

Joint inversion of PP and SS precursor waveforms and Rayleigh wave phase velocities for global mantle transition zone structure

F. Bissig¹, A. Khan^{1,2} and D. Giardini¹

¹*Institute of Geophysics, Department of Earth Sciences, ETH Zürich, 8092 Zürich, Switzerland. E-mail: felix.bissig@erdw.ethz.ch ; amir.khan@erdw.ethz.ch*

²*Physik Institut, Mathematisch-Naturwissenschaftliche Fakultät, Universität Zürich, 8006 Zürich, Switzerland*

Accepted 2202 November 22. Received 2022 October 15; in original form 2022 May 4

SUMMARY

We have compiled a new data set of global PP and SS precursor waveforms that we jointly invert in combination with fundamental-mode and higher-order Rayleigh-wave phase velocities for upper mantle and mantle transition zone (MTZ) structure. We observe clear S410S, S520S, S660S and P410P precursor arrivals, but not P660P, because of interfering phases. Traveltimes and amplitudes of precursor phases reflect a complex interplay of data and modelling factors, implying that MTZ structure is best resolved through direct inversion of waveforms. To model waveforms as accurately as possible, we account for effects arising from data processing, shallow structure, incoherent stacking, attenuation and source effects, among others. As part of the inversion, we consider two independent model parametrizations to obtain quantitative insights into the seismic and thermochemical constitution of the MTZ. These include a ‘classical’ seismic parametrization based on a layered seismic velocity structure and a thermodynamic parametrization, where seismic profiles are self-consistently built from mineral physics data. The results show lateral variations in thermal, compositional and discontinuity structure that partly correlate with tectonic setting. The mantle beneath continents and subduction zones is found to be colder in comparison to oceans and hotspots as reflected in MTZ thickness. In terms of composition, we find that subduction zones are enriched in basalt. Mid-MTZ structure shows a trend from simple sub-ocean single- to complex circum-Pacific subduction-zone-related dual-discontinuity structure—the possible signature of oceanic crustal transport to the MTZ. Statistical analysis indicates that a mechanically mixed mantle matches seismic data better than an equilibrated mantle across $\sim 2/3$ of the globe. Finally, while a large part of the seismic data can be matched by an iso-chemical and adiabatic mantle, complexities within the MTZ are not entirely captured by this assumption.

Key words: Composition and structure of the mantle; Body waves; Surface waves and free oscillations; Waveform inversion; Mantle discontinuities; PP and SS precursors.

1 INTRODUCTION

The mantle transition zone (MTZ), which separates the upper and lower mantle, exerts a fundamental control on the material flux across it, as evidenced in seismic tomography images that show subducting slabs stagnating in the lower MTZ at some locations and passing through at others (e.g. Fukao & Obayashi 2013). The MTZ is characterized by seismic discontinuities at global average depths of 410, 520 and 660 km, hereinafter referred to as ‘410’, ‘520’ and ‘660’, respectively. These discontinuities are attributed to transformations of olivine and its high-pressure polymorphs: olivine \rightarrow wadsleyite (410), wadsleyite \rightarrow ringwoodite

(520) and ringwoodite \rightarrow bridgmanite + ferropericlasite (660) (e.g. Ita & Stixrude 1992).

The pressure and temperature dependence of these mineral transformations can be used to determine lateral variations in mantle temperature by mapping topographic undulations on the discontinuities (e.g. Helffrich & Bina 1994; Ritsema *et al.* 2009; Khan *et al.* 2013; Waszek *et al.* 2021). In contrast, the absolute seismic velocity change at the discontinuities is a proxy for olivine content, and thus mantle composition (e.g. Agee 1998). Consequently, studying the MTZ and its bounding discontinuities in greater detail by seismological means allows for insights into mantle thermochemical structure and ultimately informs us of the processes by which

the mantle convects (e.g. Khan *et al.* 2009; Simmons *et al.* 2009; Afonso *et al.* 2019; Munch *et al.* 2020; Fullea *et al.* 2021; Bissig *et al.* 2022).

To characterize MTZ discontinuities, in particular the 410 and 660, a plethora of seismic waves have been used, including refracted waves (e.g. Niazi & Anderson 1965; Grand & Helmberger 1984; Tajima *et al.* 2009; Bissig *et al.* 2022), *P*-to-*s* and *S*-to-*p*-wave conversions (e.g. Langston 1979; Chevrot *et al.* 1999; Farra & Vinnik 2000; Tauzin *et al.* 2008; Munch *et al.* 2020; Bissig *et al.* 2021), ScS reverberations (e.g. Revenaugh & Jordan 1991) and precursors to underside reflected PKP waves (e.g. Engdahl & Flinn 1969; Benz & Vidale 1993), *P* waves (e.g. Flanagan & Shearer 1999; Rost & Weber 2002) and *S* waves (e.g. Shearer 1991; Flanagan & Shearer 1998).

The spatially most coherent long-wavelength images of global MTZ structure have been provided by *S* waves that bounce off the underside of the 410 and 660 and arrive as precursors to the SS phase. The SS precursors have typically been employed to estimate MTZ thickness and topography (e.g. Flanagan & Shearer 1998; Gu *et al.* 2003; Deuss 2009; Houser 2016; Guo & Zhou 2020; Waszek *et al.* 2021), which, in turn, can be used to track "cold" signatures associated with subduction zones. While information on discontinuity depths is extracted from precursor traveltimes, amplitudes are used to quantify the impedance contrasts at the respective discontinuities (e.g. Shearer & Flanagan 1999; Lawrence & Shearer 2006; Yu *et al.* 2018). Similarly, PP precursors reflected at the 410 have been used to map discontinuity topography (e.g. Flanagan & Shearer 1999) and reflectivity (e.g. Chambers *et al.* 2005a), although the generally more noisy character of PP precursors limits their usefulness. Additional precursor arrivals are generally interpreted as underside reflections at discontinuities within the upper mantle (e.g. Gu *et al.* 2001; Rost & Weber 2001), MTZ (e.g. Deuss & Woodhouse 2001; Tian *et al.* 2020) and lower mantle (e.g. Waszek *et al.* 2018; Saki *et al.* 2022).

Here, we build a new set of global PP and SS waveform stacks. In contrast to previous studies, where precursor traveltimes and amplitudes are usually analysed separately, we invert precursor waveforms to simultaneously resolve seismic velocity, discontinuity depth, and impedance contrast. Moreover, where precursor studies generally map traveltimes to discontinuity depth by relying on existing tomographic velocity models, we jointly invert precursor waveforms and surface wave data for upper mantle and MTZ structure. For this purpose, we follow a two-pronged approach: we first rely on a 'classical' seismic parametrization, after which, and independently thereof, we consider a thermodynamic parametrization to construct self-consistent seismic velocity models from mineral physics data following our earlier work (e.g. Khan *et al.* 2009, 2013; Munch *et al.* 2020; Bissig *et al.* 2022). This enables us to obtain insights on the thermochemical makeup of Earth's mantle from both direct inversion and from comparison of the two parametrizations.

This study is separated into a methods and an application part. We first detail the selection and processing of PP and SS precursor waveforms to compile a global data set (Section 2). The Rayleigh wave phase velocity data set is described in Section 3. Following this, we provide a detailed description of forward and inverse modelling (Sections 4 and 5) with particular emphasis on careful modelling of precursor waveforms without which inversion of the latter would not be possible. As part of the application (Sections 6 and 7), we present and discuss results from the joint inversion of the precursor waveform and surface-wave data sets, that provide us with improved information on the nature, composition, thermal state and seismic structure of the MTZ.

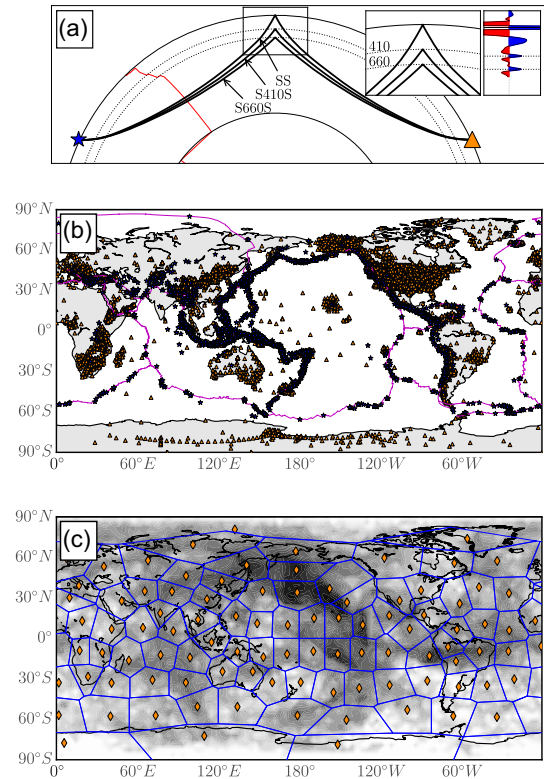


Figure 1. Overview of ray paths and location of events, stations, and bounce points of the precursor data set considered in this study. (a) Ray paths of the SS phase and underside reflections at the 410 (S410S) and 660 (S660S) for an event (blue star) at 35 km depth and a station (orange triangle) at 130° epicentral distance for model IASP91 (red profile; Kennett & Engdahl 1991). The insets show details of the bounce point region near the 410 and 660 discontinuities (dotted lines) and a synthetic SS waveform with prominent SS, S410S and S660S phases (blue wiggles; cf. dotted lines). (b) Global configuration of stations (orange triangles) and events (blue stars) used in the present study. Plate boundaries from USGS are plotted in magenta. (c) The grey-shaded map illustrates the logarithmic bounce point density, where white/black areas indicate low/high density, respectively. Plotted on top are the outlines of the Voronoi cells (blue grid) used for binning waveforms and the respective mean bounce point locations (orange).

2 PRECURSOR DATA PROCESSING

P and *S* waves are reflected at the free surface and arrive as PP and SS phases at the receiver. Analogously, sharp discontinuities in seismic properties, such as at 410 and 660 km depth, cause additional underside-reflected *P* and *S* phases, labelled PdP and SdS, respectively, where 'd' denotes the discontinuity depth. These reflections appear as weaker precursors to the free surface reflection on a seismogram. This is illustrated in Fig. 1(a), which shows SS and SdS ray paths and, in the inset, a synthetic waveform, indicating the arrivals (in blue) of S410S and S660S. A similar image is obtained for the case of PP waves, which shows that the ray paths of the precursors are very akin to those of the PP and SS waves in regions away from the bounce point. Therefore, the relative traveltimes and amplitudes of the precursors and the corresponding free surface reflection are, to first order, mainly sensitive to mantle structure close to the bounce point. The traveltime is affected by discontinuity depth and the velocity structure above. In contrast, the relative amplitudes primarily depend on the incidence angle and the impedance contrast at the discontinuity. The frequent occurrence

of earthquakes around the Pacific and the large number of seismic stations on nearby continents (Fig. 1b) result in a particularly good coverage of bounce points beneath the Pacific (Fig. 1c) and a reasonable global coverage of bounce points as also observed in earlier analyses (e.g. Shearer 1993; Gossler & Kind 1996; Flanagan & Shearer 1998; Gu & Dziewoński 2002; Chambers *et al.* 2005b; Houser 2016; Waszek *et al.* 2021).

2.1 Data selection

We consider three-component waveforms from 2616 events and 10 132 broad-band stations that are available from IRIS (Incorporated Research Institutions for Seismology; *cf.* Fig. 1b). Earthquake selection is based on the global centroid moment tensor catalogue (GCMT; Dziewoński *et al.* 1981; Ekström *et al.* 2012) and relies on the following criteria (*cf.* Deuss 2009):

- (i) Events occurred between 1990 and 2020.
- (ii) In order to generate strong signals, while reducing complexity and length of the source-time function, moment magnitude lies between 6.0 and 7.0.
- (iii) The focal depth is limited to values shallower than 75 km to decrease the influence of depth phases (i.e. pP and sS).
- (iv) The epicentral distance range (Δ) ranges from 80° – 180° .

2.2 Waveform processing

We first remove the instrument response from all waveforms to obtain displacement, after which the waveforms are rotated to the vertical-radial-transverse (ZRT) coordinate system, detrended and filtered between 1 and 75 s with a second-order Butterworth band-pass filter. The vertical and transverse components are used for analysis of PP and SS precursors, respectively. We trim PP and SS waveforms to a window spanning -200 to 100 s and -300 to 100 s, respectively, around the theoretical arrival time of the free surface reflection in IASP91 (Kennett & Engdahl 1991). In order to normalize the source signature, we use iterative time-domain deconvolution, which is commonly employed for computing receiver functions (Ligorria & Ammon 1999). In doing so, we deconvolve the free surface reflection wavelet (± 50 s around its theoretical arrival time in IASP91) from the entire waveform (*cf.* Heit *et al.* 2010) and normalize the trace to the maximum amplitude within ± 20 s to the free surface reflection. Subsequently, we filter the waveforms in a passband from 15 to 50/60 s period for PP/SS, respectively. While PP waveforms are sometimes filtered at shorter periods (e.g. Deuss 2009), here we focus on longer periods in order to reduce computational costs in the inversion. We correct for the differential moveout between precursor arrivals by stretching and compressing the time axis of each trace to a reference slowness of 7.5 s deg^{-1} ($\Delta \sim 103^\circ$) for PP and 12.0 s deg^{-1} ($\Delta \sim 135^\circ$) for SS, respectively. Finally, waveforms are only retained for further consideration if the signal-to-noise ratio (SNR), i.e., the ratio of the maximum absolute amplitude within ± 30 s to the free surface reflection ('signal') and the maximum absolute amplitude within a window starting 50 s before the 660 precursor to 50 s after the 410 precursor ('noise'), is greater than 3.

2.3 Stacking and uncertainty estimate

Because the amplitudes of precursors are typically on the level of a few per cent compared to the free surface reflection, waveform stacking is used to enhance precursor signals. To do so, we

first collect waveforms in common depth point bins defined by the bounce point locations. There exist various approaches for the design of bin geometries, including overlapping circles of different radii (typically 10° ; e.g. Schmerr & Garnero 2006) or polygons to combine waveforms over regional scales (e.g. Chambers *et al.* 2005a). Here, we define a set of 96 Voronoi cells (*cf.* Fig. 1c) to bin data. Instead of using a grid of regularly distributed cells, we adjust the centre points such that the cell boundaries roughly follow the continent–ocean border in order to reduce effects from shallow structure during stacking. Alternatively, lateral variability could also be addressed through data-driven 'adaptive stacking' (Waszek *et al.* 2021), where Voronoi cells are iteratively adjusted so as to minimize stack uncertainty. Using 96 points results in an average cell-radius of ~ 1300 km, which represents a good compromise between spatial resolution and the computational cost of the inversion, which increases with the number of stacks considered.

For each cap, we stack all waveforms with bounce points located within it. Fig. 2 shows synthetic (a–b) and observed (c–d) PP and SS waveform sections before stacking, respectively. Topside-reflected waves and core phases (PKIKP, ScS and their precursors) are clearly seen to obscure pP and sS in the sections. Here, interference effects are reduced by (1) rejecting epicentral distances $\Delta > 140^\circ$ for PP and $\Delta > 160^\circ$ for SS waveforms, respectively and (2) suppressing waveform-portions containing interfering phases in the stacking process ('muting'). The muting windows are shown in Fig. S9 and are designed to cover time-distance windows associated with interferences observed in a wide range of 1-D models that are proposed in the inversion. Therefore, the choice of windows is more conservative than it would be on the basis of a global 1-D seismic reference model (e.g. IASP91). A quantitative assessment of the impact of interfering phases on the final stacks, that is traveltimes and amplitude of precursor phases, is shown in Fig. S1. The largest influence is seen for the S660S phase.

Finally, the waveform stack is convolved with a cosine-taper of 5 s-length to remove abrupt steps in the stack that are due to muting. To obtain an estimate of uncertainty, we use the bootstrap resampling approach (Efron & Tibshirani 1991): the uncertainty of a stack is equal to the standard deviation of 100 stacks formed by randomly recombining the pre-stack waveforms. In contrast to previous precursor studies, we neither apply a pre- nor post-stack traveltimes correction to account for crust and upper mantle velocities and instead invert for shallow structure by including surface wave data.

2.4 Analysis of precursor data set

In the following, we provide a description of our precursor data set in terms of traveltimes and amplitude and comparison to other precursor studies.

2.4.1 A systematic search for discontinuities

Figs 3(a) and (b) show the PP and SS precursor waveform stacks. To systematically search for reflections associated with positive velocity jumps, we pick traveltimes and amplitudes of all maxima in the stacks, which are shown in Figs 3(c) and (d).

