Doctorate Dissertation (Censored)

博士論文　（要約）

Imaging slabs in the Earth's mantle transition zone and D″ layer beneath Central America and its vicinity using waveform inversion

（波形インバージョンによる中央アメリカおよびその近傍下のマントル遷移層と最下部マントルにおける沈み込んだスラブのイメージング）

A Dissertation Submitted for Degree of Doctor of Philosophy

December 2018

平成 30 年 12 月博士（理学）申請

Department of Earth and Planetary Science, Graduate School of Science,

The University of Tokyo

東京大学大学院理学系研究科地球惑星科学専攻

BORGEAUD DIT AVOCAT Anselme Francois Emile

ボルジョ　ディ　アボカ　アンセルム　フランソア　エミル

Abstract

The mantle transition zone (MTZ; between ~410–660 km depth) and the D″ region (from the core-mantle boundary to ~400 km above it), the two boundary layers in the Earth's mantle, play a major role in governing the modality of convection in the mantle. Previous work using travel-time tomography reported large variability of convection modality of slabs in and near the MTZ. Also, the length scale of velocity anomalies in the D″ region, and thus of convection if a thermal origin is assumed, is still being debated. In order to improve our understanding of how slabs interact with the 660 km discontinuity and the thermal boundary layer in the D″ region, I use waveform inversion to image the detailed 3-D S-velocity structure in the D″ region and MTZ beneath Central America and vicinity. The region beneath Central America is of particular geodynamical interest, since the paleo- and present Pacific plates have been subducting beneath the western margin of Pangaea since ~250 Ma, which implies that paleoslabs could have reached the lowermost mantle and are still present in the MTZ. I obtain high-resolution images because of the dense sampling by seismic waves due to the full deployment of the USArray broadband seismic stations in the conterminous US during 2004–2015. In the D″ region, I find evidence for two distinct paleoslabs possibly corresponding to the Farallon slab, and remnants from intra-oceanic subduction, and for chemically distinct denser material just above the core-mantle boundary. In the transition zone, I find complex subduction modality of the Cocos slab, possibly due to the thermal structure of the Cocos slab, and its interaction with a lower-mantle plume, suggesting the importance of properties of the slab itself on the convection modality in and near the MTZ, and of the current state of convection in the lower-mantle. This work favors whole mantle convection (at least in the study region) with slabs sinking to the CMB possibly creating iron-rich heterogeneities by chemical

differentiation, and triggering upwelling flow of hot material that could in turn interact with younger slabs in the MTZ.

Table of Contents

Acknowledgements

I am deeply grateful to my PhD supervisor Kenji Kawai for his teaching, guidance, and support during this PhD, and to Robert J. Geller, who supervised me during my Master degree and continued to do so during my PhD. They both taught me what it is to be a researcher. I would also like to thank my two seniors Kensuke Konishi, and Kei Hasegawa, for their support and guidance during my Master degree and part of this PhD. I thank my co-authors Kawai Kenji, Robert J. Geller, Kensuke Konishi, Yuki Suzuki, and Lina Yamaya. I thank the members of Kenji Kawai's lab: Yuki Suzuki, Lina Yamaya, Tsubasa Ue, and Hanaya Okuda for their comments and advices during the lab meetings. I thank the older members of Robert J. Geller's lab, Nozomu Takeuchi, and Tatsuhiko Hara for helpful discussions. I thank the professors of the Earthquake Research Institute, especially Satoru Honda, Takehiko Hiraga, Hitoshi Kawakatsu, Nozomu Takeuchi, and Takashi Iidaka, for their teachings and comments during my Master and PhD degree. I thank the professors of the Solid Earth group, especially Ataru Sakuraba, Sato Masahiko, Keisuke Nishida, Tsuyoshi Iizuka, Kei Hirose, Simon Wallis, Ryosuke Ando, and Satoshi Ide for their helpful remarks, encouragements, and advices. I thank my classmates Yuki Hibiya, and Keita Itano, and the members of room 712, Julie Maury, Pierre Romanet, Miki Aso, Lisa Kaneko, Izumi Mashino, Kei Hasegawa, Masahiko Sato, Naoto Mizuno, and Junki Komori for the good times I spent with them during these five years at the University of Tokyo. I would also like to thank professors outside of the University of Tokyo, especially Nobumasa Funamori, Taku Tsuchiya, Jun Tsuchiya, Maxim Ballmer, Takashi Nakagawa, Satoshi Kaneshima, Toshiki Ohtaki, Christine Houser, and John Hernlund for their teachings, advices, and encouragements. I also wish to thank Yukako Baba, Shio Watanabe, and Nami Tonegawa for their great administrative support during my PhD. Finally, I would

like to thank all the members of Tokyo Kung Fu Dojo, with whom I had the chance to share three years of strict but rewarding training under the supervision of Sifu Sato.

I probably chose the path to being a researcher thanks to my Physics high-school teacher Jean-Charles Delacrétaz, to which I would like to express my profound gratitude. Preparations for this “adventure” to Japan started at the EPFL in 2013 with my friend Michele Oliosi, who is currently also suffering to finish his PhD in Kyoto University! I am deeply grateful to him. This PhD would have been difficult to complete without the boundless support of my parents, Emmanuelle and Jean-Jacques, and my little brother Tanguy, from Switzerland and during their many visits to Tokyo, and from the rest of my family. Last but not least, I am deeply grateful to my girlfriend Tzu-Yu for her continuous support and encouragements during these three years of PhD.

Chapter 1 General introduction

Seismic tomography, named by analogy with similar imaging techniques in medicine, is the imaging of the elastic (i.e., propagation velocity of seismic waves), and anelastic (i.e., attenuation) structure of the Earth's interior. Seismic tomography uses the information accumulated by seismic waves as they interact with the Earth's structure along propagation paths from earthquakes' sources to seismic stations (receivers) at the Earth's surface. Seismic tomography is an inverse problem, where the parameters to be inferred (unknowns) are the velocity (or anelastic) structure in the Earth, and the quantity to be minimized is the difference (called residuals) between seismic observables (e.g., travel-times, amplitudes, velocity dispersion), and synthetic seismic data (called "synthetics") calculated for a given velocity (and anelastic) model of the Earth.

Three-dimensional (3-D) seismic tomography, i.e., the imaging of lateral heterogeneities in the Earth, was introduced by the pioneering work of Aki et al. (1977), and Dziewonski et al. (1977), who used recorded travel-times of direct P-waves, and the laws of geometrical optics (ray theory) to constrain the 3-D structure. Since then, improvements in 1) the quantity of available seismic data (e.g., deployment of new seismic stations, and "arrays" of seismic stations), 2) theoretical framework for the formulation of the inverse problem (e.g., "kernels" computed using full-wave theory, without using geometrical optics approximations), 3) methods for the calculation of synthetic data (e.g., 3-D wave propagation using the full-wave theory), and 4) computational power have led to the imaging of increasingly smaller scale features of the Earth's interior. Below, I give an overview of the progress of seismic tomography since the above pioneering work, with particular focus on the main question addressed in this dissertation, which is: what can seismic data tell us about the state of

convection of the Earth's mantle? In particular, I will focus on so called "high-resolution" tomographic models, which are typically 1) obtained either by travel-time tomography, or by regional waveform inversion, and 2) parametrized in terms of blocks (voxels) or local functions (in contrast to a global parametrization in terms of spherical harmonics, whose models have typically lower resolution of the order of ~500–1000 km and whose primary focus is not on the detailed structure of slabs). We note that other detailed reviews can be found in previous works for regional tomography (Thurber & Aki, 1987), global tomography (Romanowicz, 1991, 2003), and complementary methods to tomography, such as array analysis (Lay & Garnero, 2011).

1.1 Early works: elucidating the convection style of the mantle

Aki et al. (1977), and Dziewonski et al. (1977) first applied inverse theory to infer the regional 3-D structure of the lithosphere beneath an array of seismic stations, and the global structure of the Earth's mantle, respectively, using travel-time residuals of teleseismic direct P-waves. Below, I describe in some detail the work of Aki et al. (1977), since the inverse problem theory used in their work is nearly identical to that used in following works.

Aki et al. (1977) parametrized the lithosphere beneath the stations in 3-D blocks (voxels), in which the slowness (inverse of velocity) is allowed to vary from a reference 1-D model (obtained for this region in previous studies) and used ray theory to relate travel-time residuals $\delta \boldsymbol{d}$ (the data) to fractional slowness perturbations $\delta \boldsymbol{m}$ (the unknown model parameters), as in eq. (1)

$$\delta \boldsymbol{d} = \boldsymbol{A} \delta \boldsymbol{m}, \tag{1}$$

where

$$A_{ij} = \frac{\bar{d}_{ij}}{\bar{v}_j}, 1 \le i \le N, 1 \le j \le M \quad (2)$$

is called the kernel (or Fréchet kernel, or partial derivative matrix), $\bar{d}_{ij}$ is the distance traveled by ray i in block j with the ray computed in the reference model, and $\bar{v}_j$ is the average velocity in the depth layer of block j. Since the number of data (N) is generally larger than the number of model parameters (M), eq. (1) is over-determined, and the aim of the tomographic inversion is to find the model perturbation $\delta\boldsymbol{m}$ that minimizes the quantity $\|\boldsymbol{A}\delta\boldsymbol{m} - \delta\boldsymbol{d}\|$, where $\|\cdot\|$ is a norm, which is usually the L_2 norm. This implies that $\delta\boldsymbol{m}$ is the solution of the least-square eq. (3)

$$\boldsymbol{A}^T\boldsymbol{A}\delta\boldsymbol{m} = \boldsymbol{A}^T\delta\boldsymbol{d}, \quad (3)$$

Because of the poor ray coverage of some of the model blocks, $\boldsymbol{A}^T\boldsymbol{A}$ is in general singular. In this case, eq. (3) is solved using the generalized inverse $(\boldsymbol{A}^T\boldsymbol{A})^{-g}$ given by the SVD decomposition (Lanczos, 1961), as in eq. (4)

$$(\boldsymbol{A}^T\boldsymbol{A})^{-g} = \boldsymbol{V}\boldsymbol{\Lambda}_p^{-1}\boldsymbol{V}^T, \quad (4)$$

where $\boldsymbol{\Lambda}_p^{-1}$ is the inverse of the diagonal matrix of non-zero eigenvalues, and $\boldsymbol{V}$ is the matrix of corresponding eigenvectors. Aki et al. (1977) estimated the resolution and model errors using the resolution matrix, and the model covariance matrix (see section 2.1), which are still used in recent models. Eqs. (1)–(4) are the basic equations for seismic tomography. The fundamental difference between different tomographic works is the definition of the data vector $\delta\boldsymbol{d}$ (travel-time residuals for the case of travel-time tomography; waveform residuals for the case of waveform inversion, as discussed in section 2.1), and of the partial derivate kernel $\boldsymbol{A}$ in eq. (1) (using ray theory as in Aki et al. 1977, and more recently using finite frequency travel-time kernels, or waveform kernels for waveform inversion).

Aki et al. (1977) noted that the use of the reference model to compute $\bar{d}_{ij}$ is generally justified by Fermat's principle, which states that the travel-time is stationary with respect to small perturbations in the raypath (see the Appendix of Aki et al., 1977), which is valid if the reference model is close enough to the perturbed model. This approximation coming from the linearization of the inverse equation is usually not explicitly verified by tomographers (even for recent models), but was shown to be inaccurate for raypaths turning in the (shallow) upper-mantle, because of strong radial heterogeneities in the upper mantle (Grand, 1987), and because raypaths are concentrated in the high-velocity slab (for subduction zone earthquakes) and avoid the low-velocity mantle wedge (Simmons et al., 2012).

The pioneering work for global tomographic model is that of Dziewonski et al. (1977). Dziewonski et al. (1977) formulated the inverse problem similarly to Aki et al. (1977), but parametrized their model using spherical harmonics up to degree 3 (resolution of ~6700 km at the Earth's surface), and used ~700,000 travel-time residuals for direct P-waves recorded at epicentral distances between 27–105˚, and PKIKP phase at epicentral distance greater than 160˚ recorded worldwide and published in the Bulletins of the International Seismological Center (ISC). They didn't use the P-waves for distance smaller than 27˚ because of possibility of phase misidentification due to multipathing in the upper mantle, but this resulted in poor resolution in the upper mantle. Dziewonski et al. (1977) reported for the first time the strong negative correlation at spherical harmonics degree 2 and 3 between velocity anomalies in the lower mantle below ~1100 km, and geoid (gravity) anomalies, and suggested that this could be explained with a negative conversion factor between velocity and density anomalies (assuming a rigid Earth). However, Hager (1984) later showed that for a viscous Earth, positive density anomalies can actually induce negative geoid anomalies due to the displacement of the surface

and CMB ("dynamic topography"), if the viscosity of the lower-mantle is at least ~30 times that of the upper mantle. Hager (1984) estimated the amount of density anomalies in the mantle based on the location of slabs worldwide given by hypocenters of deep earthquakes, and showed that spherical harmonics degree 2–9 of the geoid anomalies can be relatively well explained if cold and denser slabs extend to the lower-mantle to depths of 700–1200 km (the actual depth is not well constrained). Motivated by the observation of Dziewonski et al. (1977), Hager (1984) thus for the first time used constraints from seismology (location of deep Earthquakes) to place constraints on the rheology of the Earth's mantle.

The same year, Dziewonski et al. (1984) extend the model of Dziewonski et al. (1977) by increasing the spherical harmonics expansion up to degree 6 (i.e., resolution of ~3300 km at the Earth's surface), and established the presence of a ring of high-velocity anomaly regions around the Pacific beneath long-lived subduction zones, which could be remnant of cold slabs that sank down to the core-mantle boundary (CMB). However, the poor resolution of their model in the upper mantle and transition zone did not allow them to reach a definite conclusion on the link between upper-mantle and lowermost mantle high-velocity anomalies (i.e., on whether seismic data support the existence of whole-mantle convection). Dziewonski et al. (1984) also established the dominant spherical harmonics degree-2 pattern of lower-mantle (P-) velocity anomalies, with high-velocities beneath the circum-Pacific, and low-velocities beneath the Pacific and southern Africa. Using this observation, Richards and Engebretson (1992), and Engebretson et al. (1992) computed the amount and distribution of subducted lithosphere between 0–180 Ma based on their previously published plate reconstruction model, and showed that it correlates well with the lower-mantle seismic structure of Dziewonski et al. (1984) (for spherical harmonics degree 2 and 3). Based on this, Engebretson et al. (1992)

proposed that the high-velocity anomalies in the lower-mantle correspond to a "lithospheric graveyard" of slabs accumulated in the lower-mantle.

The dataset of P-wave travel-time residuals used by Dziewonski et al. (1984), which is reported in the ISC Bulletin, was the major dataset used in global P-wave tomography until the work of Engdahl et al. (1998). Engdahl et al. (1998) published an improved catalog of P-wave travel-time residuals (called the EHB catalog) by relocating ~100,000 events in the period 1964–1995 using travel-time residuals reported to the ISC and to the US Geological Survey's National Earthquake Information Center (NEIC) by 1) including depth (pP, sP, ocean-reflected pwP), and core (PKiKP, PKPdf) phases; 2) using an improved reference Earth model, AK135 (Kennett et al., 1995), to compute travel-time residuals (previously, travel-time residuals reported to ISC were with respect to the Jeffrey-Bullen model). Engdahl et al. (1998) re-identified a significant portion of later phases (i.e., phases that arrive after the direct P-wave, such as pP, PcP, sP) that were incorrectly identified in the ISC bulletins, significantly improving the reliability of such phases and making them more easily usable for tomography. The EHB catalog is still used in recent P-velocity models (e.g., Fukao and Obayashi, 2013), and now contains more than 11 million travel-time residuals for the direct P-wave.

In contrast to P-velocity models, which rely on the large dataset of travel-time residuals in the ISC (or later, EHB) catalog, "high-resolution" S-velocity models of the mantle relied on travel-time residuals carefully picked by each author. The reason is that, although ISC Bulletins report S-wave residuals, these are significantly noisier than the P-wave residuals because the stations reporting to ISC are short period vertical component seismometers (dominant period of P-wave is typically 1 s, but that of S-wave is longer, ~4 s) (Masters et al., 2000); hence ISC S-wave residuals were usually not used for tomography (except in a few studies, see below).

Also, ISC S-wave residuals were truncated at $\pm$7.5 s, which is ~2 times less than the largest S-wave residuals for the lower-mantle, leading to underestimated S-velocity anomalies in the lower-mantle for the few models based on ISC S-wave residuals (see below) (Masters et al., 2000).

The first "high-resolution" S-velocity model of the upper and lower-mantle (down to ~1700 km depth) is that of Grand (1987), who inferred the S-velocity structure beneath Central America and the US, using S and SS phases travel-time residuals for records at epicentral distance between 8–80˚. In contrast to the P-wave model of Dziewonski et al. (1984), the use of the S-phase at epicentral distances smaller than 25˚ and of SS phases allowed Grand (1987) to constrain the upper-mantle structure as well as the lower-mantle structure. Grand (1987) measured travel-time residuals of the direct S-wave by picking its onset time on the seismograms, and used cross-correlation of synthetic seismograms (synthetics) with observed records to measure the travel-time residual of later S phases (S-wave triplications, see below) and SS phases. His model was parametrized using constant-velocity blocks of lateral dimensions of ~500 km by 500 km (~6 times smaller than in the model of Dziewonski et al., 1984). The model of Grand (1987) is the first tomographic model to show the presence of a tabular high-velocity anomaly extending nearly vertically from ~700 km to ~1700 km (the model has poorer resolution below ~1600 km depth) beneath Central America, which he associated with the subduction of the Farallon plate.

We note that earlier studies modelling variations in travel-times of deep earthquakes as function of the takeoff angle (called "residual sphere analysis") had reported that slabs beneath the Sea of Okhotsk (Jordan, 1977; Creager & Jordan, 1984), and Mariana, Japan, and Izu-Bonin (Creager & Jordan, 1986) penetrates into the lower-mantle down to more than ~1000

km depth with a steep dip angle and nearly no deformation at ~660 km depth. However, these observed travel-time time variations where later shown to be mostly due to the (relatively small scale) 3-D structure of the mantle outside of the slab, weakening these reported evidences for simple penetration of slabs into the lower-mantle (see the review of Lay, 1994).

The S-waves at epicentral distances less than ~25–30˚ used by Grand (1987) have multiple arrivals separated by typically ~0–10 s corresponding to multiple paths in the upper mantle and MTZ due to strong S-velocity discontinuities at ~410 km and ~660 km depth. Travel-times of later-arriving phases are difficult to measure, since the time-separation between phases is small and they frequently overlap. Also, later phases are prone to phase misidentification that could strongly bias the inferred model, so that most authors use direct (first-arrival) phases only. Grand (1987) (and his more recent models) is one of the few authors that use later arriving S-wave triplications in his travel-time tomographic model. Another work that used triplicated phases in P-wave travel-time tomography is Takeuchi et al. (2014).

Shortly after the work of Grand (1987), Inoue et al. (1990) derived the first "high-resolution" global P-velocity model of the whole mantle by using ~2 million direct P-wave travel-time residuals from the ISC Bulletins at epicentral distance < 95˚ (including epicentral distances smaller than 25˚). Their model was parametrized in 32,768 constant-velocity voxels of dimension 5.625˚ by 5.625˚ (~570 km at the Earth's surface, i.e., much smaller than the ~3300 km of Dziewonski et al. 1984, comparable to the size of blocks used by Grand 1987, and ~2.5–5 times the size of voxels used in the most recent P-velocity models, e.g., Obayashi & Fukao, 2013) in 16 depth layers. The model of Inoue et al. (1990) confirmed the high-velocity anomaly beneath Central America extending from ~630–2250 km depth, and showed high-velocity anomalies beneath the Japan and Tonga trench extending from the surface to

~1000 km depth (the depth extent of the slab at the Japanese trench is discussed again below). Inoue et al. (1990) also observed strong low-velocity anomalies in the mantle wedge in both regions, and a strong low-velocity anomaly beneath the Pacific slab at the Japan trench, which are features that are also found in recent models (e.g., Obayashi & Fukao, 2013). Grand (1987) and Inoue et al. (1990) thus showed some of the first (reliable) direct evidence for penetration of (some) slabs into the lower-mantle. Note, however, that because neither Grand (1987) nor Inoue et al. (1990) used core-reflected phases (ScS, PcP), or core phases (PKP, SKS), their models have poorer resolution in the lowermost mantle (as they are constrained only by direct P- and S-waves). It was thus not possible, based on their models, to tell whether slabs sink down to the CMB, or if the lowermost mantle is isolated from the rest of the lower-mantle as will be suggested later by van der Hilst et al. (1997).

Fukao et al. (1992) extended the work of Inoue et al. (1990) in the region around the Japanese Islands (where the number of data is large so that the resolution is high) by reducing the size of blocks used to parametrized the model to 150 x 150 km in the lateral direction, and 50 km in the vertical direction (i.e., ~1–1.5 times the size of blocks used in recent models). Fukao et al. (1992) imaged high-velocity anomalies extending from the Earth's surface and deflected horizontally at ~670 km depth (above the upper mantle) beneath the southern Kurile, Japan, and Izu-Bonin arc, with some "blobs" of strong high-velocity anomaly beneath the Japan arc extending to ~800 km depth in the lower-mantle beneath the horizontal part of the slab. In this region, relatively strong high-velocity anomalies are present in the lowermost mantle (beneath eastern China), but the connection between these and the upper-mantle slabs is not clear (the lower-mantle in this region is dominated by intermittent high-velocity anomalies, and broad-scale weak high-velocity anomalies, but there is no clear slab-like tabular

high-velocity anomaly in the lower-mantle). We note that this differ from the results of Inoue et al. (1990) who claimed that slabs beneath Japan reach to ~1000 km depth into the lower-mantle, and shows the difficulty of imaging slabs, which are small-scale structures. Fukao et al. (1992) interpreted these stagnant slabs as a "megalith" resulting from the buckling of a slab as it interacts with the 670 km discontinuity (Ringwood & Irifune, 1988), and eventually sinks into the lower mantle because of excess weight if the megalith is cold enough. Fukao et al. (1992) also found a slab beneath the Java arc that extends to ~1200 km depth.

Note that in more recent fully dynamic numerical simulations (i.e., where plate velocity at the surface is not imposed as a boundary condition, but evolves in response to local forces), buckling of slabs above the 660 km discontinuity is significantly weaker than in simulations where surface plate velocities are imposed as boundary conditions (Christensen, 1996), which would make it more difficult to form a megalith (Quinteros et al., 2010).

Grand (1994) improved the model of Grand (1987) by including several additional phases: the core-reflected ScS phase that greatly improves resolution in the lowermost mantle, and the multibounce phases at the Earth surface SSS and SSSS that improve resolution for the shallow upper-mantle structure. He also reduced the size of blocks to parametrize the model to 275 x 275 km of lateral dimensions (from 500 km in his previous model). His model clearly showed the presence of high-velocity anomalies beneath the Americas extending from at least ~750 km down to the core-mantle boundary (CMB), deeper than in Grand (1987) due to the use of the core-reflected ScS phase (the MTZ is the poorest resolved region in his model, which may explain why this high-velocity anomaly does not seem to extend to shallower depths). This allowed him to estimate the sinking velocity of slabs in the lower-mantle to be ~1–2 cm/yr, assuming that the high-velocity structure at the CMB correspond to paleoslabs that were

subducting beneath the Americas at ~150 Ma. We note that this estimate is in agreement with more recent work (van der Meer et al., 2010; Domeier et al., 2016).

As mentioned above, Engdahl et al. (1998) published an improved dataset of P-wave travel time residuals (the EHB catalog), including later phases such as the depth phases pP, sP, and the ocean-surface reflected pwP, and the core phase PKP. Depth phases (pP, sP, pwP) give additional constraints on the shallow structure above the hypocenters of the events, while the PKP core phase gives additional constraints on the lowermost mantle. The first tomographic model to use this improved dataset is the global model of van der Hilst et al. (1997), who used all of the available above-mentioned P-wave phases (direct and later phases) in the EHB catalog, and parametrized their model using constant-velocity blocks of lateral dimension 2˚ x 2˚ (~200 km at the Earth's surface, which is ~1–2 times the size of blocks in recent models). Their model showed continuous subduction into the lower-mantle in several of the long-lived subduction zones: 1) beneath Central America down to the CMB (Farallon slab), confirming the observations of Grand (1994), 2) beneath southern Eurasia down to ~1500 km depth (Aegean slab; as discussed later, this slab was imaged deeper in a subsequent S-velocity model), 3) beneath Tonga-Kermadec down to ~1400 km depth. Beneath the northwestern Pacific, slabs are deflected horizontally in the MTZ, in general agreement with the model of Fukao et al. (1992). While one cross-section in van der Hilst et al. (1997) suggest continuous subduction beneath central Japan, Fukao et al. (2001) later argued that the high-velocity anomaly in the top of the lower-mantle is much weaker than that of the slab in the upper-mantle, and that thus there is no strong evidence for direct penetration of the slab into the lower-mantle beneath central Japan.

Based on the global model of van der Hilst (1997), van der Hilst et al. (1999) argued for compositional stratification in the bottom 1000 km depth of the mantle. This was based on 1) a shift from short (spatial) wavelength slab-like tabular high-velocity anomalies in the upper-mantle and top of the lower-mantle to long-wavelength broad high-velocity anomalies below ~1700 km depth, and 2) the fact that the depth of this change in the wavelength of high-velocity anomalies corresponds roughly to an increase in the bulk- to S-velocity ratio reported in the global joint S- and P-velocity tomography study of Kennett et al. (1998). High bulk- to S-velocity ratio in the lower-mantle was thought to indicate the presence of a chemically distinct layer, suggesting that convection in this layer is isolated from the rest of the mantle.

However, Masters et al. (2000) argued that the anomalously high bulk to S-velocity ratio in Kennett et al. (1998) was due to the use of S-wave residuals reported in the ISC Bulletins, which were clipped at $\pm$7.5 s, approximately 2 times smaller than the maximum direct S-wave residual measured by Masters et al. (2000) for S-waves that samples the lower-mantle, resulting in underestimated S-velocity perturbations. In contrast, Masters et al. (2000) found a higher bulk- to S-velocity ratio values that can be explained a thermal origin of velocity anomalies (Karato, 1993), except in the lowermost 500 km of the mantle, where the bulk- and S-velocity become negatively correlated. We note that this negative correlation is now thought to indicate the presence of Mg-post-perovskite (Mg-pPv) (Murakami et al., 2004; Tsuchiya et al., 2004) in the lowermost mantle (Wentzcovitch et al., 2006; Koelemeijer et al., 2018).

We conclude this subsection on earlier tomographic models by discussing the work of Grand et al. (1997) and Fukao et al. (2001). Grand et al. (1997) presented a new high-resolution global S-velocity model obtained using the methods of Grand (1994), and compared it to the model of van der Hilst et al. (1997). They showed the first good agreement between global P-

and S-velocity models for small scale (high) velocity anomalies. The two models showed good agreement on most of the slab structure that extends into the lower-mantle discussed by van der Hilst et al. (1997). In particular, the Farallon slab extends down to the lowermost mantle in both models. There are some differences, though, for instance for the Aegean slab that extends down to the lowermost mantle in the S-velocity model of Grand et al. (1997), but seems to disappear below ~1800 km depth in the model of van der Hilst et al. (1997).

Fukao et al. (2001) reviewed the fate of slabs in the circum-Pacific region based on their recent P-velocity model, on the P-velocity model of van der Hilst (1997), and on two low-resolution global S-velocity models. Fukao et al. (2001) argued for a majority of slabs in the circum-Pacific region stagnating (e.g., beneath central Japan) or stopping (e.g., beneath the Marianas) between 660–1000 km depth. 1000 km depth corresponds to the lower limit of the transition region as defined originally by Bullen (1963). Although there seems to be no strong global seismic discontinuity at ~1000 km depth (or mineral phase transition) that could justify stagnation of slabs at ~1000 km depth, it had recently been shown that there might be a seismic discontinuity around this depth (~920 km depth) beneath some subduction zones, e.g., beneath Tonga, and beneath the Japan and Flores sea (Java) (Kawakatsu & Niu, 1994). A recent global study of mid-mantle reflectors using SS phase precursors (long period and sensitive to horizontal reflectors of more than ~500 km in size) find no global lower-mantle discontinuity, but find evidences for regional reflectors extending over ~1500 km laterally at depths of ~850, and ~1050 km in two “neutral” regions (i.e., where tomographic models report no significant velocity anomalies), which they interpret as possibly due to (regional) compositional layering or ancient slabs trapped due to (regional) viscosity layering (Waszek et al., 2018). Also, other studies using S-to-P converted wave to detect small-scale scatterers (dimensions of ~10 km) in

the lower-mantle found such scatterers beneath subduction zones in the circum-Pacific regions beneath the currently subducting Pacific and Nazca slabs down to ~1800 km depth (Kaneshima & Helffrich, 2010; Kaneshima, 2016), which they interpreted as possible evidence for the presence of basaltic composition heterogeneities entrained into the lower-mantle by past subduction. Such scatterers are not detected beneath ~1800 km depth, and in the western Pacific region their distribution peaks between ~800–1100 km depth, which could indicate layering, but could also be explained by the strong decrease in the S-velocity contrast between basaltic and pyrolitic composition from ~3% at ~700 km depth to ~1% at ~1300 km depth (Tsuchiya, 2011). Finally, as reviewed by Kaneshima (2016), there are several other studies using different kind of scattered waves (e.g., P-to-P scattering) that detected scatterers in the lower half of the lower-mantle, suggesting the presence of subducted material deep into the lower-mantle.

1.2 Developments: finite-frequency and body-wave waveform tomography

The tomographic models presented in the previous section are all based on ray-theoretical interpretations of the observed travel-time residuals. In this framework, it is assumed that these residuals are caused by the 3-D structure along the raypath only, with no contribution from 3-D structure away from the raypath. This is exact only at infinitely high frequencies, but is justified by the fact that the dominant period of P-, and S-waves used in the above-mentioned models is high: ~1 s (wavelength of ~10 km), and ~4 s (wavelength of ~30 km), respectively (Masters et al., 2000), and by the fact that the part of the waveform near the onset time corresponds to the higher-frequency content (lower-frequencies typically travel with a smaller phase velocity due to attenuation).

However, when the size of heterogeneities is comparable to the wavelength, diffraction effects (not taken into account by ray theory) become important (e.g., "wavefront healing": Nolet & Dahlen, 2000), and a more accurate theory for the interpretation of travel-times is necessary to resolve small-scale structures of size comparable to seismic wavelengths. Furthermore, the use of waveform data instead of travel-time data can significantly increase the degree of constraint on the structure, since waveforms contain broadband information (i.e., the shape) and amplitude information; however, waveform inversion requires the use of full-wave theory to be sufficiently accurate.

Several theoretical frameworks were developed for the inversion of waveform data (Tarantola & Valette, 1982; Nolet, 1990; Geller & Hara, 1993; Li & Romanowicz, 1995; Fichtner et al., 2008), and of "finite-frequency" travel-times, which take into account the sensitivity of travel-times away from the raypath, and are measured by cross-correlation of synthetics with observed waveforms (Dahlen et al., 2000; Hung et al., 2000).

Again, we focus here on "high-resolution" tomographic models, obtained using relatively short period body-waves (or surface-wave overtones, or a combination of both). There were several earlier studies that used waveform data for the case of long-period body-waves (e.g., Su et al., 1994), or surface-waves (e.g., Woodhouse & Dziewonski, 1984), but these have typically low resolution and did not aim to look at the detailed structure of slabs.

One of the first "high-resolution" models to use finite-frequency information was obtained by waveform inversion of surface-wave overtones and S-wave triplications to image the more detailed structure of the Farallon plate in the upper mantle and MTZ beneath North America (van der Lee & Nolet, 1997a, 1997b).

Hung et al. (2005) used finite-frequency travel-time kernels (called "banana-doughnut" kernels because of their particular shape) to image more detailed structure of the lowermost mantle (D″ region) beneath Central America; their study revealed much smaller scale structure of the broad high-velocity anomalies at the CMB in this region than had commonly been imaged in previous global models (Grand et al., 1997; van der Hilst et al., 1997), suggesting the existence of smaller scale convection that could suggest stiff and cold slabs at the CMB.

Recently, Kawai et al. (2014), and Konishi et al. (2014) developed methods for localized waveform inversion for the 3-D S-velocity structure of the D″ region. Suzuki et al. (2016), and Borgeaud et al. (2017) applied these methods to large datasets of waveforms from the full deployment of the USArray (a large array of ~2500 seismic stations) in the contiguous US during 2004–2015 to image the 3-D S-velocity structure beneath the Northern Pacific, and Central America and the Caribbean, respectively. These studies confirmed the presence of strong high-velocity anomalies in the D″ region beneath two long-lived subduction zones (corresponding to paleoslabs), but also the presence of strong low-velocity anomalies just above the CMB extending vertically around the high-velocity anomalies, resulting in the inference of a smaller-scale convection pattern than had been possible in previous global studies. Compared to the model of Hung et al. (2005), the model of Borgeaud et al. (2017) shows stronger low-velocity anomalies just above the CMB, and more consistent vertical features that can be interpreted as slabs or possible small-scale plumes originating at the CMB.

Takeuchi (2007) published the first global S-velocity model using the exact formulation of waveform kernel of Geller and Hara (1993). His model confirmed the stagnation of slabs at 660 km beneath Central Japan. Although this is not discussed in Takeuchi (2007), his model also seems to show the presence of slabs beneath Kermadec down to more than ~1200 km, and

beneath Java down to more than ~1500 km (in general agreement with Fukao et al., 2001, but in apparent disagreement with Fukao & Obayashi, 2013), but also beneath northern Kurile down to ~2000 km (in apparent disagreement with Fukao et al., 2001).

Sigloch (2008) published the first model using finite-frequency tomography with travel-time and amplitude kernels measured in several frequency bands to increase the number of constraints. We note that earlier finite-frequency models typically used a single frequency, and did not use amplitude information. On the other hand, models based on waveform inversion (e.g. Takeuchi, 2007) intrinsically use broadband information and amplitude information. Sigloch (2008) showed complex subduction of the Farallon plate beneath North America, which she linked to the Laramide orogeny at ~50–70 Ma. In her model, the Farallon slab penetrates deep into the lower-mantle, in agreement with earlier studies (Grand et al. 1997; van der Hilst et al., 1997).

