

Non-linear surface wave phase velocity inversion based on ray theory

Marianne Bruneton,¹ Véronique Farra,² Helle Anette Pedersen¹
and the SVEKALAPKO Seismic Tomography Working Group

¹Laboratoire de Géophysique Interne et Tectonophysique, Observatoire des Sciences de l'Univers de Grenoble, BP 53, 38041 Grenoble cedex 09, France

²Département de Sismologie, Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris cedex 05, France. E-mail: mbruneto@obs.ujf-grenoble.fr

Accepted 2002 June 12. Received 2002 June 6; in original form 2001 August 10

SUMMARY

The development of temporary and permanent broad-band seismic arrays reinforces the need for advanced interpretation techniques in surface-wave analysis. We present a new method based on 2-D paraxial ray theory of inverting teleseismic surface-wave phase information and constructing phase velocity maps on a regional scale. Measurements of local phase velocities and propagation directions of Rayleigh waves taken from full waveform synthetic seismograms are used to validate the ray theory for smooth structures on a regional scale. Curved wavefronts created by heterogeneous structure outside the study area are taken into account through joint inversion for the phase velocity field and the shape of the incoming wavefronts. In the forward ray tracing procedure, the curved wavefronts are introduced through the boundary conditions by equating the slowness vector of the ray at the edge of the study region with the known gradient of the arrival time of the wave. To make the inverse problem non-singular we constrain the parameters in the inversion primarily by applying a smoothness criteria on the velocity field and on the incoming wave-field. Inversions of synthetic data sets computed by direct ray tracing and by full waveform modelling show that for 100 km spacing between stations the minimum size of structure that we can image is approximately 150 km. Heterogeneities with a size approximately equal to the wavelength are reconstructed by the ray-based inversion even though velocity variations are underestimated due to the wave-field smoothing of the structures. A minimum signal-to-noise ratio of 3.5 is necessary in order to correctly retrieve the phase velocity field. Inversion of a subset of the SVEKALAPKO data for 60 s period demonstrates the applicability of the method on real data.

Key words: lithosphere, ray theory, seismic tomography, surface waves.

1 INTRODUCTION

Fundamental mode surface waves of 10 to 100 s period provide information on the shear-velocity of the upper 200 km of the Earth. Such studies are useful to complement lithospheric models obtained using body-wave tomography as these models are mostly based on *P*-wave traveltimes inversion and are therefore sensitive to the *P*-velocity structure. *S*-wave tomography is difficult to conduct as low velocity zones, which are more pronounced in *S* than *P*-wave velocity, introduce non-linearity into the inversion. Surface waves therefore provide a more reliable means to retrieve the *S*-wave structure. *P* and surface wave tomographies are also complementary in terms of resolution as the former has very good lateral resolution and the latter better vertical resolution.

To date, inversions for lithospheric structures using surface waves on a regional scale have mostly been based either on the classical two-station method (see, for example, Calcagnile 1991), or on network analysis. One alternative approach is to measure local phase velocities by analyzing data from small aperture arrays (Cotte *et al.*

2000). Larger scale arrays have been used for regional studies using various methods such as the 2-D group velocity ray-based tomography of Yanovskaya *et al.* (1998). Several inversion methods are based on waveform modelling, for example the path-average approximation with modal summations (Nolet 1990; Zielhuis & Nolet 1994; Simons *et al.* 1999), linearized scattering (Snieder 1988a,b), mode coupling (Maupin 1988; Marquering & Snieder 1996; Marquering *et al.* 1996) and multiple scattering using Born's approximation (Friederich & Wielandt 1995; Friederich 1998; Pollitz 1999). The last three cited studies have the advantage that the shapes of the incoming wave-fields need not be planar, so reducing artefacts due to non-planar wavefronts incident upon the array. However, they use the acoustic approximation and can be applied only to the vertical component of the seismograms, that is to the Rayleigh wave (Friederich *et al.* 1993). As van der Lee (1998) shows, one should be cautious in interpreting the entire waveform of a seismogram in terms of heterogeneities along the propagation path. Amplitude anomalies of intermediate-period fundamental-mode Rayleigh waves are dominated by earthquake source parameters and heterogeneous

structures near the source, unlike phase anomalies which mostly depend on the structure along the wave path. Yanovskaya (1996) proposes a tomographic method based on inversion of azimuthal anomalies measured from polarization analysis, while still remaining within the framework of geometrical ray theory.

The aim of the present study is to invert the data set recorded during the SVEKALAPKO deep seismic experiment (SVEKALAPKO Seismic Tomography Working Group 2001), to determine the lithospheric structure of the Baltic Shield and its lateral variations beneath Finland. Improved knowledge of this structure should advance our understanding of plate-tectonic processes in the early history of the Earth, as it contains traces of a 1.9 Ga collision (for a summary of the geology and tectonic feature of the Fennoscandian Shield, see for example Korja *et al.* 1993).

The lateral lithospheric heterogeneities in the Fennoscandian Shield are generally considered to be smooth. Several deep seismic sounding profiles have been carried out in the region (Luosto 1997), revealing, in particular, an anomalously deep Moho at approximately 60 km depth under the eastern part of the Ladoga-Bothnian Bay zone, which is also associated with very high seismic velocities in the lower crust. The lithosphere has been shown to be at least 170 km thick beneath the area (Babuška *et al.* 1988; Calcagnile 1991).

To obtain 2-D phase velocity maps of the individual Rayleigh and Love modes, we adapted the paraxial ray tracing method developed by Farra (1990, 1993). The use of ray theory limits us to smoothly varying structures as mode coupling is not taken into account. We extend Farra's method to take into account non-planar incoming wave-fields and simultaneously invert for the phase velocity field and the shape of the incident wavefronts. In the first part of the paper, we demonstrate the validity of ray theory for surface wave propagation in smooth regional structures—local propagation directions and velocities are calculated from synthetic seismograms computed using the Indirect Boundary Element Method (IBEM) for multilayered media (Pedersen *et al.* 1996), and compared with the results expected from ray theory. We then define the ray tracing procedure and the inversion method we use to invert for the phase velocity field. Finally, we present numerical examples from inversion of synthetic data sets computed by ray tracing and IBEM, as well as a preliminary inversion of a subset of the SVEKALAPKO data.

2 APPLICABILITY OF RAY THEORY

A major drawback to the use of ray theory is that it derives from a high frequency approximation and therefore does not take into account heterogeneities smaller than the wavelength of the incident wave. In this section, we justify the use of ray theory to obtain 2-D velocity maps of structures of the type encountered beneath SVEKALAPKO through analysis of full waveform synthetic seismograms. To compute the synthetic seismograms, we use a 2-D structure composed of a single layer (the crust) over a half space (the mantle). The model shown in Fig. 1 is a simplified north–south cross-section of the Moho depth map obtained by Luosto (1990) from a compilation of seismic profiles in Fennoscandia. The thickness of the crust varies smoothly from 44 to 60 km in the central part of the region, in an area 350 km wide. On each side of the trough, the crust thins out to 43 km over a distance of 100 km. In this model, we propagate a plane Rayleigh wave with a backazimuth of 60 degrees (that is, coming from approximately the north-east). Synthetic seismograms are computed using the Indirect Boundary Element Method (IBEM) developed by Sánchez-Sesma & Campillo, and extended to 2.5-D multilayered media (Pedersen *et al.* 1996). IBEM solves the

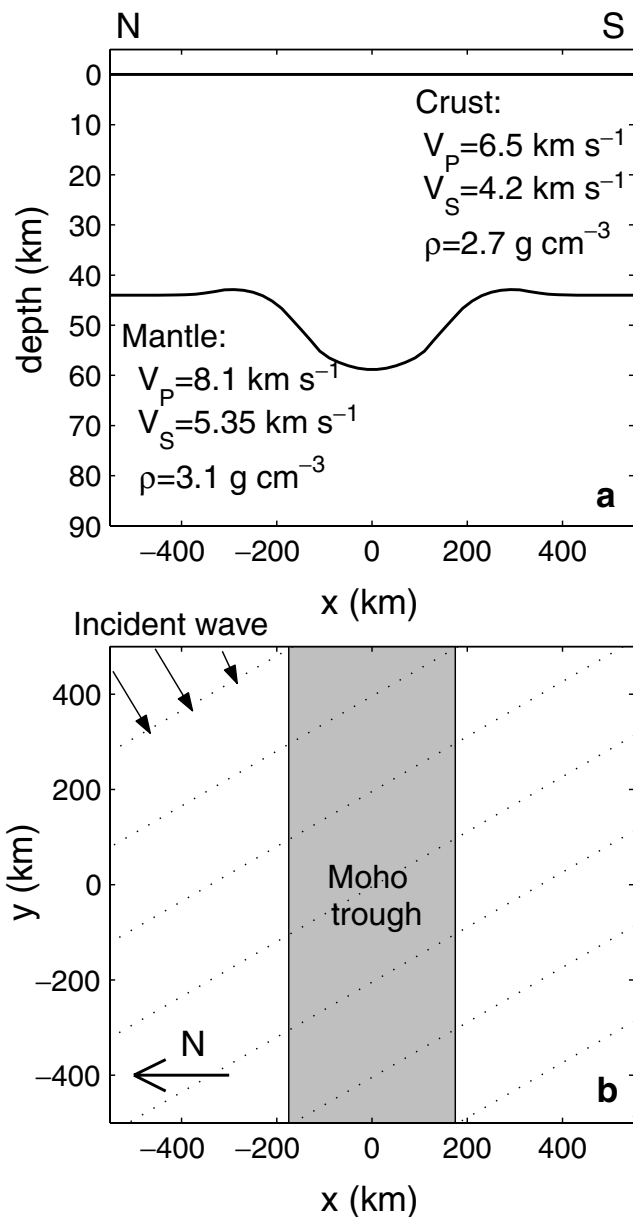


Figure 1. 2-D crustal model used to compute synthetic seismograms by the Indirect Boundary Elements Method. (a) North-South cross-section; (b) map view of the propagating surface wave, black arrows give the initial direction of the wave, dotted lines are successive wavefronts.

full wave equation so multiple scattering and all coupling effects between waves are taken into account. Even though lateral variations in the model are smooth, they are distributed over distances comparable to a wavelength (varying between 40 and 460 km in the 10–100 s period interval).

2.1 Comparison of measured and predicted phase velocities

We compute seismograms for a receiver geometry similar to that of the SVEKALAPKO experiment, a regular grid with 90 km between pairs of adjacent stations. The time delays between pairs of traces are computed in two steps. The first step consists of finding the best fitting dispersion curve for all the data from one event using the

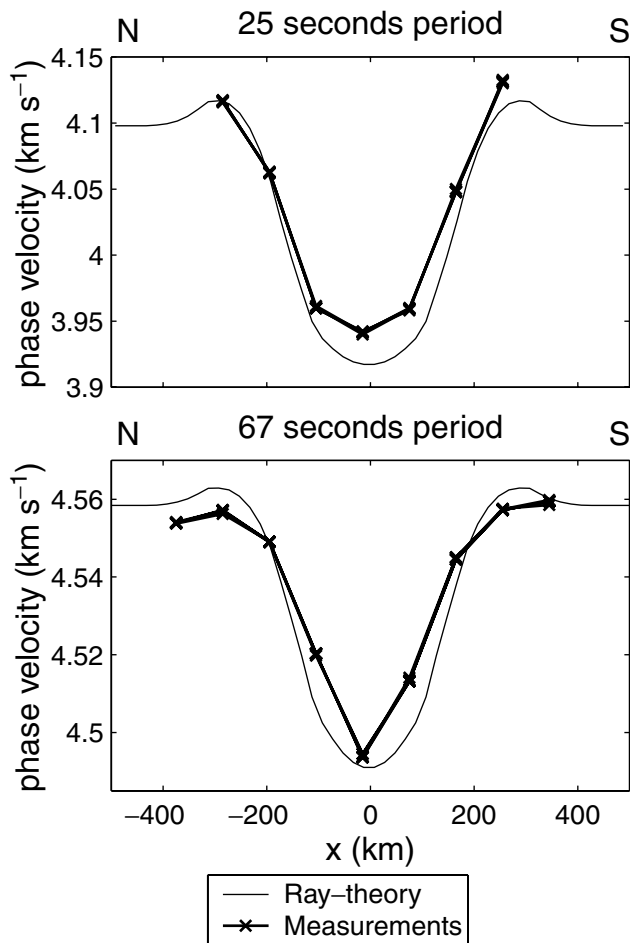


Figure 2. Phase velocities obtained from synthetic seismograms computed with IBEM, on the structure of Fig. 1.

slant-stack method developed by McMechan & Yedlin (1981) by computing the image of the wave-fields in the (p, ω) plane, where p is the slowness and ω the frequency. The maximum amplitudes at each frequency define the best fitting phase velocity curve. We correct the synthetic seismograms for this dispersion, and then compute the remaining small time delays for each pair of traces using Wiener filtering. Only frequencies with high coherency (>0.9) are considered. To obtain the total delay between the original seismograms we add the delays computed at each of the two steps. The phase velocity between each neighbouring station pair is calculated as the inverse of the gradient of the time delay.

