Accuracy and resolution of the travel–time tomography

Luděk Klimeš

Department of Geophysics, Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic, Fax: +420-2-2191 1214, E-mail: klimes@seis.karlov.mff.cuni.cz

Summary

The geological medium is assumed to be a representation of a random medium described in terms of the first two statistical moments: the mean value and the medium correlation function. The proposed approach to the seismic travel–time tomography enables to derive the equations for the inversion of travel times without additional subjective a priori information, and to estimate the accuracy and resolution of the resulting seismic model of a geological structure.

Keywords

Seismic model, travel–time tomographic inversion, covariance functions and matrices, smoothing, resolution.

1 Introduction

Seismic (macro) model approximates the spatial distribution of material properties of the geological medium.

It is obvious that the geological structures cannot be definitely described in terms of a finite number of model parameters: going from large to small scales, more and more details of the medium may be discovered. From this point of view, the seismic model can never represent all details of the medium, no matter how numerous are seismic data.

On the other hand, we cannot consider individual small details of the medium to be mutually independent. Such a medium, with statistical properties of a white noise, would be unpredictable from a finite number of seismic data and seismic inversion would be nonsense. We must assume that the probability of the same or similar material properties in two points increases with decreasing distance of the points. Such a property of the medium may, in the roughest approximation, be described in terms of the medium correlation function.

In this paper, the geological medium is assumed to be a representation of a random medium with known statistical properties. We assume to know some travel times corresponding to the geological medium (i.e. to the representation of a random medium) and determine the seismic model as the "best" approximation to the (possibly smoothed) geological medium. The "best" approximation means the smallest Lebesgue L_2 norm of the standard deviations between the model and the (possibly smoothed) geological medium.

This approach enables to arrive at the final formulae for linearized travel–time tomography and to separate the smoothing effects of (a) the inversion for non–smoothed medium, (b) the inversion for smoothed medium, (c) the finite model parametrization, and (d) the additional a priori smoothing information specified in terms of the Sobolev norm.

The Einstein summation over the repeating indices is used throughout the paper.

In: Seismic Waves in Complex 3-D Structures, Report 6 (Department of Geophysics, Charles University, Prague, September 1997) [*Considerably revised version of the corresponding paper of* Report 4 (September 1996)]

2 Formulation

2.1 Random slowness

Assume that the slowness $u(\mathbf{x})$ in the geological structure is a realization of the random quantity with the first statistical moment (*mean value*)

$$u^0(\mathbf{x}) = \langle u(\mathbf{x}) \rangle \tag{1}$$

and the second statistical moment

$$\langle u(\mathbf{x}_1) u(\mathbf{x}_2) \rangle \quad . \tag{2}$$

The *medium correlation function* (also called *medium covariance function*) of the slowness is

$$C(\mathbf{x}_1, \mathbf{x}_2) = \langle u(\mathbf{x}_1) u(\mathbf{x}_2) \rangle - u^0(\mathbf{x}_1) u^0(\mathbf{x}_2) \quad . \tag{3}$$

Here $u^0(\mathbf{x})$ is the a priori known deterministic component of the random medium serving as the background slowness when estimating the medium correlation function $C(\mathbf{x}_1, \mathbf{x}_2)$, which should be determined prior to the proposed tomographic inversion.

We shall see that the determination of $u^0(\mathbf{x})$ by means of very strongly overdetermined least squares is consistent with the proposed tomographic inversion. The least–square fitting of $u^0(\mathbf{x})$ must be very strongly overdetermined in order to leave a sufficient amount of information for the estimation of the medium correlation function $C(\mathbf{x}_1, \mathbf{x}_2)$.

The determination of $u^0(\mathbf{x})$, parametrized by several parameters, and $C(\mathbf{x}_1, \mathbf{x}_2)$, parametrized by one or few parameters, should be carried out prior to the proposed tomographic inversion. The determination of $u^0(\mathbf{x})$ and $C(\mathbf{x}_1, \mathbf{x}_2)$ is demonstrated by Klimeš (1996). Various types of the medium correlation functions $C(\mathbf{x}_1, \mathbf{x}_2)$ are systematically listed by Klimeš (1997).

2.2 Rays and travel times

The *field travel times* T_I are the integrals of the slowness $u(\mathbf{x})$ along rays Ω_I plus the *picking errors*,

$$T_I = \int S_I(\mathbf{x}) u(\mathbf{x}) \, \mathrm{d}\mathbf{x} + \delta T_I \quad , \tag{4}$$

where the integral is taken over the model volume. *Uppercase subscripts* are used hereinafter to denote travel times and the corresponding rays.

