event as

$$M^* = M_0^* \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$  

(41)

If the three orthonormal vectors (actually the normal vector $n$ and the slip direction vector $\nu$) are known, we can transform the local-coordinate representation of CMT in eq. (41) to the global-coordinate representation as

$$M(x^*; M_0^*) = R^T(x^*) M^* R(x^*).$$

(42)

Here, $R$ is the $3 \times 3$ rotation matrix composed of the three orthonormal vectors.
In Case 1, where every fault plane is aligned with the plane of maximum shear stress there, the normal vector \( \mathbf{n}(x_r^*) \) is directly determined from the true stress tensor \( \mathbf{T}(x_r^*) = [\tau_{ij}(x_r^*)] \) by solving the associated eigenvalue problem. Given the normal vector \( \mathbf{n}(x_r^*) \) and the true stress tensor \( \mathbf{T}(x_r^*) \), we can calculate the slip direction vector \( \nu(x_r^*) \) as

\[
\nu(x_r^*) = \mathbf{n}(x_r^*) \times [\mathbf{F}(x_r^*) \times \mathbf{n}(x_r^*)] / \| \mathbf{n}(x_r^*) \times [\mathbf{F}(x_r^*) \times \mathbf{n}(x_r^*)] \|.
\]

(43)

where \( \mathbf{F}(x_r^*) \) is the stress vector acting on the fault plane, defined by Cauchy’s formula as

\[
\mathbf{F}(x_r^*) = \mathbf{T}(x_r^*) \mathbf{n}(x_r^*).
\]

(44)

In Case 2, where fault orientation is random, we generate the random distribution of fault strike \((0^\circ \leq \theta \leq 360^\circ)\) and fault dip \((0^\circ \leq \delta \leq 90^\circ)\) for the set of 1400 events with the method proposed by Kagan (2005). Here, the fault strike is measured from the \(x\)-axis clockwise, and the fault dip is measured from the \(x\)-\(y\) plane downward. By using the fault strike \(\theta^n\) and the fault dip \(\delta^n\) the normal vector \(\mathbf{n}(x_r^*)\) is written as

\[
\mathbf{n}(x_r^*) = (-\sin \theta^n \sin \delta^n, \cos \theta^n \sin \delta^n, -\cos \delta^n).
\]

(45)

Given the normal vector \(\mathbf{n}(x_r^*)\) and the true stress tensor \(\mathbf{T}(x_r^*)\), we can calculate the slip direction vector \(\nu(x_r^*)\) by eq. (43). In Fig. 5 we show the distribution of fault strike, fault dip and fault slip direction ( rake ) generated through the above process.

Now, following eq. (42), we can transform the local-coordinate representation of CMT in eq. (41) to the global-coordinate representation. After the transformation, we normalize the CMT by its seismic moment \(M_0\). Finally, we add Gaussian errors with zero mean and variance \(\sigma^2 = 4.0 \times 10^{-2}\) to the elements of the normalized CMT in both cases, Cases 1 and 2. It should be noted that the theoretical CMT calculated from eq. (39) and the CMT data obtained by eq. (42) are different from each other. The difference is larger in Case 2 than in Case 1, because in Case 2 the pre-existing faults have random orientations. As a result, the effective values of \(\sigma^2\) become \(6.0 \times 10^{-2}\) and \(2.4 \times 10^{-1}\) in Case 1 and Case 2, respectively, which are 1.5 and 6 times as large as the given variance.

### 4.2 Numerical experiments

Now we apply the CMT data inversion method to the two synthetic data sets. In both cases the 1-D isseismic stress field \((-165 \text{ km} \leq x \leq 165 \text{ km})\) is represented by the superposition of 63 normalized cubic B-splines with 5 km equally spaced local supports as shown in Fig. 6. The effective stress-release distance \(L\) of the 1-D weighting function in eq. (40) is taken to be twice of the fault dimension of each seismic event, calculated from the empirical relation between seismic moment \(M_0\) and fault dimension \(L_f\):

\[
M_0 = \frac{\pi}{8} \Delta \tau L_f^3.
\]

(46)

Here, the average stress drop \(\Delta \tau\) is assumed to be 3 MPa for intraplate small- and medium-sized earthquakes.