Fig. 2 displays the phase velocity curves obtained for periods of 25 and 67 s (approximately 100 and 300 km wavelength) along different profiles perpendicular to the structure. They are compared to the theoretical phase velocities calculated assuming for every point a horizontally stratified medium with the properties of the model immediately below the point, using the software published by Herrmann (1985). At 25 s period as well as at 67 s, the measured phase velocities correctly display the geometry of the structure. However, the structure is smoothed; the minimum velocity is slightly larger than predicted by theory. For the longest wavelength (300 km), the structure is also simplified and the small bumps on either side of the trough are not resolved. At 25 s period, velocities slightly higher than theoretically predicted appear immediately down-stream the structure, possibly due to the interaction of the primary Rayleigh wave with Rayleigh-Love-Rayleigh converted waves.

2.2 Comparison of measured and predicted propagation directions

We also measured the propagation direction of the wave on the synthetic seismograms. The backazimuth angle is measured using the array analysis developed by Cotte *et al.* (2000) following the work of Barker *et al.* (1996) and Poupinet *et al.* (1984). For each measurement, we use a dense array of 5×5 stations, with 3 km between neighbouring station pairs.

Fig. 3 displays the measured backazimuth in six frequency bands. On each graph, the propagation angles given by Snell's law for the lowest and highest predicted phase velocities within the frequency interval are also displayed.

The measured backazimuth is in close agreement with the theoretical deviation in all frequency bands. The largest difference between measured and theoretical values is half a degree and appears for the higher frequencies in waves that have crossed the heterogeneous structure. As expected, the structure is smoothed at long periods where the heterogeneity scale is small as compared to the wavelength.

The part of the structure where the Moho becomes shallower is detected at all frequencies, but displaced by a few kilometers from its real position, probably due to the interaction of the incoming waves with those reflected from this interface. Notwithstanding these small differences, ray theory seems to provide a good approximation to the behaviour of Rayleigh waves propagating across a smooth structure, even though the variation takes place over distances of the order of the wavelength.

3 RAY TRACING FORMULATION

Our aim is to obtain the regional distribution of phase velocity as a function of frequency by inverting time delays of surface waves measured on a 2-D network of receivers, repeated for each wave mode. We first need a formulation of the forward problem: how to compute the arrival times of the wave at every receiver. We assume that we know the phase velocity at a given frequency throughout a rectangular region containing the 2-D network of receivers. We then suppose that a surface wave is incident upon this region. Surface waves are known to be generally non-planar (Friederich *et al.* 1994; Friederich 1998), so we characterize the wavefront by its arrival time at the boundary of the model. The corresponding rays going from the edge of the box to every receiver have to be traced in order to compute the arrival times at the receivers.

The ray tracing method used here is derived from Farra & Madariaga (1987) and Farra (1990). In Section 3.1 we briefly present this method of paraxial ray tracing, and in Section 3.2 we present the theory necessary for its extension to the geometrical setup used here, that is, stations inside the study region and non-planar incoming waves.

3.1 Paraxial ray tracing

The high frequency approximation of the wave equation leads to the eikonal equation:

$$(\nabla_x T)^2 = c^{-2} = u^2 \quad (1)$$

where T is the arrival time, c the phase velocity, u the slowness, and ∇_x the gradient in the horizontal plane. Eq. (1) belongs to the family of Hamilton-Jacobi equations $H(\mathbf{x}, \mathbf{p}, \tau) = 0$ (Červený 1989). The Hamiltonian H is a function of position \mathbf{x} , slowness vector $\mathbf{p} = \nabla_x T$, and sampling parameter τ along the ray. As the problem is

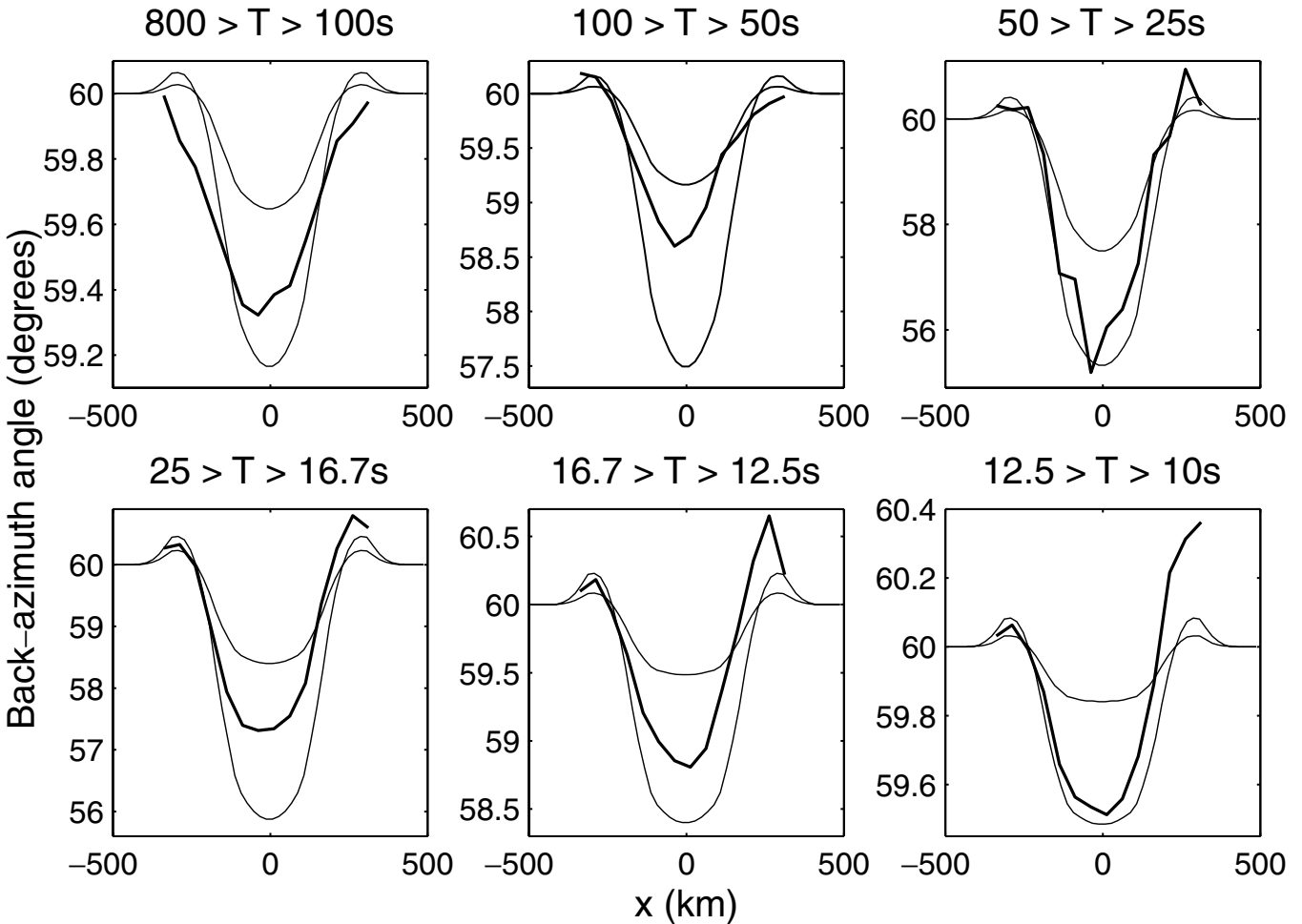


Figure 3. Measured propagation angle on synthetic seismograms across the 2-D structure of Fig. 1. The period interval is given above each plot, thin lines are the extreme values within each period interval according to ray theory. Note that the backazimuth scale is different in each plot.

two-dimensional, $\mathbf{x} = (x_1, x_2)$ and $\mathbf{p} = (p_1, p_2)$. The Hamiltonian may take many different forms (Červený 1989) and we choose here that used by Farra (1990):

$$H(\mathbf{x}, \mathbf{p}, \tau) = \frac{1}{2}(\mathbf{p}^2 - u^2(\mathbf{x})) \quad (2)$$

The corresponding sampling parameter τ has units $\text{km}^2 \text{s}^{-1}$ and is related to traveltime by $dT = u^2 d\tau$. This parameter is chosen instead of the more commonly used traveltime (in s) or arclength (in km) because the elements of the Hamiltonian are separated, depending on slowness and position, which simplifies the computation of the derivatives (Farra 1993).

The Method of Characteristics (Courant & Hilbert 1966) is used to solve the eikonal equation. The rays, described in terms of position $\mathbf{x}(\tau)$ and slowness vector $\mathbf{p}(\tau)$ satisfy the canonical equations:

$$\begin{aligned} \dot{\mathbf{x}} &= \nabla_{\mathbf{p}} H = \mathbf{p} \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{x}} H = \frac{1}{2} \nabla_{\mathbf{x}} u^2 \end{aligned} \quad (3)$$

where dot means derivative with respect to τ and $\nabla_{\mathbf{p}}$ is the gradient with respect to the slowness vector.

Suppose that a ray has been traced in the known medium using eq. (3). Around this ray, called the central ray, we can obtain neighbouring rays by means of first order perturbation theory. Paraxial rays deviate from the reference ray by small perturbations in position

($\mathbf{x} = \mathbf{x}_0 + \delta\mathbf{x}$) and slowness vector ($\mathbf{p} = \mathbf{p}_0 + \delta\mathbf{p}$). The perturbations of position and slowness satisfy the paraxial ray tracing equations deduced from (3):

$$\begin{aligned} \delta\dot{\mathbf{x}} &= \nabla_{\mathbf{x}} \nabla_{\mathbf{p}} H \delta\mathbf{x} + \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} H \delta\mathbf{p} = \delta\mathbf{p} \\ \delta\dot{\mathbf{p}} &= -\nabla_{\mathbf{x}} \nabla_{\mathbf{x}} H \delta\mathbf{x} - \nabla_{\mathbf{p}} \nabla_{\mathbf{x}} H \delta\mathbf{p} = \mathbf{U} \delta\mathbf{x} \end{aligned} \quad (4)$$

where \mathbf{U} is a 2×2 matrix whose elements $U_{ij} = \frac{1}{2} \frac{\partial^2 u^2}{\partial x_i \partial x_j}$ are calculated along the reference ray.

The paraxial propagator of system (4) is used to compute the evolution of $\delta\mathbf{x}$ and $\delta\mathbf{p}$ along the ray.