From the point of view of the high–frequency asymptotic ray theory, the *ray characteristic function* (called also "travel–time sensitivity function", or "sampling function") $S_I(\mathbf{x})$ of ray Ω_I has the form of 2-D Dirac distribution $\delta(q^1, q^2)$ in ray–centred coordinates (q^1, q^2, q^3), with q^1 and q^2 being the Cartesian coordinates in the plane perpendicular to the ray parametrized by arclength q_3,

$$S_I(\mathbf{x}(q^1, q^2, q^3)) = \delta(q^1, q^2) \tag{5}$$

for the values of q^3 between the source and the receiver. Let us note that the ray characteristic function may also describe a "thick" ray in a general form of

$$S_I(\mathbf{x}(q^1, q^2, q^3)) = f_I(q^1, q^2, q^3) \left[\iint f_I(q^1, q^2, q^3) \mathrm{d}q^1 \mathrm{d}q^2 \right]^{-1} \quad . \tag{6}$$

For example, it may describe the paraxial effective Fresnel volume in the Gaussian form of

$$f_I(q^1,q^2,q^3) = \exp\Big(-\omega\tfrac{1}{2}\sum_{\alpha=1}^{2}\sum_{\beta=1}^{2} q^\alpha |M|_{\alpha\beta}(q^3) q^\beta\Big) \tag{7}$$

with ω being a given constant and $|M|_{\alpha\beta}(q^3)$ being the absolute value of the 2×2 matrix of the second travel–time derivatives with respect to q^1 and q^2.

We assume zero mean values of the picking errors,

$$\langle \delta T_I \rangle = 0 \quad , \tag{8}$$

and introduce the covariance matrix of picking errors (data covariance matrix)

$$T_{IJ} = \langle \delta T_I \; \delta T_J \rangle \tag{9}$$

describing errors δT_I. We assume no correlation between picking errors δT_I and exact slowness $u(\mathbf{x})$,

$$\langle \delta T_I \; u(\mathbf{x}) \rangle = 0 \quad . \tag{10}$$

2.3 Seismic model

Slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model is considered to be an approximation to the smoothed version $\overline{u}(\mathbf{x})$ of the slowness $u(\mathbf{x})$ of the medium,

$$\overline{u}(\mathbf{x}) = \int \overline{R}(\mathbf{x},\mathbf{x}') \; u(\mathbf{x}') \; \mathrm{d}\mathbf{x}' \quad , \tag{11}$$

where $\overline{R}(\mathbf{x}_1,\mathbf{x}_2)$ is the given *smoothing function.* The smoothing function may express a priori ideas on the resolution of the model and may strongly influence the local standard deviation between the model and the geological medium, because the small details in the medium are much worse resolved than larger inhomogeneities. The *smoothing function may also be taken in the form of the Dirac distribution,* $\overline{R}(\mathbf{x}_1,\mathbf{x}_2) = \delta(\mathbf{x}_1-\mathbf{x}_2)$. In such a case, no a priori smoothing is applied during the tomographic travel–time inversion and the slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model is an approximation to the slowness $u(\mathbf{x})$ of the medium. A possible smoothing may then be applied to the model a posteriori. This approach may be, to some extent, limited only by an insufficiently coarse parametrization of the model and other functions.

The model slowness $\widetilde{u}(\mathbf{x})$ is dependent on the exact slowness $u(\mathbf{x})$ through travel times T_I. We assume a linear dependence of the model on travel times,

$$\widetilde{u}(\mathbf{x}) = R^0(\mathbf{x}) + R_K(\mathbf{x}) \left[T_K - T_K^0\right] \quad , \tag{12}$$

where we have formally denoted

$$T_I^0 = \int S_I(\mathbf{x}) u^0(\mathbf{x}) \; \mathrm{d}\mathbf{x} \tag{13}$$

the integrals with the ray characteristic functions $S_I(\mathbf{x})$ corresponding to exact rays Ω_I (reference travel times). Note that, because of (4) and (13) with (1) and (8),

$$\langle T_K - T_K^0 \rangle = 0 \quad . \tag{14}$$

We assume here that the differences between $u(\mathbf{x})$, $\overline{u}(\mathbf{x})$ and $\widetilde{u}(\mathbf{x})$ are sufficiently small that the corresponding travel times may be calculated along the same rays within the accuracy not distorting the covariance matrix Θ_{IJ} of travel times, which will be introduced later on. In other words, we assume that the *linearization* of travel times with respect to medium perturbations may be applied in the *vicinity of the correct model.*

After insertion of (4) and (13), equation (12) reads

$$\widetilde{u}(\mathbf{x}) = R^0(\mathbf{x}) + \int R(\mathbf{x},\mathbf{x}') \left[u(\mathbf{x}') - u^0(\mathbf{x}')\right] \, \mathrm{d}\mathbf{x}' + R_K(\mathbf{x})\delta T_K \quad , \tag{15}$$

where

$$R(\mathbf{x}_1,\mathbf{x}_2) = R_K(\mathbf{x}_1)S_K(\mathbf{x}_2) \quad , \tag{16}$$

is the *resolution function*, linearly projecting the exact slowness $u(\mathbf{x})$ onto the model $\widetilde{u}(\mathbf{x})$.

The resolution function may be constrained by zero to several a priori conditions indexed here by α,

$$\int R(\mathbf{x},\mathbf{x}')u^\alpha(\mathbf{x}') \, \mathrm{d}\mathbf{x}' = u^\alpha(\mathbf{x}) \quad . \tag{17}$$

These constraints express the condition that the travel–time inversion without smoothing gives exact results in the structures which differences from $u^0(\mathbf{x})$ lie in a linear subspace generated by few functions $u^\alpha(\mathbf{x})$. We shall see later that these conditions are equivalent to adjusting the mean slowness $u^0(\mathbf{x})$ by a linear combination of functions $u^\alpha(\mathbf{x})$, fitted by the least squares.