SS waveforms. SS waveform stacks exhibit a series of robust maxima (Fig. 3c), that are also clearly visible in the waveforms (Fig. 3a; *cf.* green lines) and correspond to the underside reflections at the 410 and 660 around -160 and -230 s, respectively. A third arrival of smaller amplitude, yet persistent across the entire data set,

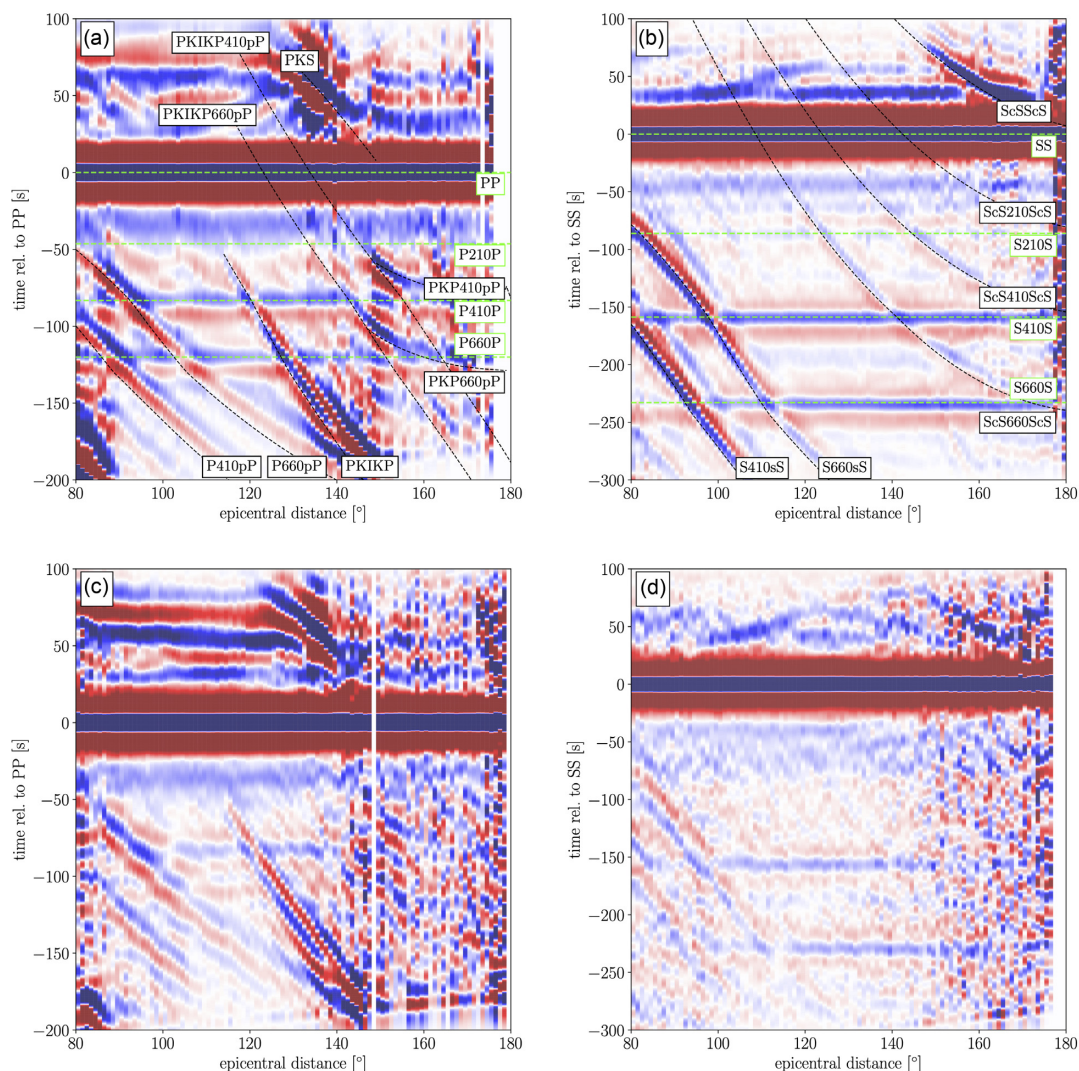


Figure 2. Record sections of synthetic (top panels) and observed (bottom panels) waveform stacks. Synthetic Instaseis (van Driel *et al.* 2015) PP (a) and SS waveform stacks (b), respectively, are computed for model IASP91 (Kennett & Engdahl 1991), 100 randomly picked events, and 10 stations per event. Partial stacks are computed as a function of epicentral distance in 1° -bins. Blue/red colours indicate positive/negative amplitudes, respectively. Traveltime curves for PP/SS and their underside reflections (precursors) are shown in green and interfering phases in black, respectively. Observed PP (c) and SS (d) waveform sections, respectively, illustrated for the cap with the largest number of bounce points (cap #22) without muting of interfering phases.

i.e., globally present, is associated with the mid-MTZ discontinuity at ~ 520 km depth (around -180 s). While the 520 is most likely associated with the dissociation of wadsleyite, which is expected to occur over a wider pressure range in comparison to the 410 and 660 (e.g., Akaogi *et al.* 1989), its global existence has been debated because its observation was restricted to regional scales (Gu *et al.* 1998) and it was only sparsely observed in short-period data (e.g. Benz & Vidale 1993; Chevrot *et al.* 1999). Yet, the recent SS precursor study by Tian *et al.* (2020) revealed the global existence of a discontinuity at 520 km depth. Discrepancies with previous studies are attributable to a larger data set and better sensitivity of long-period data to the 520.

In relation to the 520, we also observe more waveform complexities around the 520-arrival with a decrease in the number of traces. Similar observations were made by Deuss & Woodhouse (2001), who proposed the presence of two separate reflectors at 500–515 km and 551–566 km depth, respectively. The second discontinuity, referred to as ‘560’, has been attributed to the

garnet→perovskite transition, which, in contrast to the thermally sensitive wadsleyite→ringwoodite transition, is suggested to be governed by the Ca content (Saikia *et al.* 2008). Hence, its detection could be related to localized enrichment of subducted oceanic crust (Tian *et al.* 2020).

In addition to a complex mid-MTZ, additional maxima of small amplitudes are detected before the S660S-arrival, that may be related to mineralogical complexities of the 660 arising from the post-garnet transition (e.g. Deuss *et al.* 2006).

PP waveforms. PP precursor stacks (Figs 3b and d) show clear reflections at the 410 around -85 s, while the 520 around -100 s is less apparent but nevertheless visible. Many stacks show an increase in amplitude at the IASP91-predicted P660P-traveltime around -120 s, yet consistent detection of this precursor is difficult. The non-detection of the P660P in global precursor data sets is a well-known and somewhat puzzling issue (e.g. Shearer & Flanagan 1999; Rost & Weber 2002), since it is at odds with the predictions from mineral physics and seismic reference models of the MTZ.

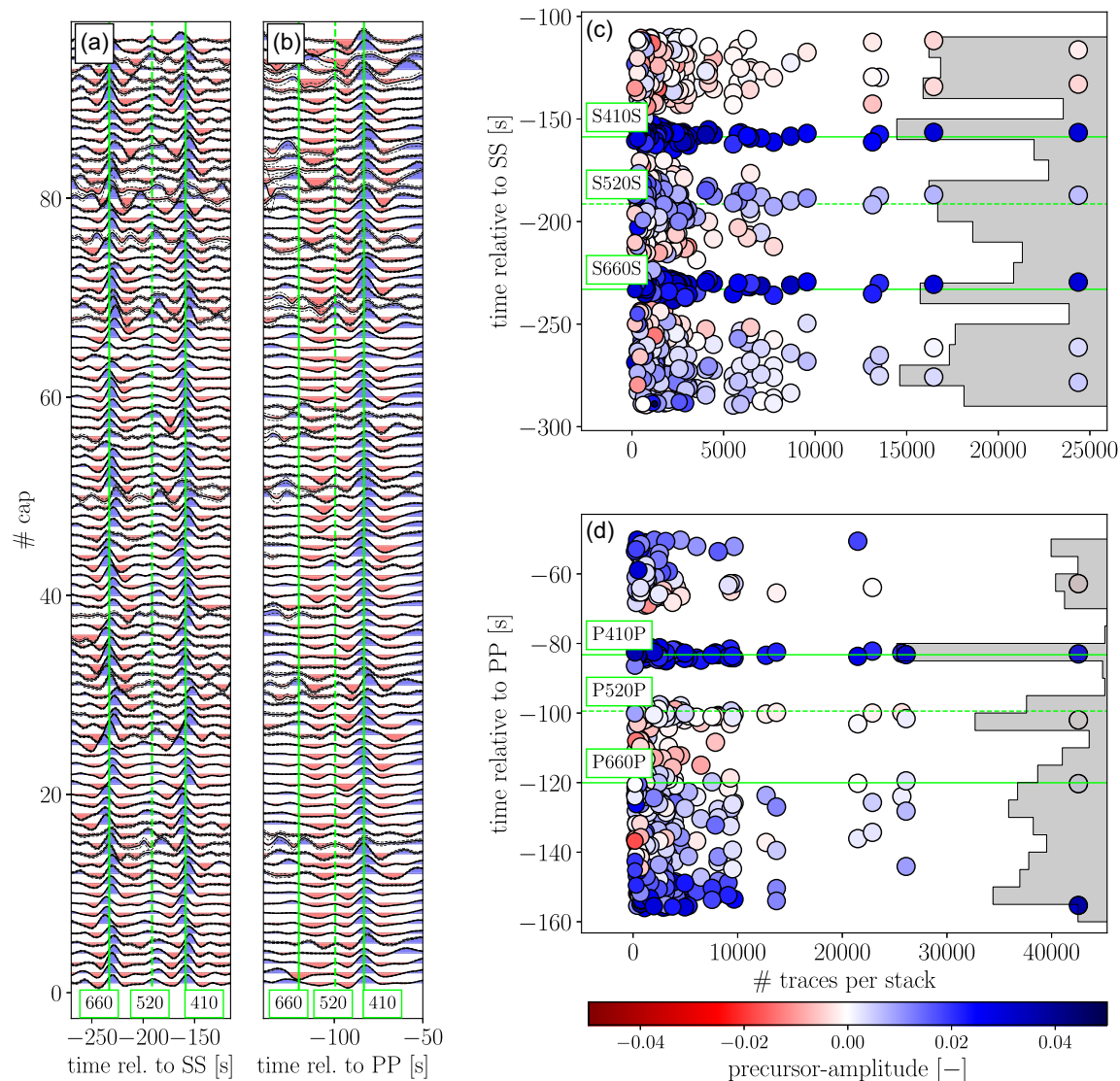


Figure 3. Overview of observed precursor phases. Observed PP (a) and SS waveform stacks (b), respectively, for the caps shown in Fig. 1(c) and the numbering detailed in Table S1. Bootstrap uncertainties are shown as dashed lines. Blue/red colours represent positive/negative amplitudes, respectively. Systematic search for SdS (c) and PdP (d) reflections. Amplitudes (colour-coded) of local maxima in the observed waveform stacks are plotted as a function of picked traveltime (ordinate; relative to the free surface reflection) and number of traces per stack (abscissa). Larger/smaller dots represent smaller/larger bootstrap uncertainties associated with the detected reflection, respectively. The total number of detected reflections as a function of traveltime is displayed by the grey histogram. In all plots, theoretical arrival times of the 410-, 520- and 660-underside reflections (P410P/S410S, P520P/S520S, P660P/S660S) in the seismic reference model IASP91 (Kennett & Engdahl 1991) are shown in green.

Several propositions have been offered. Estabrook & Kind (1996) suggested that the impedance contrast in seismic velocities at the 660 is smaller for *P* waves than for *S* waves, which would result in much weaker P660P phases. In contrast, Deuss *et al.* (2006) argued, on the basis of regional detections, for a strongly variable character of the 660 over short scales that results in a vanishing P660P because of incoherent stacking when performed over larger areas. In line herewith, Day & Deuss (2013) observed lateral undulations on the 660 in short-period P'P'-precursors, which cancel in long-period P'P'-stacks and in P660P-stacks. Lessing *et al.* (2015) further pointed to the importance of compositional variability alongside localized topography of the 660 as a means of influencing P660P observations, while Waszek *et al.* (2021) has argued for P660P observations at locations beneath the Pacific ocean that were only

matched by a mechanically mixed and extremely hot mantle.

Here, we take the purely observational point of view that the P660P is only weakly observable over a very narrow distance range because of the strong interference of other phases. This is exemplified in Fig. 2 for a representative cap (see Fig. S2 for additional caps), where the synthetic waveform section for model IASP91 (panel a) indicates a clear P660P, in contrast to the observations (panel c). In the observed waveform section, the interfering top-side reflection (P660pP) is smeared out over a wider time-distance window, considerably narrowing the range for observing the P660P. This severely reduces the robustness of P660P detections, making it very difficult to use quantitatively. Accordingly, we do not consider the P660P here (see Houser *et al.* 2008, for a similar point of view).

2.4.2 Traveltimes of the 410- and 660-reflections

To analyse our observations of S410S, S660S and P410P more quantitatively, we analyse traveltimes and amplitudes of the 410 and 660 arrivals within a ± 10 s window around the IASP91-predicted arrival time. We compare picked traveltime residuals against theoretical values for a model of crust (Crust1.0; Laske *et al.* 2013) and mantle (GyPSuM; Simmons *et al.* 2010) to assess the effect of shallow structure on traveltimes. The comparison (see Fig. S3) shows that (1) regions with delayed/advanced 410 and 660 precursor times are associated with anomalously short/long crustal and upper mantle traveltimes, respectively and (2) the combined effect of crust and upper mantle structure is of the same order as the observed residuals. These observations highlight the importance of accounting for crust and upper mantle structure when interpreting absolute precursor traveltimes in terms of discontinuity topography. Here, this is accomplished by simultaneously inverting for crust, upper mantle, and discontinuity structure. In contrast, differential S410S–S660S traveltime residuals are almost decorrelated from theoretical MTZ traveltimes. This indicates that observed differential traveltimes are dominated by MTZ thickness and only to a lesser extent by lateral velocity variations within the MTZ (see Chevrot *et al.* 1999, for a similar conclusion).

The MTZ thickness signature in differential S410S–S660S traveltime residuals becomes more evident when interpolating the observations globally (Fig. 4a; see also Fig. S4 for a denser grid). We observe larger differential traveltimes around the Pacific Ocean that reflect the signature of downwelling slabs, including Eurasia, whereas smaller differential traveltimes are seen beneath the Pacific, Atlantic and Africa. This large-scale pattern generally agrees with previous studies (e.g. Shearer 1993; Flanagan & Shearer 1998; Deuss 2009; Tian *et al.* 2020). On average, we detect positive and negative differential traveltimes under continents and oceans, respectively (*cf.* Fig. S3d), that may indicate a surface tectonic control on MTZ structure (*cf.* Gossler & Kind 1996). Subtle differences between Figs 4(a) and S4 (e.g. Northwestern Pacific) illustrate the non-uniqueness of the grid used for stacking. This could be addressed in more detail, for example by adaptive stacking (Waszek *et al.* 2021).

Observed P410P and S410S traveltime residuals relative to IASP91 are further compared in Fig. 4(b) (shown in blue). A positive correlation ($\chi = 0.644$) between P410P and S410S traveltimes is apparent, implying that both precursor types sense the same lateral variations in shallow velocity structure and 410-discontinuity depth (*cf.* Flanagan & Shearer 1999; Chambers *et al.* 2005b).

2.4.3 Amplitude of the 410- and 660-reflections

In Fig. 4(b) (shown in orange), we compare amplitudes of the P410P and S410S phases. While a traveltime correlation is present (Fig. 4b, in blue), there appears to be no correlation between amplitudes ($\chi = 0.004$). This suggests that P410P and S410S are affected differently by, for example uncorrelated *P*- and *S*-wave impedance contrasts. In Table 1, we list the amplitudes measured here and those reported by Shearer & Flanagan (1999) and Chambers *et al.* (2005a). Absolute amplitude variations among the various studies can be attributed to differences in data processing, amplitude measurement techniques and study regions. We find that P410P amplitudes are characterized by a larger variance compared to S410S amplitudes, as also reported earlier by Shearer & Flanagan (1999) and Chambers *et al.* (2005a). This led Chambers *et al.* (2005a) to suggest that S410S impedance contrasts vary over shorter length-scales relative

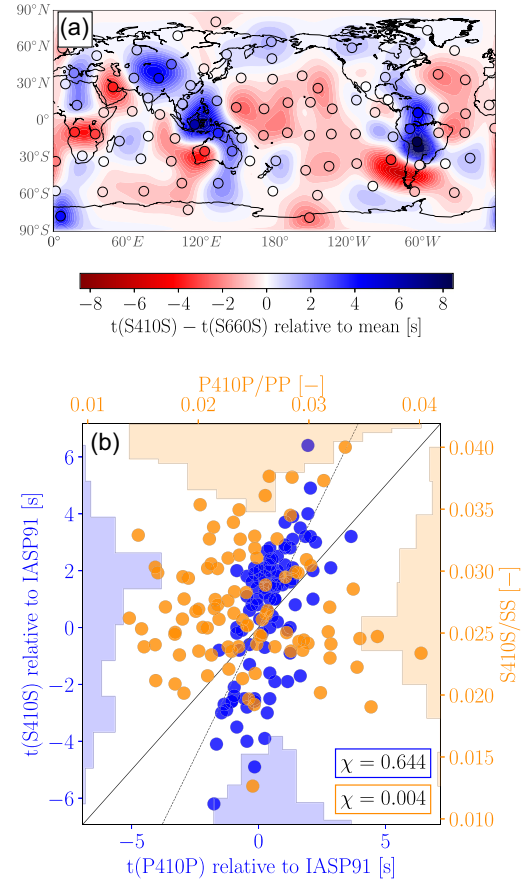


Figure 4. Observed precursor traveltimes and amplitudes. (a) Map of interpolated differential S410S–S660S traveltimes relative to the global mean. Each cap is represented by a dot at the corresponding mean bounce point location. (b) Comparison of P410P and S410S traveltimes (blue) and amplitudes (orange). A 1:1-relation is depicted by the solid line. The dashed line represents the approximate *P*- over *S*-wave velocity ratio of the upper mantle (slope of 1.82). Pearson's correlation coefficient is denoted by χ .

to P410P on account of compositional effects (e.g. water, melt or other chemical heterogeneities). We will return to the question of variations in mantle chemistry in Section 7.

