Extended time-term method for identifying lateral structural variations from seismic refraction data

Takaya Iwasaki

Earthquake Research Institute, the University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-0032, Japan

A time-term method is a simple travel time inversion for seismic refraction crustal studies. Although this method is based on simple travel time calculation, it is robust in determining a basement velocity under the complicated surface structure. In the present paper, this method is extended to detect lateral velocity variation in the crystalline crust. We consider a model where a basement is composed of segments with a different velocity. Defining a likelihood function for unknown time-terms and refractor velocities, their best estimates are obtained by minimizing Akaike's Bayesian Information Criterion (ABIC). To realize accurate velocity estimation, we also incorporate an effect of vertical velocity gradient within the refractor, which was ignored in the conventional algorithms. Numerical tests undertaken for realistic crustal structures show that our method successfully detects lateral velocity variation in a basement even for a case with undulated surface layers. It is particularly stable and reliable if the surface layers are not thick and their velocities are more than 30–40% smaller than the refractor velocities. Then, the refractor velocities are almost decoupled to the time-terms in the process of inversion. The incorporation of the effect of the velocity gradient is inevitably important when travel time data at far distances are included in the inversion.

1. Introduction

In the last two decades, intensive seismic refraction experiments have been undertaken in the Japanese Islands (Yoshii, 1994). Usually, a refraction profile crosses several geological units with different physical properties. Such structural variations provide important keys in understanding crustal evolution and deformation processes of the surveyed area. In the Japanese Islands, however, the upper crust is often covered with highly deformed sedimentary layers, for which a useful technique of ray-tracing (e.g. Červený *et al.*, 1977) sometimes does not work. This prevents us from investigating the lateral velocity change in the crystalline part of the crust. The purpose of this paper is to provide a robust and reliable method for detecting lateral heterogeneity of the upper crust by modifying a "time-term method".

The time-term method is a simple travel-time inversion developed by Scheidegger and Willmore (1957) and was widely used for seismic refraction crustal studies in the 1960–70s (Willmore and Bancroft, 1960; Berry and West, 1966; Meru, 1966; Smith *et al.*, 1966; Yoshii and Asano, 1972). This method has several advantages in spite of its crude travel-time approximation. The first one is the computational stability. A sophisticated travel-time inversion based on the ray-tracing method is sometimes too sensitive to small-scale structural changes because of its high frequency approximation. Such instability never happens in the time-term method because of its simple travel-time calculation. The second advantage is the computational fastness due to the linear observation equations. Moreover, this method is

quite robust for estimating a refractor velocity. Hence, with appropriate modifications, the time-term method has a potentiality to be a useful tool for detecting lateral velocity variation of a refractor. To identify subtle velocity variation, which sometimes occurs in the crystalline part of the crust, an inversion with a sophisticated error analysis is inevitably important.

This paper presents an algorithm of the time-term method based on an inverse theory with the use of Akaike's Bayesian Information Criterion (ABIC) (Akaike, 1980; Yabuki and Matsu'ura, 1992). By a similar approach, Koketsu and Higashi (1992) studied the 3-D basement topography in the Tokyo metropolitan area, Japan. Our algorithm, on the other hand, aims at detecting lateral variation in a basement velocity. For the accurate velocity estimation, the effect of velocity gradient in the basement is incorporated. The validity and effectiveness of our method are intensively examined through several numerical tests for realistic crustal models.

2. Mathematical Formulation
2.1 Observation equations

Consider a structure model where a basement (a refractor) is covered with a surface layer (Fig. 1). We take m_s shots and m_r receivers on a profile line extending from $x = 0$ to L. The basement beneath the profile is divided into m_v segments with a velocity of V_k ($k = 1, \ldots, m_v$). Then, a travel time , t_{ij}, of a refracted wave between the i-th shot and the j-th receiver is approximated as

$$t_{ij} = a_i + b_j + \sum_{k=K_i}^{L_j} R_{ij}^k / V_k. \tag{1}$$
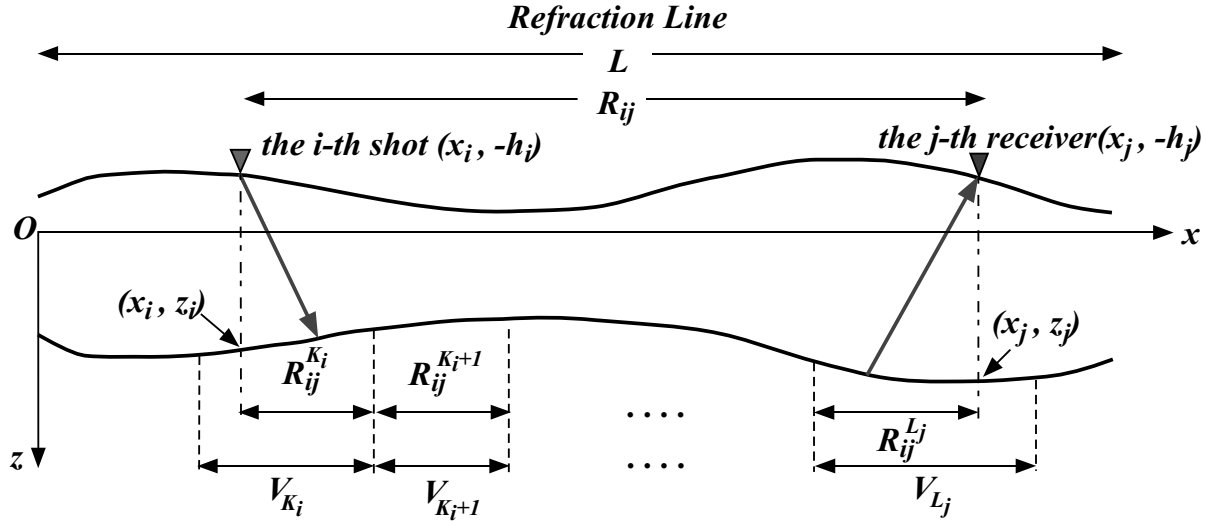
Refraction Line

Fig. 1. Geometry and notations in the present study. A refractor along a profile line is divided into segments with a different velocity.

Here, a_i and b_j are called *time-terms* of the i-th shot and the j-th receiver, respectively, and R_{ij}^k is the path length over which the ray from the K_i to the L_j-th velocity segment overlaps the k-th one (Fig. 1). Denoting locations of the shot and receiver as x_i and x'_j, a_i and b_j are expressed as

$$a_i = \int_{-h_i}^{z_i} \sqrt{v^{-2}(x_i, z) - U_i^{-2}} dz,$$
$$b_j = \int_{-h'_j}^{z'_j} \sqrt{v^{-2}(x'_j, z) - U_j'^{-2}} dz, \quad (2)$$

where $v(x, z)$ is a velocity function above the refractor, and h_i (h'_j), z_i (z'_j) and U_i (U'_j) are a topographic height, a basement depth and a refractor velocity at the shot (receiver) point, respectively. Equation (1) can be solved with some constraints by a least squares method (Meru, 1966; Hearn and Clayton, 1986) or a damped least-squares method (Pavlis, 1986).

In this paper, we formulate an algorithm of inversion using Bayes' theorem (Yabuki and Matsu'ura, 1992). For this purpose, we rewrite Eq. (1) into a stochastic expression,

$$\mathbf{t} = \mathbf{Am} + \mathbf{e}, \quad (3)$$

where

$$\mathbf{t}^T = (t_{11}, t_{12}, \ldots, t_{m_s m_r}),$$
$$\mathbf{m}^T = (a_1, a_2, \ldots, a_{m_s}; b_1, b_2, \ldots, b_{m_r}; \quad (4)$$
$$V_1^{-1}, V_2^{-1}, \ldots, V_{m_v}^{-1}),$$

and $\mathbf{e}$ represents random errors in travel time data, $\mathbf{t}$, caused in our measurement and inadequate modelling. We take $\mathbf{e}$ to be Gaussian with zero mean and covariance, $\sigma^2 \mathbf{E}$,

$$\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{E}), \quad (5)$$

where σ^2 is an unknown scaling factor (hyperparameter). Hereafter, we denote dimensions of $\mathbf{t}$ and $\mathbf{m}$ in Eq. (3) by n and m, respectively. Assuming that the p-th element of the

vector $\mathbf{t}$ corresponds to the travel time between the i_p-th shot and the j_p-th receiver, the elements of A_{pq} of the matrix A are expressed as

$$A_{pq} = \begin{cases} 1 & (q = i_p \text{ or } q = m_s + j_p), \\ R_{i_p j_p}^{q-(m_s+m_r)} & (m_s + m_r + K_{i_p} \\ & \leq q \leq m_s + m_r + L_{j_p}), \\ 0 & (\text{otherwise}). \end{cases} \quad (6)$$

From Eqs. (3) and (5), a stochastic model relating the data, $\mathbf{t}$, and the model parameter, $\mathbf{m}$, is given as

$$p(\mathbf{t} \mid \mathbf{m}; \sigma^2) = (2\pi\sigma^2)^{-n/2} |\mathbf{E}|^{-1/2}$$
$$\times \exp\left[-\frac{1}{2\sigma^2} (\mathbf{t} - \mathbf{Am})^T \mathbf{E}^{-1} (\mathbf{t} - \mathbf{Am}) \right], \quad (7)$$

where $|\mathbf{E}|$ is an absolute value of the determinant of $\mathbf{E}$.

2.2 Prior constraints

A solution for Eq. (1) or (3) is not uniquely determined. For a certain solution of a_{0i} and b_{0j}, $a_{0i} + c$ and $b_{0j} - c$ also satisfy the same equation for an arbitrary constant of c. In a usual seismic refraction experiment, we set a receiver in the vicinity of a certain shot point. Hence, we can impose the following constraints for n_c pairs of shot and receiver,

$$a_{i_p} - b_{j_p} = \beta_p \quad (p = 1, \ldots, n_c), \quad (8)$$

where β_p is a travel time difference between the i_p-th shot and the j_p-th receiver. If these shot and receiver points completely coincide with each other, β_p is set to zero. In a marine profile with surface shots and ocean bottom seismograms, β_p represents a travel time delay for the water column beneath the i_p-th shot.