Obayashi et al. (2013) updated the model of Fukao et al. (2001) by using finite-frequency travel-time kernels. Based on this model, Fukao and Obayashi (2013) argued that stagnation of slabs between 660–1000 km depth is the dominant convection style in the circum-pacific region. Some important changes compared to the work of Fuako et al. (2001) were for slabs beneath Java, and Kermadec, previously seen as subducting down to ~1200, and 1400 km depth, respectively, but both shown as extending horizontally at ~1000 km depth in the more recent model of Obayashi et al. (2013). We further note that they interpreted the Farallon slab beneath Peru, and Chile, as trapped at ~1000 km, and stagnating at ~660 km, respectively. However, a more recent regional travel-time tomography model suggest that the Farallon slab penetrates into the lower-mantle beneath both Peru and Chile, and reach at least 900 km depth (where their model stops), suggesting that the fate of the Farallon slab in this region is still uncertain

(resolution tests in Obayashi et al., 2013, indeed show relatively poor resolution in the lower-mantle in this region).

In summary, previous work in seismic tomography established that slabs in some locations (currently) stagnate above the upper-mantle (all authors agree on the stagnation of slabs from southern Kurile to Izu-Bonin), while they penetrate deep (i.e., more than ~1000 km depth) into the lower-mantle (e.g. Aegean slab beneath south-east Europe) and reach the CMB in some other locations (e.g., beneath Central America). For some subduction zones, however, there seem to be disagreements between different authors on the maximum penetration depth of slabs. This is the case for slabs beneath Java, Kermadec, South-America, and possibly Kamchatka. Also, the tectonic history is not always considered when interpreting the fate of slabs beneath some subduction zones. For instance, the slab beneath the Mariana trench seems to stop at ~1000 km depth (Fukao & Obayashi, 2013), but this probably does not imply that it is "trapped" at ~1000 km depth, since subduction at the Mariana trench is thought to have initiated around 50 Ma (Seno & Maruyama, 1984), which is in broad agreement with a slab reaching ~1000 km depth.

Finally, the fate of slabs in the deeper part of the lower-mantle is not clear. Previous global models all noted an increase in the length-scale (spatial wavelength) of high-velocity anomalies in the lowermost mantle beneath long-lived subduction zones (e.g., in the circum-Pacific region), which has sometimes been interpreted as evidence for a separate convective layer in the lowermost mantle (van der Hilst et al., 1999), or for an increase in viscosity around ~2000 km depth (Forte & Mitrovica, 2001). However, regional studies seem to show smaller-scale structures in the lowermost mantle, typically with stronger small-scale low-velocity anomalies (e.g., Hung et al., 2005; Kawai et al., 2014).

To summarize the reported fate of slabs, I show in Figure 1 a global view of subduction zones labeled following the classification of Fukao and Obayashi (2013), which is as follows: 1) slabs stagnating at 660 km depth, 2) slabs that penetrate into the lower-mantle, 3) slabs that are trapped in the uppermost lower mantle, between 660–1000 km depth (this includes both slabs that seems to be deflected horizontally at ~1000 km depth, and slabs that seems to stop abruptly at ~1000 km depth, but without horizontal deflection, as beneath the Mariana arc), 4) slabs that penetrate deep into the lower mantle. Classifications are coded by colors from red to blue (see caption in Figure 1). Subduction zones with multiple labels correspond to regions with different studies reporting different subduction modality for the same slab (the color of the cross-section line indicates the deepest penetration depth of the corresponding slab as reported in at least one previous study).

1.3 Global correlation between slabs and past plate boundaries

Finally, we note that recently, several studies have aimed to establish the correlation between high-velocity anomalies in the Earth's mantle obtained by tomographic models, and location of past plate boundaries. Some of these studies are based on a single global P-velocity model inferred using travel-time tomography and a 3-D initial model by Amaru (2007) (van der Meer et al., 2010, 2018), while other studies use several (~10) S- (Domeier et al., 2016) and P-velocity models (Shephard et al., 2017).

Domeier et al. (2016) found a global correlation between the location of past subduction zones from 10–200 Ma, and of slabs in the lower-mantle as imaged in several previous tomographic models, down to at least ~2300 km depth, assuming that slabs sink vertically. This suggests penetration of slabs deep into the lower-mantle on a global scale, although the fact that slabs in the lower-mantle typically appears as broadened high-velocity anomalies

(probably a combination of an increase in viscosity in the lower-mantle, of the complexities induced at 660 km depth, and of the limited resolution of tomographic models) implies that this correlation cannot represents well the complexities in convection modality around 660–1000 km depth discussed in previous tomographic models (e.g. Fukao & Obayashi, 2013).

Van der Meer et al. (2010, 2018) conducted a more detailed (but perhaps less robust due to the use of a single tomographic model) comparison between high-velocity anomalies in the lower-mantle, and past subduction zones. Van der Meer et al. (2018) classify 94 slabs, giving their top and bottom depths, and the ages of initiation and cessation of subduction, relying on previous interpretations and existing geophysical records. Around half of the slabs classified by van der Meer et al. (2018) reside in the lower-mantle only. One particularity of their interpretation is that they propose that several of these lower-mantle paleoslabs are remnants from past intra-oceanic subduction zones, which are not included in plate reconstruction models (e.g., Müller et al., 2016).

1.4 Full waveform tomography

Due to the increase in computational power in the last two decades, several studies have started to perform waveform inversion using 3-D reference models and computing exact synthetics for 3-D models using the full-wave theory, which is commonly call "full-waveform tomography" (we note that this is a confusing designation, since many works using full-waveform tomography do not actually use all the information in the waveforms, but often use only the finite-frequency travel-time information). French and Romanowicz (2014) published the first global model obtained using a hybrid approach by computing exact 3-D synthetics at each iteration, and efficient approximate partial derivative kernels using the NACT (Li & Romanowicz, 1995), a waveform misfit (L2 norm of the waveform residual), and waveforms

down to a period of ~32 s. Bozdag et al. (2016) published the first global model obtained using the so-called adjoint tomography method (Tromp et al., 2005) by using the 3-D wave-propagation code SPECFEM3D_GLOBE (Komatitsch et al., 2015) for both the forward simulations of synthetics and the computation of the partial derivative kernel at each iteration, which required a significant amount of computational power (~19,000 GPU). Bozdag et al. (2016) used a phase misfit (finite-frequency travel-time information), and waveforms down to a period of ~17 s. Zhu et al. (2015, 2017) used adjoint tomography and SPECDEM3D_GLOBE (Komatitsch et al., 2015) to image the regional upper-mantle and transition zone structure beneath Europe, and the US, respectively. They used a phase and amplitude misfit criterion, and waveforms down to a period of 15 s. Tao et al. (2018) used adjoint tomography with SPECFEM3D_GLOBE (Komatitsch et al., 2015), and a cross-correlation misfit criterion to make use of the information contained in S-wave triplications to image the upper-mantle and MTZ beneath Japan and eastern China. Tao et al. (2018) used waveforms down to a period of 8 s.

We note that the formulation of the inverse problem when using adjoint tomography differs from that introduced by Aki et al. (1977) in eq. (3), who solve the normal equation written using the $\boldsymbol{A^T A}$ matrix, sometimes called the approximate hessian. In adjoint tomography, the misfit is minimized using a gradient descent based approach, which has a slower convergence rate than when formulating the inverse problem using the $\boldsymbol{A^T A}$ matrix (Chen et al. 2007). The reason is that adjoint tomography does not have access to the $\boldsymbol{A^T A}$ matrix, since for each event only one back-propagated wavefield is computed (with the source term being given by the sum over all receivers of each individual adjoint source), while kernels for all individual records are required to compute the $\boldsymbol{A^T A}$ matrix. On the other hand, this makes the computation of kernels

in adjoint tomography significantly faster (i.e., of the order of the computation cost for the forward propagation), because there is no need to compute the back-propagated wavefield for each receiver (which are often of the order of thousands).

Finally, we note that although the several iterations performed in the works using adjoint tomography with SPECFEM3D_GLOBE (Komatitsch et al., 2015) account for the non-linearity of the inverse problem, some other issues, such as the dependence on the initial model, still remain. In this respect, the methods for waveform inversion used in this study might allow more efficient quantification of the dependence on the initial (1-D) model, because 1) the computation of the partial derivative kernel is faster than for one iteration of adjoint tomography (due to the use of a 1-D reference model and wavefield interpolation), and 2) one iteration using the $\boldsymbol{A^T A}$ matrix should be equivalent to several iterations using the adjoint kernel (Chen et al., 2007).

1.5 This study

As discussed previously (and shown in Figure 1), not all authors agree on the fate of slabs in the Earth's mantle. Some authors have argued that most of the slabs are trapped between 660–1000 km depth (e.g., Fukao & Obayashi, 2013), but different tomographic models sometime show disagreement in the maximum penetration depth of slabs into the lower-mantle, and in their connection to high-velocity anomalies in the lowermost mantle. Also, global analysis of multiple tomographic models shows a global correlation between past plate boundaries between 10–180 Ma, and slabs in the lower-mantle down to ~2300 km depth (Domeier et al., 2016), suggesting that slabs generally penetrate deep into the lower-mantle, but with the connection to high-velocity anomalies in the lowermost mantle still uncertain. Finally, the (spatial) wavelength of velocity anomalies in the lowermost mantle (related to the

convection style in the lowermost mantle) is also most likely smaller than reported by previous global studies (as suggested by regional studies, e.g., Hung et al, 2005; Kawai et al., 2014). The above-mentioned disagreement in the fate of slabs is due to the limited resolution of tomographic models (and also to uncertainties in past plate boundaries and past plate motion).

In this dissertation, in order to improve our understanding of the fate of slabs, I image the 3-D S-velocity structure in the D″ region, and in and near the MTZ beneath Central America and its vicinity with higher resolution than in previous tomographic models. As discussed previously, these two regions are important in investigating the fate of slabs, since these are the two boundary layers of the Earth's mantle. The imaging of a possible barrier to convection at ~1000 km depth as suggested by previous studies (Fukao & Obayashi, 2013) is a topic for future studies. Improvement over previous studies comes from 1) the use of a large dataset made available from the recent full deployment of the USArray transportable network in the contiguous US from 2004–2015, 2) the use of waveform inversion based on the exact formulation of partial derivatives of Geller and Hara (1993), as developed recently for the case of the D″ region (Kawai et al., 2014; Konishi et al., 2014), and 3) for the MTZ study, the use of S-wave triplications waveforms. The merits of using waveform inversion are discussed in sections 3.2 and 4.2, and the merits of using S-wave triplications are discussed in section 4.2.

For the case of the D″ region (Chapter 3), this is the first study to use a large waveform dataset from the full deployment of the USArray in the contiguous US to image the high-resolution 3-D S-velocity structure of the D″ region beneath Central America and the Caribbean. Compared to Kawai et al. (2014) and Konishi et al. (2014), I also introduce some small technical improvements, such as weighting the data to roughly equalize the uneven distribution of azimuths and epicentral distances that arises due to the broad extent of the full

USArray. I also conducted for the first time in our group's work several additional tests to check the resolution and robustness of the models inferred by waveform inversion. One of these tests is the "nonlinear checkerboard test", that allows testing the accuracy of the linear (Born) approximation made when formulating the inverse problem; this test is computationally intensive due to the need to compute exact synthetics for a 3-D checkerboard pattern (using the full-wave propagation code SPECFEM3D GLOBE; Komatitsch et al., 2015).

For the case of the MTZ (Chapter 4), this work represents the first application of our group's localized 3-D waveform inversion method to infer the high-resolution 3-D S-velocity structure of the MTZ. This required several improvements, in particular to compute partial derivatives more efficiently and accurately, and a new way to compute corrections for shallow structures (see section 2.2). Improvements in the efficiency and accuracy of partial derivatives will also facilitate future inversions for the 3-D structure of the D″ region.

I conclude this work by discussing the implications of the inferred models of the MTZ and D″ for the viscosity contrast between the upper- and lower-mantle (Chapter 5).

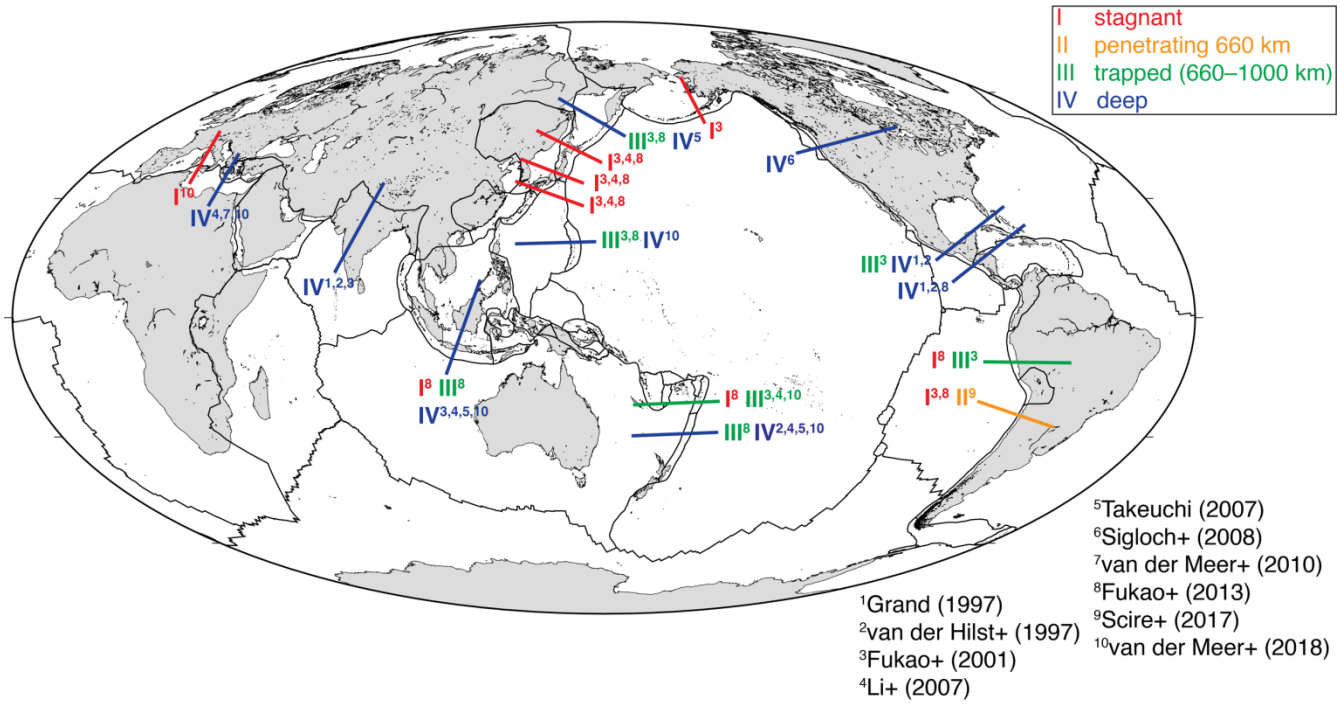

Figure 1. Global view of the fate of slabs as imaged by several tomographic models. Subduction zones (showed by colored lines with labels) are classified following Fukao and Obayashi (2013), depending on whether (I) the slab stagnates above the lower-mantle (at ~660 km depth), (II) the slab penetrates into the lower-mantle, (III) the slab seems to stagnate, or stop between 660–1000 km depth, and (IV) the slab reaches depth larger than 1000 km depth. Colors of the lines showing subduction zones are based on the deepest reported penetration depth of the slab. We note the disagreement in the fate of slabs for several subduction zones, due to the limited resolution of tomographic models, although all authors agree on stagnation of slabs above the lower-mantle from southern Kurile to Izu-Bonin, and of deep penetration of slabs beneath Central America and the US.

This page is intentionally left blank.

Chapter 2 Method

2.1 Inverse problem

In general, the synthetic wavefield $\boldsymbol{u}_{syn}(\boldsymbol{r})$ at a seismic station located at a point $\boldsymbol{r}$ at the Earth's surface is a nonlinear functional of the Earth's model $\boldsymbol{m}$ (eq. (5))

$$\boldsymbol{u}_{syn}(\boldsymbol{r}) = \boldsymbol{g}[\boldsymbol{m}](\boldsymbol{r}). \tag{5}$$

In general, the recorded wavefield $\boldsymbol{u}_{obs}(\boldsymbol{r})$ differs from the synthetic wavefield $\boldsymbol{u}_{syn}(\boldsymbol{r})$ by a quantity called the waveform residual (eq. (6))

$$\boldsymbol{r} = \boldsymbol{u}_{obs} - \boldsymbol{u}_{syn}. \tag{6}$$

Here, we assume that we have a large number of seismic records cut around the seismic phases used for the inversion (e.g., S and ScS for the case of D″, and the direct S- and S-triplications for the case of the MTZ) for a total of N time-windows. The residual $\boldsymbol{r}$ is built by juxtaposition of the N time-windows. If the functional $\boldsymbol{g}$ in eq. (5) is not too nonlinear, and if we already have a good guess of the initial Earth's model $\boldsymbol{m}^0$, eq. (6) can be linearized around $\boldsymbol{m}^0$, as in eq. (7):

$$\begin{aligned} \boldsymbol{r} = \boldsymbol{u}_{obs} - \boldsymbol{g}[\boldsymbol{m}^0 + \delta\boldsymbol{m}](\boldsymbol{r}) \approx \boldsymbol{u}_{obs} - \boldsymbol{g}[\boldsymbol{m}^0](\boldsymbol{r}) - \frac{\delta \boldsymbol{g}}{\delta \boldsymbol{m}}[\boldsymbol{m}^0](r)\delta\boldsymbol{m} \\ \Leftrightarrow \boldsymbol{r} \approx \delta\boldsymbol{d} - \frac{\delta \boldsymbol{g}}{\delta \boldsymbol{m}}[\boldsymbol{m}^0](r)\delta\boldsymbol{m}, \end{aligned} \tag{7}$$

where $\delta\boldsymbol{m}$ is the vector of perturbations from the initial model $\boldsymbol{m}^0$, and $\delta\boldsymbol{d} = \boldsymbol{u}_{obs} - \boldsymbol{g}[\boldsymbol{m}^0](\boldsymbol{r})$ is the waveform residual for the initial model. The inverse problem seeks to find $\delta\boldsymbol{m}$ so that $\boldsymbol{r} = 0$. In practice, the Earth's model $\delta\boldsymbol{m}$ is developed using a finite basis of M elements (constant-velocity blocks in this work), and the residual $\delta\boldsymbol{d}$ is discretized by sampling the waveforms with a sampling frequency of 1 Hz, and contains NK points, where

$K \approx 80$ represents the average number of points in a time-window. The discretized version of eq. (7) for $\boldsymbol{r} = 0$ is given by eq. (8) (assuming a strict equality)

$$\boldsymbol{A}\delta\boldsymbol{m} = \delta\boldsymbol{d}, \tag{8}$$

where $\boldsymbol{A}$ is called the partial derivative kernel

$$A_{ij} = \frac{\delta g_i}{\delta m_j}. \tag{9}$$

If the inverse problem is well posed, we have $M \ll NK$, i.e., the number of model parameters is much smaller than the number of data points. This means that eq. (8) does not have a solution, and the best we can do is to minimize $\|\boldsymbol{r}\|^2$, i.e. finding the least-square solution of eq. (8). We have

$$\|\boldsymbol{r}\|^2 = (\boldsymbol{A}\delta\boldsymbol{m} - \delta\boldsymbol{d})^T\boldsymbol{W}(\boldsymbol{A}\delta\boldsymbol{m} - \delta\boldsymbol{d}), \tag{10}$$

Where we introduced the (diagonal) data weighting matrix $\boldsymbol{W}$, used to roughly equalize the contribution of each record, and defined as in eq. (11)

$$W_{ij} = \delta_{ij}\frac{1}{\max\limits_{k\in\{T_i\}}(\boldsymbol{u}_{obs})_k}, \tag{11}$$

where $\{T_i\}$ is the set of indices of data points that are contained in the time-window to which data point i belongs to. For the case of the D″ layer, we also weight the data to equalize the uneven azimuthal distribution of records (see section 3.2 for details). Setting the gradient of eq. (10) to zero leads to the least-square equation that defines the inverse problem eq. (12)

$$\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}\delta\boldsymbol{m} = \boldsymbol{A}^T\boldsymbol{W}\delta\boldsymbol{d} \tag{12}$$

In this dissertation, the synthetics (eq. (5)) and partial derivatives (eq. (9)) are computed for initial 1-D models (PREM for the case of D″, and AK135 for the case of the MTZ) using the DSM (Kawai et al., 2006).

2.1.1 Regularization

It is well known that eq. (12) has a unique solution only if the matrix $\boldsymbol{A^T W A}$ is symmetric and positive definite. This is not the case if either 1) some column of the $\boldsymbol{A^T W A}$ matrix contains only zeros (i.e. some model parameters are outside of the region where the data have sensitivity), or 2) if some column of the $\boldsymbol{A^T W A}$ matrix are linearly dependent. Condition (1) is easy to avoid by defining the model only in regions sampled by the data (especially when using finite-frequency kernels that have sensitivity in a finite volume around the raypath). Condition (2), however, can be violated if the model is overparametrized, i.e., if the resolution of the whole dataset is smaller than the size of the blocks used to parametrized the model. In this case, it is necessary to modify eq. (7) to make the matrix on the left-hand side invertible. The simplest way to do it is by adding a constant that multiplies the identity matrix, which leads to the damped least-square equation eq. (13)

$$(\boldsymbol{A^T W A} + \lambda \boldsymbol{I})\delta \boldsymbol{m} = \boldsymbol{A^T W}\delta \boldsymbol{d}, \tag{13}$$

where λ is a damping factor. Another possibility is to add some a priori assumption on the solution, such as that it must be smoother than a typical length scale h, by introducing a smoothing operator $\boldsymbol{C_m}$ defined as in eq. (14)

$$(\boldsymbol{C_m})_{ij} = c_{ij} \exp\left(-\frac{\Delta_{\mathrm{ij}}^2}{h^2}\right), \tag{14}$$

where Δ_{ij} is the epicentral distance (in degrees) between model parameters i and model parameter j, and h is the horizontal correlation length (in degrees). The factor c_{ij} is computed as in eq. (15)

$$c_{ij} = \alpha \frac{Tr(\boldsymbol{A^T W A})}{M}, \tag{15}$$

where Tr denotes the trace, M is the number of model parameters (and is also the dimension of the matrix $\boldsymbol{A}$), and α controls the amount of smoothing and damping.

Using $\boldsymbol{C}_m$, the inverse problem becomes

$$(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \boldsymbol{C}_m^{-1})\delta\boldsymbol{m} = \boldsymbol{A}^T\boldsymbol{W}\delta\boldsymbol{d}. \tag{16}$$

It is possible to re-write eq. (16) in the form of eq. (13) by letting

$$\boldsymbol{C_m} = \boldsymbol{L}\boldsymbol{L}^T, \tag{17}$$

$$\boldsymbol{B} = \boldsymbol{W}^{1/2}\boldsymbol{A}\boldsymbol{L}, \tag{18}$$

$$\boldsymbol{x} = \boldsymbol{L}^{-1}\delta\boldsymbol{m}, \tag{19}$$

$$\boldsymbol{b} = \boldsymbol{W}^{1/2}\delta\boldsymbol{d}. \tag{20}$$

which leads to eq. (21)

$$(\boldsymbol{B}^T\boldsymbol{B} + \boldsymbol{I})\boldsymbol{x} = \boldsymbol{B}^T\boldsymbol{b}. \tag{21}$$

Comparison of eqs. (13) and (21) shows that $c_{ii} \approx \lambda^{-2}$. Eq. (21) is more advantageous than eq. (13), since the Cholesky decomposition of $\boldsymbol{C_m}$ (eq. (17)) is faster and more stable numerically to compute than its inverse.

We note that there are other possible choices to impose a priori constraints on the solution, such as the first or second (spatial) derivative operators applied to $\delta\boldsymbol{m}$ (e.g., Inoue et al., 1990; Takeuchi et al., 2007).

In this dissertation, I formulate the inverse problem using the usual (not damped) least square eq. (12) for the case of the D″ region, and the regularized least-square eq. (16) for the case of the MTZ. The reason is that for the case of the MTZ, I use more model parameters of smaller lateral dimensions than for the case of D″ (and slightly less data). This requires regularization to obtain a smooth model of the MTZ without relatively strong small-scale

variations that are probably artefacts due to the fact that the kernel cannot resolve well each individual voxel separately.

2.1.2 Conjugate gradient

We solve eq. (21) using a truncated conjugate gradient method in which we truncate the expansion at the n^{th} conjugate gradient (CG) vector (see Kawai et al., 2014), with n chosen to minimize a modified version of the Akaike Information Criterion (Akaike, 1977).

The conjugate gradient method is an iterative method to solve a system of linear equations defined by a positive definite matrix, which is the case of the inverse problem eq. (21). For simplicity, we re-write eq. (21) as in eq. (22)

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{b}, \tag{22}$$

where $\boldsymbol{A}^T\boldsymbol{A}$ is a M×M matrix. The conjugate gradient method consists in finding the solution to eq. (22) iteratively using descent directions given by vectors $\boldsymbol{p}_i$ mutually conjugate to each other, which is the case if

$$\boldsymbol{p}_i^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{p}_j^T = 0, \forall i, j \in \{1, 2, .., M\} \tag{23}$$

The $\{\boldsymbol{p}_i\}$ form a base of $\mathbb{R}^M$, so that the solution $\boldsymbol{x}$ to eq. (22) can be written as a linear combination using this basis. The method starts with an initial vector $\boldsymbol{p}_0$ given by the gradient eq. (22), assuming that the initial solution $\boldsymbol{x}_0 = 0$,

$$\boldsymbol{p}_0 = \boldsymbol{A}^T\boldsymbol{b}, \tag{24}$$

and iterates over $k = 0, 1, .., N \leq M$, defining the next conjugate gradient vector $\boldsymbol{p}_k$ as close as possible to the residual $\boldsymbol{r}_k$

$$\boldsymbol{r}_k = \boldsymbol{A}^T\boldsymbol{b} - \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}_k \tag{25}$$

$$\boldsymbol{p}_k = \boldsymbol{r}_k - \sum_{i<k} \frac{\boldsymbol{p}_i^T\mathbf{A}^T\boldsymbol{A}\boldsymbol{r}_k}{\boldsymbol{p}_i^T\mathbf{A}^T\boldsymbol{A}\boldsymbol{p}_i}\boldsymbol{p}_i. \tag{26}$$

The solution at step $k + 1$ is then given by

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k \boldsymbol{p}_k, \tag{27}$$

where

$$\alpha_k = \frac{\boldsymbol{p}_k^T \boldsymbol{r}_k}{\boldsymbol{p}_k^T \boldsymbol{A}^T \boldsymbol{A} \boldsymbol{p}_k} \tag{28}$$

We note that in eqs. (25), (26), and (28), we do not have to compute the matrix multiplication $\mathbf{A}^{\mathrm{T}}\boldsymbol{A}$, but just the matrix-vector products $\boldsymbol{A}\boldsymbol{x}_k$, $\boldsymbol{A}\boldsymbol{p}_i$ and $\boldsymbol{A}\boldsymbol{r}_k$. Hence, the conjugate gradient method avoids the computation of $\mathbf{A}^{\mathrm{T}}\boldsymbol{A}$, which is computationally intensive when the number of parameters becomes large.

The models inferred in this study are for $n = 6$. A small number of vectors minimizes the solution error, and avoids overfitting the data. It does not, however, explicitly impose a smoothness condition, which is why we introduce regularization (eq. (14)).

2.1.3 Variance and AIC

For the regular (without damping) inverse problem (eq. (12)), the variance reduction (labeled VR) is defined as in eq. (29)

$$\begin{aligned} \text{VR}\ (\%) &= \frac{\left|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{s}-\boldsymbol{u})\right|^2}{\left|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{u}\right|^2} \cdot 100 = \frac{\left|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{A}\delta\boldsymbol{m}-\delta\boldsymbol{d})\right|^2}{\left|\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{u}\right|^2} \cdot 100 \\ &= (\delta\boldsymbol{m}^T\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}\delta\boldsymbol{m} - 2\delta\boldsymbol{m}^T\boldsymbol{W}\boldsymbol{A}^T\delta\boldsymbol{d} + \left|\boldsymbol{W}^{\frac{1}{2}}\delta\boldsymbol{d}\right|^2) \cdot 100 \end{aligned} \tag{29}$$

Using this definition, the VR for the data is 100 %.

For the case of the regularized least square (eq. (21)), the variance reduction (labeled VR′) is defined as in eq. (30)

$$\begin{aligned} \text{VR'}\ (\%) &= [\boldsymbol{x}^T(\boldsymbol{B}^T\boldsymbol{B}+\boldsymbol{I})\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{B}^T\boldsymbol{b} + |\boldsymbol{b}|^2] \cdot 100 \\ &= [\delta\boldsymbol{m}^T(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}+\boldsymbol{C}_{\boldsymbol{m}}^{-1})\delta\boldsymbol{m} - 2\delta\boldsymbol{m}^T\boldsymbol{W}\boldsymbol{A}^T\delta\boldsymbol{d} + \left|\boldsymbol{W}^{\frac{1}{2}}\delta\boldsymbol{d}\right|^2] \cdot 100 \\ &= \text{VR'} + \delta\boldsymbol{m}^T\boldsymbol{C}_{\boldsymbol{m}}^{-1}\delta\boldsymbol{m} \cdot 100. \end{aligned} \tag{30}$$

Since $\boldsymbol{C_m^{-1}}$ is positive definite, we note that VR' $>$ VR.

We also define the incremental variance reduction (labeled inc. VR) as the difference between the VR for synthetics for the initial 1-D model (AK135), and that for the inferred 3-D model

$$inc.VR = VR^{initial} - VR^{final}. \tag{31}$$

Using this definition, a positive inc. VR means that synthetics for the final model better fit the data than those for the initial model.

As mentioned in section 2.1.2, the conjugate gradient method used to solve the inverse problem is truncated to the first n conjugate gradient vectors based on a modified version of the AIC criterion (Akaike, 1977). The (modified) AIC_n^α for the first n CG vectors and the empirical redundancy parameter α is given by eq. (32)

$$AIC_n^\alpha = N^\alpha ln(2\pi) + N^\alpha ln(VR_n) + N^\alpha + 2(n+1), \tag{32}$$

with

$$N^\alpha = \frac{N}{\alpha} = \frac{1}{\alpha}\frac{T_{min}}{2} \cdot (\text{Number of data points at 1Hz sampling}), \tag{33}$$

where N is the minimum number of sampling points to represent the data vector $\delta \boldsymbol{d}$ (i.e., two points per minimum period).

2.2 Partial derivatives

I summarize the formulation of the partial derivatives, as given by Geller and Hara (1993). For the k^{th} earthquake and a receiver at position $\boldsymbol{r}^{(p)}$, the partial derivative of the i^{th} component (in locally Cartesian spherical coordinates) of the wavefield u_i for a perturbation in shear modulus at position $\boldsymbol{r}^{(l)}$ inside the Earth is given by eq. (34)

$$\frac{\partial u_i^{(k)}\left(\boldsymbol{r}^{(p)}\right)}{\partial \mu_l} = -\left[u_{j,q}^{(k)}\left(\boldsymbol{r}^{(l)}\right)\right]^*\left[C_{jqrs}^{(l)}\right]^* \eta_{ri,s}^{(p)}\left(\boldsymbol{r}^{(l)}\right), \tag{34}$$

where u_j is the j^{th} component of the forward-propagated wavefield from the earthquake source, η_{ri} is r^{th} component of the back-propagated wavefield for a single-force point source in the i^{th} direction at the receiver, and $C_{jqrs}^{(l)} = \frac{1}{2}(\delta_{jr}\delta_{qs} + \delta_{qr}\delta_{js})$ is the elastic tensor for a unit perturbation in the shear modulus. Repeating indices are summed, and derivatives with respect to the i^{th} coordinate are denoted by a comma $,i$. Since I use only the transverse component, we neglect the spheroidal components. In this case, the transverse component is given by u_3.

2.2.1 Three-point interpolation

I use the Lagrange polynomials to interpolate the wavefield from three points at the Earth's surface at a constant azimuth and epicentral distances $\theta_1, \theta_2,$ and θ_3. Given the three values of the wavefield $u_i \equiv u(\theta_i)$ $(i = 0, 1, 2)$ at the three epicentral distances θ_i, we can write a second-order polynomial $p(\theta)$ that takes the values u_i at θ_i using the Lagrange base

$$p(\theta) = u_0\varphi_0(\theta) + u_1\varphi_1(\theta) + u_2\varphi_2(\theta), \tag{35}$$

with

$$\begin{aligned}
\varphi_0(\theta) &= \frac{(\theta - \theta_1)(\theta - \theta_2)}{(\theta_0 - \theta_1)(\theta_0 - \theta_2)}, \\
\varphi_1(\theta) &= \frac{(\theta - \theta_0)(\theta - \theta_2)}{(\theta_1 - \theta_0)(\theta_1 - \theta_2)}, \\
\varphi_2(\theta) &= \frac{(\theta - \theta_0)(\theta - \theta_1)}{(\theta_2 - \theta_0)(\theta_2 - \theta_1)}.
\end{aligned} \tag{36}$$

In the case of three evenly spaced epicentral distance points $\theta_{i+1} - \theta_i \equiv h$, and defining $h_i = \theta - \theta_i$ $(i = 0, 1, 2)$ we can write

$$p(\theta) = \frac{1}{h^2}\left(h_1h_2u_0 - h_0h_2u_1 + \frac{h_0h_1}{2}u_2\right). \tag{37}$$

The error for the three-point interpolation decreases in $O(|h|^3)$.

2.2.2 Wavefield catalog

Significant computational and hard disk savings can be achieved by exploiting the fact that the initial model is spherically symmetric. Indeed, due to the spherical symmetry of the initial model, the wavefield for a given earthquake recorded at several receivers with the same epicentral distance but different azimuths can be computed only once for the given epicentral distance, and the azimuthal dependence can be computed analytically, as detailed below.

