Geophys. J. Int. (2004) **158,** 169–178 doi: 10.1111/j.1365-246X.2004.02287.x

On interpolation functions in traveltime tomography

Masatoshi Miyazawa[1,*] and Mamoru Kato[2]

[1]*Department of Geophysics, Graduate School of Science, Kyoto University, Kyoto, Japan. E-mail: linen@rcep.dpri.kyoto-u.ac.jp*
[2]*Earth Science Group, Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan. E-mail: mkato@gaia.h.kyoto-u.ac.jp*

Accepted 2004 February 16. Received 2004 February 16; in original form 2002 July 3

SUMMARY

Grid representation of 3-D media and linear interpolation functions have been widely used in recent seismic traveltime tomographic studies. Either velocity or slowness is used as a model parameter in such studies, which are by definition mutually reciprocal. Field representations in grid-based models are, however, mutually inconsistent when velocity and slowness respectively are used, as widely used interpolation functions do not hold the reciprocal relationship between velocity and slowness fields. As a consequence these two field representations lead to two tomographic results from one set of data, and the difference is systematic and cannot be disregarded. We propose a criterion for evaluating this systematic difference in the tomographic results. We first derive a general analytical expression for interpolation functions for traveltime tomographic studies. This interpolation scheme uses norm estimation in distance measurement, and equates grid representation with block representation in the limit. The systematic difference in the two representations could be avoided by carefully formulating the tomographic problem.

Key words: body waves, ray tracing, seismic ray theory, tomography, traveltime.

1 INTRODUCTION

3-D heterogeneity in the interior of the Earth is often represented by a finite number of grid nodes or blocks. In grid models, geophysical quantities such as seismic velocity are specified at grid nodes to represent the model. Values at points other than the grid nodes are derived from those at adjacent grid nodes through interpolation. In block models, the entire model space is filled with cellular blocks, for each of which a geophysical quantity is specified. Artificial first-order discontinuities occur at block boundaries in block representations, while the model varies continuously in grid representations.

Velocity v and slowness s are by definition mutually reciprocal, and both are used in seismic traveltime tomography. When we discuss resultant tomographic models, velocity is usually the parameter of choice. Slowness is often the parameter of choice in computation; traveltime is expressed by the path integral of slowness along the ray path, and slowness perturbation Δs and traveltime residuals have a linear relation to first order. The velocity perturbation Δv is related to the slowness perturbation $\Delta s \simeq -\Delta v/v^2$ when $\Delta v \ll v$. Δv can be directly inverted for from traveltime residuals (Spakman 1991). In a popular program package based on Thurber (1983), slowness is used as the model parameter, and a combination of a grid representation and a linear interpolation function is used to represent the velocity media.

In grid-based traveltime tomographic models, reciprocity between velocity and slowness does not hold in the model space because the linear interpolation functions that are commonly used in such studies do not retain a reciprocal relation. To illustrate this, let us consider two cases of interpolation: one with velocities at grid nodes being specified as the model parameter (V-grid), and the other with slownesses at grid nodes being specified as the model parameter (S-grid). For simplicity, we consider a 1-D case with two grid nodes x_i ($i = 1, 2$). Velocity v_i and slowness s_i at x_i satisfy reciprocity, $v_i s_i = 1$, by definition. At an arbitrary point x ($x_1 < x < x_2$), the commonly used linear interpolation function (e.g. Thurber 1983; Zhao *et al.* 1992) gives

$$v(x) = v_1 \left| \frac{x - x_2}{x_1 - x_2} \right| + v_2 \left| \frac{x - x_1}{x_1 - x_2} \right|. \tag{1}$$

$s(x)$ is similarly given, and $v(x)s(x) \neq 1$ unless $v_1 = v_2$ (or, simply, let us remind ourselves that the mean of α and β and the reciprocal of the mean of two reciprocals, $1/\alpha$ and $1/\beta$, do not agree with each other unless $\alpha = \beta$). We hereafter call the occurrence of $vs \neq 1$ in the field that is due to interpolation 'reciprocal failure'.

When reciprocal failure occurs, ray paths and traveltimes with the V-grid and with S-grid do not agree with each other even when the same grid specification is used. This leads to the tomographic results depending on whichever parameter, velocity or slowness, is solved

*Now at: Research Center for Earthquake Prediction, Disaster Prevention Research Institute, Kyoto University, Japan.

for in the computation. In this paper we evaluate how reciprocal failure propagates in traveltime tomography by examining properties of interpolation functions. We present and discuss a general analytical expression for interpolation functions for traveltime tomography, which uses norm estimation to measure distances in model spaces. Our numerical examples demonstrate that reciprocal failure induces a difference in the tomographic results between V-grid and S-grid models, which is systematic and could be significant.

2 NORM ESTIMATION AND INTERPOLATION

Let us consider a general interpolation function $f(x)$ in an interval between two grid nodes $[x_1, x_2]$. Hereafter we assume with no loss of generality that $x_1 < x_2$ and $f(x_1) > f(x_2)$. We assume that the arithmetic mean applies for $f(x)$ at the mid-point, i.e.

$$f(x_\mathrm{m}) = \frac{f(x_1) + f(x_2)}{2} \quad \text{for} \quad x_\mathrm{m} = \frac{x_1 + x_2}{2}. \tag{2}$$

An interpolation function in $[x_1, x_\mathrm{m}]$ is represented as a line or a curve connecting $(x_1, f(x_1))$ and $(x_\mathrm{m}, f(x_\mathrm{m}))$, where $f(x)$ is unique for x. Geometrically, any point X on a curve which connects two points A and B (with $\angle\mathrm{AOB} = \pi/2$) is specified by an angle θ ($= \angle\mathrm{AOX}$) and a distance r from the origin O (Fig. 1) The distance r is defined by the length of the pth norm, $\|\mathrm{OX}\|_p = r$, where

$$\|x\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}, \quad p \geq 1. \tag{3}$$

