Transition zone imaging beneath hotspots in the Pacific Ocean using SS precursors

Lisanne Jagt

May 10, 2017

MSc thesis

Supervisors:
Dr. Arwen Deuss
Dr. Ivan Pires de Vasconcelos

Abstract

Transition zone discontinuities at roughly 410 and 660 km depth result from phase transitions in mantle minerals. Depending on local temperature and composition, these seismic discontinuities are either depressed or elevated relative to a global average. According to mineral physics, hot olivine-dominated regions should have a thinner transition zone. However, in hot garnet-dominated regions both the 410- and 660-discontinuities may be depressed, resulting in no net change in transition zone thickness. Nevertheless, transition zone topography can provide essential clues for hotspot origin depth, which puts constraints on mantle convection. Hotspots fed by mantle plumes from the lower mantle is an argument in favour of whole-mantle convection. We have studied the transition zone in the mid-Pacific Ocean covering three known hotspots: Marqueses, Tahiti and Pitcairn. The transition zone is imaged with precursors of the SS-wave that reflect off the discontinuities. Imaging through slowness-time stacking is complicated by the large X-shaped Fresnel zone of underside reflections and presumed narrowness of mantle plumes. Therefore reverse time migration using the scalar wave equation and imaging techniques common in exploration seismics are applied to the same data set. Stacking results show that the transition zone is slightly thinner than average near the three hotspots, which might indicate a mantle plume origin with olivine-dominated discontinuities. Migrated images are blurred due to limited aperture, among other factors, but may show topography of the 410-discontinuity.

Contents

1 Introduction

1.1 Mineralogy, hotspots and mantle convection

The mantle transition zone is bounded by two first-order global seismic discontinuities at approximately 410 and 660 km depth. It separates the upper mantle from the lower mantle. The 410-discontinuity is defined by a phase transition of olivine to wadsleyite. This reaction has a positive Clapeyron slope, meaning that for a higher temperature, the phase change occurs at higher pressure, and vice versa (Figure 1). So mineral physics predicts a deeper 410-discontinuity in hot regions. For the 410-discontinuity, water may be an important factor. Water or melt could increase the width of the phase transition to 35 km (Karato, 2011). Wadsleyite transforms to ringwoodite at rougly 520 km depth, leading to a more intermittent seismic discontinuity (Deuss, 2009). At 660 km depth, ringwoodite breaks down to form perovskite and magnesiowüstite (Ringwood, 1975; Ita & Stixrude, 1992). This reaction has a negative Clapeyron slope, causing the 660-discontinuity to be more shallow in high temperature regions (Figure 2).

The mineralogy of the mantle consists of 40-60% olivine (Ringwood, 1975), so chemical reactions in other minerals may contribute to discontinuity topography and width. In a pyrolitic mantle, phase transitions in garnet and pyroxene may interfere, especially for the 660-discontinuity. High-pressure experiments showed that the transition from majorite garnet (to ilmenite at medium to low temperature) to Mg-perovskite occurs at similar depth ranges (660-700 km) and has a positive Clapeyron slope (Figure 2; Hirose, 2002). This reaction could become dominant over the ringwoodite to perovskite and magnesiowüstite transition in higher temperature regions. So in anomalously hot garnet-rich regions, a high temperature can correspond to a deeper 660-discontinuity and a transition zone of average thickness. Such anomalously hot regions are in theory mantle plumes, which have temperatures exceeding 1800°C required for the transition from majorite garnet to Mg-perovskite to be dominant (Hirose, 2002), although exact conditions at which majorite garnet is stable are still debatable (Yu et al., 2011).

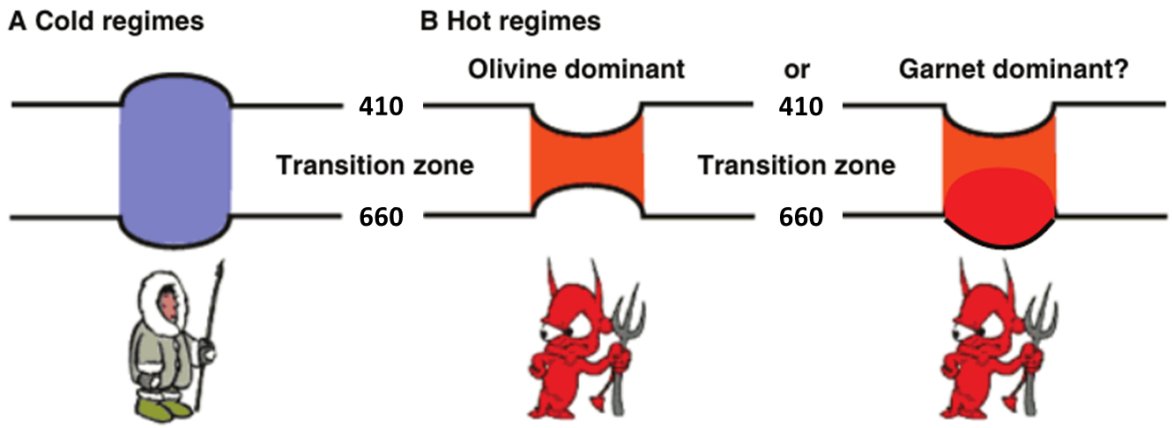

Figure 1: Cartoon showing different mantle transition zone configurations for two thermal regimes: A) Cold regime (e.g. subduction zone) where the transition zone is thicker. B) Hot regime (e.g. mantle plume) where the transition zone is thinner for purely olivine-controlled discontinuities, or of average thickness if garnet is involved. Deuss (2007), adapted from http://www.mantleplumes.org.

Whether the phase transition causing the 660-discontinuity has a positive or negative Clapeyron slope affects resistance to mass flow going through the transition zone (Hirose, 2002). Negative Clapeyron slopes have a tendency to impede mass flow, whereas positive slopes tend to promote mass flow (Davies, 2000). For hot bodies crossing a phase transition with positive

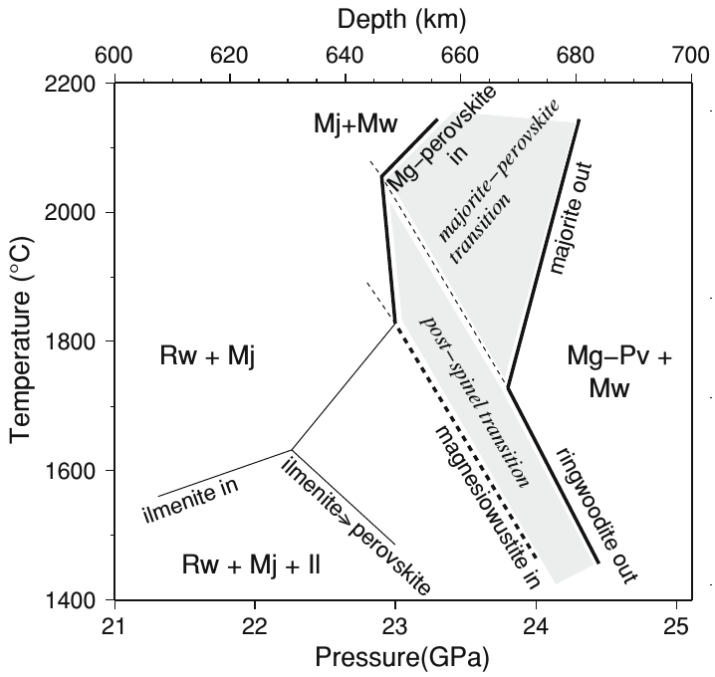

Figure 2: Phase diagram for a pyrolitic mantle and depth range 600 to 700 km, after Hirose (2002), showing the complex nature of the 660-discontinuity. For low temperatures, Mg-perovskite and magnesiowüstite are formed from ringwoodite over a relatively narrow depth range. The Clapeyron slope is negative. For high temperatures, Mg-perovskite forms from majorite garnet over a broader depth interval with a positive Clapeyron slope. Deuss (2009).

Clapeyron slope, transformation to the low-density phase occurs deeper in the hot body than in its colder surroundings. This creates a region where the low-density phase in the hot body exists next to the high-density phase, causing positive buoyancy force (upwards) and aiding its ascent. A negative Clapeyron slope results in a negative buoyancy force (downwards) in hot bodies, impeding its ascent and possibly causing ponding of hot material below the transition depth. The sign of the Clapeyron slope therefore has implications for mantle convection.

Using the depth of transition zone discontinuities as a thermometer is more complicated for the 660-discontinuity. Transition zone thickness, which does not depend strongly on mantle velocity corrections, has often been used as thermometer. Considering only olivine phase transitions, a thin transition zone corresponds to warm regions and a thick transition zone to cold regions (Figure 1). Absolute discontinuity depths have become more important, since knowledge about the complicated nature of the 660-discontinuity has added more complexity to the temperature-transition zone thickness relation. Because of this complexity, the absolute depth of the 410-discontinuity is more reliable as source of information regarding regional temperatures than the 660-discontinuity. Garnet-dominated 660-discontinuities may have been observed in some studies. In a global SS precursor study, Deuss (2007) found that two-thirds of 26 examined hotspots had a deeper 410-discontinuity paired with either a deep or shallow 660-discontinuity, consistent with hot regions in garnet-dominated or olivine-dominated mantle. A recent receiver function study found both discontinuities to be depressed beneath Iceland, of which the deeper (660 km) was interpreted as garnet-controlled (Jenkins et al., 2016).

Hotspots are the surface expressions of either hot rising mantle plumes caused by thermal convection in the mantle, or anomalously hot regions near the Earth's surface, resulting in volcanic activity. They are not necessarily close to spreading ridges. The origin of hotspots could be shallow (upper mantle), intermediate (transition zone) or deep (lower mantle; D" layer in particular), according to the classification by Courtillot et al. (2003). An intermediate-to-deep origin is associated with mantle plumes and would leave an imprint on the topography of the transition zone discontinuities, whereas a shallow origin would not. Transition zone observations can thus contribute to the debate on the existence of mantle plumes, along with tomography, bathymetry and mineralogy. Mantle plumes are difficult to resolve in seismic tomography due

to wavefront healing, since they are (supposedly) narrow and waves are more likely to travel through the faster surroundings. Delayed arrivals of waves that did travel through the plume are masked by earlier arrivals (Nolet et al., 2007). Hotspot origin depths can add useful insights to the discussion on whole-mantle convection versus layered mantle convection. Deep origins for the hot rising material feeding hotspots are an argument in favour of whole-mantle convection. It implies that the discontinuities have not formed a barrier to mass flow from the lower to the upper mantle.

1.1.1 Hotspots in the study area

The Pacific Ocean has a large low shear wave velocity province (LLSVP) in the lower mantle. It is a characteristic feature in images generated by seismic tomography (e.g. Ritsema et al., 1999). This superswell could feed secondary hotspots, which have origin depths just below the transition zone (Courtillot et al., 2003). As mentioned before, the other two types of hotspot proposed by Courtillot et al. (2003) are upper mantle hotspots, possibly linked to lithospheric break-up, and primary hotspots, anchored to the deepest lower mantle. The three Pacific hotspots in the study area of the south-central Pacific are Tahiti, Pitcairn and Marqueses. They all have at least two out of five characteristics for hotspots with a deep origin according to Courtillot et al. (2003). All three have a hotspot track, of which Marqueses has a long-lived track with potentially a flood basalt body or oceanic plateau at the start of its hotspot track (Clouard & Bonneville, 2001). All three hotspots have high buoyancy fluxes. Pitcairn and Tahiti presumably have high $^3He/^4He$ ratios, corresponding to material from a more primitive mantle reservoir. On the other hand, Marqueses has a low $^3He/^4He$ ratio, which might indicate a more shallow origin. None of them have a slow S-wave velocity anomaly at 500 km depth in the tomographic model S20RTS of Ritsema et al. (1999), whereas Tahiti is shown to have a strong plume emanating from the Pacific LLSVP by Montelli et al. (2006). Courtillot et al. (2003) argue that Pitcairn may join Tahiti in this class of hotpots originating at superswells. Mainly because of its long-lived hotspot track, Marqueses could be a so-called primary plume, originating at the core-mantle boundary, but the low helium isotope ratio is an argument refuting that statement.

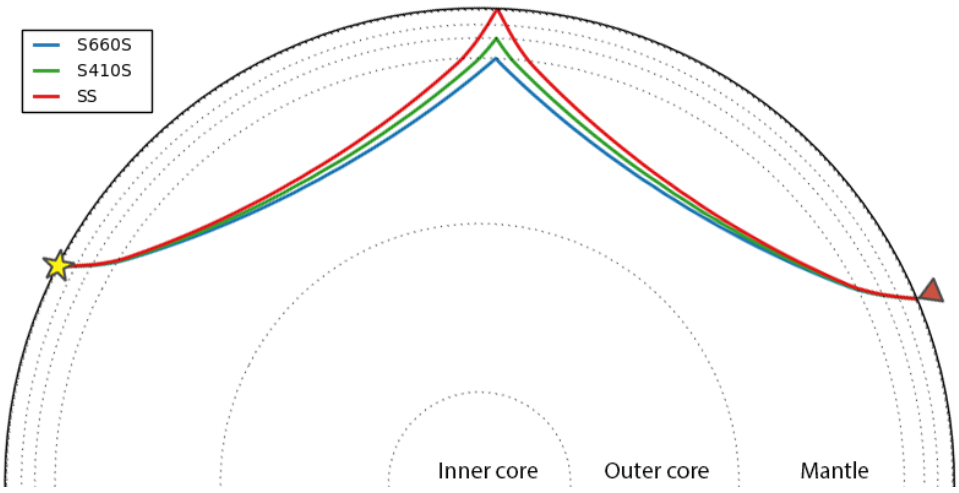

Figure 3: Specular ray paths of SS, S410S and S660S through PREM for 130° epicentral distance. From shallow to deep, the dashed lines represent the Moho, 220-discontinuity, 410-discontinuity, 660-discontinuity, core-mantle boundary and inner-outer core boundary. Note that ray paths vary the most near the bounce point, which is therefore the area that they are most sensitive to.