In the DSM and for the computation of the partial derivatives (Geller & Hara, 1993; Kawai et al., 2006), the wavefield is expended using vector spherical harmonics defined as in eq. (38)

$$\begin{aligned} \boldsymbol{S}^1_{lm}(\theta,\phi) &= (Y_{lm}(\theta,\phi),0,0), \\ \boldsymbol{S}^2_{lm}(\theta,\phi) &= \left(0,\frac{1}{\mathcal{L}}\frac{\partial Y_{lm}(\theta,\phi)}{\partial\theta},\frac{1}{\mathcal{L}\sin\theta}\frac{\partial Y_{lm}(\theta,\phi)}{\partial\phi}\right), \\ \boldsymbol{T}_{lm}(\theta,\phi) &= \left(0,\frac{1}{\mathcal{L}\sin\theta}\frac{\partial Y_{lm}(\theta,\phi)}{\partial\phi},-\frac{1}{\mathcal{L}}\frac{\partial Y_{lm}(\theta,\phi)}{\partial\theta}\right), \end{aligned} \tag{38}$$

with

$$Y_{lm}(\theta,\phi) = N_l^m P_l^m(\cos\theta)e^{im\phi}. \tag{39}$$

In eq. (39), N_l^m is a normalization factor, and $P_l^m(x)$ is the associated Legendre polynomial of angular order l and azimuthal order m. Synthetic seismograms are written using vector spherical harmonics, while the forward and backward propagated wavefields for the computation of the partial derivatives are written using vector spherical harmonics and their first order partial derivatives (in spherical coordinates), hence in terms of the second order derivatives of $Y_{lm}(\theta,\phi)$.

Therefore, the only dependence on the azimuth ϕ comes from the term $e^{im\phi}$. It is then possible to separate the dependence on ϕ during the computation of the synthetics and partial derivatives, and to write

$$\boldsymbol{u}(r,\theta,\phi) = \sum_{m=-2}^{2} \widetilde{\boldsymbol{u}}^{m}(r,\theta)e^{im\phi}, \tag{40}$$

with $\boldsymbol{u}$ being either the synthetics, forward propagated or back propagated wavefield, and $\widetilde{\boldsymbol{u}}^{m}$ the solution when ignoring the term $e^{im\phi}$ in the equations (i.e., $\boldsymbol{u}(r,\theta,\phi=0)$). We note that for the case of the back-propagated field, the excitation source is a single force, and $\widetilde{\boldsymbol{u}}^{m} = 0$ for $|m| = 2$. Furthermore, for the case of toroidal component (SH), $\widetilde{\boldsymbol{u}}^{0} = 0$.

It becomes then advantageous to pre-compute the epicentral distance-dependent part of the wavefield on a 1-D grid of epicentral distances and interpolate it for each receiver using three-point interpolation (see above). The azimuthal dependence can then be computed simply using eq. (40). In Figure 2, we verify the accuracy of the interpolation of the back-propagated wavefield. We compute the partial derivatives for the earthquake #1 in Table A1 (focal depth of 104.7 km) recorded at a station at epicentral distance of ~15˚ and for a perturbation point located at the turning point of the S-wave triplications turning inside the MTZ (at 412 km depth). In Figure 2a, we show the "exact" (i.e. computed without interpolation) partial derivative (top trace in black line) in a time-window of duration 200 s starting 20 s before the arrival of the direct S-wave. We also show the residual partial derivatives computed as the difference between the "exact" partial, and the partials computed using interpolation with a grid interval varying from $d\theta = 0.01^{\circ}$ to 0.5°. All the partial derivatives are filtered using the same bandpass filter as in the inversion (20–100 s). We observe that the error when using interpolation becomes significant for $d\theta > \sim 0.1^{\circ}$. Given that the velocity of the S-wave at depth ~412 km is ~5 km/s, the minimum local wavelength is ~100 km (~1˚). This suggests that the error when using interpolation is becoming significant when $d\theta$ becomes greater than

around one tenth of the minimum wavelength. In Figure 2b, I show the relative error (eq. (41)) of the partial derivatives as function of the epicentral distance grid interval $d\theta$.

$$error\ (\%) = \frac{|\boldsymbol{u} - \boldsymbol{u}_{d\theta}|}{|\boldsymbol{u}|} \cdot 100, \tag{41}$$

where $\boldsymbol{u}$ is the "exact" partial derivative (without using interpolation), and $u_{d\theta}$ is the interpolated partial derivative from a catalog with epicentral increment $d\theta$. We note that the error decreases as $\sim d\theta^{3.12}$, which agrees with the theoretical error estimate for three-points interpolation.

Figure 2 shows that a grid interval of 0.01 degrees is sufficient to obtain accurate synthetics (filtered between 20–200 s) after interpolation. We note that interpolation is not efficient for the synthetics and forward-propagated wavefield, since the wavefield catalog would have to be computed for each earthquake (because of the different source depths). However, interpolation can greatly reduce the computational time for the back-propagated wavefield, since the source is always at 0 km depth. Hence the back-propagated wavefield catalog has to be computed only once irrespectively of the number of stations. For the case of the MTZ, the range of epicentral distances between the source (station) and receiver (perturbation point) for the back-propagated wavefield is 0.01–85.00°. With a grid interval of 0.01 degrees, this means that the back-propagated wavefield catalog is computed for 8500 epicentral distance points.

The ratio of the total number of records of the back-propagated wavefield without using interpolation to the number of records in the catalog of back-propagated wavefield is thus

$$\frac{\#horizontal\ perturbation\ points\ \cdot \#stations}{8500} = \frac{1080 \cdot 1471}{8500} \approx 186.9 \tag{42}$$

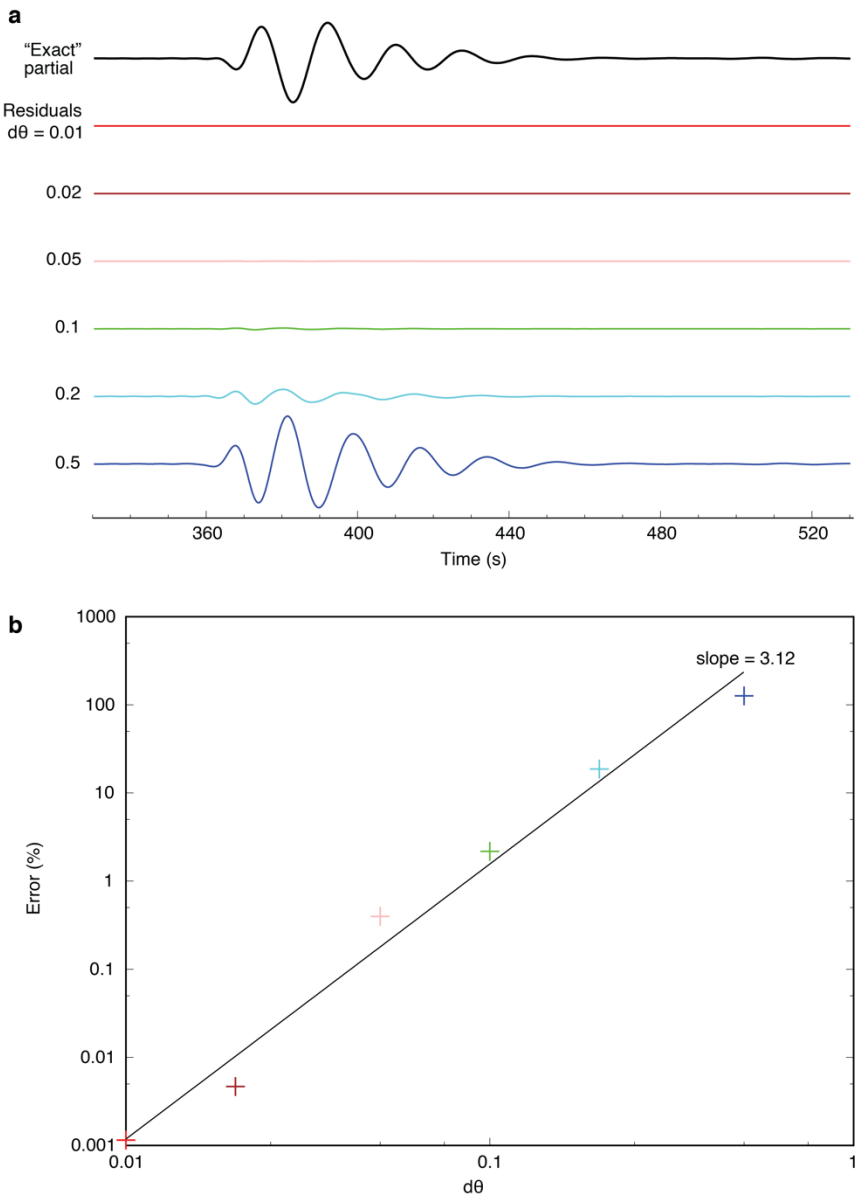

Figure 2. Interpolation error for the catalog of back-propagated wavefield. Interpolation error for a partial derivative computed at the deepest turning point of the S-wave for event #1 in Table A2 recorded at 15˚ of epicentral distance. a) "Exact" partial (top black trace; without using interpolation), and residuals (remaining colored traces), defined as the difference between the exact partial, and the interpolated partials for epicentral distance increments $d\theta =$ 0.01 to 0.5. b) Relative error (eq. (41)) for the residual traces shown in (a).

Since DSM scales linearly with the number of records, this means that computation of the back-propagated wavefield using interpolation is ~186.9 times more efficient than without using interpolation. The disk space requirement is also reduced, but by a factor 2 less (i.e. 93.5, for the case of the back-propagated SH field), since $\tilde{\boldsymbol{u}}^m$ in eq. (40) has to be stored for $m = -1, 1$. We note that the error due to interpolation when using a grid interval of 0.01 degrees is ~0.001%, which is comparable to the error in the DSM computations due to the truncation at a finite angular order in the spherical harmonics expansion. Figure 2 suggests that, at the cost of a slight loss in accuracy, a grid interval of up to 0.05 degrees is a possible choice (with an error of ~0.8 %), which would make the interpolation ~934.5 times more efficient (and the storage cost ~467.3 times less) than without interpolation, in this particular case.

2.2.3 Numerical accuracy of partial derivatives

In this work (and in the previous works of our group), the model parameters are constant-velocity blocks (voxels). The partial derivative with respect to the shear modulus μ for a particular block is computed at the center of the block (at position r_c), and multiplied by its volume V, which is an approximation for the exact partial derivative given by eq. (43)

$$\int_V d^3r \frac{\partial u}{\partial \mu}(r) \approx V \frac{\partial u}{\partial \mu}(r_c). \tag{43}$$

The improvements in computational efficiency of the partial derivatives due to the use of the wavefield catalog (section 2.2.2) make it easier to obtain more accurate partial derivative kernels by 1) computing the kernel on a finer mesh of voxel (than used in the actual inversion), and 2) summing adjacent voxels to obtain a more accurate evaluation of the integral in eq. (43).

Figure 3 and Figure 4 show kernels for the D″ region (of lateral dimension 5˚ x 5˚), and for the MTZ (of lateral dimensions 2˚ x 2˚), respectively, with three levels of accuracy. The first column (panels a) shows the original kernel for partial derivatives computed at the center

of each voxel, i.e. using eq. (43). The second (panels b), and third (panels c) columns, show more accurate kernels computed by increasing the number of lateral quadrature points within each voxel to compute the integral in eq. (43) (the number of mesh points in the vertical direction is kept constant): panels b) 4 points located at the center of voxels of lateral dimensions half the original voxel size; panels c) 8 points located at the center of voxels of lateral dimensions one fourth of the original voxel size. The differences between panels (a) to (c) come from variations in the kernel amplitude that are smaller in scale than the size of the voxels used in the inversion.

In this dissertation, the inferred models of the D″ region are obtained using a kernel computed on a 5˚ x 5˚ mesh using one point at the center of each voxel, as in Figure 3a, while the kernel for the MTZ is computed on a 2˚ x 2˚ mesh using four points in each voxel to obtain a more accurate kernel, as in Figure 4b. The effects of a more accurate kernel for the MTZ on the inferred model and on the nonlinear checkerboard test are discussed in details section 1.1, and show that 1) the amplitude of the inferred model using a more accurate kernel is larger than that using a less accurate kernel and better match the amplitude of the input checkerboard model; 2) when using the more accurate kernel, the region of the model near the sources is better resolved in the nonlinear checkerboard test, and is also most likely improved in the actual MTZ model since it better shows slabs in this region. For the case of D″, the effect of using a more accurate kernel is possibly an important topic for future work, but we expect that the results will not be significantly different than the ones presented in this dissertation (as for the case of the MTZ).

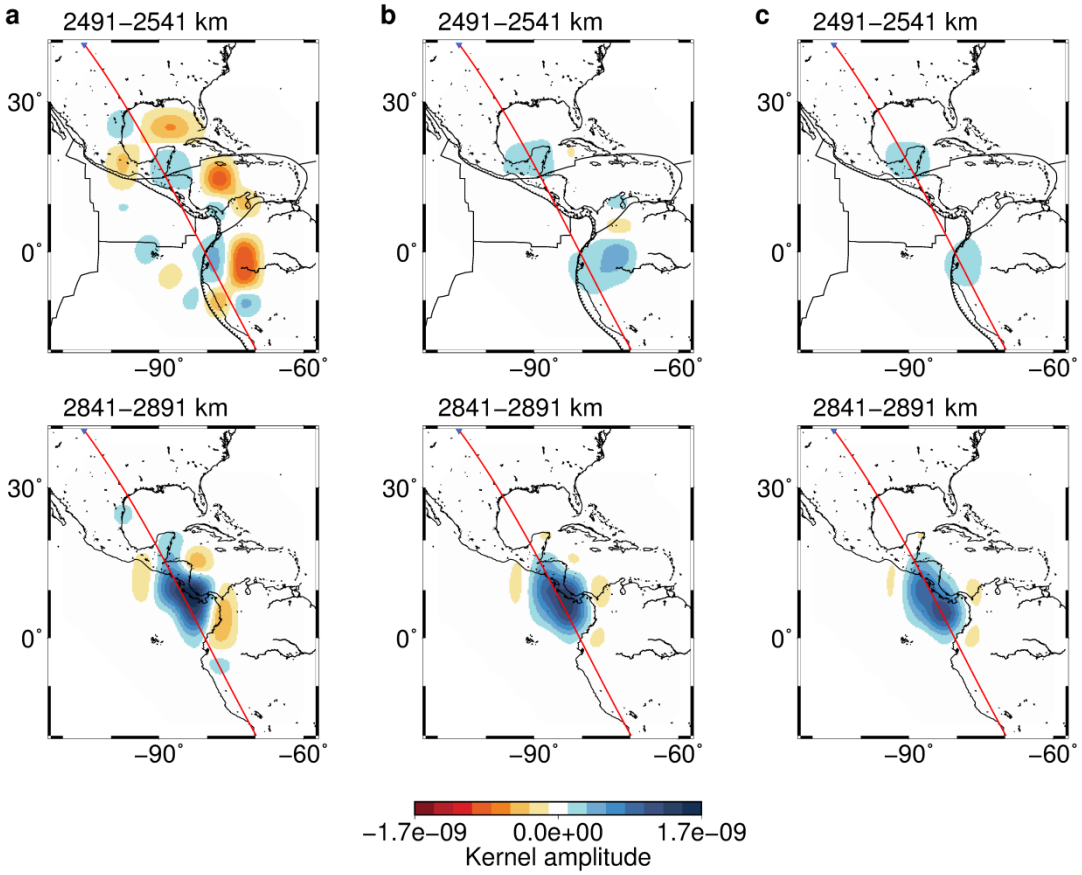

Figure 3. Accuracy of the kernel for D″. Kernels are for a mesh of voxels of lateral dimension 5˚ x 5˚ (as for the inversion for the D″ structure in Chapter 3), and filtered between 12.5–100 s. Columns from left to right show the kernel amplitude ~7 s after the arrival of the ScS phase at depths 2491–2541 km, and 2841–2891 km with increasing accuracy: a) original kernel computed on a 5˚ x 5˚ grid; b) more accurate kernel computed on a 2.5˚ x 2.5˚ grid, then summed to obtain a 5˚ x 5˚ grid; c) even more accurate kernel computed on 1.25˚ x 1.25˚ grid, then summed to obtain a 5˚ x 5˚ grid.

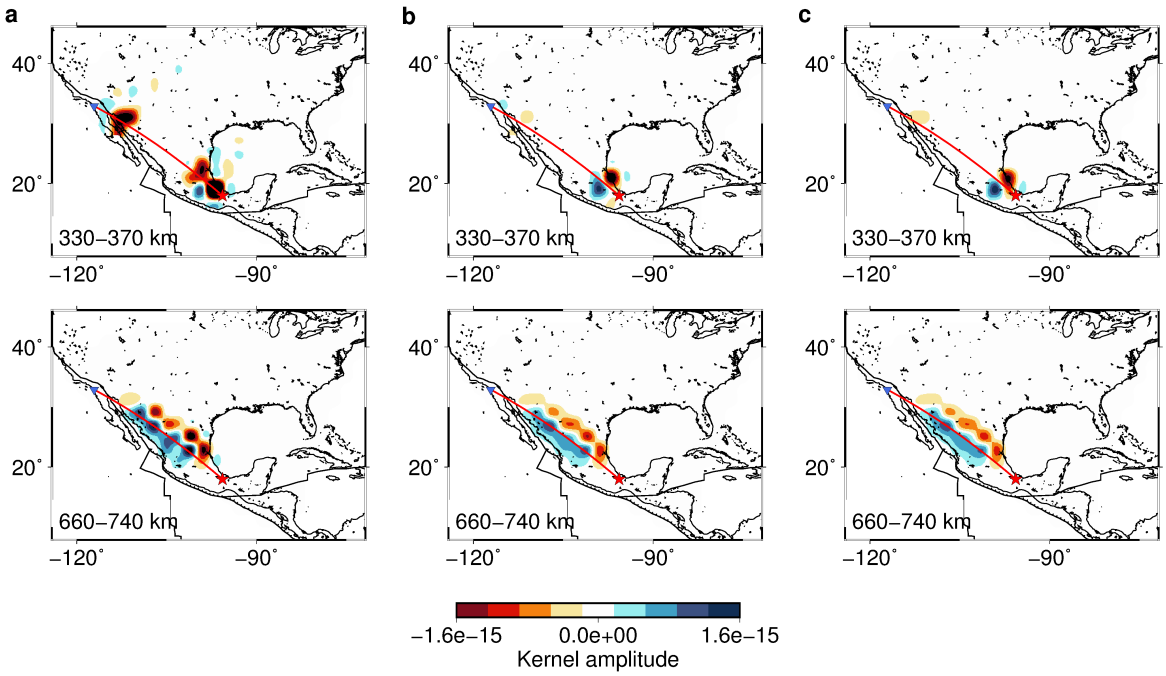

Figure 4. Accuracy of the kernel for the MTZ. Kernels are for a mesh of voxels of lateral dimension 2˚ x 2˚ (as in the inversion for the MTZ structure in Chapter 1), and filtered between 20–100 s. Columns from left to right show the kernel amplitude ~5 s after the arrival of the direct S wave at depths 330–370 km, and 660–740 km with increasing accuracy: a) original kernel computed on a 2˚ x 2˚ grid; b) more accurate kernel computed on a 1˚ x 1˚ grid, then summed to obtain a 2˚ x 2˚ grid; c) even more accurate kernel computed on 0.5˚ x 0.5˚ grid, then summed to obtain a 2˚ x 2˚ grid.

2.3 Data correction

It is well known from previous tomographic studies that the shallow structure (~0–300 km) is strongly heterogeneous, with S-velocity anomalies that can reach ~10%. For such strong velocity anomalies, it is not possible to approximate the wavefield as a linear perturbation to the wavefield for the initial 1-D model.

In this work, I conduct localized waveform inversion for the 3-D S-velocity structure of the D″ region (between 0–400 km above the CMB), and of the MTZ (between 330–820 km depth), without modeling the part of the Earth above the target region. It is therefore necessary to correct the data to account for the effect of the 3-D structure outside (i.e., above) D″, and the MTZ, respectively.

For the case of D″, I use the autopick method developed by Fuji et al. (2010), which uses the S phase as a reference phase to correct the data for the travel-time residual due to the structure close to the source and receiver (Figure 5).

For the case of the MTZ, it is not possible to use the direct S phase as a reference phase, since it samples the target region. Corrections for the shallow structure are computed by using previous tomographic models. Details of the computation, and tests of the effectiveness of these corrections are shown in section 4.3.2.

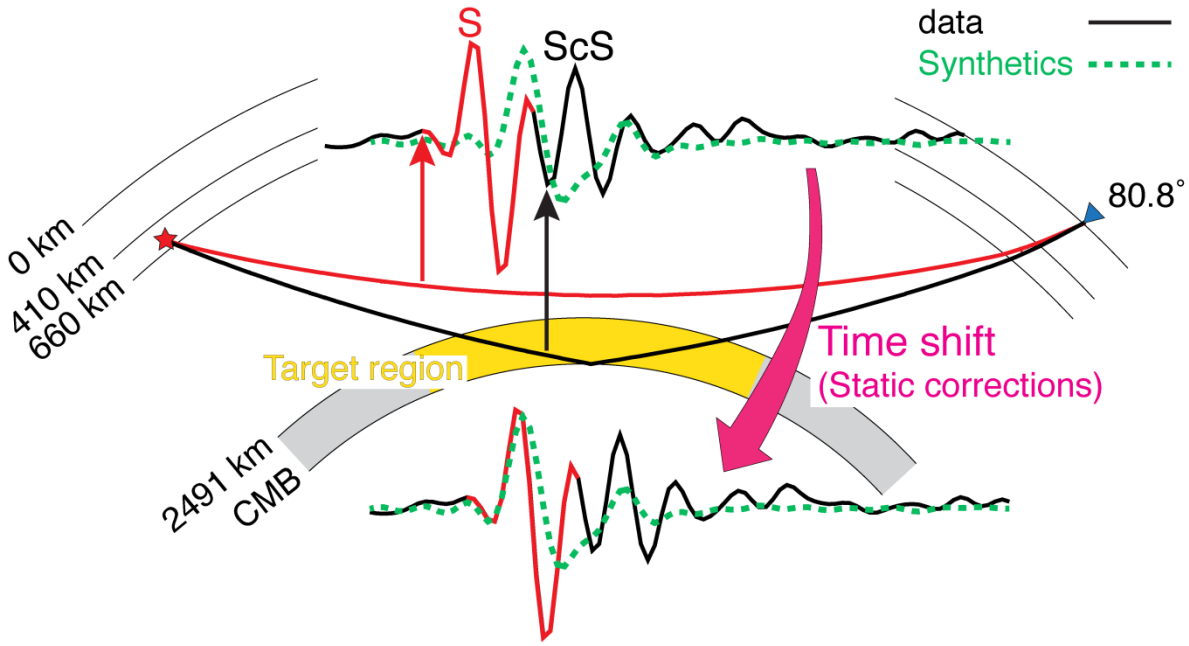

Figure 5. Corrections for the shallow structure for the case of D″. The direct S-wave (red line) turns above the target region in D″ (shown in yellow), and is used as a reference phase to correct the data for the 3-D heterogeneities near the source (red star) and the station (blue inversed triangle).

2.4 Validation of inferred models

As for any kind of seismic tomography, it is necessary to validate the models obtained using waveform inversion to check that the resolution of the partial derivative kernel is enough to resolve the smallest scale 3-D variations in the inferred model, and to check the robustness of inferred models with respect to possible source of errors due to 1) data errors, and 2) unmodeled properties (e.g., unmodeled part of the Earth in the case of local tomography).

2.4.1 Resolution tests

The (linear) checkerboard test is the most widespread test to check the resolving power of the kernel. It is performed by computing approximate input synthetics $\boldsymbol{s}^c$ for a 3-D checkerboard pattern of alternating high- and low-velocity anomalies by multiplying the kernel with the checkerboard pattern anomalies, as in eq. (44)

$$\boldsymbol{s}^c = \boldsymbol{s}^0 + \boldsymbol{A}\delta\boldsymbol{m}^c \Rightarrow \delta\boldsymbol{d} = \boldsymbol{s}^c - \boldsymbol{s}^0 = \boldsymbol{A}\delta\boldsymbol{m}^c, \tag{44}$$

and then performing an inversion using $\boldsymbol{s}^c$ as input data. Substituting eq. (44) in the inverse problem eq. (16) yields eq. (45)

$$(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A} + \boldsymbol{C}_m^{-1})\delta\boldsymbol{m} = \boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}\delta\boldsymbol{m}^c. \tag{45}$$

If $\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}$ is invertible and if smoothing constraints are not imposed (i.e., $\boldsymbol{C}_m^{-1} = 0$, as for the case of the D″ layer in this work), then the solution of eq. (45) is exactly $\delta\boldsymbol{m}^c$ (i.e., the checkerboard pattern is resolved perfectly). However, since we use only the first n CG vectors to solve eq. (45) (as in the actual inversion), the truncated solution will be different from $\delta\boldsymbol{m}^c$, and will show how well can the first n CG vectors resolve the checkerboard anomalies. Similarly, if $\boldsymbol{C}_m^{-1} \neq 0$, the solution of eq. (45) will be different from $\delta\boldsymbol{m}^c$, which allows us to test how the resolution changes when imposing smoothing constraints. We note that since the CG vectors depend also on the residual vector $\delta\boldsymbol{d}$, using the same number of CG vectors for

the checkerboard test and for the actual inversion does not ensure that the nominal resolution is the same. A discussion on this topic is given for the case of the MTZ in section 1.1.

The checkerboard test has been criticized by some authors, mostly because it does not check the nonlinearity of eq. (5) that arises because of the relatively strong deviations between the initial model and the inferred 3-D model (Rawlinson & Spakman, 2016). Nonlinearity can, however, be checked by computing exact synthetics for the input checkerboard pattern using a 3-D wave propagation code such as SPECFEM3D GLOBE (Komatitsch et al., 2015). We call this test "nonlinear checkerboard test". Although the linear checkerboard test is widely used in seismic tomography, its nonlinear version is not common. One of the reasons why nonlinear checkerboard tests are usually not performed is that they require the computation of exact synthetics for the 3-D structure, which is computationally intensive. Also, for tomographic studies that perform several iteration for the actual inversion (French & Romanowicz, 2014), a similar number of iterations would have to be performed for the nonlinear checkerboard test, which means that the nonlinear checkerboard test has the same computational cost as the actual inversion. We perform the nonlinear checkerboard test for the case of the D″ layers (section 3.4) and of the MTZ (section 1.1), and show that the linear approximation is reasonable for these two cases (for perturbations of strength 3.5%, and ~2–4%, respectively).

A computationally less intensive version of the nonlinear checkerboard test can be obtained by computing synthetics for an input layered 1-D structure, which we call "synthetic block test". This test also allows to easily check the improvement of the fit to the input synthetic data for the recovered model. Since the travel-time anomaly caused by such block pattern is nearly zero, this test also allows to show the usefulness of using amplitude information to add constraints to the inferred structure (as shown for the case of the D″ layer in section 3.4).

The fourth and last type of resolution test performed in this work is a point-spread function test. The point-spread function test is similar to the checkerboard test, with the only difference that the input synthetics are computed for a point anomaly (i.e., only one voxel is perturbed) rather than for checkerboard pattern anomalies. This test allows to quantify the (nominal) amount of spreading (or smearing) from a point anomaly to other voxels in the model, and is equivalent to computing a particular column of the resolution matrix, defined as in eq. (46)

$$\delta \boldsymbol{m}^{est} = \boldsymbol{R}\delta \boldsymbol{m}^{true}, \tag{46}$$

where $\delta \boldsymbol{m}^{est}$ and $\delta \boldsymbol{m}^{true}$ are the inferred, and actual model perturbations, respectively. Indeed, for instance if the point anomaly corresponds to the j^{th} voxel in the vector of model parameters, then $\delta \boldsymbol{m}^{true} = \hat{\boldsymbol{e}}_{\boldsymbol{j}}$, which implies that $\delta \boldsymbol{m}^{est} = \boldsymbol{R}\hat{\boldsymbol{e}}_{\boldsymbol{j}} = \boldsymbol{R}_j$, the j^{th} column of $\boldsymbol{R}$.

We note that, as discussed by Rawlinson and Spakman (2016), these kind of resolution tests obtained by solving the inverse problem for a synthetic input are currently widely used, even though current computers allow for the computation of the full resolution matrix for most of the current tomographic models. However, even the full resolution matrix does not tell us about the "actual" resolution of the tomographic inversion, because the actual inverse problem is non-linear (e.g., Romanowicz, 1991). Resolution tests using synthetic input can, on the other hand, test the validity of linearizing the inverse problem if the synthetic input is computed by solving the (non-linear) forward problem (as discussed above).

2.4.2 Robustness tests

Even when the resolution is good, artifacts in the inferred model can arise because of errors in the data (e.g., noise, misorientation of the seismometer), or in unmodeled structure or source parameters (e.g., hypocenter, moment tensor, source-time function). In this work, we test the robustness of inferred models for errors in the unmodeled structure above the target region (see

section 2.3), which we believe is the major source of possible artifacts in the inferred models, and for the effect of (gaussian) noise in the data (see section 3.4). We do not check the errors in the model due to possibly inaccurate source parameters. However, a recent study showed that waveform inversion for the structure of D″ is most likely not so sensitive to errors in the source parameters (Yamaya et al., 2018), as long as the minimum period used in the inversion is larger, or comparable to the duration of the source-time function of the earthquakes used (i.e., that only earthquake with relatively small magnitude are used), which is the case for this work.

The robustness tests performed for the case of the D″ layer and MTZ are described in details in sections 3.4, and 1.1, respectively, so that we will only enumerates them here. We tested: 1) the orthogonality between partial derivatives in the target region and partial derivatives near the sources, 2) a joint inversion for the structure of D″ and the shallow structure (0–350 km depth) to explicitly show that the structure in D″ is nearly not affected by the shallow structure, 3) the changes in the model when inverting with 50% of the data randomly picked among the full dataset (jackknife test), 4) the effect of gaussian noise on the inferred model by performing a checkerboard test with noise added to the input synthetics, 5) the dependence of the inferred model on the choice of the initial model, 6) the tradeoff between velocity anomalies in the MTZ and errors in the depths of the 410 and 660 km discontinuities.

Chapter 3 Imaging paleoslabs in the D″ layer beneath Central America and the Caribbean using waveform inversion

3.1 Introduction

The purpose of this chapter is to obtain high resolution 3-D images of the S-velocity structure in the D″ layer beneath Central America and the Caribbean to search for evidence of paleoslabs above the CMB and for evidence of small-scale low-velocity anomalies that might suggest chemical heterogeneity (Garnero & McNamara, 2008).

Recently, dense seismic arrays like the USArray, which includes many portable stations that have steadily been moved eastward to cover the entire conterminous area of the U.S., are providing excellent data for high-resolution imaging of localized regions of D″ using waveform inversion. We recently conducted two small-scale feasibility tests of this method to invert for 3-D S-velocity structure in the D″ layer beneath central America (Kawai et al., 2014) and the western Pacific (Konishi et al., 2014), and applied the same method to a much larger dataset to invert for the 3-D S-velocity structure in D″ beneath the Northern Pacific (Suzuki et al., 2016). In the present study, I use the full USArray to obtain dense coverage of the D″ layer beneath central America and the Caribbean. The use of short-period (up to 8 s) waveforms makes it possible to image small-scale structure with finer resolution than travel-time tomography or global waveform inversion studies.

The D″ layer at the base of the mantle is, after the Earth's crust and uppermost mantle, the second most seismically laterally heterogeneous region of the Earth's mantle (Ritsema et al., 2011; Tkalčić et al., 2015). Strong large low-velocity provinces (LLVPs) beneath Africa and the South Pacific, and high-velocity regions beneath the circum-Pacific, are the large-scale

features found ubiquitously by travel-time tomography or global waveform inversion studies (French & Romanowicz, 2014; Ritsema et al., 2011).

Low seismic velocity regions in the lowermost mantle can be explained by high temperatures, chemically distinct material, or a combination of the two. Pyrolite is widely thought to be the average composition of the lower mantle (Wentzcovitch et al., 2004; Wang et al., 2015), but the details of the bulk composition of the lower mantle remain controversial (Ringwood, 1989; Cottaar et al., 2014). Chemical compositions with increased amounts of impurities such as Fe and Al have lower shear velocities than pyrolite (Tsuchiya, 2011; Zhang et al., 2016). Such chemical heterogeneity, resulting from exchanges between the core and the mantle, from partial melting in the TBL, from basalt entrained to the base of the mantle by past subduction, or as long-lived remnants of chemical differentiation in the early Earth, is expected at the CMB (Garnero & McNamara, 2008). However, the extent to which chemical anomalies contribute to the lowermost mantle seismic structure is still unclear; the LLVPs, which could possibly be large-scale chemically distinct regions, have been variously suggested to be due to temperature anomalies only (Kawai & Tsuchiya, 2009; Schuberth & Bunge, 2009; Davies et al., 2012; Konishi et al., 2014), or to be chemically distinct from the rest of the lower mantle (Frost & Rost, 2014; Garnero et al., 2016). High-velocity anomalies in the lowermost mantle can be explained by the combined effect of low temperatures and the bridgmanite (abbreviated as MgPv below) to Mg-post-perovskite (abbreviated as MgPPv below) phase transition (Murakami et al., 2004; Tsuchiya et al., 2004) (see Hirose et al., 2017, for a recent review on post-perovskite).

It is generally thought that high-velocity anomalies inferred from seismic tomography in the upper mantle and in the upper part of the lower mantle correspond to remnants of past

subduction (Bunge & Grand, 2000; Fukao & Obayashi, 2013). Seismic tomography shows high-velocity anomalies continuous from the upper mantle down to the 660 km discontinuity, where the negative Clapeyron slope of spinel decomposition and a jump in viscosity make some slabs stagnate, while others can be seen to descend to the mid-mantle (~1800 km depth). In the lowermost ~500 km of the mantle, strong and broad high-velocity anomalies are seen beneath slabs that clearly extend to the mid-mantle (Grand, 2002; Ritsema et al., 2011; French & Romanowicz, 2014).

Evidence supporting the existence of paleoslabs at the CMB includes reports of continuous high-velocity anomalies from the transition zone down to the CMB beneath North America (Farallon slab) and beneath the southern Indian Ocean (Simmons et al., 2015) and the fact that high-velocity anomalies in the lowermost mantle are generally consistent with the location of ancient slabs (Bunge et al., 1998). Estimates of the global average sinking rate of slabs in the lower mantle based on seismic tomography vary from 1.1 cm/yr to 1.9 cm/yr (van der Meer et al., 2010; Domeier et al., 2016). Subducted material older than 160-260 Ma might thus have reached the CMB. Slabs at the CMB are strongly heated and probably disintegrate in ~100 Myr (Tackley, 2011); thus material older than ~260-360 Ma is unlikely to cause high-velocity anomalies at the CMB.

