



Triplicated P-wave measurements for waveform tomography of the mantle transition zone

S. Stähler, K. Sigloch, T. Nissen-Meyer

► To cite this version:

S. Stähler, K. Sigloch, T. Nissen-Meyer. Triplicated P-wave measurements for waveform tomography of the mantle transition zone. *Solid Earth*, 2012, 3 (2), pp.339-354. 10.5194/se-3-339-2012 . hal-03425596

HAL Id: hal-03425596

<https://hal.science/hal-03425596>

Submitted on 30 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Triplicated P-wave measurements for waveform tomography of the mantle transition zone

S. C. Stähler¹, K. Sigloch¹, and T. Nissen-Meyer²

¹Dept. of Earth and Environmental Sciences, Ludwig-Maximilians-Universität (LMU), München, Germany

²Institute of Geophysics, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland

Correspondence to: S. C. Stähler (staehler@geophysik.uni-muenchen.de)

Received: 18 May 2012 – Published in Solid Earth Discuss.: 2 July 2012

Revised: 15 September 2012 – Accepted: 2 October 2012 – Published: 7 November 2012

Abstract. Triplicated body waves sample the mantle transition zone more extensively than any other wave type, and interact strongly with the discontinuities at 410 km and 660 km. Since the seismograms bear a strong imprint of these geodynamically interesting features, it is highly desirable to invert them for structure of the transition zone. This has rarely been attempted, due to a mismatch between the complex and band-limited data and the (ray-theoretical) modelling methods. Here we present a data processing and modelling strategy to harness such broadband seismograms for finite-frequency tomography. We include triplicated P-waves (epicentral distance range between 14 and 30°) across their entire broadband frequency range, for both deep and shallow sources. We show that it is possible to predict the complex sequence of arrivals in these seismograms, but only after a careful effort to estimate source time functions and other source parameters from data, variables that strongly influence the waveforms. Modelled and observed waveforms then yield decent cross-correlation fits, from which we measure finite-frequency traveltime anomalies. We discuss two such data sets, for North America and Europe, and conclude that their signal quality and azimuthal coverage should be adequate for tomographic inversion. In order to compute sensitivity kernels at the pertinent high body wave frequencies, we use fully numerical forward modelling of the seismic wavefield through a spherically symmetric Earth.

1 Introduction

The mantle transition zone (MTZ) is of great interest geodynamically, since its properties determine the extent to which material and heat gets exchanged between the upper and lower mantle. In the seismological view, the MTZ extends from the discontinuity at 410 km depth to one at 660 km – both discontinuities are characterized by marked jumps in seismic velocity. The sharpness and topographic undulations of these discontinuities can be linked to mineral physics laboratory experiments in order to infer material properties and mantle rheology.

The seismic waves that sample the MTZ most extensively are regional body waves, i.e., refracted waves that turn within the MTZ. Travelling only moderate distances of $\Delta \approx 1500$ –3100 km (i.e. 14–29°), these waves are recorded strong and clear on seismic stations, delivering by far the highest signal to noise ratio of any wave type that could be used to study the transition zone. Yet these phases have rarely been used in general, and for seismic tomography in particular, in sharp contrast to teleseismic phases ($\Delta > 30^\circ$). The reason is that these regional waves generate more complex signals than teleseismic ones as they have interacted more extensively with the MTZ discontinuities. Such observations do not lend themselves to abstraction into isolated pulses and the associated, idealized modelling by ray theory. Conversely we may suspect that if we succeed at modelling and inverting these waveforms, we will be able to learn a great deal about the mantle structures that are leaving such a strong imprint on them. Here we demonstrate that this should be possible.

Regional P and S waves are commonly termed triplicated waves, since every interaction with a discontinuity spawns three distinguishable phases. We investigate triplicated P-waves, which occur at epicentral distances of 14° to 29° , and sample the MTZ in some interval halfway between source and receiver. Our aim is to use them in finite-frequency waveform inversion for transition zone structure.

Finite-frequency modelling as originally conceived by Dahlen et al. (2000) is feasible across the entire relevant frequency range of body waves, but has been limited to interpreting direct and reflected teleseismic phases (Montelli et al., 2004; Sigloch et al., 2008; Zhou et al., 2004, 2006; Tian et al., 2011), due to the applicability limitations of paraxial ray tracing for computing sensitivity kernels. Here we present kernels that overcome this limitation. They are obtained from fully numerical forward computations of the seismic wavefield, using the spectral element code of (Nissen-Meyer et al., 2007b). Exploitation of the near-spherical symmetry of the 3-D Earth uses the currently available computational resources very efficiently, allowing access to the entire relevant broadband range of the wavefield (0.03 to 1 Hz), like the original Dahlen method.

Triplicated body waves carry a very strong imprint of their interaction with transition zone discontinuities, which is both an advantage and a challenge. Not only do the triplicated arrivals overlap each other in time, due to the finite-frequency nature of real data, but for shallow earthquakes they additionally overlap the depth phases pP and sP, which get triplicated themselves. Hence an integral part of modelling the waveforms is the careful estimation of source parameters (since they determine the shape of the synthetic Green's function), and of the source time function.

The differential moveout of the triplicated phases has since long been used to derive one-dimensional velocity models (Grand, 1984), which required a large number of seismic stations in a narrow azimuth range. Modelling of individual triplicated waveforms has been successfully used to sample localized heterogeneities in the MTZ (Tajima and Grand, 1995; Melbourne and Helmberger, 2002; Tajima et al., 2009). A few studies that included picked arrival times into ray theoretical inversions demonstrated the potential of triplicated waves for tomography (Grand, 2002), but so far they have rarely been used. Only very recently, Zhu et al. (2012) used triplicated phases, amongst others, for a European regional tomography, albeit at lower frequencies.

Regional body waves are complementary to all phases currently used for MTZ studies. Teleseismic body waves also offer good signal-to-noise ratios, but have comparably low sensitivity in the MTZ, which they traverse at steep angles. Hence they constrain the MTZ beneath sources and/or station, whereas triplicated waves sample it extensively midway between sources and receivers. So it would be highly beneficial to combine them in one inversion.

Other methods for sampling transition zone discontinuities have much lower signal-to-noise ratios (SNR). Receiver functions exploit P-to-S or S-to-P converted energy of body waves, which require stacking numerous seismograms. Their migration from the time domain to depth depends on a velocity structure model, which either must have been obtained independently or is neglected, with corresponding systematic errors in the result.

PP and SS precursors, i.e., body waves reflected at the undersides of MTZ discontinuities, have also been used in structural studies, (Shearer and Masters, 1992; Thomas and Billen, 2009; Deuss et al., 2006; Deuss, 2009). Their SNR is low, so that only the rare strong earthquakes that are recorded on seismic arrays can be used. Distinguishing between the signal of an undulated discontinuity and a volumetric velocity perturbation is challenging (Chaljub and Tarantola, 1997), as for receiver functions.

Surface waves have lower image resolution than body waves. Only their higher modes have significant sensitivity to the MTZ, but higher modes carry little energy, and are more difficult to process. Hence surface wave tomography is largely limited to depths above the MTZ.

We start with a discussion of the nature of triplicated body waves, and of their expression in actual seismic broadband data (Sect. 2.1). Section 2.2 and 2.3 demonstrate how we model these waveforms; this covers the computations of Green's functions and the inversion for source time functions and source mechanisms. We then explain the concept of wavefield kernels (Sect. 3), and show how passband-filtering to different frequency bands significantly increases the resolution in the transition zone, by comprehensively exploiting information across the entire broadband range (Sect. 3.3). The sizeable and well-instrumented continents of Europe and North America currently offer the most favourable source-receiver combinations of criss-crossing regional body waves for tomography. Section 4 discusses the two data sets that we have assembled for these two regions, with a focus on the *USArray*. We conclude with a discussion of the results, and the prospects of inverting these data for MTZ structure (Sects. 5 and 6).

2 Regional P-waves

2.1 Triplications

The term *triplication* refers to three seismic wave phases that have similar ray parameters and arrive closely spaced in time. The concept is rooted in ray theory. While our data processing and kernel computations are targeted at interpreting the full waveforms, it is useful to consider their ray theoretical approximation first, in order to appreciate the nature of triplicated waves. It assumes that the wave from one source to one receiver travels along an infinitesimally narrow ray path, which is described by the eikonal equation (e.g., Kennett,

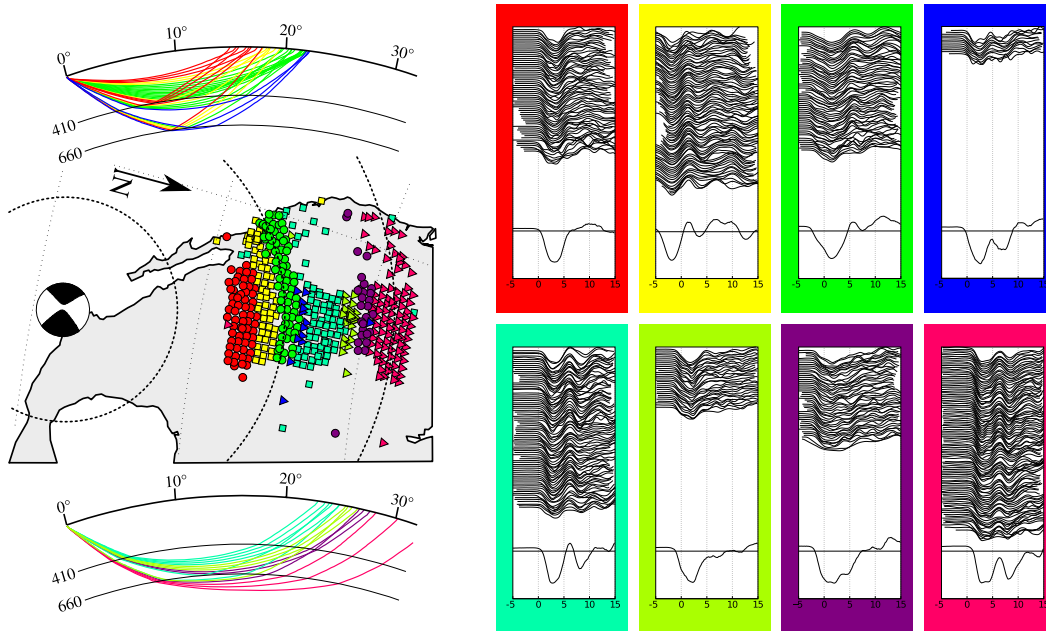


Fig. 1. Observed waveforms and nominal ray paths of regional P-waves, for a magnitude 6.4 earthquake off the coast of Mexico (2008/09/24 02:33:05, Lat: 17.61 Lon: −105.50, Depth: 6 km). Ray colouring in the left figure corresponds to the colouring of the waveform clusters in the right figure. Our automated clustering algorithm sorts the different kinds of triplicated phases into different groups, based on similarity of the broadband waveforms. The waveform below each seismogram group is the waveform stack. This empirical sorting into groups coincides with the epicentral distance bins predicted by theory (see text for details). This demonstrates that we are dealing with a robust signal that waveform tomography should be able to interpret.