In Figs 7(b) and (c) we show the patterns of inverted stress fields in Case 1 and Case 2, respectively, together with the pattern of the true stress field (Fig. 7(a)). Here, the focal spheres represent the earthquake mechanisms expected from the true or inverted stress fields. The colour scale indicates the closeness \(C\) of the inverted stress tensor \(\hat{\tau}_{ij}\) to the true stress tensor \(\tau_{ij}\), defined by

\[
C = \frac{\sum_{i=1}^{3} \sum_{j=1}^{3} \hat{\tau}_{ij} \tau_{ij}}{\left( \sum_{i=1}^{3} \sum_{j=1}^{3} \hat{\tau}_{ij}^2 \right)^{\frac{1}{2}} \left( \sum_{i=1}^{3} \sum_{j=1}^{3} \tau_{ij}^2 \right)^{\frac{1}{2}}}.
\]

(47)

In both cases, Case 1 and Case 2, the \(C\)-values are almost 1 over the whole region. This indicates that the patterns of the inverted stress fields almost completely agree with the pattern of the true stress field. The \(C\)-values in Case 2 are a little smaller than those in Case 1, because the synthetic data in Case 2 include larger noise than those in Case 1.

The inverted stress field strongly depends on the choice of the value of the hyper-parameter \(\sigma^2\) that controls the structure of the Bayesian model. In order to demonstrate this point we show the inverted stress fields for three different values of \(\sigma^2\) in Case 1. First, we show the inverted stress tensor components for the optimum value \(\hat{\sigma}^2 = 1.0 \times 10^4\) in Fig. 8, which should be compared with the true stress tensor components in Fig. 3. From the comparison of Fig. 8 with Fig. 3 we can see that the optimum solution explains properly the characteristic patterns of the true stress field. In Figs 9 and 10 we show the inverted stress tensor components for the inappropriate, too large or too small, values of \(\sigma^2\), respectively. When the value of \(\sigma^2\) is too large (by two orders larger than the optimum value), the solution explains the gross feature of the true
CMT data inversion using a Bayesian information

Figure 6. The normalized cubic B-splines to parameterize the 1-D seismogenic stress field for numerical experiments. In the present case M is taken to be 63.

Figure 7. The patterns of seismogenic stress fields represented by focal spheres. (a) The true stress field. (b) The inverted stress field in Case 1 (aligned orientations). (c) The inverted stress field in Case 2 (random orientations). The focal spheres represent the earthquake mechanisms expected from the true or inverted stress fields. The colour scale indicates the closeness $C$ of the inverted stress tensor to the true stress tensor, defined in eq. (47) of the text.

stress field but not the short wavelength variation. When the value of $\alpha^2$ is too small (by two orders smaller than the optimum value), on the other hand, the solution too much amplifies the short wavelength variation that comes from noise in data. The point is that all of them are the maximum likelihood solutions. The use of ABIC enables us to objectively select the optimum solution from among a number of maximum likelihood solutions.

The optimum values of $\sigma^2$ calculated from eq. (37) are $3.98 \times 10^{-2}$ in Case 1 and $1.89 \times 10^{-1}$ in Case 2, which almost agree with the true values of $6.00 \times 10^{-2}$ in Case 1 and $2.40 \times 10^{-1}$ in Case 2. This means that the optimum value of the hyper-parameter $\alpha^2$ has been correctly determined from observed (synthetic) data by using ABIC (Yabuki & Matsu’ura 1992). From the validity tests on the two extreme cases, Case 1 (aligned fault orientations) and Case 2 (random fault orientations), we may conclude that the present CMT data inversion method gives us unbiased reliable results for the pattern of seismogenic stress fields unless the distribution of fault orientations has a bias to the plane of maximum shear stress.

Figure 8. The inverted stress tensor components in Case 1: The optimum value of the hyper-parameter $\alpha^2$.

Figure 9. The inverted stress tensor components in Case 1: The too large value of the hyper-parameter $\alpha^2$. 

C⃝ 2007 The Authors, GJI, 172, 674–685
Journal compilation C⃝ 2007 RAS
5 APPLICATION TO ACTUAL CMT DATA IN NORTHEAST JAPAN

In this section we apply the method of CMT data inversion to the actual data in northeast Japan, where the Pacific Plate is descending beneath the North American Plate at the Japan Trench in the direction of N68°W as shown in Fig. 11(a). Our purpose is to reveal the across-arc stress pattern associated with plate subduction. For this purpose we consider a sectorial area bounded by grey broken lines, and take a vertical plane parallel to plate convergence at its centre. On the vertical plane we take the x-axis horizontally northward and the z-axis vertically downward. The origin \((x = 0)\) is located at the trench axis. Then, we took out about 1000 seismic events with magnitude 3.5–5.0 in the sectorial area from the National Research Institute for Earth Science and Disaster Prevention (NIED) Seismic Moment Tensor Catalogue (1997–2005), which are indicated by grey dots in Fig. 11(a). All of these events are projected on the vertical plane as shown in Fig. 11(b). We use the CMT data of these events for inversion analysis. In Fig. 12 we show the magnitude-frequency diagram of seismic events in northeast Japan \((N35°–40°, E137°–145°)\) for reference. From this diagram we can confirm that seismic events in the magnitude range of 3.5–5.0 follow the Gutenberg–Richter’s magnitude-frequency distribution.

