1 Influence of the amount and the geometry of the geofluid on the seismic velocity

The seismic wave velocities ($V_P$ and $V_S$) of a solid-fluid composite media are generally expressed as functions of the intrinsic elastic parameters of the solid framework and the pore-filled fluid, the fluid volume fraction, and the pore geometry (e.g., O’Connell and Budiansky 1974; Mavko 1980; Takei 1998). Recently, Takei (2002) proposed a unified formulation of $V_P$ and $V_S$ as a function of the effective aspect ratio ($\alpha$) and the fluid volume fraction, that is, the porosity ($\phi$), of the fluid-filled pores, which can be applied to a wide variety of pore shapes. The functions are expressed as

$$V_P = \frac{\sqrt{K_{eff}}/k + (4\gamma/3)N/\mu}{\sqrt{1 + 4\gamma/3}} \sqrt{\rho/\rho_0},$$  
(A.1)

$$V_S = \frac{\sqrt{N/\mu}}{\sqrt{\rho/\rho_0}},$$  
(A.2)

where

$$\frac{K_{eff}}{k} = \frac{K_b}{k} + \frac{(1 - K_b/k)^2}{1 - \phi - K_b/k + \phi k/k_f},$$  
(A.3)

and $\gamma = \mu/k$. Here, $V_P^0$ and $V_S^0$ are the seismic wave velocities of the dry rock; $k$, $\mu$, and $\rho$ are the bulk modulus, the shear modulus, and the density of the solid phase; $k_f$ is the bulk modulus of the fluid phase; $\phi$ and $\rho_f$ are the fluid volume fraction and the total density of the solid and fluid portions; $K_b$ and $N$ are the bulk and shear moduli of the solid framework; and $K_{eff}$ is the effective bulk modulus of the solid-fluid composite media.

The ratios of the bulk and shear moduli of the solid framework to those of the solid phase can be written using $\phi$ and $\alpha$ as

$$\frac{K_b}{k} = 1 - \phi \Lambda_{K_b}(\alpha),$$  
(A.4)

$$\frac{N}{\mu} = 1 - \phi \Lambda_N(\alpha),$$  
(A.5)

where the slopes, $\Lambda_{K_b}$ and $\Lambda_N$, are functions of $\alpha$ and take positive values larger than unity. The explicit form of the function is shown in Figure 3 of Takei (2002).

Assuming that the type and compositions of the rock and fluid are fixed and that the thermal and pressure effects are negligible, these functions can be simplified, as follows:

$$V_P = f_P(\phi, \alpha),$$  
(A.6)

$$V_S = f_S(\phi, \alpha).$$  
(A.7)

In this study, we calculated these functions for each set of $\phi$ and $\alpha$, as follows: $\Lambda_{K_b}$ and $\Lambda_N$ were calculated using the fitting curves for Figure 3 in Takei (2002), and then $V_P$ and $V_S$ were obtained by substituting $\phi$ and Eqs. (A.4) and (A.5) into the constitutive function of Eqs. (A.1) and (A.2).

2 Estimation of the hyperparameters

In this study, the set of hyperparameters that minimize the free energy function $F(\theta)$ was obtained by the method of steepest descent. The slope of the free energy $F(\theta)$ with respect to the hyperparameter $\sigma^2$ can be rewritten as

$$\frac{\partial F(\theta)}{\partial \sigma^2} = \frac{\int \int \frac{\partial E(\phi, \alpha; \theta)}{\partial \sigma^2} \exp \{-E(\phi, \alpha; \theta)\} d\phi d\alpha}{\int \int \exp \{-E(\phi, \alpha; \theta)\} d\phi d\alpha},$$  
(A.8)

where $\langle f(x) \rangle_{p(x)}$ indicates the expectation value of $A$ for the probability distribution $p(x)$. In this study, we used the Metropolis algorithm (Metropolis et al. 1953), which is a type of Markov chain Monte Carlo (MCMC) method, in order to numerically calculate the expectation values of $\partial E(\phi, \alpha; \theta)/\partial \sigma^2$ for the posterior probability distribution $p(\phi, \alpha|V_P, V_S)$. In the calculations of the MCMC method, numerous candidate sets of $\phi$ and $\alpha$ are generated to produce the posterior probability distribution $p(\phi, \alpha|V_P, V_S)$. As a result, during the process of estimating the hyperparameters, we were able to concurrently find the solutions of $\phi$ and $\alpha$ with the greatest likelihood, that is, the maximum a posteriori (MAP) solution, which maximizes $p(\phi, \alpha|V_P, V_S)$.

3 Metropolis algorithm

Markov chain Monte Carlo (MCMC) methods are very efficient mathematical tools for the numerical calculation of multidimensional integrals or expectation values. In this study, a MCMC method was implemented by the Metropolis algorithm (Metropolis et al. 1953) in order to estimate the slopes of the free energy $F(\theta)$ that were used in the method of steepest descent (Eq. (A.8)).

