Surv Geophys (2009) 30:377–405
DOI 10.1007/s10712-009-9071-5

Upper Mantle Imaging with Array Recordings of Converted and Scattered Teleseismic Waves

Stéphane Rondenay

Received: 17 November 2008 / Accepted: 14 April 2009 /

Abstract This paper provides a review of array-based imaging techniques that use converted and scattered teleseismic waves. It addresses various aspects of the imaging process, from the preprocessing of the data to the application of the imaging algorithms. The reviewed techniques form a continuum with respect to the level of complexity adopted in the treatment of the scattering problem. On one end of the spectrum, images may be produced by simple stacking of normalized *P*-to-*S* conversion records (i.e., receiver functions), which are binned according to station or common conversion points (CCP) and mapped to depth. Finer resolution can be achieved through the stacking of singly scattered wavefields along diffraction hyperbolae to recover relative scattering intensity/potential at individual points through a 2-D or 3-D model space. Moving to higher levels of complexity, we find methods that involve inversion/backprojection of scattered teleseismic wavefield to recover estimates of localized material property perturbations with respect to an a priori background model.

Keywords Receiver functions · Teleseismic · Migration · Imaging lithosphere · Upper mantle

1 Introduction

Over the past 30 years, converted seismic waves have been used extensively to identify and characterize discontinuities in material properties in the Earth. The analysis of converted waves, along with seismic tomography, form the core of methodologies used in seismic imaging of the solid Earth at regional and global scales. While tomography is sensitive to volumetric changes in material properties, converted waves are sensitive to rapid variations in material properties (i.e., changes in seismic velocities and density occurring over a distance that is of the same order as the wavelength of the seismic waves).

S. Rondenay (✉)
Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA 02139, USA
e-mail: rondenay@mit.edu

Therefore, the two families of methods provide complementary means of characterizing the structure of the Earth's subsurface.

A number of methodologies have been developed over the years to analyse converted seismic waves, ranging from single station applications to high-resolution imaging using dense arrays of broadband seismometers. The objective of this paper is to provide a review of these approaches and highlight recent developments in the field of high-resolution teleseismic imaging. Such developments have been made possible by the increased availability of teleseismic data recorded at dense broadband seismic arrays.

The principles of teleseismic, converted wave analysis are rooted in the ray-theoretical treatment of seismic waves at welded boundaries. Waves encountering a discontinuity in material properties are partitioned into transmitted and reflected waves of various polarizations in order to satisfy boundary conditions (continuity of traction and displacement) at the discontinuity. An example of this phenomenon is shown schematically in Fig. 1, where a *P*-wave incident upon a sub-horizontal discontinuity in *S*-wave velocity and/or density partitions into upward and downward propagating *P* and SV waves. The same phenomenon occurs for an incident *S*-wave, and the partitioning of the energy between the different waves is given by the Zoeppritz equations (e.g., Aki and Richards 2002). In this simple example, the depth of a discontinuity can be estimated based on the time delay between the transmitted incident (*P*) and converted (*Ps*) waves, and the magnitude of the discontinuity can be estimated from the relative amplitude between these two waves.

The phenomenon described in Fig. 1 was first exploited to characterize crustal and upper mantle discontinuities through synthetic modelling of *S*-to-*P* and *P*-to-*S* converted waves (e.g., Båath and Stefánsson 1966; Burdick and Langston 1977). It was then used as a direct imaging tool in landmark studies by Vinnik (1977) and Langston (1979). To increase the signal to noise ratio of converted phases, these authors combined records from multiple sources by stacking traces that were source-normalized and time-shifted according to incidence angle. The term receiver function (RF) was introduced by Langston (1979) to describe these normalized records of converted waves and their stacks.

A large number of RF studies have been based on the analysis of data recorded at single surface locations, for example at permanent or semi-permanent seismic stations. This is a desirable strategy because such stations provide a high degree of data redundancy and

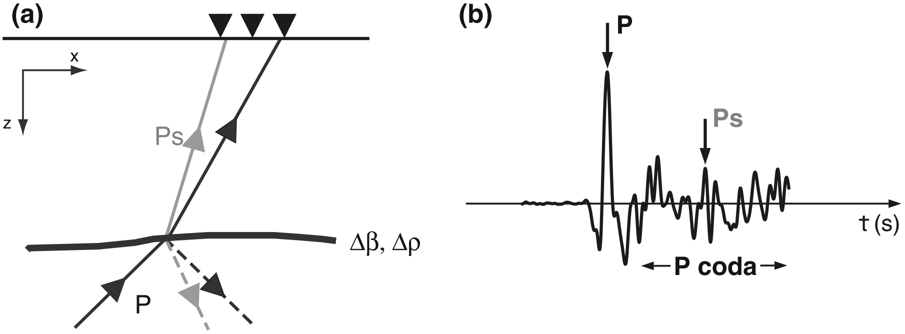


Fig. 1 Schematic diagram showing **a** the partitioning of a *P*-wave incident upon a sub-horizontal discontinuity in *S*-wave velocity and/or density, and **b** the resulting seismogram. The converted, upward propagating SV-wave is the basis of the receiver function technique. The time delay between the incident (*P*) and converted (SV, often denoted *Ps*) waves is proportional to the depth of the discontinuity, whereas the amplitude of *Ps* relative to *P* is indicative of the magnitude of the perturbation

therefore yield robust observations. However, these studies usually suffer from limited spatial resolution due to the sparse distribution of permanent stations. Another issue with this approach is that it assumes a one-dimensional (1-D) model geometry, i.e., planar and horizontal discontinuities over the region sampled by all incident rays recorded by a station (Fig. 2a).

This 1-D simplifying assumption has proven valid to first order for the characterization of major discontinuities of the Earth's subsurface: for example the Moho, the lithosphere–asthenosphere boundary, or the 410 and 660 km discontinuities bounding the mantle transition zone (e.g., Vinnik 1977; Bostock 1998; Rychert et al. 2005). However, there exist many subsurface structures that occur over small scale-lengths (i.e., same order as the wavelengths of teleseismic waves). These perturbations do not conform to the 1-D assumptions and they cause diffraction of the incident wave (Fig. 2b). Methods that use diffracted waves to image subsurface perturbations have been employed for over 50 years in seismic reflection imaging (see, e.g., Yilmaz 2001). They are based on the backprojection to depth of the diffracted wavefield recorded at the surface and, as such, are collectively referred to as migration techniques. These techniques allow for a more general treatment of subsurface discontinuities, but they require a comprehensive spatial sampling of the diffracted signal at the surface through dense arrays of recorders—something that has not been traditionally possible in passive seismic experiments.

In recent years, a notable increase in availability and fidelity of broadband seismic instruments has opened the way to realizing dense passive arrays that yield teleseismic data amenable to migration techniques. A range of methodologies has been developed to address this problem. On one end of this range, images may be produced by simple stacking of normalized *P*-to-*S* conversion records (i.e., receiver functions—RF), which are binned according to common conversion points (CCP) and mapped to depth (Dueker and Sheehan 1997; Kosarev et al. 1999). Finer resolution can be achieved through the stacking of singly scattered wavefields along diffraction hyperbolae to recover relative scattering

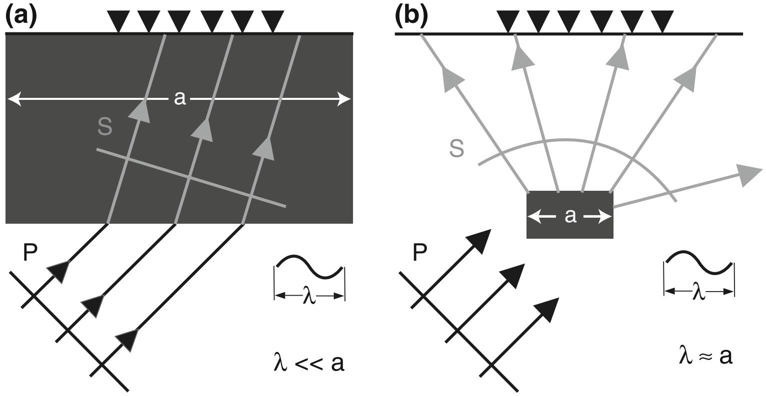


Fig. 2 Behaviour of converted phases as a function of the wavelength of the incident wave (λ) and the scale-length of the structure where the conversion occurs *a*. Both panels represent vertical profiles through the earth, with the incident *P*-wave and converted *S*-waves denoted by *black* and *gray* vectors, respectively. The *P*-to-*S* conversion occurs at a structure denoted by a *black rectangle*, and stations at the surface are denoted by *black triangles*. **a** 1-D geometry characterized by $\lambda \ll a$ and boundaries that are planar and horizontal. A planar incident wavefield produces an equally planar converted wavefield. **b** 2-D/3-D geometry characterized by $\lambda \sim a$. A planar incident wavefield produces converted waves that are radially scattered from the perturbations

intensity/potential at individual points through a 2-D or 3-D model space (Revenaugh 1995; Ryberg and Weber 2000; Sheehan et al. 2000). Moving to higher levels of complexity, methods involving inversion/backprojection operators (Bostock and Rondenay 1999; Bostock et al. 2001; Poppeliers and Pavlis 2003a) and full 3-D waveform inversion (Frederiksen and Revenaugh 2004) of scattered teleseismic body waves recover either the scattering potential or estimates of localized material property perturbations with respect to an a priori background model.

This paper first discusses the type of data used in converted-wave imaging, and reviews the main preprocessing steps necessary to render these data amenable to analysis. It then presents various array-based imaging techniques that have been developed in recent years, ranging for the juxtaposition of 1-D profiles to the full pre-stack migration of diffracted teleseismic waves.

2 Teleseismic *P*-Coda

In principle, scattering caused by any seismic wave incident upon a region of interest can be used for seismic imaging. The incident waves are either *P* or *S*. They can originate from either local or distant (teleseismic) sources, and represent direct or indirect arrivals (e.g., *P* or *PP* waves). In practice, however, a vast majority of methodologies aimed at imaging lithospheric and mantle structure use the coda of teleseismic *P*-waves. For the purpose of this paper, teleseismic *P*-waves are defined as waves recorded at epicentral distances $\geq$30°, such that they are not triplicated by the mantle transition zone.

P-waves are favoured because they are the first arrival in a seismogram and therefore possess a coda that is less likely to be contaminated by other primary phases. Consequently, the coda mainly comprises scattered waves that represent precisely the input data required in the imaging methodologies discussed here.

Teleseismic sources are favoured over local ones for various reasons: (1) the *P*-wave and its coda are better isolated from other primary incident waves as the time delay between phases tends to increases with distance; (2) the polarization direction and ray parameter of the incident wave are better constrained as they are much less affected by standard location errors on sources at teleseismic distances than on local earthquakes; and (3) the incident teleseismic wavefront can be considered planar at regional scales of several hundreds of kilometers, i.e., the polarization direction and ray parameters of the incident wave are nearly constant across regional arrays, which greatly simplifies the implementation of imaging algorithms.

To summarize, teleseismic sources afford superior constraints on the geometrical attributes of the incident wave and the *P*-coda affords superior isolation of the scattered signal. This explains why *P*-to-*S* (*Ps*) phases are more commonly used in converted-wave imaging. However, there has been increasing interest in using alternative incident wavefields. For example, *S*-to-*P* (*Sp*) scattering preceding teleseismic *S*-waves has proven useful for imaging discontinuities in the 80–150 km depth range, which are usually obscured by multiples in the *P*-coda (e.g., Vinnik and Farra 2002; Yuan et al. 2006; Rychert et al. 2007). Moreover, the combined analysis of *Ps* and *Sp* receiver functions (Rychert et al. 2007) increases the robustness of the resulting images as the two types of waves provide independent information about the discontinuities, with *Ps* waves occurring in the coda of *P* and *Sp* waves appearing as precursors to *S*. Analyses of *Sp* scattering have been conducted both at single stations (e.g., Rychert et al. 2007) and at dense arrays of broadband seismographs (e.g., Li et al. 2007).

The methodologies and examples presented in this paper focus mainly on scattered waves in the coda of *P*, because these waves are still the most practical and commonly used for high-resolution upper mantle imaging. However, the reader should be aware that in principle, but not always in practice, most of these methodologies are applicable to scattered phases originating from any incident wave (e.g., *PP, PcP, S, SKS*, etc.).

3 Partitioning of the Signal

Imaging techniques based on teleseismic converted waves require isolation and separation of the incident and scattered wavefields. A first step in this process that is common to all approaches, both at single stations and at arrays, consists of recasting the three orthogonal components of the seismogram (usually N–S, E–W, and vertical) into radial, transverse and vertical components:

$$\begin{pmatrix} \mathrm{R} \\ \mathrm{T} \\ \mathrm{Z} \end{pmatrix} = \begin{pmatrix} -\cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & -\cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{N} \\ \mathrm{E} \\ \mathrm{Z} \end{pmatrix}, \tag{1}$$

where N, E, and Z are the North-South, East-West, and vertical components, respectively; R and T are the radial and transverse components; and γ is the backazimuth of the incident ray (Figs. 3a, b, 4a, b).

Starting from the R–T–Z components, the signal can be partitioned into estimates of the incident and scattered wavefields following one of three approaches. These approaches were developed for 1-D RF analysis where discontinuities are assumed to be planar, horizontal and isotropic, yielding incident *P*-waves and converted SV-waves that are confined to the sagittal plane (i.e., the vertical plane comprising the source and the receiver). Accordingly, we first introduce them in the context of 1-D receiver functions and then assess their applicability to 2/3-D scattered wave imaging.

