# Bent-ray traveltime tomography and migration without ray tracing

P. E. Ecoublet, S. C. Singh, C. H. Chapman, G. M. Jackson

*Geophys. J. Int.* (2002) **149,** 633–645

# Bent-ray traveltime tomography and migration without ray tracing

P. E. Ecoublet,[2] S. C. Singh,[1,2] C. H. Chapman[3] and G. M. Jackson[4]

[1] *Laboratoire de Géosciences Marines, Institut de Physique du Globe de Paris, 4 Place Jussieu,* 75252 *Paris Cedex 5, France. E-mail: singh@ipgp.jussieu.fr*
[2] *Bullard Laboratories, Department of Earth Sciences, University of Cambridge, Madingley Road, Cambridge* CB3 0EZ, *UK*
[3] *Schlumberger Cambridge Research, High Cross, Madingley Road, Cambridge* CB3 0EL, *UK*
[4] *Geoscience Research Centre, Elf UK, London, UK*

Accepted 2001 November 21. Received 2001 July 23; in original form 2000 August 2

## SUMMARY

We present a new method of traveltime tomography. In this method, the traveltime between source and receiver is described by an analytical function, which consists of a series expansion of geometrical coordinates of the source and receiver locations. As the traveltime is derived from the eikonal equation, the analytical function must also satisfy the eikonal equation. This condition imposes a strong constraint on the uniqueness of the analytical function. The coefficients of the series expansion are estimated by minimizing the misfit between the observed and the analytical time function in a least-squares sense. Once the coefficients of the series expansion are known, the eikonal equation, which turns out to also be in the form of a series expansion, provides the velocity in the medium. Thus there are two analytical functions, one defining the traveltime and the other defining the slowness, and they can be used for pre-stack depth migration and velocity model definition. The feasibility of this approach is first tested on a synthetic data set and then applied to a real data set. This new method of tomography and pre-stack migration provides a significant gain in computation time compared with ray-tracing techniques. The method can easily be extended to incorporate reflection data and has potential for solving 3-D seismic reflection and global seismology inverse problems.

**Key words:** bent rays, inversion, migration, ray tracing, tomography.



## INTRODUCTION

The goal of seismic tomography is to determine the large-scale velocity structure of the subsurface of the Earth using traveltime measurements made on the surface. The term tomography was first introduced in medical imaging to describe image reconstruction from an X-ray line integral (Tanabe 1971; Mersereau 1976; Kak 1979; Louis & Natterer 1983). Dines & Lytle (1979) used the term computerized geophysical tomography in analogy with X-ray tomography to determine the velocity distribution between two boreholes from first-arrival traveltimes and assuming a straight ray path between sources and receivers. The straight ray approximation enables a fast implementation of inversion methods, similar to those used in X-ray tomography (Gordon 1974; Censor 1983), but it does not account for the ray path dependence on velocity in the medium, which leads to a non-linear relationship between the data and the model parameters. The non-linearity can be taken into account to some extent by using an iterative linearized inversion method. A common approach consists of starting from an initial velocity model and iteratively updating this model in such a way that the traveltime data fit the synthetic traveltimes. A review of traveltime tomographic methods is given by Worthington (1984) and a comparison of various inversion techniques is discussed by Phillips & Fehler (1991).

Ray-based tomography has been applied to various problems (Bording *et al.* 1987), but a number of problems remain for recovering an accurate and unique image of the Earth from traveltime tomography. Unlike medical imaging, where the measurements are made from all directions of the body that is to be imaged, seismic experiments illuminate the investigated medium from a maximum of three sides providing non-uniform and incomplete data coverage. Consequently, fewer independent traveltime data than unknown model parameters are available, leading to a non-unique solution, i.e. several different velocity models fit the data (Carrion 1991). The non-uniqueness of the solution may be constrained by using prior information on the model parameters (Jackson 1979). Berryman (1990) used Fermat's principle to constrain the solution. Nevertheless, the computation of traveltime, either using ray-tracing techniques or solving the eikonal equation with a numerical method (Vidale 1990; Moser 1991), constitutes a time-consuming process in traveltime tomography, particularly in three dimensions.

An alternative approach to ray-tracing tomography and numerical methods was introduced by Bates *et al.* (1991) and was applied to 1-D tomographic inversion by Enright *et al.* (1992). This method deals not with rays but with a functional description of the traveltimes that must satisfy theoretical relationships derived from Fermat's principle. Which means that both non-linearity and non-uniqueness are addressed implicitly by this method. In this

paper, we extend this approach to solve a 2-D traveltime tomography problem. In this new method, the traveltime between two points is defined by an analytical function in the form of a series expansion of source and receiver locations. The unknown coefficients of the series expansion are estimated by minimizing the difference between the observed traveltimes and the traveltimes calculated with the analytical function. After estimation of the unknown coefficient, the eikonal equation, derived from the traveltime function, is used to compute the inverse of the square of the velocity. Furthermore, the analytical traveltime function can be used to compute traveltimes between two points for implementing pre-stack depth migration. In the next few sections, we present the theory behind the new method, and then test the feasibility of the method on a synthetic example before applying it to a real data set.

## THE TRAVELTIME FUNCTION

The seismic traveltime between a source $S$ and a receiver $R$, denoted by $T(R, S)$, is related to the slowness $s(P)$ at the point $P$ along the ray path by the integral

$$T(R, S) = \int_{L(s)} s(P)\, dl, \tag{1}$$

where $L(s)$ is the ray path between the source and the receiver, and $dl$ is the arc length along the ray. The traveltime is a solution of the eikonal equation given by

$$\nabla_P T(P, S) \cdot \nabla_P T(P, S) = s^2(P), \tag{2}$$

where $\nabla_P$ is the gradient operator at $P$.

The traveltime $T(R, S)$ can either be computed using a ray-tracing method from eq. (1) or a numerical method from eq. (2) (e.g. Vidale 1990). Here, we assume that $T(R, S)$ can be defined by an analytical function. However, this function must satisfy a number of invariant properties. For example, the first condition requires the traveltime function to vanish for a common source and receiver location $S$, i.e.

$$T(S, S) = 0. \tag{3}$$

The reciprocity condition states that the traveltime is invariant when the source and receiver locations are interchanged,

$$T(R, S) = T(S, R). \tag{4}$$

The eikonal equation defined by eq. (2) must remain valid for a common source and receiver point $P$, i.e.

$$\nabla_P T(P, P) \cdot \nabla_P T(P, P) = s^2(P), \tag{5}$$

where $\nabla_P T(P, P) = \nabla_P T(P, S)|_{S=P}$.

It should be noted that the left-hand side of eq. (2) is a function of two points, the source location $S$ and the observation point $P$, whereas the right-hand side of eq. (2), which contains the slowness at the observation point $P$, is only a function of $P$. Therefore, taking the gradient with respect to $S$ of the eikonal equation (2) yields

$$\nabla_S[\nabla_P T(P, S) \cdot \nabla_P T(P, S)] = 0. \tag{6}$$

Eq. (6) imposes a uniqueness constraint on the value of slowness at $P$, which is independent of any source point $S$ and receiver locations. For a common source–receiver point $P$, eq. (6) gives a boundary condition of the form

$$\nabla_S[\nabla_P T(P, P) \cdot \nabla_P T(P, P)] = 0. \tag{7}$$

Eqs (6) and (7) are valid for any point $P$ and $S$, and provide a strong constraint on the analytical function.

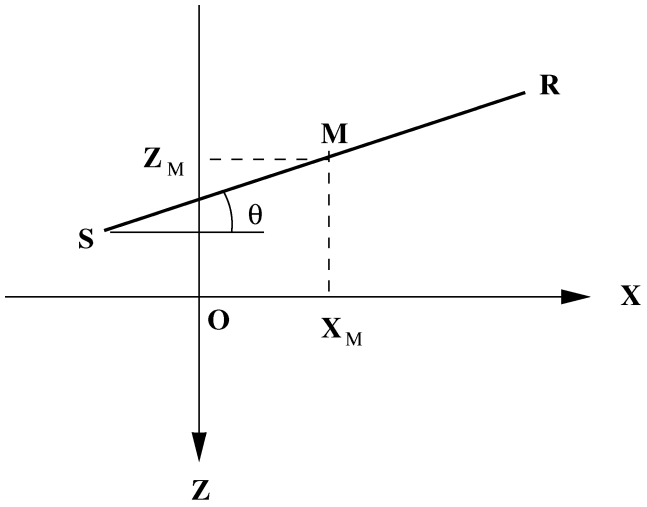


**Figure 1.** Geometrical parameters for 2-D tomography. The source–receiver pair $(S, R)$ is uniquely defined by four parameters: $X$ and $Z$ positions of the mid-point $M$, the orientation $\theta$ with respect to the $X$-axis, and the source–receiver distance $d(R, S)$.

A distinction can be made between two sets of equations; those related to data and those independent of data. We call the first set of eqs (3), (4), (6) and (7) hard constraints on the analytical function as they are independent of data, and the second set of eqs (1), (2) and (5) soft constraints as they are dependent of data $T$.

### Traveltime as a series expansion

The formulation of the 2-D problem requires a choice of parametrization that best satisfies the 'hard' constraints. The following parameters define the source $S(X_S, Z_S)$ and receiver $R(X_R, Z_R)$ locations uniquely in the 2-D space (Fig. 1). The source–receiver mid-point, $M(X_M, Z_M)$, can be defined as

$$X_M = \frac{X_S + X_R}{2} \quad \text{and} \quad Z_M = \frac{Z_S + Z_R}{2}, \tag{8}$$

the source–receiver distance $(d)$ is defined as

$$d = \left[(X_R - X_S)^2 + (Z_R - Z_S)^2\right]^{1/2}, \tag{9}$$

and the angle $\theta[-\pi/2, \pi/2]$ as

$$\theta = \arcsin\left(\frac{Z_R - Z_S}{d}\right). \tag{10}$$

The traveltime function can then be written as a series expansion of these four parameters $(X_M, Z_M, d, \theta)$ as

$$T(R, S) = \sum_{l=0}^{L} \sum_{m=0}^{M} \sum_{n=-N}^{N} \sum_{p=0}^{P} C_{lmnp} P_l(X_M) P_m(Z_M) \exp(i2n\theta) d^p, \tag{11}$$

where $C_{lmnp}$ are the unknown complex coefficients of the series expansion, $P_l$ is the $l$th-order Chebyshev polynomial (Arfken 1985), and $L$, $M$, $N$, $P$ are the number of coefficients associated with each term of the series expansion. The term related to the symbol $L$ controls lateral variations of velocity and that to the symbol $M$ controls the smoothness of the velocity variation with depth. Chebyshev polynomials are used in place of a power-series expansion as they are orthogonal in the interval $[-1, +1]$. Any other basis function can

be used to define the traveltime function. The boundary condition on traveltime for a common source and receiver location, i.e. $d = 0$ (eq. 3) yields

$$\sum_{l=0}^{L}\sum_{m=0}^{M}\sum_{n=-N}^{N} C_{lmn0} P_l(X_M) P_m(Z_M) \exp(i2n\theta) = 0, \tag{12}$$

and consequently

$$C_{lmn0} = 0. \tag{13}$$

The condition of traveltime reciprocity (eq. 4) is satisfied explicitly by the traveltime series expansion, since the angle $\theta$ remains unchanged by swapping over the source and receiver locations.