The rays that are solution of eq. (4) also have to satisfy another relation derived from the perturbation of the eikonal equation (2):

$$\delta H(\tau) = \nabla_{\mathbf{p}} H \cdot \delta\mathbf{p} + \nabla_{\mathbf{x}} H \cdot \delta\mathbf{x} = 0 \quad (5)$$

where the partial derivatives of H are taken along the reference ray. As δH is invariant along any solution of system (4), eq. (5) merely needs to be satisfied at $\tau = 0$.

3.2 Two-point ray tracing

Our aim is to compute the ray going from a curved wavefront, known on the edges of a rectangular region, through a medium of known velocity, to a given receiver, so as to compute the arrival time at

the receiver. To simplify the computation, the rays are traced from $\tau = 0$ at the receiver back to the edge of the box.

The boundary conditions are as follows: first, the ray goes through the receiver position \mathbf{x}_r at $\tau = 0$, that is

$$\mathbf{x}(0) = \mathbf{x}_r \quad (6)$$

Second, the projections of the slowness vector of the ray and of the arrival time gradient onto the edge of the model must be equal. The arrival time T_0 of the incident wave at the edge of the model is known, its definition along the edge allowing the wavefront to be non-planar. The slowness vector of a ray intersecting the edge of the model across a boundary parallel to the x_i axis satisfies the boundary condition:

$$p_i = -\frac{\partial T_0}{\partial x_i} \quad (7)$$

where we use the definition of the slowness vector $\mathbf{p} = \nabla_{\mathbf{x}} T$, and the negative sign comes from the fact that the ray is traced from the station to the edge of the model.

The reference ray with initial slowness vector $\mathbf{p}_0(0)$ in the direction given by the theoretical backazimuth of the wave is first traced from the receiver to the edge of the medium. All quantities on this central ray are identified by the subscript $_0$. This ray satisfies the boundary condition (6), but not necessarily (7) as four initial scalar conditions are needed to define the ray at $\tau = 0$ ($x_1(0)$, $x_2(0)$, $p_1(0)$ and $p_2(0)$), and we have chosen $\mathbf{p}_0(0)$ arbitrarily.

We therefore calculate the paraxial ray associated with this reference ray which satisfies both boundary conditions (6) and (7) to first order. The four required initial conditions of the paraxial ray may be obtained from the following boundary conditions:

(1 + 2) The paraxial ray as well as the central ray goes through the receiver at $\tau = 0$ so that:

$$\delta x_1(0) = 0 \quad \text{and} \quad \delta x_2(0) = 0 \quad (8)$$

(3) From eqs (5) and (8) and taking into account the canonical equation (3), we obtain

$$\mathbf{p}_0(0) \cdot \delta \mathbf{p}(0) = 0 \quad (9)$$

(4) The last expression is less straightforward. Let us denote by τ_i and τ'_i the τ -values of the reference ray and the paraxial ray at their respective intersection points with the edge of the model (Fig. 4). From the values $\mathbf{x}_0(\tau_i)$ and $\mathbf{p}_0(\tau_i)$ of the central ray on the edge of the model, the values $\mathbf{x}(\tau'_i)$ and $\mathbf{p}(\tau'_i)$ of the paraxial ray at the same edge can be derived from a projection. Let \mathbf{n} be the unit vector normal to the edge and let us define $d\mathbf{x} = \mathbf{x}(\tau'_i) - \mathbf{x}_0(\tau_i)$ and $d\mathbf{p} = \mathbf{p}(\tau'_i) - \mathbf{p}_0(\tau_i)$. To first order, the relation between $(d\mathbf{x}, d\mathbf{p})$ and $(\delta \mathbf{x}(\tau_i), \delta \mathbf{p}(\tau_i))$ is (Farra 1990)

$$\begin{aligned} d\mathbf{x} &= \Pi_1 \delta \mathbf{x}(\tau_i) \\ d\mathbf{p} &= \Pi_2 \delta \mathbf{x}(\tau_i) + \delta \mathbf{p}(\tau_i) \end{aligned} \quad (10)$$

where the projection matrices Π_1 and Π_2 are defined by:

$$\Pi_1 = \mathbf{I} - \frac{\nabla_{\mathbf{p}} H \mathbf{n}^T}{\nabla_{\mathbf{p}} H \cdot \mathbf{n}} \quad \text{and} \quad \Pi_2 = \frac{\nabla_{\mathbf{x}} H \mathbf{n}^T}{\nabla_{\mathbf{p}} H \cdot \mathbf{n}} \quad (11)$$

\mathbf{I} is a 2×2 identity matrix; \mathbf{n}^T is the transpose of vector \mathbf{n} so that $(\nabla_{\mathbf{p}} H \mathbf{n}^T)_{ij} = (\nabla_{\mathbf{p}} H)_i \mathbf{n}_j$. All computations are made on the central ray at $(\mathbf{x}_0(\tau_i), \mathbf{p}_0(\tau_i))$. Replacing p_i by $p_{0i} + dp_i$ and $\frac{\partial T_0}{\partial x_i}(\mathbf{x})$ by $\frac{\partial T_0}{\partial x_i}(\mathbf{x}_0 + d\mathbf{x})$ in eq. (7), and expanding $\partial T_0 / \partial x_i$ to first order, we obtain the final condition:

$$\frac{\partial^2 T_0(\mathbf{x}_0)}{\partial x_i^2} dx_i + dp_i = -\frac{\partial T_0(\mathbf{x}_0)}{\partial x_i} - p_{0i} \quad (12)$$

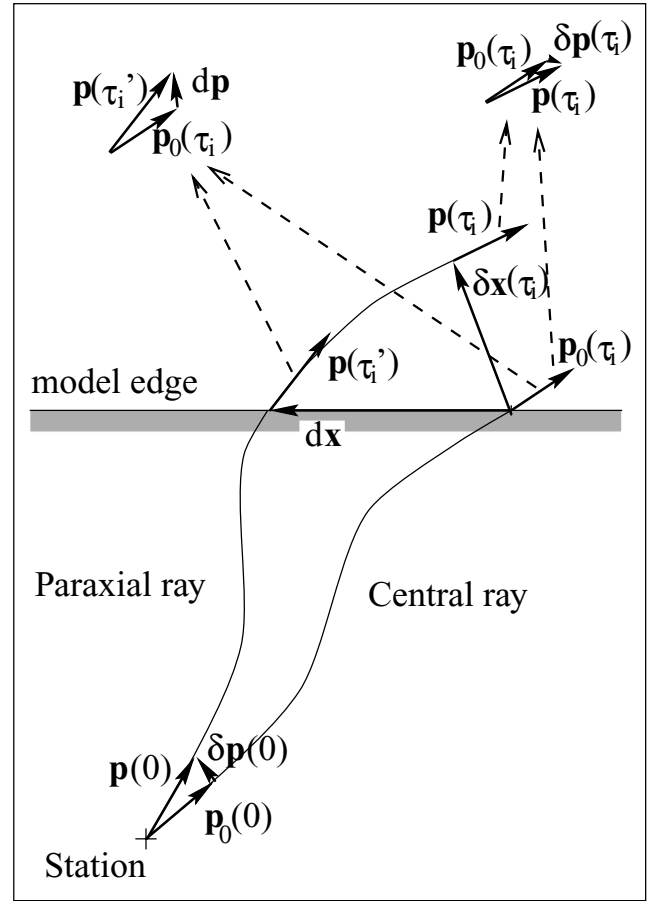


Figure 4. Paraxial ray tracing. \mathbf{p} : slowness vector; \mathbf{x} : position vector; τ : sampling parameter; see main text for the definition of the other terms.

The two components of $\delta \mathbf{p}(0)$ may be obtained from relations (9) and (12), using the paraxial propagator of system (4) between τ_i and 0 and relation (10).

Due to the use of first order approximations, the two-point ray tracing procedure is iterative. We first trace the ray with initial slowness vector $\mathbf{p}_0(0)$, of which the direction is given by the theoretical backazimuth of the wave, to the edge of the model; then the paraxial propagator is computed. The initial conditions of the paraxial ray are determined to satisfy the boundary conditions (8), (9) and (12). Then, a fresh iteration starts with $\mathbf{p}_0(0) + \delta \mathbf{p}(0)$ as the revised initial slowness vector for the reference ray. The process stops when the central ray satisfies the boundary conditions (6) and (7) within predefined convergence criteria.

Finally, the traveltime of the wave at the station is derived by integration along the ray:

$$T = \int_{\text{ray}} \mathbf{p} \cdot \dot{\mathbf{x}} d\tau = \int_{\text{ray}} u^2(\mathbf{x}(\tau)) d\tau \quad (13)$$

4 INVERSE PROBLEM

The data are the arrival times T_i^{obs} measured for each event at each receiver. We denote by T_i^c the traveltimes calculated with the ray tracing procedure explained above, using as input a 2-D velocity model and for each wavefront the arrival time T_0 at the edges of the study region. The inverse problem consists of estimating the model

\mathbf{m} that explains the observed traveltimes in a least-squares sense, that is, that minimizes the least-squares misfit function S defined as:

$$S(\mathbf{m}) = [\mathbf{T}^{\text{obs}} - \mathbf{T}^c(\mathbf{m})]^T \mathbf{C}_T^{-1} [\mathbf{T}^{\text{obs}} - \mathbf{T}^c(\mathbf{m})] + [\mathbf{F}_m - \mathbf{F}(\mathbf{m})]^T \mathbf{C}_F^{-1} [\mathbf{F}_m - \mathbf{F}(\mathbf{m})] \quad (14)$$

where \mathbf{T}^{obs} and $\mathbf{T}^c(\mathbf{m})$ are vectors containing the observed and calculated traveltimes respectively, \mathbf{C}_T is the covariance matrix of observed traveltime data and \mathbf{m} is a vector containing the velocity model, or more precisely the square of the slowness in a grid covering the study area, and the arrival times T_0 of the incoming wavefront at the appropriate edges of the model. To constrain the solution, we introduce *a priori* values \mathbf{F}_m of some parameter combinations $\mathbf{F}(\mathbf{m})$ (see Section 4.3). The model covariance matrix \mathbf{C}_F describes the uncertainties in the *a priori* values \mathbf{F}_m (Farra & Madariaga 1988).

The non-linear least-squares problem (14) can be solved iteratively by the Gauss-Newton method. The functions $\mathbf{T}^c(\mathbf{m})$ and $\mathbf{F}(\mathbf{m})$ in expression (14) are linearized around a current model \mathbf{m}_0 to obtain the quadratic approximation $E(\Delta\mathbf{m})$ of the least-squares misfit function S :

$$E(\Delta\mathbf{m}) = [\Delta\mathbf{T} - \mathbf{A}_T \Delta\mathbf{m}]^T \mathbf{C}_T^{-1} [\Delta\mathbf{T} - \mathbf{A}_T \Delta\mathbf{m}] + [\Delta\mathbf{F}_m - \mathbf{A}_F \Delta\mathbf{m}]^T \mathbf{C}_F^{-1} [\Delta\mathbf{F}_m - \mathbf{A}_F \Delta\mathbf{m}] \quad (15)$$

where $\Delta\mathbf{T} = \mathbf{T}^{\text{obs}} - \mathbf{T}^c(\mathbf{m}_0)$ contains the arrival time residuals, that is, the differences between the measured and the computed arrival times; $\Delta\mathbf{m}$ is the parameter perturbation; and $\mathbf{A}_T = \partial\mathbf{T}^c/\partial\mathbf{m}$ is a matrix containing the partial derivatives of the traveltime with respect to the parameters, also known as Fréchet derivatives (see Section 4.2). The vector $\Delta\mathbf{F}_m = \mathbf{F}_m - \mathbf{F}(\mathbf{m}_0)$ and the matrix $\mathbf{A}_F = \partial\mathbf{F}/\partial\mathbf{m}$ describe additional constraints such as *a priori* information and regularization, which we describe in more detail in Section 4.3. As the data are assumed to be uncorrelated, we set the non-diagonal terms of matrices \mathbf{C}_T and \mathbf{C}_F to zero. The diagonal terms are computed from uncertainties in the traveltime measurements (see Section 6 for estimation on real data) and *a priori* values of the combinations of the parameters $\mathbf{F}(\mathbf{m})$.