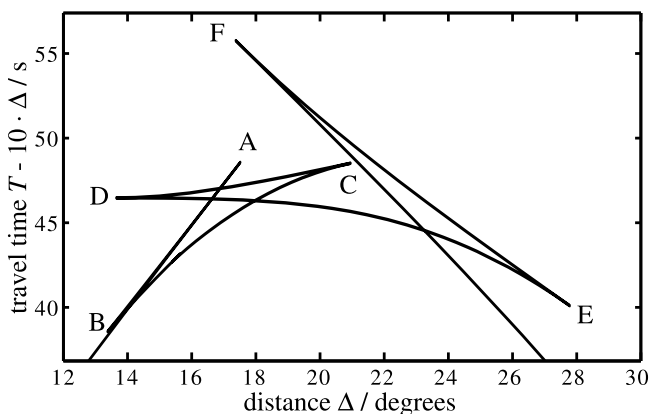


Fig. 2. Triplication of P-wave traveltimes in IASP91. The triplication C to D is produced by the discontinuity at 410 km depth, the one from E to F by the discontinuity at 660 km. The AB triplication originates from the 210 km discontinuity, which is only of second order in IASP91. This triplication can hardly be observed, especially since it overlaps with the one from 410 km.

2001). Strictly speaking, the approximation only holds for the case of a non-dispersive medium. Even when the medium is not intrinsically dispersive, velocity heterogeneities may introduce a frequency-dependent group velocity, which is the motivation for finite-frequency tomography.

In a layered velocity model $c(r)$, the traveltime τ and the angular distance Δ of a direct phase (e.g., a P-wave) depend only on the ray parameter \wp .

$$\Delta(\wp) = \int \frac{dr}{r \sqrt{r^2/(\wp c(r))^2 - 1}} \quad (1)$$

$$\tau(\wp) = \int \frac{dr}{c(r) \sqrt{1 - (\wp c(r))^2/r^2}} \quad (2)$$

If the velocity gradient with respect to depth is smooth, τ and Δ increase monotonously with decreasing \wp , i.e. stations at a larger distance from the earthquake record the P-phase at later times. If the velocity profile contains strong gradients, or even discontinuities with positive velocity jumps, $\frac{d\Delta}{d\wp}$ changes sign, so that rays of smaller \wp arrive closer to the source¹. Because $\frac{d\tau(\wp)}{d\wp} < 0$ in either case, two rays arrive at the same distance at different times. If $\frac{dc}{dr} < 0$, $\Delta(\wp)$ is still continuous. The turning points of $\Delta(\wp)$ mark the transition (with decreasing \wp) first from a ray travelling above the discontinuity to one being critically reflected at the discontinuity and then to one travelling below it. In the traveltime-distance relation $\tau(\Delta)$, which is also continuous, a pattern called triplication emerges (Fig. 2).

¹The literature does not provide any hard constraints for a triplication to occur. However, it is clear that a discontinuity in $c(r)$ or dc/dr will be sufficient.

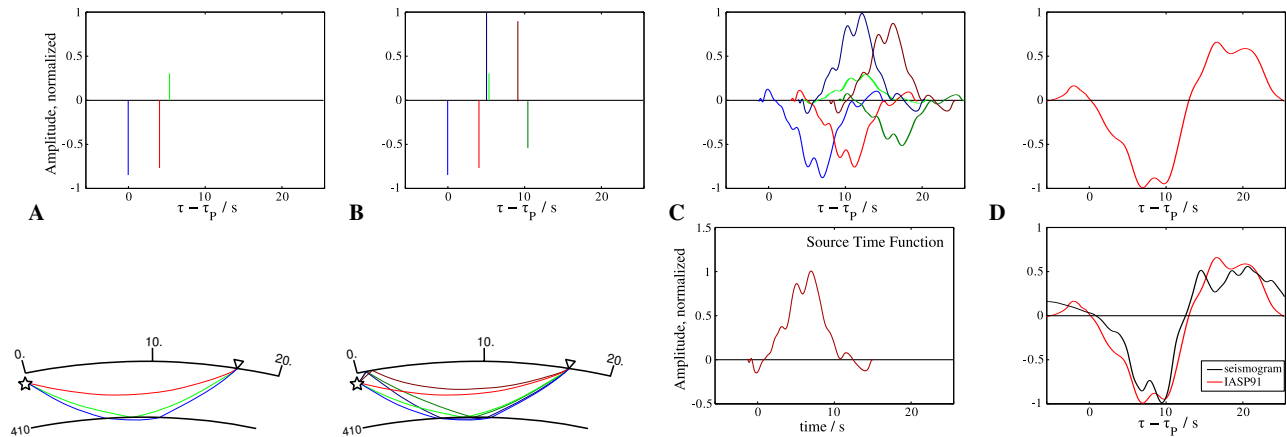


Fig. 3. Schematic composition of a triplicated broadband P-wave signal: **(A)** Arrival times (top row) and nominal ray paths (bottom row) of the direct P-waves. Three pulses arrive, two of which are refracted above and below the discontinuity (here 410), and the third reflected by it. **(B)** If the earthquake is shallow, additional depth phases arrive (pP and sP), which get triplicated themselves. **(C)** Real-world seismograms do not resemble sequences of dirac pulses. Rather, the pulses of the Green's functions are dispersed and the source time function (bottom row) convolves into all of them. **(D)** The sum of all contributions is the predicted seismogram (top), which may be compared to the observed seismogram (bottom row), in order to extract measurements for waveform tomography. Even though the waveform is complex, it can be modelled sufficiently well using a layered background model.

The velocity discontinuities in seismic velocity in the upper mantle (at 410 and 660 km depth in the *IASP91* reference model (Kennett and Engdahl, 1991), henceforth termed 410 and 660, produce triplications: Neglecting crustal phases and the weakly developed 210-km discontinuity, we receive one P phase up to a distance of 14°, which has travelled through the upper mantle. At 14° another P-phase arrives 8 seconds later (point D in Fig. 2), being reflected at the 410. This branch bifurcates into one reflected path (DC), and one refracted below (DE). The refraction branch continues to 28° distance (E), while the uppermost-mantle branch meets with the reflected one at 21.5° (C) and ends there. From 17.5° (F) on there are also reflection and refraction branches from the 660. The 660-reflection branch (EF) meets with the 410-refraction branch (DF) at 27.8° (E), where they both end. Only the 660 refraction branch can be measured at distances exceeding 27.8°.

With this conceptual knowledge, consider the real data in Fig. 1. It shows broadband P-waveforms in a 20 s window, starting 5 s before the first arrival predicted by *IASP91*. The 692 waveforms from stations in the western United States are aligned (VanDecar and Crosson, 1990) and clustered by similarity as described by Sigloch and Nolet (2006). This clustering results purely from the waveform shapes and does not use meta-information such as station distance or azimuth from the source. The clusters show a clear zonation by distance, caused by the triplications from the transition zone discontinuities. The figure can best be understood by first considering the most distant group of P-waveforms (pink). The stations are mainly located beyond 28° epicentral distance, so that the signal contains no triplications. The secondary pulse, arriv-

ing 7 s after the first one, can therefore be recognized as a reflection from the ocean surface near the source (pwP).

In the distance bin closest to the source, from ≈15° to 18° (red), the triplicated phases overlap so that just one broadened pulse arrives. In the second group (yellow), the arrivals are further spread so that a first arriving pulse, a surface reflected phase (at 5 s) and multiple triplicated phases can be distinguished. Additionally, we observe an arrival ≈10 s after the first one, which is a reflection from the 660 (FE branch in Fig. 2). In the third group (green), this reflection is already closer to the first arrival, and the first-arriving pulse is a superposition of the DE branch from the transition zone and the BC phase from the upper mantle and depth phases. The next group (dark blue) is quite heterogeneous and marks the transition from 410 triplications to 660 triplications.

In the cyan group, beginning at 22°, the refractions above and below the 660 arrive almost at the same time and collapse into a single pulse. The reflected phase is usually quite weak and not visible at all here. The two clearly separated arrivals hence correspond to P (two overlapping refractions) and pwP (another two refractions). In the next two groups (grass green and purple), the refracted phase from above and below the 660 begin to separate, so that they spill into pwP once more. The most distant group (pink) shows clearly separated P and pwP pulses no longer affected by triplications. The shape of this teleseismic waveform does not change much out to ≈85° distance.

Hand-picked arrival times of the different triplication branches have been used in a few tomographic studies (Grand, 1994, 2002), although this approach fails to take into account the finite frequencies of seismic waves. As Fig. 1 demonstrates, the triplications are often not clearly separated,

due to several factors. The finite duration of the source rupture has a low-passing effect. Earth's intrinsic attenuation disperses the P-wave pulse, which does not contain much energy above 1 Hz. Additionally, waves of a finite wavelength λ are influenced by scatterers off of the direct path, if the detour to reach the scatterer is less than $\lambda/2$, which is the definition of the Fresnel zone. This argument is continued in Sect. 3.

2.2 Waveform modeling

We download broadband seismograms from the *IRIS* and *ORFEUS* data management centres. The waveforms are corrected and band-passed between 0.01 Hz and 3.5 Hz, detrended, and transferred to ground displacement by deconvolving the instrument responses. A time window is cropped from 10 s before to 25 s after the theoretical P arrival. Noisy seismograms are singled out by a clustering algorithm and removed manually.

In order to do cross-correlation measurements for waveform tomography, it is necessary to compare a synthetic waveform to the observed seismogram. Synthetic seismograms are calculated using the reflectivity method of Fuchs and Müller (1971). As a reference model we use *IASP91* (Kennett and Engdahl, 1991), together with the density and intrinsic attenuation of *PREM* (Dziewoński, 1981).