For simplicity, we assume $\partial \|\mathrm{OX}\|_p / \partial\theta = 0$. The trajectory of X forms an arc of a generalized circle (or ellipse) connecting A and B. This arc satisfies $\|\mathrm{OX}'\|_p = 1$, where OX′ indicates proper normalization of OX by |OA| and |OB|. By placing $(x_1, f(x_1))$, $(x_\mathrm{m}, f(x_\mathrm{m}))$ and $(x_1, f(x_\mathrm{m}))$ at A, B and O, respectively, for $[x_1, x_\mathrm{m}]$, and similarly $(x_2, f(x_2))$, $(x_\mathrm{m}, f(x_\mathrm{m}))$ and $(x_2, f(x_\mathrm{m}))$ at A, B and O, respectively, for $[x_\mathrm{m}, x_2]$ (Figs 1 and 2), we derive equations for the trajectory of X on $[x_1, x_2]$,

$$\begin{aligned} 1 = \|\mathrm{OX}'\|_p &= \left\| \left(\frac{|f(x_\mathrm{m}) - f(x)|}{|(f(x_1) - f(x_2))/2|}, \frac{|x_i - x|}{|(x_1 - x_2)/2|} \right) \right\|_p \\ &= \left[\left(\frac{2|f(x_\mathrm{m}) - f(x)|}{|f(x_2) - f(x_1)|} \right)^p + \left(\frac{2|x_i - x|}{|x_2 - x_1|} \right)^p \right]^{1/p} \end{aligned} \tag{4}$$

for $(x_\mathrm{m} - x)(x_i - x) \leq 0$ for $i = 1, 2$.

Using trigonometry with θ, such as

$$\frac{2|f(x_\mathrm{m}) - f(x)|}{|f(x_2) - f(x_1)|} = \cos^{2/p}\theta, \tag{5}$$

$$\frac{2|x_i - x|}{|x_2 - x_1|} = \sin^{2/p}\theta, \tag{6}$$

we derive our interpolation function as a function of p,

$$f(x) = \begin{cases} \dfrac{1 + \cos^{2/p}\theta}{2} f(x_1) + \dfrac{1 - \cos^{2/p}\theta}{2} f(x_2) & \text{for } x_1 \leq x < x_\mathrm{m} \\ \dfrac{1 - \cos^{2/p}\theta'}{2} f(x_1) + \dfrac{1 + \cos^{2/p}\theta'}{2} f(x_2) & \text{for } x_\mathrm{m} < x \leq x_2, \end{cases} \tag{7}$$

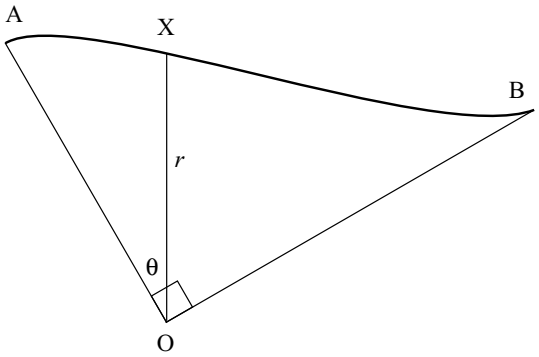

Figure 1. Geometric representation of a 1-D interpolation function. Point X is located on the curve connecting A and B, and is specified with r and θ, where $\mathrm{OX} = r$ and $\angle\mathrm{AOX} = \theta$ ($\angle\mathrm{AOB} = \pi/2$).

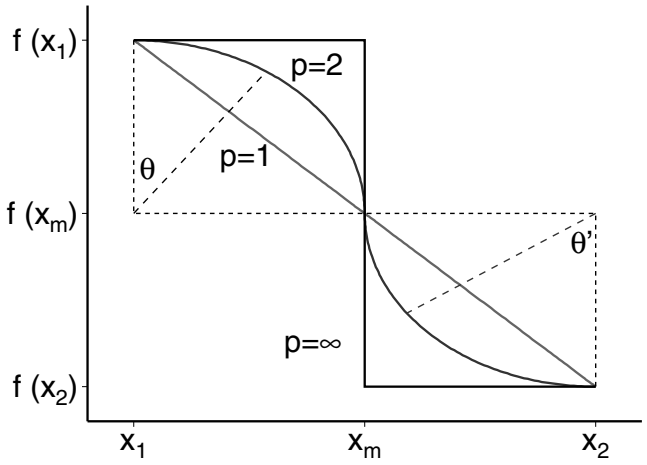

Figure 2. Interpolation function $f(x)$ for x ($x_1 \le x \le x_2$) for $p = 1$, 2 and ∞ (eq. 7). θ and θ' for $p = 2$ (Euclidean norm) are shown as examples.

where

$$x = \begin{cases} \left(1 - \dfrac{\sin^{2/p}\theta}{2}\right) x_1 + \dfrac{\sin^{2/p}\theta}{2} x_2 & \text{for } x_1 \le x < x_{\mathrm{m}} \\ \dfrac{\sin^{2/p}\theta'}{2} x_1 + \left(1 - \dfrac{\sin^{2/p}\theta'}{2}\right) x_2 & \text{for } x_{\mathrm{m}} < x \le x_2, \end{cases} \tag{8}$$

and $0 \le \theta, \theta' < \pi/2$ (Fig. 2). This is a natural extension of commonly used interpolation functions; those with $p = 1$ and $p = 2$ are interpolation functions with absolute value norm and Euclidean norm, respectively (see also fig. 1.06a of Parker 1994).

3 INTERPOLATION OF VELOCITY AND SLOWNESS

We evaluate the relationship between velocity interpolated with the V-grid, $v_p^{\mathrm{V}}(x)$, and slowness interpolated with the S-grid, $s_p^{\mathrm{S}}(x)$, in respect of $\xi = v_2/v_1$ with our interpolation function. First we introduce R as a measure of reciprocal failure, which we define as $v_p^{\mathrm{V}}(x)\, s_p^{\mathrm{S}}(x)$, or equivalently $v_p^{\mathrm{V}}(x)/v_p^{\mathrm{S}}(x)$. We hereafter assume $0 < \xi < 1$ (i.e. $v_1 > v_2$) with no loss of generality. From eqs (7) and (8),

$$R = \frac{1 + \cos^{4/p}\theta}{2} + \frac{1 - \cos^{4/p}\theta}{4}\left(\frac{1}{\xi} + \xi\right) \tag{9}$$

for $x_1 \le x \le x_2$. Here we substitute θ for θ', since R is symmetrical with regard to the mid-point (Fig. 3). R is ideally always unity, but is actually larger than unity. A smaller p gives a larger R. R takes the maximum value at the mid-point, x_{m}.