2.4 Parametrization of the model and resolution function

Assume that the slowness distribution $\widetilde{u}(\mathbf{x})$ in the seismic model, which is the best possible approximation to $\overline{u}(\mathbf{x})$, is composed of a finite set of *basis functions* $B_i(\mathbf{x})$ indexed hereinafter by *lowercase subscripts*,

$$\boxed{\widetilde{u}(\mathbf{x}) = B_i(\mathbf{x}) \; u_i \quad .} \tag{18}$$

For example, tricubic B-splines may be employed for the basis functions $B_i(\mathbf{x})$. Comparing now (18) with (12), we see that functions $R^0(\mathbf{x})$ and $R_K(\mathbf{x})$ should also be composed of the basis functions,

$$R^0(\mathbf{x}) = B_i(\mathbf{x})R^0_i \quad , \tag{19}$$

$$R_K(\mathbf{x}) = B_i(\mathbf{x})R_{iK} \quad , \tag{20}$$

that the resolution function (16) takes the form of

$$R(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)R_{iI}S_I(\mathbf{x}_2) \quad , \tag{21}$$

and that equation (12) can be expressed in the component form of

$$\boxed{u_i = R^0_i + R_{iI} \left[T_I - T^0_I\right] \quad .} \tag{22}$$

Assume that functions $u^\alpha(\mathbf{x})$ may also be composed of the same basis functions,

$$u^\alpha(\mathbf{x}) = u^\alpha_i \; B_i(\mathbf{x}) \quad . \tag{23}$$

Introducing now the variation

$$G_{iJ} = \int B_i(\mathbf{x})S_J(\mathbf{x}) \, \mathrm{d}\mathbf{x} \tag{24}$$

of the J^{th} synthetic travel time with respect to the i^{th} model slowness parameter, constraints (17) with (23) and (21) take the form of

$$R_{iI} T_I^\alpha = u_i^\alpha \tag{25}$$

with

$$T_I^\alpha = G_{jI} u_j^\alpha \quad . \tag{26}$$

Note that if the decomposition in the form of (23) takes force also for $u^0(\mathbf{x})$, equation (26) may be used to evaluate T_I^0.

2.5 Model covariance matrix

The covariance matrix describing the deviation of the model slowness $\widetilde{u}(\mathbf{x})$ from the smoothed exact slowness $\overline{u}(\mathbf{x})$ is

$$\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = \langle [\widetilde{u}(\mathbf{x}_1) - \overline{u}(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - \overline{u}(\mathbf{x}_2)] \rangle \quad . \tag{27}$$

Inserting (11) and (15) into (27), and considering (1), (8) and (10), we arrive at

$$\begin{aligned}
\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint \left[R(\mathbf{x}_1, \mathbf{x}_1') - \overline{R}(\mathbf{x}_1, \mathbf{x}_1')\right] \left[R(\mathbf{x}_2, \mathbf{x}_2') - \overline{R}(\mathbf{x}_2, \mathbf{x}_2')\right] \\
& \times \langle \left[u(\mathbf{x}_1') - u^0(\mathbf{x}_1')\right] \left[u(\mathbf{x}_2') - u^0(\mathbf{x}_2')\right] \rangle \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' \\
& + R_I(\mathbf{x}_1) R_J(\mathbf{x}_2) \langle \delta T_I \; \delta T_J \rangle \\
& + \left[R^0(\mathbf{x}_1) - \int \overline{R}(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1') \, \mathrm{d}\mathbf{x}_1'\right] \left[R^0(\mathbf{x}_2) - \int \overline{R}(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2') \, \mathrm{d}\mathbf{x}_2'\right] \quad .
\end{aligned} \tag{28}$$

Using then (3), (9), and (20), we arrive at

$$\begin{aligned}
\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint \left[R(\mathbf{x}_1, \mathbf{x}_1') - \overline{R}(\mathbf{x}_1, \mathbf{x}_1')\right] \left[R(\mathbf{x}_2, \mathbf{x}_2') - \overline{R}(\mathbf{x}_2, \mathbf{x}_2')\right] C(\mathbf{x}_1', \mathbf{x}_2') \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' \\
& + B_i(\mathbf{x}_1) B_j(\mathbf{x}_2) R_{iI} R_{jJ} T_{IJ} \\
& + \left[R^0(\mathbf{x}_1) - \int \overline{R}(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1') \, \mathrm{d}\mathbf{x}_1'\right] \left[R^0(\mathbf{x}_2) - \int \overline{R}(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2') \, \mathrm{d}\mathbf{x}_2'\right] \quad .
\end{aligned} \tag{29}$$