Fig. 3 demonstrates the challenge of modelling regional P-waveforms. Within a short time window, three to five triplicated phases arrive (Fig. 3a). If the earthquake is shallow, as most earthquakes are, the surface-reflected phases pP, sP or pwP arrive within a few seconds of P, and are themselves triplicated (Fig. 3b). For an earthquake at 20° epicentral distance, ten phases arrive within less than 9 s, triplicated by both the 410 and the 660 discontinuities. The polarity of the reflected phases is negative compared to the refracted phases, and the depth phases may have reversed polarity depending on the source plane orientation. Hence the overall waveform is highly sensitive to the exact depth, mechanism, and distance of the earthquake.

Additionally, the finite duration of the source process is imprinted on the seismogram, which is the convolution of the moment rate function $\dot{m}(t)$, termed *source time function* (STF), with the Green's function $\mathbf{G}(\mathbf{r}_s, \mathbf{r}_r, \tau)$. Since duration of the STF is usually of several seconds, it can change the waveform completely.

If the maximum frequency in the seismograms is $f \ll 1/T$, where T is the duration of earthquake rupture, then the source time function is sometimes approximated by a Dirac delta, a triangular function or a Gaussian. This is usually done in surface wave tomography and long-period waveform inversion. We use earthquakes of moment magnitudes between 5.5 and 7.0, where T is between one and several tens of seconds, and we want to invert up to dominant periods of up to 1 s. Hence this approximation is too rough for our purposes, and we need explicit estimates of each STF in order to construct the matched filter.

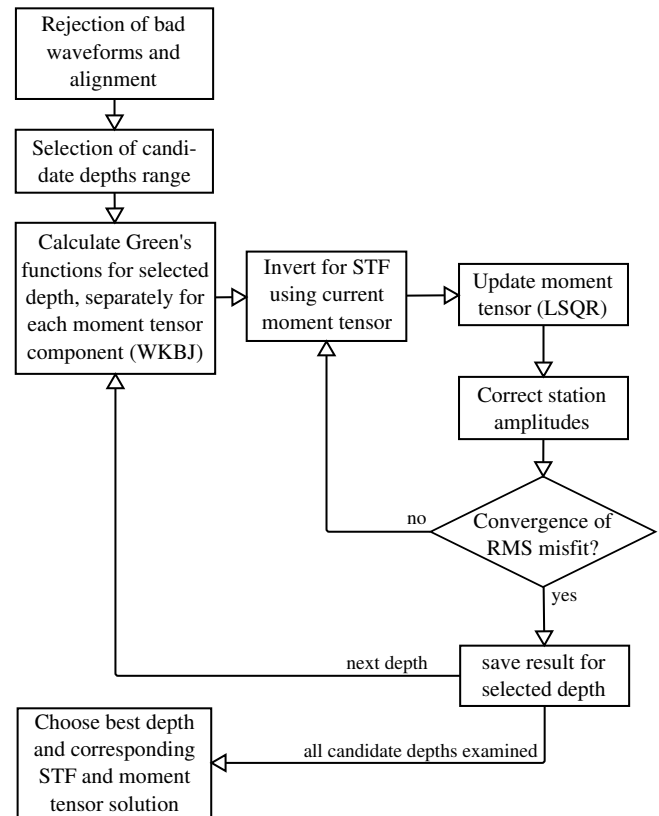


Fig. 4. Flowchart of the source inversion procedure as described in (Sigloch and Nolet, 2006).

2.3 Source inversion

As seen in the previous section, waveform tomography requires an inversion for the temporal and spatial parameters of the earthquake sources prior to the actual tomographic inversion. For the deconvolution of the STF, we use teleseismic P-waves rather than the complex, triplicated regional waves. We briefly describe the procedure here and in Fig. 4, for details see Sigloch and Nolet (2006): We remove obviously problematic stations and align all waveforms to the arrival of the P-phase (VanDecar and Crosson, 1990). We then choose a reasonable candidate depth range to survey, 1–50 km for shallow events and *NEIC* depth ± 30 km for deep events. Then we execute the following scheme for each candidate depth: First a joint deconvolution of the synthetic seismograms, calculated with the *NEIC* moment tensor M_0 from the measured seismograms is done, resulting in an STF estimate $\dot{m}(\tau)$. Source orientation is assumed to be constant during the rupture, so that $\dot{m}(\tau)$ is identical for all components of the moment tensor. Second, with this STF, an update for the moment tensor δM is calculated and the amplitudes of all stations are corrected individually. The new moment tensor $M_0 + \delta M$ and the amplitude corrections are used to derive a new STF estimation and this is repeated, until the

RMS misfit between synthetics and broadband seismograms has converged.

After all depths have been treated, we manually choose the depth, at which the RMS misfit is minimal and the STF does not contain any significant negative parts, which would be unphysical. The STF and moment tensor results for this “most likely” depth are retained for kernel calculation and tomography.

The teleseismic Green’s functions are calculated by the *WKBJ* code of Chapman (1978), using *IASP91* as reference model. For source inversion we use only waveforms from globally distributed, reliable broadband stations (*GSN* and *Geoscope* networks). For typical earthquakes, around 40 of these stations are located within teleseismic range. Since STF deconvolution is a numerically sensitive operation, we deliberately use only this small, high-quality ensemble, rather than all available broadband stations. The spatial distribution of these permanent, international network stations is relatively even, whereas a deconvolution from all *IRIS* stations would always be dominated by the 1000+ receivers located in North America. Their waveforms and misfits are highly correlated, so that the additional information content with regard to the source is low. However, all available stations will later be used for tomography.

We always attempt to invert for a single STF that fits all global teleseismic data, but in about one out of three earthquakes, we need to allow for two or more regional STF solutions, e.g., when the European station cluster cannot be fit to the same STF as the North American cluster. This may be due to structure close to the source, like a subducting slab or be an effect of source directivity. Our view of a “source time function” is pragmatic: we want it to absorb all signal that is common to all seismograms, even when that signal does not derive from the source rupture *sensu stricto*. The remaining signal can then be interpreted as an imprint of the structure along the wave-path.

3 Sensitivities

3.1 Waveform tomography – direct phases

The concept of arrival time of a phase, which may be picked manually by an analyst, assumes a broadband minimum-phased signal. In a heterogeneous Earth, scattering off of the “direct” path adds a small, frequency-dependent component to the direct waveform. Hence scattering introduces a non-linear dispersion even if the medium is perfectly viscoelastic (Dahlen et al., 2000; Nolet, 2008). This dispersion reflects the scale-dependent interaction of different finite wavelengths with the mantle heterogeneities that we want to image. Hence it embodies the information that waveform inversion captures above and beyond the ray theoretical approximation. We measure this dispersion by the method of matched filtering (Sigloch and Nolet, 2006).

A predicted waveform $\mathbf{u}(t)$ is synthesized from six partial Green’s functions $\mathbf{G}_j(\mathbf{r}_s, \mathbf{r}_r, t)$, weighted by the six independent components of the moment tensor M_j , and convolved by the source time function $\dot{m}(t)$. The broadband $\mathbf{u}(t)$ is immediately bandpass filtered to $\mathbf{u}_k(t)$, through convolution with a filter response $f_k(t)$ (k is the frequency band index). Hence the finite-frequency synthetic is

$$\mathbf{u}_k(t) = \sum_{j=1}^6 \mathbf{G}_j(\mathbf{r}_s, \mathbf{r}_r, t) \cdot \dot{m}(t) \cdot M_j \cdot f_k(t), \quad (3)$$

which must be compared to an accordingly filtered observed waveform $\mathbf{u}_k^o(t)$. For this we parametrize $\mathbf{u}_k(t)$ by the two observables (misfit measures) that we want to estimate: the traveltime anomaly δT_k and the amplitude anomaly δA_k ,

$$\hat{\mathbf{u}}_k(t) = A_k \mathbf{u}^o(t - \delta T_k). \quad (4)$$

$\hat{\mathbf{u}}_k(t)$ is the matched filter, and the optimal δT_k , δA_k are obtained by minimizing the RMS misfit between $\hat{\mathbf{u}}_k(t)$ and $\mathbf{u}_k^o(t)$. This is equivalent to finding the time shift δT that maximizes the cross correlation between $\hat{\mathbf{u}}_k(t)$ and $\mathbf{u}_k^o(t)$.

With this method, we obtain up to $2 \cdot k$ frequency-dependent misfit observables δT_k , δA_k from one broadband P-wave seismogram. The derivative of these misfits with respect to the Earth model can be expressed in so-called *sensitivity kernels* $K_i(\mathbf{r}_x)$. These represent the sensitivity of the traveltime misfit or the amplitude misfit with respect to changes in P-wave velocity $\Delta V_P / V_P$ at a given point \mathbf{r}_x . The traveltime anomaly δT_k is then modelled as

$$\delta T_k = \int_{\oplus} K_k(\mathbf{r}_x) \frac{\Delta V_P}{V_P}(\mathbf{r}_x) d^3 \mathbf{r}_x. \quad (5)$$

Obviously, this kernel has to be calculated using a reference velocity model. Equation 5 and the construction of the kernel K_k contain the assumption that the traveltime anomaly δT_k originates exclusively from single scattering off of anomalies $\Delta V_P / V_P$, which quantify the difference between the reference model and true Earth structure. This so-called Born approximation is justified by the observation that the magnitude of lateral mantle heterogeneities $\Delta V_P / V_P$ is small, typically on the order of a few percent, so that multiple scattering can be neglected. Kernel K_k is the first Fréchet derivative of δT_k towards $\Delta V_P(\mathbf{r}_x)$ (Marquering et al., 1999).

Dahlen et al. (2000) proposed a fast algorithm to derive kernels for teleseismic body waves by paraxial ray tracing. This effectively uses ray theory to go beyond the limits of ray theory, since it synthesizes sensitivity kernels of finite volume from the interaction of an infinity of rays. The method has been applied to teleseismic body waves with great success (Montelli et al., 2004; Sigloch et al., 2008) and was extended to surface waves (Zhou et al., 2004, 2006; Tian et al., 2011) but its scope is limited to phases without caustics, diffractions or other wave effects (Dahlen et al., 2000; Nissen-Meyer et al., 2007a).