R being non-unity causes disagreement of traveltimes along $[x_1, x_2]$ between the V-grid and the S-grid. From eqs (7) and (8), we derive the ratio between traveltimes derived from the V-grid, T_{V}, and from the S-grid, T_{S}, (see Appendix A) as

$$\frac{T_{\mathrm{V}}}{T_{\mathrm{S}}} = \frac{8\xi}{x_2 - x_1} \int_0^{(x_2 - x_1)/2} \left\{ (1+\xi)^2 - (1-\xi)^2 \left[1 - \left(\frac{2x}{x_2 - x_1} \right)^p \right]^{2/p} \right\}^{-1} dx. \tag{10}$$

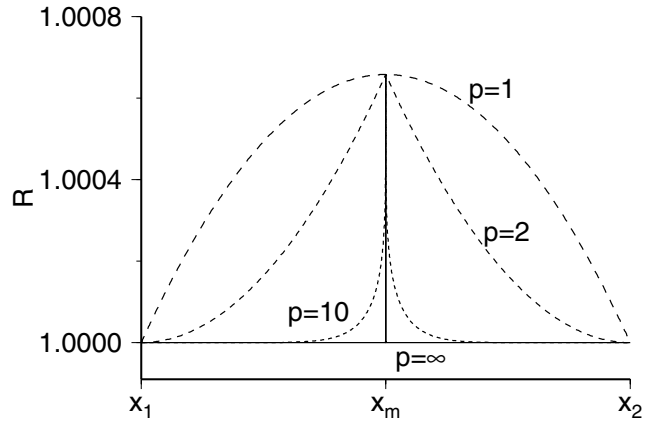

Figure 3. $R = v(x)\, s(x)$ as a function of p and x. $v(x)$ is derived from v_1 and v_2 (fixed values at two gridpoints), and $s(x)$ is derived from interpolation of s_1 ($=1/v_1$), and s_2 ($= 1/v_2$). $R = 1$ when reciprocity is satisfied. $\xi(=v_2/v_1) = 0.95$ is arbitrarily chosen for this plot.

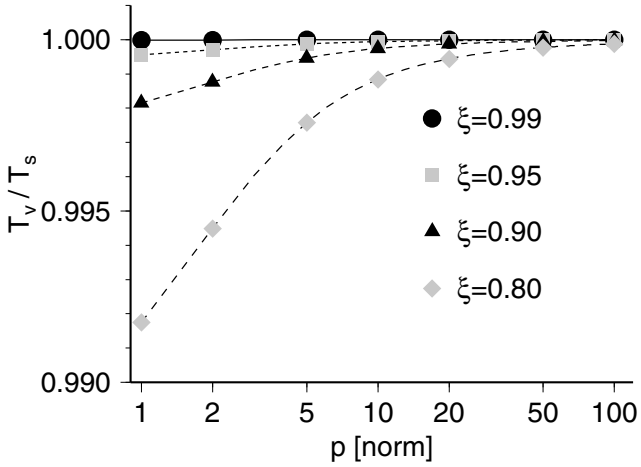

Figure 4. Ratio between calculated traveltimes from velocity and slowness grid models. T_V/T_S, which should be unity when the reciprocity is satisfied, monotonically converges to unity with an increase of ξ and p.

For example, the absolute value norm ($p = 1$) gives

$$\frac{T_V}{T_S} = \frac{2\log\xi}{\xi - 1/\xi}. \tag{11}$$

T_V/T_S is always less than unity (Fig. 4), which is a natural consequence of $R > 1$ (eq. 9 and Fig. 3). T_V/T_S converges to unity for large ξ and/or p.

Regardless of what value ξ takes, both R and T_V/T_S gradually approach unity as p increases (Figs 3 and 4). When p is infinitely large, the norm becomes the maximum norm,

$$\|x\|_\infty \equiv \max(|x_n|). \tag{12}$$

Use of this maximum norm always guarantees $R = 1$ (except at x_m, which is a singularity), and $T_V/T_S = 1$. Hence, $p = \infty$ should be applied to eq. (7) as the most suitable interpolation function in the grid representation when the reciprocal relation between velocity and slowness should be strictly satisfied in the entire model space. Grid representation with an interpolation using the $p = \infty$ norm is actually a block representation; a value at an arbitrary point in $[x_1, x_2]$ is equal to the value at the nearest gridpoint (Fig. 2). Spakman & Bijwaard (2001) argued that induction of artificial boundaries may not be a drawback of block representation. Successful applications of the block representation include recent whole mantle traveltime tomography (e.g. van der Hilst *et al.* 1997; Kennett *et al.* 1998; Bijwaard *et al.* 1998; Widiyantoro *et al.* 2000; Spakman & Bijwaard 2001).