We also incorporate initial guess, $\mathbf{m}_0$, on the model parameters of $\mathbf{m}$,

$$\mathbf{m} = \mathbf{m}_0. \quad (9)$$

Equations (8) and (9) are combined into a stochastic form,

$$c_1 = B_1 m + d_1, \tag{10}$$

where

$$c_1^T = (\beta_1, \beta_2, \ldots, \beta_{n_c}; m_{o1}, m_{o2}, \ldots, m_{om}), \tag{11}$$

and B_1 is a $(n_c + m) \times m$ matrix expressed by

$$B_{1_{pq}} = \begin{cases} 1 & (1 \le p \le n_c, \ q = i_p), \\ -1 & (1 \le p \le n_c, \ q = m_s + j_p), \\ 1 & (n_c + 1 \le p \le n_c + m, \\ & q = p - n_c), \\ 0 & (\text{otherwise}) \end{cases} \tag{12}$$

In Eq. (10), d_1 represents Gaussian errors with covariance of $\rho_1^2 D_1$,

$$d_1 \sim N(0, \rho_1^2 D_1), \tag{13}$$

where ρ_1^2 is a hyperparameter. Denoting the rank of the matrix D_1 by r_1, the probability density function (pdf) for Eq. (10) is written as

$$p(m; \rho_1^2) = (2\pi\rho_1^2)^{-r_1/2} |D_1|^{-1/2}$$
$$\times \exp\left[-\frac{1}{2\rho_1^2} (c_1 - B_1 m)^T D_1^{-1} (c_1 - B_1 m) \right]. \tag{14}$$

2.3 Smoothness constraint

As shown in Section 3, a least squares or a damped least squares solution for Eq. (3) sometimes shows unexpected oscillation due to data noise. In order to avoid this effect, we impose the following smoothness conditions on the time-terms, b_j,

$$b_{j-1} - 2b_j + b_{j+1} = 0 \qquad (j = 2, \ldots, m_r - 1), \tag{15}$$

which is rewritten into the following stochastic expression,

$$0 = B_2 m + d_2, \tag{16}$$

with

$$d_2 \sim N(0, \rho_2^2 D_2), \tag{17}$$

where ρ_2^2 is a hyperparameter controlling the covariance of D_2. In Eq. (16), B_2 is a $m_r - 2 \times m$ matrix whose elements are given by

$$B_{2_{pq}} = \begin{cases} 1 & (q = m_s + p), \\ -2 & (q = m_s + p + 1), \\ 1 & (q = m_s + p + 2), \\ 0 & (\text{otherwise}) \end{cases} \tag{18}$$

Denoting the rank of the matrix D_2 by r_2, the pdf for Eq. (16) is written as

$$p(m; \rho_2^2) = (2\pi\rho_2^2)^{-r_2/2} |D_2|^{-1/2}$$
$$\times \exp\left[-\frac{1}{2\rho_2^2} (B_2 m)^T D_2^{-1} (B_2 m) \right]. \tag{19}$$

In a case with equally spaced receivers, B_2 is a discrete form of a differential operator, d^2/dx^2. Then, $(B_2 m)^T D_2^{-1} (B_2 m)$ in Eq. (19) corresponds to a quantity, Φ,

$$\Phi = \int_0^L \left(\frac{d^2 b(\xi)}{d\xi^2} \right)^2 w(\xi) d\xi, \tag{20}$$

where $b(\xi)$ is a continuous time-term function, and $w(\xi)$ a weight function. This is called a roughness penalty incorporated in a likelihood function of an inversion analysis to suppress the oscillation of model parameters (e.g. Koketsu and Higashi, 1992; Yabuki and Matsu'ura, 1992).

2.4 Inversion procedure

According to Bayes' theorem, a likelihood function for m is constructed from the prior pdfs in Eqs. (14) and (19) and the data distribution in Eq. (7),

$$l(m; \sigma^2, \rho_1^2, \rho_2^2 \mid d)$$
$$= p(t \mid m; \sigma^2) p(m; \rho_1^2) p(m; \rho_2^2). \tag{21}$$

Introducing new parameters of $\alpha_1 = \sigma/\rho_1$ and $\alpha_2 = \sigma/\rho_2$, it is rewritten as

$$l(m; \sigma^2, \alpha_1^2, \alpha_2^2 \mid d) = (2\pi\sigma^2)^{-(n+r_1+r_2)/2} (\alpha_1^2)^{r_1/2} (\alpha_2^2)^{r_2/2}$$
$$\times |E|^{-1/2} |D_1|^{-1/2} |D_2|^{-1/2}$$
$$\times \exp\left[-\frac{1}{2\sigma^2} s(m) \right], \tag{22}$$

with

$$s(m) = (t - Am)^T E^{-1} (t - Am)$$
$$+ \alpha_1^2 (c_1 - B_1 m)^T D_1^{-1} (c_1 - B_1 m)$$
$$+ \alpha_2^2 (B_2 m)^T D_2^{-1} (B_2 m). \tag{23}$$

With σ^2, α_1^2 and α_2^2 fixed at certain values, the above likelihood function has a maximum at

$$m^* = (A^T E^{-1} A + \alpha_1^2 B_1^T D_1^{-1} B_1 + \alpha_2^2 B_2^T D_2^{-1} B_2)^{-1}$$
$$\times (A^T E^{-1} t + \alpha_1^2 B_1^T D_1^{-1} c_1). \tag{24}$$

This is a well-known damped least squares solution. A fundamental problem for this solution is how to find out the best hyperparameters (damping factors). From the principle of the entropy maximization, Akaike (1980) proposed a Bayesian Information Criterion (ABIC) for determining the best hyperparameters,

$$\text{ABIC} = (-2) L(\sigma^2, \alpha_1^2, \alpha_2^2), \tag{25}$$

with

$$L(\sigma^2, \alpha_1^2, \alpha_2^2) = \int l(m; \sigma^2, \alpha_1^2, \alpha_2^2 \mid d) dm. \tag{26}$$

Here, $L(\sigma^2, \alpha_1^2, \alpha_2^2)$ is called a marginal likelihood of the hyperparameters. The best estimates of σ^2, α_1^2 and α_2^2 are obtained by minimizing ABIC. This requires

$$\partial L(\sigma^2, \alpha_1^2, \alpha_2^2)/\partial\sigma^2 = \partial L(\sigma^2, \alpha_1^2, \alpha_2^2)/\partial\alpha_1^2$$
$$= \partial L(\sigma^2, \alpha_1^2, \alpha_2^2)/\partial\alpha_2^2 = 0. \tag{27}$$

From the first equation of Eqs. (27), we obtain

$$\sigma^2 = s(\boldsymbol{m}^*)/(n + r_1 + r_2 - m), \qquad (28)$$

and

$$\begin{aligned}
\text{ABIC}(\alpha_1^2, \alpha_2^2) &= (n + r_1 + r_2 - m)\log s(\boldsymbol{m}^*) \\
&\quad - r_1 \log(\alpha_1^2) - r_2 \log(\alpha_2^2) \\
&\quad + \log|(A^T E^{-1} A + \alpha_1^2 B_1^T D_1^{-1} B_1 \\
&\quad + \alpha_2^2 B_2^T D_2^{-1} B_2)^{-1}| + C. \qquad (29)
\end{aligned}$$

Denoting the value of α_1^2 and α_2^2 minimizing Eq. (29) by $\hat{\alpha}_1^2$ and $\hat{\alpha}_2^2$, the best estimates of m and σ^2 are respectively expressed as

$$\begin{aligned}
\hat{\boldsymbol{m}} &= (A^T E^{-1} A + \hat{\alpha}_1^2 B_1^T D_1^{-1} B_1 + \hat{\alpha}_2^2 B_2^T D_2^{-1} B_2)^{-1} \\
&\quad \times (A^T E^{-1} t + \hat{\alpha}_1^2 B_1^t D_1^{-1} c_1), \qquad (30)
\end{aligned}$$

and

$$\hat{\sigma}^2 = s(\hat{\boldsymbol{m}})/(n + r_1 + r_2 - m). \qquad (31)$$

Our computer program directly calculates the inverse matrix appearing in Eq. (30), from which the resolution and covariance matrices are constructed. Estimation errors of $\hat{\boldsymbol{m}}$, which are essentially important for identifying the lateral variation of refractor velocity, are evaluated from this covariance matrix.

2.5 Velocity gradient

In the conventional time-term method, the velocity gradient of the refractor is not taken into account. For the accurate velocity estimation, we incorporate this effect after Zhao (1993). In a case with a velocity function of $v = v_0 + \eta z$, a travel time, t, at an offset of R is written by

$$t = \frac{2}{\eta} \sinh^{-1}\left(\frac{\eta}{2v_0} R\right). \qquad (32)$$

For $\eta R/2v_0 \ll 1$, it is expanded as

$$t \sim \frac{R}{v_0} - \frac{R^3}{24}\left(\frac{\eta^2}{v_0^3}\right). \qquad (33)$$

The second term in the above equation expresses the correction term for the velocity gradient. After Eq. (33), we modify Eq. (3) to approximate a case with lateral velocity variation,

$$t_{ij} = a_i + b_j + \sum_{k=K_i}^{L_j} R_{ij}^k / V_k - \frac{R_{ij}^3}{24}\left(\frac{\bar{\eta}^2}{\bar{V}^3}\right), \qquad (34)$$

where R_{ij} is an offset distance (Fig. 1), and $\bar{V}$ and $\bar{\eta}$ are the velocity and the velocity gradient averaged over the profile line, respectively. Adding the quantity of $\bar{\eta}^2/\bar{V}^3$ to our unknowns, $\boldsymbol{m}$, an optimum solution is obtained from Eq. (30) with slight modification of matrix A using Eq. (34). The validity of the approximation in Eq. (34) will be discussed in Subsection 3.4.

Table 1. Initial guesses and errors for model parameters.

	Initial guess	Uncertainty
Time-term	0.1 sec	0.2 sec
Velocity	6.0 km/s	1.0 km/s
Velocity Gradient	0.02 s^{-1}	0.02 s^{-1}

3. Numerical Experiment
3.1 Deficiencies in the conventional methods

We start with a model with a laterally uniform basement velocity (Model 1, Fig. 2(a)), where the uppermost layers have the same degree of complexity as observed in Japan (e.g. Iwasaki et al., 1994, 1998, 1999). For 5 shot points (S1–S5) and 114 receiver points on a 170-km profile, synthetic data are made from travel times of diving wave using a ray-tracing method (Iwasaki, 1989, see Fig. 2(b)). If a diving wave is not found for a certain pair of shot and receiver due to the complex layer geometry, we adopt a travel time of head wave after checking its accuracy. The errors of the data are taken to be Gaussian with zero mean, whose standard deviations are set 0.05, 0.10 and 0.15 s according to the offset ranges of 0–60, 60–120 and 120–170 km, respectively. Here, we adopt a relatively sparse receiver distribution (~1.5 km) and large noise in the data (0.05–0.15 s) to examine the effectiveness of our algorithm under rather poor experimental conditions. For example, the recent seismic experiments in Japan were undertaken with denser spacing (0.5–1 km) and less timing error (0.01–0.05 s). Initial guesses and their uncertainties used in our tests are listed in Table 1. As a measure of the time-term recovery, we define a quantity Δb, which is a root-mean-square (r.m.s.) of time-term deviation (the difference between obtained and assumed time-terms). We also present a r.m.s. of travel time residual (Δt_s) to indicate a degree of data fitness.