3.2 Waveform tomography – arbitrary phases

Triplicated waveforms, with their interactions of refracted and reflected phases around the discontinuity, are not adequately modelled by the ray theoretical formalism. The caustics at turning points A–F (Fig. 2) would lead to infinite amplitudes at the corresponding distances. Moreover, the ray-centred approach by Dahlen et al. (2000) works strictly speaking only in a continuous velocity model. Hence we need to calculate the kernels from the full wavefield instead.

Computing the first order perturbation $\delta \mathbf{u}(t)$ of a seismic waveform $\mathbf{u}(t)$ involves the wavefield from source to every possible scattering location $\vec{\mathbf{u}}(t, \mathbf{r})$, and of the scattered wavefield to the receiver. Thanks to source-receiver reciprocity, this second wavefield may instead be replaced by the back-propagating wavefield $\overleftarrow{\mathbf{u}}(t, \mathbf{r})$ from receiver to scatterers, which is generally much cheaper computationally (Nissen-Meyer et al., 2007a, eqn. 10). This is conceptually similar to the adjoint method (Tromp et al., 2005).

$$\delta u = - \int_{\oplus} \left[\delta \rho \vec{v}_i \cdot \vec{v}_i + \vec{E}_{ij} \cdot \overleftarrow{E}_{kl} \delta C_{ijkl} \right] d^3 \mathbf{r}. \quad (6)$$

$\vec{v} = \partial_t \vec{\mathbf{u}}$ is the velocity field of the forward propagating wave and $\vec{E}_{ij} = \frac{1}{2} \left[\frac{\partial u_i}{\partial r_j} + \frac{\partial u_j}{\partial r_i} \right]$ is the strain tensor. Note that this is index notation, so summation over repeat indices is implied. The arrows serve as a reminder of the forward and backward nature of the wavefields and strains. For an isotropic medium, the dependence on the Lamé parameters is

$$\delta u = - \int_{\oplus} \left[\delta \rho \vec{v}_i \cdot \vec{v}_i + \vec{E}_{ii} \cdot \overleftarrow{E}_{jj} \delta \lambda + 2 \vec{E}_{ij} \cdot \overleftarrow{E}_{ij} \delta \mu \right] d^3 \mathbf{r}. \quad (7)$$

If the inversion is for δV_P and δV_S rather than for the Lamé parameters $\delta \lambda$, $\delta \mu$, we replace them:

$$\delta \mu = V_S^2 \delta \rho + 2 V_S \rho \delta V_S \quad (8)$$

$$\delta \lambda = \left(V_P^2 - 2 V_S^2 \right) \delta \rho + 2 V_P \rho \delta V_P - 4 V_S \rho \delta V_S \quad (9)$$

Replacing $\delta \mu$ and $\delta \lambda$ in Eq. (7), we have

$$\begin{aligned} \delta u = & \quad (10) \\ & - \int_{\oplus} \left[\delta \rho \left(\vec{v}_i \cdot \vec{v}_i + \left(V_P^2 - 2 V_S^2 \right) \cdot \vec{E}_{ii} \cdot \overleftarrow{E}_{jj} + V_S^2 \cdot \vec{E}_{ij} \cdot \overleftarrow{E}_{ij} \right) \right. \\ & + \delta V_P \left(2 \rho V_P \cdot \vec{E}_{ii} \cdot \overleftarrow{E}_{jj} \right) \\ & \left. + \delta V_S \left(-4 \rho V_S \cdot \vec{E}_{ii} \cdot \overleftarrow{E}_{jj} + 2 \rho V_S \cdot \vec{E}_{ij} \cdot \overleftarrow{E}_{ij} \right) \right] d^3 \mathbf{r}. \end{aligned}$$

Equation 11 shows that V_P -kernels are straightforward to calculate because they only involve the diagonal elements of the strain tensor, which is equivalent to the divergence of the displacements: $\vec{E}_{ii} = \frac{\partial u_i}{\partial r_i} = \nabla \cdot \vec{\mathbf{u}}$. Hence for P-wave tomography,

we need to store only the displacement field $\mathbf{u}(\mathbf{r})$, rather than the full strain tensor $E_{ij}(\mathbf{r})$. The exact expression of the kernel now depends on the chosen misfit criterion. We prefer the cross-correlation traveltime misfit measured on one component i (in the case of P-waves usually BHZ), defined as

$$\delta T = - \frac{\Re \int_0^\infty \omega i \overline{u_i(\omega)} \delta u_i(\omega) d\omega}{\int_0^\infty \omega^2 \overline{u_i(\omega)} u_i(\omega) d\omega}. \quad (11)$$

With $u_i(\omega) = G_{ij}(\mathbf{r}_s, \mathbf{r}_r, \omega) \dot{m}(\omega) M_j$, we can calculate the sensitivity of δT w.r.t. δV_P . Using the definition of the kernel in Eq. (5), we arrive at

$$\begin{aligned} K(\mathbf{r}_x) = & \sum_{j=1}^6 \frac{2 V_P(\mathbf{r}_x) \rho(\mathbf{r}_x)}{\int_0^\infty \omega^2 |G_{sr,ij}(\omega) \dot{m}(\omega)|^2 d\omega} \cdot M_j \\ & \cdot \Re \int_0^\infty i \omega \overline{G_{sr,ij}(\omega)} |\dot{m}(\omega)|^2 \\ & \nabla \vec{G}_j(\mathbf{r}_x, \omega) \nabla \overleftarrow{G}_j(\mathbf{r}_x, h\omega) d\omega. \end{aligned} \quad (12)$$

The term $|\dot{m}(\omega)|$, which Dahlen et al. (2000) originally denoted as the source term, contains the source spectrum, but also the bandpass filters. $\mathbf{G}_{sr}(\omega)$, the Green's function from source to receiver, introduces any intrinsic attenuation of the reference Earth model into the kernel.

3.3 Spherical Earth kernels

The calculation of a global wavefield at a dominant period of 5 s requires around 10^4 CPUh when a full 3-D forward solver like *SPECFEM* (Komatitsch and Tromp, 2002) is used. For a realistic iterative global tomography using $> 10^5$ waveforms, the calculation cost would be $\gg 10^9$ CPUh, which is completely prohibitive. In a spherically symmetric background model, this cost can be reduced dramatically: 3-D wavefields in a layered Earth can be computed at the cost of the equivalent 2-D wavefields, since the symmetry implies that one dimension may be calculated analytically (Nissen-Meyer et al., 2007a). Forward wavefields are pre-computed for reasonable depth increments (e.g. increments of 1 km for depths from 0 to 100 km, and increments of 10 km from 100 to 700 km, requiring 160 simulations). The backward wavefield needs to be calculated just once, assuming that all receivers are located at the surface or within one P-wavelength of it. The 160 forward calculations need to be done four times (Nissen-Meyer et al., 2007a, p. 1057) for

1. a M_{zz} monopole source
2. a $\frac{1}{2} (M_{xx} + M_{yy})$ monopole source
3. a M_{xz} dipole source
4. a M_{xy} quadrupole source.

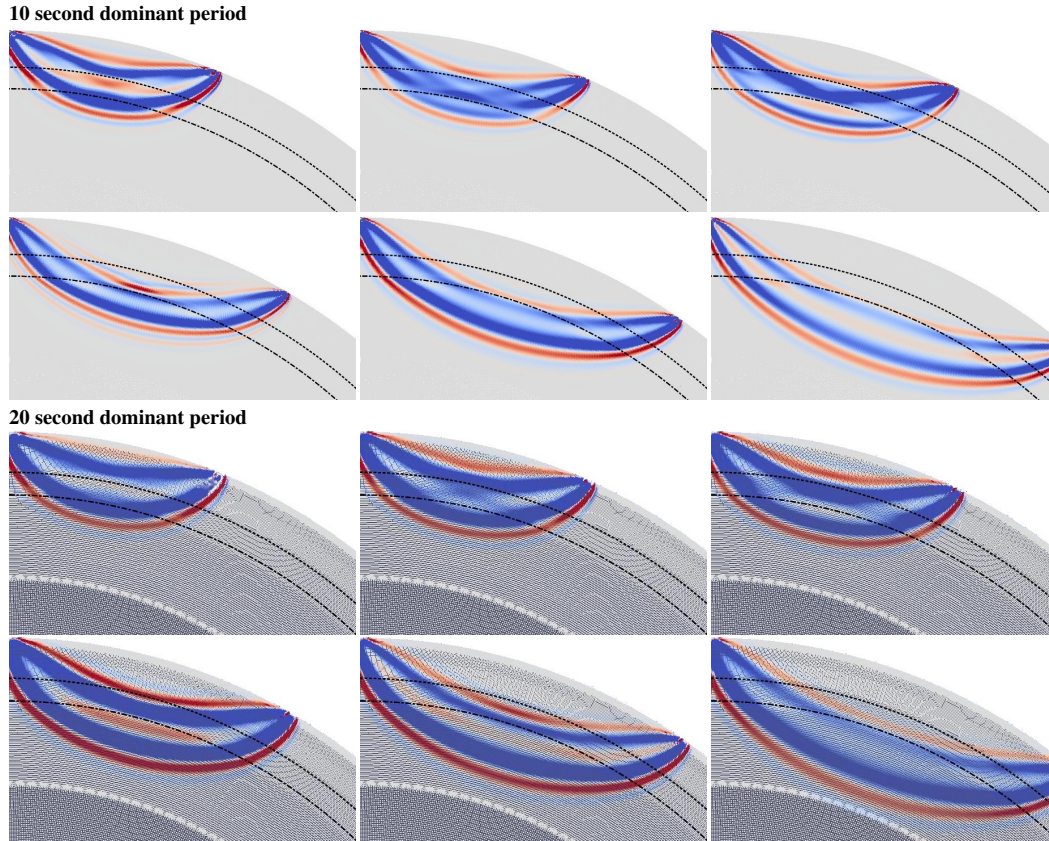


Fig. 5. Traveltime sensitivity kernels for $\Delta V_p/V_p$, calculated by the Axisem spectral element code. Sensitivities are for a cross-correlation traveltime misfit, in the time window of 5 s before to 15 s after the estimated arrival. Dominant period is 10 s for the upper two rows, and 20 s for the lower two rows. Distances (from top left): 22°, 24°, 26°, 30°, 35°, 40°. Background model: IASP91; source: explosion; receiver: z-component. The 410-discontinuity is marked by the dashed line, the 660 by the dashed-dotted line. Note that most kernels do not feature the famous doughnut hole, which is filled in by the sensitivities to the two discontinuities.