In addition to the aforementioned issues, measured amplitudes can also be affected by incoherent stacking of time-shifted waveforms and the amount and quality of data. The former can be understood as arising from pre-stack traveltime shifts between waveforms that originate in topography variations and velocity heterogeneities in the crust and upper mantle (e.g. Revenaugh & Jordan 1991; Shearer 1991). Data amount and quality is reflected in the robustness of a stack as evidenced by a decrease in bootstrap uncertainty as a function of the number of stacked traces (Fig. S5).

To investigate the effect arising from discontinuity topography and/or velocity heterogeneities on observed amplitudes, we follow the approach of, for example Shearer & Flanagan (1999) and estimate the variability of precursor traveltimes by computing the standard deviation of the precursor traveltime distributions. We find global mean standard deviations of 0.88 s for P410P, 1.8 s for S410S and 2.4 s for S660S phases, respectively (Fig. S6A). Regional variations in traveltime variance (Fig. S7) reflect differences in data coverage and, possibly, lateral variations in length-scale of topographic undulations. Overall, standard deviations of S660S traveltimes are larger, implying stronger topography on the 660 than on the 410,

Table 1. Comparison of P410P and S410S amplitudes from this study, Shearer & Flanagan (1999), Chambers *et al.* (2005a) and Waszek *et al.* (2021).

		P410P/PP	S410S/SS
This study	0.024 ± 0.0054	0.027 ± 0.0047	
Shearer & Flanagan (1999)	$\sim 0.01\text{--}0.05$	$\sim 0.02\text{--}0.05$	
Chambers <i>et al.</i> (2005a)	(Continents)	$\sim 0.01\text{--}0.02$	$\sim 0.02\text{--}0.04$
	(Oceans)	$\sim 0.015\text{--}0.035$	$\sim 0.02\text{--}0.035$
Waszek <i>et al.</i> (2021)	0.033 ± 0.008	0.038 ± 0.006	

in agreement with other global and regional studies (e.g. Flanagan & Shearer 1998; Gu *et al.* 2003; Zheng *et al.* 2015). To account for effects associated with incoherent stacking, we estimate amplitude correction factors from standard deviations of traveltimes (see Section 4.2.4).

2.4.4 Summary of traveltime and amplitude observations

We can summarize the main observations from our precursor data set as follows:

- (i) We observe clear S410S, S660S and S520S phases, including the S560S.
- (ii) We see clear P410P, but not the P660P because of interfering phases.
- (iii) Absolute precursor traveltimes depend strongly on surface topography, crust, and mantle structure, requiring either accurate corrections or, as implemented here, simultaneous inversion for crust and upper mantle structure.
- (iv) While the traveltimes of P410P and S410S phases correlate well, their amplitudes do not, representing a more complex interplay of data and modelling factors (to be detailed in Section 4.2).
- (v) Traveltimes of S660S exhibit larger variability than those of S410S, indicating stronger topography on the 660 than 410.

3 SURFACE WAVE DATA

Inverting precursor waveforms for MTZ structure suffers from a trade-off between mantle velocity structure and discontinuity depth. This is illustrated in Section S4.1, where results from inversion of our PP and SS precursor data set are presented. As a consequence of the velocity–depth trade-off, a strong correlation ($\chi = 0.766$) between inverted 410- and 660-discontinuity depth is seen to exist. A similar observation was made in receiver function analysis by van Stiphout *et al.* (2019). To circumvent this, Gu *et al.* (2003) simultaneously inverted surface and body wave data including differential SS precursor traveltimes. Not surprisingly, Gu *et al.* (2003) found significant differences in inverted 410- and 660-topography relative to previous precursor studies that only relied on traveltime corrections for upper mantle structure.

Here, we follow previous work in the field (e.g. Calò *et al.* 2016; Munch *et al.* 2020) and consider Rayleigh wave dispersion data in the inversion to improve resolution in the upper mantle and to reduce trade-off between velocity and discontinuity topography. We rely on the global Rayleigh wave phase velocity maps, including uncertainties, of Durand *et al.* (2015). These maps are the result of an inversion and hence are affected by regularization. The phase velocities are defined on a $2^\circ \times 2^\circ$ -grid and, at each grid-node, include 48 fundamental-mode and overtone (to fifth order) phase velocities covering a period range of 40–240 s. To compute phase

velocity dispersion curves for each of the 96 caps, we interpolate the original phase velocities to the bounce point coordinates and average them for each mode, period and cap separately. For each cap, the respective uncertainty is estimated by (1) extracting 100 samples at each bounce point from a normal distribution characterized by mean and standard deviation corresponding to the reported phase velocities and errors, respectively, at the particular location and (2) computing the standard deviation of all samples within the considered cap. In addition to improving resolution in the upper mantle and MTZ, this approach offers the important advantage of obviating the need for model-dependent corrections to velocity structure and instead allows for a self-consistent modelling scheme. Synthetic Rayleigh wave phase velocities are computed for isotropic models using the spectral-element normal-mode code *specnm* (Kemper *et al.* 2021) that takes into account attenuation and self-gravitation.

4 MODEL PARAMETRIZATION AND FORWARD MODELLING

4.1 Model parametrization

To explore mantle velocity and discontinuity structure that fit our precursor waveform stacks, we use two independent model parametrizations. The first, ‘seismic’ parametrization is based on the elastic parameters themselves, while the second, ‘thermodynamic’ parametrization additionally considers mineral physics data to self-consistently build seismic profiles (e.g. Afonso *et al.* 2008; Khan *et al.* 2009; Fulla *et al.* 2011; Munch *et al.* 2020; Bissig *et al.* 2022). Attenuation and crustal structure are accounted for in both parametrizations. An overview of all model parameters introduced in the following is given in Table 2 and shown in Fig. 5.

4.1.1 Seismic parametrization

To first order, precursor traveltimes and amplitudes sense seismic velocities and impedance contrast at the discontinuities. To link P - (V_P) and S -wave velocities (V_S), we parametrize models in terms of density (ρ) and bulk (κ) and shear moduli (μ):

$$V_P = \sqrt{\frac{\kappa + \frac{4}{3}\mu}{\rho}}; \quad V_S = \sqrt{\frac{\mu}{\rho}} \quad (1)$$

P - and S -wave impedance contrasts (R_P , R_S) are given by (e.g. Shearer & Flanagan 1999):

$$R_P \approx \Delta\rho + \Delta V_P; \quad R_S \approx \Delta\rho + \Delta V_S, \quad (2)$$

where $\Delta\rho$, ΔV_P and ΔV_S are small fractional jumps in density and P - and S -wave velocity, respectively. Each seismic profile is defined by six variable mantle depth nodes (Fig. 5a, Table 2). Nodes cannot

Table 2. Overview of model parameters, their quantity (#), and prior model parameter values and probability distributions (uniform/normal). The prior constraints for the elastic properties in the seismic inversions are given in terms of (percental) jumps across discontinuities, that is $\delta\rho$, δV_P and δV_S with ρ , V_S and V_P being density, S - and P -wave velocities, respectively.

Description	Parameter	#	Prior
crust			
Depth of crustal base (km)	Z_c	1	
S -wave velocity (km s^{-1})	V_S^c	1	cap-dependent
Density- and P -to- S -wave	ρ/V_S	1	(Crust1.0 (Laske <i>et al.</i> 2013);
Velocity crustal ratios	V_P/V_S	1	<i>cf.</i> Fig. S19)
Crustal traveltimes correction (s)	$T_{\text{corr}}^{\text{PP}}$	1	
	$T_{\text{corr}}^{\text{SS}}$	1	
Mantle (seismic parametrization)			
Discontinuity depth (km)			
‘210’-node	Z_1	1	70–300, uniform
‘410’-node	Z_2	1	70–600, uniform
‘520’-node	Z_3	1	475–600, uniform
‘560’-node	Z_4	1	475–600, uniform
‘660’-node	Z_5	1	475–820, uniform
‘sub-660’-node	Z_6	1	700–820, uniform
Elastic parameters:			
Bulk modulus (GPa)	κ_j	6	$\delta V_{P,S}(Z_2, Z_5) = 10$ per cent
Shear modulus (GPa)	μ_j	6	$\delta V_{P,S}(Z_3, Z_4) = 3$ per cent
density (kg m^{-3})	ρ_j	6	$\delta\rho(Z_2) = 6$ per cent, $\delta\rho(Z_3, Z_4) = 2.5$ per cent, $\delta\rho(Z_5) = 9$ per cent + no negative gradients below Z_2 + prior on absolute values
Shear attenuation quality factor (–):			
surface to Z_c	Q_μ^1	1	100–600, uniform
Z_1 to Z_5	Q_μ^2	1	50–350, uniform
Z_6 to bottom	Q_μ^3	1	200–600, uniform
Mantle (thermodynamic parametrization)			
Basalt fraction (–)	f	1	0–0.5, uniform
Lithospheric temperature ($^{\circ}\text{C}$)	T_{lit}	1	950–1600, uniform
Lithospheric thickness (km)	Z_{lit}	1	70–300, uniform
Crustal basal temperature ($^{\circ}\text{C}$)	T_c	1	thermal gradient in crust \geq lithosphere
Source			
Moment tensor (Nm)	M_i^{PP}	3	cap-dependent
	M_i^{SS}	2	(likelihood values from source
Focal depth (km)	$Z_{\text{src}}^{\text{PP}}$	1	inversion and observed focal
	$Z_{\text{src}}^{\text{SS}}$	1	depth distribution)
Amplitude correction factors			
P410P/PP (–)	A(P410P)	1	1.05 ± 0.031 , normal
S410S/SS (–)	A(S410S)	1	1.14 ± 0.094 , normal
S660S/SS (–)	A(S660S)	1	1.23 ± 0.106 , normal

move past each other, for example the 410 always remains above the 660. MTZ-discontinuities corresponding to the 410, 520, 560 and 660, are all represented by a first-order discontinuity for which we vary the jumps in elastic properties. While the 410 and 660 have been observed to be sharp (<5 km; e.g. Benz & Vidale 1993), the more gradual mid-MTZ discontinuities (e.g. Xu *et al.* 1998) can be modelled by concatenating the 520- and 560-nodes. At the base of the crust and the ‘210-node’, we perturb absolute values of elastic properties. Similarly to Lawrence & Shearer (2006), elastic properties in between and above the MTZ-nodes are computed by using the corresponding gradients from IASP91 (Kennett & Engdahl 1991). Gradients are variable in the upper mantle (above the ‘210-node’ in Fig. 5a) and below the 660, whereas negative velocity gradients are

only permitted in the uppermost mantle (above the ‘210-node’). At and below the ‘sub-660’ depth node, properties are set to IASP91.

4.1.2 Thermodynamic parametrization

A second parametrization follows our earlier efforts (e.g. Khan *et al.* 2009; Munch *et al.* 2020; Bissig *et al.* 2022) and is rooted in petrological and geodynamic models of the Earth’s mantle.

Partial melting of the mantle at mid-ocean ridges creates a chemically stratified oceanic lithosphere composed of a fertile, basaltic crust and a depleted, harzburgitic mantle layer. The segregation of basalt from harzburgite in subducted oceanic lithosphere and subsequent mixing produces chemical heterogeneities (e.g. Stixrude &

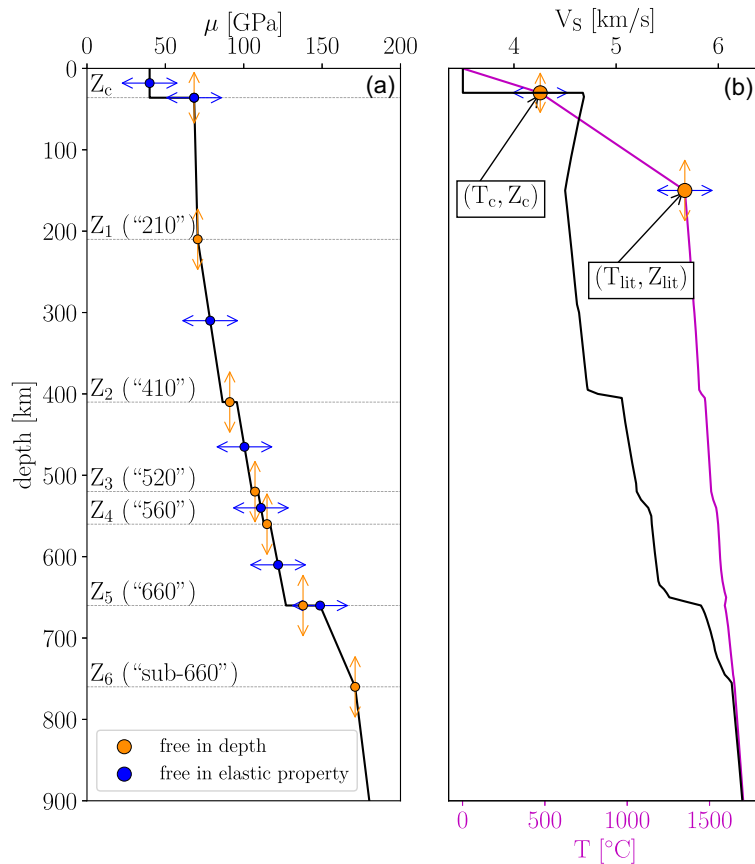


Figure 5. Illustration of seismic (a) and thermodynamic (b) model parametrizations, respectively. (a) Profile of shear modulus (μ), where dots indicate segments and nodes that are variable in elastic properties and depth, respectively. (b) Geotherm (magenta) where the adiabatic mantle is overlain by a 150 km thick, conductive lithosphere with basal temperature of 1350 °C. Free thermal parameters are indicated by dots. The associated S -wave velocity (V_S) profile for a pyrolytic mantle is plotted in black. See Table 2 for an overview of all model parameters.

Table 3. Major element compositions of mid-ocean ridge basalt (MORB; \mathbf{X}_B) and harzburgite (\mathbf{X}_H) in wt per cent (from Khan *et al.* 2009).

Component	MORB, \mathbf{X}_B	Harzburgite, \mathbf{X}_H
CaO	13.05	0.5
FeO	7.68	7.83
MgO	10.49	46.36
Al ₂ O ₃	16.08	0.65
SiO ₂	50.39	43.64
Na ₂ O	1.87	0.01

Lithgow-Bertelloni 2012). Depending on the amount of equilibration between basalt and harzburgite, the mantle can either be described by a mechanical mixture (MM) or equilibrium assemblage (EA) (Xu *et al.* 2008). Using basaltic and harzburgitic end-member compositions, \mathbf{X}_B and \mathbf{X}_H , respectively, we can express the resultant mineralogy, Φ , by:

$$\text{EA} : \Phi[f\mathbf{X}_B + (1 - f)\mathbf{X}_H], \quad (3)$$

$$\text{MM} : f\Phi[\mathbf{X}_B] + (1 - f)\Phi[\mathbf{X}_H], \quad (4)$$

for the two mixing models, respectively, and variable basalt fraction, f . End-member compositions in the Na₂O–CaO–FeO–MgO–Al₂O₃–SiO₂ model chemical system are listed in Table 3. To determine stable mineralogy and corresponding physical properties for a set of pressure and temperature conditions, we use Gibbs’ free energy minimization and equation-of-state modelling using the

software package Perplex_X (Connolly 2009). Equivalent properties for a mechanically mixed mantle are obtained using the software package PHEMGP (Zunino *et al.* 2011).

The so-computed density and elastic moduli are uncertain by ~ 0.5 per cent and ~ 1 –2 per cent, respectively (Connolly & Khan 2016). In modelling pressure, we assume hydrostatic equilibrium and integrate the load from the surface. The geotherm (*cf.* Fig. 5b) is modelled assuming a conductive lithosphere overlying an adiabatic mantle. We split the lithospheric geotherm into separate crust and mantle segments, parametrized by surface temperature (held constant at 2 °C), crustal thickness and basal temperature (Z_c and T_c) and lithosphere thickness and basal temperature (Z_{lit} and T_{lit}), respectively. The sublithospheric mantle is adiabatic when convection dominates the conductive heat transport in the mantle (Katsura 2022). Here, the adiabatic geotherm is self-consistently computed from the entropy at the base of the lithosphere. The immediate advantage of an adiabat as implemented here over a more general geotherm is a reduction in the number of model parameters.