4 INTERPOLATION IN THREE-DIMENSIONAL GRIDS

Our interpolation function is easily extended to 3-D grids using the same scheme. One popular (e.g. Thurber 1983; Zhao *et al.* 1992) linear interpolation function in three dimensions is

$$f(x,y,z) = \sum_{i,j,k=1}^{2} f(x_i,y_j,z_k)\left(1-\left|\frac{x-x_i}{x_1-x_2}\right|\right)\left(1-\left|\frac{y-y_j}{y_1-y_2}\right|\right)\left(1-\left|\frac{z-z_k}{z_1-z_2}\right|\right), \tag{13}$$

where (x, y, z) is located in a volume defined by eight gridpoints (x_i, y_j, z_k) for $i, j, k = 1, 2$. At the centre of these eight points, the mid-point rule (2) should apply;

$$(x_m, y_m, z_m) = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}, \frac{z_1+z_2}{2}\right), \tag{14}$$

$$f(x_m, y_m, z_m) = \sum_{i,j,k=1}^{2} \frac{1}{2^3} f(x_i, y_j, z_k). \tag{15}$$

By replacing the absolute norm in eq. (13) with the pth norm, we derive

$$f(x,y,z) = \sum_{i,j,k=1}^{2} \frac{1}{2^3} f(x_i,y_j,z_k)(1+\epsilon_i\cos^{2/p}\theta_x)(1+\epsilon_j\cos^{2/p}\theta_y)(1+\epsilon_k\cos^{2/p}\theta_z), \tag{16}$$

where

$$\epsilon_l = \begin{cases} 1 \quad \text{for} \begin{cases} l = 1,\ w_1 \le w < w_{\rm m} & \text{or} \\ l = 2,\ w_{\rm m} < w \le w_2 \end{cases} \\ -1 \ \text{for} \begin{cases} l = 1,\ w_{\rm m} < w \le w_2 & \text{or} \\ l = 2,\ w_1 \le w < w_{\rm m} \end{cases} \end{cases} \tag{17}$$

and

$$\theta_w = \arcsin \left| \frac{2(w_l - w)}{w_2 - w_1} \right|^{p/2} \begin{cases} l = 1 \ \text{for} & w_1 \le w < w_{\rm m} \\ l = 2 \ \text{for} & w_{\rm m} < w \le w_2 \end{cases} \tag{18}$$

for $(w, l) = (x, i)$, (y, j) and (z, k).

By comparing eqs (7) and (16), it is evident that dependence on $\cos^{2/p}\theta$ terms, which is the source of reciprocal failure, is raised to the third power in 3-D cases. Although the effect appears to be relatively small when considering only two adjacent nodes, these effects are integrated along a ray path, which may lead to an indispensable difference in the computed traveltime. As is shown for the 1-D case, velocity interpolated with the V-grid, $v_p^{\rm V}(x, y, z)$, and slowness interpolated with the S-grid, $s_p^{\rm S}(x, y, z)$, always satisfy the reciprocal relationship when the maximum norm is applied.

5 SYSTEMATIC DIFFERENCE IN TOMOGRAPHIC MODELS

Reciprocal failure consequently causes mutually inconsistent tomographic results in two parametrizations, the V-grid and the S-grid. The traveltime residual ΔT for a given ray path is expressed in integral form as

$$\Delta T = \int_{\rm path} -\frac{\Delta v}{v^2} dl \tag{19}$$

in the V-grid, or

$$\Delta T = \int_{\rm path} \Delta s dl \tag{20}$$

in the S-grid, where Δv and Δs are the velocity and slowness perturbation respectively. Δv and Δs satisfy

$$\Delta s = -\frac{\Delta v}{v^2} \quad \text{for} \quad \Delta v \ll v. \tag{21}$$

Discretized forms of eqs (19) and (20) are

$$\Delta T_{\rm V} = M_{\rm V} \Delta v \tag{22}$$

and

$$\Delta T_{\rm S} = M_{\rm S} \Delta s, \tag{23}$$

respectively, where $M_{\rm V}$ and $M_{\rm S}$ indicate discretization of the corresponding Fréchet kernel operators, and $\Delta T_{\rm V}$ and $\Delta T_{\rm S}$ are traveltime residuals calculated in the respective grids. In seismic tomography, the velocity perturbation Δv and slowness perturbation Δs are usually obtained by iteratively solving eqs (22) and (23) respectively. When the same interpolation function is used in both V- and S-grids and calculated ray paths are identical, we evidently have a linear relationship between two matrix operators,

$$M_{\rm S} = -v^2 M_{\rm V}. \tag{24}$$

Traveltime residuals in these equations (eqs 22 and 23) are formally defined as

$$\Delta T_{\rm V} = T^{\rm obs} - T_{\rm V}^{\rm cal} \tag{25}$$

$$\Delta T_{\rm S} = T^{\rm obs} - T_{\rm S}^{\rm cal}, \tag{26}$$

where $T^{\rm obs}$ is observed traveltime and $T_{\rm V}^{\rm cal}$ and $T_{\rm S}^{\rm cal}$ are traveltimes calculated in V- and S-grids respectively. $T_{\rm V}^{\rm cal}$ does not always agree with $T_{\rm S}^{\rm cal}$, and the difference $\Delta T^{\rm err}$, which we define as

$$\Delta T^{\rm err} = T_{\rm V}^{\rm cal} - T_{\rm S}^{\rm cal} = \Delta T_{\rm S} - \Delta T_{\rm V}, \tag{27}$$

depends both on initial models and their perturbations, primarily as a function of ray paths. Reciprocal failure does not occur in block models, and $T_{\rm V}^{\rm cal} = T_{\rm S}^{\rm cal}$, or $\Delta T^{\rm err} = 0$, is always guaranteed. This makes eqs (22) and (23) equivalent, and as a result Δv solved for V-grid and Δs for S-grid are mutually consistent. In grid models, $\Delta T^{\rm err}$ is generally non-zero due to reciprocal failure, and solutions of eqs (22) and (23) are mutually inconsistent.

Let us consider traveltime tomography in 1-D model space, which we solve both in V-grid and S-grid models using eqs (22) and (23). When locations of the sources and observation stations are fixed, ray paths in both models are always identical, which validates eq. (24). We can derive the slowness perturbation using eqs (22), (23), (24) and (27),

$$\begin{aligned}\Delta s &= M_{\mathrm{S}}^{-1}\Delta T_{\mathrm{S}} \\ &= -M_{\mathrm{V}}^{-1}\left(\Delta T_{\mathrm{V}} + \Delta T^{\mathrm{err}}\right)/v^2 \\ &= -\Delta v/v^2 - M_V^{-1}\Delta T^{\mathrm{err}}/v^2.\end{aligned} \tag{28}$$

Eq. (28) is inconsistent with eq. (21), and this implies that the tomographic results of eq. (22) disagree with those of eq. (23) unless $M_{\mathrm{V}}^{-1}\Delta T^{\mathrm{err}}/v^2 = 0$, or more specifically, $\Delta T^{\mathrm{err}} = 0$.