### The eikonal equation

The expression for the eikonal equation at a unique point $P$ (eq. 5) takes the form (Appendix A)

$$\nabla_P T(P, P) \cdot \nabla_P T(P, P) = \left[\sum_{l=0}^{L}\sum_{m=0}^{M}\sum_{n=-N}^{N} C_{lmn1} P_l(X_P) P_m(Z_P) \exp(i2n\theta)\right]^2 = s^2(P). \tag{14}$$

The condition given by the eikonal equation on the invariance of the slowness at $P$ (eq. 7) imposes eq. (14) to be independent of $\theta$. In order to satisfy this condition, the coefficients related to $\theta$ must vanish, i.e.

$$C_{lmn1} = 0, \quad \forall n \neq 0. \tag{15}$$

Incorporating the condition (15) into eq. (14), we obtain

$$\nabla_P T(P, P) \cdot \nabla_P T(P, P) = \left[\sum_{l=0}^{L}\sum_{m=0}^{M} C_{lm01} P_l(X_P) P_m(Z_P)\right]^2 = s^2(P). \tag{16}$$

Since the slowness at any point $P$ is always positive, we can say that

$$s(P) = \sum_{l=0}^{L}\sum_{m=0}^{M} C_{lm01} P_l(X_P) P_m(Z_P). \tag{17}$$

The real value of the slowness is ensured by imposing the following condition on the coefficients $C_{lm01}$:

$$\mathrm{Im}\,(C_{lm01}) = 0 \quad \text{and} \quad \mathrm{Re}\,(C_{lm01}) = R_{lm} \quad \forall l, m. \tag{18}$$

Eq. (17) defines the slowness at any point $P$ in the medium as a 2-D series expansion of its coordinates $(X_P, Z_P)$. If the coefficients $C_{lmnp}$ are known, eq. (17) can be used to reconstruct the slowness in the medium. By incorporating the conditions (13), (15) and (18) into the traveltime eq. (11), the traveltime function can be written as

$$T(R, S) = \sum_{l=0}^{L}\sum_{m=0}^{M} R_{lm} P_l(X_M) P_m(Z_M)\, d + \sum_{l=0}^{L}\sum_{m=0}^{M}\sum_{n=-N}^{N}\sum_{p=2}^{P} C_{lmnp} P_l(X_M) P_m(Z_M) \exp(i2n\theta)\, d^p. \tag{19}$$

The substitution of the slowness eq. (17) into eq. (19) gives the traveltime equation as

$$T(R, S) = s(M)\, d + \sum_{l=0}^{L}\sum_{m=0}^{M}\sum_{n=-N}^{N}\sum_{p=2}^{P} C_{lmnp} P_l(X_M) P_m(Z_M) \exp(i2n\theta)\, d^p. \tag{20}$$

Thus the traveltime between a source $S$ and a receiver $R$ is the sum of the traveltime through a medium at constant slowness, the slowness of the mid-point $s(M)$, plus a power-Fourier series of geometrical coordinates of the source and the receiver. The first term corresponds to a straight ray path approximation, and the second term can be thought as a correction to this approximation incorporating ray bending in a heterogeneous medium.

### The constraint equations

Condition (6), stating the independence of the eikonal equation with respect to the source location, yields two partial derivative equations

$$\frac{\partial T(P,S)}{\partial X_P}\frac{\partial}{\partial X_S}\left[\frac{\partial T(P,S)}{\partial X_P}\right] + \frac{\partial T(P,S)}{\partial Z_P}\frac{\partial}{\partial X_S}\left[\frac{\partial T(P,S)}{\partial Z_P}\right] = 0, \tag{21a}$$

and

$$\frac{\partial T(P,S)}{\partial X_P}\frac{\partial}{\partial Z_S}\left[\frac{\partial T(P,S)}{\partial X_P}\right] + \frac{\partial T(P,S)}{\partial Z_P}\frac{\partial}{\partial Z_S}\left[\frac{\partial T(P,S)}{\partial Z_P}\right] = 0. \tag{21b}$$

The equation of constraint (21) define a feasibility domain, independent of the data, in which the traveltime function is a solution of the eikonal equation and therefore complies with the ray theory. The traveltime function defined by eq. (20) does not satisfy these two partial derivative equations explicitly. However, in order for the unknown coefficients of the series expansion to satisfy the constraints (21a) and (21b), these equations are generated for a chosen set of $S$ and $P$ points, and are solved numerically.

## TOMOGRAPHY

### The misfit function

The unknown coefficients of the series expansion are estimated such that the analytical traveltime function fits the data and satisfies the constraint equations. This is achieved by minimizing a misfit function that accounts for the difference between the observed and the calculated traveltimes as well as the difference between the constraint equations and the *a priori* information, for example, information on the velocity from a sonic log. Such *a priori* information is treated as data. To simplify the notation, we write $\mathbf{m}$ as a vector containing the unknown coefficients of the series expansion, i.e.

$$\mathbf{m} = (C_{lmnp})^{\mathrm{T}}. \tag{22}$$

We assume that both data and modelling errors follow a Gaussian distribution, and use a probabilistic definition of the misfit function, where an *a posteriori* probability density function on the model parameters PDF ($\mathbf{m}$) can be written as

$$\mathrm{PDF}\,(\mathbf{m}) = \text{constant} \times \exp\left[-\frac{1}{2}\left\{[\mathbf{D}_{\mathrm{cal}}(\mathbf{m}) - \mathbf{D}_{\mathrm{obs}}]^{\mathrm{T}}\mathbf{C}_{\mathrm{D}}^{-1} \times [\mathbf{D}_{\mathrm{cal}}(\mathbf{m}) - \mathbf{D}_{\mathrm{obs}}] + \mathbf{fc}(\mathbf{m})^{\mathrm{T}}\mathbf{C}_{\mathbf{fc}}^{-1}\mathbf{fc}(\mathbf{m})\right\}\right]. \tag{23}$$

The vector $\mathbf{D}_{\rm cal}(\mathbf{m})$ contains $N_{\rm t}$ traveltimes $\mathbf{T}_{\rm cal}$ defined by eq. (19), and $N_{\rm s}$ slownesses $\mathbf{s}_{\rm cal}$ defined by eq. (17), where

$$\mathbf{D}_{\rm cal}(\mathbf{m}) = [\mathbf{T}_{\rm cal}(\mathbf{m}), \mathbf{s}_{\rm cal}(\mathbf{m})]^{\rm T}. \tag{24}$$

The data vector $\mathbf{D}_{\rm obs}$ contains the data, $N_{\rm t}$ observed traveltimes $\mathbf{T}_{\rm obs}$, and $N_{\rm s}$ known slowness values $\mathbf{s}_{\rm obs}$, where

$$\mathbf{D}_{\rm obs} = (\mathbf{T}_{\rm obs}, \mathbf{s}_{\rm obs})^{\rm T}. \tag{25}$$

The vector $\mathbf{fc}(\mathbf{m})$ contains $N_{fc}$ constraint equations (21) defining the invariance of the eikonal equation with respect to the source location.

$$\mathbf{fc}(\mathbf{m}) = \{\nabla_S[\nabla_P \mathbf{T}_{\rm cal}(\mathbf{m}) \cdot \nabla_P \mathbf{T}_{\rm cal}(\mathbf{m})]\}^{\rm T}. \tag{26}$$

Maximizing the *a posteriori* probability density function amounts to minimizing the argument of the exponential, that is to say minimizing the misfit function $Q(\mathbf{m})$ defined by

$$Q(\mathbf{m}) = \left\{[\mathbf{D}_{\rm cal}(\mathbf{m}) - \mathbf{D}_{\rm obs}]^{\rm T}\mathbf{C}_{\rm D}^{-1}[\mathbf{D}_{\rm cal}(\mathbf{m}) - \mathbf{D}_{\rm obs}] + \mathbf{fc}(\mathbf{m})^{\rm T}\mathbf{C}_{\mathbf{fc}}^{-1}\mathbf{fc}(\mathbf{m})\right\}. \tag{27}$$

$N_{\rm D}$ is the total number of data points, $N_{\rm D} = N_{\rm t} + N_{\rm s}$, and $\mathbf{C}_{\rm D}$ is a data error covariance matrix. The diagonal terms of $\mathbf{C}_{\rm D}$ contain the data variance, denoted by $\sigma_{\rm t}^2$ and $\sigma_{\rm s}^2$ and can be defined as $C_{\rm D}^{ii} = \sigma_{\rm t}^{i^2}, \forall i = 1, N_{\rm t}$, $C_{\rm D}^{ii} = \sigma_{\rm s}^{i^2}, \forall i = N_{\rm t}+1, N_{\rm s}$, and $C_{\rm D}^{ii} = 0, \forall i \neq j$. $\mathbf{C}_{\mathbf{fc}}$ is a covariance matrix describing the parametrization errors defined as $C_{fc}^{ii} = \sigma_{fc}^{i^2}, \forall i = 1, N_{fc}$ and $C_{fc}^{ii} = 0, \forall i \neq j$.

### The initial model

Although the traveltime series expansion is a linear function of the unknown coefficients, the constraint equations introduce non-linear relationships between the coefficients and the data. Consequently, the misfit function becomes non-quadratic; its minimum cannot be reached in a single step, and an iterative optimization process has to be implemented to find the minimum of the misfit $Q(\mathbf{m})$. The initial model parameters, denoted as $\mathbf{m}_0$ for iteration zero, are set to zero except for the first coefficient $C_{0001}^0$, which is defined as the average slowness $\bar{s}$, i.e.

$$C_{0001}^0 = \frac{1}{N_{\rm t}} \sum_{i=1}^{N_{\rm t}} \frac{T_{\rm obs}^i}{d^i} = \bar{s}, \tag{28a}$$

and

$$C_{lmnp}^0 = 0, \quad \forall (l, m, n, p) \neq (0, 0, 0, 1), \tag{28b}$$

where $d^i$ is the source–receiver distance and $C_{lmnp}^0$ are the initial coefficients of the traveltime series expansion at iteration zero.