4.1.3 Attenuation

There is little consensus on the radial shear attenuation structure of the Earth (*cf.* Romanowicz & Mitchell 2015). Synthetic tests reveal that radial profiles of shear quality factors (Q_μ) (e.g. Dziewoński & Anderson 1981; Oki & Shearer 2008; Hwang *et al.* 2011) result in variable precursor amplitudes and phase velocities (*cf.* Fig. S14).

Therefore, to account for uncertainties related herewith, we vary Q_μ during the inversion.

In case of the seismic parametrization, we define Q_μ in three layers: crust, upper mantle and MTZ, and lower mantle. In contrast, the thermodynamic parametrization permits the use of the laboratory-based extended Burgers model (Jackson & Faul 2010) to compute Q_μ self-consistently from the geotherm. The implementation and rheological parameters used are those of Bagheri *et al.* (2019). Parameter values are listed in Table S3. For both parametrizations, bulk attenuation is held constant after PREM ($Q_k = 57\,823$; Dziewoński & Anderson 1981).

4.1.4 Crust

It is common practice in SS and PP precursor studies to assume that effects related to crustal structure and surface topography are restricted to (1) the relative traveltimes of precursors and the free surface reflection and (2) can be accounted for by applying a traveltime correction based on a crustal model at the location of the bounce point. To verify these assumptions, we perform a 3-D synthetic test using the software-package Salvus (Afanasiev *et al.* 2019, for details see Section S2.2). A comparison of waveform stacks computed for models with different crustal thickness below source, bounce point, and receiver is shown in Fig. 6. We observe that both traveltimes and amplitudes (in particular for PP) of the 410 and 660 precursors are affected by crustal structure below the bounce point. Therefore, we account for crustal effects by:

1. Modelling the crust as a single layer that is representative of the crust at the bounce point, whose seismic properties (P - and S -wave velocities, density and thickness) are based on the crystalline layers in Crust1.0 (Laske *et al.* 2013).
2. Applying traveltime corrections based on the differential two-way traveltimes between the simplified one-layer models used here and the multilayer profiles of Crust1.0 including surface topography or bathymetry.
3. Inverting for elastic properties, thickness and traveltime corrections for each cap (bounce point) separately (*cf.* Table 2).

4.2 Forward modelling

4.2.1 Waveform modelling

In this study, we invert PP and SS precursor waveforms directly for the seismic structure of the MTZ. For this, we follow our previous work (e.g. Munch *et al.* 2018; Bissig *et al.* 2022) and use the reflectivity method (Fuchs & Müller 1971) to model synthetic waveforms that are processed similarly to observed data. Consistent processing of both synthetic and observed data was shown by Munch *et al.* (2018) to be, in addition to accurate waveform modelling, the most important element when inverting waveforms.

The use of synthetic seismograms to reproduce stacks analogously to data differs from most previous approaches in precursor waveform inversions that relied either on ray-theoretical (e.g. Shearer 1996; Lawrence & Shearer 2006) or Green's function-based methods (Dokht *et al.* 2016). An exception is the work by Waszek *et al.* (2021), who also stacked reflectivity seismograms and analysed the resulting precursor traveltimes and amplitudes, but did not invert waveforms.

Waveform amplitudes and traveltimes are affected by a number of issues, including: (1) impedance contrast and depth of discontinuities; (2) incidence angles at the discontinuities and geometrical

spreading; (3) attenuation; (4) crust and upper mantle velocity structure; (5) source effects and (6) incoherent stacking. In the following, we consider each of these modelling issues in more detail.

4.2.2 Geometrical spreading, attenuation, and shallow structure

To account for amplitude variations due to variable incidence angles at the discontinuities and geometrical spreading, we take into account the observed slowness distribution for each cap by computing individual waveforms for 60 separate slowness values and then weighing the synthetic waveform amplitudes by the respective value in the slowness distribution prior to stacking (*cf.* Munch *et al.* 2018). Shear attenuation and shallow structure are accounted for by inverting for the appropriate model parameters (Section 4.1 and Table 2).

4.2.3 Source effects

To remove interference effects due to the depth phase, we restrict focal depths to <75 km. To further eliminate signatures of the radiation pattern and source–time function, we deconvolve the free surface-reflected wavelet from the waveform, assuming that the rays reflected at the discontinuities and at the free surface have very similar take-off angles at the source (*cf.* Fig. 1a). Despite these measures, a comparison of synthetic stacks (*cf.* Fig. S16) shows that source parameters can still affect amplitudes, in particular of PP precursors. Consequently, we also vary moment tensor components and source depths.

4.2.4 Amplitude correction factors

As noted in Section 2.4.3, incoherent stacking can reduce observed precursor amplitudes. To quantitatively assess the effect of traveltime shifts on stacked amplitudes, we build stacks from two synthetic waveform sets: with and without time shifts. The ratio between the resulting stacks gives amplitude correction factors for each precursory phase, that are used to correct synthetic waveform stacks during the inversion. To create the time-shifted waveform set, we use the estimated traveltime standard deviations from the observed precursor stacks (*cf.* Section 2.4.3). Since it is difficult to reliably estimate standard deviations, because they depend on data quality/amount, the stacking grid, discontinuity topography and velocity heterogeneities, we estimate standard deviations for different grids and data selection criteria and compute the distributions of amplitude correction factors instead of single values (*cf.* Fig. S6b). The mean and standard deviation of each distribution serves as a prior in the inversion, allowing us to invert for amplitude correction factors as well.

4.2.5 Modelling limitations

Contributions from anisotropy and 3-D mantle structure can also affect waveforms. Azimuthal anisotropy in seismic velocities can lead to variations in amplitudes and traveltimes of precursor arrivals that sample distinct backazimuth windows. Huang *et al.* (2019), for example, found from the analysis of SS precursors that azimuthal anisotropy, on a global scale, is on the order of 1 per cent in the upper mantle and <1 per cent within the MTZ, although stronger signals (up to 3–4 per cent) were seen in subduction zones. They concluded that the impact of azimuthal anisotropy on derived MTZ thickness maps is minor (± 3 km).

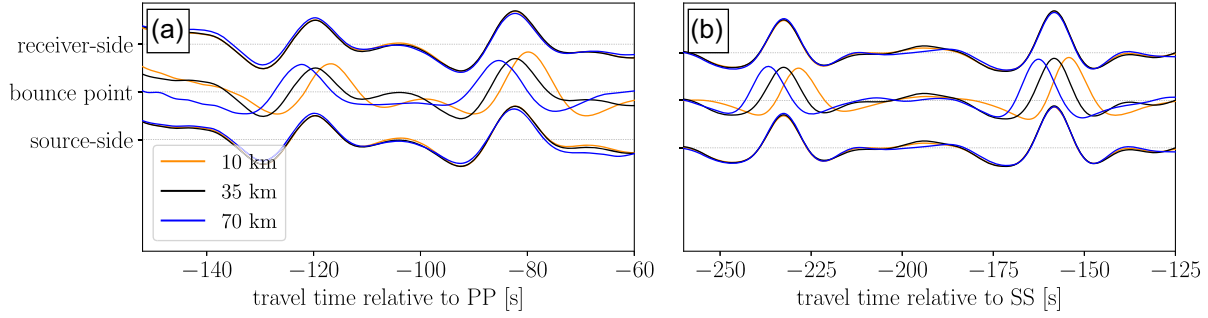


Figure 6. 3-D crustal effects on waveform stacks. Shown are synthetic PP (a) and SS waveform stacks (b) for 3-D models with different crustal thickness (colour-coded) beneath the receiver, bounce point and source (from top to bottom). See Section S.2.2 for computational details.

There is evidence for radial anisotropy in the upper mantle, which peaks at about 8 per cent at 100–200 km depth and most probably vanishes below 300 km (e.g. Dziewoński & Anderson 1981; Beghein *et al.* 2006; Kustowski *et al.* 2008; Khan *et al.* 2009; Lekić & Romanowicz 2011; Chang *et al.* 2015). While the SS waveforms studied here are horizontally polarized, Rayleigh wave dispersion data sense the velocity of vertically polarized waves. However, a synthetic test (*cf.* Fig. S15) illustrates that changes in fundamental-mode phase velocities due to variations in radial anisotropy are at least one order of magnitude smaller than data uncertainties and isotropic variations. Therefore, we leave anisotropy in precursor waveform inversion for a future study.

3-D effects on precursor traveltimes have been suggested to arise from the presence of interfering phases (Koroni *et al.* 2017). Here, we minimize these effects by muting these prior to stacking. Apart from traveltimes, Bai & Ritsema (2013) have measured synthetic S410S and S660S amplitudes that range from 0.035–0.042 and 0.065–0.070, respectively, for different 3-D velocity models. However, they showed that the modelled variations in amplitudes are smaller by a factor of 3 in comparison to the observations. Accordingly, contributions from, for example impedance contrasts across a discontinuity, have a bigger impact on observed amplitudes than 3-D structure and we therefore ignore the latter here.

5 INVERSE PROBLEM

5.1 Formulation of the inverse problem

Within the Bayesian framework, the solution to the inverse problem, $\mathbf{d} = \mathbf{g}(\mathbf{m})$, where \mathbf{m} is the model parameter vector, \mathbf{g} the forward modelling operator and \mathbf{d} the data vector, is given by (Mosegaard & Tarantola 1995):

$$p(\mathbf{d}, \mathbf{m}) \propto h(\mathbf{m})\mathcal{L}(\mathbf{d}, \mathbf{m}), \quad (5)$$

where probability density functions are used to encode prior information on model parameters, $h(\mathbf{m})$, the misfit between predicted and observed data (through the likelihood function $\mathcal{L}(\mathbf{d}, \mathbf{m})$) and the posterior distribution, $p(\mathbf{d}, \mathbf{m})$, respectively. Model parameters and prior information are listed in Table 2. The likelihood function is given by

$$\mathcal{L}_i(\mathbf{d}, \mathbf{m}) = \exp(-\Phi_i), \quad (6)$$

where the misfit between observed and synthetic data is given by a L_2 -norm-function:

$$\Phi_i^{L_2} = \frac{1}{N} \sum_j \frac{\|\mathbf{d}_j^{\text{obs}} - \mathbf{d}_j^{\text{syn}}\|^2}{2\sigma_j^2}. \quad (7)$$

$\mathbf{d}_j^{\text{syn}}$ and $\mathbf{d}_j^{\text{obs}}$ are synthetic and observed data, respectively, σ_j is the uncertainty of measurement j , and N denotes the number of total measurements for each data subset i . We split the data set into 6 subsets, where the first subset comprises the phase velocities and the other 5 subsets correspond to 5 separate time windows of the waveform stacks. For SS, we consider 3 windows, two for S410S and S660S, respectively, and a third for the signal in-between, that is S520S and/or S560S. For PP, we use a narrow window around the P410P and a broader window spanning the theoretical arrivals of the P520P and P660P. Due to phase interference and the short epicentral distance windows over which P520P and P660P are visible, we rely less on the P520P and P660P waveforms. As for the Rayleigh wave phase velocities, each of these have equal weight, but because the fundamental-mode branch consists of more phase velocity points than the overtones, its relative weight in the misfit function is larger. The combined likelihood function is given by the conjunction of the likelihood functions associated with each data subset:

$$\mathcal{L}(\mathbf{d}, \mathbf{m}) = \prod_i \mathcal{L}_i(\mathbf{d}, \mathbf{m}). \quad (8)$$

5.2 Inversion strategy

For each cap, we follow a step-wise inversion strategy:

(i) *Source inversion*: we use Instas eis (van Driel *et al.* 2015) and model PP and SS stacks for IASP91 (Kennett & Engdahl 1991) and 500 randomly picked event-station configurations. The stacks are then inverted separately for pseudo-event moment tensor components and source depths to account for source effects (Section 4.2.3). For this, we rely on the Covariance Matrix Adaption Evolution Strategy (CMAES) by Hansen & Ostermeier (2001), a global optimization method. Based on the returned source parameters and their associated misfit-values, a probability distribution over the parameter space is estimated, which serves as a prior in the following inversions. This step is the same for both, seismic and thermodynamic, parametrizations. Source inversion results are presented in Section S3.1.

(ii) *Structural inversion I* (*‘optimization stage’*): observed waveform stacks and dispersion curves are inverted for structural and source parameters using the cascaded Metropolis–Hastings algorithm (Mosegaard & Tarantola 1995; Bissig *et al.* 2021). This approach results in a diminished trade-off between the different data types and waveform parts and thereby results in an overall improved fit of the entire data set. The optimization stage is characterized by a coarse exploration of the model space by using large step sizes. Typically, this stage requires two runs, where the second optimization is based on a selection of well-fitting models from the first

inversion. Ultimately, a collection of well-fitting models from the optimization stage are used as starting models in the next stage.

(iii) *Structural inversion II ('sampling stage')*: the inversion setting from inversion stage I is adjusted by (1) reducing step sizes and (2) running multiple chains several times in order to sample 10^4 s of models for each cap. In total, every 10th model is retained, resulting in 10^3 s of models for each cap.

To achieve good acceptance rates (~ 30 – 50 per cent) for reasonable step sizes, a constant scaling of the misfit values has proven useful, which is determined by trial-and-error for each cap separately. Although useful from a practical point of view, the scaling can affect the width of the posterior distribution. The results presented in the following sections are based on the models collected in the sampling stage. To visualize the spatial pattern of model parameters, we use the movie strategy (e.g. Tarantola 2005; Khan *et al.* 2013) and produce interpolated maps of the mean and maximum likelihood models as well as the upper and lower bound of the 50 per cent-highest probability interval. In what follows, we restrict ourselves to show mean models and include the movies in Figs S38–S49.

6 SEISMIC INVERSIONS: RESULTS AND DISCUSSION

6.1 Datafit

In Fig. 7, we show a representative example (cap #54) of waveform and Rayleigh wave dispersion datafit. We separate the datafits obtained using the seismic (Figs 7a–d) and thermodynamic (Figs 7e–h) parametrizations and presently focus on the former. The datafit for all caps is shown in Figs S25–S34. In general, we find the data to be well-fit for the various seismic phases considered.

6.2 Seismic velocity structure

Interpolated maps of mean *S*-wave velocity structure are shown and compared to the tomographic model by Durand *et al.* (2015) in Fig. 8 at different depths (more depth slices are shown in Fig. S35). At 100 km depth (Fig. 8a), the *S*-wave velocity patterns are strongly correlated with surface tectonics (*cf.* plate boundaries in Fig. 8a) and are indicative of shallow thermal variations as seen elsewhere (e.g. Afonso *et al.* 2019; Debayle *et al.* 2020). We see negative velocity anomalies beneath mid-ocean ridges (e.g. East Pacific rise) and the backarc basins of subduction zones (e.g. Western Pacific). In contrast, positive anomalies are associated with old, thickened oceanic lithosphere (e.g. central Pacific) and most continental parts. Although the spatial pattern of anomalies is in general agreement with that seen in tomographic models (Fig. 8b; e.g. Lekić & Romanowicz 2011; Ritsema *et al.* 2011; Schaeffer & Lebedev 2013), amplitudes differ. The coarse grid used here acts to smooth the velocity maps and tends to reduce the amplitudes from typically ± 5 – 7 per cent seen in tomographic models to ± 4 per cent found here. Additional differences might relate to cap geometry and inversion method through, for example regularization parameters.

At 600 km depth (Fig. 8c), the amplitudes of velocity perturbations generally decrease and the spatial pattern tends to be decorrelated from surface tectonics (Khan *et al.* 2011; de Viron *et al.* 2021). Our models map high-velocity anomalies underneath South-Eastern Asia, South America and parts of Africa, which are possibly related to slabs of active and ancient subduction systems. Most dominant low-velocity anomalies are imaged beneath the Eastern

and Southwestern Pacific as well as underneath the Indian ocean and Central Asia. Our final model is affected by volumetric velocity perturbations sensed by surface waves and the detailed discontinuity structure (depth, velocity jumps) determined by precursors, which is opposite to most tomographic models, where discontinuity depths are fixed in the inversion. This may explain the weaker agreement with tomographic models at 600 km depth (Fig. 8d) relative to what is observed at 100 km depth.

6.3 Discontinuity depths and mantle transition zone thickness

An overview of the global MTZ structure in the form of *S*-wave impedance contrasts (*cf.* eq. 2) and discontinuity depths is shown in Fig. 9 based on the results from all caps. In the following, we first discuss the discontinuity depths, followed by impedance contrasts.