In 3-D model space, when the interpolation function is based on finite pth-norm, failure of velocity–slowness reciprocity between V-grid and S-grid models occurs except on grid nodes, ray paths in the two models do not agree with each other, and resultant perturbations Δv and Δs are generally incompatible. Even when differences in the ray paths in the two grid models are negligible, tomographic results are affected by the reciprocity problem as eq. (28) states.

The norm of the traveltime residual $\|\Delta T\|$ usually decreases as the model is updated through iteration, and the relative importance of $\|\Delta T^{\mathrm{err}}/T^{\mathrm{obs}}\|$ compared with $\|\Delta T/T^{\mathrm{obs}}\|$ becomes progressively larger. $\|\Delta T^{\mathrm{err}}/T^{\mathrm{obs}}\|$ and $\|\Delta T/T^{\mathrm{obs}}\|$ (ΔT being either ΔT_{V} or ΔT_{S}) are controlled by specific aspects of individual tomographic problems, such as distribution and geometry of ray paths and the quality of the observations. When the differences between ray paths are very small and $\|\Delta T^{\mathrm{err}}/T^{\mathrm{obs}}\| \ll \|\Delta T/T^{\mathrm{obs}}\|$ is satisfied, reciprocity failure has a negligible influence on tomography and results in both grid models should be compatible. In block models, the reciprocal relationship between velocity and slowness is always valid; ray paths do not vary with parametrization, either with velocity or slowness, and results are consistent with each other.

6 NUMERICAL EXAMPLES

To illustrate whether reciprocity failure actually has a significant effect on tomographic results when conditions are met, i.e. $\|\Delta T^{\mathrm{err}}/T^{\mathrm{obs}}\| \ll \|\Delta T/T^{\mathrm{obs}}\|$, we present numerical examples of grid-based traveltime tomography. Our 2-D examples simulate a regional tomography with teleseismic observations, and we solve the problem with both V-grid and S-grid models.

We set a rectangular region of 140 km × 100 km, which is represented by 165 equally spaced grid nodes. Interpolation is based on a norm of $p = 1$ in both V-grid and S-grid. An inclined zone of +5 per cent velocity anomaly in the upper mantle (30–100 km) is superposed on the reference 1-D structure, which represents a subducting plate (Fig. 5a). We randomly place earthquakes and observation sites, and calculate mock data traveltimes for 1000 ray paths, T^{obs}, using a slowness field interpolated with an S-grid. No artificial noise term is added to mock data. Ray paths and locations of earthquakes are fixed in the following procedure.

Our initial model is the reference 1-D model (Fig. 5b), and we invert our observations for the model perturbations, either velocity or slowness. Our solutions are internally consistent; in a case where velocity is the parameter solved for, we use the V-grid for both forward

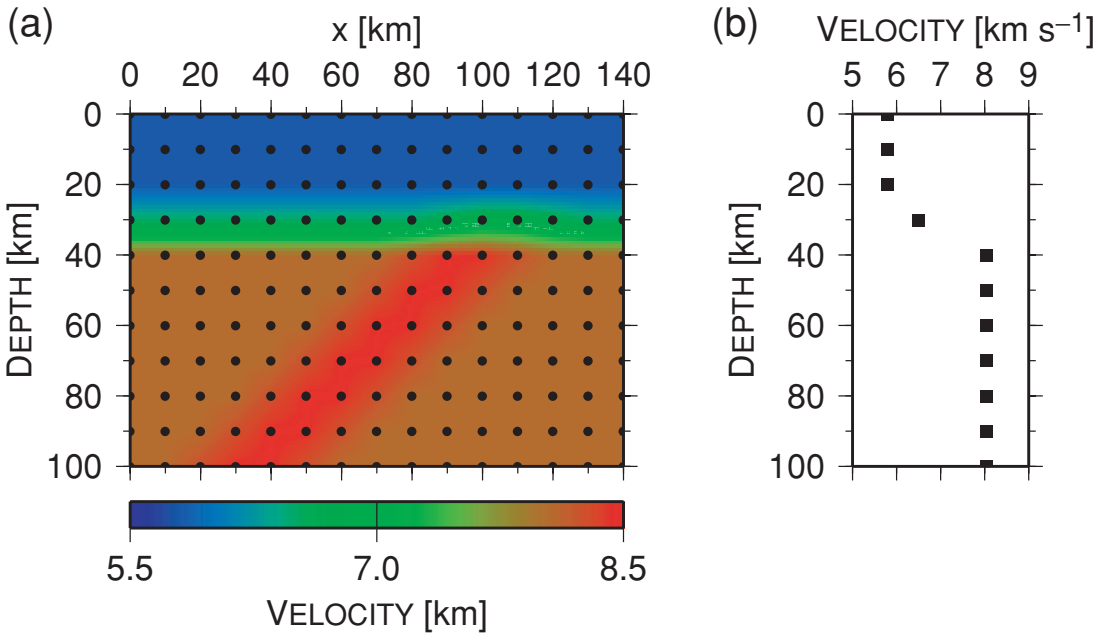

Figure 5. Geometry of our 2-D numerical experiments. A rectangular region of 140 km × 100 km is assumed (a), which is represented by 165 equally spaced grid nodes (grid spacing is 10 km). We simulate a subducting plate by an inclined zone of +5 per cent velocity anomaly in the upper mantle, and superimpose it on the reference 1-D structure (b). The reference 1-D model is used as our initial model.