### Model parametrization

As in any inversion, it is extremely important to choose appropriate model parameters, which in our case are the values of $L, M, N, P$. There are two approaches to parametrizing models: over parametrization or under parametrization. Over parametrization may lead to an unstable inversion algorithm, but this can be avoided by the use a smoothing criteria (McCaughey & Singh 1997). Under parametrization requires significant human input and hence is very subjective. Here, we use a combination of these two approaches. Depending upon the data coverage, we start with a small number of parameters, and then increase them as inversion proceeds until the algorithm becomes unstable and start diverging. Since the number of parameters is small (only four in the 2-D case) and the algorithm is efficient, we can easily explore optimum values of $L, M, N, P$ using a trial-and-error method.

### Optimization

The minimum of the misfit function is obtained by implementing an iterative Gauss–Newton method (Scales 1985). The inverse of the Hessian matrix is computed using a singular-value decomposition (SVD) (Golub & Reinsch 1970; Press *et al.* 1992). The larger eigenvalues are associated with the dominant features of the model parameters, whereas the smaller eigenvalues are dominated by noise. Numerical instabilities during inversion can be avoided by not including the eigenvalues smaller than a cut-off value, $\lambda_{\rm cut}$. The minimum of the misfit is reached when the gradient of the misfit function vanishes, i.e. when the increment $\Delta\mathbf{m}_N$ tends to zero. The convergence criteria are based on the value of $\Delta\mathbf{m}_N$ compared with a threshold. The final solution, denoted by $\mathbf{m}_*$, is reached when the increment $\Delta\mathbf{m}_N$ becomes smaller than a threshold value that can be determined numerically.

### Slowness image reconstruction

Once the model parameters, i.e. the unknown coefficients of the series expansion, have been determined, the slowness distribution in the investigated media can be easily obtained using eq. (17). Since the Chebyshev polynomials in eq. (17) cannot represent sharp discontinuities, the final slowness model is always smooth. The degree of smoothness will depend on the number of coefficients defining the series expansion.

## ERROR AND RESOLUTION ANALYSES

No inversion is complete without having error and resolution in the final model. Unlike in conventional tomography, where such an analysis is carried on the final velocity model, the error and resolution analyses need to be first performed on the coefficients of the series expansion. Prior to performing the conventional error analysis, which requires the *a posteriori* probability function (ppd) to be a Gaussian function, we first need to prove that the ppd we obtain using our method is indeed a Gaussian. This requires that provided the initial model $\mathbf{m}_0$ is a solution of the constraint equations (i.e. if it is within the feasibility region defined by the theoretical constraints), the *a posteriori* probability density function eq. (23) is a Gaussian, and therefore the algorithm converges towards the maximum-likelihood solution. An initial model defined by the average constant slowness computed from the traveltime data lies within the feasibility domain and therefore ensures convergence toward the global minimum of the misfit function. Eq. (B10) (see Appendix B) confirms that the ppd is a Gaussian function, whereas eqs (B13) and (B14) satisfy the condition required by the initial model.

The result of the stochastic inversion provides statistical information on the estimated model parameters by calculating the *a posteriori* covariance matrix, denoted by $\mathbf{C}_*$, and given by

$$\mathbf{C}_* = \left(\mathbf{G}^{\rm T}\mathbf{C}_{\rm D}^{-1}\mathbf{G} + \mathbf{F}_0^{\rm T}\mathbf{C}_{\mathbf{fc}}^{-1}\mathbf{F}_0\right)^{-1} = 2\mathbf{H}_*^{-1}. \tag{29}$$

From the *a posteriori* covariance matrix, of which a detailed calculation is given in Appendix B, uncertainties of slowness reconstruction and traveltime calculation can be evaluated.

## PRE-STACK DEPTH MIGRATION

The traveltime series expansion resulting from the tomographic inversion can be used to compute of traveltimes and ray path between two points. We propose to use this result for implementing a diffraction pre-stack migration without ray tracing. Kirchoff migration (Yilmaz 1987) requires intensive traveltime computation of the diffracted wave between each subsurface scattering point and each source and receiver. This operation, usually achieved by shooting rays through the velocity model (Schneider 1978; Berryhill 1979), constitutes the most expensive step of the Kirchoff migration. However, a two-point traveltime computation can be accurately and rapidly performed using the traveltime series expansion resulting from the tomographic inversion. The smooth and continuous slowness model, defined by the slowness function, is particularly appropriate to perform a pre-stack migration of seismic data. The computation time required to compute traveltimes with the series expansion depends on the number of coefficients composing the series expansion. Another advantage of our method is that traveltimes can be computed at the time of migration, hence reducing the memory requirement of input/output (I/O) operations.

### Computation of the incidence angle of the ray

Limited-aperture data produces migration smiles that spread the image of the actual scatterers. These artefacts can be reduced by performing a limited aperture migration that requires computation of the incidence angle of the ray on arriving and on leaving each scattering point. Rays for which the incidence angle at the scattering point exceed the extreme values controlled by the bounded slope are rejected.

Let us consider a ray emanating from a source $S$ to a point $P$ (Fig. 2). The gradient operator applied to the traveltime $T(P, S)$ gives the ray parameter at $P$, denoted by $\mathbf{p}$, such as

$$\nabla_P T(P, S) = \mathbf{p} = s(P)\frac{\mathbf{dr}}{ds}, \tag{30}$$

where $s(P)$ is the slowness at $P$, $\mathbf{dr}$ is the tangent along the ray with length $ds$. If $dx$ and $dz$ are the projections of $\mathbf{dr}$ along the $X$ and $Z$ axes, $\mathbf{dr} = (dx, dz)^{\mathrm{T}}$, the incidence angle $\theta_z$, defined as the angle between $\mathbf{dr}$ and the $Z$-axis such as $dz = ds\cos(\theta_z)$, is given by

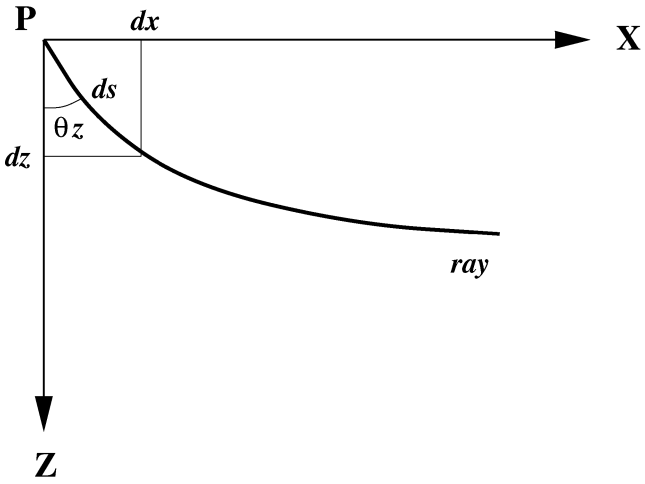


**Figure 2.** The incidence angle of the ray at $P$, $\theta z$, defined as the angle between the ray and the vertical axis, is computed from the ray parameter $\mathbf{p}$ and the slowness $s(P)$.

$$\theta_z = \arccos\left[\frac{1}{s(P)}\frac{\partial T(P, S)}{\partial Z_P}\right]. \tag{31}$$

The directions of the ray, on arriving and on leaving each subsurface point, are computed from the traveltime function and so that unwanted rays can be eliminated when summing the diffracted energy. Incorporating aperture limitation aims to reduce artefacts and thus to improve the result of the migration.

## SYNTHETIC DATA ANALYSIS

A synthetic experiment was designed to investigate the feasibility of our method. We consider a vertical seismic profile (VSP) configuration because a real data set was available to test our algorithm. The VSP experiment (Fig. 3) consists of three sources located on the surface at −2.0, 0.0 and 2.0 km. Receivers are arranged along a borehole centred at 0.0 km that splits into a straight well and a bent well around 3.2 km depth. The velocity model is defined by eight layers, each bounded by planar interfaces. The top layer is a 150 m thick water layer with a velocity of 1.5 km $s^{-1}$. The second layer, between 150 and 2000 m depth, is defined by a linearly increasing velocity with depth from 1.8 to 2.5 km $s^{-1}$. The next six layers are defined by a constant velocity in each layer and bounded by horizontal interfaces except for the seventh layer which has a dip of 7°. The synthetic data set constitutes of 284 traveltimes generated by tracing rays through the velocity model. Out of 284 traveltimes, 102 traveltimes were computed for the source at −2.0 km and receivers located in the well between 2857 and 3907 m depth, 82 traveltimes for the source at 0.0 km and receivers from 2248 m down to a depth of 3734 m, and 102 traveltimes for the source at 2.0 km and receivers from 2897 m down to 3904 m depth. The ray coverage, marked by straight white lines, clearly shows the narrow ray coverage.

We have assumed that the horizontal variation in velocity can be sufficiently described by two parameters ($L = 2$) allowing for a linear velocity variation. The variation of velocity with depth is defined by eight parameters ($M = 8$), i.e. by a seventh-order series expansion. Three parameters describe the angular dependence of the traveltime ($N = 3$), and six parameters describe the traveltime as a function of the source–receiver offset ($P = 6$). Altogether, 496 real coefficients define the traveltime series expansion eq. (20). As the velocity in the water layer is assumed to be known, the slowness function is computed at 102 different points regularly spaced between the surface and 150 m depth. Note that the slowness function in eq. (17) defined by a subset of 16 coefficients. The initial average slowness computed from the observed traveltimes is 0.412 s $km^{-1}$ (average velocity of 2.427 km $s^{-1}$), and the initial root mean square (rms) error between the observed and calculated traveltimes is 50.368 ms.