Figure 2(c) shows time-terms and velocities determined from conventional methods (a least squares method and a damped least squares method) with no smoothness constraints. Travel times calculated from the solutions are compared with synthetic data in Fig. 2(d). Here, we use the data at offsets of 5–100 km to exclude the effect of the velocity gradient. In the least squares solution, Eqs. (3) and (8) are simultaneously solved using an iterative method (the Gauss-Seidel method) proposed by Meru (1966). This procedure was widely used for seismic refraction studies in Japan (e.g. Yoshii and Asano, 1972). The damped least squares solution is obtained from Eq. (24) taking $\alpha_1 = 1.0$ and $\alpha_2 = 0.0$. Although both of the methods satisfactorily estimate the refractor velocity, the time-terms show unexpected oscillation arising from the data noise. Large travel time discrepancies for the least squares method ($\Delta t_s = 0.095$ s) come from the Gauss-Seidel method, in which the weight factor for the individual observation equation cannot be adjusted from the data quality. Actually, a large time-term discrepancy at a shot point of S3 yields the systematic travel time difference for this shot (Figs. 2(c) and (d)).

As stated in Subsection 2.4, the most serious problem in the damped least squares method is how to choose the appropriate damping parameters (hyperparameters). For sev-
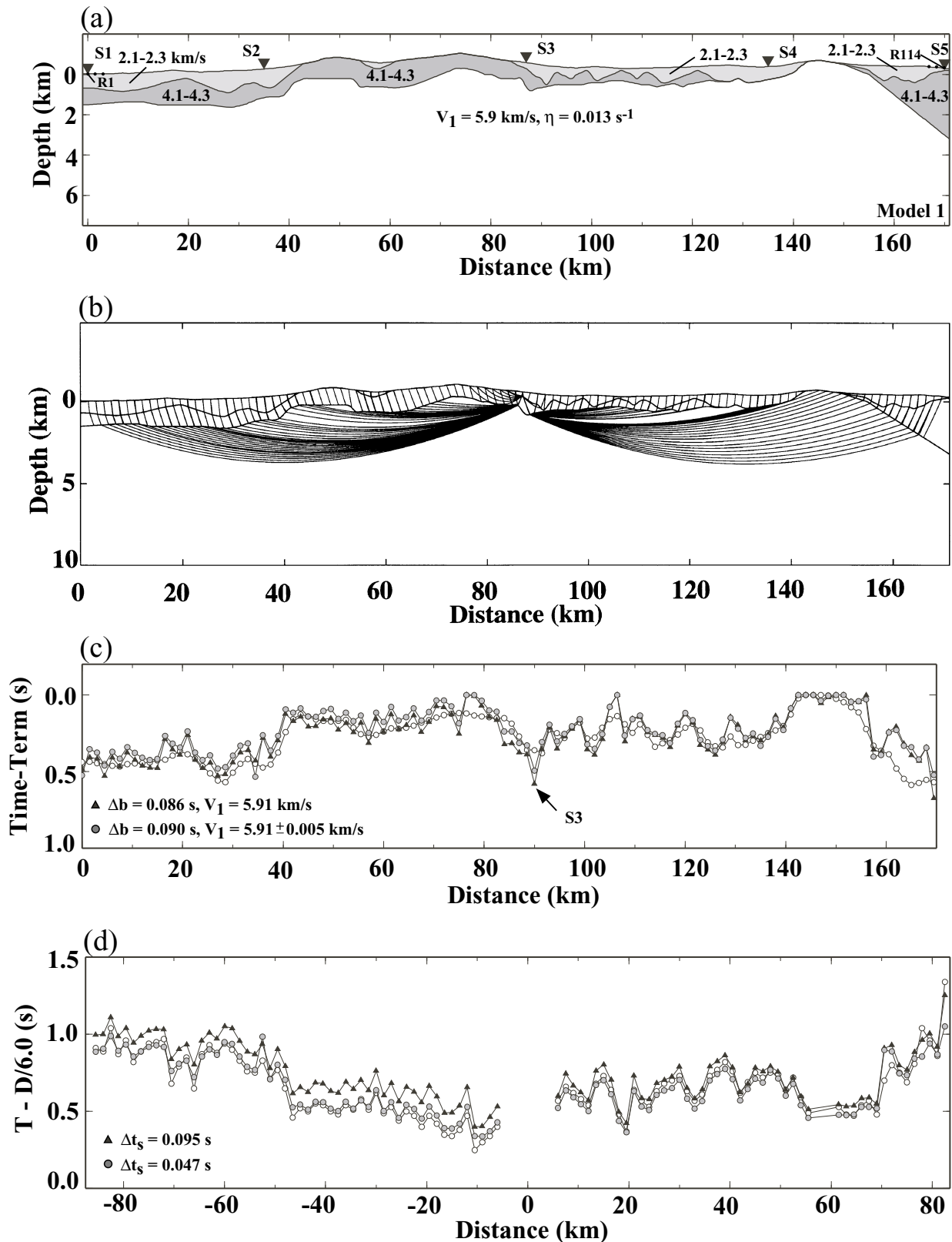
Fig. 2. Numerical test for a laterally uniform velocity model. The effect of velocity gradient is not taken into account in the inversion. (a) Structure model (Model 1). Surface layers are modelled from complex crustal structures in the Japanese Islands (e.g. Iwasaki *et al.*, 1994, 1998, 1999). The velocity at the top of the basement, V_1, and velocity gradient, η, are set 5.9 km/s and 0.013 s^{-1}, respectively. Shot and receiver points are also indicated. The receiver points with a spacing of 1.5 km are numbered from R1 to R114. (b) Ray diagrams for S3. Both of diving and head waves are shown. (c) Time-terms obtained by the least squares method (solid triangles) and the damped least squares method (gray circles), respectively. Open circles indicate exact time-terms. In either case, Δb, which is a root-mean-square (r.m.s.) of time-term deviation (the difference between obtained and exact time-terms), is shown. An arrow indicates time-terms of S3. Note unexpected oscillations of time-terms caused by the noisy data. (d) Travel times for S3 calculated from the solutions in Fig. 2(c). Solid triangles and gray circles correspond to the same cases as in Fig. 2(c). Open circles indicate synthetic data. A r.m.s. of travel time residual (Δt_s) is shown for the individual case. Note large travel time discrepancy in the case of the least squares method (solid triangles).
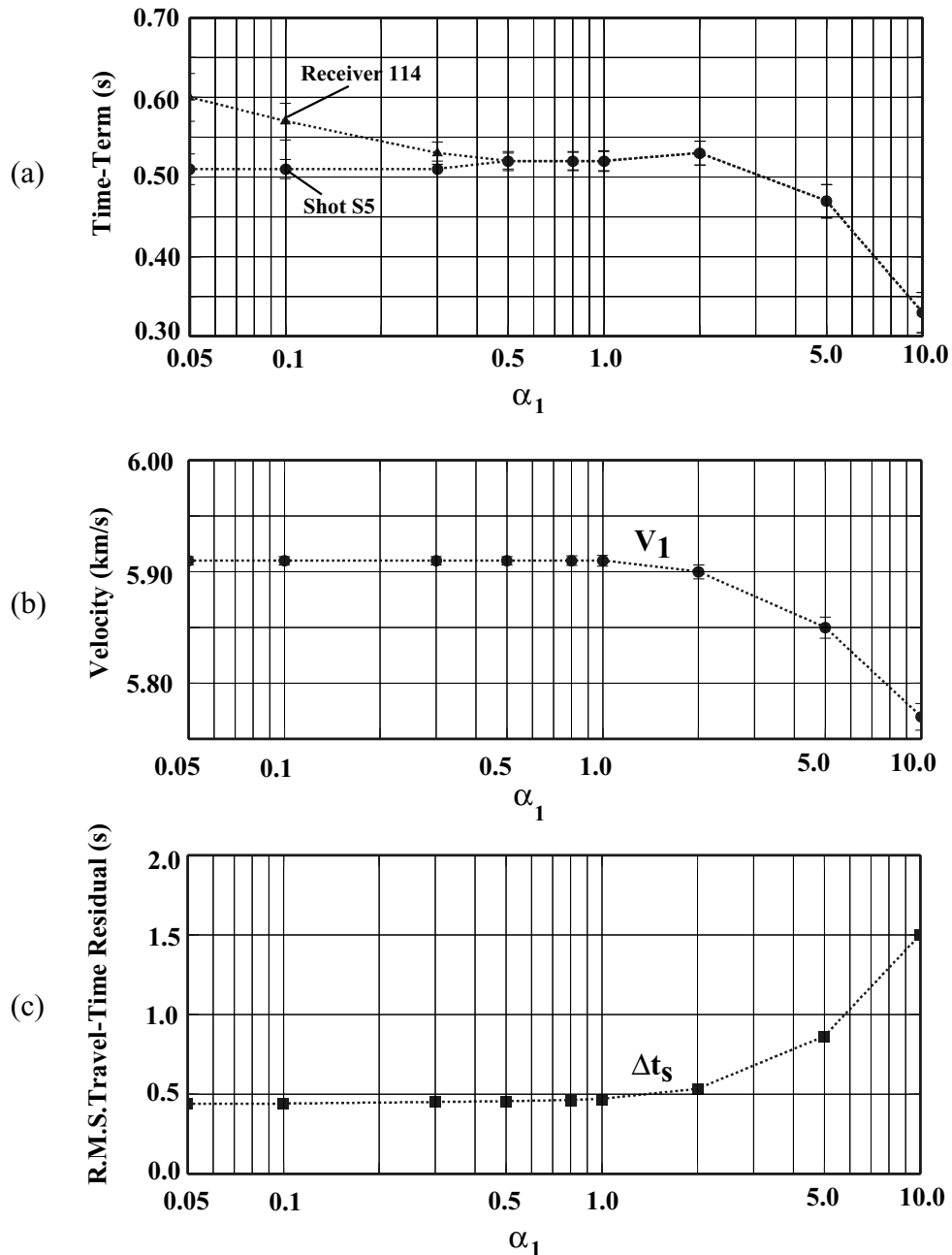
Fig. 3. Variations of time-terms, refractor velocity and r.m.s. travel time residual with respect to α_1 (damping parameter) in the damped least squares solutions for Model 1 ($V_1 = 5.9$ km/s, $\eta = 0.013$ s^{-1}). The effect of the velocity gradient is not taken into account. (a) Time-terms for a shot point S5 (solid circles) and a receiver point R114 (solid triangles), which is located in the vicinity of S5. Their time-terms are related by the constraint of Eq. (8). Note a large time-term change at $\alpha > 5.0$ because of the strong control from the initial guess (see also Figs. 3(b) and (c)). In a range of $\alpha > 0.3$ time-terms at S5 and R114 show significant discrepancy because the constraint from Eq. (8) is weakened. (b) Refractor velocity. (c) R.m.s. travel time residual (Δt_s).

eral values of α_1, we show a time-term at the shot point of S5, a refractor velocity and a r.m.s. of travel time residuals (Fig. 3). To check the constraint from Eq. (8), Fig. 3(a) also presents time-terms of a receiver of R114 taken in the vicinity of S5. Large travel time residuals for $\alpha_1 > 5.0$ (Fig. 3(c)) are due to the strong control from prior constraints (Figs. 3(a) and (b)). The large time-term difference between S5 and R114 for $\alpha_1 < 0.3$ indicates that the constraint of Eq. (8) is clearly broken.