The first approach relies on the near-vertical incidence of *P*-waves to assume that the incident wavefield is confined to the vertical component (Z) whereas the converted waves are entirely contained in the radial (R) component. These assumptions are acceptable to first order and have been widely employed in teleseismic imaging studies, but there is in fact considerable signal leakage between the R and Z components. This is because the actual incidence angle of *P*-waves at the surface typically ranges between ~10° and 30°, which is not vertical and implies that the incident and *Ps* wavefields are recorded on both vertical and radial components (Figs. 3b, 4b).

The second approach partially addresses the signal leakage issue by rotating the R–T–Z components in the direction of polarization of the incident *P*-wave (L) and its perpendicular (Q) in the R–Z plane (e.g., Vinnik 1977; Kind et al. 1995):

$$\begin{pmatrix} \mathrm{L} \\ \mathrm{Q} \\ \mathrm{T} \end{pmatrix} = \begin{pmatrix} \cos i_\alpha & \sin i_\alpha & 0 \\ -\sin i_\alpha & \cos i_\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{Z} \\ \mathrm{R} \\ \mathrm{T} \end{pmatrix}, \tag{2}$$

where $i_\alpha = \arcsin(p\alpha)$ is the angle between the ray of the incident *P*-wave and the vertical axis. This transformation requires estimates of the near-surface *P*-wave velocity α, which can be obtained from global reference models or regional tomographic models, and of the ray parameter p of the incident wave, which can be obtained from tau-p tables (e.g., Crotwell et al. 1999). A schematic diagram of this transformation is presented in Fig. 3c. It shows that the incident wavefield is completely isolated in the L component provided that

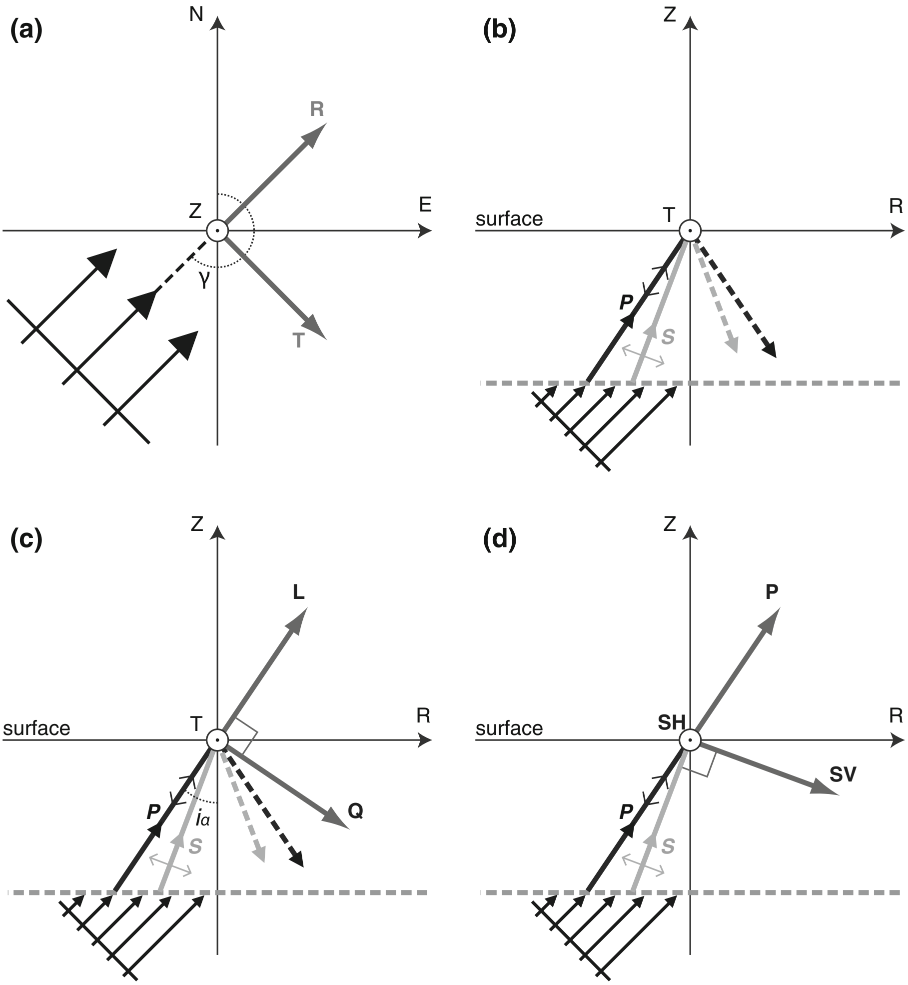


Fig. 3 Partitioning of teleseismic waves for imaging purposes: Projections axes. **a** Plan view of a teleseismic wave incident from the SW. Axes N, E, and Z denote North, East and vertical directions, respectively. The recording point is located at the intersection of the axes. Angle γ is the backazimuth of the incident wave. R and T denote the radial and transverse components, respectively. **b** Vertical profile through the sagittal plane, showing an incident wave (*bottom left*) that interacts with a discontinuity denoted by the horizontal *dashed line*. Upgoing and downgoing (free-surface reflected) waves are denoted by *solid and dashed lines*, respectively. The propagation direction of the waves is indicated by *solid arrows*, and their polarization is shown by the *open arrows*. (**c**) L–Q–T transformation, where L is parallel to the upgoing *P*-wave and Q is perpendicular to this direction. (**d**) *P*–SV–SH transformation, where P is parallel to the upgoing *P*-wave, SV is perpendicular to the upgoing *S*-wave, SH is parallel to T, and the effects of the free-surface are suppressed. *P*–SV–SH transformation does not involve simple rotation/projection of the recorded signal, but actual separation into distinct wavefields (see Eq.3). The various wavefield transformations are illustrated with a synthetic data example in Fig. 4

all assumptions are valid, the parameter estimates are exact, and the effects of the free surface are ignored. The synthetic example of Fig. 4c demonstrates that, in practice, a small pulse related to the incident arrival is still present in the Q component. This pulse is associated with the free-surface reflection of the incident wave. In the L–Q–T

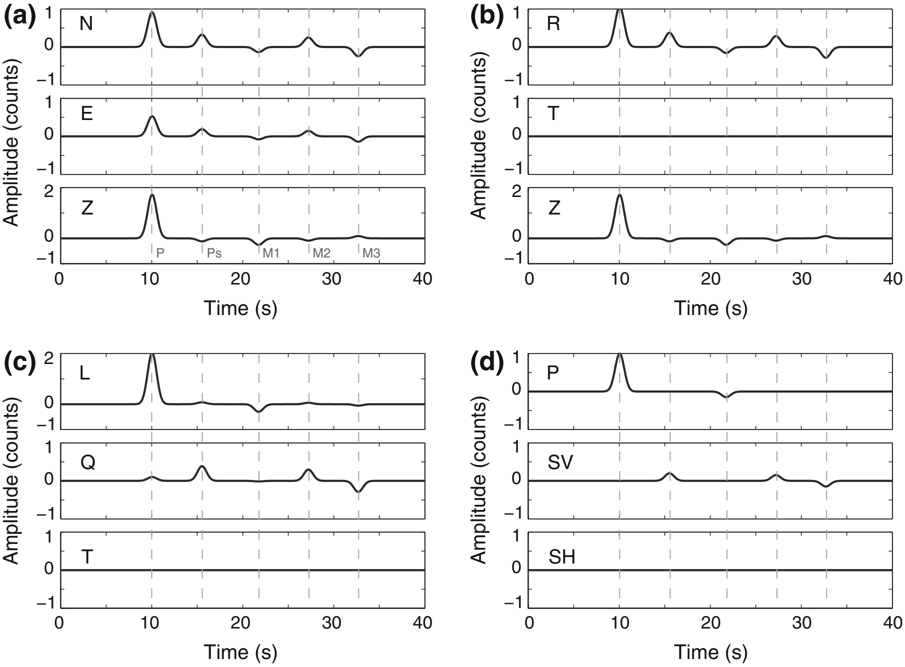


Fig. 4 Partitioning of teleseismic waves for imaging purposes: synthetic data example. The dataset is constructed using the RAYSUM package (Frederiksen and Bostock 2000), for an incident *P*-wave with backazimuth $\gamma = 210°$ and a ray parameter $p = 0.08$ s/km, recorded at a single surface location. The model sampled by the incident wave is isotropic, and consists of a 40 km-thick horizontal layer ($\alpha_0 = 6.0$ km/s, $\beta_0 = 3.4$ km/s, $\rho_0 = 2{,}600$ kg/m^3) over a half-space ($\alpha_1 = 8.1$ km/s, $\beta_1 = 4.5$ km/s, $\rho_1 = 3{,}500$ kg/m^3). **a** N–E–Z components. **b** R–T–Z components. **c** L–Q–T components. **d** *P*–SV–SH components. Observed phases: incident *P*-wave (*P*), *P*-to-*S* conversion from the discontinuity at 40 km depth (*Ps*), free-surface multiples (*M1*, *M2*, *M3*). See Fig. 3 for a definition of the various coordinate systems

transformation, *Ps* waves converted at horizontal discontinuities are mostly confined to the Q component, however they leak into the L component since the signal is projected using the incident angle of the *P* wave as opposed to that of the *S* wave (see Fig. 4c).

In the third approach, the R–T–Z components are converted using the free-surface transfer matrix (Kennett 1991). This transformation rotates the wavefield in the polarization directions of the incident *P*-wave and corresponding SV/SH waves for a 1-D layered model, and suppresses the effects of the free-surface (i.e., downgoing waves reflected off the free-surface). It is expressed as follows (e.g., Bostock and Rondenay 1999):

$$\begin{pmatrix} P \\ \mathrm{SV} \\ \mathrm{SH} \end{pmatrix} = \begin{pmatrix} \frac{\beta^2 p^2 - 1/2}{\alpha q_\alpha} & \frac{p\beta^2}{\alpha} & 0 \\ p\beta & \frac{1/2 - \beta^2 p^2}{\beta q_\beta} & 0 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} \mathrm{Z} \\ \mathrm{R} \\ \mathrm{T} \end{pmatrix}, \qquad (3)$$

where α and β are the near-surface *P*- and *S*-wave velocities, respectively, and $q_\alpha = \sqrt{1/\alpha^2 - p^2}$ and $q_\beta = \sqrt{1/\beta^2 - p^2}$ are the corresponding vertical slownesses. With this transformation, the incident *P* wavefield and the *Ps*-waves converted at horizontal discontinuities are, in theory, completely isolated from one another and signal distortion caused by downgoing waves reflected at the free surface is suppressed (Figs. 3d, 4d).

For imaging in 1-D isotropic, horizontally layered media, the second and third approaches are clearly superior to the first as they prevent leakage of the incident signal into the converted/scattered signal. In theory, the *P*–SV–SH transformation is an improvement over L–Q–T as it further prevents leakage of the converted waves into the incident signal, and it suppresses the signal associated with reflections at the free surface (Reading et al. 2003). In practice, however, the difference between these two transformations is not always significant as it may be superseded by the effects caused by departure from the assumed 1-D geometry and errors in the estimates of near-surface velocity (Svenningsen and Jacobsen 2004). In any of these transformations, the presence of coherent signal on the transverse component, i.e., T or SH, indicates that the imaged structure does not conform to the assumed 1-D geometry and isotropic behaviour—it is indicative of anisotropy, dipping layers, and 2-D or 3-D scatterers.

For the more general cases of 2-D or 3-D scattered-wave imaging that employ the Born approximation (see Sect. 7), both the L–Q–T and *P*–SV–SH transformations are useful because they isolate the incident wavefield on one of the components. However, the scattered waves are now present in all three components, and other means must be employed to separate the scattered signal from the incident wave (see Sect. 4). Finally it is worth noting that, since the *P*–SV–SH transformation requires estimates of the *P*- and *S*-wave velocities at the surface, it is possible to use Eq. 3 to estimate these velocities by finding the values of α and β that minimize the amount of leakage of the incident wavefield onto the SV–SH components (Bostock and Rondenay 1999; Svenningsen and Jacobsen 2007).

4 Separation of 2-D and 3-D Scattered Signal

A more general treatment of the scattered wavefield is afforded by considering the effects from line or point perturbations, which account for interactions with 2-D/3-D structure (e.g., dipping slab, intrusive body) in addition to horizontally layered structure. For line and point perturbations, the scattered wavefield is present in all three components resulting from any of the transformations described in Sect. 3 (including L or *P*). In this case, the separation of the 2-D and 3-D scattered signal is best achieved when the wavefields are recorded at multiple stations, i.e., when multichannel recordings from dense local or regional arrays are available. For teleseismic data, this translates into arrays with $\leq$20 km station spacing over apertures of 100–500 km.

Multichannel preprocessing approaches rely on the assumption that the incident and scattered wavefields have distinct moveouts (arrival times vs offset), and usually comprise the following steps (e.g., Rondenay et al. 2005): (1) The traces are aligned relative to the incident wavefield identified on the L/*P* components. (2) The incident wavefield is extracted either by stacking the aligned L/*P* components, or by conducting Principal Component Analysis (PCA) on the aligned L/*P* components and retaining the first or first few components. We will refer to this estimate of the incident wavefield as $\bar{L}$, or $\bar{P}$, (3) The portion of the scattered wavefield contained in components L/*P* is obtained by subtracting the incident wavefield estimates $\bar{L}$, or $\bar{P}$, from each L/*P* trace, such that $L' = \mathrm{L} - \bar{L}$, or $P' = P - \bar{P}$,. (4) The scattered wavefield is thus isolated in components L'–Q–T or in P'–SV–SH, depending on the chosen transformation. Figure 5 shows an example of multichannel preprocessing applied to data from a dense array of broadband seismographs deployed across the Hellenic subduction zone (Suckale et al. 2007).