The convergence was reached after five iterations. The traveltime misfit was reduced to 2.382 ms and the constraints misfit to 0.018. Although the slowness function follows increasing velocity down to 2.0 km depth, it does not match the velocity variations below 2.0 km where additional constraints need to be applied on the velocity model. We assume that the well log is available and use the velocity profile along the borehole as *a priori* information to constrain the inversion. 43 velocities were picked at regular spacing along the log and included in the slowness term of the misfit function. The convergence was achieved after five iterations and the resulting velocity tomogram is shown in Fig. 4. The traveltime misfit was reduced to 2.362 ms and the constraints misfit to 0.006. The slowness misfit on the *a priori* information, defined as the rms error between the calculated and the observed slownesses along the

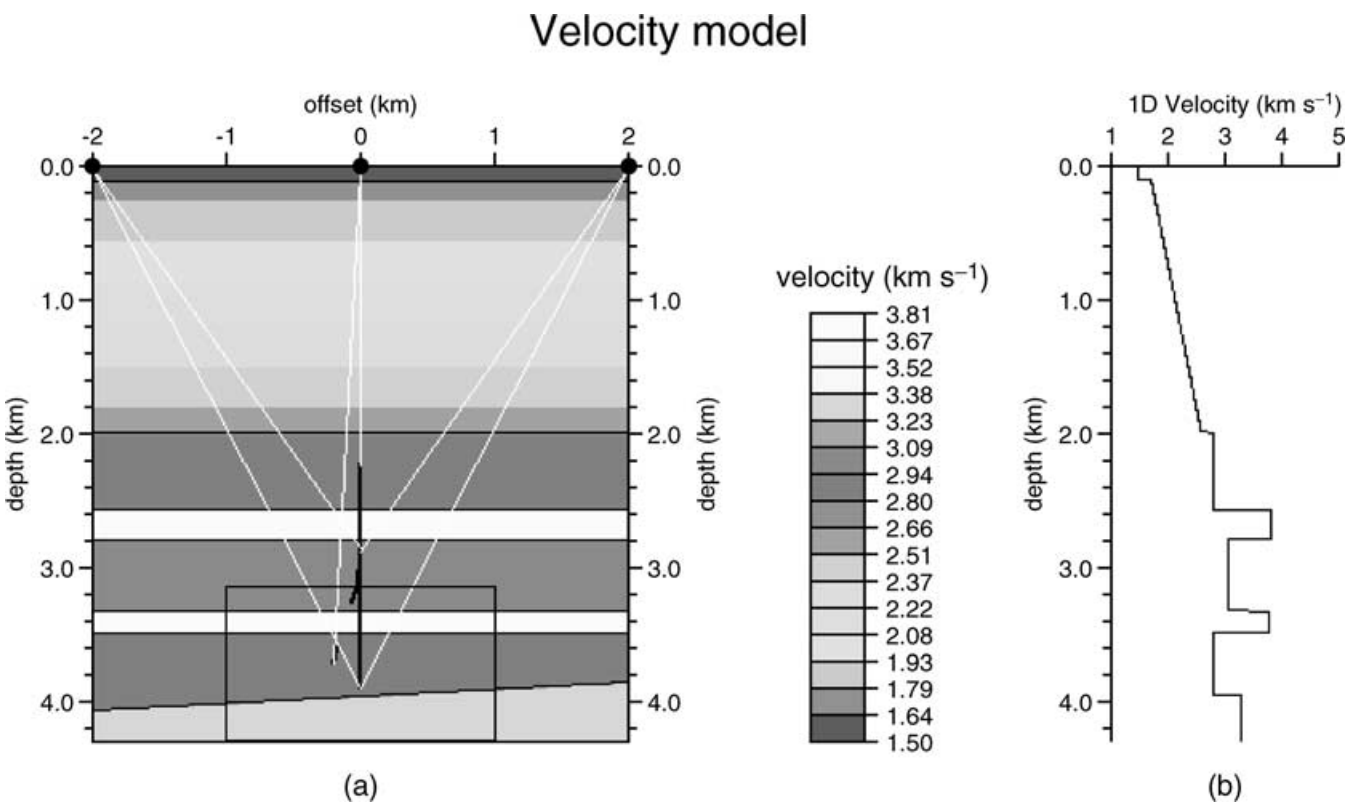


**Figure 3.** Synthetic velocity model and recording geometry. The velocity model is decomposed into eight layers. The velocity is constant in each layer except in the second layer in which the velocity increases linearly with depth. The velocity log at 0.0 km is plotted on the right of the figure. The acquisition geometry consists of three sources on the surface (black dots) and receivers in a borehole located at 0.0 km. The area between the white lines shows the ray coverage.

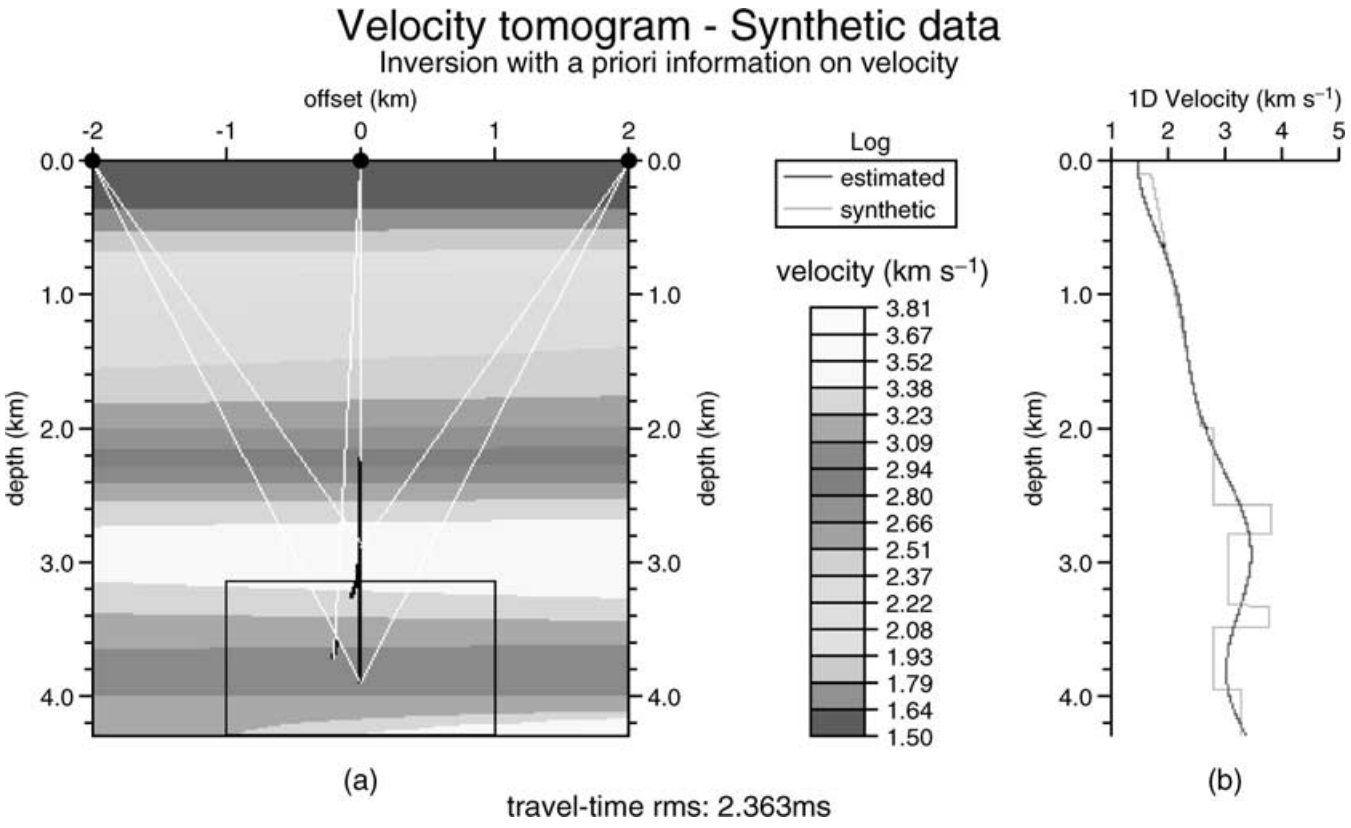


**Figure 4.** Velocity tomogram estimated with the *a priori* information on the velocity distribution. The smooth velocity tomogram matches more closely the true velocity variations.

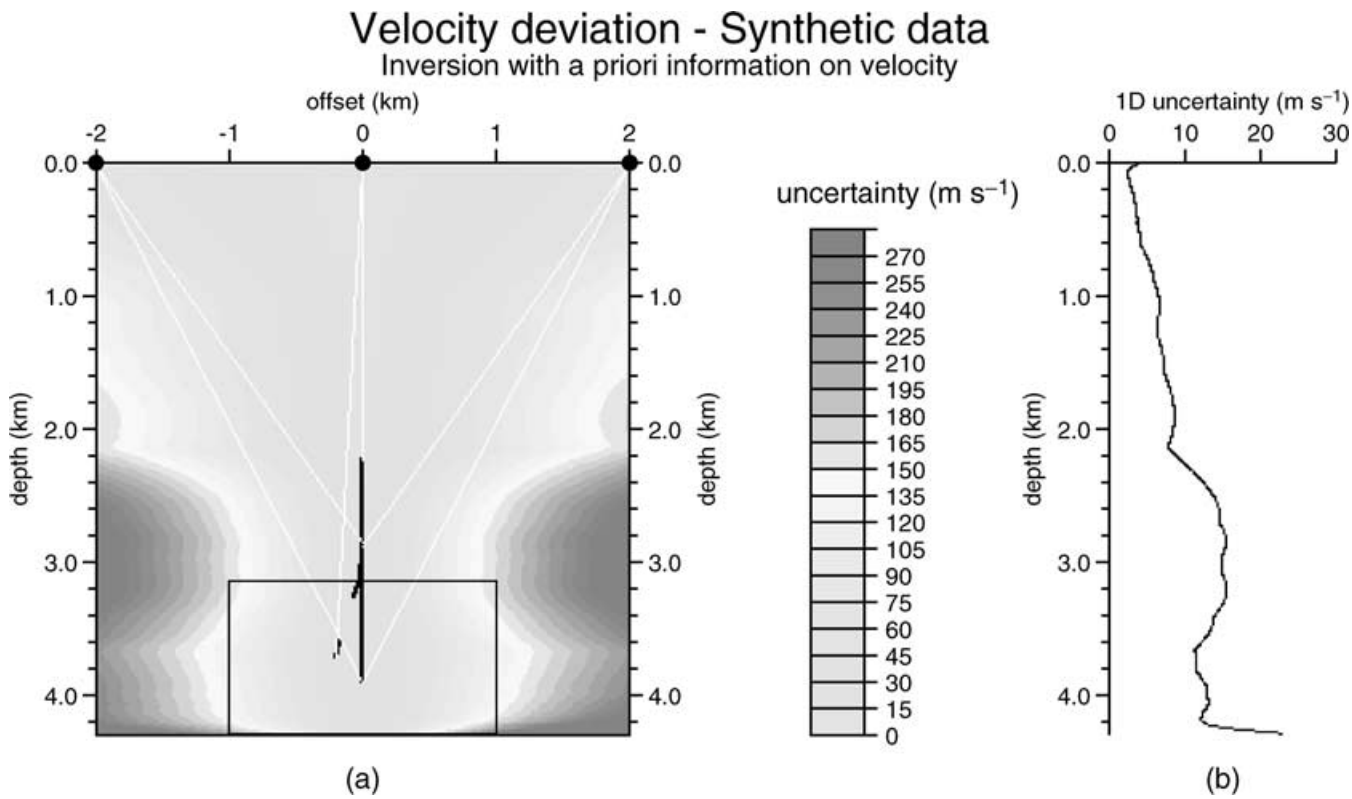


**Figure 5.** Uncertainty on the velocity reconstruction. The velocity at the vertical of the borehole is well resolved by the zero offset data. Uncertainty is limited to 150 m $s^{-1}$ within the region covered by the rays and increases rapidly where data are not available.

log was reduced to 0.033 s $km^{-1}$. Although the traveltime misfit remains similar and small, the constraints misfit is noticeably reduced by introducing the *a priori* information on the velocity. Inversion produces a smooth continuous velocity model, which is comparable to the original layered velocity model (Fig. 4).