Let us introduce the diagonal matrices σ_T and σ_F defined as $\sigma_T^2 = \mathbf{C}_T$ and $\sigma_F^2 = \mathbf{C}_F$ and the matrix \mathbf{D} as $\mathbf{D} = \begin{pmatrix} \sigma_T^{-1} \mathbf{A}_T \\ \sigma_F^{-1} \mathbf{A}_F \end{pmatrix}$. The least-squares solution of the misfit function (15)

$$\Delta\mathbf{m} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \begin{pmatrix} \sigma_T^{-1} \Delta\mathbf{T} \\ \sigma_F^{-1} \Delta\mathbf{F}_m \end{pmatrix} \quad (16)$$

is obtained by singular value decomposition of matrix \mathbf{D} . As the problem is non-linear, the process is repeated until the residuals $\Delta\mathbf{T}$ drop to the level of measurement error. At each iteration the rays are traced using the procedure described in Section 3.

An *a posteriori* error analysis is performed to evaluate a model confidence interval. This estimate is obtained by deriving the *a posteriori* covariance matrix \mathbf{C}'_M (Tarantola 1987):

$$\mathbf{C}'_M = \left(\mathbf{A}_T^T \mathbf{C}_T^{-1} \mathbf{A}_T + \mathbf{A}_F^T \mathbf{C}_F^{-1} \mathbf{A}_F \right)^{-1} \quad (17)$$

The diagonal elements of \mathbf{C}'_M are an estimate of the square of the uncertainty in the model obtained from the inversion. They depend on the array geometry and the number of data used in the inversion rather than on the model parameters.

4.1 Parametrization

Because of the form of the ray tracing equations, the parameters of the model \mathbf{m} are the square of the slowness, and the arrival times of

every incoming wavefront at the edge of the parametrized region. All these functions are modelled by B-splines (de Boor 1978; Bartels *et al.* 1988; Farra & Madariaga 1988).

The square of the slowness is interpolated using third-order 2-D B-splines:

$$u^2(x_1, x_2) = \sum_{i=1}^{N_{x_1}} \sum_{j=1}^{N_{x_2}} u_{ij}^2 B_{u_{ij}}(x_1, x_2) \quad (18)$$

where u_{ij}^2 , $i = 1, \dots, N_{x_1}$, $j = 1, \dots, N_{x_2}$ are the coefficients of B-spline interpolation and $B_{u_{ij}}$ are 2-D spline basis functions.

Each wavefront is described by its arrival time at the two edges of the model first encountered by the wave (Fig. 5). Along each edge, the arrival time is modelled using fourth order B-spline interpolation:

$$T_0(x_k) = \sum_{i=1}^{N_{l_k}} T_{0i} B_{l_i}(x_k) \quad (19)$$

where T_{0i} , $i = 1, \dots, N_{l_k}$, are the spline coefficients and B_{l_i} are the spline basis functions.

4.2 Fréchet derivatives of the arrival time

To solve the inverse problem, we need to compute the matrix \mathbf{A}_T of eq. (15) containing the partial derivatives of traveltimes. This requires us to compute the variation in arrival time due to velocity perturbations, and the variation in arrival time due to perturbation of the wavefronts, that is, the variation in the T_0 parameters.

Let us first calculate the Fréchet derivatives with respect to the square of the slowness. In the reference medium of slowness squared u_{ref}^2 we denote by R_{ref} the ray between station S and point M on the edge of the model. If the square of the slowness is modified from u_{ref}^2 to $u^2 = u_{\text{ref}}^2 + \Delta u^2$, we can trace the new ray R joining S and M in the perturbed model. This new ray may not satisfy the boundary condition (7): coherence between the slowness vector of the ray and the gradient of T_0 . We denote by R' the ray of the perturbed medium which satisfies the boundary conditions (6–7). This ray intersects the edge of the model at M' . Subscript ref indicates quantities obtained in the reference medium.

In the perturbed medium, the arrival time $T(S)$ is:

$$T(S) = T_0(M') + \int_{R'} u^2 d\tau \quad (20)$$

$$= T_0(M') + \int_R u^2 d\tau + \mathbf{p} \cdot \overrightarrow{MM'} \quad (21)$$

where $\mathbf{p} = -\nabla_x T(M)$ is the slowness vector at the edge of the model (the negative sign is due to our convention of computing the ray from the station to the edge of the box).

Fermat's principle implies that:

$$\int_R u^2 d\tau = \int_{R_{\text{ref}}} u_{\text{ref}}^2 d\tau + \int_{R_{\text{ref}}} \Delta u^2 d\tau \quad (22)$$

To first order, we also know that $T_0(M') = T_0(M) + \nabla_x T \cdot \overrightarrow{MM'}$. The relation between the arrival times at the station in the reference medium and in the perturbed medium is:

$$T(S) = T_0(M) + \int_{R_{\text{ref}}} u_{\text{ref}}^2 d\tau + \int_{R_{\text{ref}}} \Delta u^2 d\tau \quad (23)$$

$$= T_{\text{ref}}(S) + \int_{R_{\text{ref}}} \Delta u^2 d\tau \quad (24)$$

Therefore, using parametrization (18), we derive the Fréchet derivatives with respect to the square of the slowness parameters:

$$\frac{\partial T(S)}{\partial u_{ij}^2} = \int_{R_{\text{ref}}} B_{u_{ij}}(x_1, x_2) d\tau \quad (25)$$

Let us now calculate the Fréchet derivatives with respect to the arrival time T_0 of the wave at the edge of the modelled region. For the reference arrival times $T_{0\text{ref}}$ and the square of the slowness model u^2 , we denote by R the ray between the station S and the point M on the edge of the model which satisfies the boundary condition (7). Let us consider a perturbation ΔT_0 , so that $T_0 = T_{0\text{ref}} + \Delta T_0$, and the ray R' satisfying conditions (6–7) in the perturbed model. The ray R' is traced from the station S to the point M' on the edge of the study region. Starting from eq. (21) and from $T_0(M') = T_{0\text{ref}}(M) + \Delta T_0(M) + \nabla_x T \cdot \overrightarrow{MM'}$, we obtain

$$T(S) = T_{0\text{ref}}(M) + \Delta T_0(M) + \int_R u^2 d\tau \quad (26)$$

$$= T_{\text{ref}}(S) + \Delta T_0(M) \quad (27)$$

Using parametrization (19), we obtain the Fréchet derivatives with respect to the T_0 parameters:

$$\frac{\partial T(S)}{\partial T_0} = B_i(x(M)) \quad (28)$$

4.3 *A priori* information

Additional *a priori* information is usually needed to make the inverse problem non-singular. To limit the roughness of the solution, we can minimize the value of the first and second order derivatives of the square of the slowness by introducing a function $\mathbf{F}(\mathbf{m})$ in eq. (14). This function can be used to set the first and second order derivatives equal to zero with given uncertainties, denoted by σ_{F_1} and σ_{F_2} respectively. In practice, the constraints are introduced through the first and second finite differences of the B -spline parameters u_{ij}^2 .

We add two constraints on the shape of the wavefront. Each wavefront is modelled by the arrival time T_0 along two boundaries of the region, the two sets of arrival times being treated independently (Fig. 5). We therefore impose the condition that the values at the point shared by the two lines at the corner of the region are identical, to within σ_{F_3} , an amount equal to the data uncertainty. We also impose a certain degree of smoothness on the incoming wavefront. To obtain this, the derivative of the arrival time along the edge should not be very different from that given by the theoretical propagation angle ϕ :

$$\frac{\partial T_0}{\partial x_1} = -\frac{\sin \phi}{c} \quad \text{or} \quad \frac{\partial T_0}{\partial x_2} = \frac{\cos \phi}{c} \quad (29)$$

according to the edge considered. In (29), c is the wave speed at the point considered and ϕ is the theoretical backazimuth or a measured propagation direction. The x_1 and x_2 axes are oriented toward the East and South, respectively. We denote by σ_{F_4} the corresponding *a priori* uncertainty. σ_{F_4} therefore controls the upper limit of allowed variation in the local backazimuth of a wavefront, as compared to a plane wave. This upper limit depends strongly on the theoretical backazimuth of the incident wave. For $\sigma_{F_4} = 0.05 \text{ s km}^{-1}$ used in the inversions presented here, the upper limit is at least 10° . An example of curved wavefronts is shown on Fig. 8 and discussed in more detail in Section 5.2.

The strength of the constraints, related to σ_{F_1} , σ_{F_2} , σ_{F_3} and σ_{F_4} , is chosen as a compromise between stability of the inversion and data misfit.

5 NUMERICAL EXAMPLES

The inversion was applied to several synthetic examples. The station geometry is that of the SVEKALAPKO experiment (Fig. 5). This 2-D arrangement is significantly better than that usually used for 2-D body-wave tomography, as the receivers are distributed throughout the modelled area and not only along one edge. The sources are a set of 36 theoretical events evenly distributed in azimuth. We want to obtain the phase velocity structure in the rectangular region containing this network and shown in Fig. 5. In the first part of this section (5.1–5.3) we use arrival times at the stations calculated by ray tracing to study the behaviour of the inversion procedure in ideal conditions. In Section 5.4 we use full waveform modelling with and without noise to approach more realistic conditions.

At every run of the inversion the data misfit $\|\Delta \mathbf{T}\|$ is computed, that is the root mean square of the time residuals. As we know the true velocity model, we also compute the “model misfit” Δc as the root mean square of the difference in phase velocity c . To take into account the lack of resolution in some areas, the differences in velocity are weighted by the inverse of the *a posteriori* error in the velocity σ_c .

$$\Delta c(\mathbf{m}) = \sqrt{\frac{1}{\sum 1/\sigma_c^2} \sum \left(\frac{c_{\text{true}} - c(\mathbf{m})}{\sigma_c} \right)^2} \quad (30)$$

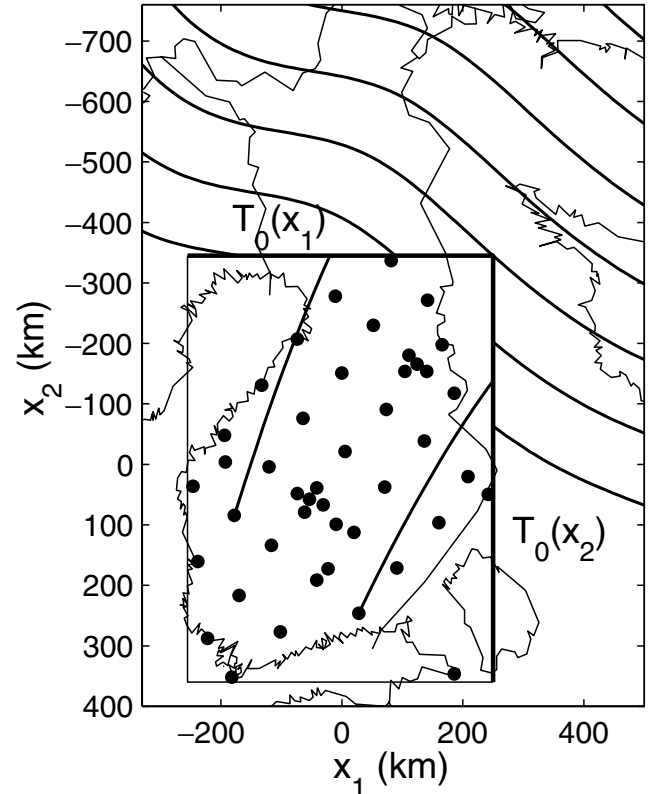


Figure 5. Setting of the problem for the SVEKALAPKO deep seismic experiment. Position of the broad-band receivers are the black dots, the rectangular box is the limit used in the computations. An example of a curved wavefront propagating from the north-east outside the model is shown together with two rays traced from the edge of the model to two stations. The shape of the wavefront is modelled by the arrival time of the wave on the box: $T_0(x_1)$ and $T_0(x_2)$.