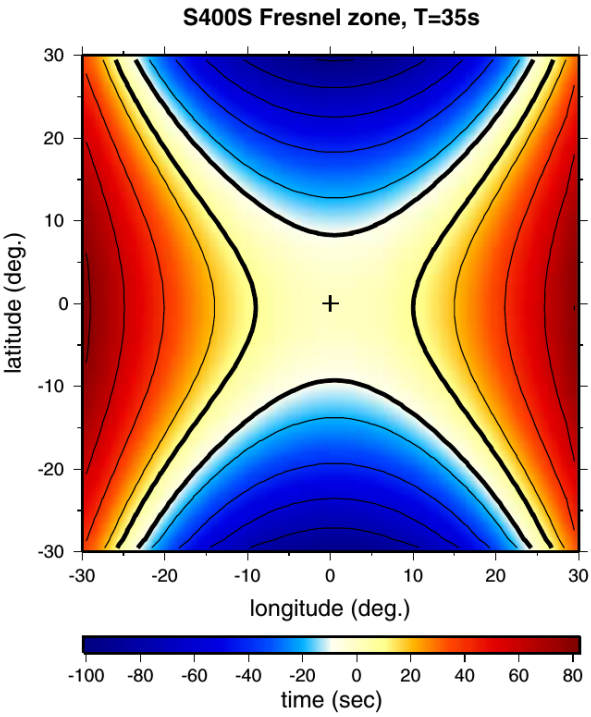

Figure 4: Fresnel zone for SS precursors with a period of 35 seconds reflected off a discontinuity at 400 km depth. The specular reflection is situated at the centre and contour lines show travel time differences with respect to the specular reflection. Thick solid line represents the Fresnel zone (travel time difference smaller than T/4). Adapted from Deuss (2009).

1.2 SS precursors and migration

The transition zone is usually imaged by underside reflections (either SS or PP precursors), ScS reverberations, or receiver functions (P-to-S conversions at interfaces). Receiver functions are sensitive to the mantle structure below receivers, whereas underside reflections are sensitive to the bounce point area midway on the source-receiver path (Figure 3). ScS reverberations are sensitive to mantle structure below receivers and source locations, as well as to the bounce point area. They are usually used in studies with small epicentral distances (10-50°) between source and receiver (Deuss et al., 2013). Due to the lack of seismic stations in oceans, underside reflections are preferred in oceanic regions. The 660-discontinuity is not always observed by PP precursors (Deuss et al., 2013), so SS precursors are the seismic phases used in this study.

A regular SS wave reflects off the Earth's surface at a bounce point roughly midway the source-receiver path (Figure 3). SS precursors follow nearly the same path as SS, but reflect off a seismic discontinuity, such as the 410- and 660-discontinuities in Figure 3. Their shorter distance results in an earlier arrival time than SS, hence the name SS *pre*cursor. The discontinuity reflection coefficients of roughly 10% (Shearer et al., 1999), caused by relatively low impedance contrasts compared to circa 100% at the surface, causes precursor amplitudes to be smaller than SS. Precursor phases of SS are named SdS, with d standing for depth of the discontinuity that they reflected off. S410S and S660S are the SS precursors that bounced off the 410- and 660-discontinuity, respectively. Using SS precursors for imaging purposes brings problems with it that have to be taken into consideration and can partly be adressed.

Ray theory assumes infinite frequency and that the rays are specular, i.e. reflect at a single point with equal angles of incidence and reflection. In reality, waves have finite frequency, causing them to propagate through a finite volume in space (Fresnel volume), instead of through an infinitely narrow ray. Finite frequency methods have therefore been gaining popularity over ray theory methods (Tromp et al., 2005). A cross-section of the Fresnel volume is the Fresnel zone. The first Fresnel zone is the area in which waves interfere constructively with each other. It is usually defined as the area around the specular (geometrical) reflection point for which

waves arrive within T/4 (T is dominant period) from the specular wave (Figure 4; Deuss, 2009). Seismic waves with a large Fresnel zone are sensitive to a large portion of a reflector, rather than a single point (Rost & Thomas, 2009).

The infinite-frequency assumption of ray theory breaks down for low-frequency underside reflections. Wave energy scattered from regions up to thousands of kilometres away from the theoretical reflection point may have travel times close to the travel time of the arrival following the specular ray path, and therefore contribute to the observed signal (Neele et al., 1997; Rost & Thomas, 2009). Underside reflections are mini-max phases, meaning that for the same epicentral distance and reflection point, waves travelling non-specular paths may arrive earlier or later than the specular ray (Cao et al., 2010). The Fresnel zone is saddle-shaped or X-shaped with the specular reflection both as maximum and minimum travel time in the centre of the saddle (Figure 4; Deuss, 2009). Energy reflected at points closer to the source or receiver along the source-receiver path (radial direction) arrives earlier than the specular SdS ray, because a larger portion of the path is spent at greater depths, where velocities are higher. On the other hand, energy reflected at points off the theoretical ray path (tangential direction) must travel longer distances and hence arrives later than the specular ray. These two effects balance out to a net similar travel time as the specular ray on the "arms" of the X-shaped Fresnel zone (Lawrence & Shearer, 2008). The Fresnel zone of SS precursors is approximately 1000 km (Neele et al., 1997), which is problematic for imaging smaller features. If discontinuity topography wavelength is smaller than the size of the Fresnel zone, the topography can be underestimated (Deuss, 2007), because the topography will be smoothed over a larger area (Niu et al., 2002). Mantle plumes are narrow features, although potentially wider than previously thought (Montelli et al., 2006), which makes it difficult to spot their impact on discontinuity topography with SS precursors. Receiver functions have a smaller Fresnel zone, but unfortunately the low density of seismic stations in oceans provides insufficient coverage, as mentioned before.

Migration methods aim to resolve the issue of wave energy being sensitive to an area by incorporating the effective size of the Fresnel zone and imaging the true reflectors. Migration is an inverse problem, since the data are known (recorded seismograms) and are used to invert for a model of mantle structure. In theory, each time sample in the data can be considered to be the sum of wave energy contributions from a constant travel time surface, i.e. an isochron (Cao et al., 2010). The most likely locations of scatterers are where these isochrons intersect (Chambers & Woodhouse, 2006).

Examples of migration methods are the Generalized Radon Transform (e.g. Cao et al., 2010) and using seismic array data (Rost & Thomas, 2009). Previously, Shearer et al. (1999) have tested a migration method for SS precursors based on the concept of Kirchhoff migration in reflection seismology. Each grid point i in their model acts as a hypothetical point scatterer of strength m_i. They posed a forward problem $\mathbf{d} = \mathbf{G}\mathbf{m}$, where $\mathbf{d}$ is a vector of observed data, $\mathbf{G}$ is a matrix describing the linear relationship between scattering strength in the model $\mathbf{m}$ and the seismic observations. The system has to be solved for $\mathbf{m}$, the model with scattering strength m_i at each grid point i. A simple backprojection method can be employed, $\mathbf{m} = \mathbf{G}^{\mathrm{T}}\mathbf{d}$, meaning that each time sample of the data is back-projected to all possible scattering points on the isochron, without any form of weighting (Shearer et al., 1999; Chambers & Woodhouse, 2006). Inversion methods can be fine-tuned by applying a weighted migration. The results of Shearer et al. (1999) for backprojection were disappointing, probably due to uneven data sampling, whereas those of Chambers and Woodhouse (2006) were more promising. Chambers & Woodhouse (2006) have divided the Earth in a set of equally spaced image points. To migrate scattered S and ScS phases, they have calculated predicted source-image point-receiver travel times for each image point and looked at the amplitudes in all seismic traces within 1 second of that predicted travel time. Summing the amplitudes for each image point provided an image

of the strength of mantle reflectors.

Backprojection, or reverse time migration, is the method of choice in this study. In reverse time migration, the recorded wave energy in the data is sent backward in time and spread out over a volume using the wave equation for wave field propagation. This constitutes the receiver wave field. Using a migration approach based on the wave equation takes finite frequency effects, such as the Fresnel zone and waveform propagation effects as function of frequency, into account. The scatterers are located where the receiver wave field interferes constructively with a source wave field propagated forward in time (Claerbout, 1971).

1.3 Aim, objectives and hypotheses

The main aim of this study is to image the transition zone discontinuities and transition zone thickness below known hotspots in the Pacific Ocean with SS precursors. This is achieved through regular slant-stacking of waveform data and allowing for finite frequency effects with reverse time migration applied to the data. Resulting images are then interpreted in terms of temperature, with the complications described before in mind, and translated to hotspot origin depth.

I expect to see variations in the two discontinuity depths in the study area, because the study area is larger than the Fresnel zone for underside reflections. Since Tahiti and Pitcairn are presumably of the same hotspot type (secondary hotspots), they probably have similar effects on discontinuity depths and transition zone thickness, unless one has a garnet-controlled and the other an olivine-controlled 660-discontinuity. In that case the temperature of the former has to be higher than the other and/or have a higher garnet content. Marqueses is thought to be a primary hotspot with a rooted mantle plume. It is difficult to distinguish between primary and secondary hotspots with discontinuity observations alone, because they are affected by both types. Therefore other kinds of observations are required to make interpretations. Migrating the data could potentially yield more realistic depths than those given by slant-stacking, because migration relaxes some of the assumptions made in slant-stacking. It depends on the limitations in data quality and coverage and of the method itself, which are unknown at this point, but will be explored.

This thesis consists of two parts. In part one I will use standard stacking techniques to visualize the precursors and get reflection depths below the bounce point region. The second part covers the reverse time migration applied to the data. It is a novel approach for regional seismology, using imaging techniques typically employed for hydrocarbon exploration purposes.

2 Part 1: Slant-stacking procedure

2.1 Method

2.1.1 Data selection and basic processing

The data selected for this study are retrieved from the database of the Incorporated Research Institutions for Seismology (IRIS). Several criteria regarding data selection have to be met.

The earthquake magnitudes should be high enough to have sufficient precursor energy and low enough to keep the source-time function relatively simple and pulse-like. A simple source-time function usually translates to an SS arrival formed by a single peak with side lobes, facilitating phase picking in a later stage. Moment magnitudes (M_W) in the ranges 6-7 (Deuss, 2009), 5.8-7.5 (Lessing et al., 2014), and > 5.8 (Saki et al., 2015) have been used before in precursor studies, among others. In this study a range of M_W=6-7 is used. If the SS arrival showed up as a clear spike in the unprocessed data and the half duration was less than 10 seconds, the event was taken into consideration. Focal mechanisms showing strike-slip faulting appeared to have more pronounced SS arrivals, possibly due to more S-wave energy being released from these types of earthquake. Bounce points of the specular rays should be located in the Pacific Ocean near the locations of Tahiti, Pitcairn and Marqueses (Courtillot et al., 2003). This makes earthquakes near Australia and New Zealand and receivers of the extensive Transportable Array (TA) in the USA particularly suitable. Station spacing of the TA is 70 km approximately, so bounce points are spaced ca. 35 km apart (Zheng et al., 2015). Three events west of Macquarie Island (southwest Pacific Ocean; south of New Zealand) are chosen for further processing (Table 1; Figure 5). Their moment magnitudes are 6.3, 6.4 and 6.5 and all half duration times are under 5 seconds. Bounce point locations are plotted in Figure 5. Unfortunately, the hotspots are located at the rims of the densely covered region. It would have been favourable to have an event with a different azimuth as well, to avoid bias of mantle structure in one direction. Alas, such earthquakes were not found.

Typical discontinuity reflection coefficients for SdS are less than 10% (Shearer et al., 1999), so in order to get at least some precursor energy, incidence angles have to be favourable. Incidence angle, and hence epicentral distance, affects the reflection coefficient (Lessing et al., 2014). For incidence angles of approximately 30° to 50° for the 660-discontinuity and 25° to 40° for the 410-discontinuity, corresponding to epicentral distance ranges of 100-180 degrees (Deuss, 2009), the reflection coefficient is higher than 0.01 (Lessing et al., 2014). In addition, for epicentral distances larger than 160 degrees and smaller than 100 degrees, other phases start interfering with SS and its minor arc precursors, such as ScSScS precursors (Zheng & Romanowicz, 2012), postcursors of Sdiff (core-diffracted S-wave) (Zheng et al., 2015), and major arc SS precursors (Deuss, 2009). Therefore the epicentral distance range is limited to 100-160 degrees. Schmerr & Garnero (2006) argue that within this range other phases still interfere with SS precursors, especially for deeper sources (∼75 km). For example, in their study they

Date	Longitude (°E)	Latitude (°N)	Depth (km)	M_W	Half-duration (s)	# Stations
06-01-2008	150.0487	-59.5987	10	6.5	4.3	510
08-09-2008	153.9955	-60.7493	10	6.3	3.9	312
06-05-2011	147.2350	-55.8442	14	6.4	4.0	459

Table 1: Information of the three events used in this study. The last column shows the number of stations (and seismograms) per event included in slant-stacking. Information is provided by IRIS and the Global Centroid Moment Tensor catalogue (Dziewonski et al., 1981; Ekström et al., 2012).

excluded the distance range 100-120 degrees from S660S arrivals, because s410sS and s660sS (topside reflections) arrive at similar times and could contaminate the stack. These interfering phases travel with different slownesses than SS precursors, so in vespagrams (or slowness-time plots) they should appear as separate phases. Furthermore, contamination by other phases does not matter too much if the amplitudes after stacking are not interpreted. However, the peak can shift up to 5 km due to contamination (Schmerr & Garnero, 2006).

The records start at the event time and end 60 minutes after the event time. Only the long period channels (∼1 Hz sampling rate) of high gain seismometers are used, since SS phases are long period waves. Ideally the north component has an azimuth of 0 degrees and the east component of 90 degrees, but this is not always the case. Data with component azimuths that differ more than 5 degrees from their ideal azimuth are removed, because they could cause trouble when the east and north components are rotated to the great circle path. Subsequently, the ZNE coordinate system is rotated to the ZRT system, with R as radial component along the great circle path and T as transverse component perpendicular to the great circle path. S-wave energy is favoured by the transverse component, so that is the component used for further processing (Deuss, 2009).