3.2 Determination of hyperparameters

In order to suppress oscillation of time-terms as in Fig. 2(c), the smoothness constraint should be imposed in the inversion scheme according to Eq. (24) or (30). In a case with two hyperparameters, however, it is almost impossible to obtain their best estimates from a priori information. The present paper proposes the determination of hyperparameters from ABIC. Figure 4 shows contour maps of ABIC and r.m.s. of travel time residual (Δt_s) with respect to α_1 and α_2. The ABIC has a clear minimum at $\alpha_1, \alpha_2 = 1.0$ (Fig. 4(a)). The travel time residual is almost constant for $\alpha_1, \alpha_2 < 0.5$ (Fig. 4(b)) owing to the limitation of resolving power of the data. The time-term solution for the best parameter set
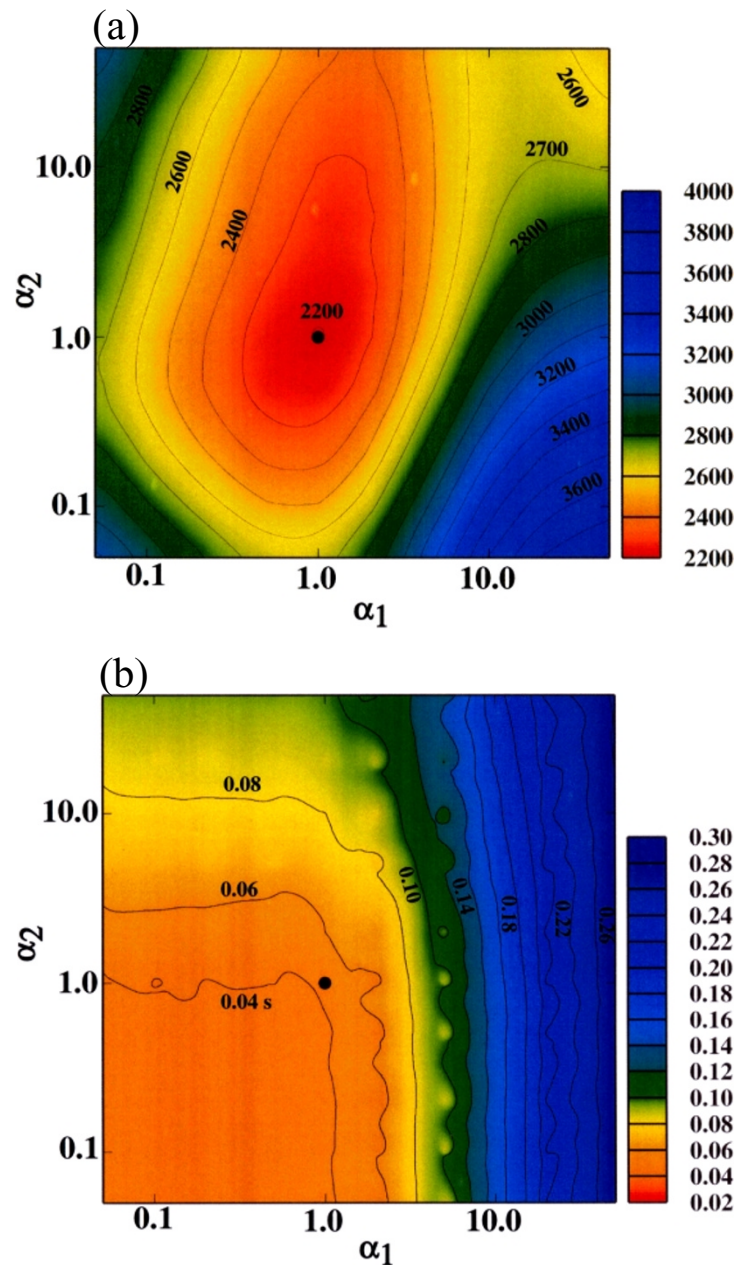
Fig. 4. Contour maps of ABIC and r.m.s. travel time residual for Model 1 (Fig. 2(a)) with respect to hyperparameters of α_1 and α_2. The effect of velocity gradient is not taken into account. (a) ABIC. The ABIC has a minimum at α_1, $\alpha_2 = 1.0$ whose location is shown by solid circle. (b) R.m.s. of travel time residual (Δt_s). The residual is almost constant at α_1, $\alpha_2 < 0.5$ due to the limitation of resolving power of data. A solid circle indicates a location (α_1, $\alpha_2 = 1.0$) which gives a minimum of ABIC (see Fig. 4(a)).

of α_1, $\alpha_2 = 1.0$ is given in Fig. 5(a). We see that the oscillation of the time-terms is effectively suppressed by the smoothness constraint, and the assumed structure is recovered well. Travel times computed from this solution explain most of the synthetic data (Fig. 5(b)). These results show that the travel time calculation in Eq. (1) works well. The average resolution of the time-terms is 0.88 for shot points and 0.66 for receiver points. On the other hand, the resolution of the refractor velocity is nearly 1.0, indicating that the velocity is almost perfectly resolved from the travel time data. The solution in Fig. 5(a) is almost independent of the initial guess. This situation, however, is broken for a more heterogeneous case as described in Section 4.

The time-terms obtained for two extreme cases of α_1, $\alpha_2 = 5.0$ and 0.05 are given in Fig. 6(a). The very smooth variation for α_1, $\alpha_2 = 5.0$ reflects the strong control from the prior information. Travel times computed from this solution show large discrepancies ($\Delta t_s = 0.105$ s) from the synthetic data (Fig. 6(b)). The time-terms for α_1, $\alpha_2 = 0.05$, on the other hand, have oscillatory feature quite similar to the case of the damped least squares method (Fig. 2(c)). The effectiveness of the ABIC criterion is clear from the comparison of these solutions to that in Fig. 5(a). Our best solution well recovers the predominant variation of the assumed time-terms without serious oscillation. In the following examples, solutions are obtained by minimizing ABIC. Their
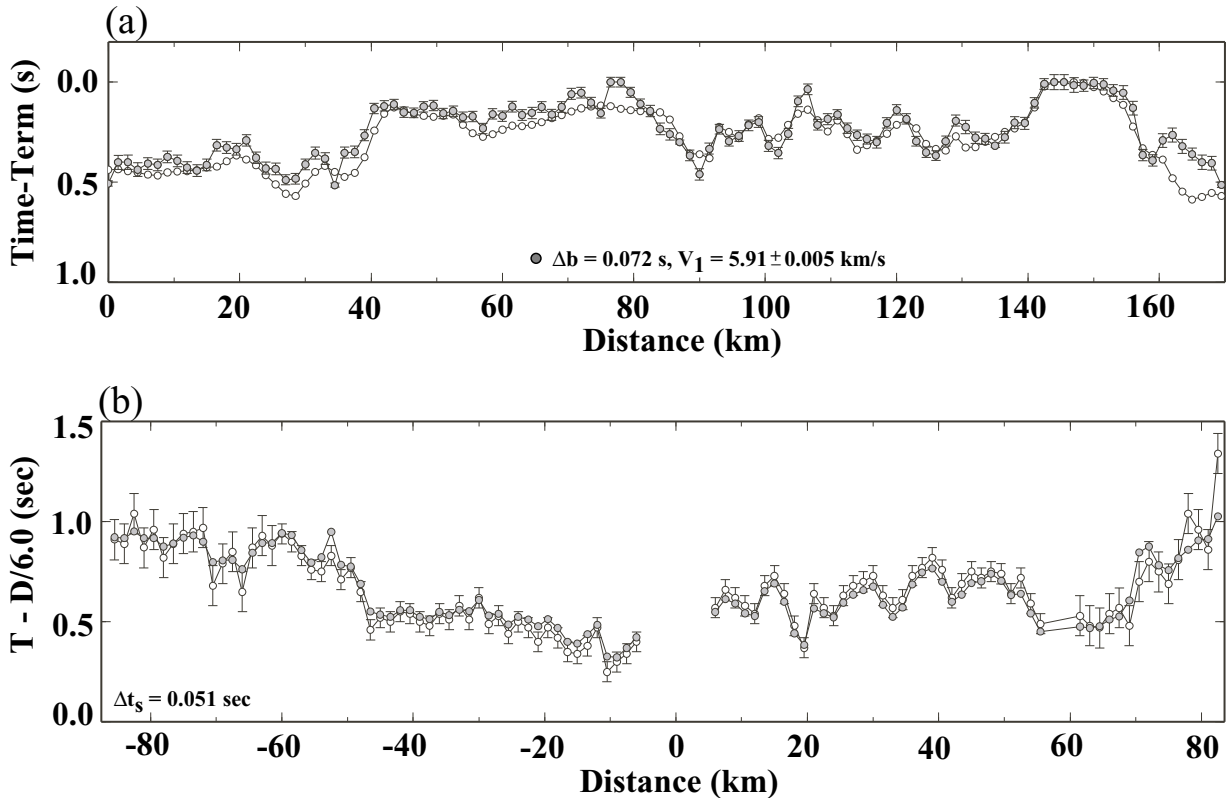
Fig. 5. Time-terms and travel-times for the best set of hyperparameters (α_1, $\alpha_2 = 1.0$) in the case of Model 1 (Fig. 2(a)). The effect of velocity gradient is not taken into account. (a) Time-terms (gray circles) with estimation errors. Open circles indicate exact time-terms. As compared with the case of Fig. 2(c), the oscillation of time-terms is effectively suppressed by the smoothness constraint. (b) Travel times calculated for S3 (gray circles) and synthetic data (open circles). Noises of the data are also shown as error bars.

behaviours with respect to α_1 and α_2 are quite similar to those in Figs. 4–6.

3.3   Lateral velocity variations in refractor

We proceed to tests for detecting lateral velocity variation in the refractor. Consider a case where the basement is divided into three geological units with a different refractor velocity (Model 2, Fig. 7(a)). A 4% velocity change in this model is comparable to those from the previous refraction experiments and seismic tomographies in Japan (e.g. Iwasaki *et al.*, 1999; Zhao *et al.*, 1992). We assume that the approximate locations of the units are known from other information such as geological observations, and fix them at 60 and 145 km from S1 in the inversion. Ray diagrams for this model are presented in Fig. 7(b). Figure 7(c) shows the solution obtained from the travel time data in an offset range of 5–100 km. In spite of the large undulation of the synthetic data, our algorithm well detects the refractor velocities of the individual geological unit. Their resolutions are nearly 1.0, indicating the velocities are almost perfectly resolved from the travel time data. The time-terms also satisfactorily coincide with the true values within an error of 0.1–0.15 s. The synthetic data are fitted well by this solution in the entire part of the profile (Fig. 7(d)). The solid triangles in Figs. 7(c) and 7(d) correspond to a case in which the same data are inverted assuming a 1-velocity segment (laterally uniform) model. It is noticed that the travel time residuals become large at offsets greater than 60 km. The time-terms, on the other hand, are almost the same as those for the three-segment case. The insensitiveness of time-terms to the basement structure and the high resolution of the refractor velocities indicate that the time-terms and refractor velocities are almost decoupled in this inversion. As discussed in Section 4, such decoupling is realized for relatively thin surface layers with a velocity 30–40% lower than the refractor velocity.