An increase in the reported amplitudes of the high-velocity anomalies in cold regions in the lowermost mantle, and thus in the visibility of hypothetical paleoslabs, seems most likely to be due to the MgPv to MgPPv phase transition (Murakami et al., 2004; Tsuchiya et al., 2004). The detailed structure of broad high-velocity anomalies in the lowermost ~500 km of the mantle cannot be resolved in current tomographic models because of their coarse parametrization in the lowermost mantle. It thus was heretofore not possible to say whether

high-velocity anomalies near the CMB result from spreading of accumulated slabs or from separate paleoslabs that followed different paths to the lower mantle, possibly originating at different subduction zones.

3.1.1 Study area

The western margin of North and South America (previously the western margin of the supercontinent Pangaea) is a region of long-lived subduction, where the oceanic Farallon plate is believed to have initiated eastward subduction around 180-207 Ma (van der Meer et al., 2010); tectonic studies suggest subduction of the Pacific Ocean beneath the west coast of Pangaea since ~250 Ma (van der Lelij et al., 2016). The Farallon plate is seen in tomographic models as an eastward dipping high-velocity feature reaching ~2000-2500 km depth (Bunge & Grand, 2000; Ren et al., 2007; van der Meer et al., 2010), and its location at those depths is in general agreement with the reconstructed plate boundary ~180 Ma (Müller et al., 2016).

Continuity of the Farallon plate (as an eastward dipping high-velocity anomaly) from ~2000-2500 km depth to the CMB beneath Central America is not a strong feature of previous models. Previous models show, however, a broad high-velocity anomaly beneath, and westward of the location of the Farallon slab at ~2000 km depth. Tearing and breaking of the slab at the 660 km discontinuity or around ~1000 km depth could explain the apparent discontinuity of the Farallon slab (Fukao & Obayashi, 2013). This may be supported by geological evidence for voluminous igneous activity ~200 Ma (van der Lelij et al., 2016). High-velocity anomalies deeper than ~2500 km depth thus might correspond to subduction older than 200 Ma at the western margin of Pangaea. In addition to subduction at the western margin of Pangea, tectonic studies suggest there were subduction zones within the Pacific Ocean, where the ocean floor subducted beneath active volcanic arcs (Johnston & Borel, 2007). Such

intra-oceanic subduction (Leat & Larter, 2003) has been suggested as a possible explanation of the high-velocity seismic structure in the deep mantle beneath the western margin of North America (Sigloch & Mihalynuk, 2013), and in the lowermost mantle beneath Central America (van der Meer et al., 2010).

3.2 Materials and methods

The dataset used in this chapter consists of ~13,000 transverse component records of ground velocity at epicentral distances $70^\circ < \Delta < 100^\circ$ from 40 deep- and intermediate-focus South American earthquakes (see Table A1) in the period 2004-2015, recorded at broadband stations of the transportable and backbone arrays of the USArray. We augment this dataset by ~1500 records from a total of 80 South American earthquakes in the period 1993-2015 (also listed in Table A1) recorded at CNSN, CANOE, IRIS/USGS, SCSN, PNSN, and BDSN networks (Figure 6).

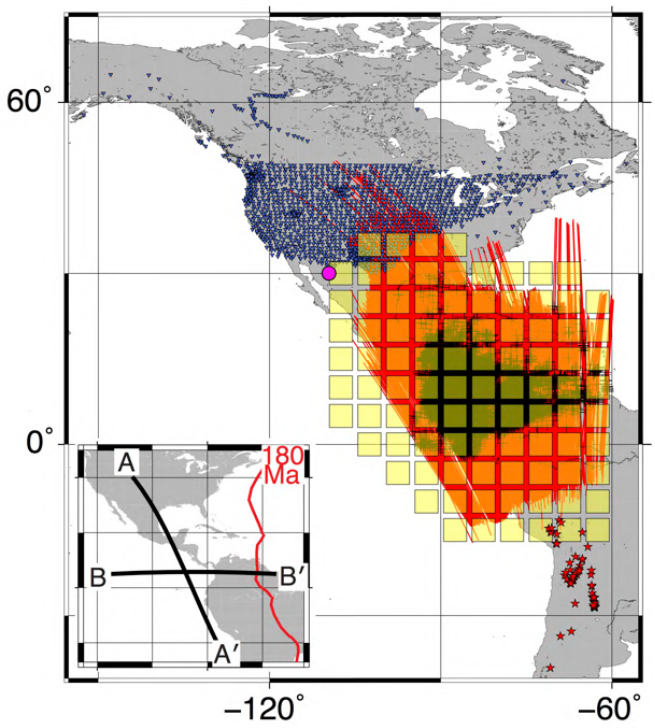

Figure 6. Target region. Waveforms from deep- and intermediate-focus earthquakes beneath South America (red stars) recorded at USArray, CNSN and CANOE seismic stations (blue inverted triangles) provide dense raypath coverage of the target region 0-400 km above the CMB (yellow squares). Red curves show ScS raypaths which sample the target region, and black crosses show ScS bounce points at the CMB. The pink solid circle at 30˚N, 110˚W shows the location for the shallow structure tradeoff test (Figure 14). The inset shows the location of the cross-sections presented in Figure 11, and the location of the Farallon plate boundary at 180 Ma (van der Meer et al., 2010).

The records used in the inversion are selected so that 1) the amplitude ratio between the observed record and the corresponding synthetics is less than 3 and greater than 0.33, and 2) the variance of the residual (i.e., the variance of the difference between the data and synthetics) is less than 300%. After selection, 7768 and 7654 records were used in the inversions for models CACAR and CACAR′, respectively. The large number of records used in this study contributes to the stability of the inversion.

Data are filtered between 8–200 s using a Butterworth bandpass filter. I cut each trace 20 s before the arrival of the direct *S* phase and 60 s after the arrival of the ScS phase. S and ScS arrivals are computed using the TauP Toolkit (Crotwell et al., 1999). ScS precursors (e.g., Scd) and post-cursors (e.g., Sbc), that are sensitive to sharp velocity contrasts in the D″ layer (Borgeaud et al., 2016), are included in the waveforms in the dataset used in this chapter. I weight each cut residual trace (observed trace minus synthetic trace) so that its maximum amplitudes are all equal to unity. I then apply a second weighting factor to the residuals to partially correct for the uneven azimuthal- and epicentral-distance distribution of the stations in the dataset used in this chapter. This second weighting factor is obtained by minimizing eq. (47) using the least-square method.

$$\sum_{i=1}^{N}(h_i - h)^2 + \lambda(h_i - 1)^2 \tag{47}$$

where i is the index of the bin, N is the number of bins, h_i is the number of records in bin i, and $\bar{h}$ is the average value of h_i. I use the damped least-squares and constrain the sum of the weighting coefficients to be equal to the number of histogram bins. Histograms for the original, and weighted dataset are shown in Figure 7**.**

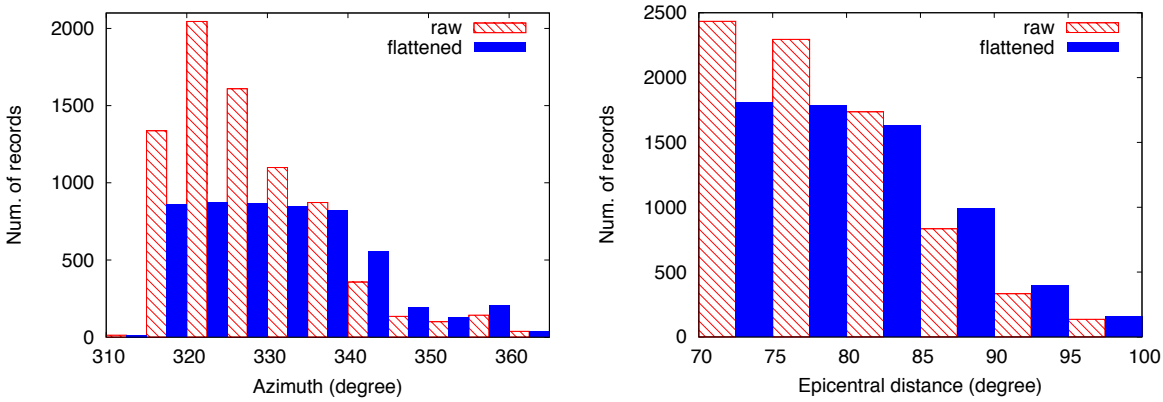

Figure 7. Data statistics. Distributions of azimuths (a) and epicentral distances (b) for the records used for the inversion. The original distribution is shown in red, and the weighted distribution is shown in blue. The records are weighted to homogenize the distribution of azimuths and epicentral distances.

Each record is also corrected for the effect of 3-D structure outside the target region by a time-shift that aligns the onset of the direct *S* phase on the record with the onset of the direct S phase on the corresponding synthetic (Fuji et al., 2010). I discard the records for which the time-shift is greater than 10 s. Although this does not correct for propagation effects due to strong heterogeneity in the upper mantle, using both S and ScS makes this dataset primarily sensitive to the lowermost mantle. In order to verify it, I conduct four different tests shown in Figure 14 to Figure 16. Figure 14 shows that the partial derivative kernel for structure in the target region is nearly completely independent from that for structure in the shallow mantle. Figure 15 shows that simultaneous inversion for the shallow upper mantle and D″ S-velocity structures 1) left the D″ essentially unchanged compared to inversion for D″ only; 2) inversion for the shallow upper mantle S-velocity structure only yields small amplitude S-velocity perturbations (~0.6%) and small variance reduction (~1%) compared to the inversion for the D″ model only (~5%). Figure 16 shows that a strong 10% velocity decrease in the depth range 24.4–220 km can be nearly completely corrected for using the above-mentioned time-shifts. Finally, Figure 17 shows that three inverted models using three different datasets selected by dividing the stations into western, central, and eastern regions, where there are different upper mantle structures (Burdick et al., 2016), are in good general agreements in regions where there is common raypath coverage.

I use methods recently developed for localized 3-D waveform inversion (Kawai et al., 2014; Konishi et al., 2014). Waveform inversion uses all the information in the waveforms, i.e., not only travel times of identified phases, but also their shape and amplitude. Since waveform inversion compares observed to synthetic waveforms directly, phases do not have to be identified individually; i.e., overlapping phases can be used. This allows the use of ScS when

it partially overlaps with S at $\Delta > 85^{\circ}$. The improvements in resolving power provided by the additional use of shape and amplitude information, and overlapping phases is discussed below.

I compute partial derivatives with respect to a spherically symmetric (1-D) initial model using the Born approximation, which gives the first-order perturbation to the wave equation. The 1-D synthetics are computed using the Direct Solution Method (DSM) (Kawai et al., 2006). Using the Born approximation with respect to a 1-D initial model significantly reduces the computation time thus allowing the use of a large dataset of relatively short period waveforms (down to 8 s).

Every inversion method uses some smoothing or regularization techniques. In this chapter, I truncate the conjugate gradient (CG) expansion for both CACAR and CACAR′ at six CG vectors based in the criterion of minimizing Akaike's Information Criterion (AIC) (see section 2.1).

In this study, I fixed the earthquake source parameters to the GCMT catalog. Later, Yamaya et al. (2018) showed that redetermination of the source parameters does not affect significantly the 3-D models obtained by the inversion. However, they found systematically smaller source time-function duration than reported in the GCMT catalog for intermediate depth and deep earthquakes, which might affect the results of waveform inversion when using either higher-frequencies, or a dataset dominated by large moment magnitude events.

3.3 Results

I conduct waveform inversion for the detailed 3-D S-velocity structure of the lowermost 400 km of the mantle beneath Central America and the Caribbean using ~7,700 transverse component records cut 20 s before the arrival of the direct S-wave, and 60 s after the arrival of the ScS phase (see section 3.2). The events are deep and intermediate focus events recorded at

epicentral distances $70^\circ < \Delta < 100^\circ$ at broadband seismic stations of the USArray, Canadian Northwest Experiment (CANOE), IRIS/USGS, SCSN, PNSN, and BDSN networks (Figure 6). The azimuthal- and epicentral-distance distribution of the stations is shown in Figure 7. The data are filtered in the period range 8-200 s using a Butterworth bandpass filter. The 3-D model is obtained by linearized inversion with respect to a spherically symmetric initial model.

I first conduct 1-D waveform inversion (Kawai et al., 2007; Konishi et al., 2009), with respect to the global reference model PREM (Dziewonski & Anderson, 1981), to infer a regional 1-D model (hereafter called PREM′) of the lowermost mantle beneath Central America (Figure 8). This model was derived by optimizing the fit of the synthetics to the full dataset over a wide range of azimuths, and thus differs somewhat from previously published 1-D models of the lowermost mantle beneath Central America that used a significantly smaller number of waveforms from a narrower range of azimuths (Kendall & Nangini, 1996; Kawai et al., 2007).

The 3-D model is parametrized with 744 3-D cells (voxels) of dimensions $5^\circ \times 5^\circ \times 50$ km (equivalent to approximately $300 \times 300 \times 50$ km at the CMB) in 8 horizontal layers of 50 km thickness, from the CMB to 400 km above the CMB (Figure 6). To study the dependence of the 3-D models on the initial 1-D model, I conducted two inversions, with respect to PREM and PREM′, respectively. We call the resulting 3-D models CACAR (an abbreviation of Central America-Caribbean) and CACAR′, respectively. Map views of these 3-D models (with the average perturbation in each layer set to zero) are shown in Figure 9, and Figure 10, respectively. The two 3-D models are in good general agreement.

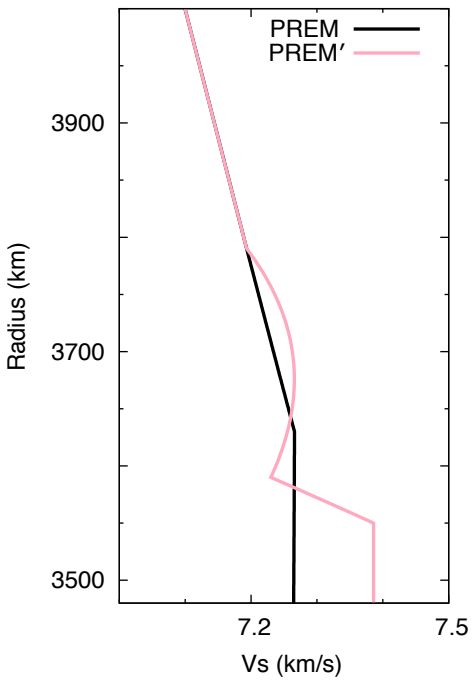

Figure 8. Initial models for 3-D inversions. I used two different 1-D models as initial models in the inversion for 3-D structure. PREM is a global reference model (Dziewonski & Anderson, 1981), and PREM′ is a local model we inferred for the lowermost 400 km of the mantle beneath Central America using 1-D waveform inversion (Kawai et al., 2007; Konishi et al., 2009) with the same dataset as for 3-D inversion, but with weighting that emphasizes raypaths beneath Central America (as compared to those beneath the Caribbean).

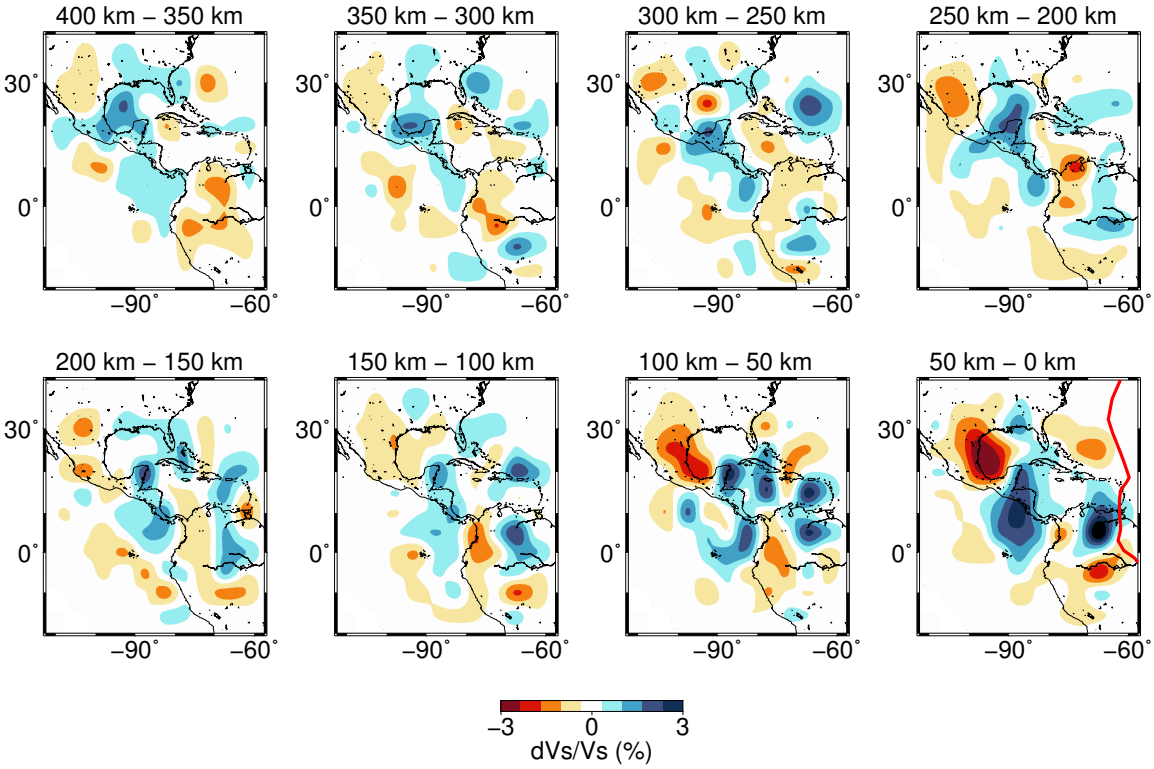

Figure 9. Model CACAR. The eight panels show the results of the inversion for the eight depth layers from 400 km above the CMB (upper left) to the CMB (lower right), with the lateral average of the 3-D perturbation set to zero in each layer. The perturbation is relative to the initial 1-D model PREM (Dziewonski & Anderson, 1981). Two distinct high-velocity regions at the CMB (lower right panel) suggest two distinct cold paleoslabs. A 3% velocity decrease beneath Mexico concentrated within 100 km of the CMB suggests the possible existence of chemically distinct material with enriched iron content (e.g., basaltic composition). Location of high- and low-velocity anomalies is generally consistent with recently inferred topography of the D″ discontinuity (Whittaker et al., 2015). The Farallon plate boundary at 180 Ma (Müller et al., 2016) is shown in red in the lower right panel. Its location is consistent with the high-velocity anomaly beneath Venezuela, but is ~1000 km away from the western high-velocity anomaly. This might indicate past intra-oceanic subduction, or breaking or tearing of an ancient paleoslab in the upper mantle (Fukao & Obayashi, 2013).

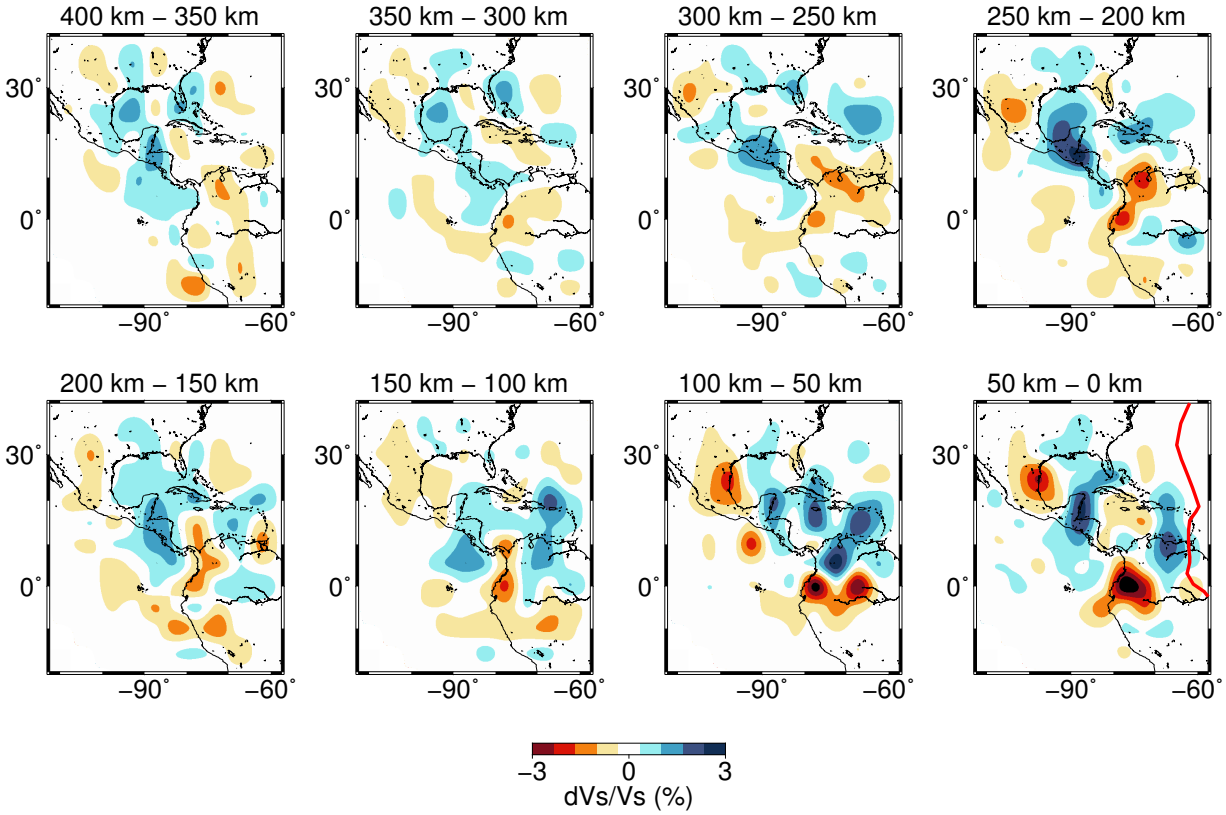

Figure 10. Model CACAR′. Same as Figure 9, but using PREM′ (Figure 8) as the initial 1-D model. The velocity perturbations are shown with respect to PREM′, with the lateral average of the 3-D perturbation set to zero in each layer. CACAR (Figure 9) and CACAR′ are in good general agreement, which suggests that the inversion is robust. The strong low-velocity anomaly beneath Mexico observed in CACAR is still present, but the strongest low-velocity anomaly is beneath Ecuador. This suggests that strong low-velocity anomalies are present at several sites around the high-velocity anomalies.

The key features of my models are (1) two distinct high-velocity anomalies just above the CMB, one beneath Central America and another beneath Venezuela, (2) strong low-velocity anomalies concentrated in the lowermost 100 km of the mantle at the edge of the high-velocity anomalies, and (3) vertically continuous low-velocity structures above the strong low-velocity anomalies at the CMB. The past Farallon plate boundary at 180 Ma (Müller et al., 2016), indicated with a red line in the lowermost right panel (0-50 km above the CMB) in Figure 9 and Figure 10, is consistent with the location of the high-velocity anomaly beneath Venezuela, while it is ~1000 km from the high-velocity anomaly beneath Central America. In the depth range 0-50 km above the CMB, the high-velocity beneath Venezuela in CACAR does not extend to the north along the past Farallon plate boundary (Figure 9). The same high-velocity beneath Venezuela in CACAR′ extends up to a few degrees north of the Caribbean Islands, but a strong high-velocity anomaly is not present to the north of the Caribbean Islands (Figure 10). However, as the altitude above the CMB increases, the high-velocity beneath Venezuela gradually extends along the past Farallon plate boundary to the north of the Caribbean Islands, and is strongest 250-300 km above the CMB (Figure 9 and Figure 10).

3.3.1 Cross-sections and comparison to previous studies

Several works in the previous decade considered this study area using finite-frequency travel-time tomography with a model parametrization with a scale similar to the one presented in this chapter (Hung et al., 2005), or migration of phases refracted by the D″ discontinuity (Thomas et al., 2004; Hutko et al., 2006), respectively; a more recent study conducted forward modeling for the topography of the D″ discontinuity beneath Central America and the Caribbean (Whittaker et al., 2015). The models presented in this chapter can also be compared to the structure in this region obtained by global waveform inversion (French & Romanowicz,

2014). From the above studies, I select the finite-frequency travel-time tomography study (Hung et al., 2005) and the global waveform inversion study (French & Romanowicz, 2014) for a detailed comparison to my models, as shown in Figure 11.

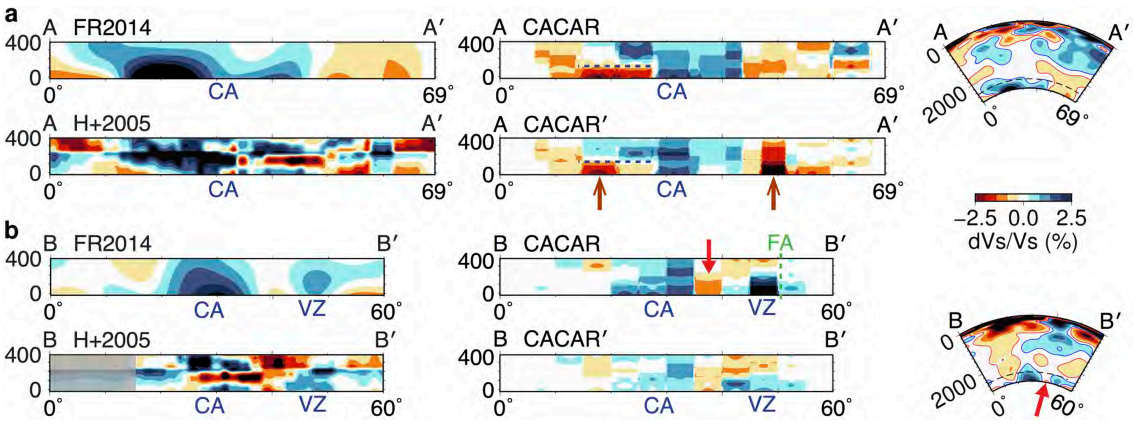

Figure 11. Cross-sections. (left column) Cross-sections through two previous models obtained by global waveform inversion (French & Romanowicz, 2014) (labeled FR2014), and by regional finite-frequency travel-time tomography (Hung et al., 2005) (labeled H+2005). (middle column) Cross-sections for my models CACAR and CACAR ′. (right column) whole-mantle cross-section through the global model FR2014. The horizontal axis of each profile shows degrees along the corresponding great circle cross-section; vertical axis shows elevation above the CMB (in km) for profiles in left and middle columns, and depth (in km) for the whole-mantle profiles in the right column. Locations of the cross-sections are shown in the inset in Figure 6. The leftmost part of the cross-section for model H+2005 in the left column of panel b has been grayed out due to lack of resolution. (Note that CACAR and CACAR′ also have little resolution in this region). (a) The strong velocity contrasts within 100 km above the CMB in CACAR and CACAR′ (blue dashed line in upper middle panel) suggest the presence of paleoslabs at the CMB and dense chemical heterogeneities. The paleoslab beneath Central America (labelled CA) is perched above a strong low-velocity anomaly (blue dashed line in upper middle panel), which might suggest dense iron-enriched material at the CMB (e.g., basaltic composition). Low-velocity vertically continuous structures (brown arrows with unfilled points in cross-section A-A′ in middle column) at the edges of CA suggest upwelling from the possibly iron-enriched material at the CMB. (b) A low-velocity vertically continuous structure (red filled arrow in cross-section B-B′ in middle column) between two distinct slabs beneath Central America (CA) and Venezuela (VZ), and connecting to a low-velocity region in the mid-mantle in the model FR2014 (right column) suggests upsplashing of hot TBL material caused by two paleoslabs CA and VZ sinking to the CMB. Past location of the Farallon plate boundary at ~180 Ma (van der Meer et al., 2010) (green vertical dashed line labelled FA) is consistent with location of VZ, but is laterally ~1000 km away from CA.

Cross-sections of the various models along two profiles (see inset of Figure 6 for locations) are shown in Figure 11. Models CACAR and CACAR′ are shown in the middle column of Figure 11. The left column of Figure 11 shows the models along these cross-sections obtained by the recent global waveform inversion (French & Romanowicz, 2014), labeled FR2014, and the models from the finite-frequency travel-time tomography study (Hung et al., 2005), labeled H+2005. Whole-mantle cross-sections through the global waveform inversion model (French & Romanowicz, 2014) are shown in the right column of Figure 11.

Figure 11a shows a northwest (A) to southeast (A′) great-circle cross-section through the high-velocity anomaly beneath Central America. As shown in the inset to Figure 6, this cross-section is roughly parallel to the reconstructed Farallon plate boundary at 180 Ma (van der Meer et al., 2010). Models CACAR and CACAR′ are shown along this cross-section in the middle column of Figure 11a. The two models differ slightly, but we find the following consistent features in both models. (1) Low-velocity anomalies are seen at both edges of the high-velocity anomaly (beneath Central America, labeled CA), marked by the brown arrows with unfilled points beneath the second panel in the middle column. (2) The low velocity anomalies are particularly strong in the lowermost 100 km of the mantle, where ~3% low-velocity anomalies are present. (3) The northern part of the high-velocity anomaly CA is perched above a 100 km thick low-velocity anomaly just above the CMB, resulting in a strong velocity contrast (5-6 % peak-to-peak over a small vertical range at the dashed blue line in the upper two panels in the middle column of Figure 11a).

Figure 11b shows a west (B) to east (B′) cross-section at 5˚N, roughly perpendicular to the past Farallon plate boundary at 160-180 Ma (van der Meer et al., 2010). For this particular cross-section CACAR′ shows anomalies with smaller amplitudes than CACAR, but the pattern

of high- and low-velocity anomalies for the two models is consistent. We see two high-velocity structures that correspond to the high-velocity anomalies beneath Central America (labeled CA) and Venezuela (labeled VZ), respectively. CA and VZ are separated by a vertically continuous low-velocity region (indicated by the red arrow).

I now compare the models inferred in this chapter to the recent global waveform inversion (French & Romanowicz, 2014), labeled FR2014, and the finite-frequency travel-time tomography study (Hung et al., 2005), labeled H+2005. The high-velocity anomaly CA is also present in the two previous models (left column of Figure 11a). The strong low-velocity regions on both edges of CA (brown arrows with unfilled points beneath the second panel in the middle column of Figure 11a) are present in the previous models (left column of Figure 11a), but are smaller in size and weaker by ~1.5%. As a result, the fact that in this region CA is perched ~100 km above a strong low-velocity zone at the CMB (dashed blue line in the upper two panels in the middle column of Figure 11a) is not evident from the previous models. We also note in passing that the high-velocity anomaly CA in Figure 11a is generally consistent with previous regional forward modeling studies (Thomas et al., 2004; Hutko et al., 2006; Whittaker et al., 2015). The low-velocity region to the right (south) of CA in the middle column of Figure 11a (marked by the brown arrow with unfilled points) connects to a larger low velocity region that was found to extend to the mid-mantle by a recent global waveform inversion study (French & Romanowicz, 2014) (right column of Figure 11a). The high-velocity VZ in the middle column of Figure 11b is covered by low-velocity material at shallower depths in the global model (right column of Figure 11b). As noted above, the location of the high-velocity anomaly beneath Venezuela is in good agreement with the past Farallon plate boundary at 180

Ma (indicated by the vertical green dashed line labeled FA), while the high-velocity zone beneath Central America (CA) is ~1000 km distant from it.

The high-velocity anomalies CA and VZ are also visible in the previous regional and global models (left column of Figure 11b). The amplitude of VZ in the previous models is, however, somewhat weaker than that of CA; the improvement in resolution of the high-velocity anomaly CA in my model is most likely due much greater number of records than the previous regional study (Hung et al., 2005), made possible by the new data from the dense transportable USArray. Also, waveform inversion allows us to also use the phase and amplitude information in the data, which most likely explains why CACAR and CACAR′ show more consistent vertical features. Part of the discrepancy between my model and the previous regional study might also be due to the latter's use of core phases (SKS), which can be affected by anisotropy, and uncertainty in the outer-core velocity structure and thus might be affected by artifacts.

The location of high- and low-velocities anomalies in CACAR and CACAR′ generally agrees with variations in the elevation of the D″ discontinuity reported on the basis of forward modeling of recent USArray data (Whittaker et al., 2015). In particular, the forward modeling study reports 1) an elevated D″ discontinuity beneath Venezuela, where we observe the high-velocity anomaly VZ, 2) a deeper D″ discontinuity in a corridor along Ecuador and Columbia that extends in the Caribbean Sea, where we see the low-velocity anomaly corridor separating VZ from CA, and 3) the highest elevation of the D″ discontinuity above the CMB where we observe the high-velocity anomaly CA. A previous study also observed this correlation (Sidorin, 1999), but CACAR and CACAR′ have finer resolution.

3.4 Resolution and robustness

The improvement in the fit of the synthetics for the 3-D models to the data is shown in Table 1. For the initial model PREM, the fit of the initial synthetics to the data (144.4%) is reduced to 79.7% after time-shifting the data to correct for 3-D structures above the target region in D″. Inversion of the corrected data yields synthetics for the final 3-D model CACAR with a fit of 74.5%, improving by 5.2% the fit of the initial synthetics for PREM. Similar fit improvements are realized for CACAR′ (Table 1). We note that the relatively small variance reduction is partly due to the fact that it is mainly the fit of the part of the waveforms around the smaller amplitude ScS phase (rather than the larger amplitude S phase) that is improved by the inversion.

The similarity between CACAR and CACAR′ suggests that the method and dataset used give robust results. In addition, I conduct various tests to evaluate and confirm the robustness and resolution of CACAR and CACAR′. I conduct two checkerboard tests showing that I can resolve an input checkerboard pattern of horizontal scale 300 x 300 km, and that the addition of artificial Gaussian noise to the input synthetics does not affect the ability to recover the input checkerboard pattern (Figure 12 and Figure 13).

Strong amplitude shallow structure (especially on the receiver side, since we use intermediate and deep events) can add scattered phases in the records we use for the inversion, and we do not want the inversion for D″ to be affected by these extraneous phases. I verify that the partial derivative kernel for structure in the target region is nearly completely independent of the partial derivative kernel for shallow (0-500 km depth) structure near the receivers. The results of this test (Figure 14) show that when I use the entire dataset, there is essentially no tradeoff between shallow structure and structure in the target region.