The *a posteriori* error in the velocity is obtained from that of the square of the slowness, σ_{u^2} , using the relation:

$$\sigma_c = c_{\text{true}}^3 \sigma_{u^2} / 2 \quad (31)$$

σ_{u^2} is the square root of the diagonal elements of matrix C'_M given by (17) corresponding to the square of the slowness parameters. In (30), the summation is made with points regularly distributed every 5 km inside the box.

The parameters of the model being of two types with different orders of magnitude ($0.1 \text{ s}^2 \text{ km}^{-2}$ for the square of the slowness and 100 s for the arrival times), we apply a normalization in all numerical examples and used as parameters in the inversion procedure the square of the slowness divided by 0.1 and the wavefront arrival times divided by 100.

5.1 Sensitivity test

The checker-board test cannot reasonably be applied to a ray tracing method as the condition of smooth structure is not respected. We therefore considered a 2-D phase velocity model of a *B*-spline interpolated checker-board. The cell width is 150 km and the maximum variation of the phase velocity is 5 per cent between 4 and 4.2 km s^{-1} (Fig. 6a). The synthetic data set consisted of the arrival times of 36 wavefronts at the receivers, computed using the ray tracing method described in Section 3. In this example, the wavefronts incident upon the structure were plane.

The starting model of the inversion procedure is homogeneous with a velocity of 4.1 km s^{-1} . The initial incoming wavefronts are plane. Rays are traced through this initial model and traveltimes computed; the root mean square of the arrival time residuals is $\|\Delta \mathbf{T}\| = 0.63 \text{ s}$.

The inverse problem consisted of evaluating the velocity field and the 36 wavefronts. The square of the slowness is represented by *B*-spline interpolation with 192 coefficients (12×16) corresponding to a regular mesh of gridpoints distributed every 50 km (about half of the mean distance between the stations). The arrival times T_0 of the 36 incident wavefronts are represented by *B*-spline interpolation with 11 and 15 coefficients (corresponding to grid points regularly distributed every 50 km), along the West-East and North-South edges of the box, that is 936 parameters. The model is then fully described by 1128 parameters. We assume an uncertainty of 1 s in the traveltime data.

If we invert the times residuals without *a priori* constraints, the velocity model obtained shows very high amplitude anomalies in the corners with a weak station coverage. We therefore introduced *a priori* information as described in Section 4.3, using the following values: $\sigma_{F_1} = 3 \times 10^{-4} \text{ s}^2 \text{ km}^{-3}$ (5 per cent variation of the velocity is allowed over a distance of 20 km), $\sigma_{F_2} = 6 \times 10^{-6} \text{ s}^2 \text{ km}^{-4}$ ($\sigma_{F_2} = \sigma_{F_1}/50$), $\sigma_{F_3} = 1 \text{ s}$, and $\sigma_{F_4} = 5 \times 10^{-2} \text{ s km}^{-1}$ (at least 10° variation is allowed in the propagation direction). The inversion process is stopped when the traveltime misfit no longer decreases.

Table 1 shows the number of data and parameters, and the initial and final values of the data and model misfits for comparison between all numerical examples. The final model (Fig. 6b) was obtained after 2 iterations, with final data misfit $\|\Delta \mathbf{T}\| = 0.34 \text{ s}$. The model misfit defined by equation (30) is 0.02 km s^{-1} (0.5 per cent of the velocity). The *a posteriori* error in the velocity model σ_c , estimated by eq. (30), shown in Fig. 6(c) has a mean of 2.6 per cent of the true velocity. It is almost constant in the central part of the box where the stations are located, with a value of approximately 2 per cent.

Table 1. Numerical values that characterize the inversions. *N*: number of time residuals; *M*: total number of lines in the inversion matrix (data and constraints); *P*: number of parameters. For definitions of data and model misfits see main text. *I*: number of iterations; (*) regularization constraints used in the first iteration(s). For values of the assumed error in the data and of the constraints see main text. cb: checker-board test; npwf: homogeneous velocity model with non-planar incoming wavefronts; moho: input velocity model computed from the Moho-depth map of Luosto (1990) for a 35 s fundamental-mode Rayleigh wave; IBEM_XXs: no-noise IBEM synthetic seismograms at XX seconds, IBEM_az_XXs: no-noise IBEM synthetic data set with a bad azimuthal coverage, IBEM_n_XXs: IBEM seismograms with noise; SVEKA_60s: real seismograms of the SVEKALAPKO seismic tomography experiment at 60 s period.

	<i>N</i>	<i>M</i>	<i>P</i>	data misfit		model misfit		<i>I</i>
				initial	final	initial	final	
				s		$\times 10^{-2} \text{ km s}^{-1}$		
cb	1573	3161	1128	0.63	0.34	5.38	2.04	2
npwf	1434	3022	1128	3.18	0.29	0	3.96	3
moho	1582	3170	1128	0.75	0.02	7.50	0.36	2
IBEM_25s	1320	2758	972	1.35	0.11	7.91	0.75	3
IBEM_33s	1320	2758	972	1.51	0.20	8.90	2.40	3*
IBEM_50s	1320	2758	972	0.88	0.25	4.90	0.90	3*
IBEM_az_33s	1320	2758	972	1.63	0.16	8.90	2.52	2
IBEM_n_25s	1320	2758	972	1.42	0.39	7.91	1.36	2
IBEM_n_33s	1320	2758	972	1.59	0.55	8.90	2.13	3*
IBEM_n_50s	1320	2758	972	1.38	1.06	4.90	7.62	3*
SVEKA_60s	1390	2942	1104	3.04	1.21	–	–	4*

When the cell size of the input model is smaller than 150 km, the data misfit hardly decreases. For 100 km cell size the checker-board is reconstructed but with smaller velocity variations than in the input model. For models with cell size larger than 150 km, the inversion result is closer to the true velocity model than in the case with 150 km cell size presented above.

The sensitivity tests tell us that structures smaller than approximately 150 km cannot really be imaged with the SVEKALAPKO experiment, even with a perfect set of recorded events. This is related to the inter-station distance in the SVEKALAPKO array, which is approximately 100 km. The north-western and south-eastern corners of the box are poorly constrained due to the lack of receivers in these regions.

5.2 Non-planar incoming wavefronts

In the following example, the incident wavefronts were randomly, but smoothly, distorted. We are not expecting backazimuth anomalies larger than 5° in Fennoscandia, considering that the region is relatively homogeneous and taking into account the azimuth anomalies measured by Levshin & Berteussen (1979) over the Barents Sea. We therefore modified the backazimuth of our 36 wavefronts with a random rotation of 5° maximum. Friederich *et al.* (1994) also observed a local curvature of the wave-fields propagating over Germany. Onto the constant rotation we superimposed random small scale variations of 30° maximum within 100 km along the wavefront. The input velocity model was homogeneous with a value of 4.0 km s^{-1} . The number of data inverted is 1434 (Table 1), which should be compared with the maximum possible of 1584 corresponding to 36 events \times 44 receivers. The missing data correspond to rays that the ray tracing procedure was not able to compute because of the perturbation of the wavefront.

The starting velocity model for the inversion was homogeneous with a velocity of 4.0 km s^{-1} . The initial wavefronts are plane and

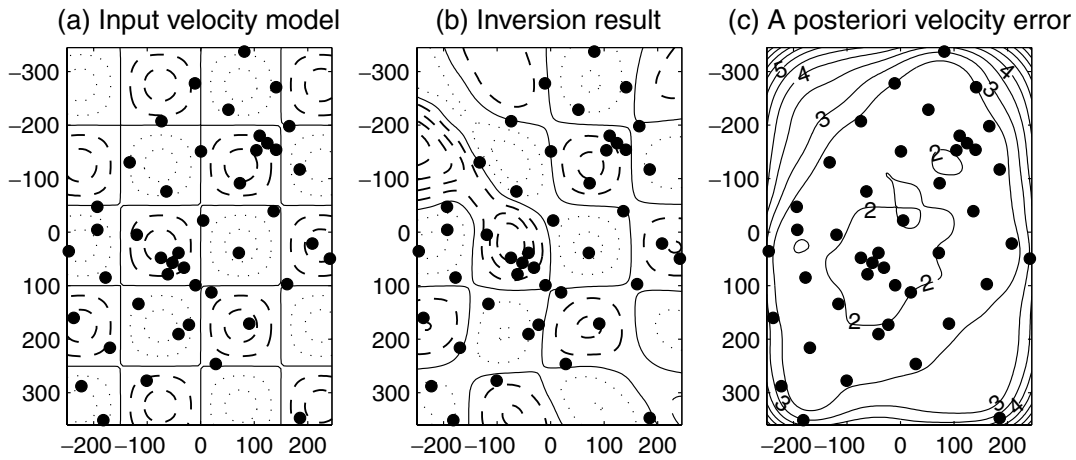


Figure 6. Sensitivity test. (a) true velocity model used to compute the arrival times. (b) velocity model obtained after 2 iterations of the inversion. Contours are every 0.04 km s^{-1} , solid line for 4.1 km s^{-1} , dotted lines for lower velocities, dashed line for larger velocities. (c) *A posteriori* velocity error as a percentage of the true velocity. Station positions are given by the black dots.

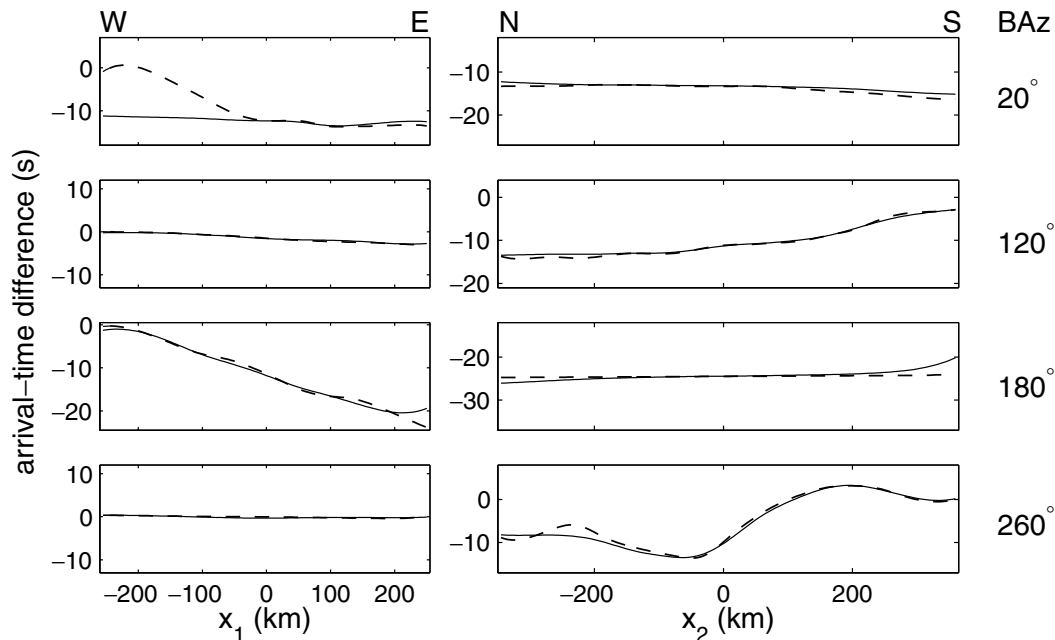


Figure 7. Four examples of non-planar wavefronts. Arrival time differences between the curved waves (input wavefront with a dashed line and inversion result with a solid line) and the plane wave coming from the theoretical backazimuth. Left panels: edge parallel to the x_1 axis; Right panels: edge parallel to the x_2 axis

incident from the theoretical backazimuth, that is, without the constant rotation. Rotation of the wavefronts as compared to the theoretical backazimuth used in the starting model leads to systematic bias of the time residuals. To obtain a minimal initial estimate of the data misfit for Table 1 we computed the root mean square of the residuals by adding a constant to every arrival time of each wavefront, so that the data misfit is minimal: $\|\Delta \mathbf{T}\| = 3.18 \text{ s}$. The parametrization used is the same as in the previous example, so the total number of parameters is 1128. The assumed data error is 1 s. The constraints used in the inversion are $\sigma_{F_1} = 3 \times 10^{-4} \text{ s}^2 \text{ km}^{-3}$, $\sigma_{F_2} = 6 \times 10^{-6} \text{ s}^2 \text{ km}^{-4}$, $\sigma_{F_3} = 1 \text{ s}$, and $\sigma_{F_4} = 5 \times 10^{-2} \text{ s}^{-1} \text{ km}$.