6.3.1 Mantle transition zone thickness: global average and spatial pattern

We find global average depths of 407.2 ± 7.5 km for the 410 and 651.5 ± 8.8 km for the 660 (*cf.* marginal histograms in Fig. 9), resulting in a total mean MTZ thickness of 244.3 ± 7.7 . While absolute discontinuity depths differ among studies (see compilation in Table 4), MTZ thickness is a more robust parameter with the global average converging on ~ 240 – 245 km. An interpolated map of mean MTZ thickness is shown in Fig. 10(a). Overall, the large-scale pattern is in agreement with the model by Tian *et al.* (2020), which is shown in Fig. 10(b), including other studies (e.g. Flanagan & Shearer 1998; Gu *et al.* 2003; Deuss 2009; Houser *et al.* 2008; Lawrence & Shearer 2008; Huang *et al.* 2019; Waszek *et al.* 2021). Significant differences in short-scale structure are, however, apparent among the various studies. This most likely relates to the considered data, their amount, quality and coverage, in addition to the manner in which velocity variations, for example inversion versus time-to-depth conversion (including differences in velocity models used for making traveltime corrections) are applied.

To better understand regional variations in MTZ thickness, we compare posterior distributions as a function of tectonic setting. To do so, we consider the global tectonic regionalization ‘GTR1’ (Jordan 1981), which distinguishes three different continental and oceanic regions: Q (Phanerozoic orogenic zones), P (Phanerozoic platforms), S (Precambrian shields and platforms), A (young ocean, <25 Ma), B (intermediate ocean, 25–100 Ma) and C (old ocean, >100 Ma). In addition to this surface tectonics-based regionalization, we characterize each cap by its relation to deeper mantle structure, i.e. hotspots/plumes and subducting slabs. For this, we use the hotspot list of Anderson & Schramm (2005) and the slab model ‘Slab2’ by Hayes *et al.* (2018) to compute, for each cap separately, the fraction of bounce points that are either close to a hotspot or a slab. The latter are distinguished based on whether they stagnate in the upper mantle or penetrate the 410. Locations with neither hotspots nor slabs are referred to as ‘normal mantle’. We show MTZ thickness histograms for the different tectonic settings in Fig. 10(c). This comparison reveals that subduction zones are characterized by a thick MTZ as opposed to normal mantle and hotspot regions, as is expected from the Clapeyron slopes of the olivine transformations (e.g. Ita & Stixrude 1992). The MTZ is overall slightly thicker/thinner underneath continents/oceans, respectively, whereas this thermally dominated surface tectonics signal is more prominent in the case of the 410 than the 660 (*cf.* values in Table 4).

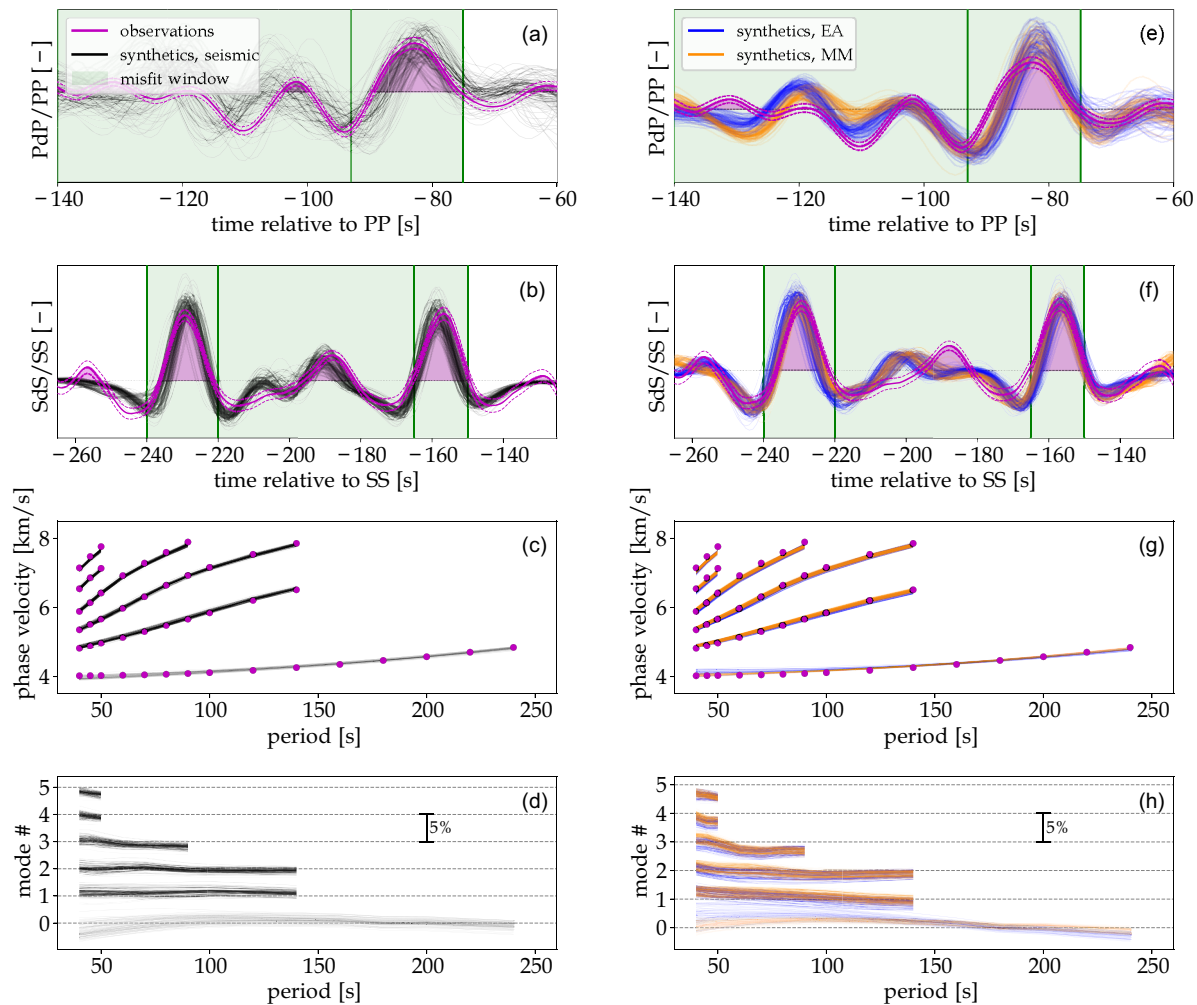


Figure 7. Datafit for seismic (left-hand panel) and thermodynamic (right-hand panel) inversions. Shown are fits of PP (a, e) and SS waveform stacks (b, f) and Rayleigh wave phase velocities (c, g) for cap #54 (Indian Ocean). Panels (d) and (h) show phase velocity residuals relative to the observations (in per cent), that is positive/negative residuals indicate higher/lower synthetic phase velocities, respectively. The thermodynamic datafit is shown for inversions based on an equilibrium assemblage (EA) and mechanical mixture (MM) model. Synthetics are randomly picked from the posterior distribution. Entire datafit sections for all caps are shown in Figs S25–S34.

This possibly implies a thermochemical decoupling of the discontinuities, as also suggested by the correlation between their mean depths ($\chi = 0.562$).

6.3.2 Mid-mantle transition zone discontinuities

In contrast to the ‘simple’ 410 and 660 discontinuities, the discontinuity-structure within the MTZ is more complex (Fig. 9). Average depths of the 520- and 560-nodes are 510.5 ± 13.8 km and 560.9 ± 16.7 km, respectively, implying larger topographic variations than on the 410 and 660. However, mapping and interpreting these values and the associated impedance contrasts is more complicated than for the 410 and 660, chiefly because the 520 is capable of producing a plethora of signals from a single (around 520 km depth) over gradual (in the depth range 520–560 km) to completely split discontinuities (e.g. 520 and 560 km depth). To avoid complications in interpreting absolute depths of mid-MTZ nodes, we use ‘the sequencer’ (www.sequencer.org) to identify trends in our seismic data set (e.g. Kim *et al.* 2020). We specifically use the sequencer to reveal structural trends in our inverted mid-MTZ reflectivity profiles. For each cap, we use the marginal posterior depth distributions of the

520- and 560-nodes (normalized to the respective 410-discontinuity depth distribution) and the associated impedance contrasts to create a profile representative of the mid-MTZ reflectivity structure. To improve the stability of the result, we only consider 50 per cent of the caps, for which mid-MTZ SS waveforms are fit very well and more than 10^3 traces are included in the stack.

A contour plot of the sequenced reflectivity profiles is shown in Fig. 11(a) with the average reflectivity profiles representative of the main trend plotted on top (coloured lines). The colour map is based on the index of the profile after sequencing and is used in Fig. 11(b) to visualize the spatial pattern of the main trend. Going from red to blue, mid-MTZ structure evolves from simple to complex, that is from a single to two well-separated discontinuities. Most of the Pacific and Atlantic (coloured in red in Fig. 11b) are characterized by a single mid-MTZ discontinuity, whereas more strongly separated discontinuities are seen underneath Eurasia and the circum-Pacific subduction zones (coloured in blue in Fig. 11b). A similar pattern was observed by Tian *et al.* (2020), who associated the detection of an additional mid-MTZ reflector, the 560, with localized oceanic crustal enrichments close to subduction zones.

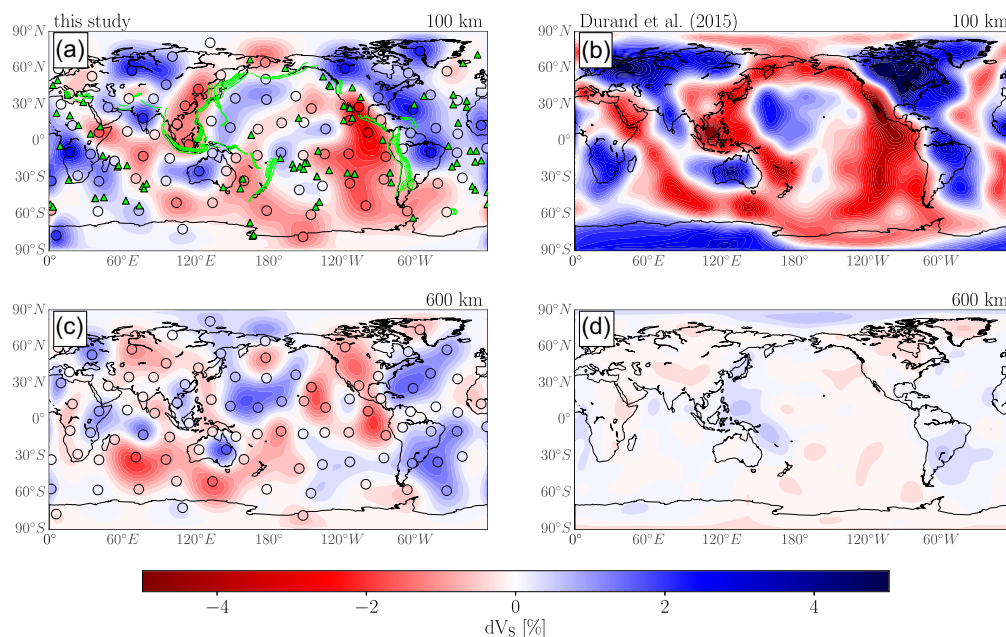


Figure 8. *S*-wave velocity variations from seismic inversions. Shown are slices through our interpolated mean model (left-hand panel) at 100 km (a) and 600 km (c) depth, respectively. For comparison, respective slices through the global tomographic model by Durand *et al.* (2015) are shown on the right (b,d). Perturbations in *S*-wave velocity, dV_S , are given in per cent with respect to the global mean at each depth. In (a), plate boundaries from USGS are plotted in green.

6.4 Impedance contrasts of discontinuities

We find average *S*-wave impedance contrasts (*P*-wave contrasts are discussed below) of 7.4 ± 1.0 for the 410 and 9.2 ± 1.4 for the 660. These values are in general agreement with previous studies (*cf.* Fig. 9; Shearer & Flanagan 1999; Chambers *et al.* 2005a; Lawrence & Shearer 2006). Spatial variations are seen in maps showing the impedance contrast at the 410 (Fig. S42), where higher values (~ 7 – 9 per cent) are largely confined beneath oceans and lower values (~ 5 – 7 per cent) underneath continents. The 660 impedance contrast (Fig. S43) is decorrelated from the 410 ($\chi = -0.225$), which points to a thermochemically decoupled discontinuity.

As seen from Fig. 4(c), the observed decorrelation of P410P- and S410S-amplitudes could imply the presence of chemical heterogeneities (Chambers *et al.* 2005a). Conversely, inverted *P*- and *S*-wave impedance contrasts correlate well ($\chi = 0.608$), favoring a similar compositional signal seen by P410P and S410S phases. This is in agreement with P410P amplitudes and *P*-wave impedance contrasts being poorly correlated ($\chi = 0.352$) and instead reflecting the importance of considering, for example incoherent stacking and crustal structure in modelling amplitudes.

Based on Fig. 9, impedance contrasts obtained here are compatible with a compositional gradient from more fertile at the 410 to pyrolytic lithologies at the 660. However, phase transformations do not result in first-order, but finite-width discontinuities, complicating compositional inference from first-order impedance contrasts. For example, Lawrence & Shearer (2006) found larger impedance contrasts at the 410 when using finite-width instead of first-order discontinuities (*cf.* LS06-3 versus LS06-1 in Fig. 9). Similarly, estimates of the impedance contrast at the 660 could be biased by the sub-660 gradient zone (post-garnet transition), which can contribute to the long-wavelength S660S-signal.

Relative to PREM, impedance contrasts at the 660 are seen to be smaller (Dziewoński & Anderson 1981). The associated density

jump is 6.1 per cent as opposed to 9.8 per cent in PREM and agrees well with estimates from precursor- (5.2 per cent; Shearer & Flanagan 1999) and normal mode-analysis (5.1–8.2 per cent; Lau & Romanowicz 2021). Resolution of the density jump can be improved by the analysis of precursor amplitude versus offset sections in densely sampled areas (AVO; e.g. Shearer & Flanagan 1999; Yu *et al.* 2018) or the inclusion of converted waves in the inversion procedure (e.g. Lawrence & Shearer 2006).

In summary, while impedance contrasts and precursor-amplitudes have been used to estimate mantle composition (e.g. Shearer & Flanagan 1999; Chambers *et al.* 2005a; Yu *et al.* 2018), it is difficult to directly relate precursor amplitudes and first-order discontinuity impedance contrasts to mantle composition. Instead, amplitudes need to be carefully modelled by relying on proper thermodynamically constrained models, which is considered in the next Section 7.

7 THERMODYNAMIC INVERSIONS: RESULTS AND DISCUSSION

7.1 Datafit

The datafit for a single cap (#54) using the thermodynamic parametrization is shown in Figs 7(e)–(h) (additional caps are shown in Figs S25–S34). As in the case of the seismic parametrization, most features of the data are fit, including the S410S, S660S and P410P phases. Less well-fit phases include mid-MTZ SS waveforms, the exact shape of the fundamental mode branch and some overtones (to be discussed in Section 7.3).

7.2 Global thermochemical structure

Lateral variations in thermal state are illustrated using mantle potential temperature, T_{pot} . The latter is computed by extrapolating the

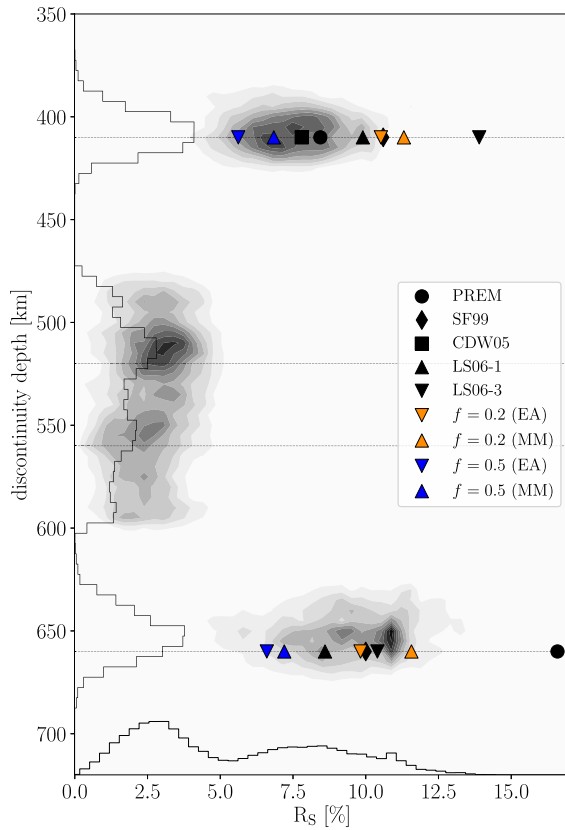


Figure 9. Global reflectivity- and discontinuity-structure of the mantle transition zone. Shown are 2-D contour-plots of the S -wave impedance contrasts (R_S ; cf. eq. 2; abscissa) against discontinuity depth (ordinate). Marginal histograms for each parameter are shown with solid black lines. For comparison, impedance contrast estimates from other studies are shown in black: PREM (Dziewoński & Anderson 1981), SF99 (Shearer & Flanagan 1999), CDW05 (Chambers *et al.* 2005a), LS06-1 and LS06-3 (results for first-order and finite-width discontinuities, respectively, from Lawrence & Shearer 2006). Coloured triangles represent impedance contrasts estimated from seismic models differing in basalt fraction (f) for an equilibrium assemblage (EA) and a mechanical mixture (MM) model, respectively. Thermal conditions are: $Z_{lit} = 150$ km, $T_{lit} = 1350^\circ\text{C}$.

inverted temperature at the lithospheric base (i.e. T_{lit}) to the Earth's surface using the inverted upper mantle thermal gradients. We find global average temperatures of $1352 \pm 69^\circ\text{C}$ for EA and $1319 \pm 53^\circ\text{C}$ for MM inversions, which encompasses the peridotite-based geotherm of Katsura (2022) with a potential temperature of 1350°C . As shown in Figs 12(a)–(c), continents and subduction zones are characterized by lower temperatures than oceans and hotspots. This is in qualitative agreement with the upper mantle thermal model WINTERC-G (Fullea *et al.* 2021, cf. Fig. S51), which has been obtained from a combined inversion of seismic, gravity, heat flow and surface elevation data. Also, temperatures below continental regions tend to decrease with age, in agreement with the trend seen by Munch *et al.* (2020) that was obtained from a similar waveform inversion of global P-to-s receiver functions as conducted here. While a recent study by Waszek *et al.* (2021), using differential SS precursor traveltimes only, found lateral thermal variations in the Pacific region that match our results, discrepancies exist in other regions. Quantitative comparison, however, is difficult on account of their set-up (differential traveltimes are first mapped to MTZ thickness and then to temperature) and the assumption of fixed mantle composition. Interestingly, we find that hotspots are characterized

by overall higher temperatures relative to normal mantle, but not necessarily by a thinner MTZ (cf. Fig. 10c). Several issues could be at play here: (1) MTZ thickness does not simply reflect temperature but also composition so that other mineral phases and their transformations (in particular around the 660) can become important (e.g. Jenkins *et al.* 2016) and (2) inverted potential temperature reflects a combination of upper mantle and MTZ structure, whereas MTZ thickness is not sensitive to upper mantle structure, indicating that not all hotspots reach the MTZ.