Suppose we now want to calculate the expectation value of the function $f(x)$ for the parameter $x$, which has probability $p(x)$. The expectation value is obtained by integrating the function $f(x)$ weighted by the probability $p(x)$, as

$$\langle f(x) \rangle_{p(x)} = \int f(x) \cdot p(x) dx.$$  
(A.9)

In MCMC methods, a multidimensional Monte Carlo integration is conducted by sampling from probability distributions based on a Markov chain random walk that has the same distribution as the desired equilibrium distribution. The Metropolis algorithm is extremely simple, as follows,

...
1. Arbitrarily choose an initial model $x_0$.

2. Randomly perturb the present guess $x_t$ to obtain a trial model $x'$. 

3. Decide the next state by comparing the present probability density $p(x_t)$ with the trial $p(x')$.

$$x_{t+1} = \begin{cases} x' & \text{if } p(x') > p(x_t), \\ x_t & \text{otherwise.} \end{cases}$$

where $r$ is a random number from the interval $[0,1]$.

4. Repeat steps 2 and 3.

In step 2, the perturbed model $x'$ is accepted as the next model if it is a better model than $x_t$, that is, if $p(x') > p(x_t)$. On the other hand, even if a trial move does not result in a better model, that is, if $p(x') < p(x_t)$, there is a finite possibility $r$ that the trial is accepted as the next move. Repeated random walks explore almost all possible models which have reasonably high $p(x)$.

In this study, we used the above Metropolis algorithm, but the computational cost was still high because numerous samplings were repeated in order to reproduce the high-dimensional probabilistic distributions. Recently, a deterministic algorithm that linearizes the system was developed in order to achieve more efficient estimation of the hyperparameters with less computational cost (Ohno et al. 2012).

4 A synthetic test for power-law type heterogeneities

We checked the effectiveness of the proposed method for the case of both Gaussian and non-Gaussian spatial heterogeneities. In this study, we consider the case in which the unknown parameters obey spatial heterogeneities that follow a power law, since the spatial heterogeneities of the earth’s interior are generally considered to be of this type (e.g., Sato et al. 2012).

Following the methods of Roth and Korn (1993) and Sato et al. (2012), a 2-D discrete inverse Fourier transform was used to generate spatial heterogeneities from power spectral densities which obey a power law. The spatial distributions of $\phi$ and $\alpha$ were synthesized using these heterogeneities (Fig. A.1a). In the same way as was done for the Gaussian case in the main text, the velocity structure, $V_P$ and $V_S$, was generated by the constructive function and by adding noise (Fig. A.1b).

As shown in Fig. A.2, each hyperparameter converges to the true value as the method of steepest descent progresses. The root-mean-square (RMS) errors of the estimated $\phi$ and $\alpha$ are shown in Fig. A.3. They decrease and converge to small values as the number of trials increases. This implies that the accuracy is increasing.

Figure A.4 (a) shows the spatial distributions of $\phi$ and $\alpha$ estimated from the above synthetic velocity structure model using the MRF model. The $\phi$ and $\alpha$ distributions calculated by numerically solving Eq. (A.7) from the observed $V_P$ and $V_S$, (the deterministic method) are shown in Fig. A.4 (b) for comparison. The distributions of $\phi$ and $\alpha$ calculated by the deterministic method are directly affected by the noise in the observed seismic velocities, and thus they are too jagged for the true distributions to be determined. Although the deterministically estimated distributions of $\phi$ and $\alpha$ are rough and have large errors, those of the MAP solutions obtained by the MRF model are much smoother and more accurate.

The results of the estimation show the effectiveness of the proposed model for power-law inhomogeneities, despite the assumption that the prior probability is Gaussian. From these results, we infer that a more precise estimation could be expected if we were to use a prior probability that corresponded to a power-law inhomogeneity. We note that further studies are needed for other types of non-Gaussian inhomogeneities.

Figure A.1: Synthetic data used in the inversion test. (a) Synthetic distributions of the porosity $\phi$ and aspect ratio $\alpha$, which are to be estimated. They were generated by random walks, and they have the variances of continuity $\sigma_0^2$ and $\sigma_0^2$, equal to 0.0012 and 0.0152, respectively. (b) Observational data of $V_P$ and $V_S$. These were obtained by using Eq. (A.7) with Gaussian noise, and the variances were set at $(\sigma_0^2, \sigma_S^2) = (0.1^2, 0.06^2)$.

Figure A.2: Estimated behavior of the hyperparameters during the method of steepest descent. (a) Variance of the continuity of the porosity ($\sigma_0^2$) and the aspect ratio ($\sigma_0^2$); and (b) variances of the noise in the observational data, $V_P$ and $V_S$. The red points indicate the true value of the variances.

References


Figure A. 3: Behavior of the root-mean-square errors of the porosity ($\phi$) and the aspect ratio ($\alpha$). The trial number is the iteration number for the method of steepest descent.

(a) MRF model

(b) Deterministic method

Figure A. 4: Estimated fluid distribution (porosity $\phi$ and aspect ratio $\alpha$). (a) MRF model (this study), and (b) deterministic method.