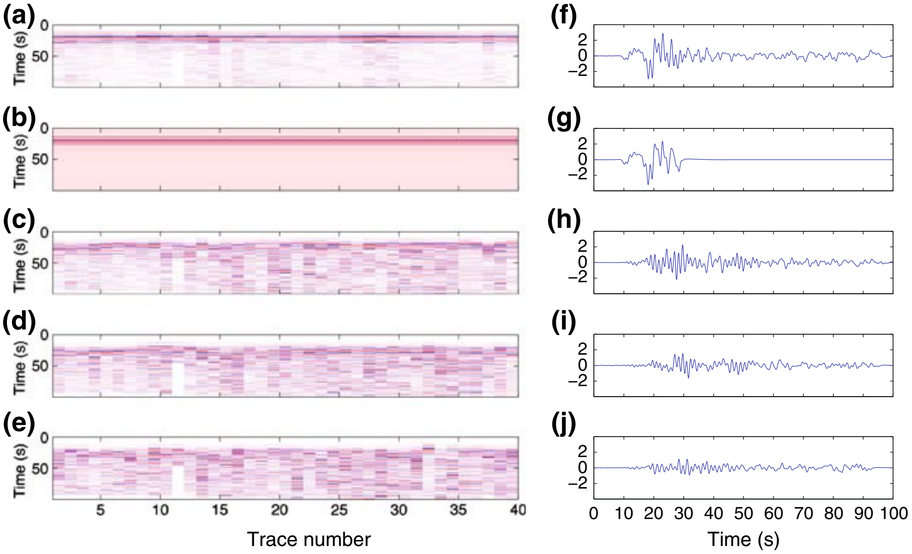


Fig. 5 Separation of incident and scattered teleseismic wavefields by multichannel processing: Example for data from an Indonesian earthquake recorded by a dense array of 40 broadband seismographs deployed across the Hellenic subduction zone (Suckale et al. 2007). *Left panels* (**a**–**e**) show data sections across the entire array, with seismograms shown as *vertical colour bands* (*red*/*blue* corresponding to positive/negative pulses). *Right panels* **f**–**j** show one seismogram from the data section to the left, i.e., trace number 5. **a**, **f** P component. Traces are normalized and aligned relative to the P-wave, which arrives at ~ 9 s. Principal Component Analysis (PCA) is used on this section to estimate the following two components of the P-wavefield: **b**, **g** the incident wavefield $\bar{P}$, which corresponds to the first principal component; and **c**, **h** the scattered wavefield P', which corresponds to the remaining principal components. **d**, **i** SV component. **e**, **j** SH component

The multichannel preprocessing approaches described above are efficient at separating the wavefields as long as these exhibit different moveouts. However, if a portion of the scattered wavefield is aligned with the incident wave, it will become incorporated into the estimate of that incident wave, resulting in a degradation of the data set. This problem occurs, for example, in the case of P-to-P backscattering (i.e., via free-surface multiples) from horizontal discontinuities in the subsurface.

5 Deconvolution

The next key step in data conditioning for any form of scattered/converted wave imaging involves normalization of the source signature for each event used in the analysis. This procedure is essential for coherently stacking scattered waveforms and ensuring that they exhibit high signal-to-noise ratio. The basis of this normalization is the forward expression:

$$w(t) * r(t) = \int_{\tau_1}^{\tau_2} w(t-\tau) r(\tau) d\tau = d(t) \tag{4}$$

in which the recorded signal d(t) is expressed as the convolution of the Earth's impulse response $r(t)$ with the combined source time function and instrument response $w(t)$. In the

Fig. 6 Moveout correction and stacking of *Ps* and *Sp* receiver functions. **a** A synthetic dataset is▶ constructed with the RAYSUM package (Frederiksen and Bostock 2000) for a range of incident ray parameters, using an isotropic velocity model consisting of a horizontal low-velocity layer (100–112 km depth) embedded in a half-space. **b** Individual radial receiver functions obtained for incident *P*-waves. The traces are aligned relative to the incident *P*-wave at 0 s. The theoretical arrival times of *Ps*-waves converted at incrementally increasing depths are shown as *dashed lines*. **c** Stacked *Ps* receiver function, obtained by summing the signal along dashed moveout curves in **b** and projecting the results onto the time axis for the reference ray parameter—here $p_0 = 0$ s/km. **d** Individual vertical receiver functions obtained for incident *S*-waves. In this case, *Sp* conversions arrive prior to the incident *S*-wave at 0 s, hence the analysis of signal in a negative time scale. **e** Stacked *Sp* receiver function—in general these are displayed in reversed time scale and polarity for direct comparison with the *Ps* signal. The moveout curves are calculated using the half-space velocities, which explains why the thickness of the layer is not accurately resolved and why conversions from the lower discontinuity are not stacked as coherently as those from the upper discontinuity in (**c**) and (**e**)

signal partitioning framework established in Sects. 3 and 4, these terms correspond to the following components: For 1-D partitioning of converted waves, (1) $w(t) = \mathrm{Z}$ and $\mathrm{d}(t) = \mathrm{R}$ or T; (2) $w(t) = \mathrm{L}$ and $\mathrm{d}(t) = \mathrm{Q}$ or T; (3) $w(t) = P$ and $\mathrm{d}(t) = \mathrm{SV}$ or SH. For 2-D and 3-D scattered signal, (4) $w(t) = \bar{L}$ and $\mathrm{d}(t) = L'$–Q–T; (5) $w(t) = \bar{P}$, and $\mathrm{d}(t) = P'$–SV–SH. This convolution can also be expressed in matrix form as (e.g., Gurrola et al. 1995):

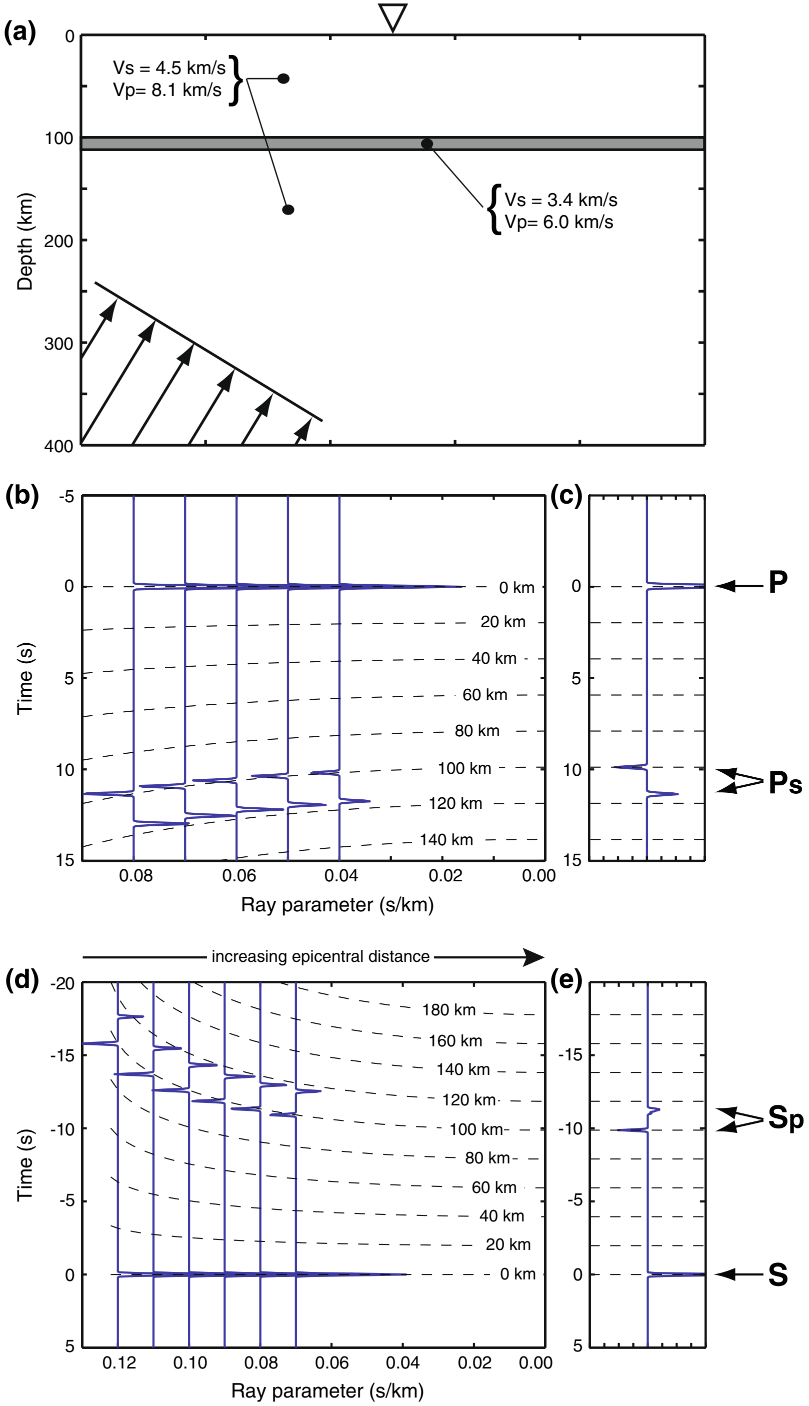
$$\mathbf{Wr} = \mathbf{d}. \tag{5}$$

If $\mathrm{d}(t)$, $r(t)$ and $w(t)$ contain n, m and p elements, respectively, then **d**, and **r** are single-column matrices containing the elements of the two time series, and **W** is an $n \times m$ matrix with elements $W_{ij} = w(t_{i\text{-}j+1})$ for $1 \leq i{-}j + 1 \leq p$ and $W_{ij} = 0$ otherwise.

The process of normalization consists of solving for $r(t)$ in Eqs. 4 and 5 by deconvolving $w(t)$ from $\mathrm{d}(t)$. In theory, if the source-instrument wavelet is known and there is no noise, it may be possible to solve this problem directly by division in the frequency domain or by computing the inverse of matrix **W** (provided that $n = m$). In practice, however, the problem is ill-posed due to the presence of noise in the data and to inaccuracies in the estimation of the source-instrument wavelet. Thus it must be solved by imposing constraints on the solution. This is commonly achieved through least-squares approaches that, in their most elementary form, search for the solution that minimizes the difference between the data and the modelled solution. In this case, one could be searching for the solution $\hat{\mathbf{r}}$ that minimizes the cost function $\|\mathbf{W}\hat{\mathbf{r}} - \mathbf{d}\|^2$, where $\|\cdot\|^2$ denotes the L_2-norm, leading to the normal solution $\hat{\mathbf{r}} = (\mathbf{W}^{\mathbf{T}}\mathbf{W})^{-1}\mathbf{W}^{\mathbf{T}}\mathbf{d}$. However, the matrix **W** is usually ill-conditioned and small eigenvalues will lead to severe amplification of the noise, causing ringing in the deconvolved signal. As a consequence, the inverse problem must be regularized. This can be done in the frequency domain by pre-whitening the amplitude spectrum of the source-instrument wavelet, leading to the expression (e.g., Langston 1979; Gurrola et al. 1995):

$$\hat{r}(\omega) = \frac{\mathrm{d}(\omega)\mathrm{w}^*(\omega)}{w(\omega)w^*(\omega) + \delta}, \tag{6}$$

where the asterisk denotes the complex conjugate, ω is the angular frequency and δ is a damping factor. Some authors have used a water level instead of the damping factor, i.e., they replace values of the denominator that fall below a certain threshold with a fixed value called the water level (e.g., Langston 1979). In RF studies, the value of this damping factor (or water level) is often chosen in an ad hoc fashion on the basis of its capacity to reduce


(a)
Depth (km)
0
100
200
300
400
Vs = 4.5 km/s
Vp= 8.1 km/s
Vs = 3.4 km/s
Vp= 6.0 km/s
(b)
Time (s)
-5
0
5
10
15
0 km
20 km
40 km
60 km
80 km
100 km
120 km
140 km
0.08
0.06
0.04
0.02
0.00
Ray parameter (s/km)
(c)
P
Ps
increasing epicentral distance
(d)
Time (s)
-20
-15
-10
-5
0
5
180 km
160 km
140 km
120 km
100 km
80 km
60 km
40 km
20 km
0 km
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Ray parameter (s/km)
(e)
Sp
S

the amount of ringing in the solution. However, there exist more objective means of determining optimal values of δ. For example, δ can be chosen on the basis of a generalized cross-validation function constructed from Eq. 6 (e.g., Bostock 1998), or it can be controlled by the pre-event noise spectrum (e.g., Berkhout 1977).

Park and Levin (2000) propose an alternative to the standard spectral division approach in which they replace the products in the numerator and denominator of Eq. 6 with multiple-taper spectral correlation operators that reduce spectral leakage usually associated with taking the FFT of finite signals. The spectral division is stabilized by adding, in the denominator, a damping factor corresponding to the spectrum estimate of the pre-event noise. The method also affords a direct means of calculating formal uncertainty estimates on the resulting RF by computing the coherence between the input signal and source wavelet estimate—a quantity that allows one to produce RF stacks in which each trace is weighted according to its robustness. This is a feature that is not readily available in other deconvolution approaches.