After 3 iterations, the final model was obtained. The data misfit is $\|\Delta \mathbf{T}\| = 0.29 \text{ s}$. The model misfit is 0.040 km s^{-1} , 1 per cent of the input phase velocity. The misfit is concentrated in the unresolved areas; if we compute the model misfit on a restricted area containing

the network, it falls below 0.6 per cent. The *a posteriori* error has the same shape as in the preceding example, with a mean value of 2.6 per cent.

Four examples of a wavefront are shown in Fig. 7. The disagreement between input and reconstructed wavefronts can be large but such misfit generally occurs in poorly resolved areas, such as the north-western corner showing a 10 s difference in the first panel of Fig. 7 for the backazimuth of 20° . Fig. 8 gives an example of a curved wavefront propagating across the station network, interpolated from the inversion result.

The mean of the *a posteriori* error on the arrival time of the wavefronts T_0 is 2.2 s. The maximum estimated error for one wavefront occurs either at the corner where the station coverage is weak, or for an edge almost parallel to the propagation direction at the end where the wavefront leaves the study region.

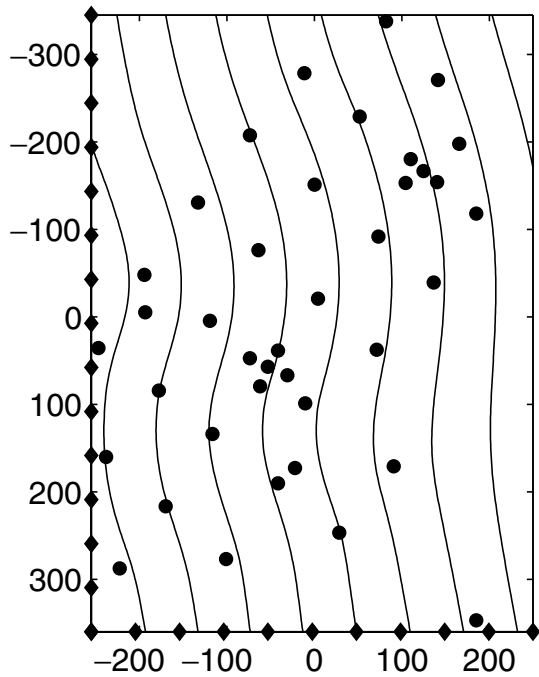


Figure 8. Propagation of the curved wavefront with theoretical backazimuth 260° through the station network. Propagation time is interpolated from the arrival time at the receivers (dots) and at two edges of the study region (diamonds). Wavefronts 15 s apart are plotted as solid lines.

The wavefront smoothness imposed in the inversion was stronger than the actual smoothness of the input wavefronts. We therefore compared the results using different values for σ_{F_4} . As a general trend, with weakening constraint on the wavefront smoothness, the data misfit increases but the model misfit decreases. For example with $\sigma_{F_4} = 0.13 \text{ s km}^{-1}$ (30° variation in the backazimuth angle), the data misfit is $\|\Delta\mathbf{T}\| = 0.46 \text{ s}$, and the model misfit is 0.020 km s^{-1} for the whole study region and 0.016 km s^{-1} in the well resolved area (0.4 per cent). On the other hand, increasing roughness of the input wavefront (small scale variations up to 40° over 100 km) deteriorates the inversion result. With $\sigma_{F_4} = 5 \times 10^{-2}$, the data misfit is 0.75 s, the model misfit is 0.053 km s^{-1} , and 0.042 km s^{-1} below the stations. Increasing σ_{F_4} does not improve the result. The curvature of 40° over 100 km is the maximum that our inversion technique is able to manage.

In real cases, the measurement of the real propagation direction in small arrays at two different locations in the network may allow us to evaluate the global rotation of the wavefront and to eliminate events with excessively perturbed wavefronts.

5.3 Variable Moho depths model

In this example, we use a more realistic velocity model derived from the Moho depth map of Luosto (1990). The physical model is composed of a single layer (the crust) over a half space (the mantle). The crust is defined by P - and S -velocities $\alpha_1 = 6.5 \text{ km s}^{-1}$, $\beta_1 = 4.2 \text{ km s}^{-1}$ and density $\rho_1 = 2700 \text{ kg m}^{-3}$, and the mantle by $\alpha_2 = 8.1 \text{ km s}^{-1}$, $\beta_2 = 5.35 \text{ km s}^{-1}$ and $\rho_2 = 3100 \text{ kg m}^{-3}$. The thickness of the crust varies from 41.7 km to 60.6 km, as derived from the interpolated Moho-depth map of Luosto (1990). At each gridpoint of the surface, the phase velocity dispersion curve is computed using the software edited by Herrmann (1985) with the 1-D model corresponding to the structure immediately below the

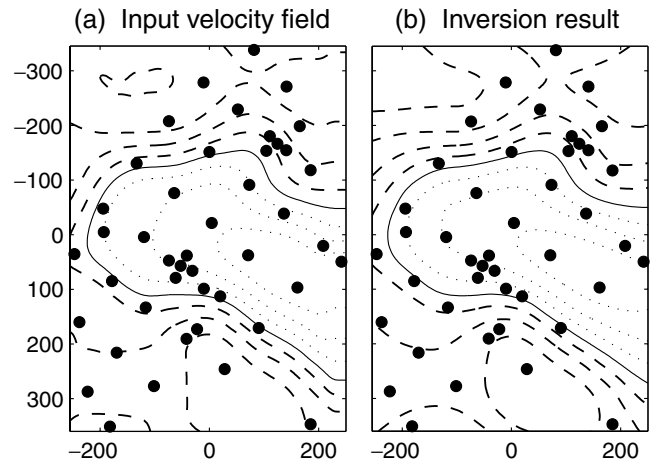


Figure 9. Inversion based on the phase velocity model derived from Luosto (1990) Moho depth map, for the Rayleigh fundamental-mode at 35 s period. (a) true velocity field. (b) inversion result after 2 iterations. Contours are every 0.03 km s^{-1} , solid line for 4.26 km s^{-1} , dotted lines for lower velocities, dashed line for larger velocities. Station positions are given by the black dots.

location under consideration. We can therefore calculate a phase velocity map for each frequency. These maps can be used as the input velocity field to compute traveltimes with our ray tracing method. As an example we used the phase velocity map corresponding to the fundamental-mode Rayleigh wave of 35 s period (Fig. 9a). At this period the phase velocity varies between 4.15 and 4.38 km s^{-1} . In reality the expected variations are lower, as the very thick crust is associated with high shear-velocities and densities in the lower crust. The sources considered were again 36 plane waves evenly distributed in azimuth.

The starting velocity model for the inversion process is homogeneous with a velocity of 4.26 km s^{-1} . The initial incoming wavefronts are plane waves. The root mean square of the traveltime residuals is $\|\Delta\mathbf{T}\| = 0.75 \text{ s}$ and the assumed error in the data is 1 s.

The parameters of the inverse problem are the 192 squares of the slowness parameters and the 936 arrival time parameters. The constraints are $\sigma_{F_1} = 3 \times 10^{-4} \text{ s}^2 \text{ km}^{-3}$, $\sigma_{F_2} = 6 \times 10^{-6} \text{ s}^2 \text{ km}^{-4}$, $\sigma_{F_3} = 1 \text{ s}$, and $\sigma_{F_4} = 5 \times 10^{-2} \text{ s km}^{-1}$. Fig. 9(b) shows the velocity model generated after 2 iterations of the inversion process. The characteristics of the inversion can be found in Table 1. The final rms of the time residuals is $\|\Delta\mathbf{T}\| = 0.02 \text{ s}$. The model misfit is 0.004 km s^{-1} (0.08 per cent of the input velocity). The *a posteriori* error on the velocity field has an average of 2.8 per cent.

The results obtained with this velocity model are particularly good and encouraging for the future interpretation of the SVEKALAPKO data.

5.4 Data computed with Indirect Boundary Element Method

Up to now all the examples were carried out using arrival times computed with the ray tracing method that is also used in the inversion. To study the performance of the inversion using more realistic input data, we computed synthetic seismograms using the Indirect Boundary Element Method (Pedersen *et al.* 1996). This method computes the complete wave-field caused by a plane surface wave obliquely incident upon a 2-D structure. The incident wave can have any backazimuth, but the computation becomes unstable when the

propagation direction is close to the axis of the structure (Pedersen *et al.* 1994). We therefore kept only 30 events, eliminating three azimuths around 90° and three around 270° . The modelled source is a fundamental-mode Rayleigh wave, with a Ricker wavelet source function centered on 0.04 Hz frequency. The structure is the simplified cross-section of the Moho-depth map from Luosto (1990) described in the beginning of Section 2 and shown in Fig. 1. As the model does not take into account the high velocity at the base of the thickened crust, the sharp velocity contrast will produce stronger diffractions than we expect in a more realistic model of the area.

For each event, we measured on the synthetic seismograms the time delays between a reference station, chosen as the first station reached by the wave, and all the other receivers. Time delays were found through the two-step algorithm described in Section 2.1. The assumed data uncertainty is 1 s.

Figs 10(a), (d) and (g) display the theoretical phase velocity models obtained for three selected periods (25, 33 and 50 s). This theoretical velocity is the phase velocity of the fundamental local mode, computed using the same procedure as in Section 2.1. However, because IBEM uses the physical 2-D model of Fig. 1 as input, and the theoretical phase velocity maps shown in Figs 10(a), (d) and (g) are not used in the calculation of the synthetic data set, it should not be considered as the ‘true’ velocity. The average theoretical phase velocity is 4.05, 4.25 and 4.46 km s⁻¹ for 25, 33 and 50 s periods respectively, resulting in wavelengths of approximately 100, 140 and 225 km.

The starting velocity models are homogeneous with velocity 4.02, 4.23 and 4.44 km s⁻¹ respectively. The initial root mean square of the arrival time residuals is between 0.88 and 1.51 s. The inverse problem consists of evaluating the 192 spline coefficients for the velocity model, and the 780 T_0 parameters describing the 30 wavefronts, that is a total of 972 parameters. The data and model misfits are given in Table 1. The final values of the constraints are $\sigma_{F_1} = 3 \times 10^{-4}$ s² km⁻³, $\sigma_{F_2} = 6 \times 10^{-6}$ s² km⁻⁴, $\sigma_{F_3} = 1$ s, and $\sigma_{F_4} = 5 \times 10^{-2}$ s km⁻¹. However, with the data set obtained for 33 and 50 s period, the inversion did not converge toward a model giving a lower data misfit. We therefore used regularization constraints for the first two iterations of the inversion process, that is, the smoothness criteria on the velocity field (for both periods) and on the wavefront (for the 50 s period data set only) were very strong in the first two iterations and relaxed for the third.