The *a posteriori* covariance matrix enables the computation of the uncertainty in the estimated velocity (Fig. 5). The data provide a strong constraint on the velocity estimation at 0.0 km offset with small uncertainty (less than 40 m $s^{-1}$) along the borehole. Lateral velocity variations have also been very well recovered even though only the second-order series expansion was used to define lateral velocity variations. The uncertainty are small in the region well covered by the rays but become large outside of this region.

### Diffraction pre-stack migration

Once the velocity has been determined, a Kirchhoff pre-stack depth migration (Jackson 1995) was performed in a limited region (shown by the box in Fig. 3), between −1.0 and 1.0 km offset, and 3.1 and 4.3 km depth. This region is discretized with 101 × 231 regularly spaced subsurface points, and the traveltimes between each source or receiver and each subsurface point are computed. Fig. 6 shows the traveltimes between the source at 0.0 km and each point in the region. Fig. 6(a) shows the traveltimes generated by tracing rays through the true velocity model, and Fig. 6(b) the traveltimes computed from the series expansion. The error, defined as the difference between the series expansion computation and the ray-traced times, is plotted in Fig. 6(c). It varies from −10 to 18 ms with an rms value of 5.778 ms. The largest errors occur where the true model has interfaces (3.3, 3.5 and 4.0 km depth), which is expected as the continuous traveltime is for the smooth velocity (Fig. 4).

Diffraction pre-stack migration of the synthetic data is performed using the traveltimes computed with our method (Fig. 7a) and that with tracing rays through the true velocity model (Fig. 7b). The migration using our method was five times faster than using the ray-tracing technique. Since the diffraction pre-stack migration was performed without any aperture limitation, the migrated images contain large smiles. The interfaces on the migrated image using our method (Fig. 8a) is slightly weak and the second interface at 3.5 km depth is slightly shifted. Nevertheless, the slope of the deepest interface is correctly imaged.

## APPLICATION TO FIELD DATA

The experimental geometry and the data type were the same as discussed in the previous section. Part of the upgoing wavefield is shown in Fig. 8. The data from the source at the −2.0 km offset reveal the existence of a strong reflector around 2 s. The data from the source at 0.0 km offset contain reflected energy at 1.3 s and around 1.8 s, but the data from the source at 2.0 km were contaminated by noise.

### Tomography

Fig. 3, our synthetic model, is the result of a traveltime inversion using a ray-tracing method. The starting model for this method consisted of flat layers with velocities determined using the direct arrival traveltimes. An interactive perturbation method (Jackson & Pawlak 1994) was used for the inversion. Both transmitted and reflected traveltimes were inverted to produce the velocity–depth model shown in Fig. 3. The results of the inversion matches the main features of

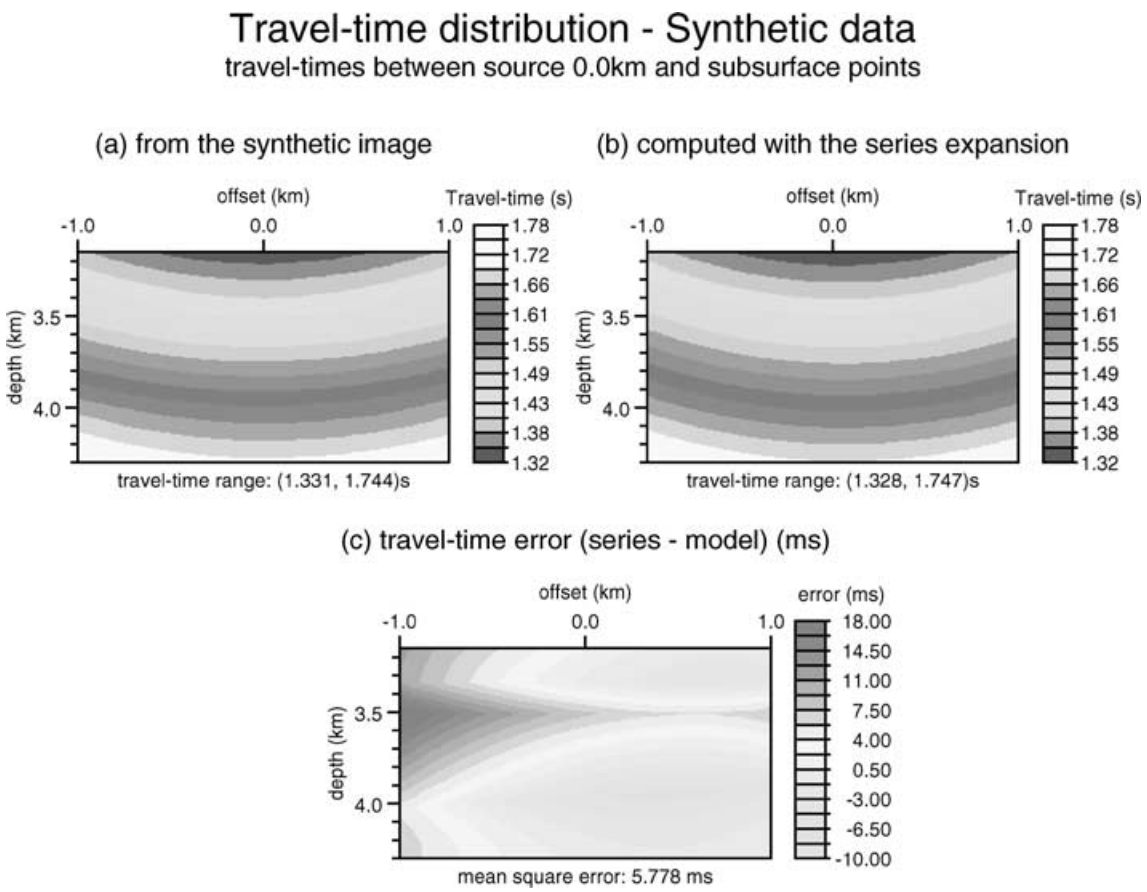


**Figure 6.** Traveltimes between the source at 0.0 km and 101 by 231 subsurface points located between −1.0 and 1.0 km offset, and 3.1 and 4.3 km depth. Times calculated by (a) ray tracing through the true model, (b) times computed with the series expansion and (c) the error in traveltime computation.

the sonic log (Fig. 9). The rms residual of transmitted traveltimes was 5.945 ms.

To perform the series expansion tomography, the velocity depth model was defined by 496 real coefficients as previously discussed for the synthetic experiment ($L = 2$, $M = 8$, $N = 3$, $P = 6$). 33 velocities were picked along the well log at a regular spacing to constrain the velocity variation with depth. After five iterations, the rms traveltime misfit achieved was 2.469 ms. The final velocity–depth model is shown in Fig. 9. As expected the inversion produced a smooth velocity variation with depth. The resultant velocity matches the general feature of the sonic log. However, the rapid changes in velocity observed between 3.3 and 3.8 km depth are not recovered.

The uncertainty in the velocity is shown in Fig. 10, which shows the correlation between the ray coverage and the confidence on velocity estimation. The velocity–depth model at 0.0 km offset is strongly

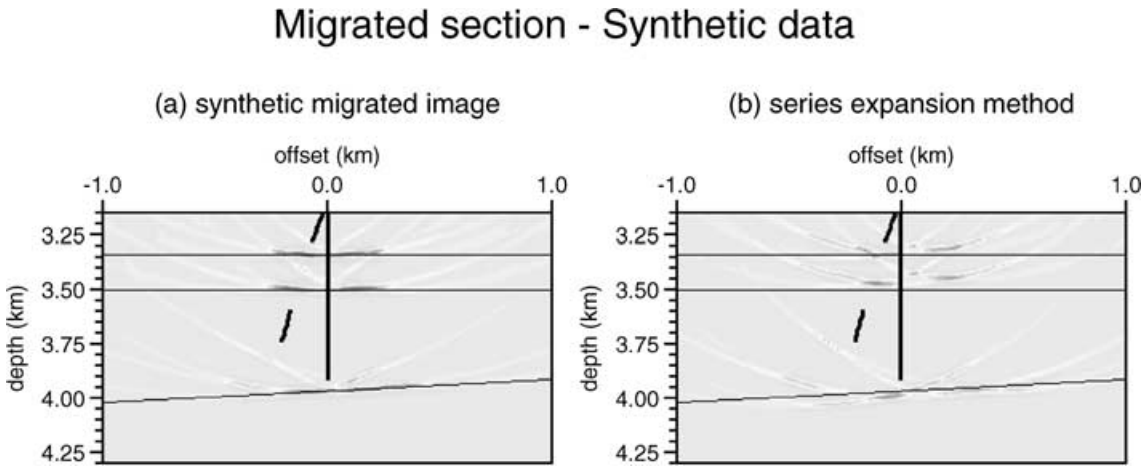


**Figure 7.** Diffraction pre-stack migration without aperture limitation. (a) Synthetic migrated image. (b) Migrated image processing times grid computed with the series expansion. The series expansion migration is five times faster than ray-tracing migration.