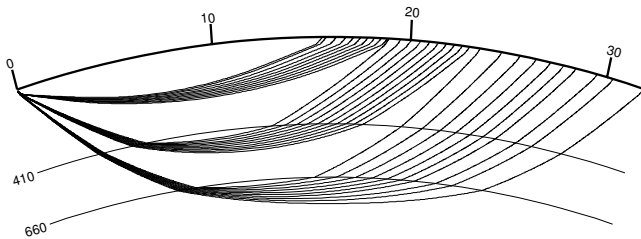


Fig. 6. Nominal ray paths of first arriving P-waves in IASP91. None of the rays bottom out in the lower part of the transition zone. Contrast this to the realistic (finite-frequency) sensitivities of Fig. 5, which extend broadly across the entire MTZ.

Since the response of a station at azimuth ϕ to a M_{xz} -source equals that of a station at $\phi - \pi/2$ to a M_{yz} -source, and response of a $(M_{xx} - M_{yy})$ -source at ϕ equals that of a M_{xy} -source at $\phi - \pi/4$, we can reconstruct the response of a spherically symmetric Earth to an arbitrary moment tensor from these four calculations. The backward propagation needs to be done twice:

1. p_z monopole source for the Z-component of the seismogram
2. p_x or p_y dipole source for the E- or N-component of the seismogram

Again, p_x or p_y are equivalent if the receiver location is rotated by $\pi/2$. Using the 3-D to 2-D reduction strategy applied in the SEM-code Axisem (Nissen-Meyer et al., 2007b, 2008), the computation cost for a 3-D global wavefield of dominant period $T = 5$ s is around 16 CPUh on an i7 machine, or $\approx 10^4$ CPUh for the entire wavefield library. Since the calculation grows at $O(1/T^3)$, a library of dominant period 1.25 s would still require less than 10^6 CPUh, which is easily feasible nowadays. From these 642 pre-computed wavefields, kernels for any arbitrary time window in the seismogram can be calculated on the fly in the frequency domain, which is memory intensive but computationally cheap.

Time window choices for these kernels are not limited to any specific phase arrivals or wave types, although it is certainly judicious to choose parts of the seismogram that contain significant seismic energy rather than noise. The window

length should be chosen with care. A window that is too long will contain more noise than necessary, and the diameter of the kernel will grow, since longer detours from the direct path are allowed. The outer parts of the kernel will contain the higher-order Fresnel zones, which oscillate rapidly and are unlikely to meaningfully contribute in the inversion. Windows too short will contain too little signal and be more prone to cycle skips. The window length should probably be no shorter than twice the dominant period of the bandpass filter. It will not be possible to calculate kernels separately for each of the triplicated phases, since in real data they overlap in time (cf. Fig. 3). However, it is neither necessary or even desirable to revert to this Dirac-type, ray-theoretical way of processing. The kernel formalism ensures the proper interpretation of the waveforms, as long as the actual measurements use the same window lengths and filters as the wavefield computations.

3.4 Kernel gallery

A selection of V_p -kernels at dominant periods of 10 s and 20 s are plotted in Fig. 5, for an explosive source at the surface and the z-component of the seismogram. Especially at 24° epicentral distance, the two frequency bands differ significantly in their sampling of the MTZ. While the 20 s kernel samples the regions directly below the 410 and the 660, the 10 s kernel samples mainly the region above the 660. This is particularly striking when compared to the sensitivity of the first arriving rays (Fig. 6), which is completely confined to the region directly below the discontinuities. The 22° and 26° kernels in the 10 s band have nearly complementary sampling characteristics in the transition zone. The 22° kernel is negative where the 26° one has its largest positive values. With dense arrays like the *USArray*, we obtain many 22° and 26° recordings for the same earthquake, which thus should yield good constraints on the MTZ. The 30° to 40° kernels in the second and forth row look more similar to teleseismic kernels, even though clear imprints of the MTZ remain even for a 40° kernel, resulting in an appearance quite different from the 40° kernel of the original Dahlen method (Dahlen et al., 2000). Geodynamically interesting regions just above the discontinuities, where effects of fluid release might be present (Ohtani, 2005), can classically be sampled only by teleseismic waves. Since they traverse the MTZ at steep incidence angles, their vertical resolution is quite limited, in contrast to triplicated phases. Topography on the discontinuities may also influence the waveforms, an effect for which sensitivities may be computed from the same wavefields (Nissen-Meyer et al., 2007a; Colombi et al., 2012). Ultimately it will be desirable to separate topography effects from those of the volumetric velocity structure, by integrating both kinds of kernels into a joint inversion.

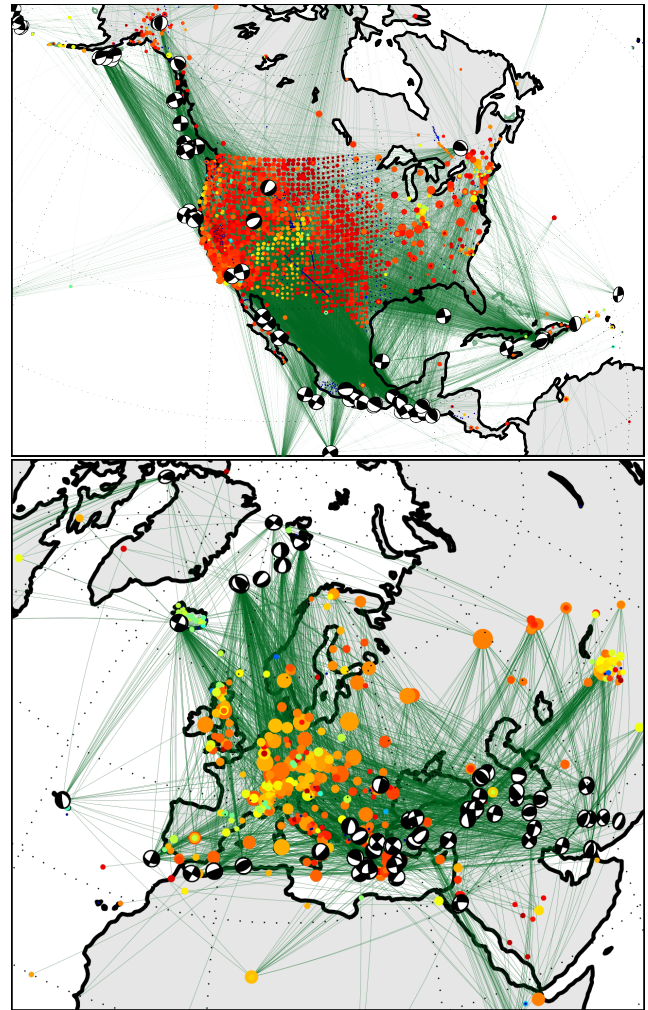


Fig. 7. Source and stations in two regional data sets for North America and Europe. The rays connect source-receiver combinations at distances between 15 and 35 degrees. Earthquake locations are shown as beachballs. Station colour codes for the goodness of fit between to observed and modelled waveforms at each station. Red denotes median cross-correlation coefficient of $CC_{\max} > 0.8$, which we consider a reasonable threshold for accepting any individual measurement. Orange and yellow colours indicate a lower median fit, which means that only some waveforms at this station may be used. Marker size is proportional to the number of events recorded by the station.

4 The data sets

We identified two regions that appear particularly suited to tomography of the transition zone using triplicated body-waves: North America and Europe. They are densely instrumented, decently surrounded by earthquake sources, and large enough for wave paths to penetrate the MTZ on their way from source regions to receivers.

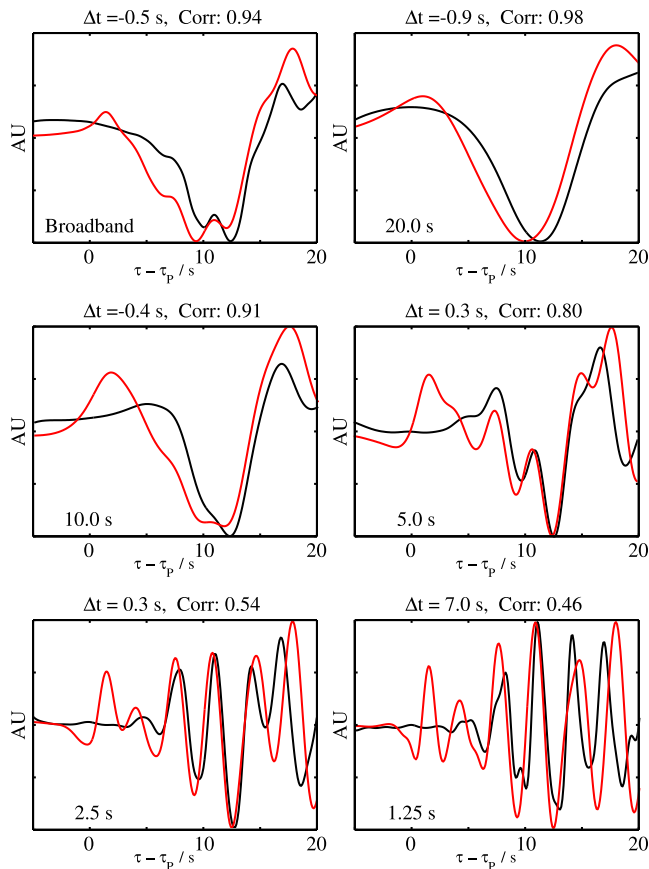


Fig. 8. The fit between a triplicated P-wave and its synthetic, filtered into frequency pass bands. The epicentral distance is 18.6° , so that the waveform contains triplications from both the 410 and 660 discontinuities. This example uses the same broadband waveforms as seen in Fig. 3; the seismic station is VALT (Mount St. Helens crater rim).

4.1 North America

Since the advent of *USArray*, North America is clearly the best-instrumented large landmass on Earth. *USArray* stations are spaced by ≈ 70 km on a regular grid and deliver superb broadband waveforms.

The data from the first year of instalment already brought new insights into the subduction history of the Farallon plate and the formation of the Rocky mountains (Sigloch et al., 2008) or to fluid transport in the Gorda subduction system (Cao and Levander, 2010). The recent move of the deployment into the Great Plains has brought the seismicities of the Guerrero subduction and along the western Canadian margin into distances of 15° to 30° , creating large numbers of crossing paths in triplication range. *USArray* recordings are supplemented by permanent networks in the US, Canada, and Mexico, and by data from temporary experiments. All North American data was downloaded from the *IRIS* data management center (DMC). Currently *IRIS* delivers about

2000 global stations per event, of which ≈ 1400 are located in larger North America, i.e. between 180° W and 45° W in the Northern Hemisphere.