Figs 12(d)–(f) compares lateral variations in composition. Oceans appear to be well-fit by a pyrolytic mantle, whereas continents show more variability and tend to more depleted compositions with age. Moreover, subduction zones are characterized by basalt-enriched compositions relative to normal mantle regions. This agrees well with seismic studies (e.g. Bissig *et al.* 2022; Tauzin *et al.* 2022) and results from geodynamic simulations (e.g. Ballmer *et al.* 2015; Yan *et al.* 2020). Hotspots appear pyrolytic in composition, possibly indicating that not all hotspots are fed by deep-rooted plumes (Courtillot *et al.* 2003) or that plumes are less important suppliers of basalt to the MTZ than downwelling slabs. An important observation that we can make based on our results is that the P410P and S410S amplitudes can be fit simultaneously by invoking lateral variations in basalt fraction and temperature without the need for additional complexities such as water or melt, in contrast to arguments by Zhou *et al.* (2022) from comparison of elastic properties of hydrous wadsleyite with seismic tomography and 410-topography models.

7.3 Comparison of seismic and thermodynamic parametrizations

In a final comparison, we show the inverted S -wave velocity profiles from both parametrizations in Fig. 13. The comparison indicates: (1) that the S -wave velocity gradients in the upper mantle and mid-to-lower MTZ obtained from the thermodynamic parametrization are steeper, that is smaller, than those from the seismic parametrization and (2) that the seismic MTZ models are more complex with double discontinuities at ~ 520 and ~ 560 km depth. The discrepancy in upper mantle velocity gradients accounts for the less well-fit fundamental-mode branch in the case of EA, whereas the velocity gradients in the MTZ account for the nature of the fourth and fifth overtone branches (cf. Figs 7g and h). The complex nature of the seismic MTZ, on the other hand, is not replicable with the thermodynamic models on account of the assumption of a compositionally homogeneous and adiabatic mantle, which explains our inability to fit the mid-MTZ waveform signal in detail (cf. Fig. 7f). In contrast, the seismic parametrization, because of more degrees of freedom, is capable of reproducing the observed waveforms and phase velocities.

To increase velocity gradients in the upper mantle and MTZ, in general, is likely to require radial chemical variations and/or deviations from an adiabatic geotherm, as indicated in seismic studies (e.g. Cobden *et al.* 2008; Cammarano *et al.* 2009; Khan *et al.* 2009; Munch *et al.* 2020; Fullea *et al.* 2021). Such variations are also seen in geodynamic models of Earth's mantle, which indicate (1) a subadiabatic mantle geotherm (Sinha & Butler 2007) and/or (2) the MTZ to be either enriched in harzburgite (e.g. Irifune *et al.* 2008), basalt (e.g. Karato 1997; Yan *et al.* 2020) or even granite (Kawai *et al.* 2013). Localized basalt enrichment from downwelling slabs has also been invoked as a means of accounting for the nature of the mid-MTZ (e.g. Saikia *et al.* 2008; Tian *et al.* 2020). In addition to

Table 4. Compilation of mantle transition zone (MTZ) discontinuity depths: 410 and 660 and MTZ thickness (ΔZ_{MTZ}).

Study		410 [km]	660 [km]	ΔZ_{MTZ} [km]
This study	Global	407.2 ± 7.5	651.5 ± 8.8	244.3 ± 7.7
	continents	402.1 ± 6.5	648.1 ± 8.6	246.1 ± 8.6
	oceans	411.0 ± 5.8	654.0 ± 8.1	243.0 ± 6.7
Shearer (1993)		413	653	240
Flanagan & Shearer (1998)		418	660	242
Gu <i>et al.</i> (2003)		409	649	240
Chambers <i>et al.</i> (2005b)		409.2	—	—
Deuss (2007)		410.1	652.6	242.5
Houser <i>et al.</i> (2008)		420	660	240
Lawrence & Shearer (2008)		—	—	242.3 ± 8.6
Huang <i>et al.</i> (2019)		416.8	661.2	244.4 ± 8.6
Waszek <i>et al.</i> (2021)	PP	414.4 ± 5.9	—	—
	SS	410.0 ± 4.7	654.1 ± 6.3	244.1 ± 6.2

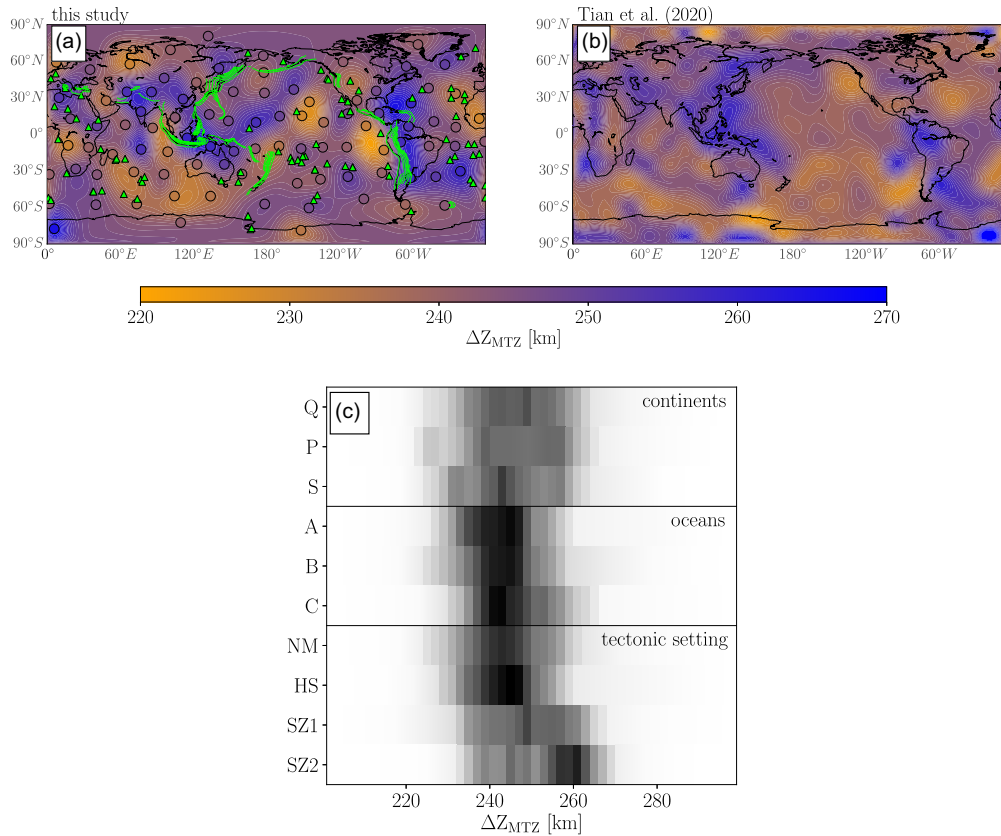


Figure 10. Mantle transition zone (MTZ) thickness. (a) Map of mean MTZ thickness (ΔZ_{MTZ}), which is compared to the global model of (Tian *et al.* 2020, b). Green triangles and lines indicate locations of hotspots (Anderson & Schramm 2005) and contours of slabs (Hayes *et al.* 2018). (c) Histograms of MTZ thickness sorted by tectonic setting. Continental (Q, P, S) and oceanic regions (A, B, C) follow the global tectonic regionalization ‘GTR1’ (Jordan 1981, *cf.* text for details). Additional results are shown for hotspots (HS; see map in a), subduction zones with a slab in the upper mantle (SZ1) or in the MTZ (SZ2; see map in a), and normal mantle (NM) away from hotspots and slabs.

temperature and composition, radial variations in grain size, which affect seismic velocities through attenuation, cannot be excluded (Faul & Jackson 2005). Besides adjusting the model parametrization, better agreement between seismic and thermodynamic models could also be achieved through improvements in the thermodynamic database and by accounting for associated uncertainties in the inversions (Connolly & Khan 2016). We leave it for a future study to investigate the exact nature of thermochemical variations required to improve the datafit of thermodynamic models.

Finally, upper mantle velocity gradients tend to be better fit by MM-based models. This is also reflected in the Bayes’ factor, which

provides a measure of the relative importance of any two physical characteristics of interest (Bernardo & Smith 1994; Khan *et al.* 2004; Bissig *et al.* 2022). The Bayes’ factor, which is computed from the sampled likelihood values (following Bissig *et al.* 2022, Fig. 12g), implies that MM appears to be more likely than EA in $\sim 2/3$ of the caps (no tectonic trend is evident). This supports the view of the mantle as a composite of equilibrated and mechanically mixed portions, in agreement with the results of Munch *et al.* (2020) and Bissig *et al.* (2022), who found that subcratonic mantle is better matched by MM and subduction zones by EA, respectively.

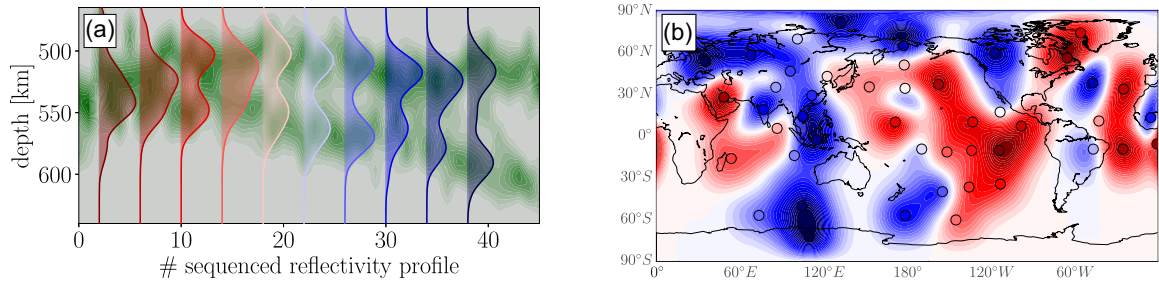


Figure 11. Spatial variation of mid-mantle transition zone structure. We apply the ‘sequencer’ to detect the main trend seen in inverted high-quality reflectivity-profiles of the structure between 465 and 640 km depth (normalized to the 410-discontinuity depth; *cf.* text for details). (a) Contour plot (shaded green in background) and average (colour-coded) of the sequenced reflectivity profiles. (b) Visualization of the spatial pattern of the sequenced profiles using the same colourmap as in (a).

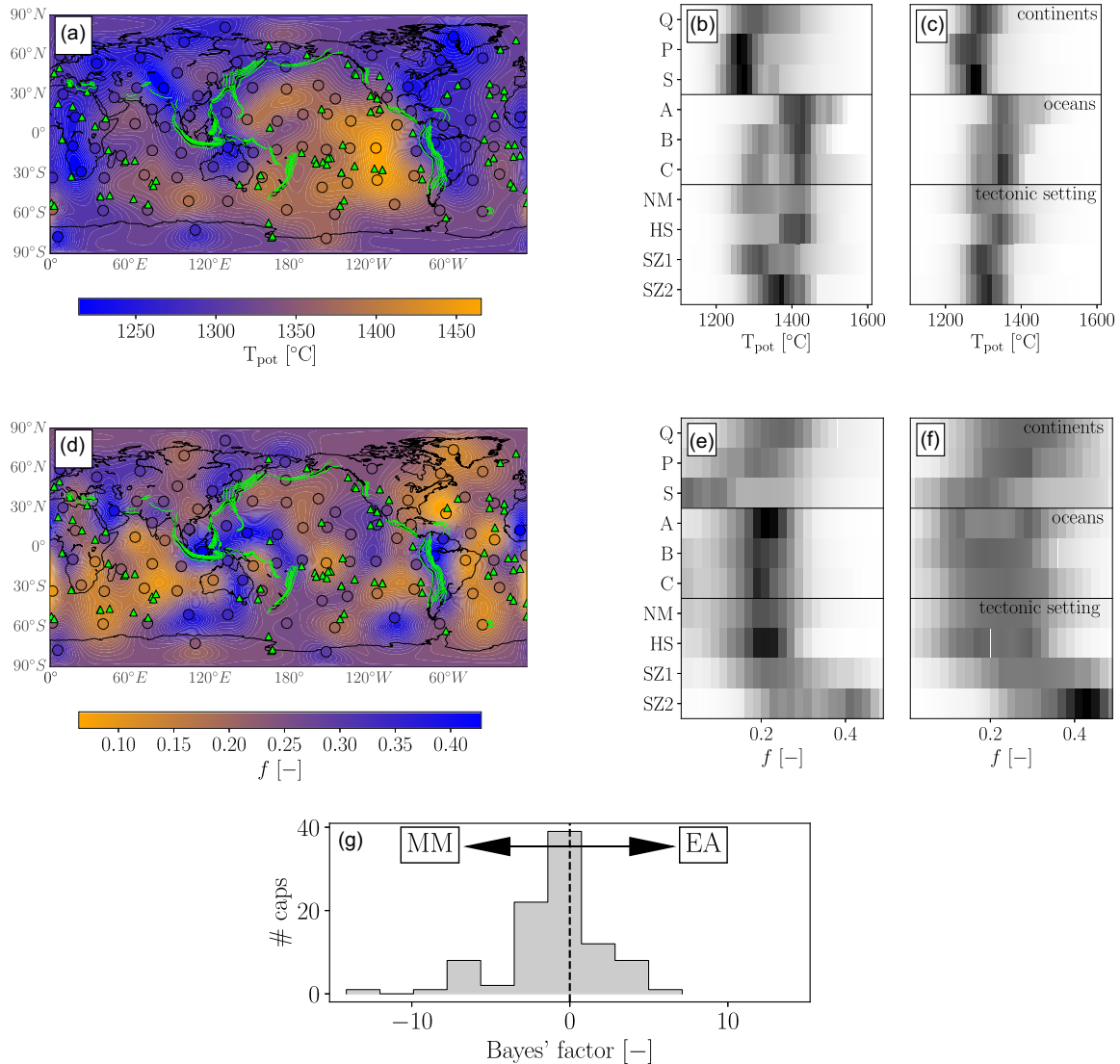


Figure 12. Thermochemical mantle structure. (a)–(c) Map of mean mantle potential temperature for a mechanically mixed mantle (T_{pot} ; a) and histograms sorted by tectonic setting for inversions using a fully equilibrated (b) and mechanically mixed mantle (c), respectively. Continental (Q, P, S) and oceanic regions (A, B, C) follow the global tectonic regionalization ‘GTR1’ (Jordan 1981, *cf.* text for details). In addition, results are shown for hotspots (HS; Anderson & Schramm 2005), subduction zones with a slab in the upper mantle (SZ1) or in the MTZ (SZ2; Hayes *et al.* 2018) and normal mantle (NM) away from hotspots and slabs. Hotspot locations and slab contours are indicated in (a) by triangles and lines, respectively. (d)–(f) Same as (a)–(c) but for composition, parametrized by basalt fraction (f). (g) Distribution of logarithmic Bayes’ factor (see main text for details) computed from the total misfit values (*cf.* eq. 8). For a Bayes’ factor < 0 , a mechanical mixture (MM) is favored over an equilibrium assemblage (EA) and vice versa for a Bayes’ factor > 0 .