The instrument response is removed by deconvolution. The data are cut -800 seconds to +400 seconds around the theoretical SS arrival, calculated using ak135 (Kennett et al., 1995) and travel time software *ttimes* (Kennett & Engdahl, 1991), and bandpass filtered using a two-way (acausal) fourth order butterworth filter with corner periods of 15 and 75 seconds,

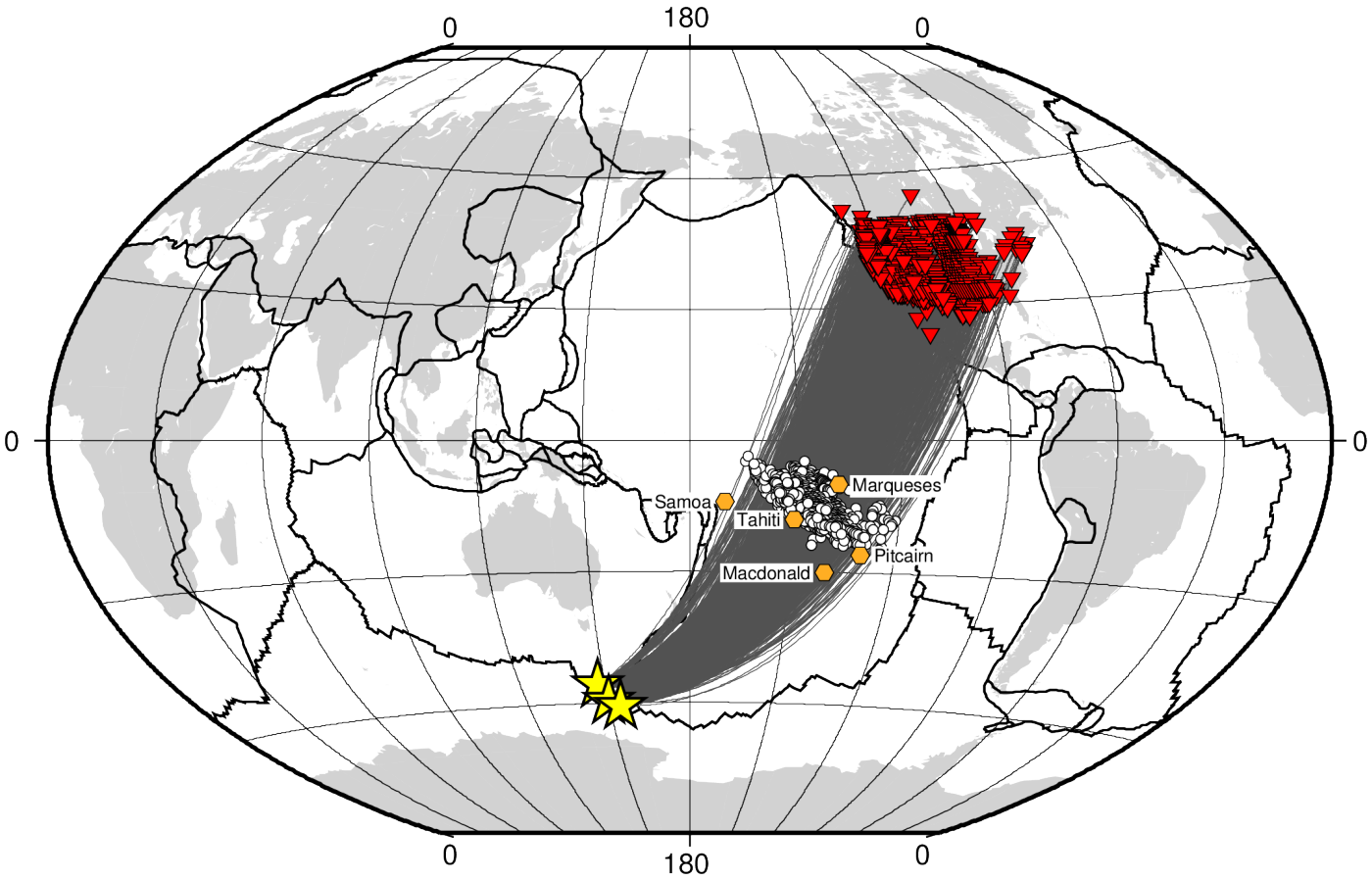

Figure 5: Stations (red triangles), sources (yellow stars) and midway points (white dots) for the most densely covered bounce point region. Nearby hotspots are indicated by orange hexagons.

which is typical for SS precursor studies (Deuss, 2009), although other corner periods have been applied as well (e.g. 20-100 s in Zheng & Romanowicz, 2012). Filtering between 10 and 100 seconds did not produce different results. The trend and mean are also removed. After filtering, seismograms with a signal-to-noise (S/N) ratio higher than 3 are kept for further processing. The signal is defined as the highest amplitude in the time window -60 to +60 seconds relative to the theoretical SS arrival. Noise is the highest amplitude in the time window of -300 to -60 seconds (i.e. precursor time window) relative to the theoretical SS arrival. The time window for determining noise is not extended beyond -300 seconds, because Sdiff arrives at approximately -400 seconds and has a much larger amplitude than SS precursors. The true SS arrivals are picked manually to ensure better quality of the picks. Polarities are reversed if necessary (for 22 out of 1281 seismograms) and the maximum amplitude of SS is normalized to unity to equalize energy across all events.

2.1.2 Stacking

The amplitude of SS precursors is typically about 5-10% of the SS pulse in individual seismograms (Figure 6), requiring stacking to enhance their visibility (Shearer et al., 1999; Zheng et al., 2015). Stacking can be done in many ways. One way is stacking with epicentral distance bins. The data are divided into epicentral distance range bins with a width of 5 degrees and simply stacked per bin without slowness corrections. This procedure yields clearly visible precursors for a global data set (Deuss, 2009), but was less successful for this regional study. Regional differences in mantle structure and not accounting for slowness variations probably made this method unsuitable for a smaller data set.

A more successful stacking method concerns making bounce point bins. The bounce points (midway on the source-receiver path) of the three events are plotted in Figure 5. The most densely sampled region lies between 0 to 30 degrees south and 190 to 240 degrees east in the

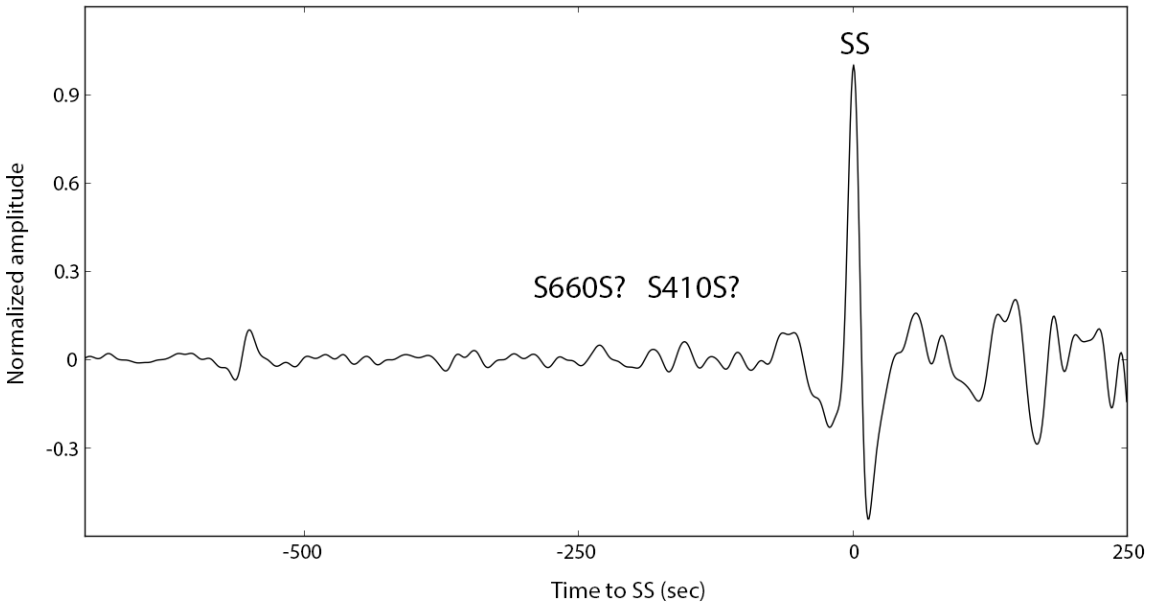

Figure 6: Example of an individual seismogram (network: TA, station: Z35A), after going through all processing steps explained in the text. SS is the large peak at zero seconds. S410S and S660S should be between -100 and -300 seconds, but do not exceed noise levels in a single seismogram.

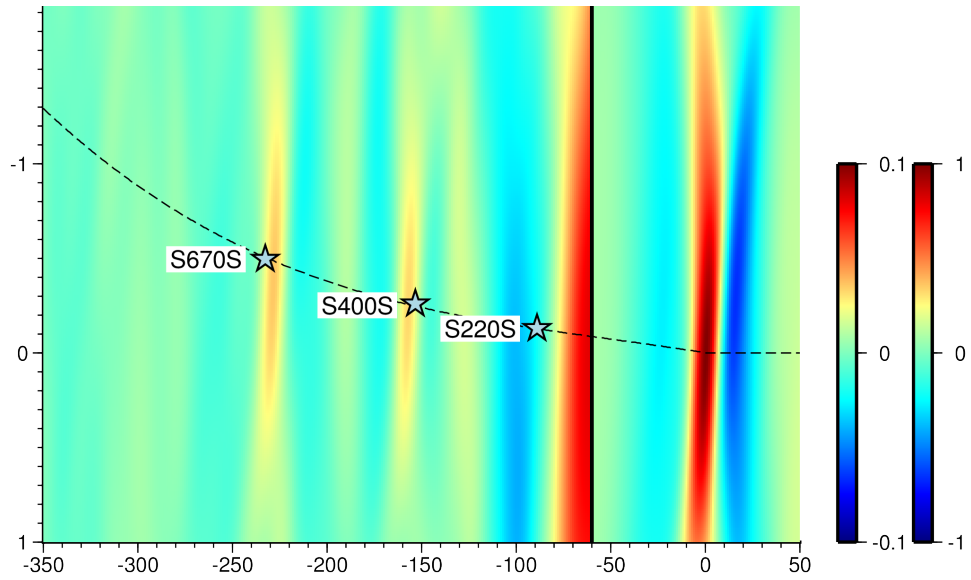

Figure 7: Vespagram for cap 9 (see Figure 8a for bin numbering) with relative slowness w.r.t. SS on the y-axis and time w.r.t. SS on the x-axis. Dashed line is the PREM profile line. Slowness-time relations for three discontinuities in PREM are annotated. Colour scale is enhanced for the part left of the solid black line (left colour bar), to visualize precursor phases.

Pacific Ocean. Hotspots from Courtillot's catalogue (2003) that are potentially covered by bounce point locations are Tahiti, Marqueses and Pitcairn (Figure 5). Samoa and Macdonald are too far away from the bounce points. This region can be divided into bounce point bins with a width of 10 degrees, roughly the same dimensions as the size of the Fresnel zone of SS precursors.

SS precursors travel a slightly different path than the SS wave, since they do not reach the Earth's surface. This difference in ray path results in a difference in slowness. Slownesses of SS precursors are smaller than the slowness of SS, because they do not travel through the slow surface layers of the Earth at the bounce point location. Slowness differences are visible in epicentral distance stacking: for larger distances, S410S and S660S arrive later with respect to the SS phase. Move out stacking can correct for slowness differences. The SS phase is the reference phase at time zero and slowness zero. The data are stacked for a range of slownesses (-2 to 1 sec/deg relative to SS slowness) with increment 0.01 sec/deg, generating vespagrams or slowness-time stacks (e.g. Deuss, 2009). A reference epicentral distance needs to be chosen for relative time shifts before slant stacking. The average of all epicentral distances in the data set (128 degrees) was chosen as reference epicentral distance. This reference epicentral distance is not ideal for some bins with a different average epicentral distance, but opting for a single reference value is more convenient in the time-to-depth conversion. The vespagrams are cross-cut by a dashed curved line representing SS precursor slownesses and arrival times in PREM (Dziewonski & Anderson, 1981; Figure 7), calculated with the TauP toolkit (Crotwell et al., 1999). Precursor arrivals should fall on this curve and have the same polarity as the SS arrival. In Figure 7 the stacked SS amplitude is positive (red), so the precursors should be positive as well. The further away from the optimum stacking slowness for the specific phases, the more diffuse the peaks become. Making a cross-section along this line results in stacked traces with clearly visible S410S and S660S precursor reflections. Time has been converted to depth using travel times of SdS phases in PREM. Time-depth relations in between those travel times and reflection depths were obtained with linear interpolation. The validity of using a

laterally homogeneous velocity model for this conversion will be discussed later. Depths of the 410- and 660-discontinuities and transition zone thickness for each cap have been inferred from the time-depth relation.

2.2 Results

The vespagrams do not show coherent contaminating phases that could cause a peak shift for S410S and S660S (Figure 7). Cross-sections through vespagrams along the PREM slowness-time curve are shown in Figure 8 for each bin. Global averages for transition zone thickness, 410-discontinuity depth and 660-discontinuity depth have been estimated at 242.5 km, 410.1 km, 652.6 km respectively (Deuss, 2007). These averages mark the centres of the colour scales in Figures 9, 10 and 11. Transition zone thickness for each bin is illustrated in Figure 9. The transition zone is 1 to 5 km thinner than the global average near hotspot locations of Marqueses, Tahiti and Pitcairn. Depths of the 410-discontinuity are plotted in Figure 10. A deeper 410-discontinuity most likely indicates a hot region. Relative abundances of garnet and olivine are not important at this depth. In the entire region southeast of the line between Marqueses and Tahiti the 410-discontinuity is relatively depressed, whereas it is elevated northwest of this line. It is not significantly deeper close to the hotspot locations. The nature of the 660-discontinuity depth (Figure 11) is a bit more complicated, as stated in the Introduction. Its topography depends on whether the discontinuity is garnet-controlled or olivine-controlled. Near Marqueses, the 660-discontinuity is 7 km more shallow than the global average, causing the transition zone to be thinner than average. Near Tahiti and Pitcairn the 660-depth is around the global average, whereas the region in between these two hotspots has a 7 km deeper 660-discontinuity.