Other means of exerting some control over the regularization parameter can be found by solving the deconvolution problem in the time domain. In this case, one seeks the solution that minimizes the regularized cost function $\|\mathbf{W}\hat{\mathbf{f}} - \mathbf{d}\|^2 + \mu^2\|\hat{\mathbf{f}}\|^2$, where μ is a Lagrange multiplier that weights the new condition of minimum energy on the impulse response. The solution in the time domain is then given by classic weighted least-squares solution (e.g., Sipkin and Lerner-Lam 1992; Gurrola et al. 1995):

$$\hat{\mathbf{f}} = (\mathbf{W^T W} + \mu^2 \mathbf{I})^{-1}\mathbf{W^T d}, \tag{7}$$

in which the problem can be solved iteratively while reducing the value of μ until the solution converges. For example, Gurrola et al. (1995) start the inversion with $\mu = 10^{10}$ and reduce its value by an order of magnitude after each iteration until the RMS misfit changes by <0.05% from one iteration to the next. The problem can be further stabilized both in the frequency and time domain by beamforming over multiple sources, i.e., simultaneously deconvolving traces from nearby sources recorded at the same station (Gurrola et al. 1995; Bostock 1998).

Time-domain deconvolution can also be performed by progressively constructing the Earth's impulse response (**r** in Eq. 6) through iterative cross-correlation between the data **d** and the source-instrument wavelet **W**, and extraction of the most highly correlated signal at each iteration (Kikuchi and Kanamori 1982; Ligorría and Ammon 1999). In an application of this approach to teleseismic receiver functions, Ligorría and Ammon (1999) construct a band-limited impulse response by adding, at each iteration, a Gaussian pulse whose amplitude is proportional to the highest cross-correlation coefficient, and whose time offset is determined by the corresponding time lag. At each new iteration, the contribution of the pulses introduced in previous steps is stripped from the recorded signal. The iterative process is repeated until the misfit between the recorded and modelled data falls beneath a certain threshold. This approach has the advantage of providing long-period stability and requiring causality in the recovered receiver functions, while not relying on any regularization scheme (Ligorría and Ammon 1999). As such, it is well tailored for automated RF analysis.

In closing this section, it is worth noting that recent efforts have been aimed at developing normalization approaches that do not rely on standard, and inherently unstable, deconvolution methods. A promising approach described by Bostock (2004) uses the minimum-phase property of the Earth's impulse response to the incident wavefield to normalize the source signature on *P*-wave seismograms.

6 Moveout Correction, Stacking and 1-D Imaging

Once traces have been normalized, they may be stacked to increase the signal-to-noise ratio and/or to capture the complexity of signal scattered from 2-D and 3-D discontinuities. In this section, we review the simplest form of converted-wave imaging using array data: stacking of receiver functions with the assumption that the background velocity model is known and that its structure is one-dimensional (i.e., horizontally layered). The stacked traces are grouped either as a function of common receivers at the surface or CCP at depth. This approach leads to a series of individual 1-D RF stacks that are juxtaposed to create pseudo 2-D profiles.

Prior to stacking, a moveout (i.e., timing) correction is applied to each trace to account for the variability in incidence angle between each incident wave. For the common case of *P*-to-SV conversions, the delay between the incident *P* wave and an SV wave converted at depth *h*, for incident angles expressed in terms of ray parameter of the incident wave *p*, is expressed as follows:

$$\begin{aligned} T_{Ps}(p,h) &= \int_{z=0}^{z=h} \left(\sqrt{1/\beta^2(z)-p^2} - \sqrt{1/\alpha^2(z)-p^2}\right)\mathrm{d}z \\ &= \int_{z=0}^{z=h} \left(q_\beta(p,z) - q_\alpha(p,z)\right)\mathrm{d}z, \end{aligned} \tag{8}$$

where $q_\beta(p,z)$ and $q_\alpha(p,z)$ are the vertical slownesses, as defined in Eq. 3, and the depth z is taken positive downward. The same expression can be used for *Sp* waves by setting the ray parameter to that of the incident *S* wave. Note that Eq. 8 and the following development assume that the incident wave is planar over the region it samples, which means that the ray parameter of the converted wave is equal to that of the incident wave (i.e., $p = p_{\{P\}} = p_{\{Ps\}}$). This assumption is valid at teleseismic distances for wave conversions occurring at structure in the upper mantle. It is applicable to signal converted at discontinuities in the ~0 to 410 km depth range, for which differences in ray parameter between the incident and converted waves result in errors of <0.05 s (i.e., <0.5 km in terms of conversion depth). This maximum value is calculated for events at 30° epicentral distance. It is far lower than the resolution of typical *P* receiver functions in the upper mantle, which is of order 2–5 km based on their frequency content and uncertainties in the background velocity model (Bostock 1999; Rychert et al. 2007). With deeper structure, it may be necessary to account for the difference in ray parameter between the incident and converted waves. For example, the use of a uniform ray parameter for signal converted at the 660 km discontinuity causes a maximum error of 0.4 s, or ~5 km in terms of conversion depth, which approaches the typical resolution of *P* receiver functions. An appropriate correction is achieved by replacing $q_\beta(p,z)$ and $q_\alpha(p,z)$ with $q_\beta(p_{\{Ps\}},z)$ and $q_\alpha(p_{\{P\}},z)$ in Eq. 8, where the ray parameters of the converted and incident waves ($p_{\{Ps\}}$ and $p_{\{P\}}$, respectively) may be estimated by ray-tracing from the source to the station through a reference Earth velocity model.

For a model consisting of discrete layers with homogeneous properties, Eq. 8 reduces to a sum over the *n* layers that are overlying the hypothetical conversion depth *h*:

$$T_{Ps}(p,h) = \sum_{i=0}^{i=n(h)} \left(q_{\beta i}(p) - q_{\alpha i}(p)\right)\Delta z_i, \tag{9}$$

where Δz_i denotes the thickness of each layer.

The moveout correction is computed by taking the difference between $T_{Ps}(p,h)$ and the time delay calculated for an arbitrary reference ray parameter p_0:

$$\Delta T_{Ps}(p,h) = T_{Ps}(p,h) - T_{Ps}(p_0,h). \tag{10}$$

This means that for each time sample t_j in a RF $\hat{r}(t_j)$, a moveout correction of $\Delta T_{Ps}(p,h_j)$ calculated for a hypothetical conversion depth h_j is applied, thus providing a direct relationship between time and conversion depth. Since the correction varies with time (or conversion depth), it has the effect of either stretching or compressing the receiver function. For a group of N traces, the correction and stack can be expressed as follows:

$$R(t) = R(h(t)) = \mathrm{IFT}\left[\sum_{k=1}^{N} \hat{r}_k(\omega)\mathrm{e}^{i\omega\Delta T_{Ps}(p_k,h(t))}\right], \tag{11}$$

where IFT denotes the inverse Fourier transform. Synthetic examples illustrating the moveout correction and stacking of *Ps* and *Sp* receiver functions for a simple 1-D model are shown in Fig. 6.

The stacked traces obtained through Eq. 11 can be plotted either as a function of time or conversion depth, and they form the basis of 1-D RF imaging. For individual stations, stacks can be formed using either all the RFs recorded by the station, or RFs binned according to ranges of backazimuths or epicentral distances (e.g., Bostock 1998; Rychert et al. 2005, 2007). This provides insight into structure underlying the station within a depth dependent aperture (hereafter referred to as RF station aperture) as shown in Fig. 7a. For array data, 1-D RFs can be juxtaposed to form pseudo 2-D profiles, but the manner in which they are stacked depends on station separation.

When station separation is larger than the RF station aperture at the depth of interest (e.g., ~20 km at Moho depth, ~70 km at 100 km depth, see Fig. 7a, b), the RF's are binned by stations and the profile is created by plotting the stacked RF's side-by-side. The lateral resolution afforded by these profiles corresponds to the station separation. An example of such a profile obtained in the SE Tibetan plateau (Xu et al. 2007) is shown in Fig. 8. The average station spacing is ~150 km and the target structure is the Moho. This

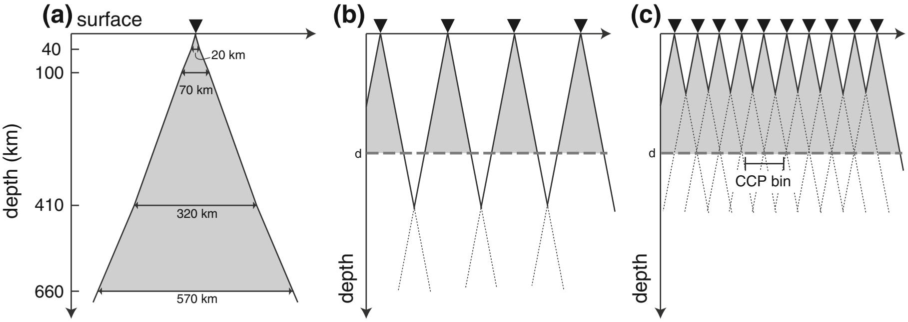


Fig. 7 Construction of a pseudo 2-D profile by juxtaposition of 1-D images. **a** RF station aperture, defined by the largest lateral conversion offset. This schematic diagram gives the approximate diameter of RF station aperture for *Ps*-waves converted at the Moho, at 100 km depth, and at transition zone discontinuities, using the Preliminary Reference Earth Model (PREM, Dziewonski and Anderson 1981). **b** Station bin profile—used when station separation is such that the RF station apertures do not overlap at the depth of the imaging target d. **c** CCP bin profile—used when station separation is such that the RF station apertures overlap at the depth of the imaging target d

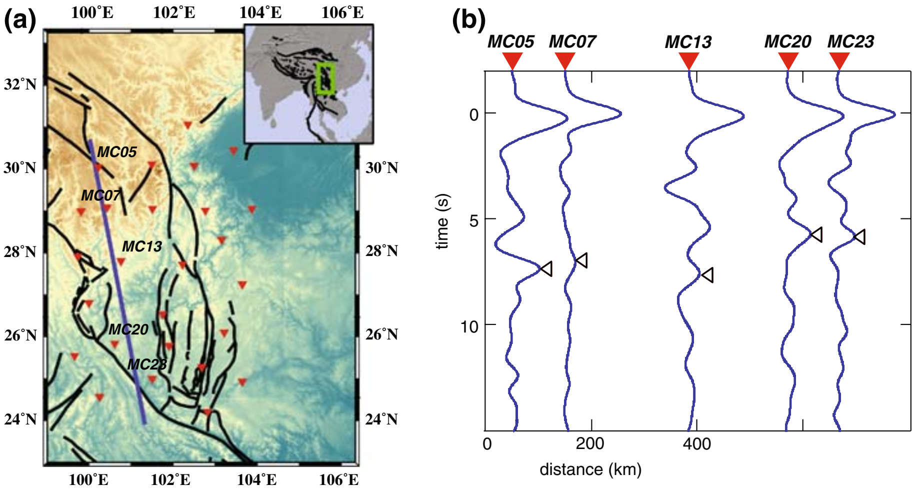


Fig. 8 Example of a station-bin RF profile for *Ps* waves recorded across the SE Tibetan Plateau (modified from Xu et al. 2007). **a** Station distribution map, with location of the study area shown in the *inset*. Stations are indicated by *red inverted triangles* and the location of the profile is indicated by a *blue line*. **b** 1-D RF section obtained by simultaneous deconvolution of converted waves binned according to station. An average of ~70 traces were used at each station. The *P*-wave arrivals are aligned at 0 s, and the *Ps* conversion from the Moho is indicated by *black arrows*. The Moho depth ranges from 60 km (7.5 s) at high elevation stations (~4000 m; MC05, MC07, MC13) to 48 km (5.7 s) at moderate elevation stations (~2000 m; MC20, MC23), reflecting isostatic compensation of surface topography

type of profile is useful for characterizing regional variations in Moho depth, but it does not capture any rapid change that may be associated with a tectonic boundary. Note that it is also possible to produce several profiles using only RF's from some backazimuthal ranges to investigate regions on either side of the profile or possible anisotropic effects.

When station separation is smaller than the RF station aperture at the depth of interest (Fig. 7a, c), the RF's can be binned as a function of CCP. The binning is done either at a single depth of interest to focus the image at that depth, or at successive depths to focus the image over the entire depth range afforded by the traces. The resulting stacks are once again plotted side by side to produce a pseudo 2-D profile. This approach, which is inspired by common midpoint (CMP) gathers in seismic reflection profiles, was introduced by Dueker and Sheehan (1997). It provides better lateral resolution than station bin profiles. The lateral resolution for this type of profile is equivalent to the Fresnel zone of the converted waves. Two examples of CCP profiles are shown in Figs. 9 and 10.

Figure 9 shows a *Ps* CCP profile across the Grenville Front, in the SE Canadian Shield (Rondenay et al. 2000). The profile outlines both regional variations in Moho depth, indicating a generally thinner crust in the Archean terranes to the north than in the Proterozoic ones to the south, and local features, such as a sudden jump in Moho depth occurring in the vicinity of the Grenville Front (GF).

Figure 10 shows an *Sp* CCP profile across the western United States taken from Li et al. (2007). Since *Sp* conversions are recorded as precursors to the incident *S*-wave, they do not overlap with crustal and Moho multiples. They are therefore well suited for imaging discontinuities in the mantle lithosphere and asthenosphere, particularly in the ~80 to 150 km depth range (i.e., the depths at which multiples may interfere with *Ps* conversions).