The results of the inversions are shown in Figs 10(b), (e) and (h). For the three frequencies, the low velocity area is reconstructed by the inversion. However, the minimum value obtained is larger than the theoretical value for the three periods selected. This is the result of the smoothing of the structure which can be expected due to the wavelength (see Section 2). The final model misfits lie between 0.008 and 0.024 km s⁻¹, the largest difference in velocity occurs either in the low velocity area where the smoothing has the strongest effect, or in the area with no stations. The average *a posteriori* error on the velocity field is 2.7, 3.0, and 3.2 per cent of the theoretical velocity respectively for 25, 33 and 50 s period.

Another problem associated with a real data set would be the azimuthal coverage. We used a data set composed of arrival times for 22 wavefronts out of the 30 wavefronts computed, for the 33 s period model. The 22 wavefronts chosen are based on the azimuthal coverage obtained by Friederich (1998) in Germany.

The starting model has a homogeneous phase velocity of 4.23 km s⁻¹. The initial data misfit is $\|\Delta\mathbf{T}\| = 1.6$ s. The final model is obtained after two iterations. The final data misfit is 0.16 s, the final model misfit is 0.025 km s⁻¹ (0.6 per cent of the input velocity). The average *a posteriori* velocity error is 3.1 per cent.

The inversion result is not really disturbed by the lack of data from the south. This is a consequence of the good station coverage of the SVEKALAPKO network.

We then added random noise to the synthetic seismograms. The measured signal-to-noise ratios are approximately 3.5, 3.7, and 2.8 for 25, 33 and 50 s period. Figs 10(c), (f) and (i) display the inversion results obtained with the data set measured on the noisy signals. Values of the misfits are given in Table 1. Regularization constraints on the velocity smoothness and on the wavefront smoothness were used for inversion of the data set obtained with 33 and 50 s period.

With a signal-to-noise ratio larger than 3.5, the inversion converges toward a solution equivalent to that obtained without noise. With a signal to noise ratio smaller than 3, the true velocity field is not recovered.

The tests using IBEM-computed synthetic seismograms ratify both the technique of time delay measurement on the seismograms and the inversion method. In the case of simple and relatively smooth structures, ray theory can succeed in imaging heterogeneities of scale close to the wavelength even though velocity variations may be underestimated. Due to the good station coverage, the lack of data incident from some backazimuths is not a problem for the inversion. The minimum signal to noise ratio allowed in data is approximately 3.5.

6 PRELIMINARY INVERSION OF THE SVEKALAPKO DATA

Finally we tested our method on real data. This section does not aim to solve the lithospheric structure of eastern Fennoscandia, but rather to demonstrate that the method can be applied successfully to real seismograms. For this preliminary inversion we use a subset of 48 high quality events (see Fig. 11), carrying out the analysis on the fundamental-mode Rayleigh wave at 60 s period. The azimuthal distribution is not as regular as in the synthetic tests but the corresponding *a posteriori* error is not very different from that obtained in the synthetic tests.

The vertical component seismograms were first filtered in the time-frequency domain using the apparent group velocity curves (Lander & Levshin 1989). For each station couple we measured the time delay using Wiener filtering on the filtered signals. Then we chose a reference station for each event, the first station encountered by the wave, unless its record was too noisy. For each station i we measured the time delay to the reference station as the median of $N - 1$ time delays, N being the number of coherent records. One delay was measured directly between the reference station and station i . The $N - 2$ others were obtained by the sum (or subtraction) of the time delay between the reference station and station j , and the time delay between station j and station i . The average standard deviation of approximately 2 s was used as data uncertainty. The number of time measurements kept for the inversion was 1390, that is 29 data per event on average.

As the number of data per event was somewhat lower than in the numerical tests, we decreased the number of spline coefficients representing the wave fronts to 8 and 11 for the West-East and North-South edges respectively (points distributed every 70 km). The number of coefficients representing the velocity model was kept to 192. The model is fully described by 1104 parameters.

The initial model for the inversion has a homogeneous velocity of 4.13 km s⁻¹. This value was obtained by fitting one dispersion curve to all the observed time delays. The sources considered were 48 plane waves coming from the backazimuths computed for a radially symmetric Earth. In the initial model the origin time for each event

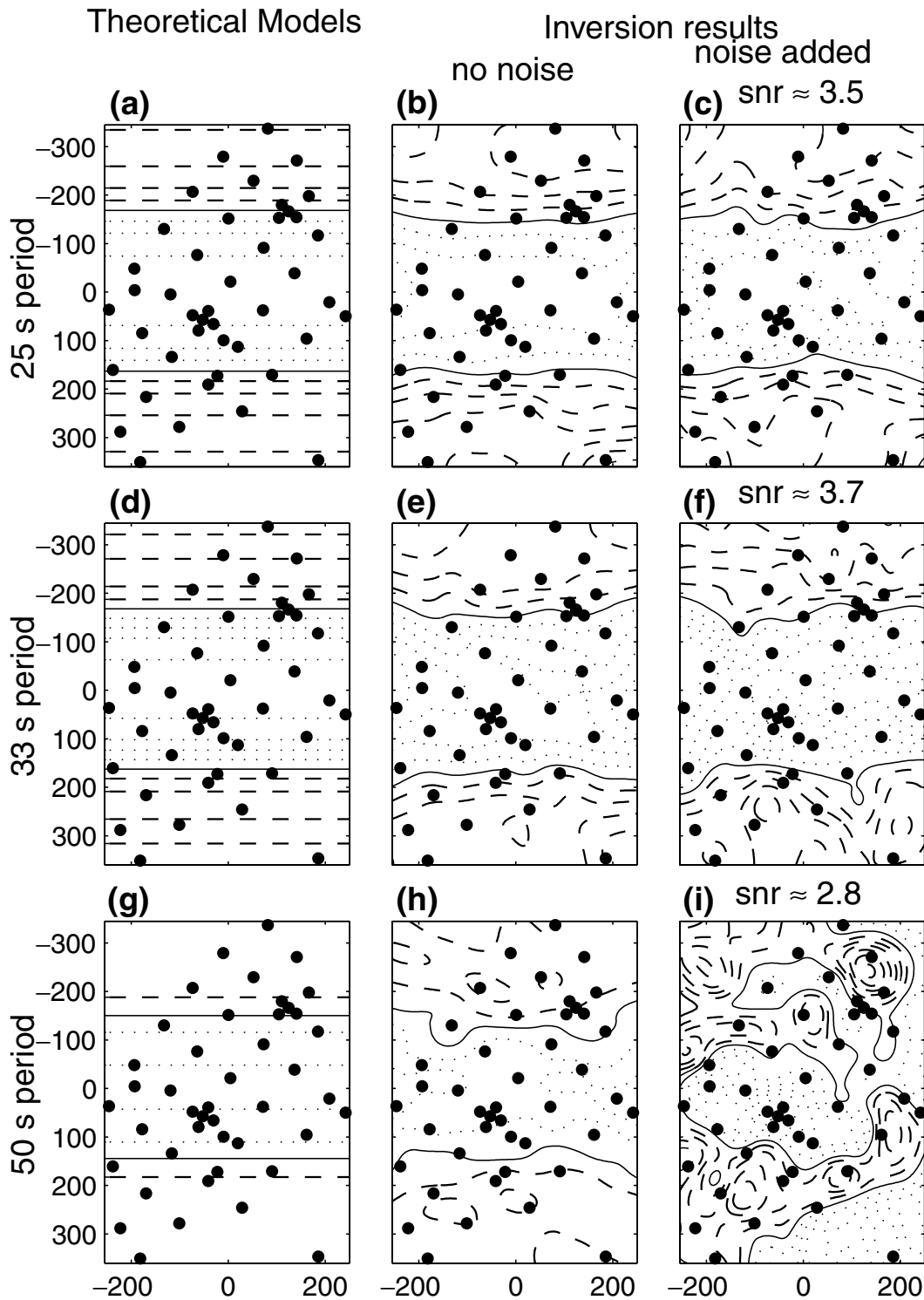


Figure 10. Inversion of time delays measured on IBEM synthetic seismograms. On each plot the contours are every 0.03 km s^{-1} , dotted lines are for lower velocities, dashed line for larger velocities. (a) (b) (c) model for 25 s period, solid line for 4.02 km s^{-1} . (d) (e) (f) 33 s period, solid line for 4.23 km s^{-1} . (g) (h) (i) 50 s period, solid line for 4.44 km s^{-1} . (a) (d) (g) theoretical local phase velocities. (b) (e) (h) models obtained from inversion of signals without noise. (c) (f) (i) models obtained from inversion of the noisy signals; snr: signal to noise ratio. Black dots give the station locations.

is situated at the corner of the study area, whereas in the real data it is located at the reference stations. This constant time shift, which can be as high as 30 s, makes the inversion unstable. We therefore used strong regularization constraints on the first iteration of the inversion process, so that the velocity model stays constant and the

wave fronts remains plane but are shifted in time. In the second iteration the constraints were lowered to $\sigma_{F_1} = 3 \times 10^{-4} \text{ s}^2 \text{ km}^{-3}$, $\sigma_{F_2} = 6 \times 10^{-6} \text{ s}^2 \text{ km}^{-4}$, $\sigma_{F_3} = 2 \text{ s}$, and $\sigma_{F_4} = 5 \times 10^{-2} \text{ s km}^{-1}$.

The time shift also influences the estimation of the initial data misfit, which we addressed by adding a constant time to each event so

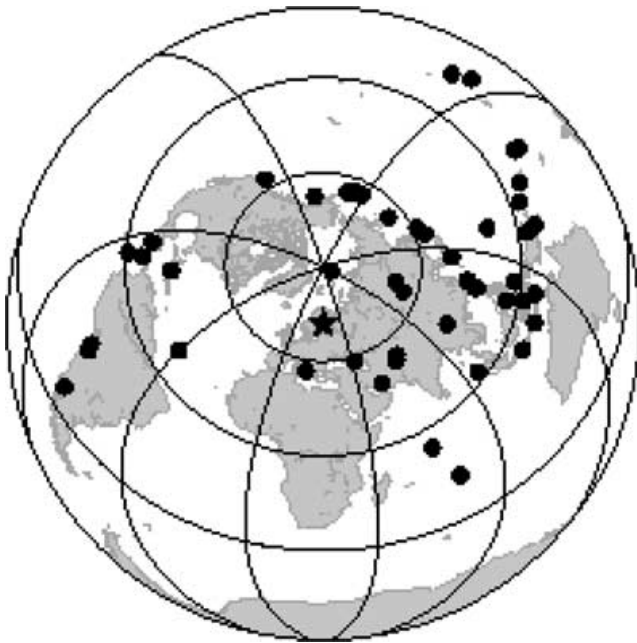


Figure 11. Distribution of the 48 events used in the inversion. Black dots: epicenters; black star: center of the network of stations.

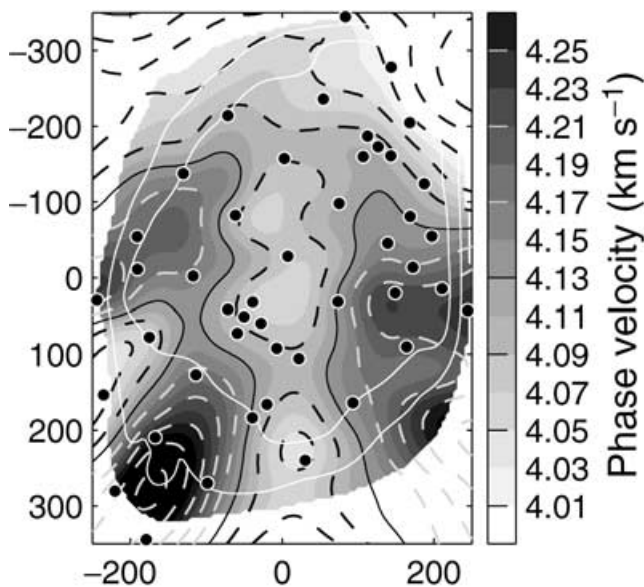


Figure 12. Phase velocity of the fundamental-mode Rayleigh wave at 60 s period, as retrieved from inversion of the time residuals measured on real seismograms from the SVEKALAPKO seismic tomography experiment. Gray levels are inside the limit of a *posteriori* error on the velocity $\sigma_c = 4$ per cent; the thick white contours corresponds to $\sigma_c = 3.5$ per cent and $\sigma_c = 3$ per cent.

that the root mean square of the time residual is minimum: $\|\Delta \mathbf{T}\| = 3.04$ s.