Insertion of (19) and (21) into (29) yields

$$\begin{aligned}
\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint \overline{R}(\mathbf{x}_1, \mathbf{x}_1') \overline{R}(\mathbf{x}_2, \mathbf{x}_2') C(\mathbf{x}_1', \mathbf{x}_2') \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' \\
& - B_i(\mathbf{x}_1) R_{iI} \iint \overline{R}(\mathbf{x}_2, \mathbf{x}_2') C(\mathbf{x}_2', \mathbf{x}_1') S_I(\mathbf{x}_1') \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' \\
& - B_j(\mathbf{x}_2) R_{jJ} \iint \overline{R}(\mathbf{x}_1, \mathbf{x}_1') C(\mathbf{x}_1', \mathbf{x}_2') S_J(\mathbf{x}_2') \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' \\
& + B_i(\mathbf{x}_1) B_j(\mathbf{x}_2) R_{iI} R_{jJ} \left[\iint S_I(\mathbf{x}_1') C(\mathbf{x}_1', \mathbf{x}_2') S_J(\mathbf{x}_2') \, \mathrm{d}\mathbf{x}_1' \, \mathrm{d}\mathbf{x}_2' + T_{IJ}\right] \\
& + \left[B_i(\mathbf{x}_1) R_i^0 - \int \overline{R}(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1') \, \mathrm{d}\mathbf{x}_1'\right] \left[B_j(\mathbf{x}_1) R_j^0 - \int \overline{R}(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2') \, \mathrm{d}\mathbf{x}_2'\right] \quad .
\end{aligned} \tag{30}$$

Introducing the *ray resolution functions*

$$\boxed{A_I(\mathbf{x}) = \langle T_I \; [u(\mathbf{x}) - u^0(\mathbf{x})] \rangle = \int C(\mathbf{x}, \mathbf{x}') S_I(\mathbf{x}') \, \mathrm{d}\mathbf{x}' \quad ,} \tag{31}$$

representing the covariances between slowness and field travel times, the a priori correlation matrix

$$S_{IJ} = \iint S_I(\mathbf{x}_1) C(\mathbf{x}_1, \mathbf{x}_2) S_J(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 = \int S_I(\mathbf{x}_1) A_J(\mathbf{x}_1)\ \mathrm{d}\mathbf{x}_1 \tag{32}$$

of the travel times, and the complete *travel–time covariance matrix*

$$\Theta_{MN} = \langle [T_M - T_M^0]\ [T_N - T_N^0] \rangle = T_{MN} + S_{MN} \quad , \tag{33}$$

equation (30) reads

$$\begin{aligned}
\widetilde{C}(\mathbf{x}_1, \mathbf{x}_2) = & \iint \overline{R}(\mathbf{x}_1, \mathbf{x}_1') \overline{R}(\mathbf{x}_2, \mathbf{x}_2') C(\mathbf{x}_1', \mathbf{x}_2')\ \mathrm{d}\mathbf{x}_1'\ \mathrm{d}\mathbf{x}_2' \\
& - B_i(\mathbf{x}_1) R_{iI} \int \overline{R}(\mathbf{x}_2, \mathbf{x}_2') A_I(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_2' \\
& - B_j(\mathbf{x}_2) R_{jJ} \int \overline{R}(\mathbf{x}_1, \mathbf{x}_1') A_J(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1' \\
& + B_i(\mathbf{x}_1) B_j(\mathbf{x}_2) R_{iI} R_{jJ} \Theta_{IJ} \\
& + \left[B_i(\mathbf{x}_1) R_i^0 - \int \overline{R}(\mathbf{x}_1, \mathbf{x}_1') u^0(\mathbf{x}_1')\ \mathrm{d}\mathbf{x}_1' \right] \left[B_j(\mathbf{x}_1) R_j^0 - \int \overline{R}(\mathbf{x}_2, \mathbf{x}_2') u^0(\mathbf{x}_2')\ \mathrm{d}\mathbf{x}_2' \right] \ .
\end{aligned} \tag{34}$$

3 Inversion

3.1 Sobolev scalar product

Covariance matrix (34) describes the deviation of the resulting model slowness $\widetilde{u}(\mathbf{x})$ from the smoothed exact slowness $\overline{u}(\mathbf{x})$. The deviation causes the differences between the synthetic wavefields calculated in the model and exact wavefields measured in geological structures. To minimize the deviation, the objective function minimizing the squared standard deviation $\widetilde{C}(\mathbf{x}, \mathbf{x})$ may be considered.

However, there is also another source of the differences between the synthetic and exact wavefields: inaccuracy of the wavefield modelling methods in too complex models. The accuracy of all numerical methods increases with an increasing smoothness of the model. The smoothness of the model may be conveniently expressed in terms of the square

$$\|\widetilde{u}\|^2 = ((\widetilde{u}, \widetilde{u})) \tag{35}$$

of the *Sobolev norm* of $\widetilde{u}(\mathbf{x})$. Here

$$\begin{aligned}
((f, g)) = & \sum_{\alpha_1} \sum_{\alpha_2} \sum_{\alpha_3} \sum_{\beta_1} \sum_{\beta_2} \sum_{\beta_3} \int W_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x}) \\
& \times \left[\left(\frac{\partial}{\partial x^1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x^2}\right)^{\alpha_2} \left(\frac{\partial}{\partial x^3}\right)^{\alpha_3} f(\mathbf{x}) \right] \left[\left(\frac{\partial}{\partial x^1}\right)^{\beta_1} \left(\frac{\partial}{\partial x^2}\right)^{\beta_2} \left(\frac{\partial}{\partial x^3}\right)^{\beta_3} g(\mathbf{x}) \right] \mathrm{d}\mathbf{x}
\end{aligned} \tag{36}$$

is the *Sobolev scalar product* of functions $f(\mathbf{x})$ and $g(\mathbf{x})$, defined in terms of the weighting functions $W_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x})$.