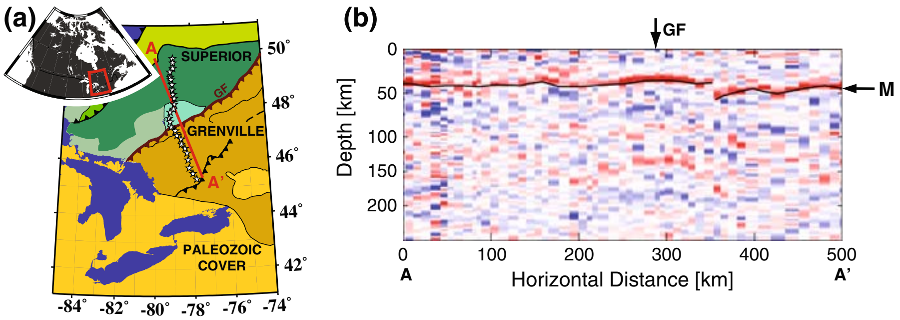


Fig. 9 Example of a CCP-bin RF profile for *Ps*-waves recorded across the SE Canadian Shield (modified from Rondenay et al. 2000). **a** Station distribution map, with location of the study area shown in the *inset*. Stations are indicated by *stars* and the location of the profile is indicated by a *red line*. Archean terranes of the Superior province to the north are separated from highly deformed Archean/Proterozoic rocks of the Grenville province to the south by the Grenville Front (GF). **b** 1-D RF section obtained by simultaneous deconvolution of converted waves binned according to CCP at 40 km depth, with variable bin size corresponding to half the station spacing along the projection line. An average of ~50 traces were used at each bin. Stacked RF's are shown as *vertical colour bands* (*red/blue* corresponding to positive/negative pulses). The *P*-wave arrivals are aligned at 0 km, and the *Ps* conversions from the Moho (M) appear as a positive pulse near 40 km depth. The crust is thicker beneath the Grenville (~45 to 50 km) than beneath the Superior (~40 km), and a Moho jump is observed 60 km to the south of the Grenville Front

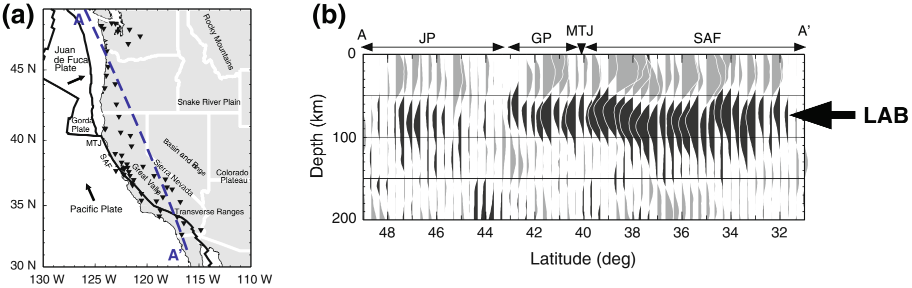


Fig. 10 Example of a CCP-bin RF profile for *Sp*-waves recorded across the western United States (modified from Li et al. 2007). **a** Map of the study area. Stations are indicated by *inverted black triangles* and the location of the profile is indicated by a *blue line*. Abbreviations: *MTJ* Mendocino Triple Junction; *SAF* San Adreas Fault. **b** 1-D RF section obtained by stacking *Sp*-waves binned according to CCP, with a bin spacing of 0.33°. A prominent negative pulse (*shaded black*) observed beneath the regions labelled GP (Gorda Plate) and SAF is indicative of a negative velocity gradient occurring in the 65–80 km depth range. This discontinuity has been interpreted by Li et al. (2007) to represent the LAB. In the northwest portion of the profile the negative signal is weaker and less coherent, perhaps reflecting conversions from structure associated with the Cascadia subduction zone (e.g., the subducting Juan de Fuca Plate—JP)

In the profile of Fig. 10, a prominent negative pulse observed beneath the regions labelled GP and SAF is indicative of a negative velocity gradient occurring in the 65–80 km depth range. This discontinuity has been interpreted by Li et al. (2007) to represent the Lithosphere–Asthenosphere Boundary (LAB).

In the approaches of 1-D RF imaging discussed above, the vertical resolution is proportional to the wavelength of the signal. The vertical resolution is defined here as both (i) the minimum distance between two discontinuities that results in distinct converted pulses, and (ii) the maximum thickness of a gradient in elastic properties that produces a converted

Fig. 11 Comparison between CCP-binning of receiver functions and teleseismic scattered-wave migration.▶ *Left column* 1-D model consisting of a 12 km thick horizontal layer embedded in a uniform half-space. *Right column* 2-D model consisting of a 12 km thick dipping layer embedded in a uniform half-space. Synthetic datasets for each model are constructed with the RAYSUM package (Frederiksen and Bostock 2000) for six incident *P*-waves propagating in the plane of the model at ray parameters 0.08, 0.06, 0.04, −0.04, −0.06 and −0.08 s/km (positive/negative values are assigned to waves incident from the left/right). The synthetic wavefields are recorded by a 500 km long array (horizontal axis) with a station spacing of 5 km. *Middle row* resulting CCP profiles. *Blue pulses* represent negative velocity contrasts (i.e., downward fast-to-slow), whereas *red pulses* represent positive velocity contrasts (i.e., downward slow-to-fast). Note that the CCP method recovers well the horizontal layer, but fails to stack coherently the signal from the dipping layer. *Bottom row* resulting migrated profiles. *Red* and *blue colours* represent slow and fast velocity perturbations (in terms of scattering potentials) relative to the background velocity model. Since the low-velocity layer is embedded in a uniform medium, it appears as a red layer in the migrated profiles. The migration technique recovers well both the horizontal and dipping layers

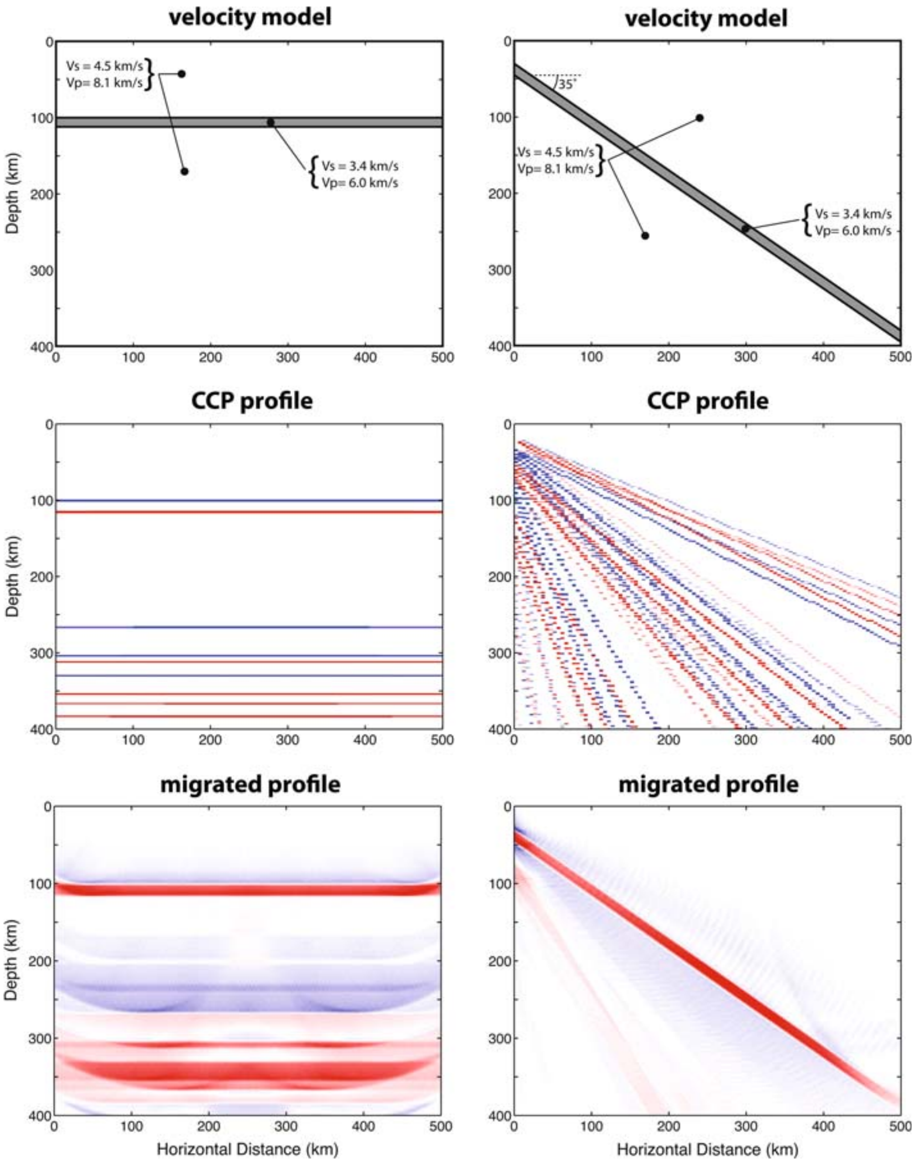
wave. Tests conducted on synthetic and field data have shown that this resolution corresponds to approximately half the wavelength ($\lambda/2$) of the converted wave (e.g., Bostock 1999; Rychert et al. 2007).

The profiles shown in Figs. 8–10 provide important constraints on crustal structure, but they rely on the fundamental assumption that discontinuities are horizontal. Violations of this assumption are likely to occur in the Earth, and may cause significant artifacts in the resulting profiles. This is similar to the limitations of CMP gathers in seismic reflection: because it is assumed that waves encounter horizontal specular reflectors, dipping layers will appear distorted (i.e., incorrect dip and lateral position) and lateral discontinuities will cause diffraction hyperbolae forming ‘bow-tie’ artifacts (e.g., Yilmaz 2001). This problem is addressed by treating 2-D and 3-D scattering with migration techniques (Fig. 11).

7 2-D/3-D Teleseismic Migration

Seismic migration involves the backpropagation to depth of a scattered wavefield recorded at the surface, with the expectation that the transposed wavefield will focus at (and therefore highlight) sources of seismic scattering in the subsurface (e.g., Yilmaz 2001). Whereas migration approaches have been developed and used in seismic reflection for more than half a century, they have been introduced in passive solid-earth seismology only over the past decade. This lag stems from the fact that the volumes of data required for migration techniques have only recently become available in passive broadband seismology. Starting in the mid 1990s, a surge in the availability and fidelity of broadband seismic recorders has allowed scientists to deploy dense station arrays yielding datasets amenable to multichannel approaches that had previously been restricted to controlled-source seismic imaging.

The main advantage of migration techniques over 1-D imaging is that they make much fewer simplifying assumptions about the geometry of subsurface structure. Instead of assuming that the Earth is made of planar, horizontal layers, teleseismic migration algorithms treat scattering from 2-D and 3-D structure, and thus consider seismic scattering in more general terms (Fig. 11). This does not mean that all the variables in the model become free parameters. Most migration techniques still rely on at least one fundamental assumption: that the background velocity model is smoothly varying and that it is known a priori. This velocity model is central to the backprojection of the scattered wavefield and therefore to the ability of the migration algorithm to illuminate structure accurately.



All the teleseismic migration techniques that have been developed to date consider only first order interactions between the incident wavefield and subsurface perturbations (i.e., they make use of the single scattering, or Born, approximation). One group of methods is based on acoustic scattering, and involves stacking of singly scattered wavefields along diffraction hyperbolae to recover relative scattering intensity/potential at individual points through a 2-D or 3-D model space (Revenaugh 1995; Ryberg and Weber 2000; Sheehan et al. 2000; Levander et al. 2005). Another group of methods based on elastic scattering combines stacking with inversion/backprojection operators (Bostock and Rondenay 1999; Bostock et al. 2001; Poppeliers and Pavlis 2003a, b) and full 3-D waveform inversion

(Frederiksen and Revenaugh 2004) of scattered waves to recover either scattering potential or estimates of localized material property perturbations relative to the background model.

In this section, we first describe teleseismic migration in general terms and then discuss techniques belonging to the two different groups defined above. We follow with a review of the resolution afforded by these techniques, with a particular focus on the benefits of including free-surface multiples in the analysis.

7.1 General Framework

Teleseismic migration techniques are usually deterministic. That is, the image is not produced by performing a blind backpropagation of the scattered wavefield, but rather by interrogating the data for potential scattered signal emanating from a regular grid of points in the subsurface—i.e., backpropagating the wavefield to specific points in model space. This procedure is represented schematically in Fig. 12 for P-to-S transmission scattering. It is based on the forward expression of the high-frequency, single-scattering approximation (Fig. 12a, b) and its inverse (Fig. 12c, d).