The final model was obtained after 4 iterations, with a final data misfit of 1.21 s (data uncertainty 2 s). Fig. 12 shows the phase velocity model obtained. The average *a posteriori* error in the velocity model is 3.4 per cent, and if we consider only the part of the model inside the 3.5 per cent contour then the mean *a posteriori* error decreases to 2.9 per cent. This is most likely an overestimate of the real error. We will use a boot-strap approach to estimate the er-

ror on the velocity field obtained from the real data set for further interpretation.

The sampling depth of the 60 s fundamental-mode Rayleigh wave is around 80 km. As the crustal thickness reaches 60 km in some areas, this phase velocity map is therefore influenced by the crustal structure. Comparing our map to the compilation of known Moho depths presented by Luosto (1997), shows that the deep crust to the west is associated with relatively low velocities, while the deep crust to the east is associated with velocities larger than the mean. However, this is also the place where the high velocity layer in the lower crust defined in Luosto (1997) is the thickest. A full interpretation of the data recorded during the deep seismic SVEKALAPKO experiment is our next planned investigation.

Even though we need to refine our inversion of real data, in particular for a better estimate of the actual error in the velocity model, our method is reliable when applied to real seismograms.

7 CONCLUSION

We present a method for tomographic inversion of teleseismic surface wave phase information based on 2-D ray theory. The inversion of the time delays is only the first step of a longer process as the dispersion curves obtained at each grid point must subsequently be inverted to obtain a 3-D shear-velocity model. Ray-based tomography was used to take into account the non-planarity of incoming wavefronts, and to stabilize the inversion we used *a priori* constraints on the smoothness of the velocity field and of the wavefronts.

The method is very robust in determining phase velocity maps in all synthetic examples. This robustness may be a consequence of using only the phase information of the signals, information which is more stable than the amplitudes of the seismograms. The numerical examples presented show that structures larger than 150 km can be obtained from phase information in the region containing the SVEKALAPKO receivers. Perturbations of the velocity model introduced by distorted wavefronts is weak. A 5° rotation of propagation direction superimposed on variations less than 30° over 100 km along the wavefront produces almost no artefact on the inverted velocity field inside the station network. The examples carried out using IBEM-computed seismograms allowed us to test both the process of time delay measurement and the inversion method. The quality of the result—even with a structure of a similar size to the wavelength—is very encouraging for their application on the SVEKALAPKO data set. The solution remains coherent with a signal to noise ratio down to 3.5, imposing objective preselection criteria for the inversion of the SVEKALAPKO data set. The inversion procedure remains stable when applied to real data, as demonstrated for the 60 s period fundamental-mode Rayleigh wave, using a subset of 48 events.

The ray tracing formulation has some major advantages. First, we can model both Rayleigh and Love waves. Second, we can easily implement discontinuities in the phase velocity model, which could be useful for imaging regional sharp structures such as the Tornquist Zone. Third, the mini-arrays of five and four stations in the middle of the SVEKALAPKO deep seismic experiment network can be used to measure locally the true propagation direction of the wavefronts, which subsequently can be used to constrain the inversion. The phase velocity beneath the mini-arrays can also be measured precisely and used as a further constraint. The lower limit of resolution depends on the inter station distance and on the wavelength. However, it results in a smoothing which is at least qualitatively predictable and can therefore be taken somewhat into account in the interpretation.

ACKNOWLEDGMENTS

We thank Tellervo Hyvönen for her help during the field work, as well as the numerous people who took part in it. The French participation received financial support from the 'Interieur de la Terre' programme of INSU and used Lithoscope and RLBM seismic stations. David Baumont provided the code used in the time-frequency filtering. Discussions with Guust Nolet and Michel Bouchon contributed to improving this paper, and two anonymous referees gave constructive comments to the original manuscript. The Editor provided a particularly careful correction of the English. The European Science Foundation financed the workshops of the SVEKALAPKO groups in Lammi, Finland. Most of the computations presented in this paper were performed at the Service Commun de Calcul Intensif de l'Observatoire de Grenoble (SCCI). We used SAC for the treatment of synthetic and real seismograms. Some figures were drawn using GMT and Xfig.

REFERENCES

- Babuška, V., Plomerová, J. & Pajdušák, P., 1988. Seismologically determined deep lithosphere structure in Fennoscandia, Proceedings in *Geologiska Föreningens i Stockholm Förhandlingar*, **110**, 380–382.
- Barker, J.S., Campillo, M., Sánchez-Sesma, F.J., Jongmans, D. & Singh, S.K., 1996. Analysis of wave propagation in the valley of Mexico from a dense array of seismometers, *Bull. seism. Soc. Am.*, **86**, 1667–1680.
- Bartels, R.H., Beatty, J.C. & Barsky, B.A., 1988. *B-splines*. Hermès Paris.
- Calcagnile, G., 1991. Deep structure of Fennoscandia from fundamental and higher mode dispersion of Rayleigh waves, *Tectonophysics*, **195**, 139–149.
- Červený, V., 1989. Ray tracing in factorized anisotropic inhomogeneous media, *Geophys. J. Int.*, **99**, 91–100.
- Cotte, N., Pedersen, H.A., Campillo, M., Farra, V. & Cansi, Y., 2000. Off-great circle propagation of intermediate period surface waves as observed on a dense array in the French Alps, *Geophys. J. Int.*, **142**, 825–840.
- Courant, R. & Hilbert, D., 1966. *Methods of Mathematical Physics*, Intersciences Publishers, Inc., New York.
- de Boor, C., 1978. *A Practical Guide to Splines*. Springer-Verlag, New York.
- Farra, V., 1990. Amplitude computation in heterogeneous media by ray perturbation theory: a finite element approach, *Geophys. J. Int.*, **103**, 341–354.
- Farra, V., 1993. Ray tracing in complex media, *J. appl. Geophys.*, **30**, 55–73.
- Farra, V. & Madariaga, R., 1987. Seismic waveform modelling in heterogeneous media by ray perturbation theory, *J. geophys. Res.*, **92**, 2697–2712.
- Farra, V. & Madariaga, R., 1988. Non-linear reflection tomography, *Geophys. J.*, **95**, 135–147.
- Friederich, W., 1998. Wave-theoretical inversion of teleseismic surface waves in a regional network: phase-velocity maps and a three-dimensional upper-mantle shear-wave-velocity model for southern Germany, *Geophys. J. Int.*, **132**, 203–225.
- Friederich, W. & Wielandt, E., 1995. Interpretation of seismic surface waves in regional networks: joint estimation of wave-field geometry and local phase velocity, Method and numerical tests, *Geophys. J. Int.*, **120**, 731–744.
- Friederich, W., Wielandt, E. & Stange, S., 1993. Multiple forward scattering of surface waves: comparison with an exact solution and Born single-scattering methods, *Geophys. J. Int.*, **112**, 264–275.
- Friederich, W., Wielandt, E. & Stange, S., 1994. Non-plane geometries of seismic surface wavefields and their implications for regional surface-wave tomography, *Geophys. J. Int.*, **119**, 931–948.
- Herrmann, R.B., 1985. *Computer Programs in Seismology, Volume III: Surface Waves in Plane Layers*, Saint Louis University, Missouri.
- Korja, A., Korja, T., Luosto, U. & Heikkinen, P., 1993. Seismic and geoelectric evidence for collisional and extensional events in the Fennoscandian Shield—implications for Precambrian crustal evolution, *Tectonophysics*, **219**, 129–152.
- Lander, A.V. & Levshin, A.L., 1989. Recording, identification, and measurement of surface wave parameters, in *Seismic Surface Waves in Laterally Inhomogeneous Earth*, V.I. Keilis-Borok (editor). Kluwer Academic Publishers, Dordrecht, 131–182.
- Levshin, A. & Berteussen, K.-A., 1979. Anomalous propagation of surface waves in the Barents Sea as inferred from NORSAR recordings, *Geophys. J. R. astr. Soc.*, **56**, 97–118.
- Luosto, U., 1990. Seismic data from the northern segment of the EGT and from the nearby profiles, in R. Freeman & St. Mueller, *The European Geotransverse (EGT) Project, Data Compilations and Synoptic Interpretation*, 53–63.
- Luosto, U., 1997. Structure of the Earth's crust in Fennoscandia as revealed from refraction and wide-angle reflection studies, *Geophysica*, **33**, 3–16.
- Marquering, H. & Snieder, R., 1996. Shear-wave velocity structure beneath Europe, the northeastern Atlantic and western Asia from waveform inversions including surface-wave mode coupling, *Geophys. J. Int.*, **127**, 283–304.
- Marquering, H., Snieder, R. & Nolet, G., 1996. Waveform inversions and the significance of surface-wave mode coupling, *Geophys. J. Int.*, **124**, 258–278.
- Maupin, V., 1988. Surface waves across 2D structures: a method based on coupled local modes, *Geophys. J.*, **93**, 173–185.
- McMechan, G.A. & Yedlin, M.J., 1981. Analysis of dispersive waves by wave field transformation, *Geophysics*, **46**, 869–874.
- Nolet, G., 1990. Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.*, **95**, 8499–8512.
- Pedersen, H.A., Sánchez-Sesma, F.J. & Campillo, M., 1994. Three-dimensional scattering by two-dimensional topographies, *Bull. seism. Soc. Am.*, **84**, 1169–1183.
- Pedersen, H.A., Maupin, V. & Campillo, M., 1996. Wave diffraction in multi-layered media with the Indirect Boundary Element Method: application to 3-D diffraction of long-period surface waves by 2-D lithospheric structures, *Geophys. J. Int.*, **125**, 545–558.
- Pollitz, F.F., 1999. Regional velocity structure in northern California from inversion of scattered seismic surface waves, *J. geophys. Res.*, **104**, 15 043–15 072.
- Poupinet, G., Ellsworth, W.L. & Fréchet, J., 1984. Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras fault, California, *J. geophys. Res.*, **89**, 5719–5731.
- Simons, F.J., Zielhuis, A. & van der Hilst, R.D., 1999. The deep structure of the Australian continent from surface wave tomography, *Lithos.*, **48**, 17–43.
- Snieder, R., 1988a. Large-scale waveform inversion of surface waves for lateral heterogeneity, 1. Theory and numerical examples, *J. geophys. Res.*, **93**, 12 055–12 065.
- Snieder, R., 1988b. Large-scale waveform inversion of surface waves for lateral heterogeneity, 2. Application to surface waves in Europe and the Mediterranean, *J. geophys. Res.*, **93**, 12 067–12 080.
- SVEKALAPKO Seismic Tomography Working Group, 2001. Seismic probing of Fennoscandian lithosphere, *EOS, Trans. Am. geophys. Un.*, **82**, 621, 628–629.
- Tarantola, A., 1987. *Inverse Problem Theory: Methods for Data Fitting and Parameter Estimation*. Elsevier, Amsterdam.
- van der Lee, S., 1998. Observations and origin of Rayleigh-wave amplitude anomalies, *Geophys. J. Int.*, **135**, 691–699.
- Yanovskaya, T.B., 1996. Ray tomography based on azimuthal anomalies, *Pure appl. Geophys.*, **148**, 319–336.
- Yanovskaya, T.B., Kizima, E.S. & Antonova, L.M., 1998. Structure of the crust in the Black Sea and adjoining regions from surface wave data, *J. Seismology*, **2**, 303–316.
- Zielhuis, A. & Nolet, G., 1994. Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.*, **117**, 695–715.