As the relative standard deviation $\sqrt{\widetilde{C}(\mathbf{x},\mathbf{x})}$ quantify the local travel–time error per a small distance of wave propagation due to the inaccuracy of the model, it is desirable to adjust the weighting functions $W_{\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3}(\mathbf{x})$ in such a way that the Sobolev norm $\|\widetilde{u}\|$ approximates the analogous local travel–time errors due to inaccuracies of the numerical methods expected to be applied in the resulting model.

In order to improve the efficiency and accuracy of numerical forward–modelling algorithms to be applied in the resulting model, it may be thus desirable to constrain the spatial slowness variations by means of adding squared Sobolev norm $\|\widetilde{u}\|^2$, quantifying the roughness of the resulting slowness distribution $\widetilde{u}(\mathbf{x})$, to the objective function y minimizing the squared standard deviation $\widetilde{C}(\mathbf{x},\mathbf{x})$ integrated over the model volume.

Unfortunately, addition of $(\!(\widetilde{u},\widetilde{u})\!)$ would violate the assumption (15) of the linear dependence of the model on travel times. Since the squared standard deviation $\widetilde{C}(\mathbf{x},\mathbf{x})$ between the medium and its model is minimized instead of the difference corresponding to a particular representation, we should treat the Sobolev norm consistently and to minimize the mean value $\langle(\!(\widetilde{u},\widetilde{u})\!)\rangle$ instead of $(\!(\widetilde{u},\widetilde{u})\!)$.

3.2 The objective function

The objective function is thus assumed in the form of

$$y = \tfrac{1}{2}\int \widetilde{C}(\mathbf{x},\mathbf{x})\ \mathrm{d}\mathbf{x} + \tfrac{1}{2}\langle(\!(\widetilde{u},\widetilde{u})\!)\rangle \quad . \tag{37}$$

Inserting (34) and (18) with (22) into (37), we arrive at

$$\begin{aligned} y = {} & \tfrac{1}{2}\iiint \overline{R}(\mathbf{x},\mathbf{x}_1)\overline{R}(\mathbf{x},\mathbf{x}_2)C(\mathbf{x}_1,\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x} \\ & - R_{iI}\iint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}_2)A_I(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 + \tfrac{1}{2}B_{ij}R_{iI}R_{jJ}\Theta_{IJ} \\ & + \tfrac{1}{2}B_{ij}R_i^0R_j^0 - R_i^0\iint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}_2)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \\ & + \tfrac{1}{2}\iiint \overline{R}(\mathbf{x},\mathbf{x}_1)\overline{R}(\mathbf{x},\mathbf{x}_2)u^0(\mathbf{x}_1)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x} \\ & + \tfrac{1}{2}(\!(B_i,B_j)\!)\langle[R_i^0 + R_{iI}\ \ (T_I - T_I^0)][R_j^0 + R_{jJ}\ \ (T_J - T_J^0)]\rangle \ , \end{aligned} \tag{38}$$

where

$$\boxed{B_{ij} = \int B_i(\mathbf{x})B_j(\mathbf{x})\ \mathrm{d}\mathbf{x}} \tag{39}$$

is the L_2 scalar product of the basis functions, and $(\!(B_i,B_j)\!)$ is the Sobolev scalar product of basis functions $B_i(\mathbf{x})$ and $B_j(\mathbf{x})$, quantifying the roughness of the resulting slowness distribution. Note that a zero Sobolev scalar product may also be considered in our approach to the travel–time tomography.

Considering (14) and (33), equation (38) reads

$$\begin{aligned} y = {} & \tfrac{1}{2}\iiint \overline{R}(\mathbf{x},\mathbf{x}_1)\overline{R}(\mathbf{x},\mathbf{x}_2)C(\mathbf{x}_1,\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x} - A_{iI}R_{iI} \\ & + \tfrac{1}{2}\left[B_{ij} + (\!(B_i,B_j)\!)\right]\left[R_{iI}R_{jJ}\Theta_{IJ} + R_i^0R_j^0\right] - A_i^0R_i^0 \\ & + \tfrac{1}{2}\iiint \overline{R}(\mathbf{x},\mathbf{x}_1)\overline{R}(\mathbf{x},\mathbf{x}_2)u^0(\mathbf{x}_1)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2\ \mathrm{d}\mathbf{x} \ , \end{aligned} \tag{40}$$

where

$$\boxed{A_i^0 = \iint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}_2)u^0(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \qquad (41)}$$

are the Lebesgue scalar products of the basis functions with the smoothed mean slowness, and

$$\boxed{\begin{aligned} A_{iJ} &= \iint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}_2)A_J(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \\ &= \iiint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}'_1)C(\mathbf{x}'_1,\mathbf{x}_2)S_J(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}'_1\ \mathrm{d}\mathbf{x}_2 \qquad (42)\end{aligned}}$$

are the Lebesgue scalar products of the basis functions with the smoothed ray resolution functions (i.e., with the covariances between the smoothed slowness and travel times).