In order to examine the detectability of velocity variation, we modify the velocities of $V_1 \sim V_3$ in Model 2 (Fig. 7(a)) as follows;

$$\begin{cases} V_1 = 5.90 - 0.10\zeta, \\ V_2 = 5.90, \\ V_3 = 5.90 + 0.15\zeta \end{cases}. \qquad (35)$$

Here, a parameter, ζ, expresses a degree of velocity variation. It is noted that a case of $\zeta = 1$ corresponds to the velocity model for Figs. 7(b)–(d). Figure 8 shows solutions for various values of ζ. In a range of $\zeta > 0.5$, the solutions are satisfactorily obtained with an almost constant travel time residual ($\Delta t_s = 0.045$ s). This indicates that the travel time approximation in Eq. (1) works well even for a case of large lateral variation. For a small value of $\zeta < 0.5$ the travel time data lose their resolving power to identify the velocity variation, particularly for V_1 and V_2.

3.4   Effect of velocity gradient

Figure 9 shows variations of solution with respect to the assumed velocity gradients (η) in Model 1, where synthetic
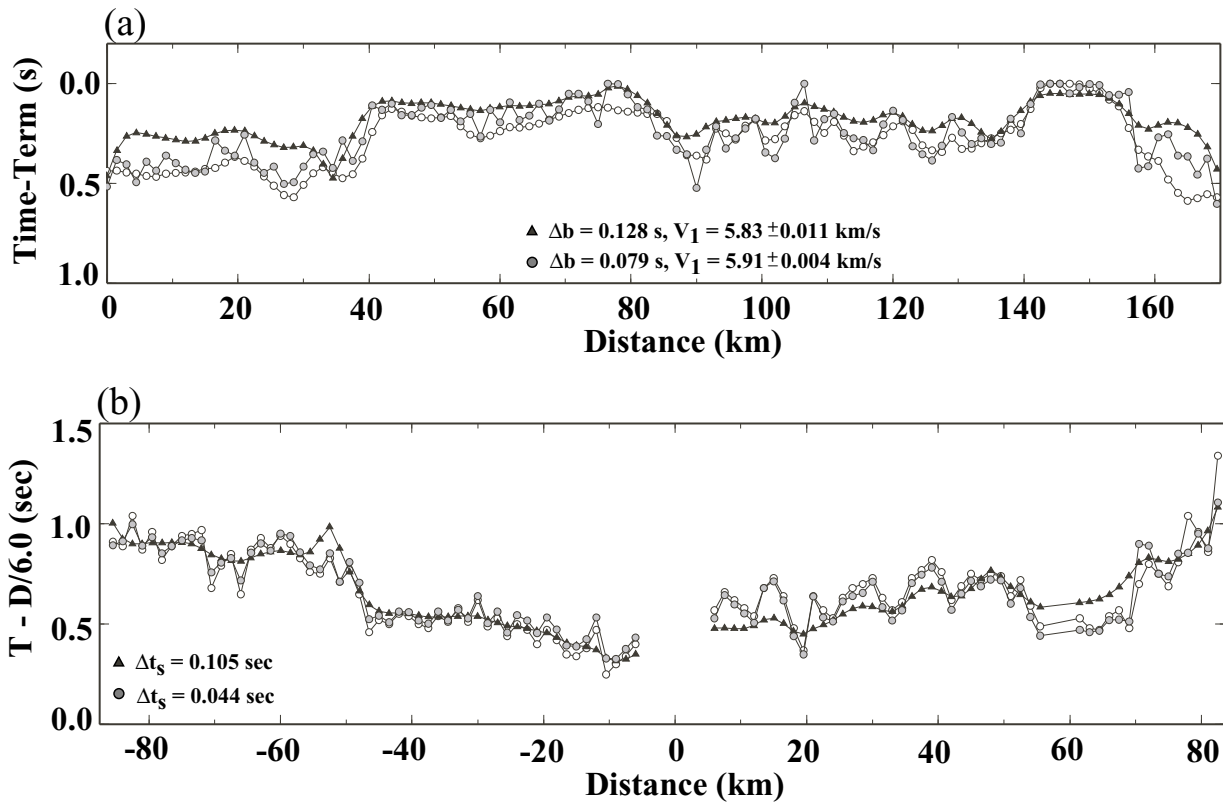
Fig. 6. Time-terms and travel-times for cases of large hyperparameters (α_1, $\alpha_2 = 5.0$) and small hyperparameters (α_1, $\alpha_2 = 0.05$) in Model 1. The effect of velocity gradient is not taken into account. (a) Time-terms. Solid triangles and gray circles correspond to cases of α_1, $\alpha_2 = 5.0$ and α_1, $\alpha_2 = 0.05$, respectively. Open circles indicate exact time-terms. The large time-term discrepancy ($\Delta b = 0.128$ s) in a case of α_1, $\alpha_2 = 5.0$ comes from the strong control by the prior constraints. The oscillatory features of time-terms for α_1, $\alpha_2 = 0.05$ is due to data noise. (b) Travel times calculated for S3. Open circles indicate synthetic data. Other symbols of gray circles and solid triangles correspond to the same cases as in Fig. 6(a).

data at offsets greater than 5 km are inverted using the formulations in Subsection 2.5. It is seen that model parameters are well recovered in a gradient range of 0.02~0.04 s⁻¹. In a case with a small gradient (<0.015 s⁻¹), we notice a trade-off between the estimated velocity and velocity gradient (Fig. 9(a)). Actually, the introduction of velocity gradient and the slight increase in refractor velocity give almost the same effect to travel times, and are not distinguishable to each other in the inversion process. For a large value of gradient, the approximation of Eq. (33) or (34) yields systematically larger travel times than exact values, leading the underestimation of gradient as in Fig. 9(a). If we neglect the effect of velocity gradient, the solutions show significant deviation from the true values, yielding large travel time residuals (solid squares in Fig. 9).

A model in Fig. 10(a) is the most general case with laterally varying velocity and velocity gradient. The velocities obtained for the individual segments are in good agreement with the assumed values (Fig. 10(b)). The estimated velocity gradient of 0.025 s⁻¹ is also consistent with an averaged value over the assumed model (0.023 s⁻¹), showing the validity of the approximation in Eq. (34). Actually, the synthetic data are well explained by this solution (Fig. 10(c)). The inversion without the velocity gradient effect, on the other hand, yields the velocities 0.08–0.09 km/s higher than the true values, and significant travel time discrepancies at distances greater than 120 km (solid triangles in Fig. 10(c)).

4. Discussion

The applicability of our method for a real structure depends on the validity of simple approximation given by Eq. (1). This approximation is broken in the following situations.

(1) Lateral velocity variation.
(2) Velocity gradient and its lateral variation.
(3) High velocity and thick surface layers.