Model	**Variance (%)**	**AIC ($\alpha = 500$)**
PREM	144.4	-
PREM with time-shifts	79.7	4231.6
CG6	74.5	4133.6
PREM′ with time-shifts	79.7	3937.0
CG6′	74.9	3855.3

Table 1. Variance and AIC. Variance (eq. (29)) of the synthetics for PREM, PREM with static corrections (time-shifts), and the final model CACAR and CACAR′ with six conjugate gradient vectors (CG6). CG6 minimizes the Akaike information criterion (AIC) computed with an empirical redundancy parameter of $\alpha = 500$ (eq. (32)). Note that PREM′ does not improve the variance of the initial synthetics because the weighting used for the 3-D inversion is different from the weighting that was used for inferring PREM′ (i.e., PREM′ has a better fit than PREM for raypaths beneath Central America, but a worse fit than PREM for raypaths beneath the Caribbean).

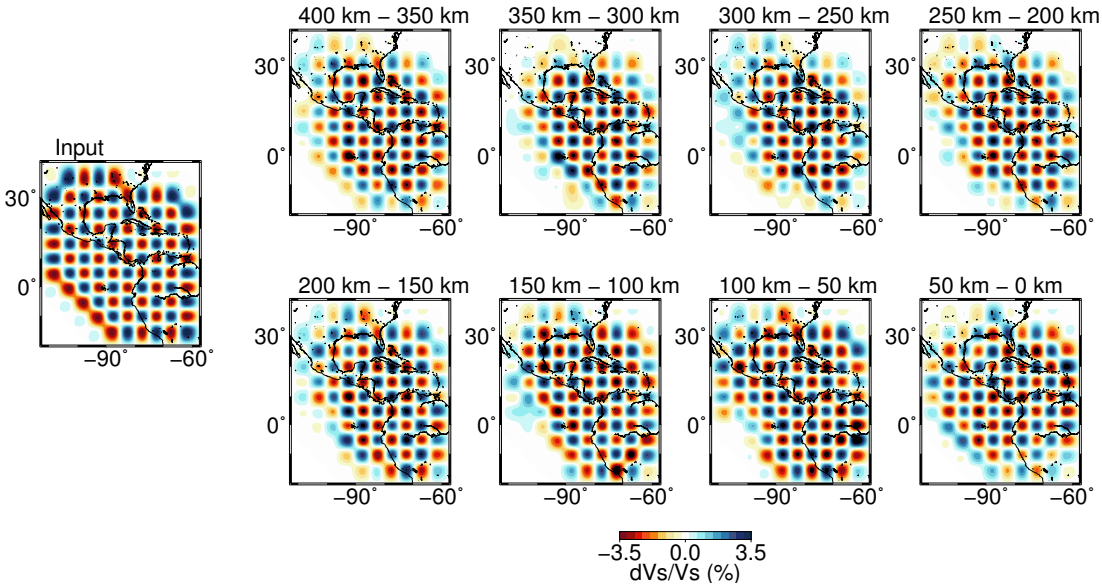

Figure 12. Checkerboard test. The input pattern for the layer 400-350 km above the CMB is shown on the left. The input pattern for the next layer (350-300 km) has the sign of all anomalies reversed, and so on for each successive layer. The dataset and the number of CG vectors for the checkerboard test is the same as for the actual inversion. This result shows that we can nominally resolve almost all voxels in the target region.

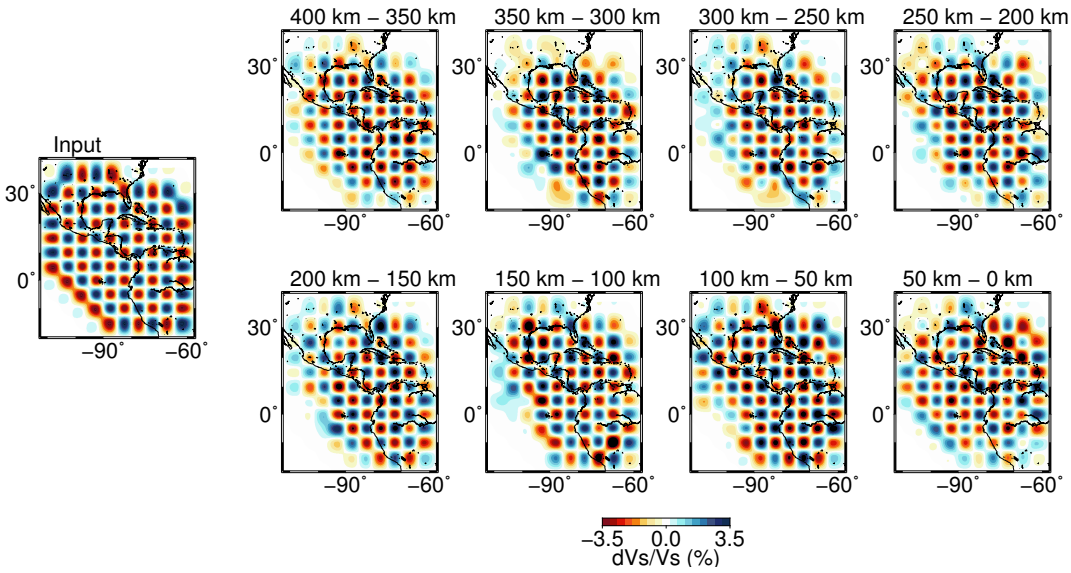

Figure 13. Checkerboard test with artificial noise. Same as in Figure 12, but with artificial Gaussian noise added to the 3-D synthetics for the checkerboard input. The noise increases the variance of the 3-D synthetics by 40%. The results are nearly identical to the checkerboard test without noise added, which suggest that the inversion is stable.

As a further check of the possibility of contamination of the inversion results due to the effect of structure outside the target region, I also perform a simultaneous inversion for the 3-D S-velocity structure in D″ (0—400 km above the CMB) and in the shallow upper mantle (25–375 km depth). For the simultaneous inversion, I apply two constant weights to the upper mantle voxels, and lowermost mantle voxels, respectively, so that the total sensitivity of those two regions is the same. The results (Figure 15) show that the S-velocity in D″ obtained 1) by joint inversion with the shallow upper mantle structure (Figure 15a) and 2) by inversion in D″ only (Figure 15b) are in nearly perfect agreement. This indicates that the inverted structure in D″ is nearly unaffected by the addition of model parameters for structure in the upper mantle, supporting the result of Figure 14 that the structure in D″ can be inverted for independently from shallow structure. Figure 15c shows the upper mantle structure obtained by simultaneous inversion with the D″ structure. The shallow upper mantle structure shows small amplitude perturbations (~0.6%), but with low-velocity anomalies in the western US, and high-velocity anomalies in the eastern US, which is generally consistent with the pattern observed by tomographic models of this region (Burdick et al., 2016). Figure 15d shows that the results of an inversion for upper mantle structure only are in good agreement with Figure 15c, which was obtained from the joint inversion. Figure 15e shows the variance reduction. The variance reduction for inversion for the shallow upper mantle structure only is small (~1%) compared to that for inversion for D″ structure only (~5%). The variance for the joint inversion is ~6%. The small variance reduction for the shallow upper mantle structure and the small amplitudes of perturbations in the inferred upper mantle structure suggest that the dataset used in this chapter does not contain much information about the shallow upper mantle structure. I thus conclude that the inversion for structure in D″ is robust.

In addition, I show in Figure 16 that the time-shift applied in the inversion using the S phase as a reference is probably sufficient to correct for strong velocity perturbations in the shallow upper mantle. For this purpose, I conduct a synthetic test using an input model with a 10% velocity decrease (with respect to PREM) in the depth range 24.4–220 km, which is otherwise identical to PREM in the rest of the mantle. As in the actual inversion, I use the S phase as reference to time-shift the input synthetics. The recovered D″ model (Figure 16) shows a maximum perturbation of 0.5% in the layer 0-50 km above the CMB, and an average perturbation within the high-velocity anomalies of less than 0.1%, which is negligible compared to the ~3% perturbations in CACAR and CACAR′.

Tomographic models of the upper mantle beneath the US (Burdick et al., 2016) show strong S-velocity variations, which might affect the inversion results. Although for the inversion in D″ we correct for the travel-time difference between synthetics and observed records, we do not correct for possible waveform distortions. In order to further test the robustness of the inversion results with respect to the upper mantle structure beneath the US, I conduct three inversions dividing the stations in the dataset into western, central, and eastern regions (Figure 17). Figure 17a shows color maps of the travel-time residuals of the direct S-wave and the amplitude ratio between the observed records and synthetics plotted at each station, and averaged in 1˚ x 1˚ bins. The western, central, and eastern datasets (labeled b, c, and d, respectively) show significant differences in travel-time and amplitude ratios, suggesting possible waveform distortion. We note that comparison with tomographic models (Burdick et al., 2016) suggests that the travel-time difference of the direct S-wave is mostly due to the upper mantle structure. The results of the inversions for the three datasets b-d are shown in Figure 17(b)-(d), respectively. Since the raypath coverage of the lowermost mantle differs for

the three datasets, the three S-velocity models in panels (b)-(d) resolve different regions of the lowermost mantle, but show good general agreement in regions where more than one model has significant raypath coverage. In particular, we note that the high-velocity anomaly CA is resolved in both models (b) and (c). This (together with Figure 14 to Figure 16) further suggests that the perturbations resolved in CACAR and CACAR′ are indeed due to the structure of the lowermost mantle, and not due to shallow structure effects.

Since the Born approximation is only strictly accurate for infinitesimal perturbations, I test the ability to recover synthetic input models with spherically symmetric perturbations (hereafter referred to as block tests) of $\pm$1% and $\pm$2% in layers 100 km- and 200 km-thick (~1.7 and 3.4 times the wavelength for 8 s waveforms); the input models are reasonably well recovered (Figure 18). The block tests also show that 1) I can resolve a 2% low-velocity anomaly 0-100 km above the CMB topped by a 2% high-velocity anomaly 100-200 km above the CMB (Figure 18c), which suggests that the strong vertical velocity contrast beneath Mexico in CACAR and CACAR′ is not an artifact; 2) the fact that the strongest high-velocity anomalies in CACAR and CACAR′ are located 0-100 km above the CMB is probably real and not an artifact of increase in sensitivity of the partial derivatives just above the CMB (Figure 18a, b); 3) CACAR might not constrain accurately the absolute amplitude of the perturbation, which might be over-estimated (Figure 18a, b). We note, however, that in Figure 18c the recovered amplitude matches that of the input model. For this case, the radially averaged perturbation in the lowermost 400 km of the mantle of the input model, and thus the S-ScS differential travel time (for $70^\circ < \Delta < \sim 85^\circ$), is nearly the same as for the initial model (PREM). This is illustrated in Figure 19, in which we show a profile of stacked waveforms for the input model of Figure 18c.

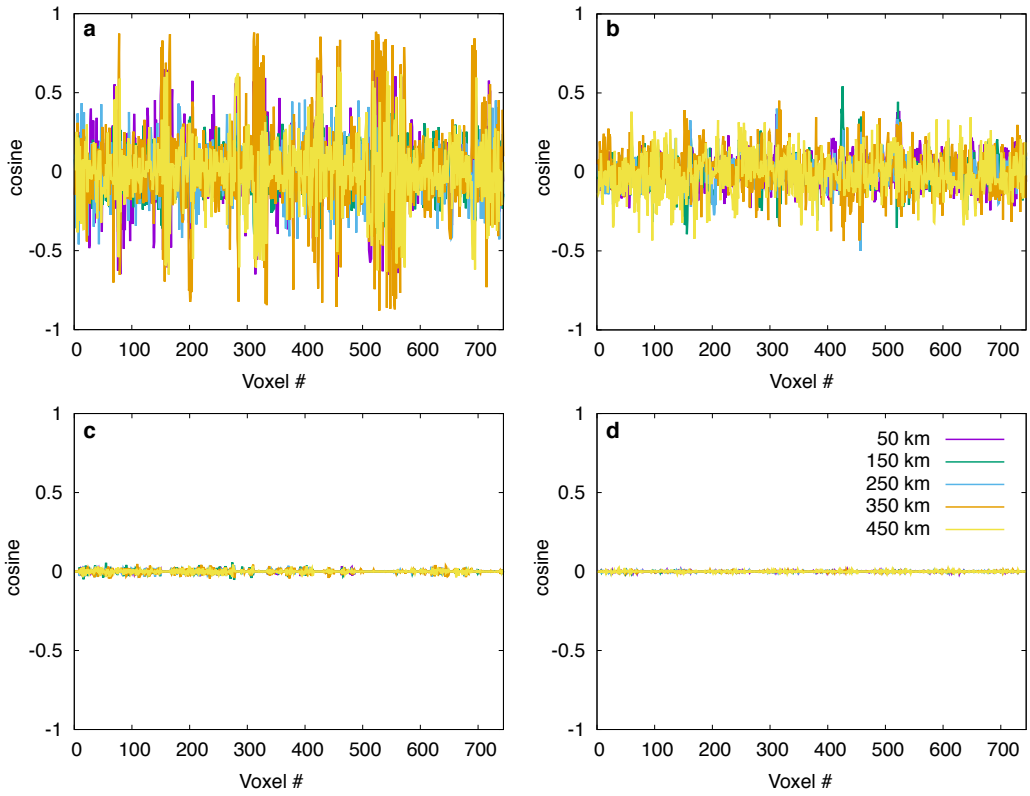

Figure 14. Tradeoff with shallow structure. Cosine of the angle between the vector of partial derivatives for each of the 744 voxels in the target region (D″) and the vector of partial derivatives for each of the 5 voxels in the upper mantle at 30˚N (see Figure 6), 110˚W and 50, 150, 250, 350, and 450 km depth. We also computed the cosine for 7 other horizontal location of the voxels in the upper mantle beneath the USArray to verify that the result does not depend on the location of the upper mantle voxels. The legend for each panel is the same as the one shown in panel d). The four panels indicate the increasing number of records used in the partial derivative vectors: a) one station, one event (1 record); b) one station, all events (4 records); c) all stations, one event (381 records); d) all stations, all events (7375 records).

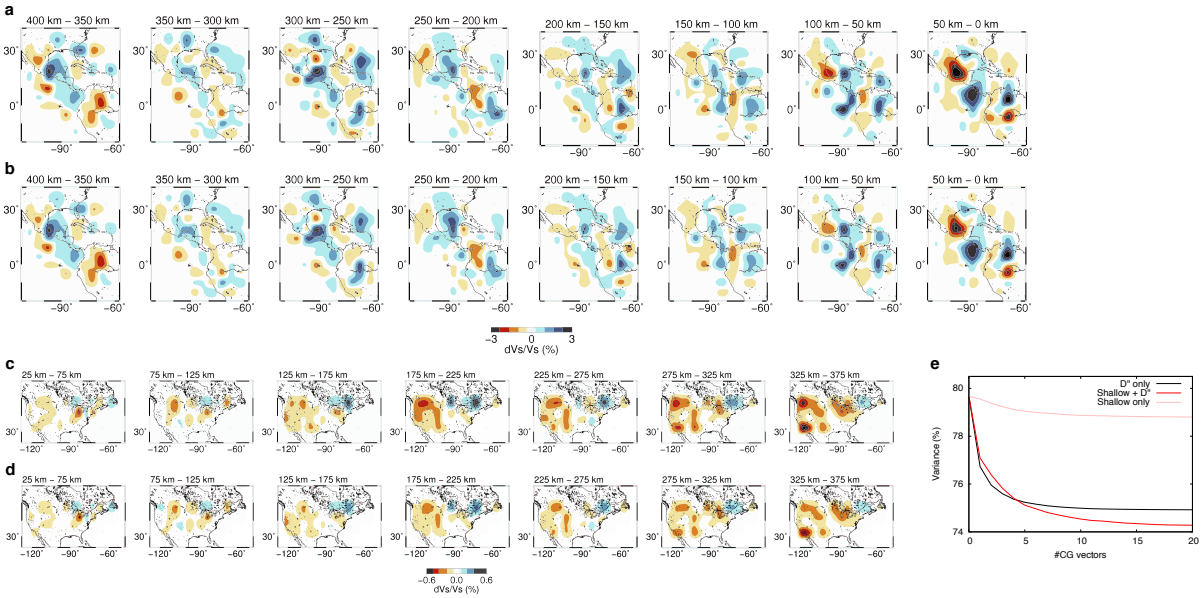

Figure 15. Simultaneous inversion for D″ and shallow upper mantle structures. (a) S-velocity in D″ from simultaneous inversion for the two depth ranges 25–375 km (shallow upper mantle), and 0–400 km above the CMB; the corresponding shallow upper mantle S-velocity structure is shown in panel (c). (b) S-velocity in D″ from inversion for D″ only. (d) S-velocity in the shallow upper mantle from inversion in the shallow upper mantle only. (e) Variance (in percent) as function of the number of model parameters (CG vectors) for the models in panels a-d. Note that the color scale for the shallow upper mantle S-velocity models (panels c and d) has a maximum amplitude of 0.6%, which is much smaller than the amplitudes of the perturbations for the D″ S-velocity models (3% in panels a and b). The variance reduction for the shallow upper mantle structure (pink curve) is small (~1%) compared to that for the D″ structure, suggesting that the dataset used in this chapter is primarily sensitive to D″. The S-velocity structures for the simultaneous inversion (panel a) and the inversion in D″ only (panel b) are nearly identical, suggesting that the dataset is robust with respect to the shallow upper mantle structure. The S-velocity models of the shallow upper mantle (panels c and d) show small amplitude perturbations, suggesting that the dataset of S and ScS waveforms has a small sensitivity to the upper mantle structure (in agreement with the small variance reduction noted above). The models in panels (a) and (b) are highly similar, as are the models in panels (c) and (d), further confirming the conclusion, based on Figure 14, that by taking a very large number of waveforms, the tradeoff between structure in the target region and in the upper mantle is effectively eliminated.

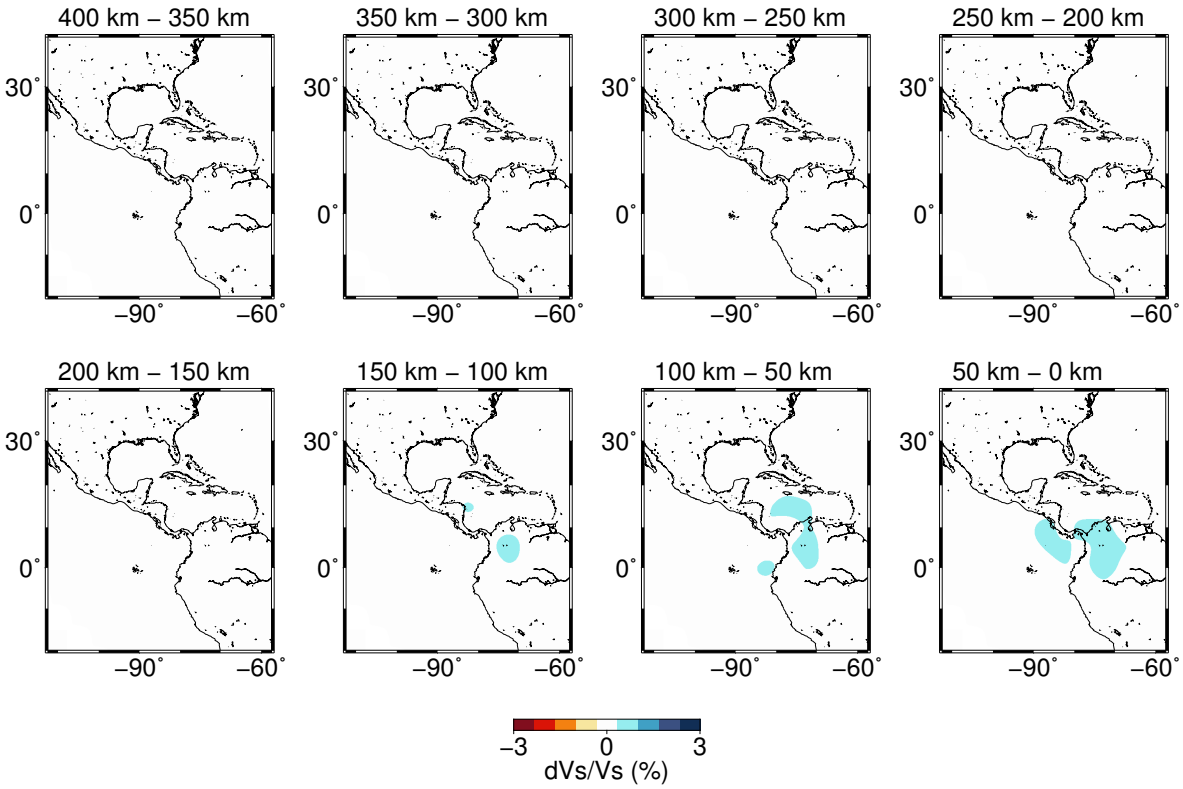

Figure 16. Possibility of contamination due to shallow structure. I invert for input synthetics for a model identical to PREM, except between 24.4 to 220 km depth, where the velocity was reduced by 10%, to test for the contamination of the D″ model by un-modelled shallow structures. As for the actual inversion, we time-shift the input synthetics to correct for the effect of the shallow structure. The recovered model within the target region shows a maximum perturbation of ~0.5%, and an average perturbation within the high-velocity regions of less than 0.1%, suggesting that the time-shift we apply is sufficient to remove the effects of un-modelled shallow structure.

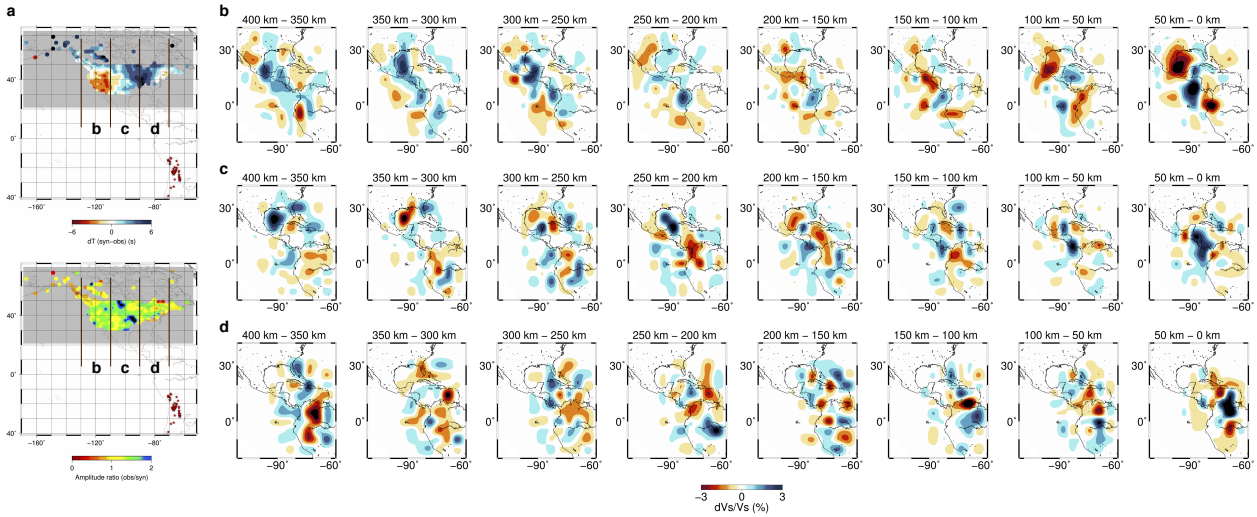

Figure 17. Inversion of western, central, and eastern datasets. (a) Color map of the travel-time residual of the direct S-wave (top), and of the amplitude ratio between observed records and synthetics, from averaging values at each station in 1˚ x 1˚ bins (bottom). Gray color means no stations. We divide the stations into western, central, and eastern datasets, labelled b, c, and d, respectively. (b-d) The three S-velocity models 0–400 km above the CMB are obtained from inversion of the three datasets b, c, and d, respectively.

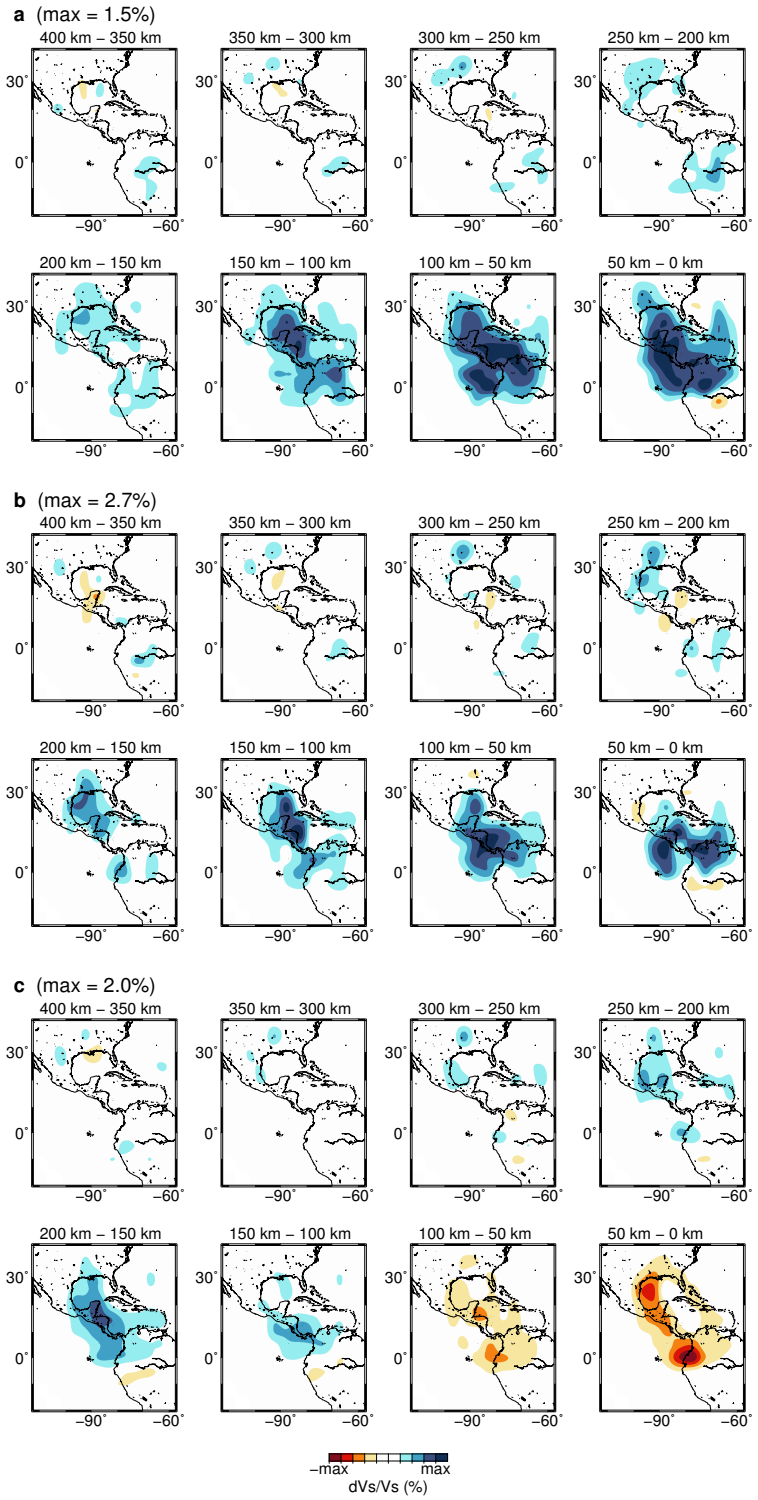

Figure 18. Block tests. 3-D inversion using as input exact 1-D synthetics computed with the DSM (without using the Born approximation) for 1-D models that differ from PREM by (a) a 1% increase 0–150 km above the CMB; (b) a 2% increase 0–150 km above the CMB; (c) a -2% decrease 0–100 km above the CMB and a 2% increase 100–200 km above the CMB. The color scale is different for panels a, b and c, and is normalized to fit the maximum amplitude of each recovered model.

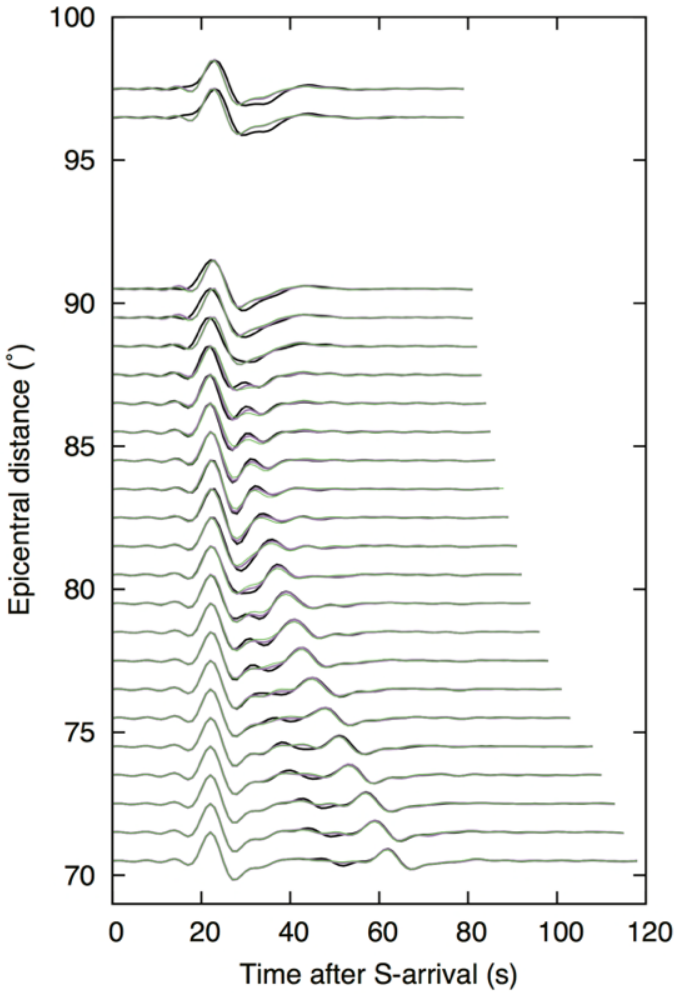

Figure 19. Record section for the block test model (c). Profile of synthetics stacked every 1˚ for the initial model PREM (green), the input model for the block test shown in Figure 18c (black), and the Born waveforms for the recovered model shown in Figure 18c (purple). S-ScS differential travel-times for the initial and the input model are nearly identical, implying that the input model is not resolvable by travel-time tomography. Instead, the waveform inversion fits the difference in amplitude between the initial and input model synthetics.

Figure 19 shows that the recovered model in Figure 18c fits the difference in amplitude between the input and initial model (PREM). Since the S-ScS differential travel-time is nearly zero, travel-time tomography cannot resolve this input model. Also, Figure 19 shows that the inversion is highly sensitive to data around epicentral distance 85˚, which cannot be used by travel-time tomography since S and ScS merge at those epicentral distances.

Figure 18b of the block test shows that the shape of the recovered high-velocity anomalies 0–50 km above the CMB is patchy, and might correlate to some extent with the two distinct high-velocity anomalies CA and VZ at the CMB (we note that when the velocity perturbation in the input block test model is smaller, as in Figure 18a, the recovered high-velocity anomaly is less patchy, which suggest that this might be due in part to the linear approximation). Figure 20 suggests that the presence of two distinct high-velocity anomalies CA and VZ at the CMB separated by a low-velocity corridor in CACAR and CACAR′ is not an artifact due to nonlinear effects, as it shows that this feature remains when using longer period waveforms, for which nonlinear effects are less significant. We compare the results of block tests for the same input model as in Figure 18b (+2% S-velocity anomaly 0–150 km above the CMB) with results of the actual inversion for three different minimum periods: 8, 12.5, and 20 s (Figure 20a, b and c, respectively) for the layer 0–150 km above the CMB. Figure 20 shows that as the minimum period increases, the recovery of the block test input model improves, but that the inferred model using actual data does not change significantly (in particular, there are still two distinct high-velocity anomalies at the CMB separated by a low-velocity anomaly). This suggest that the small-scale anomalies in models CACAR and CACAR′ are (for the most part) actual features of the data and not artifacts due to nonlinear effects (although some small artifacts are probably present).

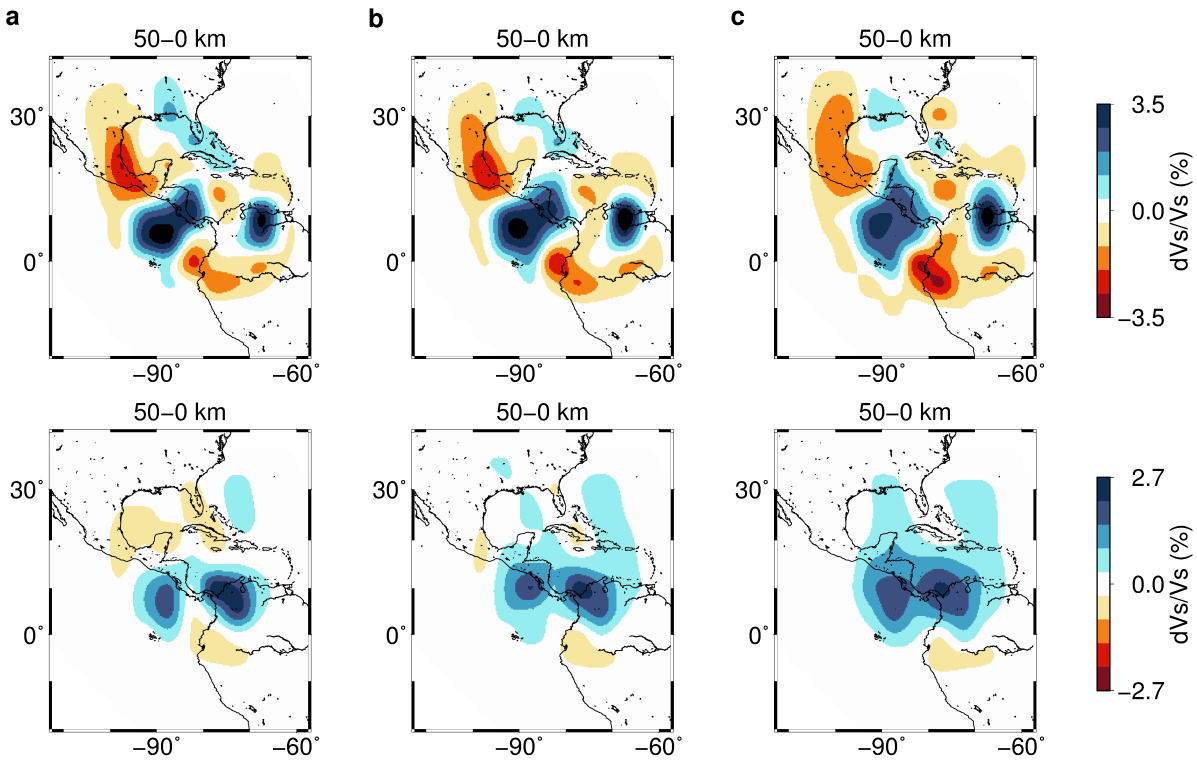

Figure 20. Results of block test for the input model in Figure 18b compared with results for the actual inversion as function of for three different minimum period: a) 8–200s, b) 12.5–200 s, and c) 20–200 s. Only the layer 0–50 km above the CMB is shown. Although the recovered high-velocity anomalies for the block test become patchy when using higher frequencies (column a), the presence of two distinct high-velocity anomalies CA and VZ separated by a low-velocity anomaly is not affected by changing the minimum period, suggesting that this is a robust feature. We note that the results in this figure for the 8–200 s filter slightly differ from those in Figure 18b, because of the use of a slightly different kernel (computed using wavefield interpolation) and dataset.