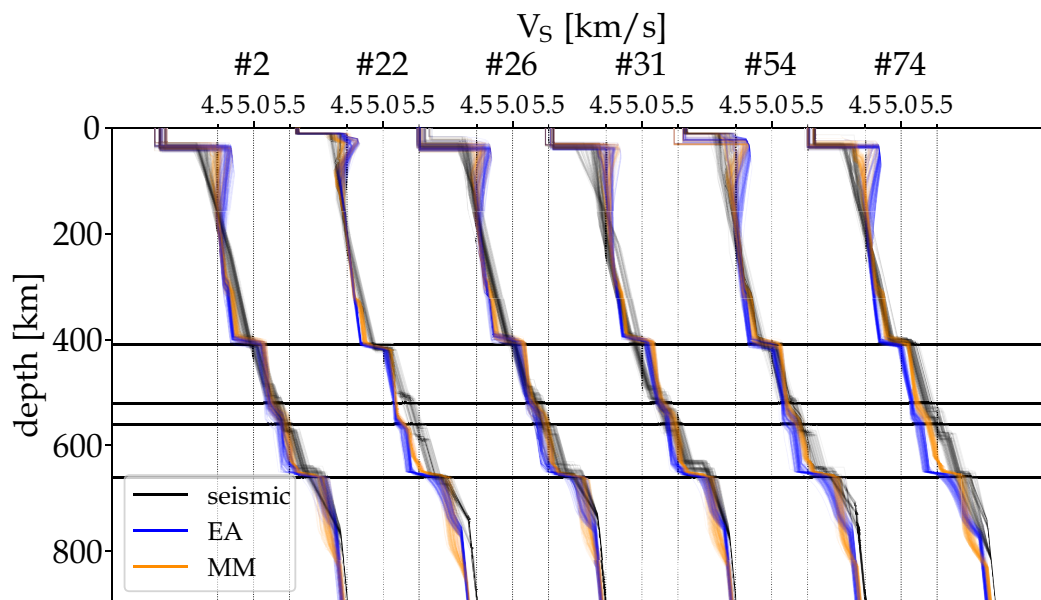


Figure 13. Sampled S -wave velocity (V_S) profiles for the seismic and thermodynamic inversions. For the latter, we rely on an equilibrium assemblage (EA) or a mechanical mixture (MM) model. 100 randomly selected profiles for each inversion are shown for caps #2 (Siberia), #22 (Northern Pacific), #26 (Eastern US), #31 (Central Africa), #54 (Indian Ocean) and #74 (Australia). Horizontal lines depict 410, 520, 560 and 660 km depth. Profiles for all caps are shown in Figs S36–S37.

8 CONCLUSIONS

In this study, we have described a method for the probabilistic inversion of PP and SS precursor waveform stacks and Rayleigh wave dispersion data for global seismic and thermochemical models of the upper mantle and mantle transition zone structure. Our set-up includes several advantages over more traditional methods: (1) simultaneous inversion for velocity variations and discontinuity depths by combining the sensitivities of surface and reflected body waves; (2) careful modelling of waveforms that account for effects arising from crustal structure, attenuation and data selection/processing, including, source characteristics and incoherent stacking; (3) inversion using seismic and thermodynamic parametrizations, which allows for more quantitative and direct inferences to be made about the thermochemical nature of the mantle and (4) the assessment of model variability by inverting for an ensemble of models in a probabilistic framework instead of single best-fitting models. Major findings include:

- (i) MTZ thickness appears to reflect surface tectonic processes and the presence of subducting slabs.
- (ii) Mid-MTZ structure trends from simple sub-ocean single to complex dual-discontinuity structure in subduction zones around the Pacific. The latter possibly represents the signature of oceanic crustal transport to the MTZ.
- (iii) Lower temperatures are observed below continents and subduction zones in comparison to oceans and hotspots, while subduction zones are enriched in basalt.
- (iv) Statistical analysis shows that a mechanically mixed mantle provides a better match to seismic data than an equilibrated mantle across $\sim 2/3$ of the globe.
- (v) While an adiabatic and isochemical mantle matches a large part of the seismic data, comparison with purely ‘seismic’ models points to the presence of additional complexities in the MTZ that require stepping beyond the assumption of chemical homogeneity and adiabaticity.

To better understand the nature of the aforementioned thermochemical complexities, improvements on both seismic and mineral physics sides are needed. First, the thermodynamic model parametrization should be extended to more complex models, that include radial variations in temperature, composition and mixing state. Secondly, additional seismic data sets, for example Love waves, receiver functions and triplicated waves, will prove useful as a means of further reducing model variability in a joint inversion. This also includes crust-sensitive data (e.g. underside reflected Moho phases) to reduce dependence on crustal model. Thirdly, improvements in measurements of the physical properties of MTZ minerals will aid in consolidating mineral physics and seismic data. Tian *et al.* (2020), for example, pointed to the necessity of studying the behaviour of (Ca, Al)-rich minor phases at MTZ-conditions to better understand mid-MTZ seismic observations.

ACKNOWLEDGMENTS

We thank Gabi Laske for editorial handling and Javier Fullea and an anonymous referee for constructive reviews. Peter Shearer and Thomas Bodin are also acknowledged for comments on an earlier version of the manuscript. This study is supported by a grant from the Swiss Federal Institute of Technology (ETH, project number ETH-05 17-1). Computations were performed on the clusters Euler (ETH) and Piz Daint (CSCS). Computations on Piz Daint were supported by the Swiss National Supercomputing Centre (CSCS) under project ID s1040. We thank Doyeon Kim for help with the sequencer and related input data processing, the Mondaic team for help with the Salvus software package (www.mondaic.com), and Dongdong Tian and Stéphanie Durand for providing MTZ thickness and seismic velocity models, respectively.

DATA AVAILABILITY

The facilities of Incorporated Research Institutions for Seismology (IRIS) Data Services, and specifically the IRIS Data Management Center, were used for access to waveforms and related metadata. IRIS Data Services are funded through the Seismological Facilities for the Advancement of Geoscience (SAGE) Award of the National Science Foundation under Cooperative Support Agreement EAR-1851048. Download and processing of seismic data were performed with ObspyDMT (Hosseini & Sigloch 2017) and Obspy (Krischer *et al.* 2015). Perple_X can be found on www.perplex.ethz.ch and phemgp on github.com/inverseproblem/Phemgp. We used the CMAES-implementation by Hansen *et al.* (2019). Maps were created by means of the Basemap Matplotlib Toolkit (matplotlib.org/basemap/). Tomographic models were downloaded from IRIS (ds.iris.edu/ds/products/emc-earthmodels/).

REFERENCES

- Afanasiev, M., Boehm, C., van Driel, M., Krischer, L., Rietmann, M., May, D.A., Knepley, M.G. & Fichtner, A., 2019. Modular and flexible spectral-element waveform modelling in two and three dimensions, *Geophys. J. Int.*, **216**(3), 1675–1692.
- Afonso, J.C., Fernández, M., Ranalli, G., Griffin, W.L. & Connolly, J.A.D., 2008. Integrated geophysical-petrological modeling of the lithosphere and sublithospheric upper mantle: Methodology and applications, *Geochem. Geophys. Geosyst.*, **9**(5), doi:10.1029/2007GC001834.
- Afonso, J.C., Salajegheh, F., Szwillus, W., Ebbing, J. & Gaina, C., 2019. A global reference model of the lithosphere and upper mantle from joint inversion and analysis of multiple data sets, *Geophys. J. Int.*, **217**, 1602–1628.
- Agee, C.B., 1998. Phase transformations and seismic structure in the upper mantle and transition zone, *Rev. Mineral.*, **37**, 165–203.
- Akaogi, M., Ito, E. & Navrotsky, A., 1989. Olivine-modified spinel-spinel transitions in the system $\text{Mg}_2\text{SiO}_4\text{--Fe}_2\text{SiO}_4$: calorimetric measurements, thermochemical calculation, and geophysical application, *J. geophys. Res.*, **94**(B11), 15 671–15 685.
- Anderson, D.L. & Schramm, K., 2005. Global hotspot maps, in *GSA SPECIAL PAPERS: Plates, Plumes and Paradigms*, Vol. **388**, pp. 19–29, eds Foulger, G., Natland, J.H., Presnall, D.C. & Anderson, D.L., Geological Society of America.
- Bagheri, A., Khan, A., Al-Attar, D., Crawford, O. & Giardini, D., 2019. Tidal response of Mars constrained from laboratory-based viscoelastic dissipation models and geophysical data, *J. geophys. Res.*, **124**, doi:10.1029/2019JE006015.
- Bai, L. & Ritsema, J., 2013. The effect of large-scale shear-velocity heterogeneity on SS precursor amplitudes, *Geophys. Res. Lett.*, **40**, 6054–6058.
- Ballmer, M.D., Schmerr, N.C., Nakagawa, T. & Ritsema, J., 2015. Compositional mantle layering revealed by slab stagnation at ~1000-km depth, *Sci. Adv.*, **1**(11), doi:10.1126/sciadv.1500815.
- Beghein, C., Trampert, J. & van Heijst, H.J., 2006. Radial anisotropy in seismic reference models of the mantle, *J. geophys. Res.*, **111**(B2), doi:10.1029/2005JB003728.
- Benz, H.M. & Vidale, J.E., 1993. Sharpness of upper-mantle discontinuities determined from high-frequency reflections, *Nature*, **365**, 147–150.
- Bernardo, J. & Smith, A., 1994. *Bayesian Theory*, John Wiley.
- Bissig, F., Khan, A. & Giardini, D., 2022. Evidence for basalt enrichment in the mantle transition zone from inversion of triplicated P- and S-waveforms, *Earth planet. Sci. Lett.*, **580**, doi:10.1016/j.epsl.2022.117387.
- Bissig, F., Khan, A., Tauzin, B., Sossi, P.A., Munch, F.D. & Giardini, D., 2021. Multifrequency inversion of Ps and Sp receiver functions: methodology and application to USArray data, *J. geophys. Res.*, **126**(2), doi:10.1029/2020JB020350.
- Calò, M., Bodin, T. & Romanowicz, B., 2016. Layered structure in the upper mantle across North America from joint inversion of long and short period seismic data, *Earth planet. Sci. Lett.*, **449**, 164–175.
- Cammarano, F., Romanowicz, B., Stixrude, L. & Lithgow-Bertelloni, C., 2009. Inferring the thermochemical structure of the upper mantle from seismic data, *Geophys. J. Int.*, **179**, 1169–1185.
- Chambers, K., Deuss, A. & Woodhouse, J.H., 2005a. Reflectivity of the 410-km discontinuity from PP and SS precursors, *J. geophys. Res.*, **110**(B2), doi:10.1029/2004JB003345.
- Chambers, K., Woodhouse, J.H. & Deuss, A., 2005b. Topography of the 410-km discontinuity from PP and SS precursors, *Earth planet. Sci. Lett.*, **235**, 610–622.
- Chang, S.-J., Ferreira, A.M.G., Ritsema, J., van Heijst, H.J. & Woodhouse, J.H., 2015. Joint inversion for global isotropic and radially anisotropic mantle structure including crustal thickness perturbations, *J. geophys. Res.*, **120**, 4278–4300.
- Chevrot, S., Vinnik, L. & Montagner, J.-P., 1999. Global-scale analysis of the mantle Pds phases, *J. geophys. Res.*, **104**(B9), 20 203–20 219.
- Cobden, L., Goes, S., Cammarano, F. & Connolly, J.A.D., 2008. Thermochemical interpretation of one-dimensional seismic reference models for the upper mantle: evidence for bias due to heterogeneity, *Geophys. J. Int.*, **175**(2), 627–648.
- Connolly, J.A.D., 2009. The geodynamic equation of state: what and how, *Geochem. Geophys. Geosyst.*, **10**(10), 1–19.
- Connolly, J.A.D. & Khan, A., 2016. Uncertainty of mantle geophysical properties computed from phase equilibrium models, *Geophys. Res. Lett.*, **43**, 1–9.
- Courtillot, V., Davaille, A., Besse, J. & Stock, J., 2003. Three distinct types of hotspots in the Earth's mantle, *Earth planet. Sci. Lett.*, **205**, 295–308.
- Day, E.A. & Deuss, A., 2013. Reconciling PP and P'P' precursor observations of a complex 660 km seismic discontinuity, *Geophys. J. Int.*, **194**, 834–838.
- de Viron, O., van Camp, M., Grabkowiak, A. & Ferreira, A.M.G., 2021. Comparing global seismic tomography models using varimax principal component analysis, *Solid Earth*, **12**, 1601–1634.
- Debayle, E., Bodin, T., Durand, S. & Ricard, Y., 2020. Seismic evidence for partial melt below tectonic plates, *Nature*, **586**, 555–559.
- Deuss, A., 2007. Seismic observations of transition zone discontinuities beneath hotspot locations, in *GSA SPECIAL PAPERS: Plates, Plumes and Planetary Processes*, Vol. **430**, eds Foulger, G.R. & Jurdy, D.M., Geological Society of America.
- Deuss, A., 2009. Global observations of mantle discontinuities using SS and PP precursors, *Surv. Geophys.*, **30**, 301–326.
- Deuss, A., Redfern, S.A.T., Chambers, K. & Woodhouse, J.H., 2006. The nature of the 660-kilometer discontinuity in Earth's mantle from global seismic observations of PP precursors, *Science*, **311**, 198–201.
- Deuss, A. & Woodhouse, J., 2001. Seismic observations of splitting of the mid-transition zone discontinuity in Earth's mantle, *Science*, **294**, 354–357.
- Dokht, R.M.H., Gu, Y.J. & Sacchi, M.D., 2016. Waveform inversion of SS precursors: an investigation of the northwestern Pacific subduction zones and intraplate volcanoes in China, *Gondwana Res.*, **40**, 77–90.
- Durand, S., Debayle, E. & Ricard, Y., 2015. Rayleigh wave phase velocity and error maps up to the fifth overtone, *Geophys. Res. Lett.*, **42**, 3266–3272.
- Dziewoński, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.*, **25**, 297–356.
- Dziewoński, A.M., Chou, T.-A. & Woodhouse, J.H., 1981. Determination of earthquake source parameters from waveform data for studies of global and regional seismicity, *J. geophys. Res.*, **86**(B4), 2825–2852.
- Efron, B. & Tibshirani, R., 1991. Statistical data analysis in the computer age, *Science*, **253**(5018), 390–395.
- Ekström, G., Nettles, M. & Dziewoński, A.M., 2012. The global CMT project 2004–2010: centroid-moment tensors for 13,017 earthquakes, *Phys. Earth planet. Inter.*, **200–201**, 1–9.
- Engdahl, E.R. & Flinn, E.A., 1969. Seismic waves reflected from discontinuities within Earth's upper mantle, *Science*, **163**(3863), 177–179.
- Estabrook, C.H. & Kind, R., 1996. The nature of the 660-kilometer upper-mantle seismic discontinuity from precursors to the PP phase, *Science*, **274**(5290), 1179–1182.