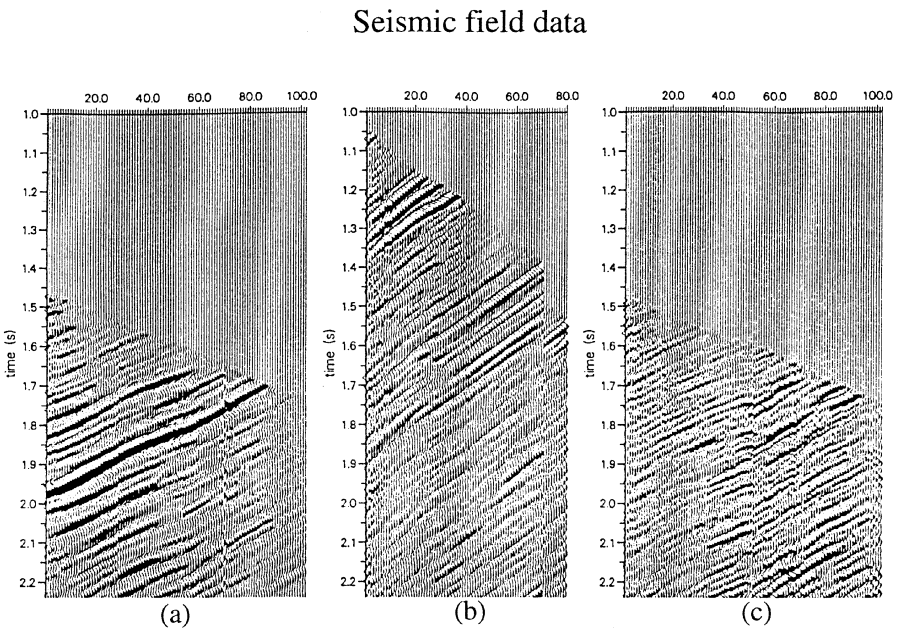


**Figure 8.** Scattered seismic data. Shots from the sources at −2 km and at 0 km contain continuous reflected energy, while the data from the source at 2 km is of poorest quality.

constrained by the zero-offset data and the *a priori* information on velocity. Uncertainties are small. The standard deviation is less than 30 m $s^{-1}$ along the borehole. Owing to the lack of data towards the bottom of the model, the uncertainties become large as the coefficients defining the velocity variation with depth are not sufficiently constrained by the data.

### Migration

As in the synthetic data example, the migration was performed in a region that covers −1.0 and 1.0 km offset, and 3.1 and 4.3 km depth. As before, the area was discretized with 101 × 231 regularly spaced points, and the traveltimes between each source–receiver

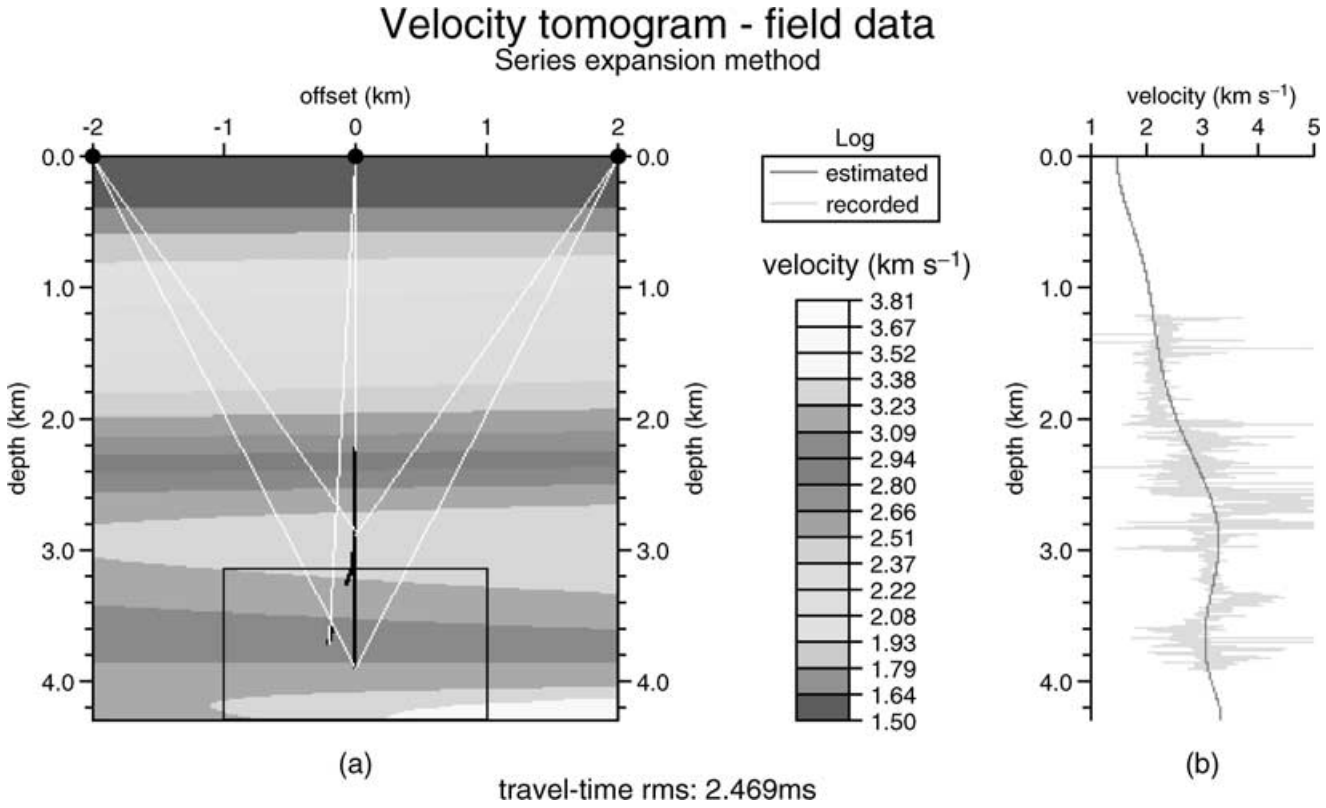


**Figure 9.** Velocity–depth model resulting from the series expansion tomography. The transmitted times are inverted with *a priori* information on the velocity given by the well log. The estimated log (dashed line) matches the sonic log (solid line).

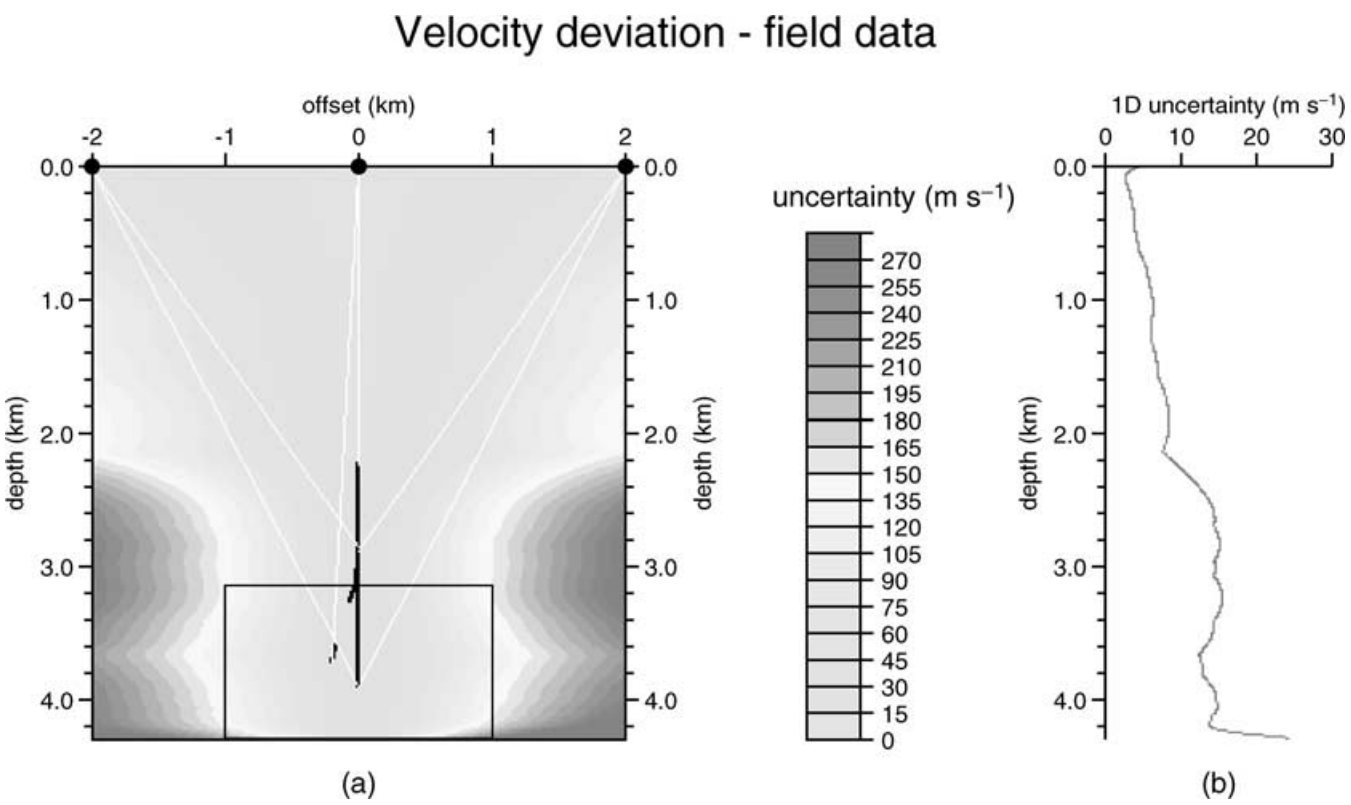


**Figure 10.** Uncertainty in the velocity reconstruction. The standard deviation is coherent with the data. The more information, the less the uncertainty.

and each subsurface point were calculated using either the ray tracing or the series expansion method. The resultant traveltimes computed from each method were very similar. The rms difference between the two traveltimes was 13.2 ms, which is about 1 per cent of the maximum traveltime. The Kirchhoff migration of the field seismic data was performed in the region defined previously by computing the traveltimes from the two methods. The migrated images are shown in Fig. 11. The strong reflection recorded from the source at −2.0 km is predominant on both migrated images, and is imaged at the same depth of 3.75 km. Both images show the same features.