Comparison of the block tests in Figure 18a and b shows the importance of using an appropriate 1-D initial model for the inversion, due to the fact that the dataset of S and ScS waveforms used is more sensitive to the vertical than to the lateral velocity structure (due to the long horizontal traveling distance of the S and ScS waves near the depth of their turning point at large epicentral distances). We note that this is most likely also the case even in studies using 3-D reference models and iterations, since the final models in these studies are usually relatively close to the initial 3-D model (e.g., Tao et al., 2018). For the case of the block test in Figure 18b, the recovery would most likely be improved if a better initial 1-D model was inferred first, and then 3-D inversion was conducted using this newly inferred 1-D model (previous studies in our group showed that such an input 1-D structure can be recovered using 1-D inversion, e.g., Konishi et al., 2009). Thus, the fact that the models CACAR and CACAR′ (which was obtained using a newly inferred initial 1-D model) are in general agreement with each other suggest that 1) the 3-D structure in CACAR and CACAR′ is robust (although the absolute velocity for these two models differ slightly, which could probably be improved by using longer periods in the inversion, as suggested by Figure 20c), and 2) that the block test in Figure 18b is one of the expected worst cases for inversion, and is probably not representative of the actual inversions in this chapter.

As an additional test of the validity of the Born approximation for inversion of S-velocity in D″, I show a "nonlinear checkerboard test" in Figure 21. The input model is shown in panel a. Synthetic seismograms for this model were computed using full 3-D wave calculation (SPECFEM3D GLOBE) (Komatitsch et al., 2015). Because of heavy computational requirements with increasing maximum frequency, we computed synthetics down to ~17 s and filtered them with a bandpass filter between 23–200 s (as compared to the 8–200 s synthetics

we used in the inversion of the actual data). I also do not include anelasticity, as the synthetics computed using SPECFEM3D GLOBE differ slightly from those computed using the DSM, whereas for the elastic case, which we use for this test, the synthetics for DSM and SPECFEM3D GLOBE for PREM are in nearly perfect agreement. Since we expect longer wavelengths to have less resolving power than shorter wavelengths, I increase the dimension of heterogeneities in the checkerboard pattern to 100 km in the vertical direction (we keep the same lateral dimension of 5˚ x 5˚ as in the inversion of actual data). The result (Figure 21b) shows that the inversion using the Born approximation underestimates the absolute amplitude of the perturbations by about 30% (~2% for the inversion result, compared to the 3% perturbation of the input pattern), but that the pattern of high- and low-velocity anomalies is reasonably well recovered.

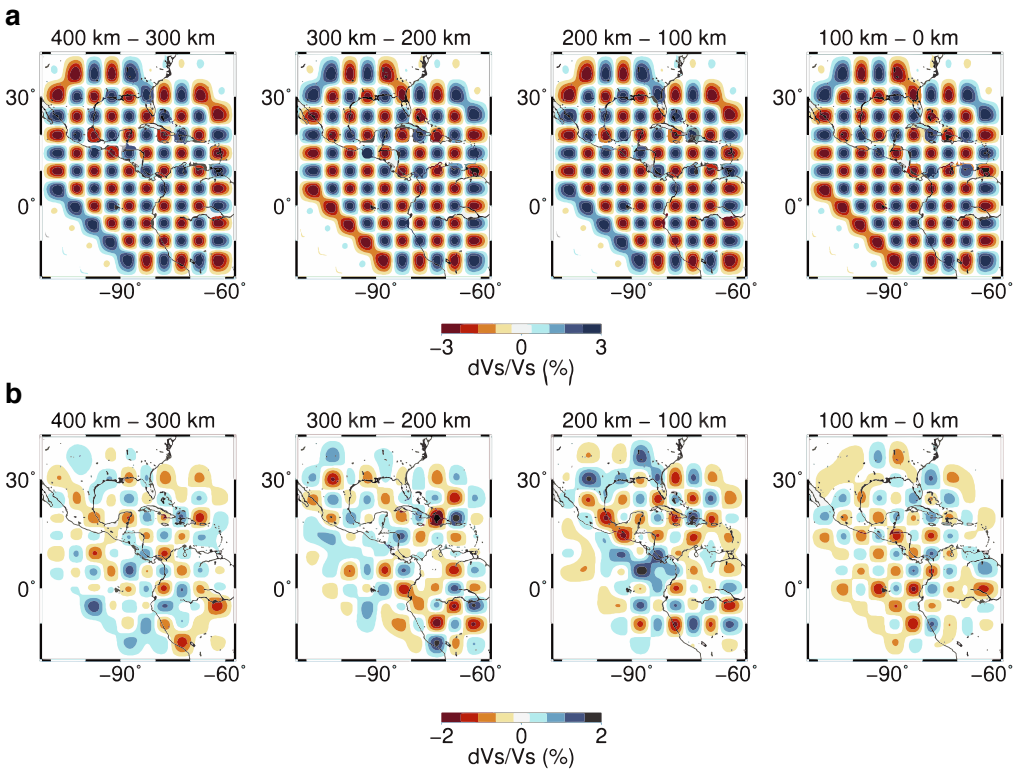

Figure 21. Nonlinear checkerboard test. In order to investigate the validity of the Born approximation used in the inversion, we conduct a "nonlinear checkerboard test" by using input synthetics computed using full 3-D wave calculation (SPECFEM3D GLOBE; Komatitsch et al., 2015). (a) Input checkerboard pattern used for the full 3-D wave calculation, with 3% heterogeneities of dimension 5˚ x 5˚ laterally x 100 km vertically. The synthetics for this input pattern are accurate down to a period of ~17 s, and filtered with a Butterworth bandpass filter between 23–200 s. (b) Recovered pattern after inversion using the partial derivatives computed using the Born approximation. Note that the maximum amplitude shown by the color scale in panel (a) (3%) differs from that in panel (b) (2%). Figure 21 shows that inversion using the Born approximation recovers reasonably well the pattern of high- and low-velocity anomalies, although the amplitude of perturbations is subject to some uncertainties.

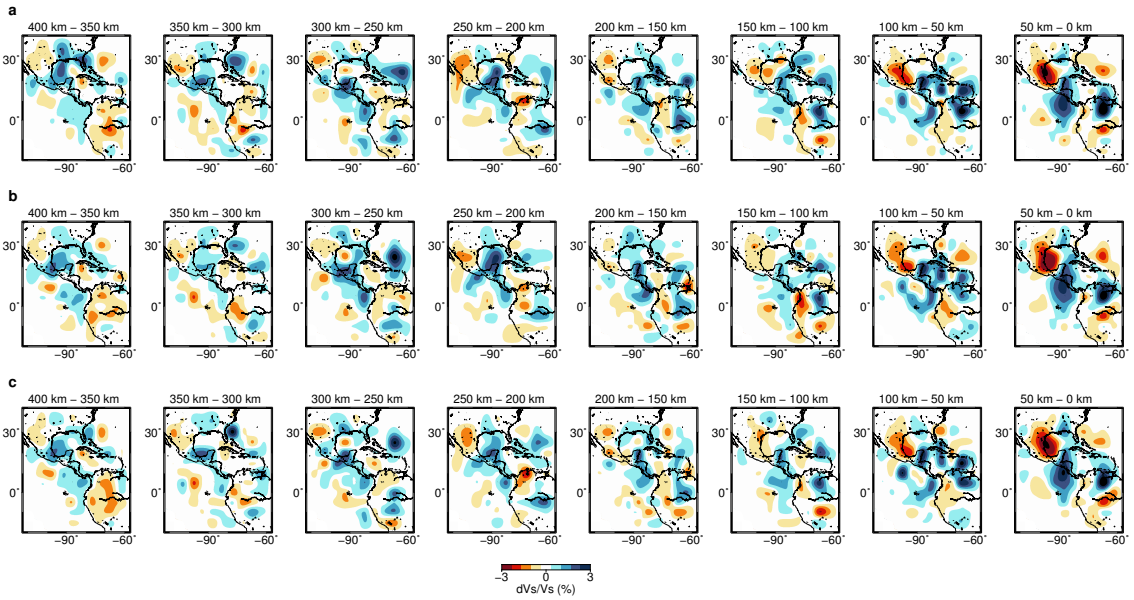

Figure 22. Jackknife test. Panels (a)-(c) show three inversion results using 50% of the records (~4000 records) randomly picked from the total dataset used for the full inversion. The weighting and number of CG vectors is the same as for the full inversion. The three models are nearly identical, which suggests that the 3-D model inferred in this chapter is robust.

To test the robustness of CACAR, I also conducted a jackknife test (Figure 22); I conducted three inversions each of which used 50% of the data randomly picked from the full dataset. The result shows that the high-velocity anomaly beneath Central America and that beneath Venezuela might be connected in the north of the target region beneath the Caribbean islands (the model inferred in this chapter is less well constrained in this region due to the smaller number of raypaths), but that a low-velocity corridor along the west coast of South America and extending to the Caribbean Sea is a robust feature; thus the two high-velocity anomalies imaged at the CMB beneath Central America and Venezuela are most likely separated. The jackknife test also shows that the strong low-velocity anomaly beneath Mexico is a robust feature of CACAR.

In order to visually confirm that the final models CACAR and CACAR′ do indeed improve the fit of the synthetics to the data, I show record sections of traces (data, and synthetics for initial and final models) for two events (#35, and #49 in Table A1) in Figure 23. Figure 23 shows that the synthetics for the final models CACAR and CACAR′ are closer to the data than the synthetics for PREM or PREM′. In particular, we note that the double-peaked S-wave at $\Delta > \sim 95^{\circ}$ in the observed traces, as well as an ScS precursor, which both indicate triplication of the S-wave due to a strong discontinuity in the D″ layer, are better reproduced by the synthetics for the final models CACAR and CACAR′ than those for the initial models. This visual check is strictly an ancillary measure for quality control.

In Figure 24, I show additional evidence for the presence of small-scale anomalies in CACAR and CACAR′, namely stacked waveforms that target three localized features in CACAR and CACAR′. Figure 24a, b, and c show stacked waveforms for raypaths sampling the high-velocity anomaly CA, the low-velocity corridor between CA and VZ, and the high-

velocity anomaly VZ, respectively. In Figure 24a, b, and c, the S-ScS travel-time residuals $\delta T_{ScS} = T_{ScS}^{obs} - T_{ScS}^{syn}$ for the initial model (PREM) are negative, positive, and negative, respectively, which correlates with the presence of the high-velocity anomalies CA, the low-velocity corridor, and the high-velocity anomaly VZ. The travel-time residual δT_{ScS} for the final 3-D model (blue traces) is clearly reduced (improved), further suggesting that the small-scale features of models CACAR and CACAR′ represent the information contained in the actual data.

Although the raypaths in the dataset used in this chapter are almost all aligned in the North-South direction, the checkerboard tests (Figure 12, Figure 13, and Figure 21) seems to show nearly no smearing along the event-receiver path (North-South direction). As an additional check of the absence of smearing, I conducted two point-spread function tests (Figure 25). The results confirm the absence of smearing in the North-South direction. This suggests that the fact that the high-velocity anomaly CA and VZ in CACAR and CACAR′ are elongated in the direction of the event-receiver path is probably a feature of the actual structure in D″ beneath Central America, and not an artifact due to smearing.

I suggest a qualitative explanation for the absence of smearing in the North-South direction (as shown in Figure 25**)** in Figure 26. The large range of epicentral distances and events latitudes in the dataset used in this chapter creates a crossing raypath geometry for ScS in a vertical cross-section. This crossing raypath geometry of ScS in the lowermost mantle is probably the reason why I can resolve individual voxels in the lowermost mantle using S and ScS waveforms.

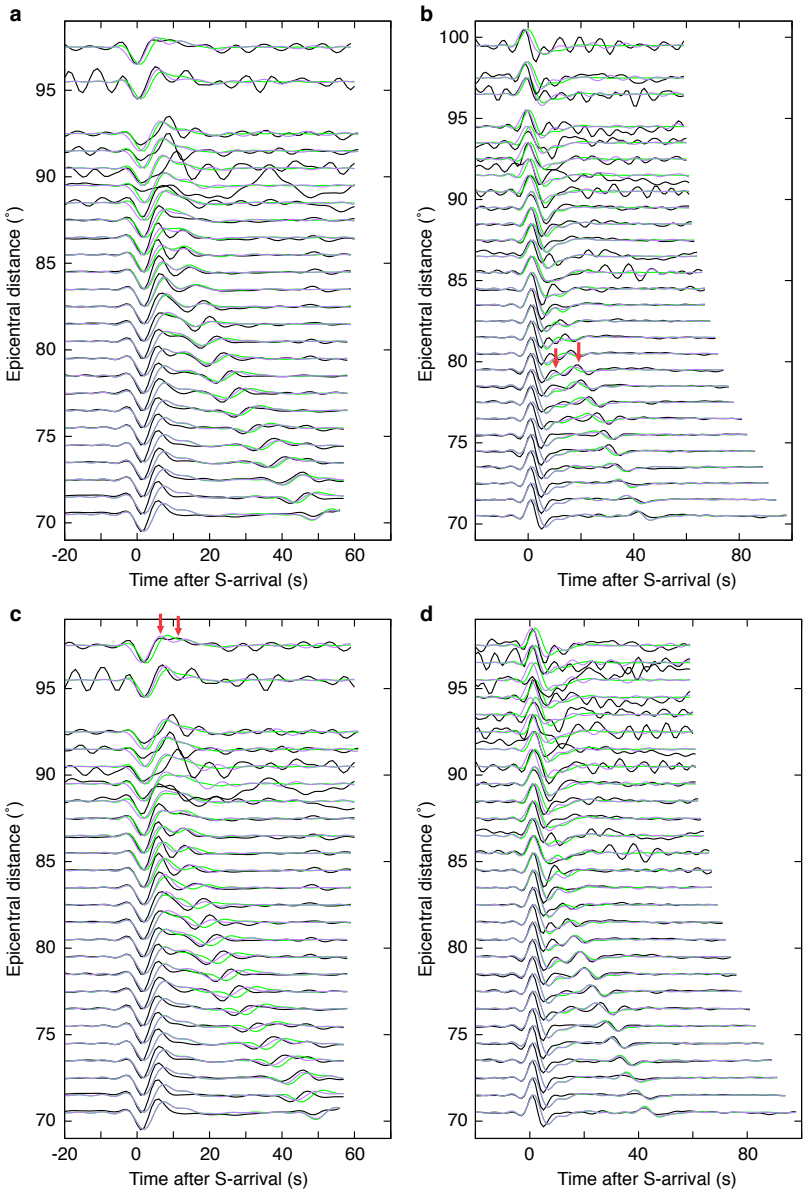

Figure 23. Quality control stacks. Record sections for the two events #35 and #49 in Table A1. The traces are stacked every 1˚ of epicentral distance (the time-shifts used in the inversion are applied to the observed data). Observed traces are shown in black. The top row shows synthetics for the initial models PREM′ (green), and final model CACAR′ (purple). The bottom row shows synthetics for the initial model PREM (green), and final models CACAR (purple). The double-peaked S-wave at $\Delta > \sim 95^{\circ}$ in the observed traces (red arrows in panel c), and ScS precursor (red arrows in panel b), which both indicate triplication of the S-wave due to a strong discontinuity in the D″ layer, is partially reproduced by the synthetics for the final models. The horizontal axis is time relative to the S-arrival time for the respective 1-D starting models, PREM and PREM′.

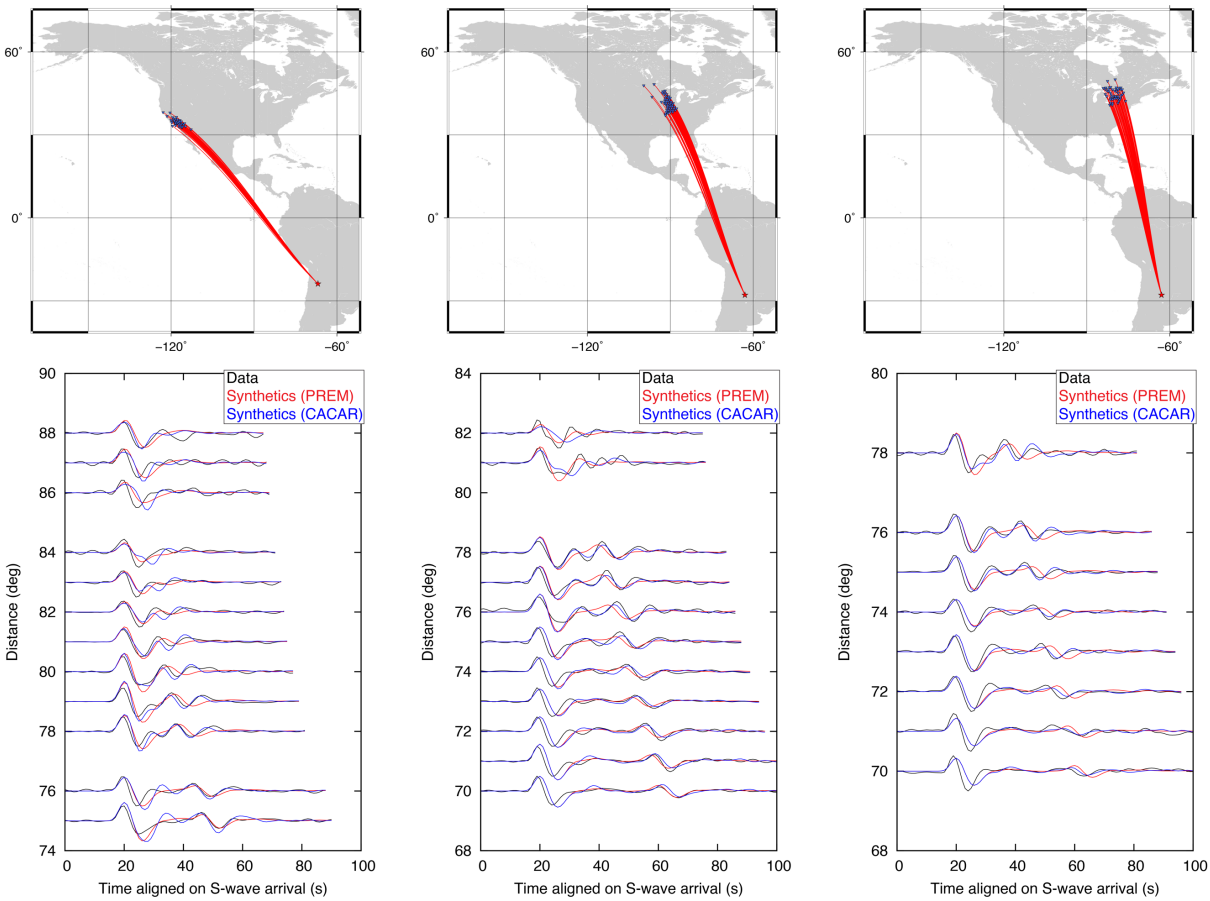

Figure 24. Localized stacks to show the improvements of synthetics for CACAR for three localized regions for event #57 in Table A1. The top row shows maps with stations and raypaths used for the stacks shown in the bottom row. Columns a–c shows raypaths sampling a) the fast region CA at the CMB; b) the slow velocity corridor between CA and VZ; c) the fast region VZ at the CMB. The sign of differential S-ScS travel-time residuals in columns a–c correlates with the sign of velocity anomalies in CACAR and CACAR′. Synthetics and data are filtered between 12.5–200 s. Note that the minimum period for the actual inversion is 8–200 s.

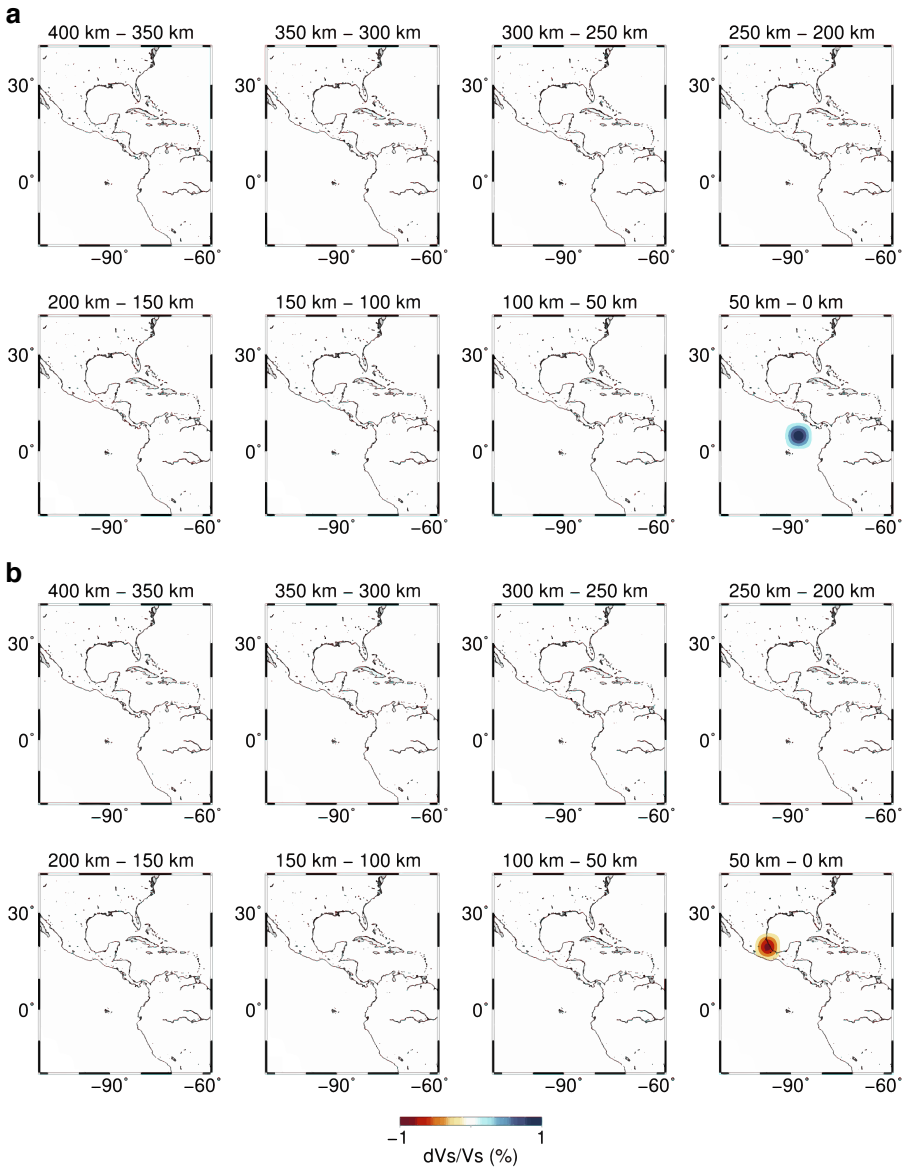

Figure 25. Point-spread function. Panels (a) and (b) show the result of inversion for a synthetic input model with only one voxel perturbed. (a) S-velocity +1% at -87.5˚ longitude, 5˚ latitude, and 25 km above the CMB; (b) S-velocity -1% at -97.5˚ longitude, 20˚ latitude, and 25 km above the CMB. The results do not show any smearing from the perturbed voxel, suggesting that the inversion can resolve voxels individually.

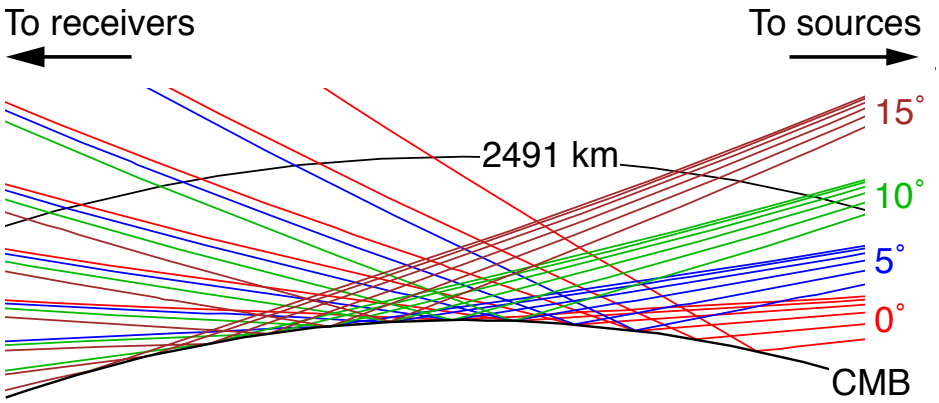

Figure 26. Crossing raypaths. A vertical cross-section shows the geometry of ScS raypaths 0–400 km (2491 km depth) above the CMB, for receivers at epicentral distances 70˚<Δ<95˚ and sources at depth 571.3 km, longitude 0˚, and latitude 0˚, 5˚, 10˚, and 15˚ (red to brown lines). The raypaths are computed using the TauP toolkit (Crotwell et al., 1999) for PREM. The range of epicentral distances and event latitudes is similar to that of the dataset used for the inversion. The large range of epicentral distance and event latitudes creates vertically crossing raypaths, which is probably the reason we can resolve heterogeneities along the event-receiver path (North-South direction).

3.5 Discussions

I used waveform inversion to image the complex, small-scale D″ structure with two distinct strong high-velocity anomalies at the CMB shown in Figure 9 to Figure 11. The inferred models CACAR and CACAR′ suggest the presence of two paleoslabs just above the CMB and dense chemical anomalies (i.e., iron-enriched material) concentrated in the lowermost 100 km of the mantle. Figure 27 shows a schematic illustration of this interpretation.

As mentioned in the introduction, geological evidence for subduction ~250 Ma beneath the western margin of Pangaea, together with an average ~1.5 cm/yr subduction rate in the lower mantle, suggests that remnants of past subduction should be found at the CMB beneath Central America and the Caribbean. The following features of CACAR and CACAR′ suggest the presence of paleoslabs: (1) ~3% high-velocity anomalies just above the CMB, (2) vertically continuous low-velocity structures at the edge of the inferred paleoslabs, (3) the good agreement between the location of the past Farallon plate boundary and the high-velocity anomaly beneath Venezuela, and (4) the correlation between the topography of the D″ discontinuity reported by previous forward modeling studies and the distribution of high- and low-velocity anomalies in CACAR and CACAR′.

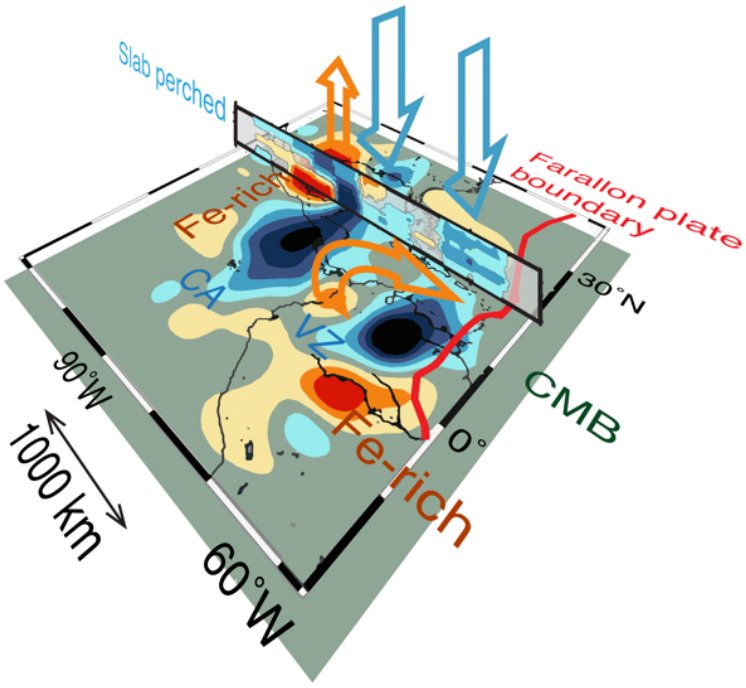

Figure 27. Possible geodynamical interpretations. Two high-velocity anomalies just above the CMB, beneath Central America (CA) and Venezuela (VZ), respectively, suggest the existence of two distinct paleoslabs that took different subduction paths to the lowermost mantle. Strong low-velocity anomalies in the lowermost 100 km of the mantle suggest dense iron-rich material (e.g., basaltic composition). Such dense low-velocity material at the CMB can also explain why slab CA is perched above the strong low-velocity beneath Mexico. Low-velocity material is upwelling from the iron-rich anomalies and between the two slabs. The Farallon plate boundary at 180 Ma (red curve) is consistent with the location of the high-velocity anomaly VZ, suggesting that the slab VZ subducted at the western margin of Pangaea.

To explain the 3% high-velocity anomalies in the inversion results, we note that a ~1.5% high-velocity anomaly can be explained by the MgPv to MgPPv phase transition, and the remaining 1.5% by a 390 K decrease in temperature using the temperature derivative for MgPPv at lowermost mantle conditions (Wentzcovitch et al., 2006). Since the temperature of 3800 K should be homogeneous at the CMB because of the vigorous convection in the outer core (Stevenson, 1981), a 390 K decrease in temperature ~25-50 km, or less, above the CMB strongly suggests the presence of cold material. An alternative (or complementary) interpretation for the high-velocity anomalies could be the presence of chemical anomalies with S-velocity faster than pyrolite. Such high-velocity composition can be obtained by an increased MgPv to ferropericlase (Fp) ratio compared to the average pyrolitic composition. The S-velocity increase in pure MgPv compared to pyrolytic composition was estimated to be ~2.5% based on theoretical calculations (Tsuchiya & Tsuchiya, 2006). Tateno et al. (2014) conducted melting experiments on peridotite to lowermost mantle pressures and showed that MgPv is a liquidous phase at pressures above ~34 GPa, which implies that MgPv-rich anomalies could have been formed in the lowermost mantle by crystallization of a past basal magma ocean (Labrosse et al., 2007).

The presence of a low-velocity vertically continuous structure (red arrow in Figure 11b) that separates the two high-velocity anomalies CA and VZ and connects to a low-velocity region that was found to extend to the mid-mantle by a recent global waveform inversion study (French & Romanowicz, 2014) is consistent with the subduction of two distinct paleoslabs. This continuous low-velocity structure from the CMB to the mid-mantle suggests that hot TBL material has been upsplashed by paleoslabs reaching the CMB, as shown in geodynamical simulations (Tackley, 2011), and upwelled to the mid-mantle.

Although MgPv-rich material at the CMB could provide an alternative explanation for the high-velocity anomalies in CACAR and CACAR′, the positive correlation on small scales between the distribution of high- and low-velocity anomalies in CACAR and CACAR′, and the lateral variations in the elevation of the D″ discontinuity reported by previous regional reflectivity and forward modeling studies (see Results section) suggests (assuming the MgPv to MgPPv phase transition) that the high-velocity anomalies CA and VZ are regions with lower-than-average temperature, and that the low-velocity corridor that separates CA and VZ is a region with higher-than-average temperatures. The resulting strong temperature gradients might be difficult to sustain without the presence of colder, and thus stiffer, slab remnants. We note this does not rule-out the possibility that a smaller part of the high-velocity perturbation in the high-velocity region could be due to MgPv/MgPPv-rich material.

CACAR and CACAR′ thus suggest the presence of a paleoslab at the CMB beneath Venezuela, which is in good agreement with the past location of the Farallon plate and might thus be a remnant of subduction at the western margin of Pangaea. Beneath Venezuela, this paleoslab is covered by a low-velocity region ~200-400 km above the CMB, but it gradually extends to the North along the Farallon plate boundary as the altitude above the CMB increases (Figure 9 and Figure 10).

The high-velocity anomaly CA beneath Central America has often been imaged by seismic studies and has been interpreted as folding or spreading of the Farallon plate at the CMB (Thomas et al., 2004; Hutko et al., 2006; Kito et al., 2008). However, as noted above, the location of this paleoslab at the CMB is ~1000 km to the west of the past Farallon plate boundary, which tends to argue in favor of two separate subduction paths in the lower mantle rather than the spreading of the Farallon paleoslab at the CMB.

I present below two possible interpretations based on CACAR and CACAR′. We note that the velocity model for the structure above the target region should be further investigated in order to verify these interpretations.

As discussed in the introduction, initiation of the subduction of the Farallon plate is estimated at 180-207 Ma (van der Meer et al., 2010) and was probably accompanied by strong igneous activity (van der Lelij et al., 2016). The two distinct paleoslabs we observe at the CMB might possibly suggest that the strong igneous activity was related to tearing or breaking of a plate subducting beneath western Pangaea, leaving two paleoslabs that sunk to the CMB, followed by the subduction of the Farallon plate itself.

Alternatively, intra-oceanic subduction of Pacific oceanic floor beneath a volcanic arc within the Pacific located around the current location of Central America, possibly the Stikinia–Quesnellia arc (Johnston & Borel, 2007), could also explain the presence of two distinct slabs at the CMB and the ~1000 km discrepancy between the location of the Farallon plate boundary and the location of the paleoslab we observe beneath Central America.