In the forward problem, a scattered wave observed at time $t_i = \mathcal{T}(\mathbf{x}, \mathbf{x}', \mathbf{p^0})$, by a station located at point $\mathbf{x}'$, for an incident P-wave with ray parameter $\mathbf{p^0}$ can be expressed as follows:

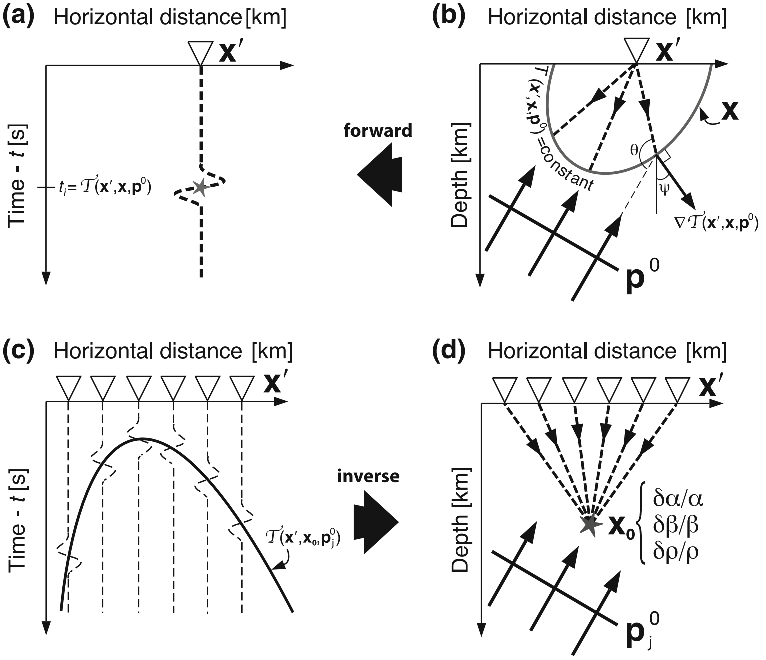


Fig. 12 Schematic representation of forward and inverse scattering problems for P-to-S scattering in terms of model space (i.e., scatterer location, *right column*) and data space (i.e., scattered wavefield as a function of station location and time, *left column*). **a**, **b** Forward problem expresses the scattered wavefield measured at a given time $t_i = \mathcal{T}(\mathbf{x}, \mathbf{x}', \mathbf{p^0})$ and station location $\mathbf{x}'$, due to all point scatterers in model space that could contribute energy to that point in data space. **b**, **c** Inverse problem involves weighted diffraction stacks along the travel-time curves corresponding to a point scatterer at $\mathbf{x}_0$ for incident waves with ray parameter $\mathbf{p}^0_j$. The backprojection operator derived in the inverse problem recovers either scattering potential or material property perturbations $(\delta\alpha/\alpha, \delta\beta/\beta, \delta\rho/\rho)$ at that point scatterer. Modified from Rondenay et al. (2005)

$$\Delta u(\mathbf{x}', \mathbf{p}^{\mathbf{0}}, t_i) = \int \mathrm{d}\mathbf{x} \mathcal{V}(\mathbf{x}, \mathbf{x}', \mathbf{p}^{\mathbf{0}}) f(\mathbf{x}) \delta[t - \mathcal{T}(\mathbf{x}, \mathbf{x}', \mathbf{p}^{\mathbf{0}})], \tag{12}$$

where Δu is the scattered wavefield, and $f(\mathbf{x})$ is the P-to-S scattering function or potential. $\mathcal{T}(\mathbf{x}, \mathbf{x}', \mathbf{p}^{\mathbf{0}})$ is the arrival time at receiver $\mathbf{x}'$ of energy scattered from $\mathbf{x}$, and $\mathcal{V}(\mathbf{x}, \mathbf{x}', \mathbf{p}^{\mathbf{0}})$ combines a series of filters and weights that include geometrical amplitudes of the incident and scattered waves. Thus, Eq. 12 means that the scattered signal observed at a given time t_i and receiver position $\mathbf{x}'$ in data space corresponds to the sum of contributions from all scatterers $\mathbf{x}$ in model space that can generate signal at that data point. These scatterers are distributed along an isochronal surface (i.e., $t_i = \mathcal{T}(\mathbf{x}, \mathbf{x}', \mathbf{p}^{\mathbf{0}})$ constant) which is defined by the ray kinematics in the reference background model. This expression was the basis for the first migration approaches used in controlled source seismology, in which the data samples were manually smeared onto isochronal surfaces corresponding to their arrival time to create images (for a historical perspective, see, e.g., Miller et al. 1987).

The inverse problem is applied to the scattered wavefield associated to one (or several) incident waves recorded at an array of stations. The data are weighted and stacked along diffraction hyperbolae corresponding to every potential scattering point along a regular grid within the model space. The diffraction hyperbolae are predicted based on the a priori background velocity model, and the value resulting from each weighted stack is mapped at its corresponding point in model space to create the image. In the most general terms, this process can be expressed as follows:

$$f(\mathbf{x}_0) = \sum_{j=1}^{N} \int \mathrm{d}\mathbf{x}' \mathcal{Y}(\mathbf{x}_0, \mathbf{x}', \mathbf{p}_j^{\mathbf{0}}) \Delta\tilde{u}[\mathbf{x}', \mathbf{p}_j^{\mathbf{0}}, t = \mathcal{T}(\mathbf{x}_0, \mathbf{x}', \mathbf{p}_j^{\mathbf{0}})], \tag{13}$$

where $f(\mathbf{x}_0)$ is the scattering potential at point $\mathbf{x}_0$ in model space, $\mathcal{Y}(\mathbf{x}_0, \mathbf{x}', \mathbf{p}_j^{\mathbf{0}})$ is a term that includes all weights and filters applied to the data (see Sects. 7.2 and 7.3 for details), and the sum is performed over N incident waves (i.e., N teleseisms). $\Delta\tilde{u}$ is the estimate of the scattered wavefield, e.g., the source-normalized L'–Q–T or P'–SV–SH wavefields obtained by preprocessing the data with the tools described in Sects. 3–5. The integral term over $\mathrm{d}\mathbf{x}'$ represents a stack over different stations of the weighted scattered wavefield at times $t = \mathcal{T}(\mathbf{x}_0, \mathbf{x}', \mathbf{p}_j^{\mathbf{0}})$ corresponding to the predicted diffraction hyperbola for point $\mathbf{x}_0$ and an incident wave defined by ray parameter $\mathbf{p}_j^{\mathbf{0}}$ (Fig. 12c, d).

Equation 13 is the basis for Kirchhoff-style migration. The two groups of methods described below, both of which rely on this equation, are distinct in how they weight the data prior to stacking. Until now, most of these approaches have been used to image 2-D structure due to the preponderance of linear over regional arrays. We therefore continue the discussion with a focus on 2-D imaging approaches, noting however that these methods are readily applicable to 3-D cases with only a minor generalization of the theory (but with a significant increase in data requirements).

7.2 Acoustic Scattered-Wave Migration

In acoustic scattered-wave migration, the incident and scattered waves are treated as scalar fields and the scattering function as a scalar operator. This means that for any type of incident wave impinging upon a point perturbation embedded in a uniform velocity model, all resulting scattered waves radiate energy uniformly in all directions. The scattering function is simply a real factor that is proportional to the magnitude of perturbation in material properties.

In the framework of Eq. 13, acoustic scattered-wave migration is performed by employing a weighting function that comprises the following elements (e.g., Revenaugh 1995; Sheehan et al. 2000): (1) a factor corresponding to the inverse of the geometrical spreading between the scattering point and the receiver, to account for the amplitude reduction with distance—i.e., proportional to $1/\sqrt{r}$ in 2-D and $1/r$ in 3-D; (2) a phase shift of $\pi/2$ for the treatment of 2-D scattering, to account for the Bessel function representing cylindrical wave propagation.

This approach has been applied to passive seismic data from local sources in a landmark study by Revenaugh (1995), and to teleseismic data for the imaging of lithospheric structure (Sheehan et al. 2000; Ryberg and Weber 2000) and discontinuities of the mantle transition zone (Levander et al. 2005). A profile produced with this approach is shown in Fig. 13. In this example from Levander et al. (2005), teleseismic data recorded across the Kaapvaal craton and neighboring terranes were migrated to image the discontinuities of the mantle transition zone beneath southern Africa. The 410 and 660 km discontinuities are clearly imaged across the profile, but shallower structure is not well resolved due to wide station spacing ($\sim$35 km) and Moho multiples.

The advantages of acoustic scattered-wave migration are the simplicity of the processing and the ability to treat simultaneously a large number of input data to produce images of scattering potential. A limitation of the approach is the lack of formal relationship between the scattering potential and actual material property perturbations in the scattering expressions (except for the P-to-P scattering mode, which is a scalar function of P-wave velocity perturbation). In particular, this limitation affects images generated by scattering modes involving S-waves (e.g., P-to-S), but can be addressed by solving the problem for elastic waves.

7.3 Elastic Scattered-Wave Migration/Inversion

Elastic scattered-wave migration is based on the elastic wave equation, in which P-waves are treated as a scalar field and S-waves as a vector field. The scattering function is now a tensor operator that relates each component of the incident wave to components of the

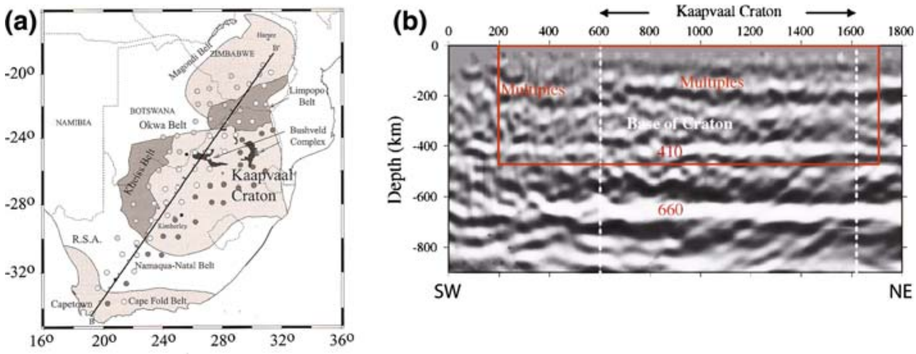


Fig. 13 Acoustic migration profile across southern Africa (modified from Levander et al. 2005). **a** Station distribution map. Stations are indicated by *circles* and the location of the profile is indicated by a *straight line*. **b** 2-D depth migration profile. *Gray shading* represents the magnitude of scattering potential, ranging from *white* (positive scattering amplitude, i.e., increase in velocity/impedance) to *black* (negative scattering amplitude, i.e., decrease in velocity/impedance). The 410 and 660 km discontinuities are clearly observed across the profile, whereas shallower discontinuities are obscured by Moho multiples

scattered wavefield. To construct it, the Green's function of the scattered wavefield is expressed as a Lippman-Schwinger volume integral, then simplified by adopting the Born approximation and the high-frequency ray approximation (e.g., Bostock and Rondenay 1999). The resulting scattering function is dependent on the angle θ between the incident and scattered wave, i.e., $f(\mathbf{x}) \to f(\mathbf{x},\theta)$, and takes the following form (e.g., for P-to-P and P-to-S scattering):

$$\begin{aligned} f^{P\to P}(\mathbf{x},\theta) =&\rho\left(2\frac{\delta\alpha}{\alpha}+\frac{\delta\beta}{\beta}\left(2\left(\frac{\beta}{\alpha}\right)^{2}(\cos 2\theta-1)\right)\right. \\ &\left.+\frac{\delta\rho}{\rho}\left(1+\cos\theta+\left(\frac{\beta}{\alpha}\right)^{2}(\cos 2\theta-1)\right)\right), \end{aligned} \tag{14}$$

$$f^{P\to S}(\mathbf{x},\theta) = \rho\left(\frac{\delta\beta}{\beta}\left(2\frac{\beta}{\alpha}\sin 2\theta\right)+\frac{\delta\rho}{\rho}\left(\sin\theta+\frac{\beta}{\alpha}\sin 2\theta\right)\right), \tag{15}$$

where α, β, and ρ represent the P-wave velocity, S-wave velocity, and density of the background model at point $\mathbf{x}$, respectively, and $\delta\alpha$, $\delta\beta$, and $\delta\rho$ are perturbations of these parameters at the same point. The corresponding radiation patterns of scattered waves are illustrated in Fig. 14. Note that the only mode of scattering that is accurately represented in the acoustic approach of Sect. 7.2 is the P-to-P scattering for a $\delta\alpha/\alpha$ perturbation.

The weighting function in Eq. 13 now comprises, in addition to the elements introduced in the acoustic treatment (i.e., geometrical spreading and phase shifting), a factor

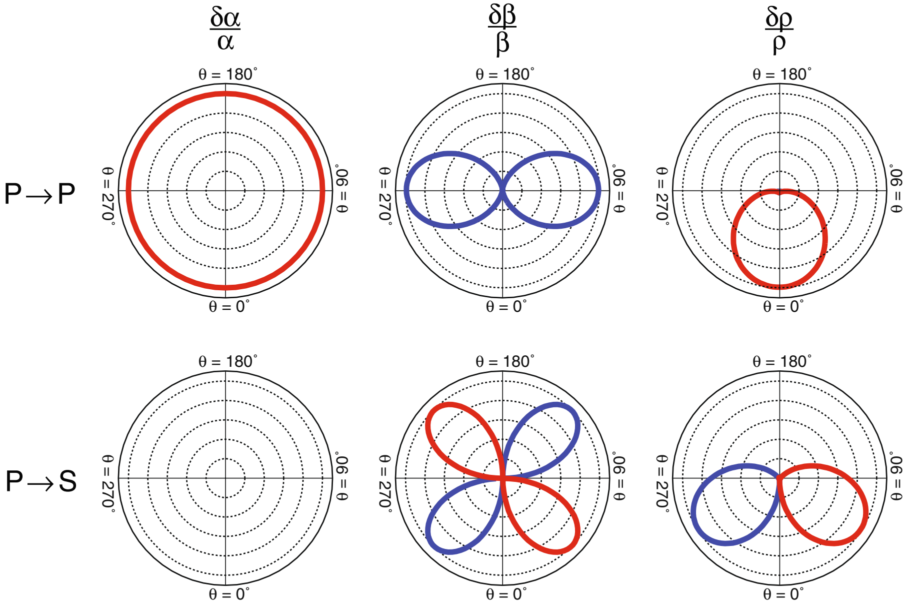


Fig. 14 Radiation patterns of P and S scattered waves produced by a vertically-incident P-wave illuminating point perturbations in seismic velocity ($\delta\alpha/\alpha$, $\delta\beta/\beta$) and density ($\delta\rho/\rho$) from below. *Polar plots* show the normalized amplitude of the scattered waves (*dotted circles mark* 0.2 amplitude increments from 0.0 in the center to 1.0 at the maximum amplitude), with *red*/*blue* denoting positive/negative amplitudes, as a function of the scattering angle θ (see Fig. 12, and Eqs. 14 and 15)

correcting for the variation in amplitude of the scattered wave as a function of the scattering angle θ. The explicit form of the backprojection operator in Eq. 13 can be found either by treating the scattering problem in terms of classical inverse theory (e.g., Bostock and Rondenay 1999), or by using the analogy between the forward scattering problem described by Fig. 12a, b and the Generalized Radon Transform (e.g., Miller et al. 1987; Bostock et al. 2001). In the latter, the backprojection operator is based on the inverse Generalized Radon Transform (GRT). Estimates of the material property perturbations are then obtained by solving a linear system based on Eqs. 14 and 15.