3.3 The minimum of the objective function with respect to the resolution function

Differentiating (40) with respect to R_i^0, we obtain the minimum objective function for

$$[B_{ij} + ((B_i, B_j))]\,R_j^0 - A_i^0 = 0 \quad , \qquad (43)$$

i.e., for

$$\boxed{R_i^0 = [B_{ij} + ((B_i, B_j))]^{-1}\,A_j^0 \quad , \qquad (44)}$$

where $[B_{ji} + ((B_j B_i))]^{-1}$ are the components of the matrix inverse to $[B_{ij} + ((B_i B_j))]$.

Differentiating (40) under constraints (25) with respect to R_{iI}, we see that the minimum objective function is reached for

$$[B_{ij} + ((B_i, B_j))]\,R_{jJ}\Theta_{JI} - A_{iI} = \Lambda_i^\alpha T_I^\alpha \quad , \qquad (45)$$

where Λ_i^α are the Lagrange multipliers.

Without constraints (25), the right–hand side of (45) is zero and the coefficients of the resolution functions have the form of

$$\boxed{R_{jJ} = [B_{ji} + ((B_j B_i))]^{-1}\,A_{iI}\Theta_{IJ}^{-1} \quad , \qquad (46)}$$

where Θ_{NM}^{-1} are the components of the matrix inverse to the travel–time covariance matrix Θ_{MN}^{-1}. The model coefficients are then dependent on travel times through equation (22).

3.4 Constraints adjusting the mean slowness

Let us now discuss the meaning of constraints (25). Multiplying (45) by $\Theta^{-1}_{IJ}[T_I - T^0_I]$ and $\Theta^{-1}_{IJ}T^\beta_J$, respectively, and taking into account (22) and (25), we arrive at

$$[B_{ij} + ((B_i, B_j))] \left[u_j - R^0_j\right] - A_{iI}\Theta^{-1}_{IJ}[T_J - T^0_J] = \Lambda^\alpha_i T^\alpha_I \Theta^{-1}_{IJ}[T_J - T^0_J] \ , \tag{47}$$

and

$$[B_{ij} + ((B_i, B_j))]\, u^\beta_j - A_{iI}\Theta^{-1}_{IJ}T^\beta_J = \Lambda^\alpha_i T^\alpha_I \Theta^{-1}_{IJ}T^\beta_J \quad . \tag{48}$$

Multiplying (48) by the "least–square" coefficients

$$\boxed{U^\beta = \left[T^\beta_M \Theta^{-1}_{MN} T^\alpha_N\right]^{-1} \ T^\alpha_P \Theta^{-1}_{PR}[T_R - T^0_R]} \tag{49}$$

and subtracting from (47), we arrive at

$$[B_{ij} + ((B_i, B_j))] \left[u_j - R^0_j - u^\alpha_j U^\alpha\right] - A_{iI}\Theta^{-1}_{IJ} \left[T_J - T^0_J - T^\alpha_J U^\alpha\right] = 0 \quad , \tag{50}$$

where the updated reference travel times $[T^0_J + T^\alpha_J U^\alpha]$ coincide with the reference travel times (13) calculated with respect to the updated mean medium

$$\boxed{U^0(\mathbf{x}) = u^0(\mathbf{x}) + u^\alpha(\mathbf{x})U^\alpha \quad .} \tag{51}$$

Comparison of equation (50) with equation (47) shows that the application of constraints (17) during the inversion is equivalent to adjusting the a priori mean slowness $u^0(\mathbf{x})$ through a linear combination (51) of functions $u^\alpha(\mathbf{x})$, fitted by the least squares according to equation (49) prior to the inversion, and then performing inversion (22) with the updated mean slowness and resolution functions (46). The least–square fitting (49) should be very strongly overdetermined. We may consider this procedure as a part of the determination of the statistical moments $u^0(\mathbf{x})$ and $C(\mathbf{x}, \mathbf{x}')$ of the medium and assume no constraints (17) during the travel–time inversion.

For example, imagine that we have no idea about the mean slowness and thus choose $u^0(\mathbf{x}) = 0$, but we wish the inversion to be exact in a homogeneous space. We may thus introduce one constraint (17) with $u^1(\mathbf{x}) = 1$. Inversion with such a constraint is equivalent to the determination of $u^0(\mathbf{x}) = \mathit{constant}$ by means of the least squares (49) and its usage in the inversion without constraints.

3.5 Discussion

Let us now decompose equation (22) with coefficients (44) and (46) into several steps to get a better insight into the proposed travel–time tomographic inversion.

The travel time differences with respect to the reference travel times in mean slowness are back–projected onto the model by the operator

$$\left[T_I - T_I^0\right] \rightarrow f(\bullet) = u^0(\bullet) + r_J(\bullet)\left[T_J - T_J^0\right] \quad , \tag{52}$$

where

$$r_J(\mathbf{x}) = A_K(\mathbf{x})\ \Theta_{KJ}^{-1} \tag{53}$$

with ray resolution functions $A_K(\mathbf{x})$ given by (31).