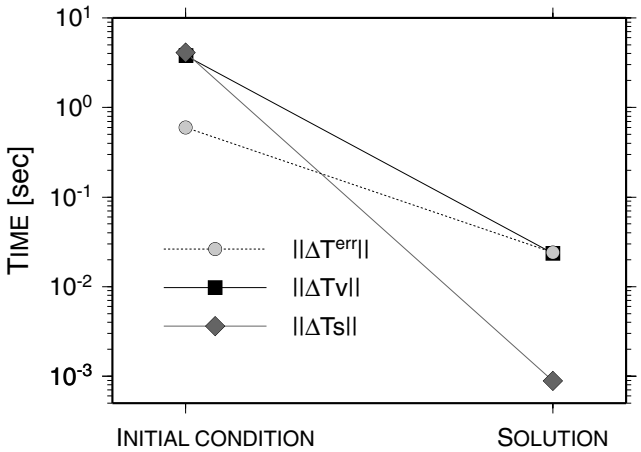

Figure 6. Traveltime residuals $\|\Delta T\|$ and the difference of traveltimes in two grids, $\|\Delta T^{\rm err}\|$, before and after the inversion. Traveltime residuals of V-grid and S-grid models are $\|\Delta T_{\rm V}\|$ and $\|\Delta T_{\rm S}\|$, respectively. In this plot, we define $\|\Delta T\| = \sqrt{\sum |T^{\rm obs} - T^{\rm cal}|^2}$ and $\|\Delta T^{\rm err}\| = \sqrt{\sum |\Delta T_{\rm V} - \Delta T_{\rm S}|^2}$, where summation is taken for all ray paths.

(traveltime computation) and inverse (computation of Fréchet kernel) problems. The velocity gradient in the initial model causes inconsistency between traveltimes in the respective grids, $T_{\rm V}^{\rm cal}$ and $T_{\rm S}^{\rm cal}$, as a result of reciprocity failure. When differences between ray paths in the two models are negligible, $\|\Delta T^{\rm err}\|$ being smaller than $\|\Delta T\|$ is the required condition for reciprocal failure not to cause bifurcation of tomographic results in two parametrizations. Initially the traveltime difference $\|\Delta T^{\rm err}\|$ is smaller than the traveltime residuals $\|\Delta T\|$ (Fig. 6).

We solve eqs (22) and (23) in the respective grids using the subroutine LSBRR in the IMSL library, which solves a linear least-squares problem with iterative refinement. For our examples, results do not depend on the choice of solver, as other solvers, including conjugate gradient types, produce nearly identical results.

In the result of for S-grid tomography (Fig. 7b) an inclined high-velocity anomaly shows up, and its shape and amplitude are consistent with the assumed velocity perturbation (Fig. 7a). For this figure (Fig. 7b), we first represent the entire field with slowness and converted it into a velocity field so that the model is not altered during drawing. Since we calculate mock data traveltimes with the S-grid, this match between Figs 7(a) and (b) indicates that our tomographic problem is fairly well-posed.

An inclined high-velocity anomaly shows up in the result with the V-grid as well (Fig. 7c). However, its amplitude is not well recovered, and small patches of low-velocity anomaly surround the high-velocity slab. Additionally, horizontal stripes of velocity anomaly appear in the shallow region. The choice of parameter to represent the field is the only difference between the two computations, and these additional features in the latter result (Fig. 7c) are evidently caused by $\|\Delta T^{\rm err}\|$ in V-grid tomography. This model appears to be dominated by negative anomalies; in the presence of velocity gradient, $T_{\rm V}^{\rm cal}$ is always smaller than $T_{\rm S}^{\rm cal}$, as is shown in eq. (10) and Fig. 4, and this shifts $\Delta T_{\rm V}$ to larger (or positive) values than $\Delta T_{\rm S}$, which in turn always perturbs the velocities to lower (or negative) values. $\|\Delta T\|$ becomes smaller than $\|\Delta T^{\rm err}\|$ in the solutions (Fig. 6), and this violates the condition that the influence of reciprocal failure is negligible.

When $T^{\rm obs}$ is calculated with the V-grid the situation reverses; small-scale velocity heterogeneities are superposed on the well-resolved slab-shaped velocity anomaly in S-grid tomography. This demonstrates that the assumption of eq. (21) does not affect our results. Neither are these results controlled by the sign of the velocity anomaly; we also test cases with a -5 per cent slab-shaped velocity anomaly and come to a similar conclusion. Therefore we can conclude that reciprocal failure is the only source of the systematic difference, that is, the resultant velocity shifts to higher values when slowness is the parameter solved for.

7 DISCUSSION AND CONCLUSION

Linear interpolation with $p = 1$ is widely used in traveltime tomography, but in the presence of heterogeneity traveltimes calculated with two different representations of the same structure, velocity and slowness grids, do not usually agree. Difference between two traveltimes should be large when strong velocity variations exist between adjacent grid nodes in the model, and slowness grids always give larger traveltimes. When a tomographic problem is formulated with one parameter, either velocity or slowness, the model is internally consistent within the respective parametrization. However, resultant models in two parametrizations are different and mutually incompatible unless careful formulation of the tomographic problem guarantees that the traveltime difference between V- and S-grid models is sufficiently small for each path.

In our numerical examples (Fig. 7) we demonstrate that use of small p actually brings about that tomographic results depend on the choice of parameter in the computations, velocity or slowness. This is consistent with our theoretical analysis which predicts that the S-grid always shifts the resultant velocity to higher values than the V-grid. The traveltime difference, $T_{\rm S}^{\rm cal} - T_{\rm V}^{\rm cal}$, that is induced by reciprocal failure, $vs \neq 1$, would be positive and accumulated for the paths. Seismic velocity generally increases with depth in the mantle (e.g. Gaherty *et al.* 1999), and the majority of ray paths sample the gradient zone in regional tomography. A similar consideration applies to interpolation in horizontal directions when the velocity structure deviates strongly from a 1-D (depth-dependent) structure, such as in subduction zones. This