The elastic scattered-wave approach has been applied to teleseismic data to image structure in subduction zones (Rondenay et al. 2008), in the continental lithosphere (Poppeliers and Pavlis 2003a, b), and in the vicinity of the core–mantle boundary (Wang et al. 2008). These studies were conducted by different groups, each of which has developed its own flavour of GRT migration. Some authors consider only transmission scattering (Poppeliers and Pavlis 2003a), others consider only reflection scattering (Wang et al. 2008), and others use a combination of both (Bostock et al. 2001). For the purpose of lithospheric imaging, the forward scattered waves (i.e., transmissions) usually correspond to P-to-S waves, whereas the back-scattered waves (i.e., reflections) correspond to top-side reflections of free-surface multiples.

A 2-D GRT profile across the eastern end of the Alaska subduction zone is shown in Fig. 15 (see Rondenay et al. 2008). In these types of applications, the $\delta\beta/\beta$ images have been shown to be more robust than the $\delta\alpha/\alpha$ images (Rondenay et al. 2001, 2005). There are two main reasons for the superiority of $\delta\beta/\beta$ images. First, they are based on a larger number of scattered waves that afford superior volume and dip resolution (see Sect. 7.4). Second, they are based on S-scattered waves that are more accurately separated from the full recorded wavefield than P-waves (see Sect. 4). The $\delta\alpha/\alpha$ and $\delta\beta/\beta$ images of Fig. 15 both show a dipping low-velocity layer corresponding to the crust of the subducted slab. The low-velocities are attributed to hydrated minerals, and the termination of the layer at depth in the $\delta\beta/\beta$ profile indicates the location where dehydration reactions occur in the slab. The continental Moho is also clearly imaged as a positive (i.e., downward slow to fast) discontinuity observed at ~25 to 40 km depth across the profile.

The advantage of elastic scattered-wave migration/inversion is that it describes the scattering interaction with the elastic wave equation, which permits a more formal

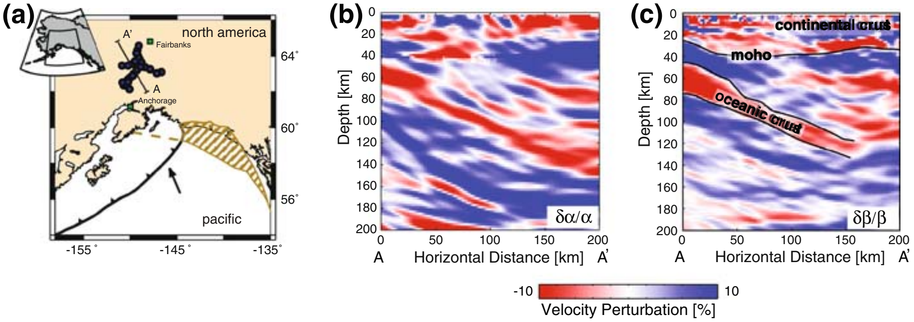


Fig. 15 2-D GRT migration profiles across the Alaska subduction zone (see Rondenay et al. 2008). **a** Station distribution map. Stations are indicated by *circles* and the location of the profile is indicated by a *straight line* labeled A–A′. **b** Profile of P-wave velocity perturbations. **c** Profile of S-wave velocity perturbations. *Red* to *blue colour* scale in profiles (**b**, **c**) represents negative (slow) to positive (fast) velocity perturbations. See text for interpretation

treatment of the problem and allows the estimation of actual material property perturbations in the subsurface. The cost of this approach, though, is that the data requirements become more stringent than for all other techniques discussed in previous sections. For the images to be robust, the sampling in sensitivity angle (ψ in Fig. 12) must be comprehensive to ensure a large range of dip resolution and to avoid signal aliasing which would prevent the recovery of perturbation estimates. Rondenay et al. (2005) provide a detailed discussion on the requirements, applicability and limitations of 2-D teleseismic GRT migration in terms of array design as a function of target depth and geometrical characteristics.

7.4 Resolution and the Benefits of Free-Surface Multiples

The resolution afforded by teleseismic acoustic/GRT migrations depends on a number of factors: station coverage, source coverage, frequency content of the scattered wavefield, validity of the assumptions (e.g., 2-D geometry, background velocities, isotropy). Means to keep track of these factors and assess how they affect the resolution of the imaged structures have been documented in Rondenay et al. (2005) and in publications discussing the various incarnations of the technique (e.g., Bostock et al. 2001; Poppeliers and Pavlis 2003a; Wang et al. 2008). Here, we focus the discussion on a quantity that is perhaps the best indicator of the migration's resolving power for a given set of field conditions: the spatial gradient of the total travel time function $\nabla\mathcal{T}(\mathbf{x},\mathbf{x}',\mathbf{p^0})$ (see Fig. 12). We first discuss the meaning of $\nabla\mathcal{T}(\mathbf{x},\mathbf{x}',\mathbf{p^0})$ and show how it can be mapped to assess the resolution at any given point in the model. We then show that the inclusion of free-surface (back-scattered) multiples in the analysis significantly improves the resolution of migrated images.

The spatial gradient of the total travel time function $\nabla\mathcal{T}(\mathbf{x},\mathbf{x}',\mathbf{p^0})$ (hereafter called 'traveltime gradient' and denoted $\nabla\mathcal{T}$ with implicit arguments) is a vector quantity that is obtained by summing the travel time gradients of the incident wave and the adjunct of the scattered wave intersecting at an image point, for a given station location. The vector is normal to the isochronal surface passing through the image point, for a given incident wave and station location (Fig. 12). It has been shown (see, e.g., Bostock et al. 2001; Rondenay et al. 2005) that the robustness of the migrated image relies on the coverage in $\nabla\mathcal{T}$. Specifically, the magnitude $|\nabla\mathcal{T}|$ represents the sensitivity of traveltime to scatterer location, whereas its orientation ψ controls dip resolution, such that structure dipping perpendicularly to sampling in $\nabla\mathcal{T}$ will be well resolved. Volume resolution is controlled by the wave-number $\omega|\nabla\mathcal{T}|$, where ω represents the angular frequency of the scattered signal (Bostock et al. 2001). Therefore, a high degree of 2-D spatial resolution will be ensured by large values of $|\nabla\mathcal{T}|$, a large signal bandwidth, and complete directional coverage in $\nabla\mathcal{T}$ (i.e., $\psi = [-\pi/2, \pi/2]$).

For an image point illuminated by several incident waves and/or sampled by an array of stations, the coverage in $\nabla\mathcal{T}$—and therefore the resolution at that point—can be assessed by plotting the vectors in an Ewald sphere (e.g., Rondenay et al. 2005). Figure 16 shows examples of Ewald sphere displays for a synthetic field setup where an incident wave from the left illuminates an image point situated in the centre of a 500 km long array, at 30 km depth. The station spacing is 5 km. Ewald spheres and corresponding $|\nabla\mathcal{T}| - \psi$ plots are shown for a series of scattering modes: $P_d\acute{s}$, $P\grave{p}_d\acute{s}$, $P\grave{s}_d\acute{s}$, and $P\grave{p}_d\acute{p}$—where the capital letter denotes the upgoing incident wave, the grave accent $\grave{}$ denotes the downgoing (free-surface reflected and possibly converted) wave, subscript $_d$ denotes the scattering interaction, and the acute accent $\acute{}$ denotes the upgoing scattered wave. For the forward scattered waves

(Fig. 16a), the theoretical setup provides a comprehensive range of dip resolution and a volume resolution proportional to 0.12ω–0.34ω km^{-1}. We note that the back-scattered waves associated with free-surface multiples afford a reduced range of dip resolution, i.e., the most extreme dips (>50°) will not be resolved accurately. However, the inclusion of back-scattered waves increases the volume resolution by a factor of ~3 to 5, especially for near horizontal discontinuities. Note also that because the backscattered waves afford a smaller range of dip resolution, they allow for a denser sampling in angle ψ at intermediate dips. Thus, they complement the sparser ψ sampling afforded in this range by $P_d\acute{s}$, and reduce the risk of spatial aliasing.

The increase in volume resolution afforded by multiples can be appreciated by inspecting the contribution from individual modes. Figure 17 shows the individual mode contributions for a dataset collected across central Alaska (same as composite mode images shown in Fig. 15). We note that the two main structures highlighted in the composite images, the subducted crust and the continental Moho, are observed coherently across the individual mode profiles (except for $P\grave{s}_d\acute{s}|_V$ in Fig. 17c), but that the $P_d\acute{s}$ image is significantly more diffuse than the images based on multiples. The $P\grave{s}_d\acute{s}|_V$ image, based on SV-to-SV back-scattering, does not resolve the structure well because all the teleseismic waves used in the analysis are obliquely incident to the profile and preferentially excite

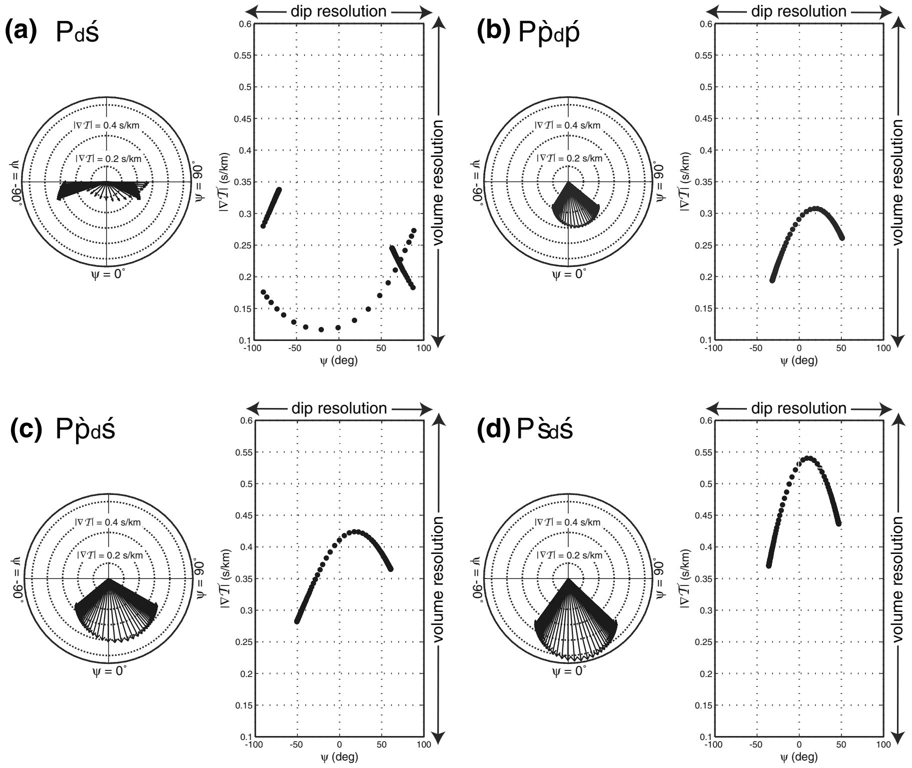


Fig. 16 Sampling in $\nabla\mathcal{T}$ for a single image point located at 30 km depth in the middle of a 500 km-long array, and for an incident wave from the left. Station spacing is 5 km. **a**–**d** $\nabla\mathcal{T}$ sampling for individual scattering modes. For each mode, the coverage in $\nabla\mathcal{T}$ is displayed in terms of an Ewald sphere (i.e., polar coordinates) in the *left panel*, and a plot of $|\nabla\mathcal{T}|$ versus ψ in the *right panel*