The first two factors were examined in Subsections 3.3 and 3.4. Here, we discuss the last situation. The time-term method basically assumes a stratified structure composed of surface layers and a basement. If the velocities of surface layers are relatively smaller (>30~40%) than the refractor velocities, the time-terms and refractor velocities are almost decoupled in the inversion process. Then, a rather high degree of lateral velocity change and/or a large velocity gradient within the refractor are satisfactorily recovered by our method (Figs. 7–10). Such decoupling, however, is not guaranteed for a structure with thick and high-velocity surface layers. In such a case, the assumption of layered structure is broken because of a smaller velocity difference between the surface layer and the basement. The increase in the total thickness and velocity of the surface layers brings another serious problem. In the time-term analysis, we usually use refracted waves beyond the crossover distance because they are observed as a first arrival. Therefore, the thick and high velocity surface layers cause the decrease of available travel
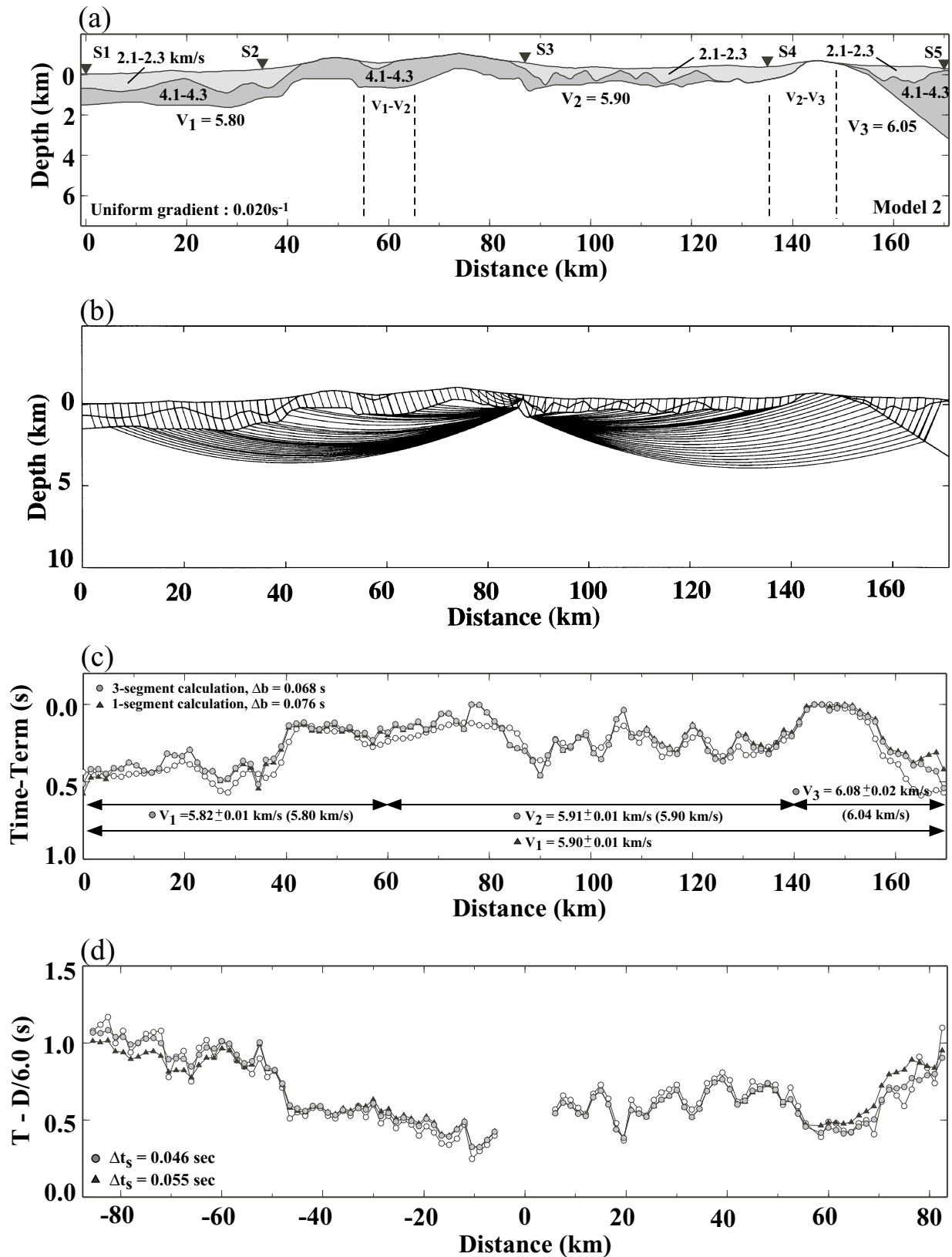
Fig. 7. Numerical test for a case with a laterally varying refractor velocity. The distribution of shots and receivers are the same as in Fig. 2(a). The effect of velocity gradient is not taken into account. (a) A structure model (Model 2) with three velocity segments. The velocity gradient is taken to be uniform (0.020 s^{-1}) (b) Ray diagrams for S3. (c) Time-terms obtained assuming three velocity segments for Model 2 (gray circles). Open circles indicate exact time-terms. A velocity calculated from the assumed model is given in parenthesis. The time-terms and velocities are satisfactorily recovered. The results from 1-velocity segment case (solid triangles) are also presented. (d) Travel times for S3 calculated from the solution in Fig. 7(c). Gray circles and solid triangles correspond to the same cases as in Fig. 7(c). Open circles indicate synthetic data.
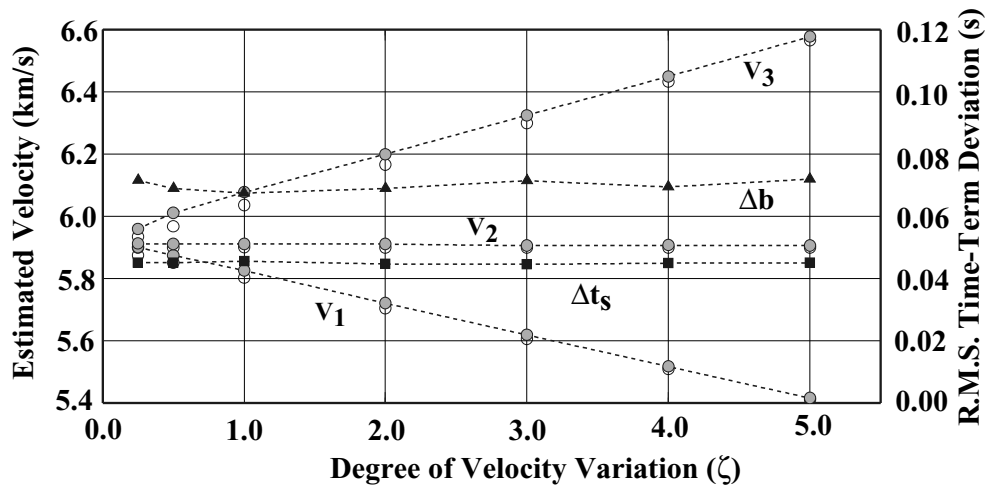
Fig. 8. Detectability for lateral velocity variation for Model 2 (Fig. 7(a)). The horizontal axis, ζ, represents a degree of lateral velocity (see Eq. (35) in text). The effect of velocity gradient is not taken into account. Open and gray circles indicate assumed and reconstructed velocities, respectively. The velocities are satisfactorily recovered except for small ζ (<0.5) where V_1 and V_2 are not resolved well. A r.m.s. of time-time deviation (Δb) for the individual calculation is shown by solid triangles.
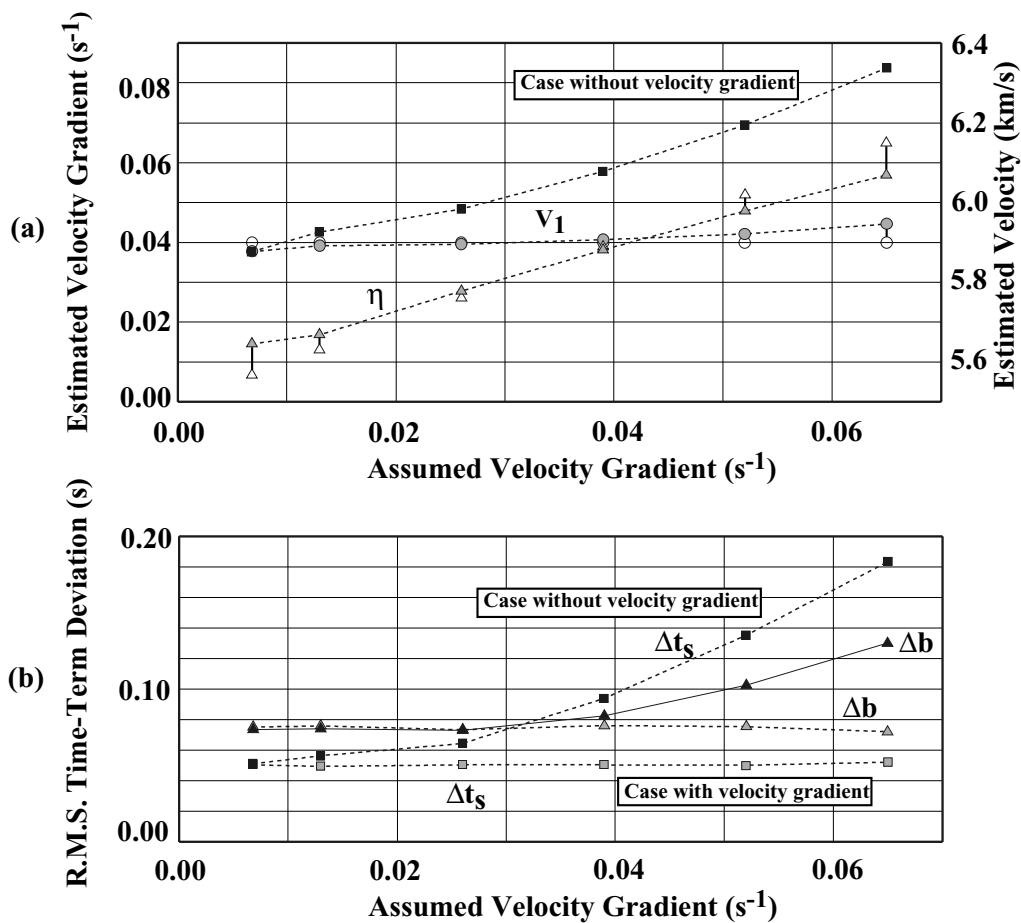

Fig. 9. Detectability for velocity gradient in Model 1. Recovery of model parameters is tested assuming various values of velocity gradient. (a) Recovery of velocity and velocity gradient. Open and gray symbols denote assumed and reconstructed values, respectively. Our method satisfactorily estimates the velocity and velocity gradient in a range of 0.02–0.05 s^{-1}. For a small gradient (<0.02 s^{-1}), on the other hand, they show slight shift from the true values. For comparison, velocities obtained without the effect of gradient are also presented by solid squares. (b) R.m.s. of time-term deviation (Δb). The time-terms are successfully obtained in the almost entire range of 0.006–0.065 s^{-1}. R.m.s. travel time residuals (Δt_s) are also shown.
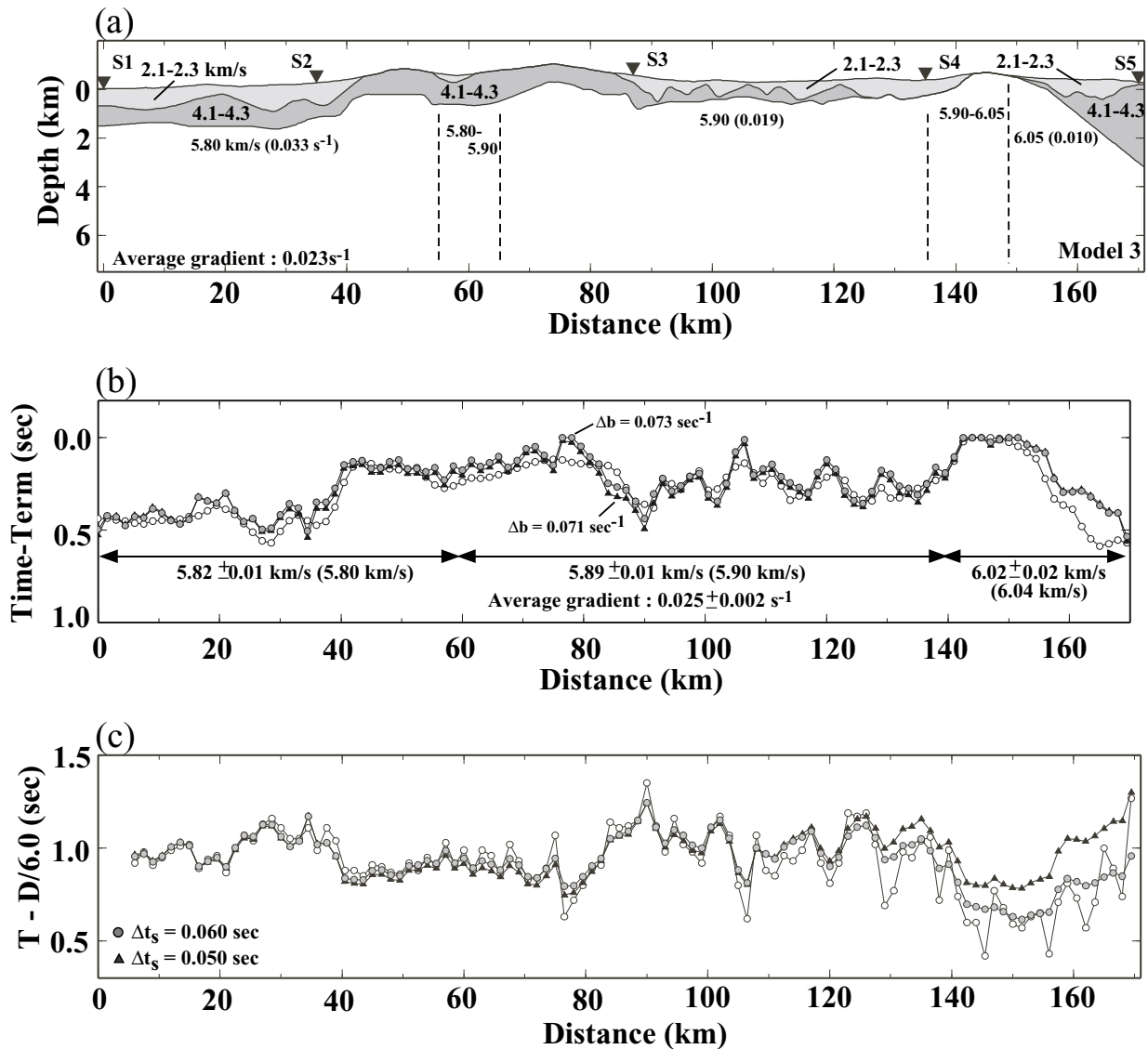
Fig. 10. Numerical test for a case with laterally varying velocity and velocity gradient (Model 3). The distribution of shots and receivers are the same as in Fig. 2(a). (a) A structure model (Model 3) composed of three segments with a different velocity and velocity gradient. A value in parenthesis is a velocity calculated from the assumed model. An average gradient of the model is 0.023 s^{-1}. (b) Reconstructed time-terms, velocities and velocity gradient. A velocity assumed for the individual segment is given in parenthesis. Gray circles show time-terms obtained for a case where the effect of velocity gradient is taken into account. Open circles indicate exact time-terms. Model parameters are satisfactorily reconstructed. Solid triangles correspond to a case without the effect of velocity gradient. (c) Travel times calculated for S1. Open circles indicate synthetic data. Other symbols are the same as in Fig. 10(b).

time data near a shot point, which directly leads the lowering of the resolving power.