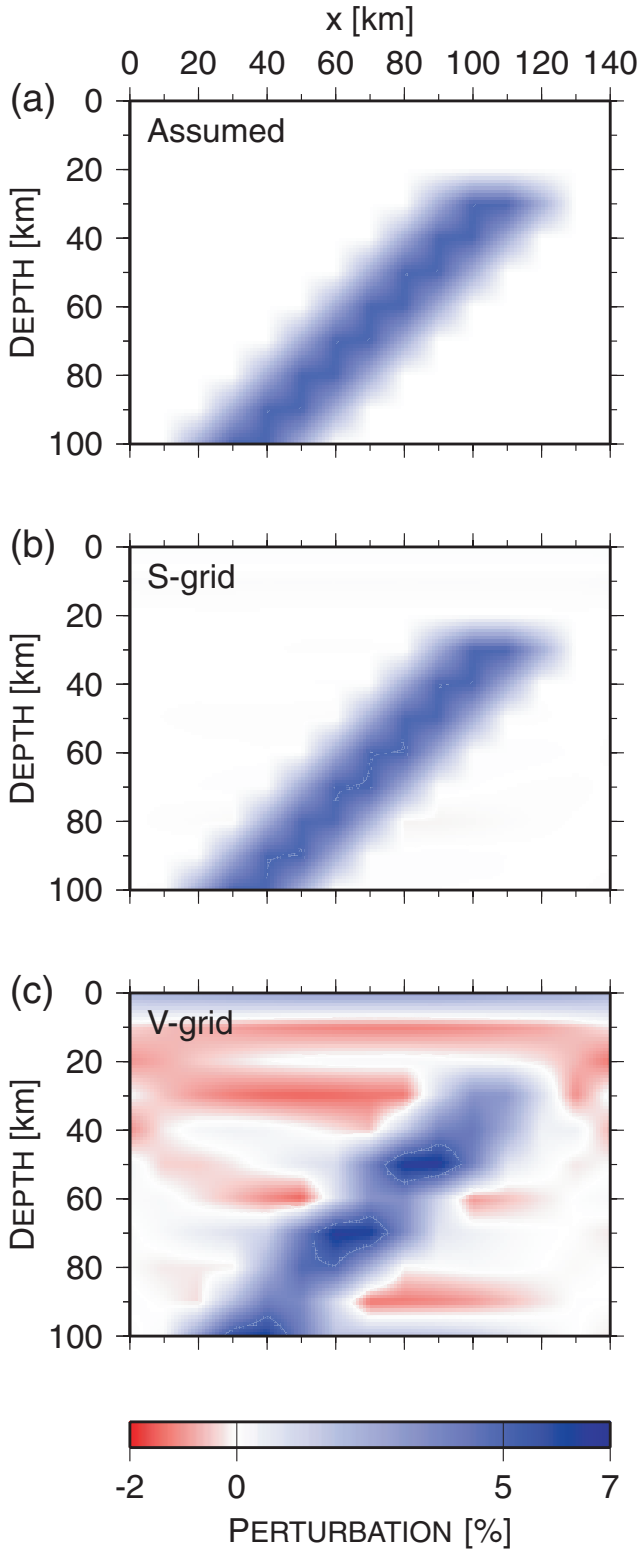

Figure 7. Results of our numerical experiments. Blue and red colours indicate positive and negative velocity anomalies respectively. The scale, shown at the bottom in per cent, is common to the three plots. We expect that any difference between the initial model and the assumed velocity model (a) will be recovered in the tomographic results under ideal conditions. The result using S-grid tomography (b) shows the well-recovered slab-shaped velocity anomaly, whose shape and amplitude are identical to the expected result. The shape and amplitude of the assumed slab are less well recovered in the result using V-grid tomography (c), and additional velocity anomalies are also present. The average velocity perturbation in the 2-D space, $\overline{\Delta v}$, in the V-grid model (c) is 0.717 km s^{-1}, i.e. smaller than $\overline{\Delta v}$ in the S-grid model (b) at 0.789 km s^{-1}, the latter of which is consistent with $\overline{\Delta v}$ in the assumed velocity model (a), 0.789 km s^{-1}.

term acts as a fictitious source of traveltime difference between two parametrizations, and will propagate onto the tomographic images (e.g the additional negative perturbations in Fig. 7(c) compared with Fig. 7(b)). How different the results for V- and S-grid tomographic models would be is dictated by the relative magnitude of $\|\Delta T^{\rm err}/T^{\rm obs}\|$ and $\|\Delta T/T^{\rm obs}\|$. $\|\Delta T/T^{\rm obs}\|$ decreases as the model is updated via inversion, and this increases the relative importance of $\|\Delta T^{\rm err}/T^{\rm obs}\|$.

Even when interpolation with $p = 1$ (i.e. the grid model) is adopted, we can lessen the inconsistency between V- and S-grids by taking $\xi \rightarrow 1$. This could be achieved by taking a small grid spacing, since small ξ occurs in regions of strong small-scale heterogeneity or of large velocity gradient. Dense grid spacing, however, is numerically expensive, and could induce instability in matrix inversion in traveltime tomography. One practical approach is to give up regularly spaced grids or cells. Irregular grids/cells have been applied to geophysical problems (e.g. Sambridge *et al.* 1995; Bijwaard *et al.* 1998; Spakman & Bijwaard 2001).

There are several other approaches for interpolation. Among them is to use a large and finite p with our scheme. This would combine the merits of two end-member models ($p = \infty$ and $p = 1$), but is computationally expensive and unrealistic for seismic traveltime tomography. Use of a geometric mean for the interpolation is also possible; $f(x) = f(x_1)^{\gamma}\ f(x_2)^{1-\gamma}$, $\gamma = |x_2 - x|/|x_2 - x_1|$, which always satisfies $R = 1$. This, however, relates traveltime and velocities at grid nodes non-linearly, and would also increase the computational burden. Other approaches to interpolation, such as that of Tarantola & Nercessian (1984) who formulated inversion in the functional approach, are also possible.