3.5.1 Chemical heterogeneities

The strong velocity contrast (5-6% peak-to-peak over less than 100 km vertically and 300 km laterally) that we observe in the lowermost 100 km of the mantle seems too strong and sharp to be explained by temperature variations only; a 5-6% velocity anomaly would require a 1300-1560 K temperature variation (Wentzcovitch et al., 2006). Chemical heterogeneities with enriched iron content have a lower seismic velocity than pyrolite and can explain strong negative anomalies (Tsuchiya & Tsuchiya, 2006). They are also denser than pyrolite and thus could remain close to the CMB. CACAR and CACAR′ show that strong low-velocity anomalies are concentrated in the lowermost 100 km of the mantle. Chemical anomalies at the

CMB might provide a more reasonable explanation for the strong velocity gradient we observe within 100 km of the CMB.

A likely candidate for chemical heterogeneities at the CMB is subducted basaltic oceanic crust (MORB) that constitutes the upper part of slabs (Irifune, 1993). MORB is a complex mineral assemblage that is enriched in calcium perovskite (CaPv) (compared to pyrolite) and contains a significant amount of calcium-ferrite type aluminous phases (Irifune, 1993), which makes the determination of its seismic velocity challenging. Although previous theoretical calculations had suggested that (cubic) CaPv had a relatively high shear modulus (Stixrude et al., 2007), leading to MORB with higher S-velocity than pyrolite in the lower-mantle (Xu et al., 2008), more recent calculations (Kawai & Tsuchiya, 2015) and sound-velocity measurements at lower-mantle pressures (Gréaux et al., 2019) showed that the shear modulus of (cubic) CaPv is lower than previously calculated, which suggestes that MORB has lower S-velocity than pyrolite in the lower-mantle (it is also denser); this could possibly explain relatively strong low-velocity anomalies near the CMB. Geodynamical simulations of a slab sinking to the CMB show that in the case where a thin basaltic layer is present just above the CMB, the center of the slab reaches the CMB, but its edges are perched above a dense basaltic layer (Tackley, 2011). Assuming that MORB indeed has a lower S-velocity than pyrolite, this would be consistent with the results presented in this chapter (Figure 11a and Figure 27) that suggest that the northwestern edge of the paleoslab CA is perched above a strong low-velocity anomaly.

The dynamical stability of a paleoslab perched above a low-velocity anomaly will depend on the density and viscosity contrast between the possible chemically distinct material at the CMB and the rest of the mantle. Some geodynamical studies suggest that such material beneath

subduction zones will always be entrained by mantle convection and form large thermo-chemical piles imaged as large low shear-velocity provinces by seismic tomography (Mulyukova et al., 2015). Other studies suggest that a denser (4.3% denser than harzburgite) primordial layer at the CMB cannot be easily entrained by mantle convection (Nakagawa & Tackley, 2015). The fact that CA is partially perched above a possibly chemically distinct low-velocity anomaly seems to suggest that the density contrast between the ambient mantle material and the chemical anomaly is relatively strong, and that that a chemical anomaly may thus not be easily entrained by mantle convection.

3.6 Conclusions

The models CACAR and CACAR′ shows that low velocity anomalies adjacent to and below high velocity anomalies interpreted as cold subducted paleoslabs seem to be connected to low velocity anomalies inferred by previous tomographic studies (French & Romanowicz, 2014). The low velocity anomalies beneath the Caribbean in the lower mantle can be interpreted as originating at the low velocity anomalies just above the CMB found in CACAR and CACAR′. These low velocity anomalies have possibly been growing below the cold paleoslabs due to increased heat transfer from the core, as suggested by some geodynamical studies (Tackley, 2011). They can be interpreted as chemical rather than thermal anomalies, due to dense primordial material or iron-rich materials chemically concentrated from pyrolytic or basaltic materials due to the heat from the core or self-generated heat in basalt. Once such concentrations of dense materials are created immediately above the CMB, they are pinned there until stirred or entrained by mantle convection, when they become the origin of upwelling flow in the lower mantle. Considering these pinning effects, significant upwelling flows in the

lower mantle found beneath hotspots by previous studies (French & Romanowicz, 2015) might be related to past subduction history.

The significant upwellings beneath the Pacific and Africa are located beneath past supercontinents, Rodinia and Gondwana, respectively (Rogers & Santosh, 2004). Since significant downwelling flow is expected in the lowermost mantle beneath the location of accumulated subduction zones such as the present East Asian zone and past supercontinental margins, self-generated heat in subducted basaltic material is likely to be the origin of the significant upwelling (Maruyama et al., 2007). The models CACAR and CACAR′ are consistent with the possibility that subducted material accumulated at the TBL produces concentrations of iron-rich material due to chemical differentiation which becomes the origin of upwelling flow from this iron-rich material pinned at the CMB. Significant downwelling from the Earth's surface due to subduction could be responsible for transporting large amounts of material enriched in radiogenic elements such as basalt that produce iron-rich materials according to chemical differentiation due to self-generated heat which are then pinned at the CMB, thus providing the origin of the hotspots. This suggests that the modality of convection in the lower mantle is controlled by plate tectonics on the Earth's surface.

As the temperature at the CMB is isothermal because of the vigorous convection of the outer core (Stevenson, 1981), knowledge of the temperature gradient just above the CMB is essential for understanding the thermal evolution of the Earth. Since the heat flux is the product of the temperature gradient and the thermal conductivity, intermittent paleoslab subduction could make the cooling rate at the surface of the outer core both spatially and temporally heterogeneous. Since significant downwelling flow and iron-rich material pinned at the CMB are expected in the lowermost mantle beneath the location of accumulated subduction zones

such as the present East Asia and past supercontinental margins, long lasting localized significant cooling of the core might affect the geodyanamo (including contributing to causing geomagnetic reversals) in the outer core.

The results presented in this chapter can be explained by a whole mantle convection model where cold and dense paleoslabs sink to the base of the mantle and trigger upwelling flow of hot and less dense material. Such a whole mantle convection is also supported by recent imaging of broad plumes originating at large low velocity provinces (LLVP) at the CMB and continuously connecting to current hotspots at the Earth's surface (French & Romanowicz, 2015). We note that the region beneath Central America and the Caribbean might not be representative of the whole Earth, and that a layered-type convection with decoupling between the upper and lower mantle (or above and below ~1000 km depth) might be appropriate for other regions, as suggested by tomographic images of stagnating slabs (Fukao & Obayashi, 2013) and geodynamical simulations (Ballmer et al., 2015).

Chapter 4 3-D S-velocity structure of the mantle transition zone beneath Central America and the Gulf of Mexico inferred using waveform inversion

4.1 Introduction

4.1.1 Tectonics of Central America

4.2 Data and model

4.3 Materials and Methods

4.3.1 Synthetics and Partials

4.3.1 Inverse problem

4.3.2 Corrections for the 3-D structure above the target region

4.4 Results

4.4.1 Morphology of the Cocos slab and a plume

4.4.2 Cross-sections and comparison to previous studies

4.4.3 Variance reduction and AIC

4.5 Robustness and resolution tests

4.6 Discussions

4.6.1 Tearing of the Cocos slab

4.6.2 Connection between the Cocos and Farallon slabs

4.6.3 Factors controlling subduction modality

4.7 Conclusions

Chapter 5 Discussion

In this work, I used waveform inversion to obtain high-resolution tomographic models of the MTZ and the D″ region beneath Central America and its vicinity. In the MTZ, I found that 1) tearing of the Cocos slab above ~410 km depth could possibly explain the disruption in Cocos plate velocity at 10 Ma (Müller et al., 2016), 2) the Cocos plate penetrates into the lower-mantle and connects continuously to the Farallon plate. In the D″ region, I found evidence for two paleoslabs, beneath Central America (imaged in previous studies), and beneath Venezuela (discussed for the first time in Borgeaud et al., 2017). The paleoslab beneath Venezuela possibly corresponds to the Farallon slab at ~180 Ma. These findings can be used to place (approximate) constraints on the sinking velocity of the slabs beneath Central America in both the upper- and lower-mantle, and thus on the viscosity contrast between the upper- and lower-mantle, as discussed in the following section.

5.1 Constraints on the viscosity of the mantle

As discussed previously, the MTZ models inferred in this work suggest that the Cocos slab at ~410 km depth results from subduction at ~10 Ma, which implies an average sinking velocity of ~4.1 cm/yr in the upper-mantle. Assuming that the sinking velocity is approximately the same in the upper-mantle and MTZ, this implies that the Cocos slab at 660 km depth results from subduction at ~16 Ma (a higher viscosity in the MTZ would slightly increase this age, but this would not significantly affect the results discussed below). We note that it is difficult to place constraints on the precise age of the Cocos plate at ~660 km depth in this region by matching the past location of the Middle America Trench to the location of high-velocity anomalies corresponding to the slab, because the slab is deflected eastward near and below ~660 km. Also, although the Cocos slab beneath Guatemala (C2) is significantly less deflected

than its northern part (C1), past location of the Middle America Trench in this region is nearly fixed during ~10–30 Ma, which means that it is not possible to estimate the age of the Cocos plate in this region by comparison with past plate boundaries.

The D″ model inferred in this dissertation suggests that remnants of the Farallon plate that was subducting at ~180 Ma currently reside at the CMB (beneath Venezuela). The MTZ model shows that the Cocos slab penetrates directly into the lower-mantle (beneath Guatemala), which implies that there is (nearly) no residence time of the Cocos slab at 660 km, and that the slab at 660 km is ~16 Myr old. This means that the Farallon slab sunk from ~660 km depth to the CMB in ~164 Myr, resulting in an average sinking velocity of ~1.4 cm/yr in the lower-mantle. This sinking rate is in agreement with previous estimates for the global average of sinking rate of ~1.2 cm/yr (van der Meer et al., 2010), and 1.1–1.9 cm/yr (Domeier et al., 2016).

The above estimates imply that the ratio of (vertical) sinking velocity between the upper- and lower-mantle is as given in eq. (48)

$$\frac{v_{sink}^{UM}}{v_{sink}^{LM}} \approx \frac{4.1\ \mathrm{cm/yr}}{1.4\ \mathrm{cm/yr}} \approx 3. \tag{48}$$

A simple equation linking slab sinking velocity to viscosity can be obtained if we assume that the gravitational force acting on the higher-density slab is equilibrated by viscous drag, which results in slabs sinking at the Stokes velocity as in eq. (49).

$$v_{stokes} = \frac{Mg}{a\nu}, \tag{49}$$

where a is a geometrical factor that roughly represents the surface of the slab in contact with the flow (i.e., roughly, the surface perpendicular to the sinking velocity). This was suggested to be the case by previous studies (e.g., Capitanio et al., 2007).

Eqs. (48) and (49) imply that if there is no thickening as the slab penetrates into the lower-mantle, the viscosity contrast between the upper- and lower-mantle is ~3. The MTZ model suggests, however, that the Cocos slab thickens when penetrating into the lower-mantle. The increase in mass of the thickened Cocos slab in the lower mantle can be estimated from the cross-sections () to be ~2–4, which means that M in eq. (49) is (roughly) ~2–4 times larger in the lower-mantle. This leads to an estimated viscosity contrast between the upper- and lower-mantle (in the region beneath Central America) of ~6–12.

This viscosity contrast is ~3–5 times smaller than that estimated from post-glacial rebound, plate divergence velocities, and geoid anomalies (Hager, 1984; Forte & Mitrovica, 2001). It is, however, in general agreement with that estimated by comparing observed plate subduction velocities to that obtained in (fully dynamic) numerical simulations of sinking slabs (Quinteros et al., 2010). Quinteros et al. (2010) also suggested that a viscosity contrast of more than ~10 significantly deforms slabs in the MTZ, and results in smaller subduction velocities than reported in past plate motion models (~2–3 cm/yr).

We note that although estimates of the viscosity of the upper-mantle based on post-glacial rebound are robust, those for the lower mantle are more uncertain, since most of the constraints come from geoid anomalies, which are due to the highly uncertain 3-D density anomalies in the Earth's mantle. Forte and Mitrovica (2001) estimated the 3-D density anomalies based on previous global S-velocity models, assuming that the velocity anomalies are due to variations in temperature and iron content, and then converting them to density anomalies. Hager (1984) used a different approach, and estimated the density anomalies in the lower-mantle by mapping slabs in the upper-mantle based on the location of deep earthquakes, and assuming a penetration depth for all slabs into the lower-mantle, and a value for the density anomaly of

slabs. In both cases, relatively large uncertainties are most likely present in the estimated 3-D density field of the lower-mantle, and thus in the inferred 1-D viscosity structure of the lower-mantle using geoid anomalies. Comparison between high-resolution tomographic models and (fully dynamic) geodynamical simulations thus seems to be a promising way to obtain more robust and detailed constraints on the viscosity structure of the mantle.

5.2 Implications for the subduction modality and tectonic history beneath Central America

The models presented in this dissertation show smaller-scale S-velocity anomalies in the D" region than previous studies based on travel-time tomography or global waveform inversion, possibly suggesting smaller scale convection in the lowermost mantle (if a thermal origin of S-velocity anomalies is assumed, as argued in Chapter 3), in contrast with previous reports of increase in length-scale of structure in the lowermost mantle (e.g., van der Hilst et al., 1997). In particular, the length scale of high-velocity anomalies at the top of the lower-mantle as seen in the MTZ models in this dissertation is comparable to that of high-velocity anomalies in the D" models CACAR and CACAR′, which could be inconsistent with a previously suggested possible strong increase in viscosity at ~2000 km depth (Forte & Mitrovica, 2001).

In CACAR and CACAR′, the two distinct high-velocity anomalies labelled CA and VZ imaged at the CMB lie beneath Central America, and Venezuela, respectively, with VZ corresponding to the location of the Farallon plate boundary at 180 Ma. In the MTZ model, the high-velocity anomaly F1 beneath the Caribbean Sea, which I interpreted as the top of the Farallon slab, is ~1000 km west of VZ in CACAR and CACAR′. We also note that F1 is ~400

km east of the Middle America Trench at 23 Ma, which probably indicates that it is deflected when penetrating into the lower-mantle. On the other hand, VZ at the CMB is just west of the Farallon plate boundary at 180 Ma. This could possibly suggest an episode of slab rollback for the Farallon slab. The other possibility is that VZ does not correspond to the Farallon slab, but is a remnant of intra-oceanic subduction within the Pacific Ocean. On the other hand, there are no high-velocity anomalies beneath Central America at the top of the lower-mantle, further suggesting that the high-velocity anomaly CA at the CMB is a remnant of past intra-oceanic subduction within the Pacific Ocean.

5.3 Future work

This work showed the potential of waveform inversion using the full-wave theory (with the only approximation being the Born approximation) to enhance the resolution of 3-D S-velocity models by using the full amount of information contained in large datasets of relatively high-frequency S, ScS, and S-wave triplication waveforms from the USArray.

The method for waveform inversion used (and improved) in this dissertation differs from that used by other groups as follows: 1) no approximation is made in the computation of the partial derivatives; 2) partial derivatives for each individual seismic record are computed, which allows computation of the full $\boldsymbol{A^T A}$ matrix (sometimes called the approximate Hessian). In contrast, this is not the case for works using adjoint tomography (because, as mentioned in Chapter 2, the computation of partial derivatives for each record in a 3-D model is currently not feasible). 3) the drawbacks are that only one iteration is performed, and that the reference model is 1-D. The formulation of the inverse problem using the $\boldsymbol{A^T A}$ matrix allows, however, faster convergence toward the best-fitting 3-D model than when using only the gradient of the misfit is used, as in adjoint tomography (Chen et al., 2007). Furthermore, we showed that the

linear (Born) approximation used in this dissertation is reasonable for the case of the MTZ and the D" region. It is nonetheless apparent that nonlinear effects are not negligible, so that future works could try to account for these higher-order terms.

An important future extension of this work is the imaging of the P-velocity structure, which could be used to place additional constraints on the origin of velocity anomalies, i.e., thermal, chemical (e.g. Karato & Karki, 2001), or possibly the presence of water for the case of the MTZ (e.g., Zhao & Ohtani, 2009). The more efficient calculation of partial derivative kernels presented in this dissertation will be useful to compute kernels for the P-waves (P-SV wavefield), which are several times more computationally intensive than for the S-waves (SH wavefield). For the case of the D" region, the PcP phase is not often used in tomographic studies because of its small amplitude, which makes it difficult to pick reliable PcP arrival times (one of the few studies using PcP travel times is that of Ren et al., 2007, but the number of PcP travel-times used is relatively small). The PcP phase has been used, however, in studies using seismic array analysis to constrain the high-resolution 1-D P-velocity structure of the D" region (Hutko et al., 2008; Cobden & Thomas, 2013; Durand et al., 2018). This suggest that the PcP phase could be used in waveform inversion with seismic array data (as the USArray), as waveform inversion does not require the relatively imprecise task of picking PcP arrival times.

For the case of the MTZ, Tao et al. (2018) recently used adjoint tomography with SPECFEM3D_GLOBE (Komatitsch et al., 2015) to infer the high-resolution S- and P-velocity of the MTZ beneath Japan and eastern China using P- and S-wave triplications, and showed improvements of the inferred model compared to the one obtained using the travel-time of the

direct S- and P-waves (Chen et al., 2015). The more efficient computation of partial derivatives presented in this dissertation using wavefield interpolation can also be implemented for the P-SV wavefield, so that a P-velocity model of the MTZ beneath Central America using the same source-receiver geometry as for the S-velocity model can be obtained in the near future. Such a P-velocity model will allow better constraint of the cause of velocity anomalies, in particular whether the low-velocity anomaly L1 (interpreted in this dissertation as a lower-mantle plume) is of thermal origin, or if it might contain water or be chemically distinct (e.g., MORB as suggested by Gréaux et al., 2019).

Finally, the methods for waveform inversion for 3-D structure used in this dissertation could be extended to account for the topography of the CMB for the case of the D" region, and that of the 660 and 410 km discontinuities for the case of the MTZ. Variations in topography give additional constraints on the presence of temperature or chemical anomalies, and are also related to the viscosity structure (Deschamps et al., 2018). Previous studies have formulated partial derivatives for the topography of solid-solid and solid-fluid boundaries (Dahlen, 2005). Colombi et al. (2014) used this formulation to invert for CMB topography using the travel-time misfit, and showed the importance of jointly considering the 3-D mantle structure when inferring the CMB topography. Since variations of the CMB topography significantly affect (relatively) high-frequency waveforms (e.g., Colombi et al., 2014), waveform inversion should allow better constraint of the CMB topography, which makes it an interesting topic for future studies.

This page is intentionally left blank.

Chapter 6 Conclusions

In this work, I used waveform inversion applied to a large number of waveforms from the recent full deployment of the USArray in the contiguous US (2004–2015) to improve the resolution of tomographic models in and near the MTZ, and in the D″ region beneath Central America and its vicinity. For the case of the MTZ, I used waveforms of S-wave triplications, which are not often used in seismic tomography, and can enhance the resolution of tomographic models in and near the MTZ. I showed that:

1) regional waveform inversion together with large seismic array dataset can improve the resolution of tomographic models, especially for the case of low-velocity anomalies that appear stronger in the models inferred in this work than in previous tomographic models (this might be due to the use of full-wave theory, without using ray-theoretical approximation, as mentioned in the Introduction);

2) in D″, I find smaller-scale heterogeneities than reported in previous global models (in general agreement with previous regional models, e.g., Hung et al., 2005; Kawai et al., 2014). The increased raypath coverage due to the most recent deployment of the USArray in the eastern US also allowed imaging, for the first time, a strong high-velocity anomaly beneath Venezuela that could correspond to the Farallon paleoslab. The correlation between the location of high- and low-velocity anomalies in the D″ models inferred in this dissertation, and variations in topography of the D″ discontinuity suggest that the velocity anomalies can be explained mainly by the combination of lateral temperature variations and the bridgemanite to pPv phase transition. The small-scale nature of temperature variations suggests the presence of colder, and thus stiffer than the surrounding mantle, slabs in the hot thermal boundary layer

above the CMB. Strong vertical velocity contrasts ~100 km above the CMB also suggest the presence of Fe-enriched, denser material at the CMB.

3) in the MTZ, I find that the Cocos slab penetrates into the lower-mantle and connects to the Farallon slab, with complex subduction modality within the MTZ showing significant variations along the Middle America Trench. These variations could be due to the thermal structure of the Cocos slab, and to the presence of a lower-mantle plume locally reducing the buoyancy of the slab. These results suggest the importance of properties of the slab itself (temperature) and of the state of convection in the lower-mantle (presence of a plume) on the subduction modality of the Cocos slab (and most-likely other slabs as well).

4) the fact that the Cocos slab penetrates into the lower-mantle, and that small-scale high-velocity anomalies are present at the CMB beneath Central America and Venezuela is in favor of whole-mantle convection in this particular region, in agreement with previous studies. However, the higher-resolution images of the D″ region obtained in this dissertation allows further argument for the presence of slabs above the CMB (see point 2). These high-resolution images also allow estimating the viscosity contrast between the upper- and lower-mantle, resulting in a value of ~9–12. This value is ~3–5 times smaller than inferred by fitting post-glacial rebound and geoid anomalies, but is probably more robust due to the large uncertainties in the 3-D density field used to fit the geoid anomalies. This value is also in general agreement with numerical studies modeling plate subduction velocities.

References

Akaike, H. (1977). An extension of the method of maximum likelihood and the stein's problem. *Annals of the Insitute of Statistical Mathematics*, *29*, 153–164.

Amaru, M. L. (2007). *Global travel time tomography with 3-D reference models*. *Geologica Ultraiectina* (Vol. 274).

Ballmer, M. D., Schmerr, N. C., Nakagawa, T., & Ritsema, J. (2015). Compositional mantle layering revealed by slab stagnation at ~1000-km depth. *Science Advances*, *1*(11). https://doi.org/10.1126/sciadv.1500815

Ballmer, M. D., Houser, C., Hernlund, J. W., Wentzcovitch, R. M., & Hirose, K. (2017). Persistence of strong silica-enriched domains in the Earth's lower mantle. *Nature Geoscience*, *10*, 236–241. https://doi.org/10.1038/NGEO2898

Van Benthem, S., Govers, R., Spakman, W., & Wortel, R. (2013). Tectonic evolution and mantle structure of the Caribbean. *Journal of Geophysical Research: Solid Earth*, *118*(6), 3019–3036. https://doi.org/10.1002/jgrb.50235

Borgeaud, A. F. E., Konishi, K., Kawai, K., & Geller, R. J. (2016). Finite frequency effects on apparent S-wave splitting in the D" layer: comparison between ray theory and full-wave synthetics. *Geophysical Journal International*, *207*(1), 12–28. https://doi.org/10.1093/gji/ggw254

Borgeaud, A. F. E., Kawai, K., Konishi, K., & Geller, R. J. (2017). Imaging paleoslabs in the D" layer beneath Central America and the Caribbean using seismic waveform inversion. *Science Advances*, *3*(11). https://doi.org/10.1126/sciadv.1602700

Bozdag, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., et al. (2016). Global adjoint tomography : first-generation model. *Geophysical Journal International*, *207*,

1739–1766. https://doi.org/10.1093/gji/ggw356

Bullen, B. (1963). *Introduction to the theory of Seismology*. New York: Cambridge University Press.

Bunge, H., & Grand, S. P. (2000). Mesozoic plate-motion history below the northeast Pacific Ocean from seismic images of the subducted Farallon slab. *Nature*, *405*, 337–340.

Bunge, H., Richards, M. A., Lithgow-bertelloni, C., Baumgardner, J. R., Grand, S. P., & Romanowicz, B. A. (1998). Time Scales and Heterogeneous Structure in Geodynamic Earth Models. *Science*, *280*(5360), 91–95. https://doi.org/DOI: 10.1126/science.280.5360.91

Burdick, S., Vernon, F. L., Martynov, V., Eakins, J., Cox, T., Tytell, J., et al. (2016). Model Update May 2016 : Upper-Mantle Heterogeneity beneath North America from Travel-Time Tomography with Global and USArray Data. *Seismological Research Letters*, *88*(2A), 319–325. https://doi.org/10.1785/0220160186

Burkett, E. R., & Billen, M. I. (2010). Three-dimensionality of slab detachment due to ridge-trench collision: Laterally simultaneous boudinage versus tear propagation. *Geochemistry, Geophysics, Geosystems*, *11*(11). https://doi.org/10.1029/2010GC003286

Capitanio, F. A., Morra, G., & Goes, S. (2007). Dynamic models of downgoing plate-buoyancy driven subduction: Subduction motions and energy dissipation. *Earth and Planetary Science Letters*, *262*(1–2), 284–297. https://doi.org/10.1016/j.epsl.2007.07.039

Chen, M., Niu, F., Liu, Q., Tromp, J., & Zheng, X. (2015). Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons. *Journal of Geophysical Research: Solid Earth*, *120*, 1762–1786. https://doi.org/10.1002/2014JB011638.Received

Chen, P., Jordan, T. H., & Zhao, L. (2007). Full three-dimensional tomography: A comparison between the scattering-integral and adjoint-wavefield methods. *Geophysical Journal International*, *170*(1), 175–181. https://doi.org/10.1111/j.1365-246X.2007.03429.x

Christensen, U. R. (1996). The influence of trench migration on slab penetration into the lower mantle. *Earth and Planetary Science Letters*, *140*(1–4), 27–39.

Christensen, U. R., & Yuen, D. A. (1985). Layered convection induced by phase transitions. *Journal of Geophysical Research: Solid Earth*, *90*(B12), 10291–10300. https://doi.org/10.1029/JB090iB12p10291

Cobden, L., & Thomas, C. (2013). The origin of D" reflections: a systematic study of seismic array data sets. *Geophysical Journal International*, *194*(2), 1091–1118. https://doi.org/10.1093/gji/ggt152

Colombi, A., Nissen-meyer, T., Boschi, L., & Giardini, D. (2014). Seismic waveform inversion for core–mantle boundary topography. *Geophysical Journal International*, *198*, 55–71. https://doi.org/10.1093/gji/ggu112

Cottaar, S., Heister, T., Rose, I., & Unterborn, C. (2014). BurnMan: A lower mantle mineral physics toolkit. *Geochemistry, Geophysics, Geosystems*, *15*(4), 1164–1179. https://doi.org/10.4088/JCP.11m07343

Couch, R., & Woodcock, S. (1981). Gravity and structure of the continental margins of southwestern Mexico and northwestern Guatemala. *Journal of Geophysical Research: Solid Earth*, *86*(B3), 1829–1840. https://doi.org/10.1029/JB086iB03p01829

Creager, K. C., & Jordan, T. H. (1984). Slab penetration into the lower mantle. *Journal of Geophysical Research: Solid Earth*, *89*(B5), 3031–3049. https://doi.org/doi.org/10.1029/JB089iB05p03031

Creager, K. C., & Jordan, T. H. (1986). Slab penetration into the lower mantle beneath the Mariana and other island arcs of the northwest Pacific. *Journal of Geophysical Research: Solid Earth*, *91*(B3), 3573–3589. https://doi.org/10.1029/JB091iB03p03573

Crotwell, H. P., Owens, T. J., & Ritsema, J. (1999). The TauP Toolkit: flexible seismic travel-time and ray-path utilities. *Seismological Research Letters*, *70*(2), 154–160. https://doi.org/https://doi.org/10.1785/gssrl.70.2.154

Dahlen, F. A. (2005). Finite-frequency sensitivity kernels for boundary topography perturbations. *Geophysical Journal International*, *162*(1), 525–540. https://doi.org/10.1111/j.1365-246X.2005.02682.x

Dahlen, F. A., Hung, S., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes—I. Theory. *Geophysical Journal International*, *141*(1), 157–174. https://doi.org/10.1007/978-3-319-06540-3

Davies, D. R., Goes, S., Davies, J. H., Schuberth, B. S. A., Bunge, H., & Ritsema, J. (2012). Reconciling dynamic and seismic models of Earth's lower mantle : The dominant role of thermal heterogeneity. *Earth and Planetary Science Letters*, *253–269*, 353–354. https://doi.org/10.1016/j.epsl.2012.08.016

Deschamps, F., Rogister, Y., & Tackley, P. J. (2018). Constraints on core-mantle boundary topography from models of thermal and thermochemical convection. *Geophysical Journal International*, *212*(1), 164–188. https://doi.org/10.1093/gji/ggx402

Domeier, M., Doubrovine, P. V, Torsvik, T. H., Spakman, W., & Bull, A. L. (2016). Global correlation of lower mantle structure and past subduction. *Geophysical Research Letters*, *43*(10), 4645–4953. https://doi.org/https://doi.org/10.1002/2016GL068827

Durand, S., Thomas, C., & Jackson, J. M. (2018). Constraints on D" beneath the North Atlantic

region from P and S travel times and amplitudes. *Geophysical Journal International*, *216*(2), 1132–1144. https://doi.org/10.1093/gji/ggy476

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, *25*(4), 297–356. https://doi.org/10.1016/0031-9201(81)90046-7

Dziewonski, A. M., Hager, B. H., & O'Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *Journal of Geophysical Research*, *82*(2), 239–255. https://doi.org/10.1029/JB082i002p00239

Ekström, G., Nettles, M., & Dziewonski, A. M. (2012). The global CMT project 2004 – 2010 : Centroid-moment tensors for 13,1017 earthquakes. *Physics of the Earth and Planetary Interiors*, *200–201*, 1–9. https://doi.org/10.1016/j.pepi.2012.04.002

Engdahl, E. R., van der Hilst, R. D., & Buland, R. (1998). Global teleseismic earthquake relocation with improved travel times and procedures for depth determination. *Bulletin of the Seismological Society of America*, *88*(3), 722–743. https://doi.org/10.1130/0-8137-2349-3.461

Engebretson, D. C., Kelley, K. P., Cashman, H. J., & Richards, M. A. (1992). 180 million years of subduction. *GSA Today*.

Fei, Y., Li, J., Hirose, K., Minarik, W., Van Orman, J., Sanloup, C., et al. (2004). A critical evaluation of pressure scale at high temperatures by in situ X-ray diffraction measurements. *Physics of the Earth and Planetary Interiors*, *143–144*, 515–526. https://doi.org/10.1016/j.pepi.2003.09.018

Ferrari, L. (2004). Slab detachment control on mafic volcanic pulse and mantle heterogeneity in central Mexico. *Geology*, *32*(1), 77–80. https://doi.org/10.1130/G19887.1

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H. (2008). Theoretical background for continental- and global-scale full-waveform inversion in the time–frequency domain. *Geophysical Journal International*, *175*(2), 665–685. https://doi.org/10.1111/j.1365-246X.2008.03923.x

Forte, A. M., & Mitrovica, J. X. (2001). Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data. *Nature*, *410*, 1049–1056.

French, S. W., & Romanowicz, B. A. (2014). Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography. *Geophysical Journal International*, *199*(3), 1303–1327. https://doi.org/10.1093/gji/ggu334

French, S. W., & Romanowicz, B. A. (2015). Broad plumes rooted at the base of the Earth's mantle beneath major hotspots. *Nature*, *525*, 95–99. https://doi.org/10.1038/nature14876

Frost, D. A., & Rost, S. (2014). The P-wave boundary of the Large-Low Shear Velocity Province beneath the Pacific. *Earth and Planetary Science Letters*, *403*, 380–392. https://doi.org/10.1016/j.epsl.2014.06.046

Fuji, N., Kawai, K., & Geller, R. J. (2010). A methodology for inversion of broadband seismic waveforms for elastic and anelastic structure and its application to the mantle transition zone beneath the Northwestern Pacific. *Physics of the Earth and Planetary Interiors*, *180*(3–4), 118–137. https://doi.org/10.1016/j.pepi.2009.10.004

Fukao, Y., & Obayashi, M. (2013). Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity. *Journal of Geophysical Research: Solid Earth*, *118*, 5920–5938. https://doi.org/10.1002/2013JB010466

Fukao, Y., Obayashi, M., Inoue, H., & Nenbai, M. (1992). Subducting slabs stagnant in the mantle transition zone. *Journal of Geophysical Research: Solid Earth*, *97*(B4), 4809–

4822. https://doi.org/10.1029/91JB02749

Fukao, Y., Widiyantoro, S., & Obayashi, M. (2001). Stagnant slabs in the upper and lower mantle transition region. *Reviews of Geophysics*, *39*(3), 291–323. https://doi.org/https://doi.org/10.1029/1999RG000068

Gao, W., Matzel, E., & Grand, S. P. (2006). Upper mantle seismic structure beneath eastern Mexico determined from P and S waveform inversion and its implications. *Journal of Geophysical Research: Solid Earth*, *111*(B08307), 1–23. https://doi.org/10.1029/2006JB004304

Garnero, E. J., & McNamara, A. K. (2008). Structure and Dynamics of Earth's Lower Mantle. *Science*, *320*(5876), 626–628. https://doi.org/10.1126/science.1148028

Garnero, E. J., Mcnamara, A. K., & Shim, S. (2016). Continent-sized anomalous zones with low seismic velocity at the base of Earth's mantle. *Nature Geoscience*, *9*, 481–489. https://doi.org/10.1038/ngeo2733

Geller, R. J., & Hara, T. (1993). Two efficient algorithms for iterative linearized inversion of seismic waveform data. *Geophysical Journal International*, *115*(3), 699–710. https://doi.org/10.1111/j.1365-246X.1993.tb01488.x

Goes, S., Capitanio, F. A., & Morra, G. (2008). Evidence of lower-mantle slab penetration phases in plate motions. *Nature*, *451*, 981–984. https://doi.org/10.1038/nature06691

Gorbatov, A., & Fukao, Y. (2005). Tomographic search for missing link between the ancient Farallon subduction and the present Cocos subduction. *Geophysical Journal International*, *160*(3), 849–854. https://doi.org/10.1111/j.1365-246X.2005.02507.x

Grand, S. P. (1987). Tomographic Inversion for Shear Velocity Beneath the North American Plate. *Journal of Geophysical Research: Solid Earth*, *92*(B13), 14065–14090.

https://doi.org/198710.1029/JB092iB13p14065

Grand, S. P. (1994). Mantle shear structure beneath the Americas and surrounding oceans. *Journal of Geophysical Research: Solid Earth*, *99*(B6), 11591–11621. https://doi.org/10.1029/94JB00042

Grand, S. P. (2002). Mantle shear-wave tomography and the fate of subducted slabs. *Philosophical Transactions of the Royal Society of London A*, *360*(1800), 2475–2491. https://doi.org/doi: 10.1098/rsta.2002.1077

Grand, S. P., van der Hilst, R. D., & Widiyantoro, S. (1997). Global seismic tomography: A snapshot of convection in the Earth. *GSA Today*, *7*(4), 1–7. https://doi.org/10.1130/GSAT01707GW.1

Gréaux, S., Irifune, T., Higo, Y., Tange, Y., Arimoto, T., Liu, Z., & Yamada, A. (2019). Sound velocity of CaSiO3 perovskite suggests the presence of basaltic crust in the Earth's lower mantle. *Nature*, *565*(7738), 218–221. https://doi.org/10.1038/s41586-018-0816-5

Griffiths, R. W., Hackney, R. I., & van der Hilst, R. D. (1995). A laboratory investigation of effects of trench migration on the descent of subducted slabs. *Earth and Planetary Science Letters*, *133*(1–2), 1–17.