We found that between 01/01/2001 and 31/12/2011, 92 regional earthquakes generated triplicated P-wave recordings of acceptable signal-to-noise ratio. This yielded 26016 unique, acceptable wave paths and broadband waveforms in total. We applied Gabor bandpass filters (bandwidth one octave) at centre periods of 20 s, 10 s, 5 s, 2.5 s, 1.25 s. Since we make cross-correlation measurements, the correlation coefficient between observed (filtered) waveforms and their synthetics using *IASP91* serves as the primary measure of goodness of fit. We denote by δT the time shift that maximizes the cross-correlation between observed seismogram and matched filter as in Eq. (4), and by CC_{\max} the correlation coefficient at this optimum time shift. Hence CC_{\max} acts as a quality measure, and δT is the actual finite-frequency observable to be interpreted by tomography. Previous experience (Sigloch and Nolet, 2006; Sigloch, 2008) lets us assume that $CC_{\lim} \geq 0.8$ is a good threshold for acceptance. If we only accept filtered waveforms with $CC_{\max} \geq 0.8$, 42 949 out of 130 080 remain (Table 1). Fitting waveforms in the five separate frequency bands thus increases the number of usable data more than fourfold, compared to fitting only the broadband seismograms. In particular, we can often accept at least one passband measurement even when the CC_{\max} of the broadband measurement is clearly too low – this salvages many important wave paths. For teleseismic waves, another frequency band below 20 s might reasonably be added, but for regional P-waves this is less interesting, since the 20 s kernels already fill up the entire depth range of the transition zone (see Fig. 5).

Figure 8 shows a comparison between matched filters and real waveforms for the seismograph station at Mount St. Helens crater, which is by no means an ideal station. Still three out of five bands have $CC_{\max} > 0.8$ and could contribute to tomography. These seismograms were calculated using *IASP91* velocity model. For an actual tomography of Northern America it may be beneficial to calculate matched filters and kernels in a regional 1-D model; for a recent model and a comparison between previous results see Chu et al. (2012).

Figure 9 shows that the average absolute traveltime anomaly δT is similar in all bands, with a standard deviation of ≈ 2.4 s. From our experience with teleseismic P-waves, we know that a large part of this δT signal is due to mislocated sources rather than to mantle heterogeneity; both contributions are jointly inverted for by our tomography software.

These δT measurements on triplicated data will be embedded into a larger data set of teleseismic P-wave data (Sigloch, 2011). The latter have better azimuthal coverage, since they include earthquakes from the east (mid-Atlantic ridge and southern Europe). Unfortunately, very few seismic sources east of North America are located at regional distances.

Table 1. Number of acceptable traveltime measurements in the various frequency bands, as a function of the chosen rejection threshold CC_{lim} . In total 26016 (NA) and 4916 (Europe) broadband seismograms were available.

Northern American data set						
CC_{lim}	0.70	0.75	0.80	0.85	0.90	0.95
20 s	19 897	18 563	16 717	14 225	10 553	5018
10 s	17 646	15 440	12 995	10 124	6317	2143
5 s	12 691	10 331	7900	5452	2861	729
2.5 s	8180	5852	3856	2235	1056	220
1.25 s	3561	2368	1481	773	258	13
total	61 975	52 554	42 949	32 809	21 045	8123
broadband	14 552	12 177	9949	7141	3715	1080

European data set						
CC_{lim}	0.70	0.75	0.80	0.85	0.90	0.95
20 s	3450	3161	2840	2423	1891	775
10 s	2953	2477	1922	1261	577	84
5 s	1556	1037	598	264	83	5
2.5 s	882	483	233	91	19	0
1.25 s	277	164	72	29	5	0
total	11 627	9311	7042	4840	2876	886
broadband	2509	1989	1377	772	301	22

4.2 Europe

Compared to North America, the European seismic networks are less homogeneous, even though the *NERIES* initiative brought a great advance. We downloaded our waveforms from the *ORFEUS* DMC. This archive currently contains around 400 seismic stations, mainly in Europe, but also several temporary installations by European agencies, e.g. in Indonesia. Figure 7 shows that data coverage is quite uneven. Several large countries like France and Poland are sparsely instrumented, other networks do not share their data yet. There have been several large temporary installations in Central Europe and Scandinavia, data which would be very interesting but are being released only slowly. The *IRIS* DMC holds hardly any broadband stations above and beyond the *ORFEUS* DMC. On the other hand, the geometry and seismicity of Europe is more favourable than for North America. Hellenic seismicity is recorded over a large azimuthal range from Spain to Russia and produces many crossing paths, together with Western Mediterranean seismicity recorded in Central Europe, and seismicity on the North Atlantic ridge between Iceland and Svalbard. One of the most interesting features in the transition zone are the slab remnants below the Northern Balkans and the Carpathians, which cannot be easily connected to the established subduction regimes. Sampling of the MTZ by triplicated waves is very good in this region, and should help in further illuminating its mantle processes. A clear difference to the North American data

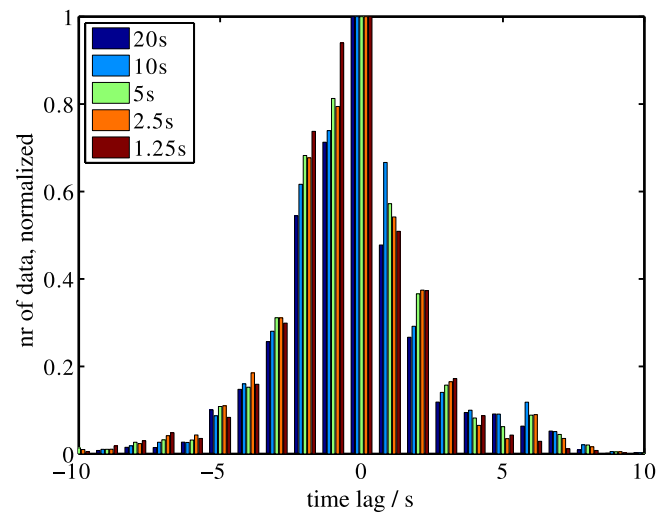


Fig. 9. Measured traveltime anomalies δT for all paths in the North American data set, in all five frequency bands (distinguished by colour). Only data whose fit exceeds $CC_{max} > 0.8$ are plotted. The histograms are normalized so that at $\delta T = 0$, each bar has value 1.

is that few earthquakes above magnitude 6 are recorded in Europe. Since we obtain best correlations between data and synthetics at around magnitude 6.5, the data quality is generally lower than in North America. In total, the number of available stations in larger Europe (defined as the region between 45° W and 75° E in the northern hemisphere) is 390. Our dataset contains 66 earthquakes, which recorded 4916 broadband seismograms (unique wave paths). 7042 out of 24 580 passband measurements have a data-synthetic fit that exceeds $CC_{max} > 0.8$. Compared to the North America, we obtain about five times fewer waveforms, and seven times fewer acceptable δT measurements. This lower average signal quality seems to be mainly due to the weaker seismicity around Europe.

4.3 Information content of triplicated P-waveforms

Before embarking on tomographic inversions, we want to convince ourselves that triplicated waveforms do indeed contain coherent and usable structural information. The broad footprint of *USArray* may be sliced up into dense seismic profiles at various back-azimuths, each featuring dozens of stations. Here we consider one such quasi linear section, generated by an earthquake from Guerrero, Mexico. Figure 10 shows the P-waveforms of 84 displacement seismograms in a range between 22 and 35° distance. The traces are color coded, where green means zero displacement, blue is negative displacement, and red is positive. Traces are time-aligned on the *IASP91*-predicted arrivals of the first P-phase.

The first arrival in the real seismograms occurs around $+2$ s for all traces, a systematic bias w.r.t. to *IASP91* that is most likely due to source mislocation. The blue triangular move-outs (e.g. to $+7$ s at 21° distance) are the triplicated phases of

P. Starting at +10 s, the whole P-arrival sequence is “echoed” (also in blue) – this is the depth phase pP, in itself triplicated, for this shallow event (12 km deep).

If the Earth were truly spherically symmetric, the P-arrivals would be aligned smoothly – if not along a vertical line of $t = 0$, then along some other smooth line of steep move-out. This is clearly not the case, signalling the kind of lateral mantle heterogeneity that seismic tomography targets. Strong unevenness is for example observed between 25 and 31° distance, where the P-onset varies between 0 s and +3 s.

The largest delays are present between 29 and 31° distance. This might be explained by a low-velocity anomaly directly below the 660, since only waves of this distance range bottom there (see Fig. 6). Recordings at more distant stations would be less influenced by this hypothetical anomaly, since they traverse it at a steeper angle, with correspondingly lower sensitivity. However, a second observation leads to another explanation: The second triplication (here arriving at ≈ 5 s at 28°), represents the phase bottoming directly above the 660. According to *IASP91*, it should be recorded only to a distance of 27.8°, but here we observe it to distances exceeding 29°. This could be explained by a depression of the 660-km discontinuity to a depth of 680 km, which would also explain the delay of the first phase at this distance. A depressed 660 in this region of subduction is plausible, since the lower temperatures of the slab would shift the phase transition to higher pressures and deeper depths. Indeed, studies from SS precursors Houser et al. (2008) showed a 660 depressed to 668–671 km in Northern Mexico.

The second subplot of Fig. 10 shows several stations at 28 to 29° distance, where the second triplication should not be recorded. We do observe the phase, but only at some of the stations (at +4 to +5 s), which suggests that the depression of the 660 is quite localized. From the azimuthal spread of the stations, we can estimate their bottoming points to be spaced only by tens of kilometres, and yet the imprint of the 660 in the seismograms is quite variable. Since the maximum frequencies in these waveforms are less than 1 Hz, the scale of the undulations on the 660 must be close to the resolution limit of these waves. The aforementioned precursor study lacked the resolution to resolve such small features.

5 Discussion

Using triplicated P-waves fills an important gap in waveform tomography. Teleseismic waves have relatively poor depth resolution in the MTZ, and little sensitivity to the discontinuities themselves. Since regional waves have this resolution, we hope to greatly increase the resolution of future P-wave models in the transition zone. So far, triplicated body waves have hardly ever been used for tomography, since they are not well modelled by classical ray theory. The appearance of the actually measured waveform has a clear finite-frequency character rather than resembling a sequence of Dirac pulses.