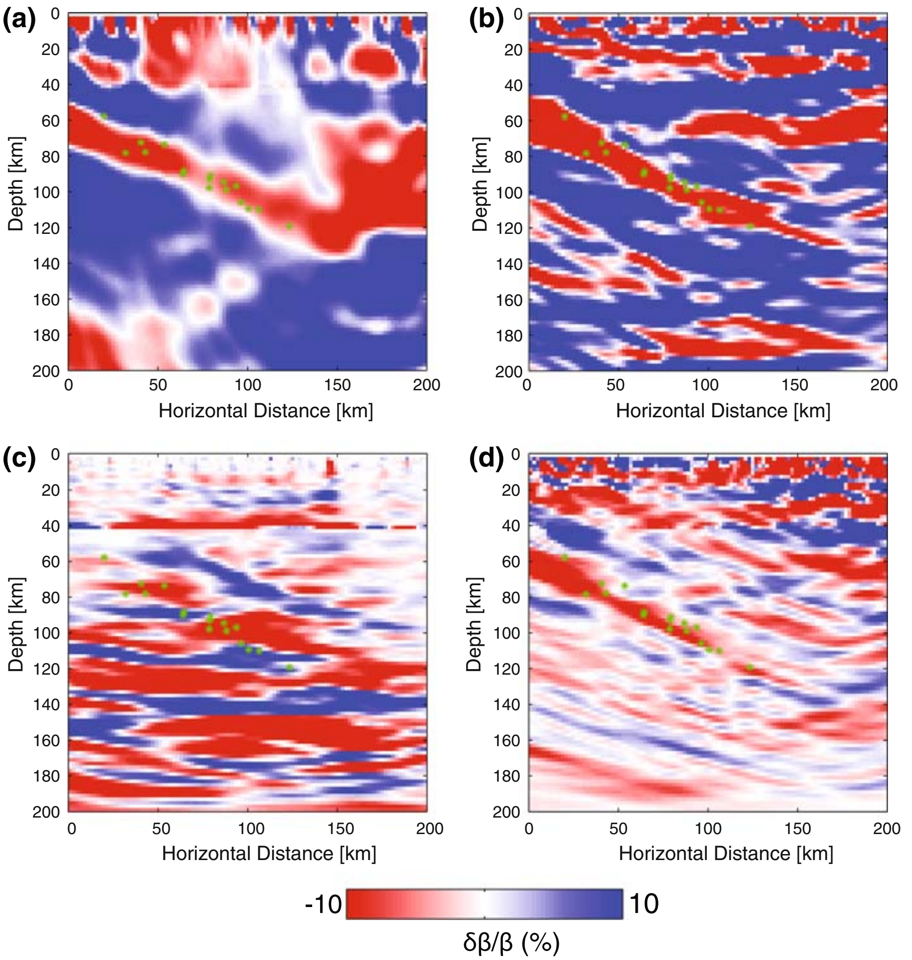


Fig. 17 Contributions of individual scattering modes in the 2-D GRT migration of teleseismic data from the Alaska subduction zone (see Fig. 15). The displayed modes are: **a** $P_d\acute{s}$, **b** $P\grave{p}_d\acute{s}$, **c** $P\grave{s}_d\acute{s}|_V$, associated with SV-to-SV back-scattered waves, and **d** $P\grave{s}_d\acute{s}|_H$, associated with SH-to-SH back-scattered waves. Note that the $P\grave{p}_d\acute{p}$ mode is not shown here because it was already displayed in Fig. 15b, as it is the only mode contributing to the $\delta\alpha/\alpha$ image. See text for analysis

SH-to-SH back-scattering. We further note, however, that the sections constructed using individual scattering modes suffer from cross-mode contamination, i.e., parallel echoes of the real structure. For example, the subducted crust observed at ~50 to 70 km depth on the left edge of the $P\grave{p}_d\acute{s}$ profile is repeated at ~100 to 120 km (Fig. 17b). This particular echo corresponds to the $P\grave{s}_d\acute{s}|_V$ scattered by the subducted crust. Fortunately, this cross-contamination is not as detrimental to the final results because the composite image tends to highlight features that are coherent from one mode to another (c.f., Fig. 15).

The synthetic and field examples presented above clearly demonstrate that the use of free-surface multiples in teleseismic migration significantly increases the resolution and robustness of upper mantle images.

8 Summary and Outlook

In this paper, we have provided a review of array-based imaging techniques that use converted and scattered teleseismic waves. These techniques form a continuum with respect to the level of complexity adopted in the treatment of the scattering problem. On one end of the spectrum, images may be produced by simple stacking of receiver functions, which are binned according to station or CCP and mapped to depth. Finer resolution can be achieved through the stacking of singly scattered wavefields along diffraction hyperbolae to recover relative scattering intensity/potential at individual points through a 2-D or 3-D model space. Moving to higher levels of complexity, we find methods that involve inversion/backprojection of scattered teleseismic wavefields to recover estimates of localized material property perturbations with respect to an a priori background model.

The existence of approaches that provide a more general treatment of seismic scattering does not make the simpler RF imaging techniques obsolete. Instead, the continuum of techniques described above should be seen as a set of tools that are designed for specific imaging objectives. This is especially important given the limitations in both hardware availability and the capacity to maintain large-scale, high station density networks. In this context, it is clear that studies aimed at investigating large-scale lithospheric structures (e.g., crustal depth variations over a 1,000 km × 1,000 km region) are better achieved currently with 1-D RF approaches applied to sparse arrays. Alternatively, high-resolution migration techniques require targeted local arrays of densely spaced stations.

Future developments in the field of teleseismic scattered-wave imaging are likely to focus on the following problems (see, e.g., Rondenay et al. 2005): (i) obtaining better estimates of the incident and scattered wavefields; (ii) 3-D imaging of upper mantle structure—including the design of appropriate recording arrays; (iii) separating completely the contribution of forward scattering modes from that of surface multiples; (iv) reducing the critical dependence of imaging algorithms on the assumed background velocity model; (v) addressing seismic anisotropy. Another important goal is to move from predictive migration approaches such as those reviewed here (which are computationally fast), to methods that backpropagate and/or invert the full wavefield recorded at the surface (which are computationally intensive; see, e.g., Pratt 1999; Tromp et al. 2008). All of these initiatives hold important potential for improving the robustness and the quality of upper mantle images obtained from teleseismic scattered waves.

Acknowledgements The author thanks Alan Levander and Xueqing Li for permission to use their figures and for providing electronic files of the originals; Karen Fischer, for insightful discussions on *Sp* receiver functions; Chuck Ammon, for input on the iterative time-domain deconvolution approach; and two anonymous reviewers for their thoughtful suggestions, which helped improve the original manuscript.

References

Aki K, Richards PG (2002) Quantitative seismology, 2nd edn. University Science Books, Sausalito, CA

Båth M, Stefánsson R (1966) *S–P* conversions at the base of the crust. Ann Geofis 19:119–130

Berkhout AJ (1977) Least-squares inverse filtering and wavelet deconvolution. Geophysics 42(7): 1369–1383

Bostock M (1999) Seismic imaging of lithospheric discontinuities and continental evolution. Lithos 48(1–4): 1–16

Bostock MG (1998) Mantle stratigraphy and evolution of the Slave province. J Geophys Res 103(B9): 21183–21200

Bostock MG (2004) Green's functions, source signatures, and the normalization of teleseismic wave fields. J Geophys Res 109:B03303. doi:10.1029/2003JB002783

Bostock MG, Rondenay S (1999) Migration of scattered teleseismic body waves. Geophys J Int 137:732–746

Bostock MG, Rondenay S, Shragge J (2001) Multiparameter two-dimensional inversion of scattered teleseismic body waves, 1, Theory for oblique incidence. J Geophys Res 106:30771–30782

Burdick LJ, Langston CA (1977) Modeling crustal structure through the use of converted phases in teleseismic body-wave forms. Bull Seismol Soc Am 67(3):677–691

Crotwell HP, Owens TJ, Ritsema J (1999) The TauP Toolkit: flexible seismic travel-time and ray-path utilities. Seismol Res Lett 70:154–160

Dueker KG, Sheehan AF (1997) Mantle discontinuity structure from midpoint stacks of converted p to s waves across the Yellowstone hotspot track. J Geophys Res 102:8313–8327

Dziewonski AM, Anderson DL (1981) Preliminary Reference Earth Model. Phys Earth Planet Int 25:297–356

Frederiksen AW, Bostock MG (2000) Modelling teleseismic waves in dipping anisotropic structures. Geophys J Int 141:401–412

Frederiksen AW, Revenaugh J (2004) Lithospheric imaging via teleseismic scattering tomography. Geophys J Int 159:978–990

Gurrola H, Baker GE, Minster JB (1995) Simultaneous time-domain deconvolution with application to the computation of receiver functions. Geophys J Int 120:537–543

Kennett BLN (1991) The removal of free surface interactions from three-component seismograms. Geophys J Int 104:153–163

Kikuchi M, Kanamori H (1982) Inversion of complex body waves. Bull Seismol Soc Am 72(2):491–506

Kind R, Kosarev GL, Petersen NV (1995) Receiver functions at the stations of the German Regional Seismic Network (GRSN). Geophys J Int 121:191–202

Kosarev G, Kind R, Sobolev SV, Yuan X, Hanka W, Oreshin S (1999) Seismic evidence for a detached Indian lithospheric mantle beneath Tibet. Science 283:1306–1309

Langston CA (1979) Structure under Mount Rainier, Washington, inferred from teleseismic body waves. J Geophys Res 84:4749–4762

Levander A, Niu F, Symes WW (2005) Imaging teleseismic P and S scattered waves using the Kirchhoff integral. In: Levander A, Nolet G (eds) Seismic Earth: array analysis of broadband seismograms, No. 157 in AGU Geophysical Monograph. AGU, Washington, DC, pp 149–169

Li X, Yuan X, Kind R (2007) The lithosphere-asthenosphere boundary beneath the western United States. Geophys J Int 170:700–710

Ligorría JP, Ammon CJ (1999) Iterative deconvolution and receiver-function estimation. Bull Seismol Soc Am 89(5):1395–1400

Miller D, Oristaglio M, Beylkin G (1987) A new slant on seismic imaging: migration and integral geometry. Geophysics 52:943–964

Park J, Levin V (2000) Receiver functions from multiple-taper spectral correlation estimates. Bull Seismol Soc Am 90(6):1507–1520

Poppeliers C, Pavlis GL (2003a) Three-dimensional, prestack, plane wave migration of teleseismic *P*-to-*S* converted phases: 1. Theory. J Geophys Res 108:2112. doi:10.1029/2001JB000216

Poppeliers C, Pavlis GL (2003b) Three-dimensional, prestack, plane wave migration of teleseismic *P*-to-*S* converted phases: 2. Stacking multiple events. J Geophys Res 108:2267. doi:10.1029/2001JB001583

Pratt RG (1999) Seismic waveform inversion in the frequency domain, Part 1: theory and verification in a physical scale model. Geophysics 64(3):888–901

Reading A, Kennett BLN, Sambridge M (2003) Improved inversion for seismic structure using transformed, *S*-wavevector receiver functions: removing the effect of the free surface. Geophys Res Lett 30(19): 1981. doi:10.1029/2003GL018090

Revenaugh J (1995) A scattered-wave image of subduction beneath the transverse ranges. Science 268: 1888–1892

Rondenay S, Bostock MG, Hearn TM, White DJ, Ellis RM (2000) Lithospheric assembly and modification of the SE Canadian Shield: Abitibi–Grenville teleseismic experiment. J Geophys Res 105(B6):13735–13754

Rondenay S, Bostock MG, Shragge J (2001) Multiparameter two-dimensional inversion of scattered teleseismic body waves, 3, Application to the Cascadia 1993 data set. J Geophys Res 106:30795–30808

Rondenay S, Bostock M, Fischer K (2005) Multichannel inversion of scattered teleseismic body waves: Practical considerations and applicability. In: Levander A, Nolet G (eds) Seismic Earth: array analysis of broadband seismograms, No. 157 in AGU Geophysical Monograph. AGU, Washington, DC, pp 187–204

Rondenay S, Abers GA, van Keken PE (2008) Seismic imaging of subduction zone metamorphism. Geology 36:275–278

Ryberg T, Weber M (2000) Receiver function arrays: a reflection seismic approach. Geophys J Int 141:1–11

Rychert CA, Fischer KM, Rondenay S (2005) A sharp lithosphere–asthenosphere boundary imaged beneath eastern North America. Nature 436:542–545

Rychert CA, Rondenay S, Fischer KM (2007) P-to-S and S-to-P imaging of a sharp lithosphere–asthenosphere boundary beneath eastern North America. J Geophys Res 112(B8):B08314. doi:10.1029/2007GL029535

Sheehan AF, Shearer PM, Gilbert HJ, Dueker KG (2000) Seismic migration processing of *P*–*SV* converted phases for mantle discontinuity structure beneath the Snake River Plain, western United States. J Geophys Res 105:19055–19065

Sipkin SA, Lerner-Lam AL (1992) Pulse-shape distortion introduced by broadband deconvolution. Bull Seismol Soc Am 82(1):238–258

Suckale J, Rondenay S, Sachpazi M, Charalampakis M, Hosa A, Royden L (2007) Imaging the southern hellenic subduction zone through migration of scattered teleseismic body waves. Eos Trans AGU 88(52):Fall Meet. Suppl., Abstract T51B-0558

Svenningsen L, Jacobsen BH (2004) Comment on Improved inversion for seismic structure using transformed, *S*-wavevector receiver functions: removing the effect of the free surface by Anya Reading, Brian Kennett, and Malcolm Sambridge. Geophys Res Lett 31:L24609. doi:10.1029/2004GL021413

Svenningsen L, Jacobsen BH (2007) Absolute *S*-velocity estimation from receiver functions. Geophys J Int 170:1089–1094

Tromp J, Komatitsch D, Liu Q (2008) Spectral-element and adjoint methods in seismology. Commun Comput Phys 3:1–32

Vinnik L (1977) Detection of waves converted from *P* to *SV* in the mantle. Phys Earth Planet Int 15:39–45

Vinnik L, Farra V (2002) Subcratonic low-velocity layer and flood basalts. Geophys Res Lett 29(4):1049. doi:10.1029/2001GL014064

Wang P, de Hoop MV, van der Hilst RD (2008) Imaging of the lowermost mantle (D$''$) and the core–mantle boundary with *SKKS* coda waves. Geophys J Int 175:103–115

Xu L, Rondenay S, van der Hilst RD (2007) Structure of the crust beneath the southeastern Tibetan Plateau from teleseismic receiver functions. Phys Earth Planet Int 165:176–193

Yilmaz O (2001) Seismic data analysis, 2nd edn. Society of Exploration Geophysicists, Tulsa, Oklahoma

Yuan X, Kind R, Li X, Wang R (2006) The *S* receiver functions: synthetics and data example. Geophys J Int 165:555–564