We should note that mixed use of two parametrizations, velocity and slowness, is not justified in most grid-based tomography. Reciprocal failure prohibits interchangeable use of velocity and slowness in any part of the tomographic computation. For example, when the ray tracing is formulated with velocity and traveltime is computed with slowness, the resultant ray paths and traveltimes on the same source–receiver geometry are internally inconsistent. Such a complicated formulation may hardly happen in the literature but could happen in program codes; frequent use of $1/v$ in traveltime computation may evoke naive use of s.

Our intention in this paper is neither to deny the results of grid-based traveltime tomography nor to insist on the absolute superiority of a block representation over a grid representation in traveltime tomography. It is a matter of which base function we choose to represent the field. We would like to note, however, that block representation is the only parametrization that always satisfies the reciprocity between velocity and slowness. Neither is it our intention to denounce grid-based tomography, either with slowness or with velocity grids; it is up to practitioners to choose the parameter used in computation. As long as the medium is properly approximated with the chosen base function and forward and inverse problems are properly solved, both parametrizations result in respectively authentic results. It is, accordingly, wrong to conclude that either of the two parametrizations, velocity or slowness, results in bias-free tomography or is superior to the other. However, our analysis demonstrates that results are systematically different between the two parametrizations, and practitioners are obliged to specify which parameter is actually solved for in presenting their tomographic model. How significant reciprocal failure would be is simultaneously controlled by other factors such as data quality and approximation in formulation and computation, and it is suggested that practitioners confirm which robust features are in their tomographic results. It is also the practitioner's obligation to confirm that mixed parametrization is not used in the computation. As for those who use tomographic models for geophysical or geological interpretation, we need to be cautious to avoid the pitfall of the parameters, for example in comparing models from separate groups or in applying model values to derive material parameters such as Poisson's ratio or elastic parameters.

ACKNOWLEDGMENTS

We appreciate constructive comments by Ichiro Nakanishi, Peiliang Xu and Shiro Ohmi. This study is partially supported by a research grant from Japan Society for the Promotion of Science Research Fellowships for Young Scientists, and the 21st Century Centers of Excellence Research Fellowships at the Disaster Prevention Research Institute, Kyoto University (MM), and by a MEXT grant-in-aid for young scientists (B), 14740266 (MK). We appreciate comments on earlier drafts by Nick Rawlinson and an anonymous reviewer, and the guidance and patience of the Editor Malcolm Sambridge through several rounds of revisions.

REFERENCES

Bijwaard, H., Spakman, W. & Engdahl E.R., 1998. Closing the gap between regional and global traveltime tomography, *J. geophys. Res.,* **103,** 30 055–30 078.

Gaherty, J.B., Kato, M. & Jordan, T.H., 1999. Seismological structure of the upper mantle: a regional comparison of seismic layering, *Phys. earth Planet. Inter.,* **110,** 21–41.

Kennett, B.L.N., Widiyantoro, S. & van der Hilst, R.D., 1998. Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. geophys. Res.,* **103,** 12 469–12 493.

Parker, R.L., 1994. *Geophysical Inverse Theory*. Princeton University Press, Princeton, NJ.

Sambridge, M., Braun, J. & McQueen, H., 1995. Geophysical parameterization and interpolation of irregular data using natural neighbours, *Geophys. J. Int.,* **122,** 837–857.

Spakman, W., 1991. Delay-time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geophys. J. Int.,* **107,** 309–332.

Spakman, W. & Bijwaard, H., 2001. Optimization of cell parameterizations for tomographic inverse problems, *Pure appl. Geophys.,* **158,** 1401–1423.

Tarantola, A. & Nercessian, A., 1984. Three-dimensional inversion without blocks, *Geophys. J. R. astr. Soc.,* **76,** 299–306.

Thurber, C.H., 1983. Earthquake locations and three-dimensional crustal structure in Coyote Lake area, central California, *J. geophys. Res.,* **88,** 8226–8236.

van der Hilst, R.D., Widiyantoro, S. & Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography, *Nature,* **386,** 578–584.

Widiyantoro, S., Gorbatov, A., Kennett, B.L.N. & Fukao, Y., 2000. Improving global shear wave traveltime tomography using three-dimensional ray tracing and iterative inversion, *Geophys. J. Int.,* **141,** 747–758.

Zhao, D., Hasegawa, A. & Horiuchi, S., 1992. Tomographic imaging of *P* and *S* wave velocity structure beneath northeastern Japan, *J. geophys. Res.,* **97,** 19 909–19 928.

APPENDIX A: CALCULUS OF THE RATIO BETWEEN V-GRID AND S-GRID TRAVELTIMES

The interpolation function (7) is rewritten by eq. (8) as

$$f(x) = \begin{cases} \dfrac{f(x_1)+f(x_2)}{2} + \dfrac{f(x_1)-f(x_2)}{2}\left[1-\left(2\dfrac{x-x_1}{x_2-x_1}\right)^p\right]^{1/p} & \text{for } x_1 \le x < x_\mathrm{m} \\ \dfrac{f(x_1)+f(x_2)}{2} - \dfrac{f(x_1)-f(x_2)}{2}\left[1-\left(2\dfrac{x_2-x}{x_2-x_1}\right)^p\right]^{1/p} & \text{for } x_\mathrm{m} < x \le x_2, \end{cases} \tag{A1}$$

where $f(x)$ is continuous at x_m as eq. (2). Traveltime T_V along $[x_1, x_2]$ with interpolation (A1) is

$$T_\mathrm{V} = 4(v_1+v_2)\int_0^{(x_2-x_1)/2} \left\{(v_1+v_2)^2 - (v_1-v_2)^2\left[1-\left(\frac{2x}{x_2-x_1}\right)^p\right]^{2/p}\right\}^{-1} dx. \tag{A2}$$

Traveltime T_S along $[x_1, x_2]$ is similarly

$$T_\mathrm{S} = \frac{1}{2}(x_2-x_1)(s_1+s_2), \tag{A3}$$

which means that the traveltime with the S-grid along the identical ray path does not depend on p. We furthermore note that T_S of the grid model is identical with that of the block model.

From eqs (A2) and (A3),

$$\frac{T_\mathrm{V}}{T_\mathrm{S}} = \text{eq.(10)}.$$