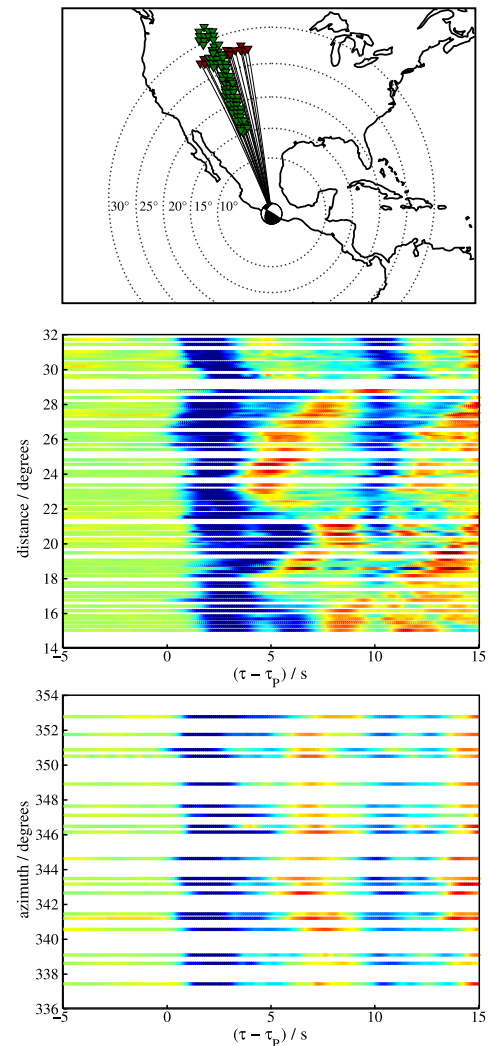


Fig. 10. Evidence of 3-D structural information contained in triplicated P-waves. Top: we extract a quasi 1-D profile of 84 *USArray* stations, recording an earthquake in Mexico (Event: 2009/04/27 16:46:28, Mag: 5.8, Lat: 17.03, Lon: -99.45, Depth: 35 km). Middle: Section of broadband seismograms at triplication distances (between 22 and 35°, green in the map plot). Z-component, colour coded, green is zero displacement, blue negative, red positive. The traces are normalized separately, and aligned to the first arriving P-wave as predicted by *IASP91*. See text for discussion. Bottom: Seismograms from the ensemble of stations at quasi-constant distance of 28 to 29° (red triangles in map plot). The second P-triplication is observed on some of the traces (at $t = 4$ –5 s) but not on others, pointing to variations on the 660-km discontinuity.

Our choice of misfit criterion is the cross-correlation traveltimes – effectively a phase shift measurement, given our relatively narrow pass-bands. Cross-correlation is the optimal strategy for the detection and estimation of a known signal (the synthetic waveform or “matched filter”) in a noisy version of the same signal (the measured waveform), where the noise is assumed to be white additive Gaussian noise in each

passband. Intuitively it might seem that cross-correlation may not be a suitable misfit for triplicated waves, since for all but the highest frequencies, the time window will necessarily encompass all three triplicated phases, each of which has different spatial sensitivity. Hence calculating one traveltimes delay on all three overlapping phases may seem unphysical. We believe that this understanding of the cross-correlation misfit is misguided by (ray-theoretical) intuition. While it is true that one cross-correlation applied to a broadband signal only calculates one delay time for the whole broadband time series, this is actually not what we propose to do. When splitting the signal into multiple frequency bands k and calculating separate δT_k , several, ideally independent measurements are done on the waveform. The corresponding sensitivity kernels describe the sensitivity of the model towards each of these measurements. It does not describe the sensitivity towards the traveltimes of one particular (ray-theoretical) phase. The kernel computation formalism ensures that the proper sensitivity for each window and filter is computed, whatever they may be, meaning a time window containing triplicated waves can be used just as any other time window. One should just not expect a ray-theoretically intuitive meaning corresponding to each measurement.

The calculation of full wavefields in a laterally heterogeneous 3-D Earth model is still prohibitively expensive for frequencies above 0.05 Hz. Hence waveform tomography must currently choose between one of two compromises:

1. In a laterally heterogeneous Earth model, calculate low frequency wavefields (< 0.05 Hz), since high frequencies are not affordable. This means focusing on surface waves and the low-frequency part of body waves. Due to the large wavelengths of body waves, their low-frequency part offers little resolution in the mantle transition zone (already at a dominant period of 20 s, the kernel fills the entire transition zone). The high cost of wavefield calculations mandates that they need not be done too often. The adjoint method proposed by Taran-tola (1984), and applied in continental-scale seismology by Tromp et al. (2005), Fichtner et al. (2009) and Zhu et al. (2012), offers an efficient solution by calculating wavefields only once per source, resulting in composite, so-called event kernels, which indicate the descent direction for the (linearised) gradient search. Several iterations are possible and customary, since wavefields can be computed in arbitrary Earth models, in particular also the updated ones.
2. In a spherically symmetric Earth model, calculate kernels that span any or all parts of the seismically relevant frequency spectrum. This broadband capacity has defined “finite-frequency tomography” since Dahlen et al. (2000). Nissen-Meyer et al. (2007b) showed that full wavefield kernel computations are feasible for frequencies up to 1 Hz, due to the extreme computational savings that result from smart exploitation of the spherical

symmetry. The use of high-frequency waves promises accordingly higher image resolution. The method is so efficient that all source-to-receiver kernels are explicitly calculated, turning the problem into one of matrix inversion, for which powerful analysis concepts and computational tools are well known. The disadvantage is that several iterations are not possible, since the Earth model would lose its spherical symmetry after the first update. This might be a serious drawback in media with heterogeneity of strong magnitude, where the single-scattering (Born) approximation starts breaking down, but where several linearised iterations might still lead significantly closer to the global misfit minimum than a single one.

We believe that the finite-frequency approach is currently better suited to tomography of the upper mantle and the transition zone, since (a) being able to use the highly resolving body waves across their entire frequency range should be a big advantage, and (b) all previous studies let us expect only relatively weak perturbations from a layered background model like *IASP91*, on the order of a few percent, so that the single-scattering approximation should be completely adequate.

A global reference model may not be best suited for a regional study. Rather, careful selection of a suitable regional reference model is a crucial step of a tomography.

We note that the mixed use of forward modelling codes (*WKB*, reflectivity method) has historical reasons and reflects an unfinished (though functional) stage of development. Ultimately all steps will be carried out with the most complete method, *Axisem*. For inversion of source time functions from teleseismic data, *WKB* has been an efficient tool since (Sigloch and Nolet, 2006). It is completely adequate for teleseismic waves, and this STF inversion step is treated as independent from the modelling of triplicated waveforms. Reflectivity is used for efficient forward modelling of the triplicated broadband seismograms (*WKB* cannot easily compute triplicated phases, and *Axisem* is not yet set up to do true broadband computations efficiently for the number of computations needed). In order to compute sensitivity kernels, the full wavefield is needed, and we obtain it from full numerical forward modelling using *Axisem* (but not yet routinely up to frequencies of 1 Hz).

Analysis of triplicated waveforms has so far been mostly applied to deep events (Tajima and Grand, 1995; Tajima et al., 2009), in order to separate the influence of the discontinuities from the depth phases. We have shown that shallow events can be modelled when source parameters are carefully estimated. Routine catalogue estimates (Global CMT (Dziewoński et al., 1981) or NEIC) do not deliver all the parameter we need (source time function), or not to the required accuracy (e.g., source depth).

However, using our own source inversion results from teleseismic P-waves, we can model the sources sufficiently well for our purposes. The ability to use shallow earthquakes

enlarges the data base enormously, since most earthquakes occur shallower than 30 km depth. In particular, good tomographic resolution requires good azimuthal coverage by sources, but few regions on Earth are surrounded by subduction zones to generate deep earthquakes from all directions.

The regional data sets are promising, both for North America and for Europe. They will be seamlessly embedded into a global inversion that also contains teleseismic P-wave measurements Sigloch (2011). A regional tomography of Europe can thus still benefit from events in eastern Asia, i.e. beyond the triplicated range. Hence the fact that the regional seismicity around Europe is weaker than in North America, is compensated by the superior azimuthal coverage at teleseismic distances. Nevertheless, the European network is nowhere as dense as the *USArray*, which will make any American mantle structure far better resolved in the foreseeable future. Average waveform cross-correlation is poorer in the European data set than in the North American one. Three possible explanations come into mind:

1. Station quality might be lower. We do not believe this is generally the case, even though the average distance to the nearest coast is smaller in Europe than in North America.
2. European mantle structure could be more complicated than under Northern America, generating a larger mismatch between observed and modelled waveforms. The tectonic history of Europe is often thought to be more complex, but tomographic studies since the advent of *USArray* have also revealed very heterogeneous mantle structure under North America (Pavlis et al., 2012; Becker, 2012). Hence it is unlikely that the true seismic velocity structure in the mantle under Europe deviates significantly more from a layered model than under North America.
3. Earthquake sources around Europe may be less suited. We think this is the main issue. In North America, we can use rather strong strike-slip events along the west coast, and numerous deep events along the Guerrero subduction zone, many of them exceeding magnitude 6. European seismicity is weaker in magnitude and tends to consist of complicated events in the Aegean subduction and along the Anatolian fault; strike-slip events along the Atlantic ridge are rather weak. Since some of these events are not contained in the *IRIS WILBER* archive and thus only the *ORFEUS* stations were available to us, our source inversion for them might be more error prone as well.

The seismic section from the Mexican earthquake demonstrates the high signal quality of *USArray* recordings (Fig. 10). The discussed strong imprint of a depressed 660 in this signal shows that we may need to be careful to properly parametrize the inversion such that depressed or elevated

discontinuities are detected and become part of the tomography result, rather than smearing into bulk velocity structure. In order to separate the two effects, we are considering the use of boundary topography kernels (Colombi et al., 2012), in addition to the volumetric velocity kernels shown in Fig. 5.

6 Conclusions

We conclude that regional body waves should be usable and useful for waveform tomography. We plan to invert these data using finite-frequency tomography and kernels from full numerical wavefield computations (Nissen-Meyer et al., 2007b), but our conclusions about the nature of the data have broader validity.