- Farra, V. & Vinnik, L., 2000. Upper mantle stratification by P and S receiver functions, *Geophys. J. Int.*, **141**(3), 699–712.
- Faul, U.H. & Jackson, I., 2005. The seismological signature of temperature and grain size variations in the upper mantle, *Earth planet. Sci. Lett.*, **234**(1–2), 119–134.
- Flanagan, M.P. & Shearer, P.M., 1998. Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors, *J. geophys. Res.*, **103**(B2), 2673–2692.
- Flanagan, M.P. & Shearer, P.M., 1999. A map of topography on the 410-km discontinuity from PP precursors, *Geophys. Res. Lett.*, **26**(5), 549–552.
- Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the reflectivity method and comparison with observations, *Geophys. J. Int.*, **23**(4), 417–433.
- Fukao, Y. & Obayashi, M., 2013. Subducted slabs stagnant above, penetrating through, and trapped below the 660 km discontinuity, *J. geophys. Res.*, **118**(11), 5920–5938.
- Fullea, J., Lebedev, S., Martinec, Z. & Celli, N.L., 2021. WINTERC-G: mapping the upper mantle thermochemical heterogeneity from coupled geophysical-petrological inversion of seismic waveforms, heat flow, surface elevation and gravity satellite data, *Geophys. J. Int.*, **226**, 146–191.
- Fullea, J., Muller, M.R. & Jones, A.G., 2011. Electrical conductivity of continental lithospheric mantle from integrated geophysical and petrological modeling: application to the Kaapvaal Craton and Rehoboth Terrane, southern Africa, *J. geophys. Res.*, **116**(B10), doi:10.1029/2011JB008544.
- Gossler, J. & Kind, R., 1996. Seismic evidence for very deep roots of continents, *Earth planet. Sci. Lett.*, **138**(1–4), 1–13.
- Grand, S.P. & Helmberger, D.V., 1984. Upper mantle shear structure of North America, *Geophys. J. Int.*, **76**(2), 399–438.
- Gu, Y.J. & Dziewoński, A.M., 2002. Global variability of transition zone thickness, *J. geophys. Res.*, **107**(B7), doi:10.1029/2001JB000489.
- Gu, Y.J., Dziewoński, A.M. & Agee, C.B., 1998. Global de-correlation of the topography of transition zone discontinuities, *Earth planet. Sci. Lett.*, **157**, 57–67.
- Gu, Y.J., Dziewoński, A.M. & Ekström, G., 2001. Preferential detection of the Lehmann discontinuity beneath continents, *Geophys. Res. Lett.*, **28**(24), 4655–4658.
- Gu, Y.J., Dziewoński, A.M. & Ekström, G., 2003. Simultaneous inversion for mantle shear velocity and topography of transition zone discontinuities, *Geophys. J. Int.*, **154**, 559–583.
- Guo, Z. & Zhou, Y., 2020. Finite-frequency imaging of the global 410- and 660-km discontinuities using SS precursors, *Geophys. J. Int.*, **220**, 1978–1994.
- Hansen, N., Akimoto, Y. & Baudis, P., 2019. CMA-ES/pycma on Github, *Zenodo*, doi:10.5281/zenodo.2559634.
- Hansen, N. & Ostermeier, A., 2001. Completely derandomized self-adaptation in evolution strategies, *Evol. Comput.*, **9**(2), 159–195.
- Hayes, G.P., Moore, G.L., Portner, D.E., Hearne, M., Flamme, H., Furtney, M. & Smoczyk, G.M., 2018. Slab2, a comprehensive subduction zone geometry model, *Science*, **362**(6410), 58–61.
- Heit, B., Yuan, X., Bianchi, M., Kind, R. & Gossler, J., 2010. Study of the lithospheric and upper-mantle discontinuities beneath eastern Asia by SS precursors, *Geophys. J. Int.*, **183**, 252–266.
- Helffrich, G.R. & Bina, C.R., 1994. Frequency dependence of the visibility and depths of mantle seismic discontinuities, *Geophys. Res. Lett.*, **21**(24), 2613–2616.
- Hosseini, K. & Sigloch, K., 2017. ObspyDMT: a Python toolbox for retrieving and processing large seismological data sets, *Solid Earth*, **8**, 1047–1070.
- Houser, C., 2016. Global seismic data reveal little water in the mantle transition zone, *Earth planet. Sci. Lett.*, **448**, 94–101.
- Houser, C., Masters, G., Shearer, P.M. & Laske, G., 2008. Shear and compressional velocity models of the mantle from cluster analysis of long-period waveforms, *Geophys. J. Int.*, **174**, 195–212.
- Huang, Q., Schmerr, N., Waszek, L. & Beghein, C., 2019. Constraints on seismic anisotropy in the mantle transition zone from long-period SS precursors, *J. geophys. Res.*, **124**, doi:10.1029/2019JB017307.
- Hwang, Y.K., Ritsema, J. & Goes, S., 2011. Global variation of body-wave attenuation in the upper mantle from teleseismic P wave and S wave spectra, *Geophys. Res. Lett.*, **38**(8), doi:10.1029/2011GL046812.
- Irifune, T., Higo, Y., Inoue, T., Kono, Y., Ohfuji, H. & Funakoshi, K., 2008. Sound velocities of majorite garnet and the composition of the mantle transition region, *Nature*, **451**, 814–817.
- Ita, J. & Stixrude, L., 1992. Petrology, elasticity, and composition of the mantle transition zone, *J. geophys. Res.*, **97**(B5), doi:10.1029/92JB00068.
- Jackson, I. & Faul, U.H., 2010. Grainsize-sensitive viscoelastic relaxation in olivine: towards a robust laboratory-based model for seismological application, *Phys. Earth planet. Inter.*, **183**(1–2), 151–163.
- Jenkins, J., Cottaar, S., White, R.S. & Deuss, A., 2016. Depressed mantle discontinuities beneath Iceland: evidence of a garnet controlled 660 km discontinuity?, *Earth planet. Sci. Lett.*, **433**, 159–168.
- Jordan, T.H., 1981. Global tectonic regionalization for seismological data analysis, *Bull. seism. Soc. Am.*, **71**(4), 1131–1141.
- Karato, S., 1997. On the separation of crustal component from subducted oceanic lithosphere near the 660 km discontinuity, *Phys. Earth planet. Inter.*, **99**(1–2), 103–111.
- Katsura, T., 2022. A revised adiabatic temperature profile for the mantle, *J. geophys. Res.*, **127**(2), doi:10.1029/2021JB023562.
- Kawai, K., Yamamoto, S., Tsuchiya, T. & Maruyama, S., 2013. The second continent: existence of granitic continental materials around the bottom of the mantle transition zone, *Geosci. Front.*, **4**(1), doi:10.1016/j.gsf.2012.08.003.
- Kemper, J., van Driel, M., Munch, F.D., Khan, A. & Giardini, D., 2021. A spectral element approach to computing normal modes, *Geophys. J. Int.*, **229**(2), 915–932.
- Kennett, B.L.N. & Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.*, **105**, 429–465.
- Khan, A., Boschi, L. & Connolly, J.A.D., 2009. On mantle chemical and thermal heterogeneities and anisotropy as mapped by inversion of global surface wave data, *J. geophys. Res.*, **114**(B9), 1–21.
- Khan, A., Boschi, L. & Connolly, J.A.D., 2011. Mapping the earth's thermochemical and anisotropic structure using global surface wave data, *J. geophys. Res.*, **116**(B1), doi:10.1029/2010JB007828.
- Khan, A., Mosegaard, K., William, J.G. & Lognonné, P., 2004. Does the Moon possess a molten core? Probing the deep lunar interior using results from LLR and Lunar Prospector, *J. geophys. Res.*, **109**(E9), doi:10.1029/2004JE002294.
- Khan, A., Zunino, A. & Deschamps, F., 2013. Upper mantle compositional variations and discontinuity topography imaged beneath Australia from Bayesian inversion of surface-wave phase velocities and thermochemical modeling, *J. geophys. Res.*, **118**, 1–22.
- Kim, D., Lekić, V., Ménard, B., Baron, D. & Taghizadeh-Popp, M., 2020. Sequencing seismograms: a panoptic view of scattering in the core-mantle boundary region, *Science*, **368**, 1223–1228.
- Koroni, M., Bozdağ, E., Paulssen, H. & Trampert, J., 2017. Sensitivity analysis of seismic waveforms to upper-mantle discontinuities using the adjoint method, *Geophys. J. Int.*, **210**, 1965–1980.
- Krischer, L., Megies, T., Barsch, R., Beyreuther, M., Lecocq, T., Caudron, C. & Wassermann, J., 2015. ObsPy: a bridge for seismology into the scientific Python ecosystem, *Comput. Sci. Discov.*, **8**(1), doi:10.1088/1749-4699/8/1/014003.
- Kustowski, B., Ekström, G. & Dziewoński, A.M., 2008. Anisotropic shear-wave velocity structure of the Earth's mantle: a global model, *J. geophys. Res.*, **113**(B6), doi:10.1029/2007JB005169.
- Langston, C.A., 1979. Structure under mount rainier, washington, inferred from teleseismic body waves, *J. geophys. Res.*, **84**(B9), 4749–4762.
- Laske, G., Masters, G., Ma, Z. & Pasyanos, M., 2013. Update on CRUST1.0 - a 1-degree Global Model of Earth's Crust, in *Proceedings of the EGU General Assembly 2013*, held 7–12 April, 2013 in Vienna, Austria, id. EGU2013-2658.
- Lau, H.C.P. & Romanowicz, B., 2021. Constraining jumps in density and elastic properties at the 660 km discontinuity using normal mode data via the Backus-Gilbert method, *Geophys. Res. Lett.*, **48**(9), doi:10.1029/2020GL092217.

- Lawrence, J.F. & Shearer, P.M., 2006. Constraining seismic velocity and density for the mantle transition zone with reflected and transmitted waveforms, *Geochem. Geophys. Geosyst.*, **7**(10), 1–19, doi:10.1029/2006GC001339.
- Lawrence, J.F. & Shearer, P.M., 2008. Imaging mantle transition zone thickness with SdS-SS finite-frequency sensitivity kernels, *Geophys. J. Int.*, **174**, 143–158.
- Lekić, V. & Romanowicz, B., 2011. Inferring upper-mantle structure by full waveform tomography with the spectral element method, *Geophys. J. Int.*, **185**(2), 799–831.
- Lessing, S., Thomas, C., Saki, M., Schmerr, N. & Vanacore, E., 2015. On the difficulties of detecting PP precursors, *Geophys. J. Int.*, **201**, 1666–1681.
- Ligorria, J.P. & Ammon, C.J., 1999. Iterative deconvolution and receiver-function estimation, *Bull. seism. Soc. Am.*, **89**(5), 1395–1400.
- Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. geophys. Res.*, **100**(B7), 12 431–12 447.
- Munch, F.D., Khan, A., Tauzin, B., van Driel, M. & Giardini, D., 2020. Seismological evidence for thermo-chemical heterogeneity in Earth's continental mantle, *Earth planet. Sci. Lett.*, **539**, 1–9.
- Munch, F.D., Khan, A., Tauzin, B., Zunino, A. & Giardini, D., 2018. Stochastic inversion of P-to-S converted waves for mantle thermal and compositional structure: methodology and application, *J. geophys. Res.*, **123**, 10 706–10 726.
- Niazi, M. & Anderson, D.L., 1965. Upper mantle structure of western North America from apparent velocities of P waves, *J. geophys. Res.*, **70**(18), 4633–4640.
- Oki, S. & Shearer, P.M., 2008. Mantle Q structure from S-P differential attenuation measurements, *J. geophys. Res.*, **113**(B12), doi:10.1029/2007JB005567.
- Revenaugh, J. & Jordan, T.H., 1991. Mantle layering from ScS reverberations: 2. The transition zone, *J. geophys. Res.*, **96**(B12), 19 763–19 780.
- 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 traveltimes and normal-mode splitting function measurements, *Geophys. J. Int.*, **184**(3), 1223–1236.
- Ritsema, J., Xu, W., Stixrude, L. & Lithgow-Bertelloni, C., 2009. Estimates of the transition zone temperature in a mechanically mixed upper mantle, *Earth planet. Sci. Lett.*, **277**(1–2), 244–252.
- Romanowicz, B.A. & Mitchell, B.J., 2015. Deep Earth structure: Q of the Earth from crust to core, in *Treatise on Geophysics*, 2nd edn, pp. 789–827, ed. Schubert, G., Elsevier.
- Rost, S. & Weber, M., 2001. A reflector at 200 km depth beneath the northwest Pacific, *Geophys. J. Int.*, **147**(1), 12–28.
- Rost, S. & Weber, M., 2002. The upper mantle transition zone discontinuities in the Pacific as determined by short-period array data, *Earth planet. Sci. Lett.*, **204**, 347–361.
- Saikia, A., Frost, D.J. & Rubie, D.C., 2008. Splitting of the 520-kilometer seismic discontinuity and chemical heterogeneity in the mantle, *Science*, **319**(5869), 1515–1518.
- Saki, M., Thomas, C. & Abreu, R., 2022. Detection and modelling of strong topography of mid-mantle structures beneath the North Atlantic, *Geophys. J. Int.*, **229**, 219–234.
- Schaeffer, A. & Lebedev, S., 2013. Global shear speed structure of the upper mantle and transition zone, *Geophys. J. Int.*, **194**(1), 417–449.
- Schmerr, N. & Garnero, E., 2006. Investigation of upper mantle discontinuity structure beneath the central Pacific using SS precursors, *J. geophys. Res.*, **111**(B8), 1–21.
- Shearer, P.M., 1991. Constraints on upper mantle discontinuities from observations of long-period reflected and converted phases, *J. geophys. Res.*, **96**(B11), 18 147–18 182.
- Shearer, P.M., 1993. Global mapping of upper mantle reflectors from long-period SS precursors, *Geophys. J. Int.*, **115**, 878–904.
- Shearer, P.M., 1996. Transition zone velocity gradients and the 520-km discontinuity, *J. geophys. Res.*, **101**(B2), 3053–3066.
- Shearer, P.M. & Flanagan, M.P., 1999. Seismic velocity and density jumps across the 410- and 660-kilometer discontinuities, *Science*, **285**, 1545–1548.
- Simmons, N.A., Forte, A.M., Boschi, L. & Grand, S.P., 2010. GyPSuM: a joint tomographic model of mantle density and seismic wave speeds, *J. geophys. Res.*, **115**(B12), doi:10.1029/2010JB007631.
- Simmons, N.A., Forte, A.M. & Grand, S.P., 2009. Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: Implications for the relative importance of thermal versus compositional heterogeneity, *Geophys. J. Int.*, **177**(3), 1284–1304.
- Sinha, G. & Butler, S.L., 2007. On the origin and significance of subadiabatic temperature gradients in the mantle, *J. geophys. Res.*, **112**(B10), doi:10.1029/2006JB004850.
- Stixrude, L. & Lithgow-Bertelloni, C., 2012. Geophysics of chemical heterogeneity in the mantle, *Annu. Rev. Earth planet. Sci.*, **40**, 569–595.
- Tajima, F., Katayama, I. & Nakagawa, T., 2009. Variable seismic structure near the 660 km discontinuity associated with stagnant slabs and geochemical implications, *Phys. Earth planet. Inter.*, **172**, 183–198.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society of Industrial and Applied Mathematics.
- Tauzin, B., Debayle, E. & Wittlinger, G., 2008. The mantle transition zone as seen by global Pds phases: no clear evidence for a thin transition zone beneath hotspots, *J. geophys. Res.*, **113**(B8), 1–17.
- Tauzin, B., Waszek, L., Ballmer, M.D., Afonso, J.C. & Bodin, T., 2022. Basaltic reservoirs in the Earth's mantle transition zone, *Proceedings of the National Academy of Sciences*, **119**(48), doi:10.1073/pnas.2209399119.
- Tian, D., Lv, M., Wei, S., Dorfman, S.M. & Shearer, P.M., 2020. Global variations of Earth's 520- and 560-km discontinuities, *Earth planet. Sci. Lett.*, **552**, doi:10.1016/j.epsl.2020.116600.
- van Driel, M., Krischer, L., Stähler, S.C., Hosseini, K. & Nissen-Meyer, T., 2015. Instaseis: instant global seismograms based on a broadband waveform database, *Solid Earth*, **6**, 701–717.
- van Stiphout, A.M., Cottaar, S. & Deuss, A., 2019. Receiver function mapping of mantle transition zone discontinuities beneath Alaska using scaled 3-D velocity corrections, *Geophys. J. Int.*, **219**, 1432–1446.
- Waszek, L., Schmerr, N.C. & Ballmer, M.D., 2018. Global observations of reflectors in the mid-mantle with implications for mantle structure and dynamics, *Nat. Commun.*, **9**(385), doi:10.1038/s41467-017-02709-4.
- Waszek, L., Tauzin, B., Schmerr, N.C., Ballmer, M.D. & Afonso, J.C., 2021. A poorly mixed mantle transition zone and its thermal state inferred from seismic waves, *Nat. Geosci.*, **14**, 949–955.
- Xu, F., Vidale, J.E., Earle, P.S. & Benz, H.M., 1998. Mantle discontinuities under southern Africa from precursors to P'P_{df}, *Geophys. Res. Lett.*, **25**(4), 571–574.
- Xu, W., Lithgow-Bertelloni, C., Stixrude, L. & Ritsema, J., 2008. The effect of bulk composition and temperature on mantle seismic structure, *Earth planet. Sci. Lett.*, **275**, 70–79.
- Yan, J., Ballmer, M.D. & Tackley, P.J., 2020. The evolution and distribution of recycled oceanic crust in the Earth's mantle: insight from geodynamic models, *Earth planet. Sci. Lett.*, **537**, doi:10.1016/j.epsl.2020.116171.
- Yu, C., Day, E.A., de Hoop, M.V., Campillo, M., Goes, S., Blythe, R.A. & van der Hilst, R.D., 2018. Compositional heterogeneity near the base of the mantle transition zone beneath Hawaii, *Nat. Commun.*, **9**(1266), 1–9.
- Zheng, Z., Ventosa, S. & Romanowicz, B., 2015. High resolution upper mantle discontinuity images across the Pacific Ocean from SS precursors using local slant stack filters, *Geophys. J. Int.*, **202**(1), 175–189.
- Zhou, W.-Y., Hao, M., Zhang, J.S., Chen, B., Wang, R. & Schmandt, B., 2022. Constraining composition and temperature variations in the mantle transition zone, *Nat. Commun.*, **13**, 1094, doi:10.1038/s41467-022-28709-7.
- Zunino, A., Connolly, J.A.D. & Khan, A., 2011. Pre-calculated phase equilibrium models for geophysical properties of the crust and mantle as a function of composition, *Geochem. Geophys. Geosyst.*, **12**(4), doi:10.1029/2010GC003304.

SUPPORTING INFORMATION

Supplementary data are available at [GJI](#) online.

suppl.data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.