In order to represent the 2-D seismogenic stress field, we divided the model region \((-130 \text{ km} \leq x \leq 330 \text{ km}, -30 \text{ km} \leq z \leq 130 \text{ km})\) into 46 x 16 subsections, and distributed 559 \((= 43 \times 13)\) bi-cubic B-splines \(\Phi_i(x, z)\) with 10 km equally spaced local supports in both the x and z directions. In the 2-D problem the CMT of a seismic event with the hypocentre \((x_i, z_i)\) and the seismic moment \(M_0\) is represented by a weighted surface integral of the true but unknown stress field \(\tau_{ij}(x, z)\) over the x-z plane:

\[
M_{ij}(x_i, z_i; M_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, z; x_i, z_i, M_0) \tau_{ij}(x, z) \, dx \, dz \quad (i = 1, 2, 3; j = 1, 2, 3),
\]

(48)

where \(p(x, z; x_i, z_i, M_0)\) is the 2-D marginal weighting function defined by

\[
p(x, z; x_i, z_i, M_0) = \int_{-\infty}^{\infty} p(x; x_i, M_0) \, dy = \frac{M_0}{2\pi L_2} \exp \left\{ -\frac{1}{2L_2^2} \left[ (x - x_i)^2 + (z - z_i)^2 \right] \right\}.
\]

(49)

In the present analysis we took the effective stress-release distance \(L\) to be twice of the fault dimension \(L\) calculated from eq. (46). We applied the inversion algorithm in the Section 3.3 to the CMT data of about 1000 seismic events that follow the Gutenberg–Richter’s magnitude-frequency distribution in the magnitude range of 3.5–5.0. In this case the optimum values of the hyper-parameters, \(\alpha^2\) and \(\sigma^2\), were determined as 1.0 and 1.25 x 10\(^{-1}\), respectively. In Fig. 13(a) we show the inverted seismogenic stress field with the focal sphere representation, where the focal mechanism patterns (backside hemisphere) expected from the inverted stress field are...
projected on the vertical plane. In Fig. 13(b) we show the directions of the maximum compressive principal stress axes projected on the vertical plane. The colour scale indicates the square root of the average variance of estimation errors for six stress components, calculated from eq. (34). From Figs 13(a) and (b) we can expect the occurrence of normal faulting with a strike parallel to the trench axis in the oceanic plate beneath the outer rise and reverse faulting in the continental crust beneath the island arc and the shallow part of the descending oceanic plate. These stress patterns are consistent with those inferred from geophysical and geological observations (e.g. Yoshii 1979).

6 DISCUSSION AND CONCLUSIONS

In order to extract reliable information about seismogenic stress fields we developed a method of CMT data inversion using ABIC. If we assume isotropic linear elasticity in source regions, we can transform the CMT data into fault slip data, but this process is always accompanied by contamination of original information and ambiguity in choosing true fault planes. This is one of the reasons why we use CMT data directly in stress inversion. Another more essential reason is that CMT data provide direct information about stress release in the elastic region surrounding the source. In contrast, as pointed out by Twiss & Unruh (1998), the information provided by fault slip data is essentially strain changes, because we need to assume some constitutive law (e.g. isotropic linear elasticity) in source regions to obtain the fault slip data.

The CMT of a seismic event is usually represented by the surface integral of 2-D moment tensor density over a rupture area. As demonstrated in the Section 2.1, this usual representation can be transformed into the volume integral of stress release over a finite elastic region surrounding the dynamic rupture area by using Gauss’ divergence theorem. From the viewpoint of earthquake generation physics, the volume integral representation of CMT is more essential, because dynamic rupture growth is controlled by energy flow into the rupture zone from the surrounding region storing elastic strain energy. By using the volume integral representation of CMT we can explicitly relate an observed CMT datum with stress release in a finite elastic region surrounding the source region without any knowledge of actual source processes.