### Limited aperture migration

Both traveltimes and the incidence angles of rays on arriving at each subsurface point can be calculated with the series expansion, enabling one to perform a limited aperture migration. The result of the limited aperture migration with an aperture limitation of 15° is shown in Fig. 12 and the corresponding sonic log is shown on the right-hand side of Fig. 12. The aperture-corrected migration obtained using ray tracing is not available for comparison with the series expansion migration. The strong positive amplitude at 3.3 km depth seems to correspond to a rapid increase of velocity at

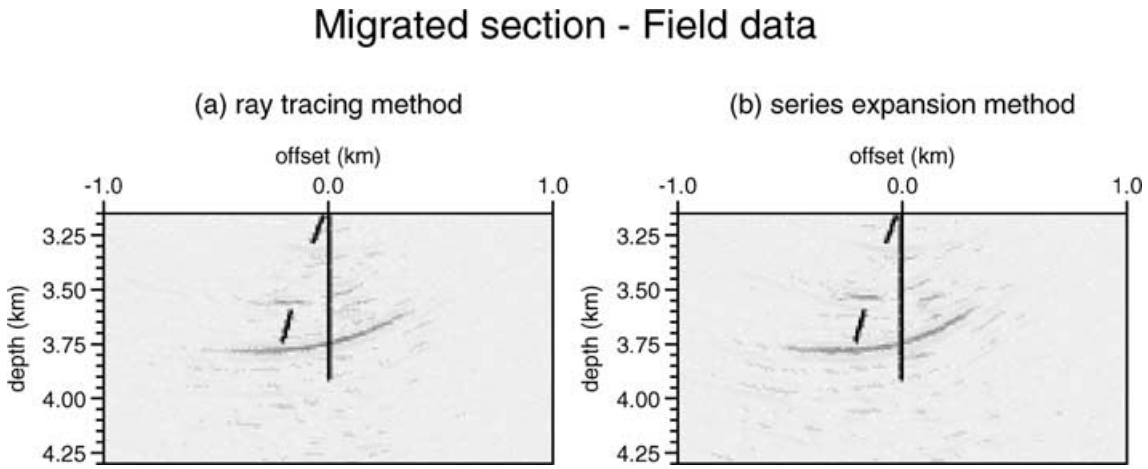


**Figure 11.** Diffraction pre-stack migration performed using (a) a ray-tracing method and (b) the series expansion method.

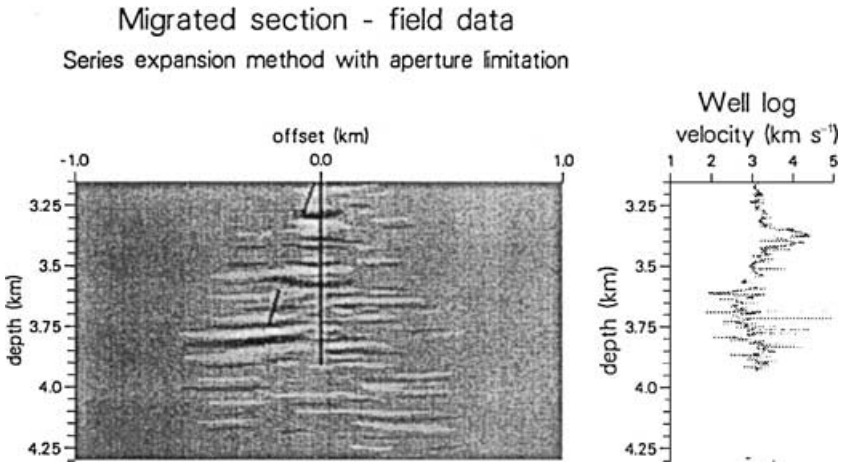


**Figure 12.** Diffraction pre-stack migration with an aperture limitation of 15°. The correlation between the migrated section and the recorded log plotted on the right, reveals that events are shifted up by 2 per cent compared with their expected locations given by the well log.

3.35 km. The succession of events appearing on the migrated section would be correlated with velocity changes on the well log if the log was shifted up by 50 m. This means that the migration velocity is about 2 per cent higher than the true velocity. Consequently, the migrated reflectors appear at a shallower depth than they actually are.

## CONCLUSIONS

(1) We have developed a new method for traveltime tomography that does not require ray tracing but takes ray bending into account.

(2) Our method provides a solution for slowness in a compact analytical form, which can be used to define large-scale models, such the velocity structure of the Earth.

(3) It also provides the traveltime function in an analytical form, which we have used for pre-stack depth migration of seismic reflection data. Such an analytical function could also be used to compute traveltimes between any source and receiver efficiently. For example, we could imagine replacing the Jeffrey–Bullen table with such an analytical function, which could be applicable for a 3-D Earth model.

## ACKNOWLEDGMENTS

This project was supported as a PhD studentship by the Geosciences Research Centre, Elf UK to PE. SCS is partially supported by NERC. The Department of Earth Sciences, University of Cambridge contribution number 7043 and IPG Paris contribution number 1819.

## REFERENCES

Arfken, G., 1985. *Mathematical Methods for Physicists,* 3rd edn, Academic Press, New York.

Bates, R.H.T., Smith, V.A. & Murch, R.D., 1991. Manageable multidimensional inverse scattering theory, *Phys. Rep.,* **201,** 185–277.

Berryhill, J.R., 1979. Wave-equation datuming, *Geophysics,* **44,** 1329–1333.

Berryman, J.G., 1989. Fermat's principle and non-linear traveltime tomography, *Phys. Rev. Lett.,* **62,** 2953–2956.

Bording, R.P., Gersztenkorn, A., Lines, L.R., Scales, J.A. & Treitel, S., 1987. Applications of seismic travel-time tomography, *Geophys. J. R. astr. Soc.,* **90,** 285–303.

Carrion, P., 1991. Dual tomography for imaging complex structures, *Geophysics,* **56,** 1395–1404.

Censor, Y., 1983. Finite series-expansion reconstruction methods, *Proc. IEEE,* **71,** 409–419.

Dines, K.A. & Lytle, R.J., 1979. Computerized geophysical tomography, *Proc. IEEE,* **67,** 1065–1073.

Enright, S.A., Dale, S.M., Smith, V.A., Murch, R.D. & Bates, R.H.T., 1992. Towards solving the bent-ray tomographic problem, *Inverse Problems,* **8,** 83–94.

Golub, G. & Reinsch, C., 1970. Singular value decomposition and least squares solution, *Num. Math.,* **14,** 403–420.

Gordon, R., 1974. A tutorial on ART, *IEEE Trans. Nucl. Sci.,* **21,** 78–93.

Jackson, D.D., 1979. The use of *a priori* data to resolve non-uniqueness in linear inversion, *Geophys. J. R. astr. Soc.,* **57,** 137–157.

Jackson, G.M., 1995. Experiences with anisotropic well seismic migration—field data example, *57th Ann. Int. Mtg, EAEG,* Expanded abstracts of papers.

Jackson, G.M. & Pawlak, F., 1994. Interactive tomography for VSP migration velocity models, *56th Ann. Int. Mtg, EAEG,* Expanded abstracts of papers.

Kak, A.C., 1979. Computerized tomography with x-ray, emission, and ultrasound sources, *Proc. IEEE,* **67,** 1245–1272.

Louis, A.K. & Natterer, F., 1983. Mathematical problems of computerized tomography, *Proc. IEEE,* **71,** 379–389.

McCaughey, M. & Singh, S.C., 1997. Simultaneous velocity and interface tomography of normal-incidence and wide-aperture seismic travel-time data, *Geophys. J. Int.*, **131,** 87–99.

Mersereau, R.M., 1976. Direct Fourier transform techniques in 3-D image reconstruction, *Comput. Biol. Med.,* **6,** 247–258.

Moser, T.J., 1991. Shortest path calculation of seismic rays, *Geophysics,* **56,** 59–67.

Phillips, W.S. & Fehler, M.C., 1991. Traveltime tomography: a comparison of popular methods, *Geophysics,* **56,** 1639–1649.

Podvin, P. & Lecomte, I., 1991. Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools, *Geophys. J. Int.*, **105,** 271–284.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical Recipes in Fortran: the Art of Scientific Computing,* 2nd edn, Cambridge University Press, Cambridge.

Scales, J.A., 1985. *Introduction to Nonlinear Optimization,* Macmillan, New York.

Schneider, W., 1978. Integral formulation for migration in two and three dimensions, *Geophysics,* **43,** 49–76.

Tanabe, K., 1971. Projection method for solving a singular system of linear equations and its applications, *Num. Math.,* **17,** 203–214.
Tarantola, A., 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation,* Elsevier, Amsterdam.
Vidale, J.E., 1990. Finite difference calculation of traveltimes in three dimensions, *Geophysics,* **55,** 521–526.
Worthington, M.H., 1984. *An Introduction to Geophysical Tomography,* First Break, November, 20–26.
Yilmaz, Ö., 1987. *Seismic Data Processing,* Society of Exploration Geophysicists.

## APPENDIX A: EIKONAL EQUATION FROM THE TRAVELTIME SERIES EXPANSION

In this Appendix, we develop the expression of the eikonal eq. (5) leading to the form given by eq. (14). The gradient of the traveltime function computed at $P$, $\nabla_P T$, has its components in the 2-D space

$$\nabla_P T = \left( \frac{\partial T}{\partial X_P}, \frac{\partial T}{\partial Z_P} \right)^{\mathrm{T}}, \tag{A1}$$

where the superscript T denotes the transpose. By substituting $\nabla_P T$ by its components into the eikonal equation formulated at a unique point $P$, eq. (5), can be written as

$$\nabla_P T(P, P) \cdot \nabla_P T(P, P) = \left( \frac{\partial T}{\partial X_P} \right)^2_{d=0} + \left( \frac{\partial T}{\partial Z_P} \right)^2_{d=0} = s^2(P), \tag{A2}$$

where $d$ is the source–receiver distance (eq. 9). The calculation of the partial derivatives of the traveltime function, eq. (11), is straightforward and yields for a null source–receiver distance

$$\left( \frac{\partial T}{\partial X_P} \right)_{d=0} = \sum_{l=0}^{L} \sum_{m=0}^{M} \sum_{n=-N}^{N} C_{lmn1} P_l(X_P) P_m(Z_P) \exp(i2n\theta) \times \left( \frac{\partial d}{\partial X_P} \right)_{d=0}, \tag{A3a}$$

and

$$\left( \frac{\partial T}{\partial Z_P} \right)_{d=0} = \sum_{l=0}^{L} \sum_{m=0}^{M} \sum_{n=-N}^{N} C_{lmn1} P_l(X_P) P_m(Z_P) \exp(i2n\theta) \times \left( \frac{\partial d}{\partial Z_P} \right)_{d=0}. \tag{A3b}$$

By substituting (A3a) and (A3b) into (A2) and considering that

$$\left( \frac{\partial d}{\partial X_P} \right)^2 + \left( \frac{\partial d}{\partial Z_P} \right)^2 = 1, \tag{A5}$$

the eikonal equation (A2) becomes

$$\nabla_P T(P, P) \cdot \nabla_P T(P, P) = \left( \sum_{l=0}^{L} \sum_{m=0}^{M} \sum_{n=-N}^{N} C_{lmn1} P_l(X_P) P_m(Z_P) \exp(i2n\theta) \right)^2 = s^2(P). \tag{A6}$$

## APPENDIX B: RESOLUTION ANALYSIS AFTER INVERSION

A resolution analysis of the inversion result is developed in this Appendix where we demonstrate that provided the starting model $\mathbf{m}_0$ is a solution of the constraint equations, the *a posteriori* probability density function eq. (23) is Gaussian, so that the algorithm converges towards the maximum-likelihood solution. This analysis provides an expression for the *a posteriori* covariance matrix that can be used to quantify uncertainties on model parameters and so on the velocity reconstruction and the traveltime computation.