2.3 Robustness of the peaks

As an estimation of the robustness of the peaks in the stacked traces over time, a 95% confidence interval for the mean (equal to the stacked trace) is calculated for each time sample in each stack. The 95% confidence interval is bounded by the sample mean plus $1.96\sigma_m$ and the mean minus $1.96\sigma_m$, where σ_m is standard error in the mean (equation 1; Squires, 2001). Standard error is a measure of the precision of the sample mean, so a smaller σ_m implies a more precisely stacked trace. This error analysis assumes a normal distribution.

$$\sigma_m \approx \left(\frac{1}{n-1}\right)^{\frac{1}{2}} s \tag{1}$$

In equation 1 n is the number of seismograms in the stack and s is the standard deviation:

$$s^2 = \frac{1}{n}\sum d_i^2 \tag{2}$$

The residual d_i is the difference between the amplitude x_i in a single trace and the stacked (mean) amplitude $\overline{x}_i$ at one time sample.

Figure 12 shows the stacked traces for all bins with the 95% confidence interval plotted as well. S410S and S670S are robust peaks in all traces. The general trend is that bins with less traces have less robust peaks. An increase in the sample size leads to a smaller standard error in the mean. Furthermore, standard error decreases with decreasing standard deviation, which happens when all individual traces have a peak of similar amplitude. A robust S520S is probably observed in bins 12-15 (Figure 12).

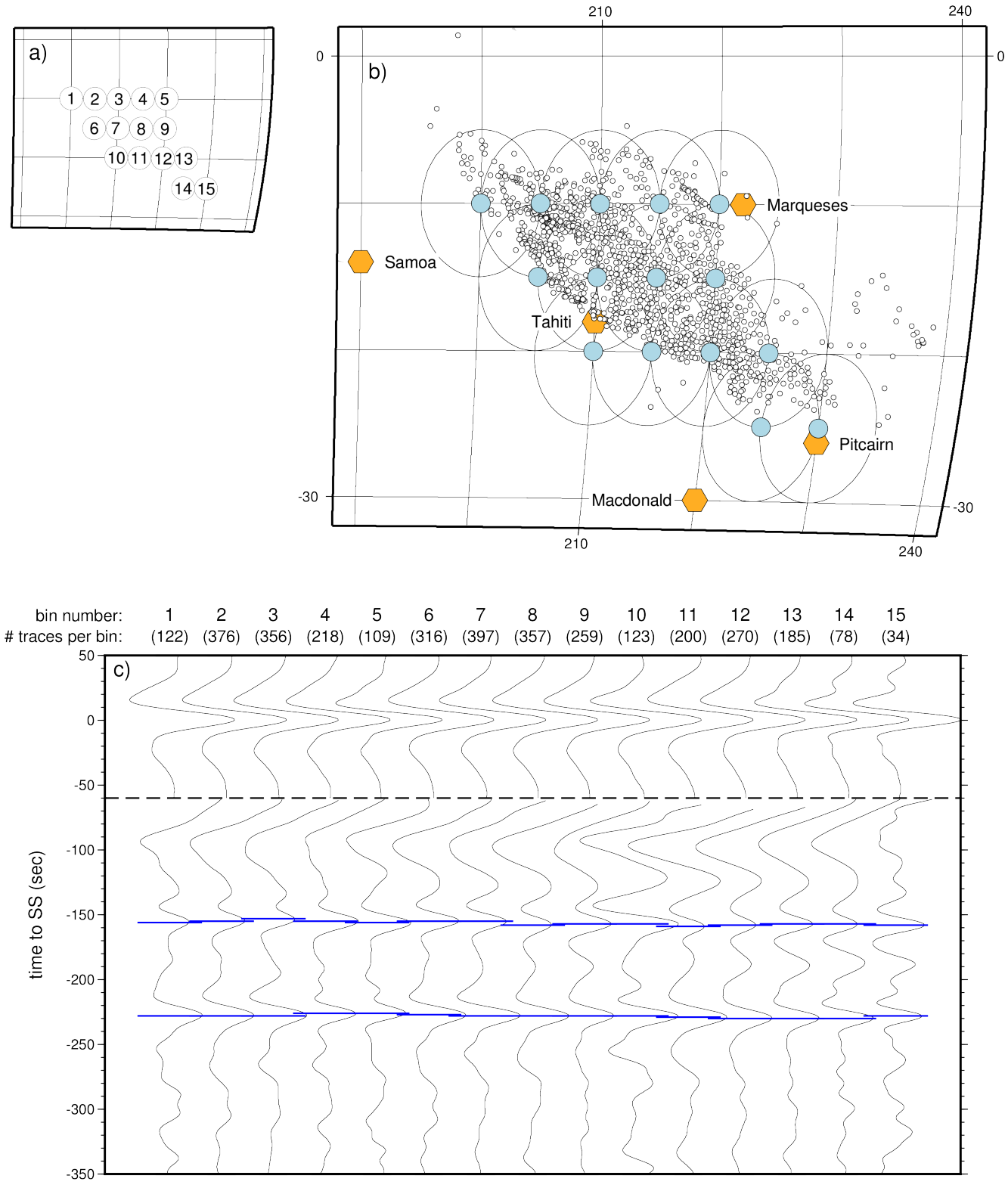

Figure 8: a) Numbering of the bins; b) Outlines (black circles) and midpoints (blue dots) of the 5˚ radius caps; c) Cross-sections through vespagrams along the PREM slowness-time curve. The 410 and 660 precursors are indicated by blue lines. Scale is enhanced below the dashed line to amplify the precursors.