The resulting back–projection of the travel–time differences is then smoothed with the smoothing function,

$$f(\bullet) \rightarrow g(\bullet) = \int \overline{R}(\bullet, \mathbf{x})\, f(\mathbf{x})\ \mathrm{d}\mathbf{x} \quad , \tag{54}$$

and its projection onto the basis functions is smoothed with the Sobolev scalar product,

$$g(\bullet) \rightarrow \widetilde{u}(\bullet) = B_i(\bullet)\ \left[B_{ij} + ((B_i, B_j))\right]^{-1} \int B_j(\mathbf{x})\, g(\mathbf{x})\ \mathrm{d}\mathbf{x} \quad . \tag{55}$$

Step (52) is independent of the a priori information contained within the smoothing function $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$, model parametrization, and smoothing expressed in terms of the Sobolev scalar product $((u_1, u_2))$. These subjective a priori parts of information are step–by–step applied a posteriori as expressed in equations (54) and (55).

Equation (55) shows that the a priori additional smoothing term $\frac{1}{2}\langle((\widetilde{u}, \widetilde{u}))\rangle$ to the objective function (37) acts just as a convolutional low–pass filter applied a posteriori to the result of the tomographic travel–time inversion performed with $((B_i, B_j)) = 0$. It is thus possible to perform the inversion without the Sobolev scalar product, i.e. with $((B_i, B_j)) = 0$. A posteriori smoothing is then on the discretion of a seismologist.

4 Special cases

4.1 Special choice of the smoothing function

Let us discuss here the very important special case of the smoothing function $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$ coinciding with the integral kernel of the projection operator onto the model subspace generated by the basis functions,

$$\overline{R}(\mathbf{x}_1, \mathbf{x}_2) = B_i(\mathbf{x}_1) B_{ij}^{-1} B_j(\mathbf{x}_1) \quad . \tag{56}$$

Equation (42) then simplifies to

$$A_{iJ} = \iint B_i(\mathbf{x}_1) C(\mathbf{x}_1, \mathbf{x}_2) S_J(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \quad . \tag{57}$$

Note that the same is in force if $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$ takes the form of the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$ or, more generally, if $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$ coincides with the Dirac distribution on the model subspace.

In a self–affine random medium, the squared standard deviation $\widetilde{C}(\mathbf{x}, \mathbf{x})$ of the resulting slowness $\widetilde{u}(\mathbf{x})$ may be very large for very narrow smoothing function $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$, and may go to infinity for $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$ approaching the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$. On the other hand, for the smoothing function in the form of (56), the model covariance matrix (34) takes the form of

$$\widetilde{C}(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)\widetilde{C}_{ij}B_j(\mathbf{x}_1) \tag{58}$$

with

$$\begin{aligned}\widetilde{C}_{ij} = C_{ij} - R_{iK}B^{-1}_{jk}A_{kK} - B^{-1}_{ik}A_{kK}R_{jK} + R_{iI}\Theta_{IJ}R_{jJ} \\ +B^{-1}_{ik}((B_k,B_m))R^0_m\ B^{-1}_{jl}((B_l,B_n))R^0_n \quad ,\end{aligned} \tag{59}$$

where

$$C_{ij} = B^{-1}_{ik}B^{-1}_{jl}\iint B_k(\mathbf{x}_1)C(\mathbf{x}_1,\mathbf{x}_2)B_l(\mathbf{x}_2)\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \tag{60}$$

are the components of the projection

$$C'(\mathbf{x}_1,\mathbf{x}_2) = B_i(\mathbf{x}_1)C_{ij}B_j(\mathbf{x}_2) \tag{61}$$

of the medium correlation function $C(\mathbf{x}_1,\mathbf{x}_2)$ onto the model space generated by the basis functions.

4.2 Special choice of the ray characteristic functions

Assume here that the ray characteristic functions $S_I(\mathbf{x})$ are replaced by their projections $S'_I(\mathbf{x})$ onto the the model subspace generated by basis functions $B_i(\mathbf{x})$,

$$S'_J(\mathbf{x}) = B_i(\mathbf{x})B^{-1}_{ik}\int B_k(\mathbf{x})S_J(\mathbf{x})\ \mathrm{d}\mathbf{x} = B_i(\mathbf{x})B^{-1}_{ik}G_{kJ} \quad . \tag{62}$$

Such a simplification may be justified, e.g., if the model parametrization is as dense as the "thick rays" (62) are mostly situated within the Fresnel volumes corresponding to frequencies significant for travel–time picking.

Integral (32) may then be expressed in terms of integrals (24) and (60),

$$S_{IJ} = G_{iI}C_{ij}G_{jJ} \quad , \tag{63}$$

and integration along rays in integral (42) may be replaced with the projection onto the basis functions,

$$A_{iJ} = \iiint B_i(\mathbf{x}_1)\overline{R}(\mathbf{x}_1,\mathbf{x}'_1)C(\mathbf{x}'_1,\mathbf{x}_2)B_l(\mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}'_1\mathrm{d}\mathbf{x}_2\ \ B^{-1}_{lj}\ G_{jJ} \quad . \tag{64}$$