After careful deconvolution of the source time function and other source parameters from teleseismic waveforms, we obtain good cross-correlations between observed and modelled triplicated seismograms, across the broadband range and even for shallow sources. Due to the much more complex nature of the waveforms, these fits are lower than what we obtain for teleseismic P-waves, but are still sufficient to assemble two decently sized tomography data sets for North America and Europe. The inclusion of these data greatly increase and complement the sensitivity to transition zone structure, and in particular to the discontinuities at 410 km and 660 km, which so far must be investigated using waves of far lower signal-to-noise ratio. The abundance and high quality of data from *USArray* make the transition zone under North America the natural target for a first waveform inversion using triplicated P-waves.

Acknowledgements. We thank F. Tajima and B. L. N. Kennett for sharing their experience on triplicated waves. G. Nolet, M. v. Driel and an anonymous reviewer helped improving the manuscript in the Solid Earth Discussion stage. All waveform data came from the *IRIS* and *ORFEUS* data management centres. S. C. S. was funded by the Munich Center for Advanced Computing (MAC/IGSSE).

Edited by: R. Carbonell

References

- Becker, T.: On recent seismic tomography for the western United States, *Geochem. Geophys. Geosyst.*, 13, 1–11, doi:10.1029/2011GC003977, 2012.
- Cao, A. and Levander, A.: High-resolution transition zone structures of the Gorda Slab beneath the western United States: Implication for deep water subduction, *J. Geophys. Res.*, 115, B07301, doi:10.1029/2009JB006876, 2010.
- Chaljub, E. and Tarantola, A.: Sensitivity of SS precursors to topography on the upper-mantle 660-km discontinuity, *Geophys. Res. Lett.*, 24, 2613–2616, doi:10.1029/97GL52693, 1997.
- Chapman, C.: A new method for computing synthetic seismograms, *Geophys. J. Roy. Astr. S.*, 54, 481–518, 1978.

- Chu, R., Schmandt, B., and Helmberger, D. V.: Upper mantle P velocity structure beneath the Midwestern United States derived from triplicated waveforms, *Geochem. Geophys. Geosyst.*, 13, 1–21, doi:10.1029/2011GC003818, 2012.
- Colombi, A., Nissen-Meyer, T., Boschi, L., and Guardini, D.: Seismic sensitivity to global boundary topography, *Geophys. J. Int.*, 191(2), 832–848, doi:10.1111/j.1365-246X.2012.05660.x, 2012.
- Dahlen, F. A., Hung, S. H., and Nolet, G.: Fréchet kernels for finite-frequency traveltimes – I. Theory, *Geophys. J. Int.*, 141, 157–174, doi:10.1046/j.1365-246X.2000.00070.x, 2000.
- Deuss, A.: Global observations of mantle discontinuities using SS and PP precursors, *Surveys in geophysics*, 30, 301–326, doi:10.1007/s10712-009-9078-y, 2009.
- Deuss, A., Redfern, S. A. T., Chambers, K., and Woodhouse, J. H.: The Nature of the 660-Kilometer Discontinuity in Earth's Mantle from Global Seismic Observations of PP Precursors, *Science*, 311, 198–201, doi:10.1126/science.1120020, 2006.
- Dziewoński, A. M.: Preliminary reference Earth model, *Phys. Earth Planet. In.*, 25, 297–356, doi:10.1016/0031-9201(81)90046-7, 1981.
- Dziewoński, A. M., Chou, T.-A., and Woodhouse, J. H.: Determination of Earthquake Source Parameters From Waveform Data for Studies of Global and Regional Seismicity, *J. Geophys. Res.*, 86, 2825–2852, doi:10.1029/JB086iB04p02825, 1981.
- Fichtner, A., Kennett, B. L. N., Igel, H., and Bunge, H.-P.: Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods, *Geophys. J. Int.*, 179, 1703–1725, doi:10.1111/j.1365-246X.2009.04368.x, 2009.
- Fuchs, K. and Müller, G.: Computation of Synthetic Seismograms with the Reflectivity Method and Comparison with Observations, *Geophys. J. Roy. Astr. S.*, 23, 417–433, 1971.
- Grand, S. P.: Upper mantle shear structure of North America, *Geophys. J. Roy. Astr. S.*, 76, 399–438, doi:10.1111/j.1365-246X.1984.tb05053.x, 1984.
- Grand, S. P.: Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.*, 99, 11591–11621, 1994.
- Grand, S. P.: Mantle shearwave tomography and the fate of subducted slabs, *Philos. T. R. Soc. A*, 360, 2475, 2002.
- Houser, C., Masters, G., Flanagan, M. and Shearer, P. M.: Determination and analysis of long-wavelength transition zone structure using SS precursors, *Geophys. J. Int.*, 174, 178–194, doi:10.1111/j.1365-246X.2008.03719.x, 2008.
- Kennett, B. L. N.: *The Seismic Wavefield: Introduction and theoretical development*, Cambridge University Press, 2001.
- Kennett, B. L. N. and Engdahl, E. R.: Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.*, 105, 429–465, doi:10.1111/j.1365-246X.1991.tb06724.x, 1991.
- Komatitsch, D. and Tromp, J.: Spectral-element simulations of global seismic wave propagation I. Validation, *Geophys. J. Int.*, 149, 390–412, 2002.
- Marquering, H., Dahlen, F. A. and Nolet, G.: Three-dimensional sensitivity kernels for finite-frequency traveltimes the banana-doughnut paradox, *Geophys. J. Int.*, 137, 805–815, doi:10.1046/j.1365-246x.1999.00837.x, 1999.
- Melbourne, T. I. and Helmberger, D. V.: Whole mantle shear structure beneath the East Pacific Rise, *J. Geophys. Res.*, 107, 2204, 2002.
- Montelli, R., Nolet, G., Dahlen, F. A., and Masters, G.: Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, 303, 338–43, doi:10.1126/science.1092485, 2004.
- Nissen-Meyer, T., Dahlen, F. A., and Fournier, A.: Spherical-earth Fréchet sensitivity kernels, *Geophys. J. Int.*, 168, 1051–1066, doi:10.1111/j.1365-246X.2006.03123.x, 2007a.
- Nissen-Meyer, T., Fournier, A., and Dahlen, F. A.: A two-dimensional spectral-element method for computing spherical-earth seismograms–I. Moment-tensor source, *Geophys. J. Int.*, 168, 1067–1092, doi:10.1111/j.1365-246X.2006.03121.x, 2007b.
- Nissen-Meyer, T., Fournier, A., and Dahlen, F. A.: A 2-D spectral-element method for computing spherical-earth seismograms–II. Waves in solid-fluid media, *Geophys. J. Int.*, 174, 873–888, doi:10.1111/j.1365-246X.2008.03813.x, 2008.
- Nolet, G.: *A Breviary of Seismic Tomography: Imaging the Interior of the Earth and Sun*, Cambridge University Press, 2008.
- Ohtani, E.: Water in the mantle, *Elements*, 1, 25, doi:10.2113/gselements.1.1.25, 2005.
- Pavlis, G. L., Sigloch, K., Burdick, S., Fouch, M. J., and Vernon, F.: Unraveling the Geometry of the Farallon Plate: Synthesis of Three-dimensional Imaging Results from USArray, *Tectonophysics*, 532–535, 82–102, doi:10.1016/j.tecto.2012.02.008, 2012.
- Shearer, P. M. and Masters, G.: Global mapping of topography on the 660km discontinuity, *Nature*, 355, 791–796, 1992.
- Sigloch, K.: Multiple-frequency body-wave tomography, Ph.D. thesis, Princeton, 2008.
- Sigloch, K.: Mantle provinces under North America from multifrequency P wave tomography, *Geochem. Geophys. Geosyst.*, 12, 27 pp., doi:10.1029/2010GC003421, 2011.
- Sigloch, K. and Nolet, G.: Measuring finite-frequency body-wave amplitudes and traveltimes, *Geophys. J. Int.*, 167, 271–287, doi:10.1111/j.1365-246X.2006.03116.x, 2006.
- Sigloch, K., McQuarrie, N., and Nolet, G.: Two-stage subduction history under North America inferred from multiple-frequency tomography, *Nat. Geosci.*, 1, 458–462, doi:10.1038/ngeo231, 2008.
- Tajima, F. and Grand, S.: Evidence of high velocity anomalies in the transition zone associated with southern Kurile subduction zone, *Geophys. Res. Lett.*, 22, 3139–3142, doi:10.1029/95GL03314, 1995.
- Tajima, F., Katayama, I., and Nakagawa, T.: Variable seismic structure near the 660 km discontinuity associated with stagnant slabs and geochemical implications, *Phys. Earth Planet. Int.*, 172, 183–198, doi:10.1016/j.pepi.2008.09.013, 2009.
- Tape, C., Liu, Q., Maggi, A., and Tromp, J.: Adjoint tomography of the southern California crust., *Science*, 325, 988–992, doi:10.1126/science.1175298, 2009.
- Tarantola, A.: Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, 49, 1259–1266, doi:10.1190/1.1441754, 1984.
- Thomas, C. and Billen, M. I.: Mantle transition zone structure along a profile in the SW Pacific: thermal and compositional variations, *Geophys. J. Int.*, 176, 113–125, 2009.
- Tian, Y., Zhou, Y., Sigloch, K., Nolet, G., and Laske, G.: Structure of North American mantle constrained by simultaneous inversion of multiple-frequency SH, SS, and Love waves, *J. Geophys. Res.*, 116, doi:10.1029/2010JB007704, 2011.
- Tromp, J., Tape, C., and Liu, Q.: Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*,

- 160, 195–216, doi:10.1111/j.1365-246X.2004.02453.x, 2005.
- VanDecar, J. C. and Crosson, R. S.: Determination of teleseismic relative phase arrival times using multi-channel cross-correlation and least squares, *Bull. Seismolog. Soc. Am.*, 80, 150–169, 1990.
- Zhou, Y., Dahlen, F. A., and Nolet, G.: Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, 158, 142–168, doi:10.1111/j.1365-246X.2004.02324.x, 2004.
- Zhou, Y., Nolet, G., Dahlen, F. A., and Laske, G.: Global upper-mantle structure from finite-frequency surface-wave tomography, *J. Geophys. Res.*, 111, B04304, doi:10.1029/2005JB003677, 2006.
- Zhu, H., Boždag, E., Peter, D., and Tromp, J.: Structure of the European upper mantle revealed by adjoint tomography, *Nat. Geosci.*, 5, 493–498, doi:10.1038/ngeo1501, 2012.