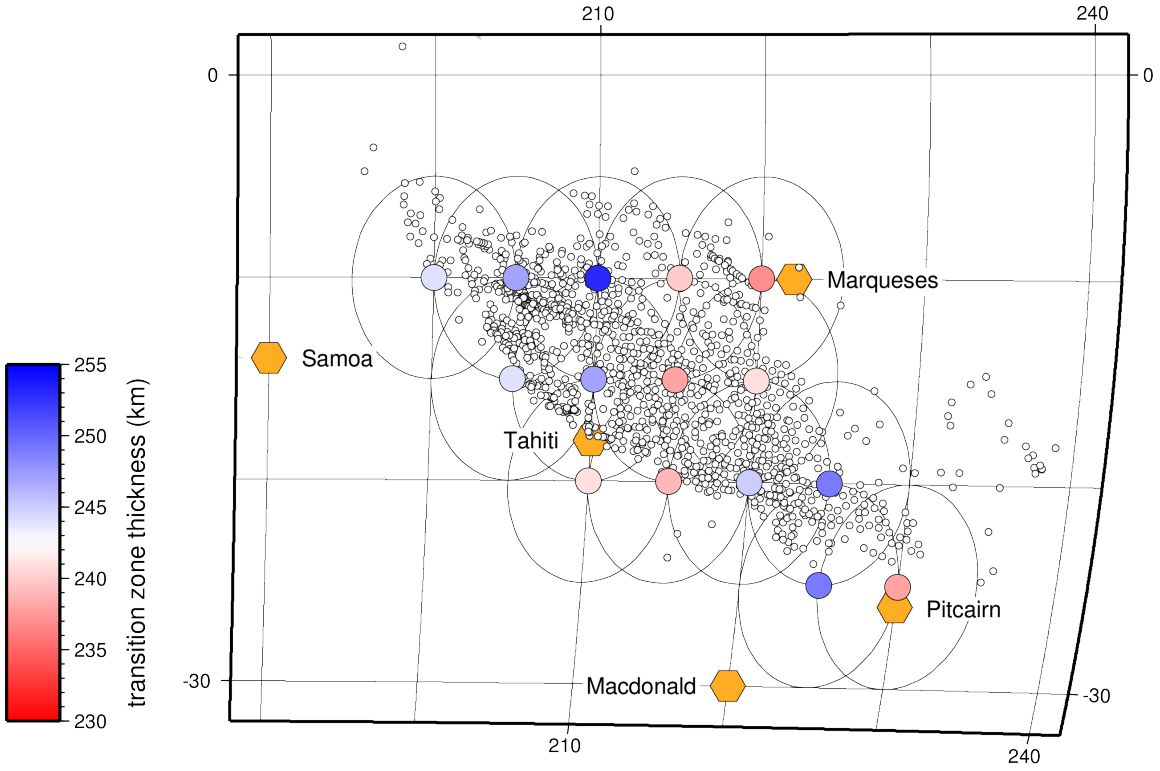

Figure 9: Transition zone thickness (km) per bin

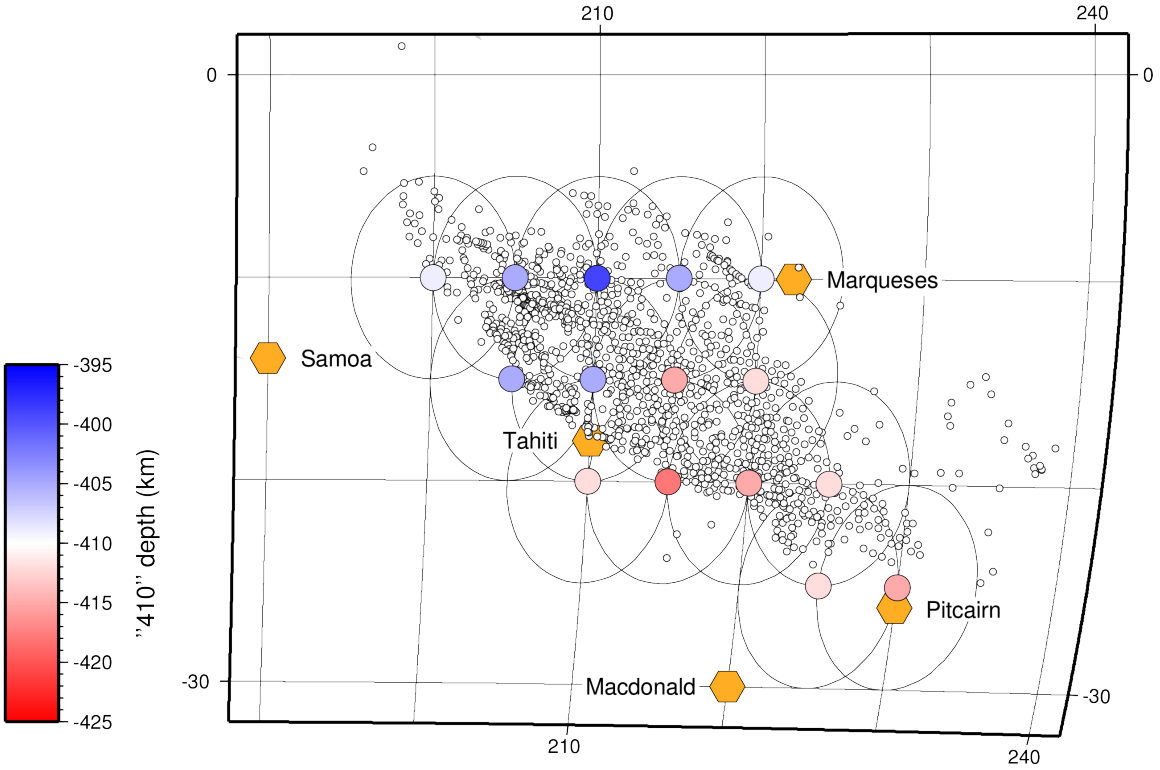

Figure 10: Depth of 410-discontinuity (km) per bin

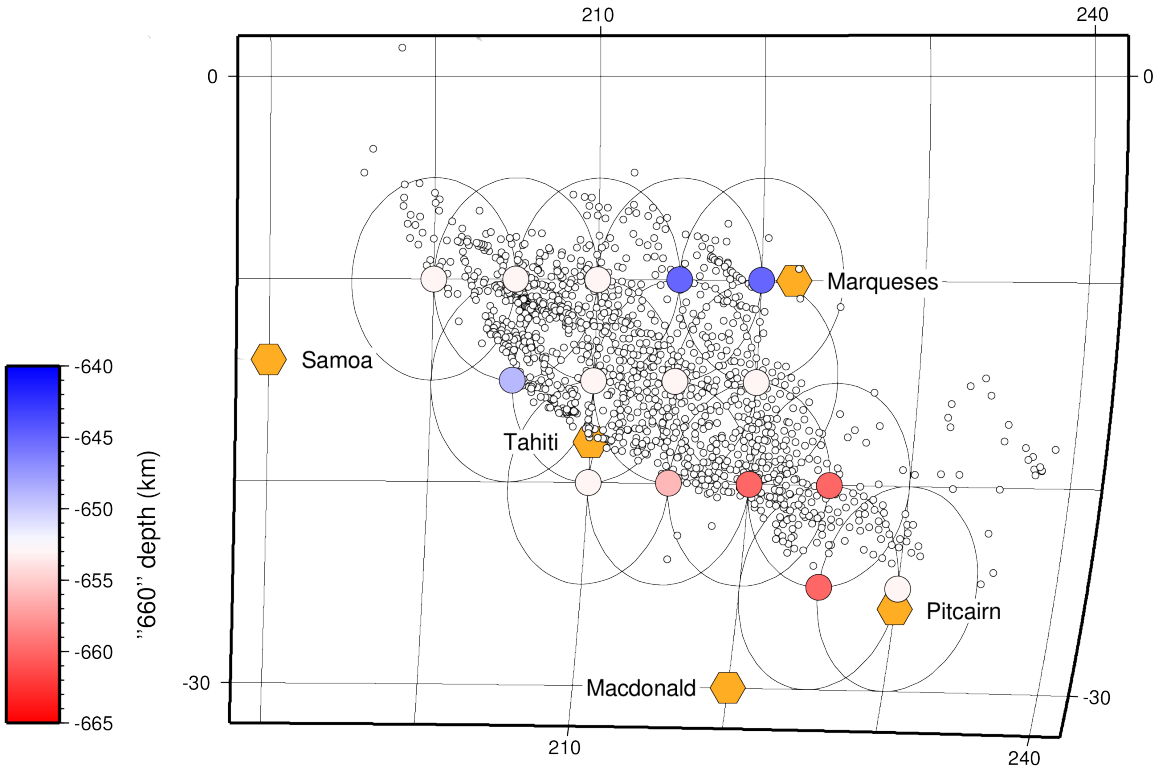

Figure 11: Depth of 660-discontinuity (km) per bin

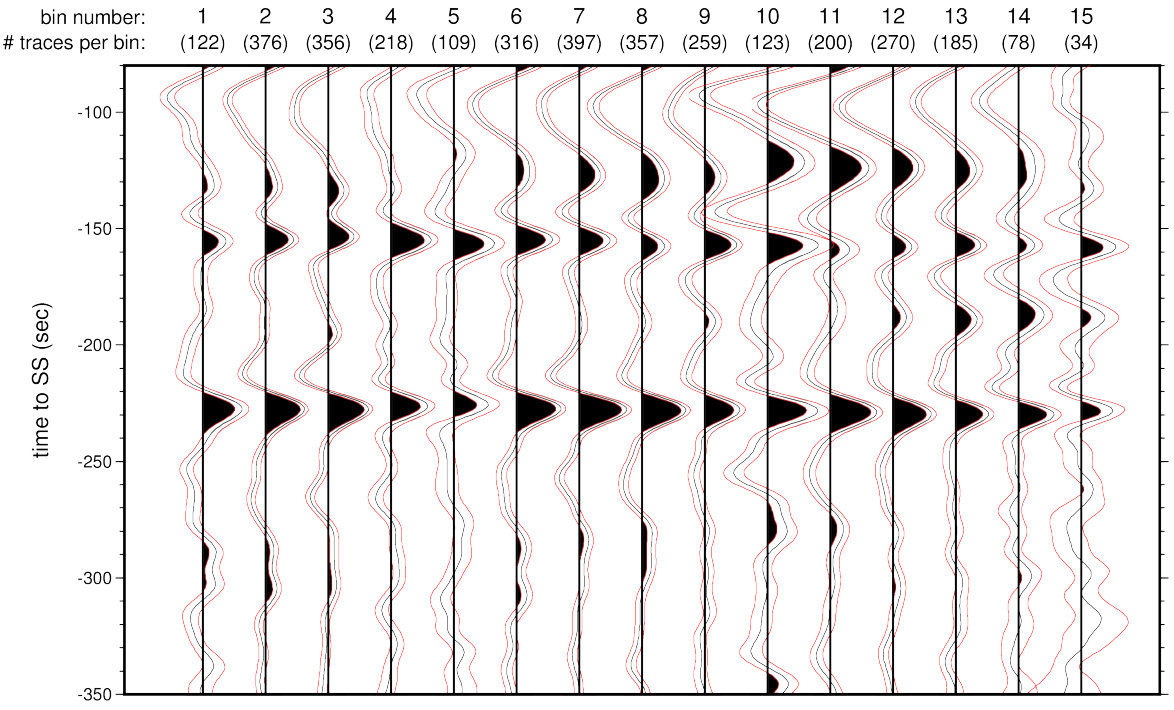

Figure 12: Stacked traces over time (black wiggles) for every bin plotted with 95% confidence interval in between the red lines. Peaks are robust if the entire confidence interval manages to surpass the baseline. Robust positive peaks are black.

3 Part 2: Reverse time migration

3.1 Method

3.1.1 Theory

The method used in this study is in essence similar to the one used in Shearer et al. (1999) mentioned in the Introduction, with the exception that here the data coverage is more dense. In this section the theory underlying the method of reverse time migration is outlined. Let us start from the beginning with the inhomogeneous scalar wave equation in the frequency domain

$$\left(\nabla^2 + \frac{\omega^2}{c(\mathbf{x})^2}\right) u(\mathbf{x}, \omega) = s(\omega) \tag{3}$$

where ∇^2 is the second order spatial differential (Laplace) operator, ω is frequency, c the wave velocity, u the wave field, and s the source function. The left hand side can be rewritten with the linear PDE operator L (equation 4), which substitutes $\nabla^2 + \omega^2/c^2$.

$$Lu = s \tag{4}$$

L depends on the medium through which the wave travels, since it depends on the wave velocity of the medium.

We assume a medium with some point scatterers. The total wave field u has linear operator $L = \nabla^2 + \omega^2/c^2$ with velocity $c = c_0 + \delta c$, whereas the background wave field u_0 (without scatterers) has linear operator $L_0 = \nabla^2 + \omega^2/c_0^2$ with velocity c_0. Therefore the wave equation for the background wave field is $L_0 u_0 = s$. Since the total wave field is a sum of the background and scattered wave fields, $u = u_0 + u_s$, the scattered wave field u_s equals the total wave field minus the background wave field: $u_s = u - u_0$. Combining this expression for u with $Lu - L_0 u_0 = 0$ (since $s - s = 0$) results in:

$$L(u_s + u_0) - L_0 u_0 = 0 \tag{5}$$

and rewritten:

$$(L - L_0)u_0 + Lu_s = 0 \tag{6}$$

Defining the scattering potential ν as, similar to Miller et al. (1987):

$$\nu = L - L_0 = \omega^2 \left[\frac{1}{c_0^2} - \frac{1}{c^2}\right] \tag{7}$$

equation 6 becomes

$$Lu_s = -\nu u_0 \tag{8}$$

Equation 8 states that the difference δc between the total velocity and background velocity creates a scattering potential ν (Miller et al., 1987) that results in a scattered wave field u_s. Velocity perturbations relative to the background model are what we are after. Therefore, the main question is: How can equation 8 be solved for δc?

We will need Green's functions to solve this problem. Green's functions are defined as the solutions to differential equations for a Dirac delta function as right hand side. If the source $s(\mathbf{x}_s, t) = \delta(\mathbf{x} - \mathbf{x}_s)\delta(t - t_0)$ is a delta pulse in space and time at $t_0 = 0$, it can be written in the frequency domain as $s(\mathbf{x}_s, \omega) = \delta(\mathbf{x} - \mathbf{x}_s)$, since the Fourier transform of $\delta(t) = 1$ for fixed ω. The wave equation then becomes $LG = \delta(\mathbf{x} - \mathbf{x}_s)$. The wave field u has been replaced by the Green's function G, which is the wave field due to a delta pulse as source. For complex media, G does not have a simple analytical solution. It is usually computed numerically. G depends on

the linear operator L, which in turn depends on the medium. A useful characteristic of Green's functions is that they are additive. They can be summed to get the total wave field.

We do not have a delta pulse as source, instead we have a complicated unknown source term. If we suppose that the source term is known, the total wave field can be obtained by convolving the Green's functions with the source term in the time domain. The convolution theorem states that for two signals $x(t)$ and $y(t)$, the product of their Fourier transforms $X(\omega)$ and $Y(\omega)$ is equal to the Fourier transform of the convolution of the two signals.

$$\mathcal{F}\left(x(t) * y(t)\right) = \mathcal{F}\left(x(t)\right)\mathcal{F}\left(y(t)\right) = X(\omega)Y(\omega) \tag{9}$$

So equivalent to convolving in the time domain is multiplying the Green's functions with the source term in the frequency domain, and summing over the volume (equation 10; Tromp et al., 2005).

$$u(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{s}, \omega) = \int_V G(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega)\, s(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega)\, \mathrm{d}^3\mathbf{x} \tag{10}$$

Equation 10 reflects the general form

$$u = \int_{\mathbb{D}} [kernel]\, [source(s)]\, \mathrm{d}^3\mathbf{x} \tag{11}$$

for the case where $Lu = s$.

As mentioned before, in reality the earthquake source-time function is not a delta pulse. However, we will assume that it is a delta pulse from now on. Following the same logic as before for a generic source term, equation 8 for a delta pulse as source becomes

$$LG_s = -\nu G_0 \tag{12}$$

The source term in equation 11 is in this case $-\nu G_0$ and G_s replaces u. The general form (equation 11) of equation 12 is therefore (see equation 7 in Miller et al. (1987))

$$G_s(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{s}, \omega) = -\int_V G(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega)\, \nu(\mathbf{x})\, G_0(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega)\, \mathrm{d}^3\mathbf{x} \tag{13}$$

for $\mathbf{x} \in V$ and $V \in \mathbb{R}^3$. The scattered wave field is induced by points where the medium differs from the background medium, reflected in the interaction between the scattering potential ν and the background wave field G_0 (Miller et al., 1987).

Born approximation

Equation 13 has two nonlinearity issues. First of all, scattered wave field G_s appears in the integrand, since $G = G_0 + G_s$ (Miller et al., 1987). The scattered wave field G_s is nonlinearly dependent on δc, because scatterers act as sources for the wave field of other scatterers (Figure 13a). To solve this first complication, equation 13 is changed by dropping the nonlinear wave field G_s from $G = G_0 + G_s$:

$$G_s(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{s}, \omega) = -\int_V G_0(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega)\, \nu(\mathbf{x})\, G_0(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega)\, \mathrm{d}^3\mathbf{x} \tag{14}$$

Now you only have the wave field from the source to point $\mathbf{x}$, $G_0(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega)$, and from point $\mathbf{x}$ to the receiver, $G_0(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega)$, to take into account. By dropping G_s, the scattered wave fields of individual scatterers do not serve as sources for other scatterers (Figure 13b).

Secondly, $\nu(\mathbf{x})$ is nonlinearly dependent on δc, since $\nu = \omega^2\left[\frac{1}{c_0^2} - \frac{1}{c^2}\right] = \omega^2\left[\frac{1}{c_0^2} - \frac{1}{(c_0+\delta c)^2}\right]$.

To solve the second problem, we rewrite the expression for the scattering potential ν to an expression linear on δc:

$$\nu(\mathbf{x}) = \omega^2 \left[\frac{1}{c_0^2} - \frac{1}{c^2}\right] = \frac{\omega^2}{c_0^2}\left(1 - \frac{1}{\left(1+\frac{\delta c}{c_0}\right)^2}\right) \tag{15}$$

with the last term rewritten as

$$1 - \frac{1}{\left(1+\frac{\delta c}{c_0}\right)^2} = 1 - \left(1+\frac{\delta c}{c_0}\right)^{-2} \tag{16}$$

The second term in expression 16 can be approximated using the first two terms of the binomial series

$$(1+x)^p = \sum_{n=0}^{\infty} \binom{p}{n} x^n = 1 + px + \frac{p(p-1)}{2!}x^2 + \cdots \tag{17}$$

which converges for $|x| < 1$, with $p = -2$ and $x = \delta c/c_0$, to

$$\left(1+\frac{\delta c}{c_0}\right)^{-2} \approx 1 - 2\frac{\delta c}{c_0} \tag{18}$$

Substituting this expression into equation 15 gives

$$\frac{\omega^2}{c_0^2}\left(1 - \frac{1}{\left(1+\frac{\delta c}{c_0}\right)^2}\right) \approx \frac{\omega^2}{c_0^2}\left(2\frac{\delta c}{c_0}\right) = 2\omega^2\frac{\delta c}{c_0^3} \tag{19}$$

$$\nu(\mathbf{x}) \approx \nu^B(\mathbf{x}) = 2\omega^2\frac{\delta c}{c_0^3} \tag{20}$$

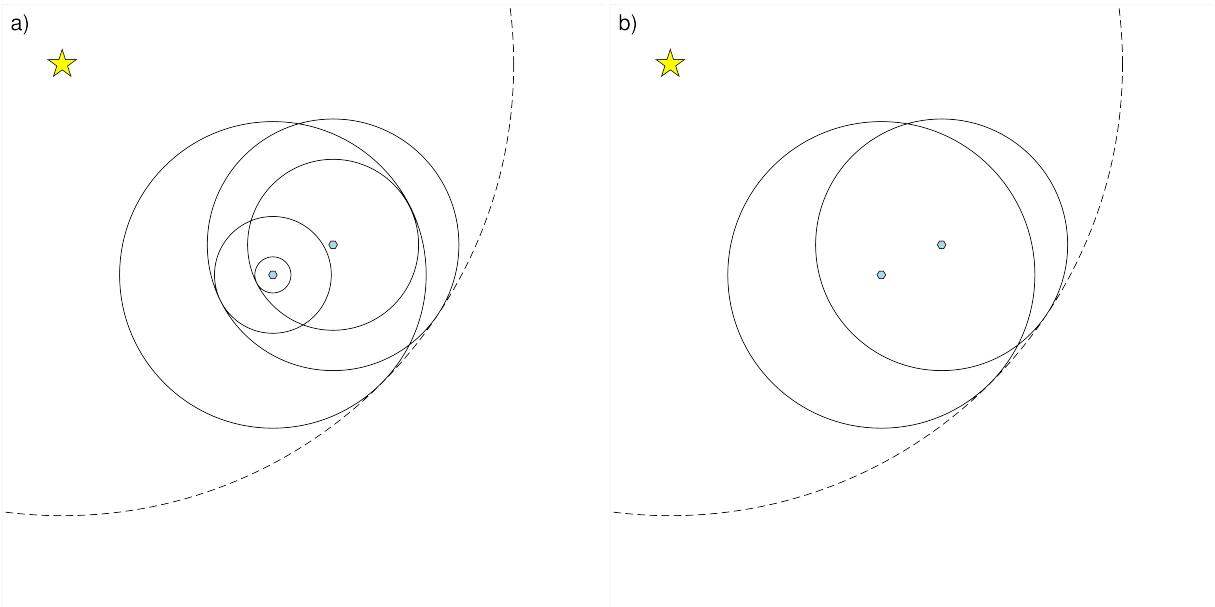

Figure 13: a) Wavefront due to a source (star) is scattered multiple times by the point scatterers (blue hexagons); b) Single scattering by the point scatterers (Born approximation). Dashed wavefront represents the background wave field. Solid circles represent the scattered wave field.