In the Section 3.1, in order to relate the stress release due to a seismic event with the true seismogenic stress field, we introduced a 3-D Gaussian-type weighting function with its peak at the hypocentre and the variance proportional to the two-thirds power of the seismic moment. Through this weighting function we extract information about the true seismogenic stress field. Since the peak value...
of the weighting function does not depend on the scale of events (seismic moment), the total amount of information obtained from a CMT solution appears to be proportional to the scale of events. However, from the consideration of modelling errors in the CMT analysis of seismic events, we may conclude that the data error of CMT is proportional to the scale of events. Therefore, the effective amount of information obtained from a CMT solution is independent of the scale of events. In other words, in CMT data inversion and also in fault slip data inversion, all seismic events have equal weight regardless of seismic moment, just like hypocentre location data. This means that the result of CMT data inversion represents the ensemble average of equal weight information from various seismic events.

The CMT data inversion method is based on the Bayesian statistical inference algorithm developed by Yabuki & Matsu'ura (1992). In this method, unlike the conventional stress inversion methods, we need not partition the study area in advance. Instead, we represent the concerned seismogenic stress field by the superposition of a finite number of tri-cubic B-splines to obtain parameterized observation equations. In general, the inversion results depend on the model used for analysis; in the present case, the number and distribution of basis functions. In order to avoid the model dependence of inversion results we need to take a sufficiently large number of basis functions (the number of model parameters). However, since the number of data is finite, the increase of the model parameters leads to numerical instability. So, introducing prior constraint on the roughness of seismogenic stress fields and combining it with observed data by Bayes’ rule, we constructed a stochastic model with hierarchic, highly flexible structure controlled by hyper-parameters. Then, the Bayesian model consists of a number of parametric models. The optimum values of hyper-parameters are objectively determined by using ABIC. In other words, the use of ABIC enables us to objectively select the optimum parametric model from among a number of possible parametric models. In the conventional methods, on the other hand, one parametric model with rigid structure is specified in advance through area partitioning. The area partitioning will be done carefully by considering hypocentre distribution and geological structure (Hardebeck & Michael 2004). Nevertheless, the conventional method cannot avoid the model dependence of the inversion results.

From the viewpoint of statistical inference, the conventional stress inversion methods can be regarded as a special case of the present inversion method except for the use of fault slip data instead of CMT data. Actually, if we take the weighting function for volume integral to be a delta function at the hypocentre, adopt tri-boxcar functions as the basis functions, and neglect the prior smoothness constraint, the present method is reduced to the conventional method. A kind of damped least squares method proposed by Hardebeck & Michael (2006) can be also regarded as a special case of the present inversion method. It is similar to the present method in using a parametric model with hierarchic, flexible structure controlled by an adjustable damping parameter in inversion analysis. If we take the weighting function for volume integral to be a delta function at the hypocentre, adopt tri-boxcar functions as the basis functions, introduce the flatness constraint instead of the smoothness constraint, and use the trade-off between data misfit and model length instead of ABIC, the present method is reduced to the Hardebeck–Michael method. The essential defect of their method is that the meaning of the trade-off criterion is not clear: the units to measure the data misfit and the model length are different from each other. Jackson (1979) and Jackson & Matsu’ura (1985) have resolved this problem by introducing the concept of prior information in Bayesian statistical inference.

In Section 4, we tested the validity of the CMT data inversion method through numerical experiments on two synthetic data sets, assuming the distribution of fault orientations to be aligned with the maximum shear stress plane in one case and to be random in the other case. In both cases the inverted patterns almost completely agree with the pattern of the true stress field. Since the actual distribution of fault orientations will be between these two extreme cases, the validity tests show that the present inversion method gives us unbiased reliable results for the pattern of seismogenic stress fields unless the distribution of fault orientations has a bias to the plane of maximum shear stress. If the distribution of fault orientations has the bias, as in the case of seismic events within the damaged zones of active faults, the present inversion method will give a biased result; that is, the principal axes of the inverted stress field will be systematically rotated from those of the true regional stress field. In reality, however, the orientations of most active faults are consistent with those of the maximum shear stress there (e.g. Thatcher & Hill 1991; Iio 1997).

In Section 5, we applied the method of CMT data inversion to actual data in northeast Japan, and obtained the stress pattern consistent with geophysical and geological observations. If we have a sufficiently long-term and wide-coverage data, the stress release pattern inverted from CMT data can be regarded as the long-term stress accumulation pattern, because stress release and accumulation must balance with each other on a long-term average.

ACKNOWLEDGMENTS

We thank Eiichi Fukuyama for his helpful comments and providing useful information on the NIED seismic moment tensor catalogue by National Research Institute for Earth Science and Disaster Prevention. We also thank Yehuda Ben-Zion and anonymous reviewers for their useful comments and suggestions to improve our manuscript.

REFERENCES