We investigate this difficulty by shifting the position of basement in Model 2. Figure 11(a) shows the variation of r.m.s. of time-term deviation (Δb) and refractor velocities for the downward shift of the basement (d_s). With increase of d_s, the solution shows larger deviation from the assumed model. In Fig. 11(b), the time-term deviation from the assumed value is shown for $d_s = 0$, 2 and 4 km. For d_s greater than 3 km, an area with large deviation extends into a wider range of 150–170 km, and the velocity of this range, V_3, shows a remarkable decrease (Fig. 11(a)). This is caused by the lowering of the resolving power of the data associated with large crossover distances (20–25 km) for shots S4 and S5. Figure 11(c) shows correlation coefficients of velocity

V_3 and time-terms, b_j that is,

$$C_j = C(b_j, V_3)/\sqrt{C(b_j, b_j)C(V_3, V_3)}, \qquad (36)$$

where $C(b_j, V_3)$ is a covariance between V_3 and b_j, and $C(V_3, V_3)$ and $C(b_j, b_j)$ are variances of V_3 and b_j, respectively. Large positive correlation of V_3 and b_j in the range of 150–170 km indicates strong coupling between these two parameters. Namely, effects by decreases in V_3 and b_j are cancel out to produce almost same the travel times. The resultant time-terms are very small, strongly dependent on the initial values of 0.1 s (Table 1). If we adopt a large initial value, for example, 1 s, the obtained time-term is larger than the true values while the velocity is considerably high (6.3–6.4 km/s). In a case with higher surface velocity of 5.0–5.6 km/s, such discrepancies appear in a wider area of 140–170
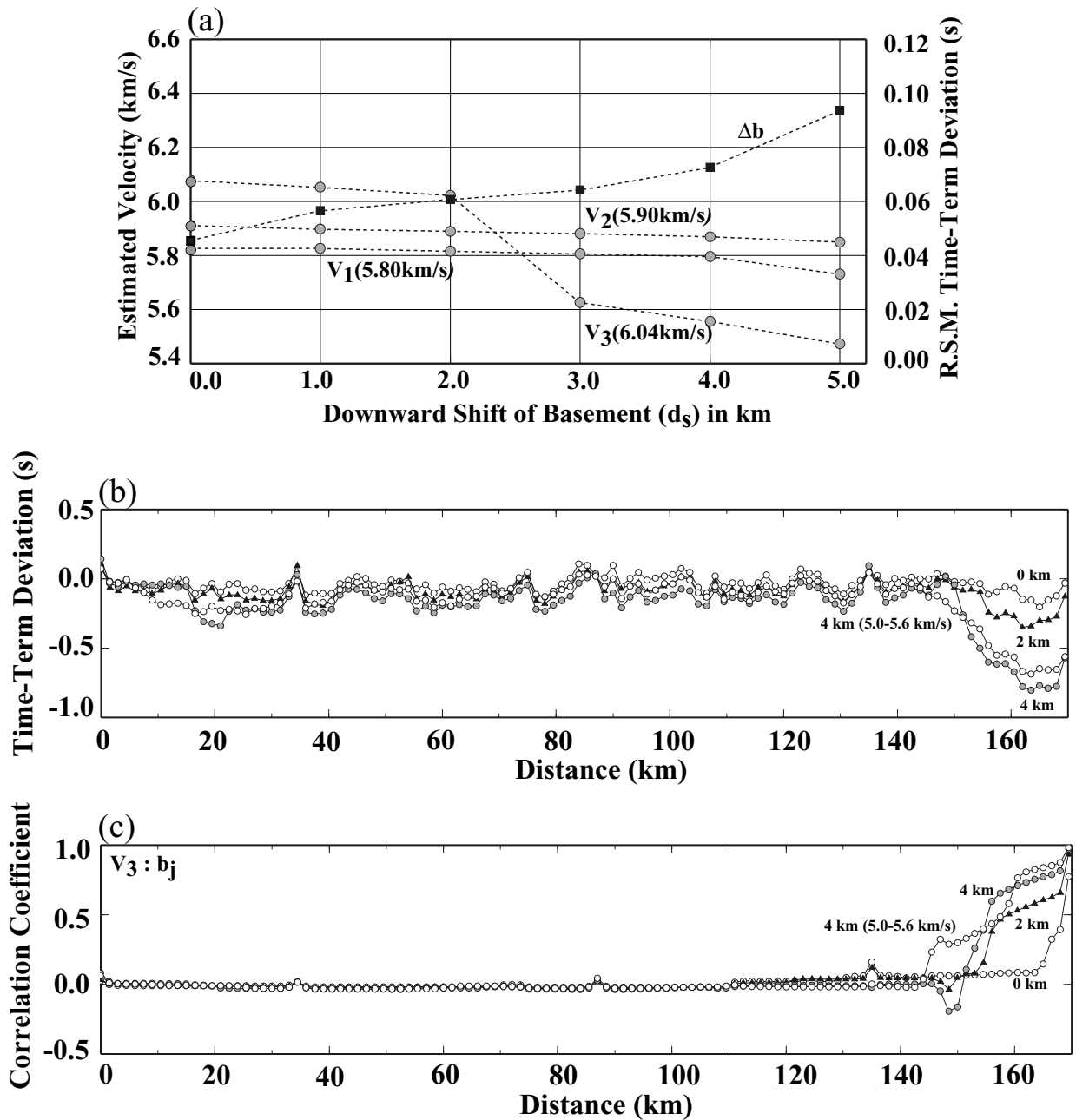
Fig. 11. Numerical tests for coupling between time-terms and velocities. For downward shift of the basement (d_s) in Model 2 (Fig. 7(a)), variations of the solution are presented. The effect of velocity gradient is not taken into account. (a) Refractor velocities (V_1–V_3) and r.m.s. of time-term deviation (Δb). A velocity calculated from the assumed model is given in parenthesis. Note a drastic decrease of V_3 (in a distance range of 150–170 km) for d_s greater than 3 km. (b) Time-term deviation (difference between obtained and assumed time-terms) for $d_s = 0$, 2 and 4 km. Note large discrepancy appearing in a range of 150–170 km which corresponds to the segment of V_3. The discrepancy extends to wider region if the velocity of the surface layer is taken to be large (a case where the basement is shifted 4 km and the velocity in the surface layer is increased to 5.0–5.6 km). (c) Correlation coefficients between the velocity, V_3 and time-term, b_j (see Eq. (36) in text). Large correlation in a range of 150–170 km indicates strong coupling of V_3 and b_j.

km (see solid triangles in Figs. 11(b) and (c)) because the crossover distance attains to nearly 30 km.

Figure 12 is a final example with a more heterogeneous structure composed of high velocity gradient part (surface layer) and low velocity gradient part (basement). The velocity change at their boundary is very small ($<0.05\sim0.1$ km/s). The lateral variation in the lower part is 8%, which is about two times larger than those in Models 2 and 3 (Figs. 7(a) and 10(a)). Ray diagrams in Fig. 12(b) indicate the diving waves penetrate a nearly 20-km depth. Fig-

ure 12(c) shows time-terms of the upper part obtained assuming four velocity segments. Large discrepancies of time-terms and refractor velocity in a range of 150–170 km are caused by the similar reason as shown in Fig. 11. In the remaining part of the profile, however, the assumed structure is well recovered, indicating that our method is still effective in spite of the large heterogeneity (Figs. 12(c) and (d)). These figures also show that the conventional method without the effect of velocity gradient is broken out in the entire part of the profile.
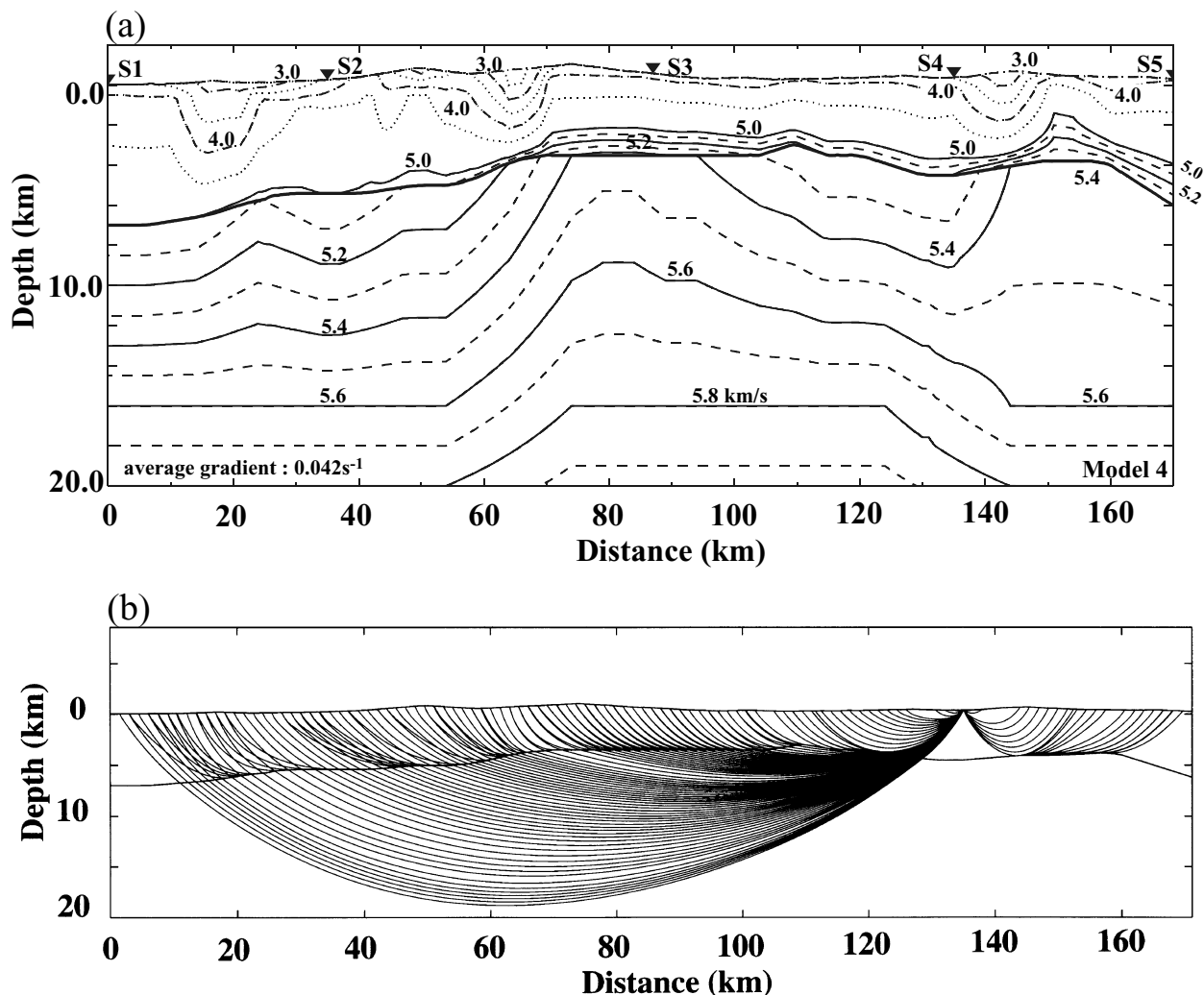
Fig. 12. Numerical test for a highly heterogeneous model. (a) A structure model (Model 4). The distribution of shots and receivers are the same as in Fig. 3(a). This model is composed of two parts with high and low velocity gradients. (b) Ray diagrams for S4. (c) Time-terms (gray circles) and velocities reconstructed with the effect of velocity gradient. Open circles indicate exact time-terms. Model parameters are well recovered in a region of 0–140 km. Velocities averaged from the assumed model are given in parenthesis. Large time-term discrepancy at distances of 150–170 km is due to the same reason (strong coupling between the velocity and time-terms) as in the case of Fig. 11. Time-terms obtained without the effect of velocity gradient (solid triangles) show large deviation from the assumed values in the entire profile. (d) Travel times calculated for S1 (gray circles) and synthetic data (open circles). The synthetic data are well fitted in spite of the failure of our method in a range of 150–170 km.