Including the third term of the binomial series (equation 17) results in a second term in the expression for the scattering potential that is only one order of magnitude smaller than the first term. Truncating the binomial series after two terms is therefore a crude approximation. These two solutions to the nonlinearity problems lead to the Born approximation and change equation 13 to:

$$G_s^B(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{s}, \omega) = -\int_V G_0(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega)\, \nu^B(\mathbf{x})\, G_0(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega)\, \mathrm{d}^3\mathbf{x} \tag{21}$$

Least squares inverse

We change the notation of the following variables in equation 21 for clarity: $G_0(\mathbf{x}_\mathrm{r}, \mathbf{x}, \omega) \Rightarrow G_0^{\mathbf{x}_\mathrm{r}}$ for the receiver wave field, $G_0(\mathbf{x}, \mathbf{x}_\mathrm{s}, \omega) \Rightarrow G_0^{\mathbf{x}_\mathrm{s}}$ for the source wave field, $\nu^B(\mathbf{x}) \Rightarrow \nu^B$ for Born scattering potential, and scattered wave field $G_s^B(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{s}, \omega) \Rightarrow d_s$ (where d stands for data). When substituting the expression for ν^B and adopting the new notation, equation 21 becomes (Dai et al., 2010):

$$d_s = -2\omega^2 \int_V \frac{1}{c_0^3}\, G_0^{\mathbf{x}_\mathrm{s}}\, G_0^{\mathbf{x}_\mathrm{r}}\, \delta c(\mathbf{x})\, \mathrm{d}^3\mathbf{x} \tag{22}$$

Equation 22 can be recognized as a matrix-vector multiplication, since $\mathbf{a} = \mathbf{M}\mathbf{b}$ is equivalent to:

$$a(x') = \int \mu(x', x'') b(x'')\, \mathrm{d}x'' \tag{23}$$

where $a(x')$ are the elements of column vector $\mathbf{a}$, $\mu(x', x'')$ the elements of matrix $\mathbf{M}$, and $b(x'')$ the elements of column vector $\mathbf{b}$. Equation 22 rewritten as matrix-vector multiplication with $\mathbf{L}$ containing $\frac{-2\omega^2}{c_0^3}\, G_0^{\mathbf{x}_\mathrm{s}}\, G_0^{\mathbf{x}_\mathrm{r}}$ and $\delta\mathbf{m}$ containing δc for every grid point is:

$$\mathbf{d_s} = \mathbf{L}\, \delta\mathbf{m} \tag{24}$$

The space domain is discretized in N points. For a unique source-receiver pair and single frequency, the scattered wave field d_s is given by:

$$d_s = \mathbf{L}\, \delta\mathbf{m} = \begin{pmatrix} \frac{-2\omega^2}{c_0^3(\mathbf{x}_{(0)})} G_0^{\mathbf{x}_\mathrm{s}}(\mathbf{x}_{(0)}) G_0^{\mathbf{x}_\mathrm{r}}(\mathbf{x}_{(0)}) & . & . & . & \frac{-2\omega^2}{c_0^3(\mathbf{x}_{(N)})} G_0^{\mathbf{x}_\mathrm{s}}(\mathbf{x}_{(N)}) G_0^{\mathbf{x}_\mathrm{r}}(\mathbf{x}_{(N)}) \end{pmatrix} \begin{pmatrix} \delta c(\mathbf{x}_{(0)}) \\ . \\ . \\ . \\ \delta c(\mathbf{x}_{(N)}) \end{pmatrix} \tag{25}$$

The vector of unknowns is $\delta\mathbf{m}$. To solve for the model $\delta\mathbf{m}$ that contains the velocity perturbations δc, the forward problem (equation 24) needs to be inverted. We first define a residual vector $\mathbf{e} = \mathbf{d_s} - \mathbf{L}\, \delta\mathbf{m}$ and then apply least squares inversion to find $\delta\mathbf{m}$ that minimizes the norm of the residual squared, $\|\mathbf{e}\|^2$. We start by writing the normal equations. Since $\mathbf{L}$ is not a square matrix and complex, we need the adjoint of $\mathbf{L}$, denoted by $\mathbf{L}^\dagger$, instead of the transpose. The adjoint of a matrix is the conjugate transpose, i.e. $(\mathbf{L}^\mathrm{T})^*$. An element of $\mathbf{L}$ is of the form $\frac{-2\omega^2}{c_0^3}\, G_0^{\mathbf{x}_\mathrm{s}}\, G_0^{\mathbf{x}_\mathrm{r}}$, whereas an element of $\mathbf{L}^\dagger$ is of the form $\frac{-2\omega^2}{c_0^3}\, (G_0^{\mathbf{x}_\mathrm{s}})^*\, (G_0^{\mathbf{x}_\mathrm{r}})^*$ with complex conjugates of the Green's functions. Multiplying equation 24 by the adjoint of $\mathbf{L}$ on both sides yields the normal equations (equation 26).

$$\mathbf{L}^\dagger\, \mathbf{d_s} = \mathbf{L}^\dagger\, \mathbf{L}\, \delta\mathbf{m} \tag{26}$$

It is followed by a multiplication by $(\mathbf{L}^\dagger\mathbf{L})^{-1}$ on both sides:

$$(\mathbf{L}^\dagger\mathbf{L})^{-1}\, \mathbf{L}^\dagger\, \mathbf{d_s} = (\mathbf{L}^\dagger\mathbf{L})^{-1}\, \mathbf{L}^\dagger\, \mathbf{L}\, \delta\mathbf{m} \tag{27}$$

Recognizing the identity matrix in $(\mathbf{L}^\dagger\mathbf{L})^{-1}\,\mathbf{L}^\dagger\,\mathbf{L}$ on the right hand side of equation 27, the least squares solution to the normal equations (equation 26) can be written as:

$$\delta\mathbf{m}^{inv} = (\mathbf{L}^\dagger\mathbf{L})^{-1}\,\mathbf{L}^\dagger\,\mathbf{d_s} \tag{28}$$

Equation 28 minimizes the norm of the residual squared in a least squares sense. The quantity $\mathbf{L}^\dagger\mathbf{L}$ is the integral blurring operator and is represented by a resolution function (Hu et al., 2001; Bai et al., 2012). If $(\mathbf{L}^\dagger\mathbf{L})^{-1}$, or $\mathbf{L}^\dagger\mathbf{L}$, is close to the identity matrix, $\delta\mathbf{m}^{inv}$ will be well resolved.

Using the proper inverse $(\mathbf{L}^\dagger\mathbf{L})^{-1}$ has some drawbacks. First of all, this quantity is highly sensitive to perturbations in the data (Bai et al., 2012). Secondly, the inverse of a matrix exists if the matrix is square, and if it has full rank, which means that the rows or columns are linearly independent. The former criterion is satisfied, since $\mathbf{L}^\dagger\mathbf{L}$ is a square matrix. The latter criterion, however, may not be satisfied or the inverse may be too computationally costly to compute (Bai et al., 2012). The simplest course of action is to approximate $(\mathbf{L}^\dagger\mathbf{L})^{-1} \approx \mathbf{I}$, i.e. to drop the resolution indicator term. In doing that, we finally arrive at the expression that can be solved for δc in $\delta\mathbf{m}^{mig}$:

$$\delta\mathbf{m}^{mig} = \mathbf{L}^\dagger\,\mathbf{d_s} \tag{29}$$

Deblurring will be applied to the migrated image at the end with a less costly method than using the inverse of the integral blurring operator. The adjoint operator $\mathbf{L}^\dagger$ is called reverse time migration operator (Dai et al., 2012) or conventional migration operator. The adjoint is safer to use than the formal inverse, because it amplifies noise to a lesser extent, tolerates imperfections in the data and does not demand full and complete information from the data (Claerbout & Fomel, 2008). Equation 29 is an important result that we will use in reverse time migration.

Summing over sources, receivers and frequencies

Next we will see how exactly equation 29 can be solved for $\delta\mathbf{m}$. For one receiver, one frequency and one source, the model value δm^{mig} at point $\mathbf{x}_\mathrm{p}$ is given by (returning to the old notation)

$$\delta m^{mig}_{s,r,\omega}(\mathbf{x}_\mathrm{p}) = \frac{-2\omega^2}{c_0^3(\mathbf{x}_\mathrm{p})}\,G_0^*(\mathbf{x}_\mathrm{p},\mathbf{x}_\mathrm{s},\omega)\,G_0^*(\mathbf{x}_\mathrm{r},\mathbf{x}_\mathrm{p},\omega)\,d_s(\mathbf{x}_\mathrm{r},\mathbf{x}_\mathrm{s},\omega) \tag{30}$$

which can be rewritten using a property of complex conjugates $\mathbf{a}^*\mathbf{b} = (\mathbf{a}\mathbf{b}^*)^*$ to

$$\delta m^{mig}_{s,r,\omega}(\mathbf{x}_\mathrm{p}) = \frac{-2\omega^2}{c_0^3(\mathbf{x}_\mathrm{p})}\,G_0^*(\mathbf{x}_\mathrm{p},\mathbf{x}_\mathrm{s},\omega)\,\left[G_0(\mathbf{x}_\mathrm{r},\mathbf{x}_\mathrm{p},\omega)\,d_s^*(\mathbf{x}_\mathrm{r},\mathbf{x}_\mathrm{s},\omega)\right]^* \tag{31}$$

In this study we have multiple sources, receivers and frequencies to sum over. For N_s sources, N_r receivers and N_ω frequency samples, the summation for model point $\mathbf{x}_\mathrm{p}$ is

$$\delta m^{mig}(\mathbf{x}_\mathrm{p}) = \sum_{i=1}^{N_s}\sum_{j=1}^{N_r}\sum_{k=1}^{N_\omega}\frac{-2(\omega^k)^2}{c_0^3(\mathbf{x}_\mathrm{p})}G_0^*(\mathbf{x}_\mathrm{p},\mathbf{x}_\mathrm{s}{}^i,\omega^k)\,\left[G_0(\mathbf{x}_\mathrm{r}{}^j,\mathbf{x}_\mathrm{p},\omega^k)\,d_s^*(\mathbf{x}_\mathrm{r}{}^j,\mathbf{x}_\mathrm{s}{}^i,\omega^k)\right]^* \tag{32}$$

The data vector $\mathbf{d_s}$ would be very large (with a size of: $N_s \cdot N_r \cdot N_\omega$) if it contained data for every source-receiver-frequency combination. Luckily, the summation over receivers is implicit in finite difference modelling. Data at the receivers belonging to one source can be injected all at once and their wave fields are added. The summation over frequencies can be simplified by evaluating the cross-correlation between source and receiver wave fields at $\tau = 0$, i.e. zero lag (Claerbout, 1971). If $f(\tau)$ is the cross-correlation function, depending on lag time τ, its inverse

Fourier transform (mapping frequency domain to the time domain) can be written in integral notation,

$$f(\tau) = \int F(\omega)\, e^{i\omega\tau} \mathrm{d}\omega \tag{33}$$

or as Riemann summation.

$$f(\tau) \approx \sum_{k=1}^{N_\omega} F(\omega_k) e^{i\omega_k \tau} \Delta\omega \tag{34}$$

By evaluating the Riemann summation at $\tau = 0$, the exponential term becomes equal to one:

$$f(\tau = 0) \approx \sum_{k=1}^{N_\omega} F(\omega_k) \Delta\omega \tag{35}$$

Neglecting scalar $\Delta\omega$ in equation 35 results in the statement: Evaluating the cross-correlation at $\tau = 0$ is equivalent to summing over all frequency samples. The actual computation is performed in the time domain, instead of the frequency domain in which all equations have been derived thus far. The imaging condition for each shot-gather is obtained by cross-correlating in time at every image location (x,z) between the source and receiver wave fields, followed by image extraction at $\tau = 0$. Summation over sources results in a final migrated image.

The idea behind all these equations is imaging reflectors in the mantle (δc), using the data and a background velocity model. As Claerbout (1971, p. 467) put it: "reflectors exist at points in the ground where the first arrival of the downgoing wave is time coincident with an upgoing wave." From the source position, propagate a "delta" pulse (a delta pulse can only be approximated numerically) in the background medium. From the receiver position, propagate the time reverse of the seismic data in the background medium. Reflectors are imaged with correlation techniques and are positioned where these two wave fields interfere constructively (Claerbout, 1971). Scatterers are thus not included in the velocity model used for FD modelling of wave field propagation.

In summary, the steps to be taken are:

1. Time-reverse the data (d_s^* in the frequency domain equals time reversal in the time domain);
2. Propagate time-reversed data from $\mathbf{x}_\mathrm{r} \Rightarrow \mathbf{x}_\mathrm{p}$ in background medium ($G_0(\mathbf{x}_\mathrm{r}, \mathbf{x}_\mathrm{p}, \omega) d_s^*$) resulting in a receiver wave field;
3. Propagate a "delta" source from $\mathbf{x}_\mathrm{s} \Rightarrow \mathbf{x}_\mathrm{p}$ in background medium, resulting in a source wave field $G_0(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{s}, \omega)$);
4. Cross-correlate $G_0(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{s}, \omega)$ with the resulting wave field of step 2 and extract the image at $\tau = 0$ (zero lag);
5. Sum over all data (sources).