4.3 Special choice of the smoothing function and ray characteristic functions

In a special case of both options (56) and (62), equation (63) may be used to evaluate integral (32). In addition, integral (42) takes a simple form of

$$A_{iJ} = B_{ik}C_{kj}G_{jJ} \quad , \tag{65}$$

compared with (57) and (64). In consequence, equation (59) simplifies to

$$\begin{aligned}\widetilde{C}_{ij} = [\delta_{ik} - R_{iI}G_{kI}]\,C_{kl}\,[\delta_{jl} - R_{jJ}G_{lJ}] + R_{iI}T_{IJ}R_{jJ} \\ +B^{-1}_{ik}((B_k,B_m))R^0_m\ B^{-1}_{jl}((B_l,B_n))R^0_n \quad .\end{aligned} \tag{66}$$

5 Some consequences

5.1 Residual travel–time covariance matrix

The covariance matrix of the travel–time residuals is

$$\widetilde{T}_{IJ} = \langle(\widetilde{T}_I - T_I)(\widetilde{T}_J - T_J)\rangle \quad , \tag{67}$$

where

$$\widetilde{T}_I = \int S_I(\mathbf{x})\widetilde{u}(\mathbf{x})\ \mathrm{d}\mathbf{x} \tag{68}$$

and T_I are given by (4). Inserting (4) and (68), equation (67) reads

$$\begin{aligned}\widetilde{T}_{IJ} = &\iint S_I(\mathbf{x}_1)S_J(\mathbf{x}_1)\langle[\widetilde{u}(\mathbf{x}_1) - u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - u(\mathbf{x}_2)]\rangle\ \mathrm{d}\mathbf{x}_1\ \mathrm{d}\mathbf{x}_2 \\ &- \int S_I(\mathbf{x})\langle\widetilde{u}(\mathbf{x})\delta T_J\rangle\ \mathrm{d}\mathbf{x} - \int\langle\delta T_I\widetilde{u}(\mathbf{x})\rangle S_J(\mathbf{x})\ \mathrm{d}\mathbf{x}\ + \langle\delta T_I\ \delta T_J\rangle \quad .\end{aligned} \tag{69}$$

Here $\langle[\widetilde{u}(\mathbf{x}_1)-u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2)-u(\mathbf{x}_2)]\rangle$ coincides with (27) for $\overline{u}(\mathbf{x}) = u(\mathbf{x})$, i.e. for smoothing function $\overline{R}(\mathbf{x}_1, \mathbf{x}_2)$ in the form of the Dirac distribution $\delta(\mathbf{x}_1 - \mathbf{x}_2)$ (no smoothing). Inserting (34) with $\overline{R}(\mathbf{x}_1, \mathbf{x}_2) = \delta(\mathbf{x}_1 - \mathbf{x}_2)$ for $\langle[\widetilde{u}(\mathbf{x}_1) - u(\mathbf{x}_1)][\widetilde{u}(\mathbf{x}_2) - u(\mathbf{x}_2)]\rangle$, and

$$\langle\widetilde{u}(\mathbf{x})\delta T_J\rangle = B_i(\mathbf{x})R_{iK}\langle\delta T_K\ \delta T_J\rangle = B_i(\mathbf{x})R_{iK}T_{KJ} \tag{70}$$

into (69), we arrive at

$$\begin{aligned}\widetilde{T}_{IJ} = S_{IJ} - G_{iI}R_{iK}S_{KJ} - S_{IL}R_{jL}G_{jJ} + G_{iI}R_{iK}\Theta_{KL}R_{jL}G_{jJ} \\ +[R^0_iG_{iI} - T^0_I][R^0_jG_{jJ} - T^0_J] - G_{iI}R_{iK}T_{KJ} - T_{IK}R_{jK}G_{jJ} + T_{IJ}\ .\end{aligned} \tag{71}$$

Thus the statistical expectation of the travel–time residuals is described by the covariance matrix

$$\widetilde{T}_{IJ} = [\delta_{IK} - G_{iI}R_{iK}]\,\Theta_{KL}\,[\delta_{LJ} - R_{jL}G_{jJ}] + [R^0_iG_{iI} - T^0_I][R^0_jG_{jJ} - T^0_J]\ , \tag{72}$$

where Θ_{KL} is given by (33).

6 Conclusions

In this paper we have proposed the approach to the seismic travel–time tomography enabling to estimate the accuracy and resolution of the resulting (macro) model of a geological structure.

Presented approach already enables to consider such a priori information like the obviously much greater probability of slow materials close to the Earth surface than deeper in the Earth, or greater probability of slow materials in the low and concave parts of the Earth periphery than in the high–elevated and convex parts.

On the other hand, we have still made various simplifications. We have considered only smooth models without structural interfaces, and have restricted the study to a single material parameter — slowness. However, we believe that this approach to the seismic travel–time tomography may further be generalized.

Acknowledgements

The author is indebted to Sergei Shapiro for fruitful discussions essential for this study.

References

Klimeš, L. (1996): Correlation function of a self–affine random medium. In: Seismic Waves in Complex 3-D Structures, Report 4, pp. 25–38, Dep. Geophys., Charles Univ., Prague.

Klimeš, L. (1997): Correlation functions of random media In: Seismic Waves in Complex 3-D Structures, Report 6, pp. 25–40, Dep. Geophys., Charles Univ., Prague.