5. Conclusions

A time-term method is a simple travel time inversion developed for seismic refraction studies. In the present paper, we modified this method based on a recently developed inversion algorithm with the use of ABIC, paying special attention for identifying lateral variations in refractor velocity. The effect of vertical velocity gradient within the refractor was also incorporated to realize accurate velocity determination.

The validity and effectiveness of our algorithm are examined through several numerical tests under realistic experimental conditions. We constructed synthetic data for the tests from crustal models with the same degree of complexity as observed in Japan. So far, the time-term method was believed to be applicable to a rather simple structure because of its crude travel time approximation. Our experiments show that this approximation works well even for a

case with undulated layers overlying a basement. It is particularly effective if the surface layers are not thick and their velocities are less than 60–70% of the basement velocity. Then, the time-terms and the basement velocities are almost decoupled, and fine structures within the basement, namely the lateral velocity variation and the velocity gradient, can be satisfactorily recovered.

The effect of vertical velocity gradient was approximately introduced into our formulation. This modification was also effective in a case where travel times at rather distant offsets are included in the analysis. Actually, the velocity gradient was well determined by our algorithm, and resultant travel time residuals were significantly reduced in the entire offset range as compared with a case of the conventional method.

The above results show the usefulness of our time-term method, which will be an effective tool of the crustal study in a geologically complex region like an island arc.
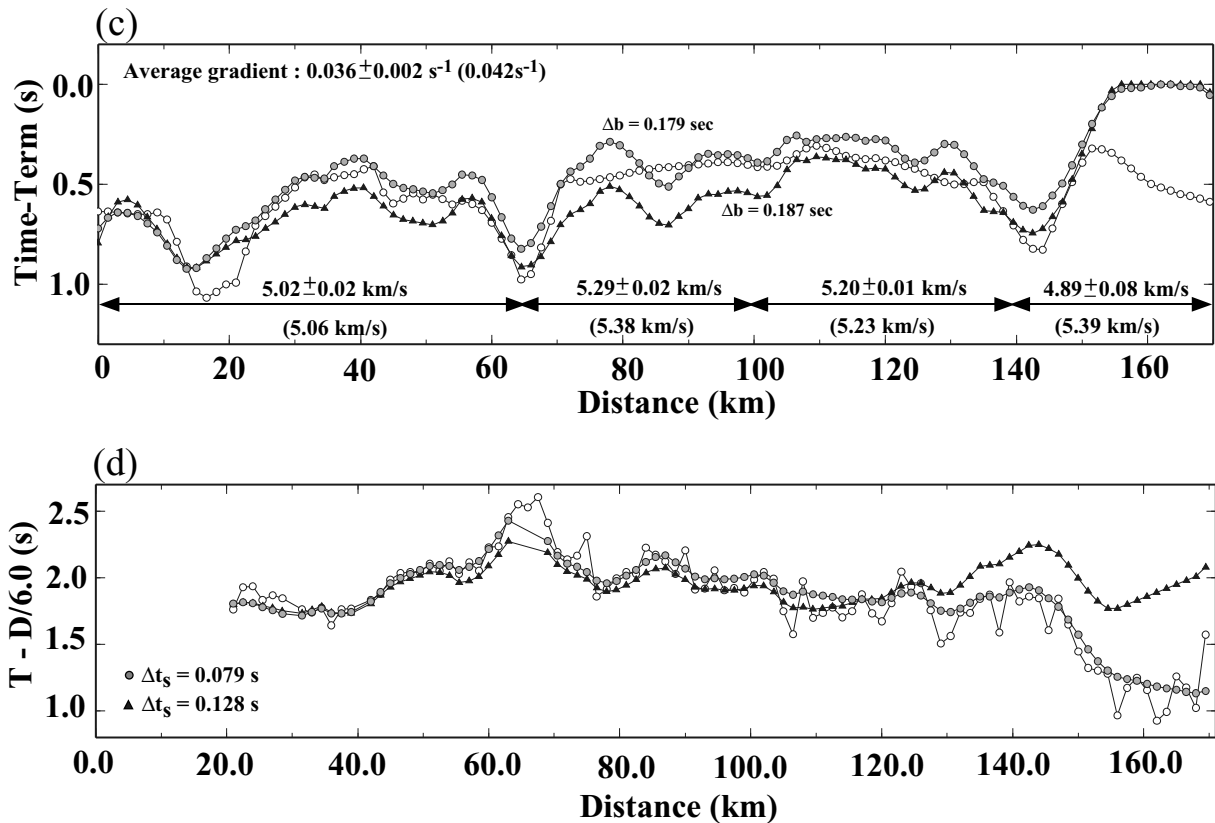
Fig. 12. (continued).

References

Akaike, H., Likelihood and the Bayes procedure, in *Bayesian Statistics*, edited by J. M. Bernard, M. H. Degroot, D. V. Lindley, and A. F. M. Smith, pp. 143–166, University Press, Valencia, Spain, 1980.

Berry, M. J. and G. F. West, An interpretation of the first arrival data of the Lake Superior experiment by the time-term method, *Bull. Seism. Soc. Am.*, **56**, 141–171, 1966.

Červený, V., I. A. Moltkov, and I. Psencik, *Ray Method in Seismology*, 214 pp., Univerzita, Karlova, Prague, 1977.

Hearn, T. M. and R. W. Clayton, Lateral velocity variations in southern California. I., Results for the upper crust from Pg waves, *Bull. Seism. Soc. Am.*, **76**, 495–509, 1986.

Iwasaki, T., Ray-tracing program for study of velocity structure by ocean bottom seismographic profiling, *J. Seismol. Soc. Jpn.*, **41**, 263–266, 1989 (in Japanese).

Iwasaki, T., T. Yoshii, T. Moriya, A. Kobayashi, M. Nishiwaki, T. Tsutsui, A. Ikami, and T. Masuda, Precise P and S wave velocity structure in the Kitakami massif, northern Honshu, Japan, from a seismic refraction experiment, *J. Geophys. Res.*, **99**, 22187–22204, 1994.

Iwasaki, T., O. Ozel, T. Moriya, S. Sakai, S. Suzuki, G. Aoki, T. Maeda, and T. Iidaka, Lateral structural variation across a collision zone in central Hokkaido, Japan, as revealed by wide-angle seismic experiment, *Geophys. J. Int.*, **132**, 435–457, 1998.

Iwasaki, T., W. Kato, T. Umino, T. Okada, S. Sekine, T. Takeda, A. Hasemi, T. Matsushima, T, Mizogami, K. Miyashita, H. Miyamachi, and T. Moriya, Crustal structure across Northern Honshu arc revealed from wide-angle seismic reflection survey, *Chikyu Monthly*, **27**, 48–55, 1999 (in Japanese).

Koketsu, K. and S. Higashi, Three-dimensional topography of the sediment/basement interface in the Tokyo metropolitan area, central Japan, *Bull. Seism. Soc. Am.*, **82**, 2328–2349, 1992.

Meru, R. F., An iterative method for solving the time-term equations, J. S. Steinhart and T. J. Smith, Ed., *Geophys. Monogr.*, **10**, pp. 495–497, Am. Geophys. Union, 1966.

Pavlis, G. L., Geotomography using refraction fan shots, *J. Geophys. Res.*, **91**, 6522–6543, 1986.

Scheidegger, A. and P. L. Willmore, The use of a least square method for the interpretation of data from seismic surveys, *Geophysics*, **22**, 9–22, 1957.

Smith, T. J., J. S. Steinhart, and L. J. Aldrich, Lake Superior crustal structure, *J. Geophys. Res.*, **71**, 1141–1172, 1966.

Willmore, P. L. and A. M. Bancroft, The time-term method approach to refraction seismology, *Geophys. J.*, **3**, 419–432, 1960.

Yabuki, T. and M. Matsu'ura, Geodetic data inversion using a Bayesian Information Criterion for spatial distribution of fault slip, *Geophys. J. Int.*, **109**, 363–375, 1992.

Yoshii, T., Crustal structure of the Japanese Islands revealed by explosion seismic observations, *Zisin*, **46**, 479–491, 1994 (in Japanese with English abstract).

Yoshii, T. and S. Asano, Time-term analysis of explosion seismic data, *J. Phys. Earth*, **20**, 47–57, 1972.

Zhao, D., A. Hasegawa, and S. Horiuchi, Tomographic imaging of P and S wave velocity structure beneath Northern Japan, *J. Geophys. Res.*, **97**, 19909–19928, 1992.

Zhao, L. S., Lateral variations and azimuthal anisotropy of Pn beneath the Basin and Range province, *J. Geophys. Res.*, **98**, 22109–22122, 1993.

T. Iwasaki (e-mail: iwasaki@eri.u-tokyo.ac.jp)