3.1.2 Finite difference model set-up

Finite difference (FD) modelling of source and receiver wave fields and extraction of the imaging condition are executed in Madagascar software. Because of subsurface imaging purposes of the software, the Earth is assumed to be spherical (in Cartesian coordinates) and acoustic, even though it is better approximated by an ellipsoid and predominantly elastic. Reverse time migration is performed on a 2-D section through the Earth. The intersection between the section plane and the Earth's surface is a great circle that runs approximately through the centres of source locations, of receivers in North-America and of bounce points in the Pacific Ocean (Figure 14). Sources and receivers have to be added to the model. Projecting source and

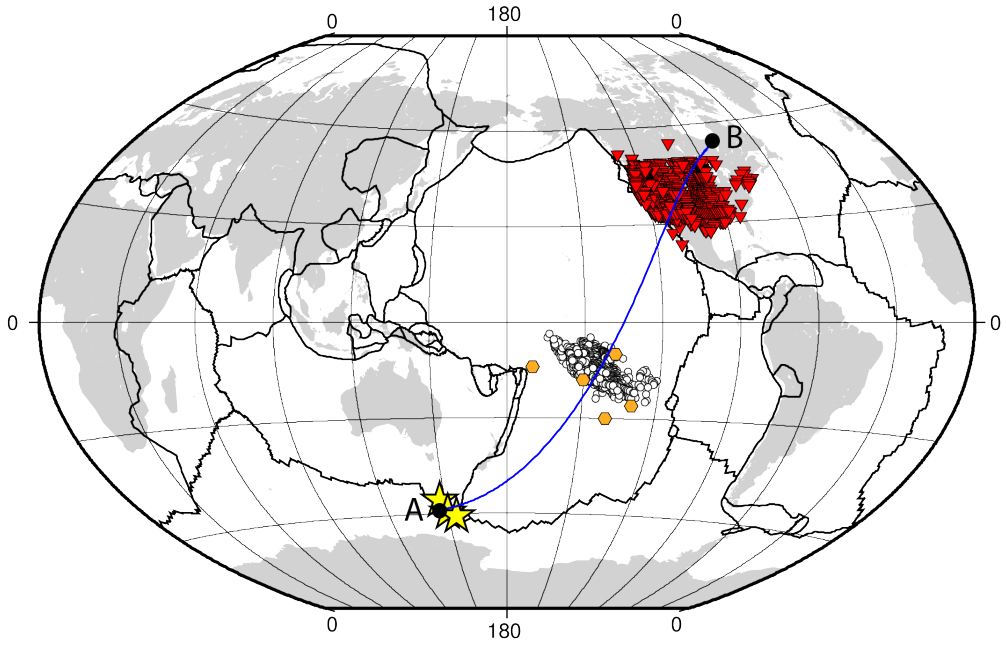

Figure 14: Blue great circle arc from point A to point B represents the intersection between the section plane and Earth's surface. All symbols have the same meaning as before.

receiver locations onto the section line along the shortest distance results in a shorter epicentral distance than the true distance. Reflecting surfaces in the imaging condition are then more shallow than they should be. Therefore, locations of sources and receivers are rotated onto the section line with the correct epicentral distance. The background velocity model has a grid size of 30 by 30 km and is based on the 3-D S-wave velocity model S40RTS (Ritsema et al., 2011). Velocities in S40RTS are velocity difference percentages with respect to PREM velocities. Equation 36 is used to obtain absolute S-wave velocites at every grid point $\mathbf{x}_\mathrm{p}$, in which the value for the fraction term is taken from S40RTS.

$$v_s(\mathbf{x}_\mathrm{p}) = v_{s,prem}(\mathbf{x}_\mathrm{p}) + v_{s,prem}(\mathbf{x}_\mathrm{p}) \cdot \frac{v_s(\mathbf{x}_\mathrm{p}) - v_{s,prem}(\mathbf{x}_\mathrm{p})}{v_{s,prem}(\mathbf{x}_\mathrm{p})} \tag{36}$$

Three background models were made:

1. Realistic model: Mantle and crust with realistic S-wave velocities computed with equation 36. Very small velocities (0.001 km/s) in the core and atmosphere (not zero, because that is computationally unstable).
2. Smoothed model: Fifteen-point triangular smoothing filter applied twice to the realistic model, causing lateral velocity heterogeneities to smear out.
3. Masked model: All mantle velocities higher than 7.3 km/s (close to the core-mantle boundary (CMB)) are set to 7.3 km/s, including the core. All mantle velocities lower than 4.4 km/s (close to the surface) are set to 4.4 km/s, including the atmosphere.

None of the models have a free surface at the top, since the top of the model intersects the centre of the Earth. Amplitudes at all model boundaries decay exponentially, so there will be no reflections off the boundaries.

The realistic model (Figure 15) has undesired reflections and multiples off the CMB and surface, because of the large sharp velocity contrasts. Those unwanted reflections will eventually enter the imaging condition and distract from the reflectors you want to image. The background

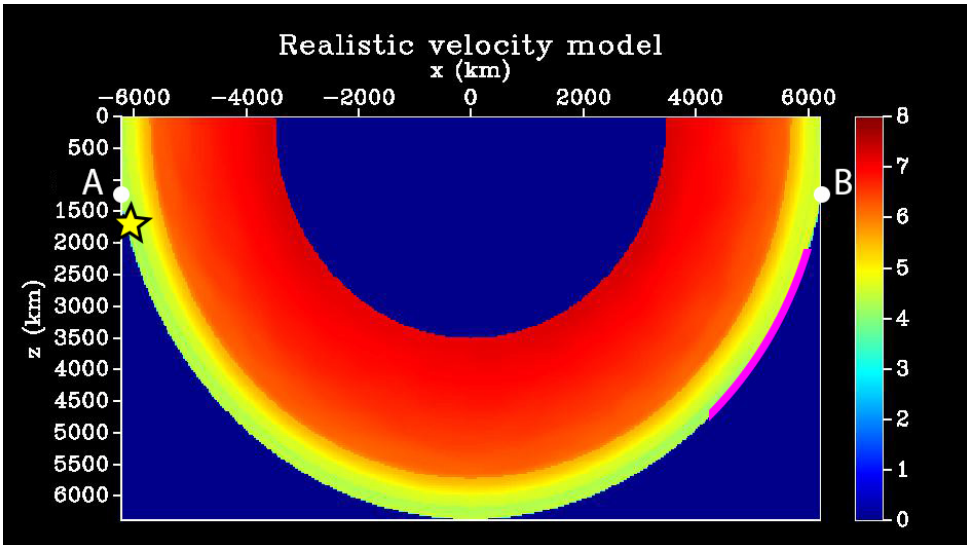

Figure 15: Realistic velocity model. Velocities are in km/s. A and B represent the ends of the section line in Figure 14. The star indicates source locations and the pink line the receiver locations.

model through which the source and receiver wave fields are propagated should not have strong reflectors. Waves in the smoothed model are partly absorbed and reemitted by the CMB and surface, which is not ideal either for the same reason. The masked model is the best background model, since it does not have sharp velocity contrasts at the CMB and surface. It does still have the transition zone discontinuities in it, which are sharp velocity contrasts as well, albeit of a lower magnitude. Those reflections carry less energy than the direct wave. A possible solution to this problem is local or general smoothing. The background model in a study by Cao et al. (2010) is reference Earth model ak135. They put gradual velocity transitions instead of velocity discontinuities at the depths of transition zone discontinuities. It is also required here, because the transition zone discontinuities are strong enough to cause reflections. Therefore, a three-point smoothing filter is applied to the masked model. All further imaging is done in the smoothed masked model (Figure 16) and the other two models are discarded.

The data injected at the receiver side are processed in a similar fashion to processing in the regular stacking procedure. I have generated four different data sets, in order of increasing processing intensity:

1. Raw transverse component data without any filtering, cutting or normalization.
2. Bandpass filtered (15-75 sec), cut -300 to +60 seconds around theoretical SS arrival (ttimes) data. This time window includes both the SS wave and its precursors.
3. Bandpass filtered (15-75 sec), cut -300 to -60 seconds around theoretical SS arrival (ttimes) data. This time window does not include the SS wave, but encompasses the precursor window.
4. Bandpass filtered (15-75 sec), cut -300 to -60 seconds around theoretical SS arrival (ttimes), normalized with respect to the SS magnitude and polarity reversed for 22 out of 1281 seismograms.

The last data set with amplitude normalization and polarity reversal is most like the processed data used in slant-stacking in the first part.

As source wavelet and delta pulse approximation, a Ricker wavelet is implemented. This type of source simulates an explosion, emitting P-wave energy through the acoustic medium. The peak frequency (most energetic frequency) of the Ricker wavelet corresponds to the dominant

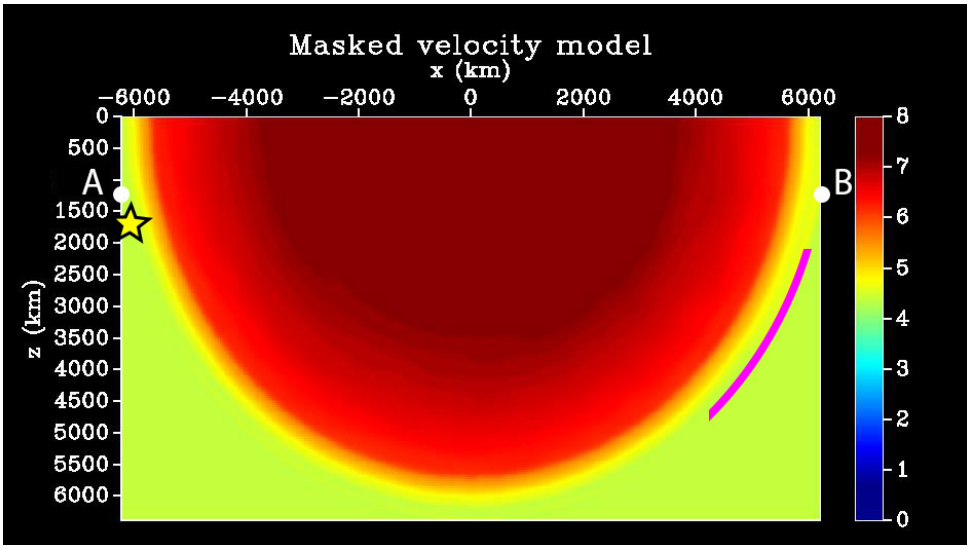

Figure 16: Masked velocity model with smoothing. Velocities are in km/s. A and B represent the ends of the section line in Figure 14. The star indicates source locations and the pink line the receiver locations.

frequency in the filtered data revealed by Fourier transformation: 0.02 Hz. To make sure that the frequency is not too low to leave an imprint on the width of reflectors in the imaging condition, several higher frequencies were tested. It turned out that 0.02 Hz is not too low. For every source, the seismic data are reversed in time and given to the receivers as input. The FD model calculates the source and receiver wave fields at every time increment $\Delta t = 1$ second. The zero-lag cross-correlation is implemented by first reversing the source wave field in time, followed by multiplication of the time-reversed source wave field and total receiver wave field at every time step. The wave field at each (x,z) coordinate is then stacked over time (Chattopadhyay & McMechan, 2008), followed by a summation over the sources.

$$image(x,z) = \sum_{sources} \sum_{time} s(x,z,t)\, r(x,z,t) \tag{37}$$

Or in the frequency domain:

$$image(x,z) = \sum_{sources} \sum_{freq} S(x,z,\omega)\, R(x,z,\omega) \tag{38}$$

The source wave field $S(x,z,\omega)$ in equation 38 corresponds to $G_0^*(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{s}{}^i, \omega^k)$ in equation 32. The receiver wave field $R(x,z,\omega)$ in equation 38 corresponds to $\sum_{j=1}^{N_r} G_0(\mathbf{x}_\mathrm{r}{}^j, \mathbf{x}_\mathrm{p}, \omega^k)\, d_s^*(\mathbf{x}_\mathrm{r}{}^j, \mathbf{x}_\mathrm{s}{}^i, \omega^k)$ in equation 32. Summation over receivers is implicit in $R(x,z,\omega)$.

3.1.3 Numerical stability

The FD scheme used here is conditionally stable, meaning that for a given grid spacing, Δt must be small enough to avoid numerical instability. When Δt used to calculate wave fields at different time steps is smaller than $\Delta t = 1$ of the data, the data must be resampled at a higher rate. Interpolation increases the input sampling rate to a higher output sampling rate. The Shannon-Nyquist sampling theorem states that a signal can be perfectly reconstructed, i.e. without aliasing, if the highest frequency in the signal is smaller than $f_s/2$, where f_s is

the sampling frequency (Shannon, 1949). The highest frequencies in the unfiltered and filtered signals are 0.5 Hz and 0.067 Hz, respectively, which are both smaller than a Nyquist frequency of 5 Hz, corresponding to a sampling frequency of 10 Hz ($\Delta t = 0.1$). For a resolution of 30x30 km, resampling the data is not required. For 10x10 km, however, resampling is required.

Another factor to take into consideration is an artefact called numerical dispersion. It is caused by interpolation errors used in approximating derivatives of the true wave field. Space and time discretization affect dispersion. Too few samples per wavelength in space and period in time lead to this numerical artifact. A source wavelet frequency of 0.02 Hz did not lead to numerical dispersion, whereas increasing the frequency to 0.04 Hz did, which is another argument in favour of using 0.02 Hz. An increase in the frequency of the source wavelet leads to less samples per wavelength and therefore to numerical dispersion. The solution therefore lies in choosing a sufficiently high resolution grid and small time step. A rule of thumb is to get at least 10 samples per wavelength in space and period in time. Fourier transformation of the data to the frequency domain showed that the dominant frequency in bandpass filtered data (15-75 sec) is 0.02 Hz, corresponding to a period of 50 seconds. So the time step should be 5 seconds or less. A time step of 1 second satisfies this condition. For an average S-wave velocity of 6 km/s, the wavelength would be 300 km. The maximum grid spacing would be 30 km. This Δx might be too large for the supposedly km-scale variations in transition zone topography that we want to image. However, decreasing the grid size would probably not lead to the desired increase in topography resolution, because of the long-period nature of the data. The highest frequency in the filtered data is approximately 0.067 Hz (period of 15 seconds) and requires a maximum grid spacing of 9 km. So for the maximum frequency there would not be 10 samples per wavelength in space in the current 30x30 km grid cells.