Hager, B. H. (1984). Subducted slabs and the geoid: Constraints on mantle rheology and flow. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 6003–6015. https://doi.org/10.1029/JB089iB07p06003

van der Hilst, R. D., & Engdahl, E. R. (1991). On ISC PP and pP data and their use in delay-time tomography of the Caribbean region. *Geophysical Journal International*, *106*(1), 169–188. https://doi.org/10.1111/j.1365-246X.1991.tb04610.x

van der Hilst, R. D., Widiyantoro, S., & Engdahl, E. R. (1997). Evidence for deep mantle

circulation from global tomography. *Nature*, *386*, 578–584. https://doi.org/10.1038/386578a0

Hirose, K., Sinmyo, R., & Hernlund, J. (2017). Perovskite in Earth's deep interior. *Science*, *358*(6364), 734–738. https://doi.org/10.1126/science.aam8561

Honda, S. (2017). Geodynamic modeling of the subduction zone around the Japanese Islands. *Monographs on Environment, Earth and Planets*, *5*(2), 35–62. https://doi.org/10.5047/meep.2017.00502.0035

Hung, S., Dahlen, F. A., & Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes—II. Examples. *Geophysical Journal International*, *141*(1), 175–203. https://doi.org/10.1007/978-3-319-06540-3

Hung, S., Garnero, E. J., Chiao, L., Kuo, B., & Lay, T. (2005). Finite frequency tomography of D" shear velocity heterogeneity beneath the Caribbean. *Journal of Geophysical Research: Solid Earth*, *110*(B7), 1–20. https://doi.org/10.1029/2004JB003373

Hutko, A. R., Lay, T., Garnero, E. J., & Revenaugh, J. (2006). Seismic detection of folded, subducted lithosphere at the core-mantle boundary. *Nature*, *441*, 333–336. https://doi.org/10.1038/nature04757

Hutko, A. R., Lay, T., Revenaugh, J., & Garnero, E. J. (2008). Anticorrelated seismic velocity anomalies from post-pervoskite in the lowermost mantle. *Science*, *320*(5879), 1070–1074. https://doi.org/10.1029/2007JF000831

Iidaka, T., & Furukawa, Y. (1994). Double seismic zone for deep earthquakes in the Izu-Bonin subduction zone. *Science*, *263*(5150), 1116–1118. https://doi.org/10.1126/science.263.5150.1116

Inoue, H., Fukao, Y., Tanabe, K., & Ogata, Y. (1990). Whole mantle P-wave travel time

tomography. *Physics of the Earth and Planetary Interiors*, *59*(4), 294–328. https://doi.org/10.1016/0031-9201(90)90236-Q

Irifune, T. (1993). Phase transformations in the earth's mantle and subducting slabs: Implications for their compositions, seismic velocity and density structures and dynamics. *Island Arc*, *2*(2), 55–71. https://doi.org/10.1111/j.1440-1738.1993.tb00074.x

Johnston, S. T., & Borel, G. D. (2007). The odyssey of the Cache Creek terrane, Canadian Cordillera: Implications for accretionary orogens, tectonic setting of Panthalassa, the Pacific superwell, and break-up of Pangea. *Earth and Planetary Science Letters*, *253*(3–4), 415–428. https://doi.org/10.1016/j.epsl.2006.11.002

Jordan, T. H. (1977). Lithospheric slab penetration into the lower mantle beneath the sea of Okhotsk. *Journal of Geophysical Research*, *43*, 473–496.

Kaneshima, S. (2016). Seismic scatterers in the mid-lower mantle. *Physics of the Earth and Planetary Interiors*, *257*, 105–114. https://doi.org/10.1016/j.pepi.2017.09.007

Kaneshima, S., & Helffrich, G. (2010). Small scale heterogeneity in the mid-lower mantle beneath the circum-Pacific area. *Physics of the Earth and Planetary Interiors*, *183*(1–2), 91–103. https://doi.org/10.1016/j.pepi.2010.03.011

Karato, S.-I., Riedel, M. R., & Yuen, D. A. (2001). Rheological structure and deformation of subducted slabs in the mantle transition zone: Implications for mantle circulation and deep earthquakes. *Physics of the Earth and Planetary Interiors*, *127*(1–4), 83–108. https://doi.org/10.1016/S0031-9201(01)00223-0

Karato, S. (1993). Importance of anelasticity in the interpretation of seismic tomography. *Geophysical Research Letters*, *20*(15), 1623–1626. https://doi.org/10.1029/93GL01767

Karato, S., & Karki, B. B. (2001). Origin of lateral variation of seismic wave velocities and

density in the deep mantle. *Journal of Geophysical Research: Solid Earth*, *106*(B10), 21771–21783. https://doi.org/10.1029/2001JB000214

Kawai, K., & Tsuchiya, T. (2009). Temperature profile in the lowermost mantle from seismological and mineral physics joint modeling. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(52), 22119–22123. https://doi.org/10.1073/pnas.0905920106

Kawai, K., & Tsuchiya, T. (2015). Elasticity of Continental Crust Around the Mantle Transition Zone. https://doi.org/10.1007/978-3-319-15627-9

Kawai, K., Takeuchi, N., & Geller, R. J. (2006). Complete synthetic seismograms up to 2 Hz for transversely isotropic spherically symmetric media. *Geophysical Journal International*, *164*(2), 411–424. https://doi.org/10.1111/j.1365-246X.2005.02829.x

Kawai, K., Takeuchi, N., Geller, R. J., & Fuji, N. (2007). Possible evidence for a double crossing phase transition in D″ beneath Central America from inversion of seismic waveforms. *Geophysical Research Letters*, *34*(9). https://doi.org/10.1029/2007GL029642

Kawai, K., Konishi, K., Geller, R. J., & Fuji, N. (2014). Methods for inversion of body-wave waveforms for localized three-dimensional seismic structure and an application to D" structure beneath Central America. *Geophysical Journal International*, *197*(1), 495–524. https://doi.org/10.1093/gji/ggt520

Kawakatsu, H., & Niu, F. (1994). Seismic evidence for a 920-km discontinuity in the mantle. *Nature*, *371*, 301–305. https://doi.org/10.1038/371301a0

Kendall, J.-M., & Nangini, C. (1996). Lateral variations in D″ below the Caribbean. *Geophysical Research Letters*, *23*(4), 399–402. https://doi.org/10.1029/95GL02659

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in

the Earth from traveltimes. *Geophysical Journal International*, *122*(1), 108–124. https://doi.org/10.1111/j.1365-246X.1995.tb03540.x

Kennett, B. L. N., Widiyantoro, S., & van der Hilst, R. D. (1998). Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle. *Journal of Geophysical Research*, *103*(B6), 12469–12493. https://doi.org/10.1029/98JB00150

King, S. D., Frost, D. J., & Rubie, D. C. (2015). Why cold slabs stagnate in the transition zone. *Geology*, *43*(3), 231–234. https://doi.org/10.1130/G36320.1

Kito, T., Thomas, C., Rietbrock, A., Garnero, E. J., Nippress, S. E. J., & Heath, A. E. (2008). Seismic evidence for a sharp lithospheric base persisting to the lowermost mantle beneath the Caribbean. *Geophysical Journal International*, *174*(3), 1019–1028. https://doi.org/10.1111/j.1365-246X.2008.03880.x

Koelemeijer, P., Schuberth, B. S. A., Davies, D. R., Deuss, A., & Ritsema, J. (2018). Constraints on the presence of post-perovskite in Earth's lowermost mantle from tomographic-geodynamic model comparisons. *Earth and Planetary Science Letters*, *494*, 226–238. https://doi.org/10.1016/j.epsl.2018.04.056

Komatitsch, D., Vilotte, J. P., Tromp, J., Afanasiev, M., Bozda, E., Charles, J., et al. (2015). SPECFEM3D Globe v7.0.0 [software], Computational Infrastructure for Geodynamics.

Konishi, K., Kawai, K., Geller, R. J., & Fuji, N. (2009). MORB in the lowermost mantle beneath the western Pacific : Evidence from waveform inversion. *Earth and Planetary Science Letters*, *278*(3–4), 219–225. https://doi.org/10.1016/j.epsl.2008.12.002

Konishi, K., Kawai, K., Geller, R. J., & Fuji, N. (2014). Waveform inversion for localized three-dimensional seismic velocity structure in the lowermost mantle beneath the Western Pacific. *Geophysical Journal International*, *199*(2), 1245–1267.

https://doi.org/10.1093/gji/ggu288

Labrosse, S., Hernlund, J. W., & Coltice, N. (2007). A crystallizing dense magma ocean at the base of the Earth's mantle. *Nature*, *450*(7171), 866–869. https://doi.org/10.1038/nature06355

Lay, T., & Garnero, E. J. (2011). Deep Mantle Seismic Modeling and Imaging. *Annual Review of Earth and Planetary Sciences*, *39*(1), 91–123. https://doi.org/10.1146/annurev-earth-040610-133354

Leat, P. T., & Larter, R. D. (2003). Intra-oceanic subduction systems: introduction. *Geological Society, London, Special Publications*, *219*, 1–17. https://doi.org/10.1016/j.jfda.2014.01.026

van der Lee, S., & Nolet, G. (1997a). Seismic image of the subducted trailing fragments of the Farallon plate. *Nature*, *386*, 266–269. https://doi.org/10.1038/386266a0

van der Lee, S., & Nolet, G. (1997b). Upper mantle S velocity structure of North America. *Journal of Geophysical Research: Solid Earth*, *102*(B10), 22815–22838. https://doi.org/10.1029/97JB01168

van der Lelij, R., Spikings, R., Ulianov, A., Chiaradia, M., & Mora, A. (2016). Palaeozoic to Early Jurassic history of the northwestern corner of Gondwana, and implications for the evolution of the Iapetus, Rheic and Pacific Oceans. *Gondwana Research*, *31*, 271–294. https://doi.org/10.1016/j.gr.2015.01.011

Li, C., Van Der Hilst, R. D., Engdahl, E. R., & Burdick, S. (2008). A new global model for P wave speed variations in Earth's mantle. *Geochemistry, Geophysics, Geosystems*, *9*(5). https://doi.org/10.1029/2007GC001806

Li, X.-D., & Romanowicz, B. A. (1995). Comparison of global waveform inversions with and

without considering cross-branch modal coupling. *Geophysical Journal International*, *121*(3), 695–709. https://doi.org/10.1111/j.1365-246X.1995.tb06432.x

Manea, M., Manea, V. C., Ferrari, L., Kostoglodov, V., & Bandy, W. L. (2005). Tectonic evolution of the Tehuantepec Ridge. *Earth and Planetary Science Letters*, *238*(1–2), 64–77. https://doi.org/10.1016/j.epsl.2005.06.060

Maruyama, S., Santosh, M., & Zhao, D. (2007). Superplume, supercontinent, and post-perovskite: Mantle dynamics and anti-plate tectonics on the Core-Mantle Boundary. *Gondwana Research*, *11*(1–2), 7–37. https://doi.org/10.1016/j.gr.2006.06.003

Masters, G., Laske, G., Bolton, H., & Dziewonski, A. M. (2000). The relative behavior of shear velocity, bluk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure. *Geophysical Monograph*, *117*, 63–87. https://doi.org/https://doi.org/10.1029/GM117p0063

van der Meer, D. G., Spakman, W., van Hinsbergen, D. J. J., Amaru, M. L., & Torsvik, T. H. (2010). Towards absolute plate motions constrained by lower-mantle slab remnants. *Nature Geoscience*, *3*(1), 36–40. https://doi.org/10.1038/ngeo708

van der Meer, D. G., van Hinsbergen, D. J. J., & Spakman, W. (2018). Atlas of the underworld: Slab remnants in the mantle, their sinking history, and a new outlook on lower mantle viscosity. *Tectonophysics*, *723*, 309–448. https://doi.org/10.1016/j.tecto.2017.10.004

Müller, R. D., Seton, M., Zahirovic, S., Williams, S. E., Matthews, K. J., Wright, N. M., et al. (2016). Ocean Basin Evolution and Global-Scale Plate Reorganization Events Since Pangea Breakup. *Annual Review of Earth and Planetary Sciences*, *44*(1), 107–138. https://doi.org/10.1146/annurev-earth-060115-012211

Mulyukova, E., Steinberger, B., Dabrowski, M., & Sobolev, S. V. (2015). Survival of LLSVPs

for billions of years in a vigorously convecting mantle : Replenishment and destruction of chemical anomaly. *Journal of Geophysical Research: Solid Earth*, *120*(5), 3824–3847. https://doi.org/10.1002/2014JB011688.We

Murakami, M., Hirose, K., Kawamura, K., Sata, N., & Yasuo, O. (2004). Post-Perovskite Phase Transition in MgSi03. *Science*, *304*(5672), 855–858. https://doi.org/DOI: 10.1126/science.1095932

Nakagawa, T., & Tackley, P. J. (2015). Influence of plate tectonic mode on the coupled thermochemical evolution of Earth's mantle and core. *Geochemistry, Geophysics, Geosystems*, *16*(10), 3400–3413. https://doi.org/10.1002/2015GC005996.Received

Nolet, G. (1990). Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs. *Journal of Geophysical Research: Solid Earth*, *95*(B6), 8499–8512. https://doi.org/10.1029/JB095iB06p08499

Nolet, G., & Dahlen, F. A. (2000). Wave front healing and the evolution of seismic delay times. *Journal of Geophysical Research: Solid Earth*, *105*(B8), 19043–19054. https://doi.org/https://doi.org/10.1029/2000JB900161

Obayashi, M., Yoshimitsu, J., Nolet, G., Fukao, Y., Shiobara, H., Sugioka, H., et al. (2013). Finite frequency whole mantle P wave tomography: Improvement of subducted slab images. *Geophysical Research Letters*, *40*(21), 5652–5657. https://doi.org/10.1002/2013GL057401

Pardo, M., & Suárez, G. (1995). Shape of the subducted Rivera and Cocos plates in southern Mexico: Seismic and tectonic implications. *Journal of Geophysical Research: Solid Earth*, *100*(B7), 12357–12373. https://doi.org/10.1029/95JB00919

Quinteros, J., Sobolev, S. V., & Popov, A. A. (2010). Viscosity in transition zone and lower

mantle: Implications for slab penetration. *Geophysical Research Letters*, *37*(9). https://doi.org/10.1029/2010GL043140

Rawlinson, N., & Spakman, W. (2016). On the use of sensitivity tests in seismic tomography. *Geophysical Journal International*, *205*(2), 1221–1243. https://doi.org/10.1093/gji/ggw084

Ren, Y., Stutzmann, E., van der Hilst, R. D., & Besse, J. (2007). Understanding seismic heterogeneities in the lower mantle beneath the Americas from seismic tomography and plate tectonic history. *Journal of Geophysical Research: Solid Earth*, *112*(B1). https://doi.org/10.1029/2005JB004154

Richards, M. A., & Engebretson, D. C. (1992). Large-scale mantle convection and the history of subduction. *Nature*, *355*, 437–440. https://doi.org/10.1038/350055a0

Ringwood, A. E. (1989). Significance of the terrestrial Mg/Si ratio. *Earth and Planetary Science Letters*, *95*(1–2), 1–7. https://doi.org/10.1016/0012-821X(89)90162-3

Ringwood, A. E., & Irifune, T. (1988). Nature of the 650-km seismic discontinuity: implications for mantle dynamics and differentiation. *Nature*, *331*, 131–136. https://doi.org/10.1038/332141a0

Ritsema, J., Deuss, A., Van Heijst, H. J., & Woodhouse, J. H. (2011). S40RTS: A degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. *Geophysical Journal International*, *184*(3), 1223–1236. https://doi.org/10.1111/j.1365-246X.2010.04884.x

Rogers, J. J. W., & Santosh, M. (2004). *Continents and Supercontinents* (Oxford Uni). New York.

Rogers, R. D., Karason, H., & van der Hilst, R. D. (2002). Epeirogenic uplift above a detached

slab in northern Central America. *Geology*, *30*(11), 1031–1034. https://doi.org/10.1130/0091-7613(2002)030<1031:EUAADS>2.0.CO;2

Romanowicz, B. A. (1991). Seismic tomography of the Earth's mantle. *Annual Review of Earth and Planetary Sciences*, *19*, 77–99. https://doi.org/10.1146/annurev.ea.19.050191.000453

Romanowicz, B. A. (2003). Global mantle tomography: Progess status in the past 10 years. *Annual Review of Earth and Planetary Sciences*, *31*(1), 303–328. https://doi.org/10.1146/annurev.earth.31.091602.113555

Schuberth, B. S. A., & Bunge, H. P. (2009). Tomographic filtering of high-resolution mantle circulation models: Can seismic heterogeneity be explained by temperature alone? *Geochemistry, Geophysics, Geosystems*, *10*(5). https://doi.org/10.1029/2009GC002401

Sdrolias, M., & Müller, R. D. (2006). Controls on back-arc basin formation. *Geochemistry, Geophysics, Geosystems*, *7*(4). https://doi.org/10.1029/2005GC001090

Seno, T., & Maruyama, S. (1984). Paleogeographic reconstruction and origin of the Philippine Sea. *Tectonics*, *102*(1–4), 53–84. https://doi.org/10.1016/0040-1951(84)90008-8

Shephard, G. E., Matthews, K. J., Hosseini, K., & Domeier, M. (2017). On the consistency of seismically imaged lower mantle slabs. *Scientific Reports*, *7*(10976). https://doi.org/10.1038/s41598-017-11039-w

Sidorin, I. (1999). Evidence for a Ubiquitous Seismic Discontinuity at the Base of the Mantle. *Science*, *286*(5443), 1326–1331. https://doi.org/10.1126/science.286.5443.1326

Sigloch, K. (2011). Mantle provinces under North America from multifrequency P wave tomography. *Geochemistry, Geophysics, Geosystems*, *12*(2). https://doi.org/10.1029/2010GC003421

Sigloch, K., & Mihalynuk, M. G. (2013). Intra-oceanic subduction shaped the assembly of Cordilleran North America. *Nature*, *496*(7443), 50–56. https://doi.org/10.1038/nature12019

Simmons, N. A., Myers, S. C., Johannesson, G., Matzel, E., & Grand, S. P. (2015). Evidence for long-lived subduction of an ancient tectonic plate beneath the southern Indian Ocean. *Geophysical Research Letters*, *42*(21), 9270–9278. https://doi.org/doi:10.1002/2015GL066237

Song, T. R. A., Helmberger, D. V., & Grand, S. P. (2004). Low-velocity zone atop the 410-km seismic discontinuity in the northwestern United States. *Nature*, *427*(6974), 530–533. https://doi.org/10.1038/nature02231

Stevenson, D. J. (1981). Models of the Earth's Core. *Science*, *214*(4521), 611–619. https://doi.org/10.1126/science.214.4521.611

Stixrude, L., Lithgow-Bertelloni, C., Kiefer, B., & Fumagalli, P. (2007). Phase stability and shear softening in CaSi O3 perovskite at high pressure. *Physical Review B - Condensed Matter and Materials Physics*, *75*(2). https://doi.org/10.1103/PhysRevB.75.024108

Su, W., Woodward, R. L., & Dziewonski, A. M. (1994). Degree 12 model of shear velocity heterogeneity in the mantle. *Journal of Geophysical Research: Solid Earth*, *99*(B4), 6945–6980. https://doi.org/10.1029/93JB03408

Suzuki, Y., Kawai, K., Geller, R. J., Borgeaud, A. F. E., & Konishi, K. (2016). Waveform inversion for 3-D S-velocity structure of D" beneath the Northern Pacific: Possible evidence for a remnant slab and a passive plume. *Earth, Planets and Space*, *68*(1). https://doi.org/10.1186/s40623-016-0576-0

Tackley, P. J. (2011). Living dead slabs in 3-D: The dynamics of compositionally-stratified

slabs entering a “slab graveyard” above the core-mantle boundary. *Physics of the Earth and Planetary Interiors*, *188*(3–4), 150–162. https://doi.org/10.1016/j.pepi.2011.04.013

Takeuchi, N. (2007). Whole mantle SH velocity model constrained by waveform inversion based on three-dimensional Born kernels. *Geophysical Journal International*, *169*(3), 1153–1163. https://doi.org/10.1111/j.1365-246X.2007.03405.x

Takeuchi, N., Kawakatsu, H., Tanaka, S., Obayashi, M., Chen, Y. J., Ning, J., et al. (2014). Upper mantle tomography in the northwestern Pacific region using triplicated P waves. *Journal of Geophysical Research: Solid Earth*, *119*(10), 7667–7685. https://doi.org/10.1002/2014JB011161

Tao, K., Grand, S. P., & Niu, F. (2017). Full-waveform inversion of triplicated data using a normalized-correlation-coefficient-based misfit function. *Geophysical Journal International*, *210*(3), 1517–1524. https://doi.org/10.1093/gji/ggx249

Tao, K., Grand, S. P., & Niu, F. (2018). Seismic structure of the upper mantle beneath eastern Asia from full waveform seismic tomography. *Geochemistry, Geophysics, Geosystems*, *19*(8), 2732–2763. https://doi.org/10.1029/2018GC007460

Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics*, *20*(2), 219–232. https://doi.org/10.1029/RG020i002p00219

Tateno, S., Hirose, K., & Ohishi, Y. (2014). Melting experiments on peridotite to lowermost mantle conditions. *Journal of Geophysical Research: Solid Earth*, *119*, 4684–4694. https://doi.org/10.1002/2014JB011176.Received

Thomas, C., Kendall, J.-M., & Lowman, J. (2004). Lower-mantle seismic discontinuities and the thermal morphology of subducted slabs. *Earth and Planetary Science Letters*, *225*(1–

2), 105–113. https://doi.org/10.1016/j.epsl.2004.05.038

Thurber, C. H., & Aki, K. (1987). Three-dimensional seismic imaging. *Annual Review of Earth and Planetary Sciences*, *15*, 115–139.

Tkalčić, H., Young, M., Muir, J. B., Davies, D. R., & Mattesini, M. (2015). Strong, multi-scale heterogeneity in Earth's lowermost mantle. *Scientific Reports*, *5*(18416). https://doi.org/10.1038/srep18416

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, *160*(1), 195–216. https://doi.org/10.1111/j.1365-246X.2004.02453.x

Tsuchiya, T. (2011). Elasticity of subducted basaltic crust at the lower mantle pressures: Insights on the nature of deep mantle heterogeneity. *Physics of the Earth and Planetary Interiors*, *188*(3–4), 142–149. https://doi.org/10.1016/j.pepi.2011.06.018

Tsuchiya, T., & Tsuchiya, J. (2006). Effect of impurity on the elasticity of perovskite and postperovskite: Velocity contrast across the postperovskite transition in (Mg,Fe,Al)(Si,Al)O3. *Geophysical Research Letters*, *33*(12), 10–13. https://doi.org/10.1029/2006GL025706

Tsuchiya, T., Tsuchiya, J., Umemoto, K., & Wentzcovitch, R. M. (2004). Phase transition in MgSiO3 perovskite in the Earth's lower mantle. *Earth and Planetary Science Letters*, *224*(3–4), 241–248. https://doi.org/10.1016/j.epsl.2004.05.017

Wang, X., Tsuchiya, T., & Hase, A. (2015). Computational support for a pyrolitic lower mantle containing ferric iron. *Nature Geoscience*, *8*(7), 556–559. https://doi.org/10.1038/ngeo2458

Waszek, L., Schmerr, N. C., & Ballmer, M. D. (2018). Global observations of reflectors in the

mid-mantle with implications for mantle structure and dynamics. *Nature Communications*, *9*(385). https://doi.org/10.1038/s41467-017-02709-4

Wentzcovitch, R. M., Karki, B. B., Cococcioni, M., & de Gironcoli, S. (2004). Thermoelastic Properties of MgSiO3-Perovskite: Insights on the Nature of the Earth's Lower Mantle. *Physical Review Letters*, *92*(1–9). https://doi.org/10.1103/PhysRevLett.93.126403

Wentzcovitch, R. M., Tsuchiya, T., & Tsuchiya, J. (2006). MgSiO3 postperovskite at D" conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3), 543–546. https://doi.org/10.1073/pnas.0506879103

Whittaker, S., Thorne, M. S., Schmerr, N. C., & Miyagi, L. (2015). Seismic array constraints on the D" discontinuity beneath Central America. *Journal of Geophysical Research: Solid Earth*, *121*(1), 152–169. https://doi.org/10.1002/2015JB012392.Received

Woodhouse, J. H., & Dziewonski, A. M. (1984). Mapping the upper mantle: Three-dimensional modeling of earth structure by inversion of seismic waveforms. *Journal of Geophysical Research: Solid Earth*, *89*(B7), 5953–5986. https://doi.org/10.1029/JB089iB07p05953

Xu, W., Lithgow-Bertelloni, C., Stixrude, L., & Ritsema, J. (2008). The effect of bulk composition and temperature on mantle seismic structure. *Earth and Planetary Science Letters*, *275*(1–2), 70–79. https://doi.org/10.1016/j.epsl.2008.08.012

Yamaya, L., Borgeaud, A. F. E., Kawai, K., Geller, R. J., & Konishi, K. (2018). Effects of redetermination of source time functions on the 3-D velocity structure inferred by waveform inversion. *Physics of the Earth and Planetary Interiors*, *282*, 117–143. https://doi.org/10.1016/j.pepi.2018.04.012

Zhang, S., Cottaar, S., Liu, T., Stackhouse, S., & Militzer, B. (2016). High-pressure,

temperature elasticity of Fe- and Al-bearing MgSiO3: Implications for the Earth's lower mantle. *Earth and Planetary Science Letters*, *434*, 264–273. https://doi.org/10.1016/j.epsl.2015.11.030

Zhao, D., & Ohtani, E. (2009). Deep slab subduction and dehydration and their geodynamic consequences: Evidence from seismology and mineral physics. *Gondwana Research*, *16*(3–4), 401–413. https://doi.org/10.1016/j.gr.2009.01.005

Zhou, Y. (2018). Anomalous mantle transition zone beneath the Yellowstone hotspot track. *Nature Geoscience*, *11*(6), 449–453. https://doi.org/10.1038/s41561-018-0126-4

Zhu, H., Bozdăg, E., & Tromp, J. (2015). Seismic structure of the European upper mantle based on adjoint tomography. *Geophysical Journal International*, *201*(1), 18–52. https://doi.org/10.1093/gji/ggu492

Zhu, H., Komatitsch, D., & Tromp, J. (2017). Radial anisotropy of the North American upper mantle based on adjoint tomography with USArray. *Geophysical Journal International*, *211*(1), 349–377. https://doi.org/10.1093/gji/ggx305

Appendix A List of Events

I show the list of events used to infer the structure of the D″ region in Table A1, and of the MTZ in Table A2. The source parameters are all as given in the GCMT catalog (Ekström et al., 2012). The duration of the triangle source time function is given in the column labeled "T (s)" in Table A2. The last two columns in Table A2 denote the variance (VR) for a given event with respect to the initial synthetics (for AK135), and the incremental variance reduction (i.e., the difference between the variance of synthetics for the initial, and final model; see section 2.1.3) using SEMUCB corrections.

Event #	Date (Y/M/D)	Longitude (°)	Latitude (°)	Depth (km)	Mw
1	1993/03/30	-67.3	-22.0	179.5	5.8
2	1993/05/24	-66.9	-23.5	231.9	6.9
3	1993/10/19	-65.7	-22.1	278.9	6.0
4	1994/01/10	-69.3	-13.3	603.6	6.9
5	1994/04/29	-63.2	-28.5	565.9	6.9
6	1994/05/10	-63.0	-28.6	603.0	6.9
7	1994/06/16	-70.4	-15.3	200.3	5.9
8	1994/08/19	-63.4	-26.7	562.6	6.4
9	1994/10/20	-70.8	-39.2	170.2	5.9
10	1994/12/07	-66.6	-23.2	246.3	5.9
11	1994/12/12	-69.7	-17.4	161.4	6.2
12	1995/02/14	-67.9	-23.6	159.5	5.9
13	1997/07/20	-66.1	-22.8	271.3	6.1
14	1997/11/28	-68.9	-13.7	600.5	6.6
15	1999/05/25	-66.5	-27.9	175.6	5.8
16	1999/09/15	-67.4	-20.7	217.5	6.4
17	2000/04/23	-63.0	-28.4	607.9	6.9
18	2000/04/23	-63.0	-28.4	610.4	6.1
19	2000/06/14	-67.1	-23.8	239.6	5.8
20	2001/06/19	-68.3	-23.0	152.6	5.9
21	2001/06/29	-66.4	-19.7	287.2	6.0
22	2001/10/26	-70.6	-14.7	221.7	5.8
23	2003/07/27	-65.2	-20.1	350.6	6.0
24	2004/03/17	-65.6	-21.2	297.0	6.1
25	2004/11/12	-63.2	-26.9	583.1	6.0
26	2004/11/13	-63.2	-26.8	582.0	5.8
27	2005/03/21	-63.5	-24.9	572.3	6.8
28	2005/03/21	-63.6	-24.7	572.2	6.4
29	2005/06/02	-67.2	-24.4	193.2	6.0
30	2005/11/17	-68.1	-22.5	155.3	6.8
31	2006/08/25	-67.2	-24.4	185.8	6.6
32	2006/09/22	-63.1	-26.9	602.4	6.0
33	2006/10/23	-65.8	-21.8	287.6	5.8

34	2006/11/13	-63.5	-26.1	573.4	6.8
35	2007/05/25	-67.4	-24.3	188.5	5.9
36	2007/07/21	-66.0	-22.3	280.2	6.4
37	2007/11/18	-66.5	-22.7	262.4	6.0
38	2008/09/03	-63.3	-26.9	571.3	6.3
39	2008/10/12	-65.2	-20.3	361.5	6.2
40	2009/07/12	-70.8	-15.3	197.1	6.1
41	2009/09/05	-70.7	-15.5	209.8	5.8
42	2009/09/30	-69.7	-15.7	257.8	5.9
43	2009/11/13	-64.2	-18.0	611.8	5.8
44	2009/11/14	-66.8	-23.0	221.2	6.2
45	2010/01/28	-67.0	-23.6	204.5	5.9
46	2010/02/12	-69.0	-33.6	173.5	5.8
47	2010/09/13	-71.1	-14.7	171.0	5.8
48	2011/01/01	-63.2	-27.0	586.0	7.0
49	2011/04/17	-63.1	-27.6	573.9	5.8
50	2011/06/08	-69.8	-17.4	150.4	5.9
51	2011/09/02	-63.1	-28.6	597.3	6.7
52	2011/11/22	-65.2	-15.4	553.8	6.6
53	2012/03/05	-63.3	-28.2	565.1	6.1
54	2012/05/28	-63.1	-28.3	591.6	6.7
55	2012/06/02	-63.6	-22.1	549.5	6.0
56	2012/11/22	-63.7	-22.9	544.4	5.9
57	2013/02/22	-63.0	-27.9	585.7	6.1
58	2014/09/24	-66.7	-23.8	227.6	6.2
59	2015/02/02	-67.1	-32.8	176.8	6.3
60	2015/02/11	-66.8	-23.1	223.0	6.7
61	2015/09/28	-66.9	-23.8	221.1	6.1

Table A1. Earthquakes used to infer the structure of the D″ region. The 35 events after 2005 (events 27-61) were all recorded at the USArray transportable and backbone broadband stations. Events 23-29 were recorded at the CANOE temporary network. Events before 2005 (events 1-26) were recorded at CNSN, CANOE, IRIS/USGS, SCSN, PNSN, and BDSN networks.

#	Event ID	Lat (°)	Lon (°)	Depth (km)	Mw	T (s)	VR AK135 (%)	Inc VR (%)
1	080704C	17.3	264.7	104.7	5.8	3.6	85.9	14.9
2	200503171337A	15.2	268.4	181.5	6.2	5.8	116.2	37.9
3	200510172159A	11.7	273.9	162.2	5.6	3.0	87.2	6.1
4	200703092101A	13.3	272.4	217.2	5.5	2.8	101.9	9.7
5	200707060109A	16.5	266.1	113.7	6.0	5.2	64.2	9.3
6	200707232230A	14.4	268.7	112.5	5.5	2.8	98.9	12.8
7	200901170257A	15.9	267.3	169.2	5.8	4.0	97.3	22.2
8	201011012316A	17.0	266.4	168.7	5.6	3.2	105.2	24.4
9	201102251307A	18.0	264.8	128.4	6.0	4.8	113.1	11.8
10	201104071311A	17.3	265.9	153.8	6.7	11.0	111.6	37.2
11	201111072235A	11.6	274.0	172.7	6.0	5.0	110.7	39.9
12	201209011801A	16.4	267.4	251.1	5.5	2.6	93.7	14.6
13	201303252302A	14.6	269.3	186.4	6.2	6.2	89.0	31.4
14	201404112029A	11.7	273.8	141.8	6.6	9.8	77.6	33.4
15	201405211006A	17.3	265.0	125.1	5.8	3.8	108.0	26.9
16	201407291046A	18.0	264.3	109.2	6.4	7.4	103.0	25.4
17	201410041916A	17.6	265.3	141.3	5.6	3.2	101.6	12.3
18	201501200659A	15.0	268.4	156.1	5.5	2.8	108.9	26.7
19	201504281856A	17.2	264.9	116.9	5.5	2.8	97.8	8.4
20	201508050913A	16.2	266.2	104.5	5.7	3.4	92.3	7.2

Table A2. Earthquakes used to infer the structure in the MTZ. “VR AK135” denotes the variance reduction for the initial synthetics. “Inc VR” denotes the incremental variance reduction between the initial synthetics, and (Born) synthetics for the final model using SEMUCB corrections (a).