### The *a posteriori* covariance matrix

For the case of a linear system, i.e. if the data are a linear function of the model parameters, the *a posteriori* probability density function PDF ($\mathbf{m}$) is a Gaussian function with a mean model $\mathbf{m}_*$ and *a posteriori* covariance matrix $\mathbf{C}_*$,

$$\mathrm{PDF}(\mathbf{m}) = \text{constant} \times \exp\left[ -\frac{1}{2} (\mathbf{m} - \mathbf{m}_*)^{\mathrm{T}} \mathbf{C}_*^{-1} (\mathbf{m} - \mathbf{m}_*) \right]. \tag{B1}$$

Although the analytical traveltime function is a linear function of the model parameters, the theoretical relationships established by the constraint equations between the model parameters are nonlinear. Thus, the *a posteriori* probability density function (B1) is not Gaussian and the inversion algorithm may converge towards a local minimum of the misfit function. Nevertheless, we demonstrate that if the initial model satisfies the constraint equations, i.e. if it is contained within the feasibility region defined by the theoretical constraints, the *a posteriori* probability density function is Gaussian, and therefore the algorithm converges towards the global minimum of the misfit function.

Eq. (24) can be written as $\mathbf{D}_{\mathrm{cal}}(\mathbf{m}) = \mathbf{G}\mathbf{m}$, where $\mathbf{G}$ is a matrix containing the geometrical terms of the series expansion. The linearization of the constraint equations $\mathbf{fc}(\mathbf{m})$ around an initial model $\mathbf{m}_0$ yields, neglecting higher-order terms,

$$\mathbf{fc}(\mathbf{m}) \approx \mathbf{fc}(\mathbf{m}_0) + \mathbf{F}_0(\mathbf{m} - \mathbf{m}_0), \tag{B2}$$

where the matrix $\mathbf{F}_0$ contains the partial derivatives of $\mathbf{fc}$ with respect to the model parameters computed at $\mathbf{m}_0$, i.e.

$$F_0^{ij} = \frac{\partial f c^i(m_0)}{\partial m_0^j}. \tag{B3}$$

By chosing the initial model $\mathbf{m}_0$, such that it is a solution of the constraint equations, which requires

$$\mathbf{fc}(\mathbf{m}_0) = 0, \tag{B4}$$

the misfit function $Q(\mathbf{m})$ takes the following form:

$$Q(\mathbf{m}) \approx (\mathbf{G}\mathbf{m} - \mathbf{D}_{\mathrm{obs}})^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} (\mathbf{G}\mathbf{m} - \mathbf{D}_{\mathrm{obs}}) + (\mathbf{m} - \mathbf{m}_0)^{\mathrm{T}} \mathbf{F}_0^{\mathrm{T}} \mathbf{C}_{\mathbf{fc}}^{-1} \mathbf{F}_0 (\mathbf{m} - \mathbf{m}_0). \tag{B5}$$

At the maximum-likelihood solution $\mathbf{m}_*$, the gradient of the misfit $\mathbf{g}_*$ vanishes, $\mathbf{g}_* = \nabla_m Q(\mathbf{m}_*) = 0$, so that

$$\mathbf{G}^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} (\mathbf{G}\mathbf{m}_* - \mathbf{D}_{\mathrm{obs}}) + \mathbf{F}_0^{\mathrm{T}} \mathbf{C}_{\mathbf{fc}}^{-1} \mathbf{F}_0 (\mathbf{m}_* - \mathbf{m}_0) = 0. \tag{B6}$$

Note that since the calculated data are linear functions of the model $\mathbf{m}$, the matrix $\mathbf{G}$ is independent of $\mathbf{m}$. After some manipulations (Tarantola 1987), the misfit function $Q(\mathbf{m})$ can be written in the form

$$Q(\mathbf{m}) \approx \frac{1}{2} (\mathbf{m} - \mathbf{m}_*)^{\mathrm{T}} \mathbf{H}_* (\mathbf{m} - \mathbf{m}_*) - \frac{1}{2} \mathbf{m}_*^{\mathrm{T}} \mathbf{H}_* \mathbf{m}_* + \mathbf{D}_{\mathrm{obs}}^{\mathrm{T}} \mathbf{C}_{\mathrm{D}}^{-1} \mathbf{D}_{\mathrm{obs}} + \mathbf{m}_0^{\mathrm{T}} \mathbf{F}_0^{\mathrm{T}} \mathbf{C}_{\mathbf{fc}}^{-1} \mathbf{F}_0 \mathbf{m}_0. \tag{B7}$$

where $\mathbf{H}_*$ is the Hessian of the misfit function calculated at the maximum-likelihood model $\mathbf{m}_*$, with

$$\mathbf{H}_* = \nabla_m \nabla_m Q(\mathbf{m}_*) = 2\left(\mathbf{G}^{\mathrm{T}}\mathbf{C}_{\mathrm{D}}^{-1}\mathbf{G} + \mathbf{F}_0^{\mathrm{T}}\mathbf{C}_{\mathbf{fc}}^{-1}\mathbf{F}_0\right). \tag{B8}$$

The last three terms of the right-hand side of eq. (B7) are constant, i.e. independent of $\mathbf{m}$, except the first term, so that

$$Q(\mathbf{m}) \approx \frac{1}{2}(\mathbf{m}-\mathbf{m}_*)^{\mathrm{T}}\mathbf{H}_*(\mathbf{m}-\mathbf{m}_*) + \text{constant}. \tag{B9}$$

By replacing the misfit function, eq. (B9), in the probability density function PDF ($\mathbf{m}$), and moving the constant terms into a factor of the exponential, we obtain

$$\mathrm{PDF}\,(\mathbf{m}) = \text{constant} \times \exp\left[-\frac{1}{4}(\mathbf{m}-\mathbf{m}_*)^{\mathrm{T}}\mathbf{H}_*(\mathbf{m}-\mathbf{m}_*)\right]. \tag{B10}$$

Eq. (B10) shows that the *a posteriori* probability density function is also Gaussian with its centre at the maximum-likelihood model $\mathbf{m}_*$ and an *a posteriori* covariance matrix, denoted $\mathbf{C}_*$, given by

$$\mathbf{C}_* = \left(\mathbf{G}^{\mathrm{T}}\mathbf{C}_{\mathrm{D}}^{-1}\mathbf{G} + \mathbf{F}_0^{\mathrm{T}}\mathbf{C}_{\mathbf{fc}}^{-1}\mathbf{F}_0\right)^{-1} = 2\mathbf{H}_*^{-1}. \tag{B11}$$

This result demonstrates that, provided the initial model satisfies the constraint equations, the *a posteriori* probability density function is Gaussian. This implies that the minimum of the misfit function is at the maximum-likelihood solution $\mathbf{m}_*$.

We now demonstrate that the initial model that is defined as the average slowness $\bar{s}$ in eq. (28) is a solution of the constraint equations and therefore lies within the feasibility domain. The initial calculated traveltime function $\mathbf{T}_{\mathrm{cal}}(\mathbf{m}_0)$ reduces to a single term

$$T^i_{\mathrm{cal}}(\mathbf{m}_0) = C^0_{0001} d^i = \bar{s} d^i, \tag{B12}$$

where $d^i$ is the source–receiver distance for the $i$th calculated traveltime. The eikonal equation, computed for the initial model $\mathbf{m}_0$, takes the form

$$\nabla_P T^i_{\mathrm{cal}}(\mathbf{m}_0) \cdot \nabla_P T^i_{\mathrm{cal}}(\mathbf{m}_0) = \bar{s}^2, \tag{B13}$$

and, consequently

$$\nabla_S \left[\nabla_P T^i_{\mathrm{cal}}(\mathbf{m}_0) \cdot \nabla_P T^i_{\mathrm{cal}}(\mathbf{m}_0)\right] = 0, \tag{B14}$$

which suggests that the initial model satisfies the condition for the constraint equations and hence the probability density function is a Gaussian.

### Estimation of uncertainty

The *a posteriori* covariance matrix, which is twice the inverse of the Hessian matrix (eq. B11), provides information on the uncertainty of the model parameters. The physical meaning of the Hessian matrix is given by the curvature of the misfit function. A high curvature of the misfit function in the neighbourhood of its minimum implies a large value for the Hessian matrix components, and therefore a small covariance. The diagonal elements of the covariance matrix contain the variance of the model parameters. As the variance of a sum is the sum of the variance of each coefficient, the uncertainty in the final slowness $s_{\mathrm{cal}}(P)$ will be the square root of the sum of the variance of each term comprising the series expansion. If $\sigma_{s_{\mathrm{cal}}(P)}$ is the uncertainty of the calculated slowness at $P$, then

$$\sigma_{s_{\mathrm{cal}}(P)} = \left[\sum_{l=0}^{L}\sum_{m=0}^{M} \sigma^2_{C_{lm01}} P_l(X_P)\, P_m(Z_P)\right]^{1/2}, \tag{B15}$$

where $\sigma_{C_{lm01}}$ is the standard deviation of the $C_{lm01}$ coefficients of the series expansion. The uncertainty in the estimated velocity is defined as the ratio of the slowness uncertainty and the slowness square. If $v_{\mathrm{cal}}(P)$ is the calculated velocity at $P$, then the uncertainty in the calculated velocity $\sigma_{v_{\mathrm{cal}}(P)}$ is

$$\sigma_{v_{\mathrm{cal}}(P)} = \frac{\sigma_{s_{\mathrm{cal}}(P)}}{s^2_{\mathrm{cal}}(P)}. \tag{B16}$$

Similarly, the uncertainty in the calculated traveltime between a source $S$ and a receiver $R$, $\sigma_{T_{\mathrm{cal}}(R,S)}$, is given by

$$\begin{aligned}\sigma_{T_{\mathrm{cal}}(R,S)} = \Bigg[&\sum_{l=0}^{L}\sum_{m=0}^{M} \sigma^2_{C_{lm01}} P_l(X_M) P_m(Z_M)\, d(R,S) \\ &+ \sum_{l=0}^{L}\sum_{m=0}^{M}\sum_{n=-N}^{N}\sum_{p=2}^{P} \sigma^2_{C_{lmnp}} P_l(X_M) P_m(Z_M) \\ &\times \exp(i2n\theta)\, d^p(R,S)\Bigg]^{1/2},\end{aligned} \tag{B17}$$

where $\sigma_{C_{lmnp}}$ is the standard deviation of the $C_{lmnp}$ coefficient of the series expansion.