For numerical stability, the Courant criterion needs to be satisfied. In 1-D the Courant criterion is defined as:

$$C = \frac{u\Delta t}{\Delta x} \leq C_{max} \tag{39}$$

It means that for a given model velocity u, the time step Δt must not be too large and grid spacing Δx not be too small. Furthermore, if the velocity model contains large velocity contrasts between a cell and neighbouring cells, the wave field will become unstable once it reaches this velocity contrast. Problems with the Courant criterion and velocity contrast were encountered during synthetic modelling experiments. Point scatterers with a large velocity contrast between the scatterer and the surrounding cells were not stable. Once the source wave field arrived at such a scatterer, the wave field became unstable. Smoothing of the point scatterers smeared out the velocity amplitude over a larger area and ensured a lower velocity gradient between cells.

3.1.4 Improving the imaging condition

Several techniques can be applied to improve the resulting imaging conditions. Differences in phase amplitudes could come from coupling between receiver and the ground (although the signals have been deconvolved for impulse response), local impedance contrasts (hard rock vs soft rock), take-off angle and geometrical spreading. The data have 3-D geometrical spreading, while the FD modelling is performed in 2-D. A brute correction for taking geometrical spreading from 3-D to 2-D is done by scaling every time sample amplitude of receiver data by a factor $\sqrt{\text{time}}$. Travel time is proportional to R in a homogeneous medium, with R being radial distance to the source. Geometrical spreading is proportional to $1/R$ in 3-D and $1/\sqrt{R}$ in 2-D. Therefore, multiplying by $\sqrt{\text{time}} \approx \sqrt{CR}$ approximates 2-D geometrical spreading.

To account for the total energy per event, the sum over sources is weighted by the total power of the signal at 19 shared receivers between all three events. For each of the four data sets, these weights are calculated separately. Visual inspection of the weighted data of all three

events in one image without clipping served as a sanity check.

Sharpening the imaging condition can be done with the triangle smoothing filter in Madagascar by subtracting an extremely smoothed version of the image from a gently smoothed version. The Gaussian smoothing filter is constructed by convolving two identical n-point triangular wavelets. The number of points in this filter determines the degree of smoothing. More points results in more smoothing, because the amplitude of one image point is spread out over a larger area.

Two additional operations that will improve the imaging condition are source illlumination and deconvolution normalization (Claerbout, 1971). Dividing the cross-correlated imaging condition by source illumination improves the accuracy of reflectivity information in the cross-correlation image. It is a way of dealing with geometrical spreading in the model. Parts of the cross-correlated image close to the source are highlighted stronger, simply because they are closer to the source and therefore have experienced less spreading of energy. Other parts that are highlighted more strongly are conduits where energy "escapes" from the model due to high local wavefront curvature. Source illumination is equal to autocorrelation at zero lag of the source wave field. It is computed in the frequency domain, because autocorrelation in the time domain did not yield a plausible result. The source wave field is firstly transferred to the frequency domain by taking the Fourier transform: $s(x,z,t) \rightarrow S(x,z,\omega)$. Then the source wave field is multiplied by its complex conjugate, yielding the power of the source wave field in the frequency domain: $S(x,z,\omega)S^*(x,z,\omega) = |S(x,z,\omega)|^2$. This operation is equivalent to taking the Fourier transform of the autocorrelation in the time domain, according to the convolution theorem. Summing over all frequencies is the same as evaluating the autocorrelation at zero lag ($\tau = 0$).

$$image(x,z) = \sum_{sources} \frac{\sum\limits_{freq} R(x,z,\omega)S^*(x,z,\omega)}{\left\langle \sum\limits_{freq} |S(x,z,\omega)|^2 \right\rangle_{smooth}} \tag{40}$$

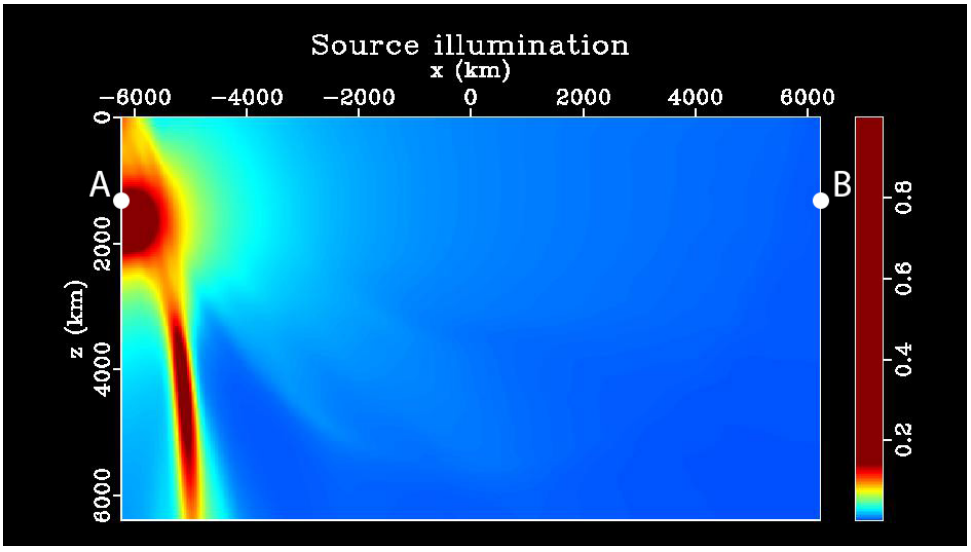

Figure 17: Smoothed and normalized source illumination map for masked velocity model, source wavelet peak frequency 0.02 Hz and source location of the 2008-06-01 earthquake. Clipping percentage is 99.

If the amplitudes in the cross-correlated image would have had any physical meaning, dividing it by the source illumination strength would have yielded an image with the same unit

(dimensionless), sign and scaling as the reflection coefficient (Chattopadhyay & McMechan, 2008). However, the absolute amplitudes in the cross-correlated image do not have any physical meaning, because the Earth's interior is not acoustic. Therefore, the source illumination map can safely be normalized by the largest value, before it is applied to the cross-correlation images. The irregularity of the source illumination map shows clearly that the velocity model is heterogeneous (Figure 17). A homogeneous velocity model would have yielded a circular symmetric source illumination map. The high-power conduit going down from the source (Figure 17) is a high curved wavefront that travels outside of the Earth.

A sharper image can be obtained by deconvolution imaging. It is linked to the inverse of the blurring operator, $(\mathbf{L}^\dagger\mathbf{L})^{-1}$, that was dropped in the derivation of an expression for reverse time migration because of its complicated and costly computation. Each diagonal element of the matrix $(\mathbf{L}^\dagger\mathbf{L})$ is a summation over sources, receivers and frequencies for one point $\mathbf{x}_\mathrm{p}$ in the model:

$$\sum_{i=1}^{N_s}\sum_{j=1}^{N_r}\sum_{k=1}^{N_\omega}(\omega^k)^4\, G_0(\mathbf{x}_\mathrm{p},{\mathbf{x}_\mathrm{s}}^i,\omega^k)\, G_0({\mathbf{x}_\mathrm{r}}^j,\mathbf{x}_\mathrm{p},\omega^k)\, G_0^*(\mathbf{x}_\mathrm{p},{\mathbf{x}_\mathrm{s}}^i,\omega^k)\, G_0^*({\mathbf{x}_\mathrm{r}}^j,\mathbf{x}_\mathrm{p},\omega^k) \tag{41}$$

Recognizing that $G_0^*(\mathbf{x}_\mathrm{p},\mathbf{x}_\mathrm{s},\omega)$ is the source wave field S, it can be rewritten as:

$$\sum_{i=1}^{N_s}\sum_{j=1}^{N_r}\sum_{k=1}^{N_\omega}(\omega^k)^4\, \left|S(\mathbf{x}_\mathrm{p},{\mathbf{x}_\mathrm{s}}^i,\omega^k)\right|^2\, \left|G_0({\mathbf{x}_\mathrm{r}}^j,\mathbf{x}_\mathrm{p},\omega^k)\right|^2 \tag{42}$$

The matrix has off-diagonal elements as well (Nemeth et al., 1999), but under the assumption that $(\mathbf{L}^\dagger\mathbf{L})$ is a diagonal matrix, its inverse $(\mathbf{L}^\dagger\mathbf{L})^{-1}$ is simply the reciprocal of equation 42.

$$\sum_{i=1}^{N_s}\sum_{j=1}^{N_r}\sum_{k=1}^{N_\omega}(\omega^k)^{-4}\, \left|S(\mathbf{x}_\mathrm{p},{\mathbf{x}_\mathrm{s}}^i,\omega^k)\right|^{-2}\, \left|G_0({\mathbf{x}_\mathrm{r}}^j,\mathbf{x}_\mathrm{p},\omega^k)\right|^{-2} \tag{43}$$

Geometrical spreading is frequency dependent, which is another reason for performing wave equation migration. It can be shown that the power of the source wave field varies with frequency. So instead of obtaining one image of source wave field power by averaging over all frequencies as in source illumination, calculating the power per frequency is more accurate.

Deconvolution is the reverse operation of convolution. It is used in signal processing when the original signal $x(t)$ is modified by the instrument or recorder to $y(t)$ (equation 44). Deconvolution filters out this instrument response by finding a wavelet $f(t)$ that becomes a delta function when convolved with $a(t)$ (equation 45).

$$y(t) = x(t) * a(t) \tag{44}$$

$$y(t) * f(t) = x(t) * a(t) * f(t) = x(t) * \delta(t) \tag{45}$$

For deconvolution imaging, we have to find a filter $f(x,z,t)$ that transforms the source wave field $s(x,z,t)$ into a delta function in time (equation 46). Equivalently, in the frequency domain (equation 47), find $F(x,z,\omega)$ that gives 1 when multiplied with the source wave field.

$$f(t) * s(t) = \delta(t) \tag{46}$$

$$F(x,z,\omega)S(x,z,\omega) = 1 \tag{47}$$

So $F(x,z,\omega)$ is the reciprocal of $S(x,z,\omega)$. Substituting $F(x,z,\omega) = 1/S(x,z,\omega)$ for $S(x,z,\omega)$ in equation 38 and assuming a sum over receivers, as is implicitly implemented in the finite difference model, expression 38 changes into:

$$image(x,z) = \sum_{sources}\sum_{freq}\frac{R(x,z,\omega)}{S(x,z,\omega)} = \sum_{sources}\sum_{freq}R(x,z,\omega)\frac{S^*(x,z,\omega)}{\left|S(x,z,\omega)\right|^2} \tag{48}$$

The extra term that has appeared in equation 48 w.r.t. equation 38 is the reciprocal of the power of the source wave field. This extra term appears as well in equation 43 for diagonal elements of $(\mathbf{L}^\dagger\mathbf{L})^{-1}$, the deblurring operator. That is why deconvolution imaging creates a sharper image. It compensates for the distortions in the images due to limitations in recording time, bandwidth and aperture by using the resolution kernel (Bai et al., 2012). However, deconvolution approaches are ill-posed, making the resulting deconvolved images sensitive to noise (Bai et al., 2012).

3.2 Results: migrating real data

After executing the five steps outlined at the end of the Theory section for each of the four different data sets, the resulting imaging conditions are analyzed. Firstly, the first data set (raw data) is migrated. The most prominent signal in this data set is a high-amplitude surface wave at 3000-3500 seconds after the earthquake origin time. Surface waves are non-scattered data, so they should ideally not be in the data used for migration. As expected, the imaging condition for migrated raw data does not show any transition zone discontinuities (Figure 18a). The surface waves are projected to the prominent banana-doughnut-shaped sensitivity kernel (Figure 18a) which does not follow the Earth's surface, probably because the model is not suitable for projecting surface waves. It looks like a sensitivity kernel for a direct wave propagating in a linear fashion from sources to receivers (Tromp et al., 2005), but since the data do not contain such an arrival, it is meaningless. The only coherent physical feature is a reflector near the Earth's surface due to the constructively interfering SS phase (following the green dashed line representing the Earth's surface in Figure 18a).

Since the imaging condition for raw data is overpowered by the projection of surface waves, the second data set with filtered SS plus precursor window is migrated next. A reflector near the location of the Earth's surface (green dashed line) that partly follows the curvature of the Earth's surface is observed in the imaging condition for a migrated filtered SS plus precursor window (data set 2; Figure 18b). This reflector appears stronger than in the previous imaging condition. Black areas in the imaging condition have negative amplitudes, whereas white areas have positive amplitude. Since nearly all data have negative SS and precursor polarities (opposed to positive peaks in the stacking approach), black lines are interpreted the reflectors. The surface reflection follows the green dashed line for only a specific part of the section. Where it deviates from the theoretical surface of the Earth due the limited extent of the receiver array, other reflectors will probably deviate from the true reflectors as well. Therefore, only parts of the imaging conditions in the black boxes in Figure 18 should be interpreted.

The SS surface reflection dominates the imaging condition, rendering the transition zone discontinuities invisible, so in data set 3 the data are cut around the precursor window only, removing the SS arrival. The imaging condition for a migrated filtered precursor window (data set 3) shows three reflectors near transition zone depths (Figure 18c). The surface reflector that appeared in the previous two imaging conditions (Figures 18a and 18b) has indeed disappeared by cutting out the SS arrival. Two out of three reflectors align with the depths of the 410- and 660-discontinuities (yellow and red dashed lines in Figure 18c, respectively). The third reflector is situated between 410 km depth and the surface at ca. 200 km depth. Normalizing w.r.t. SS and reversing polarity if needed (data set 4) does not yield a visually different image (Figure 19a), probably because only 22 out of 1281 seismograms have had their polarity reversed and amplitudes of the SS phase vary at most one order of magnitude and are thus very similar (Figure A.1).

Source illumination and deconvolution normalization, as explained in the previous section, can be applied to the imaging condition. Deconvolution imaging has not been implemented

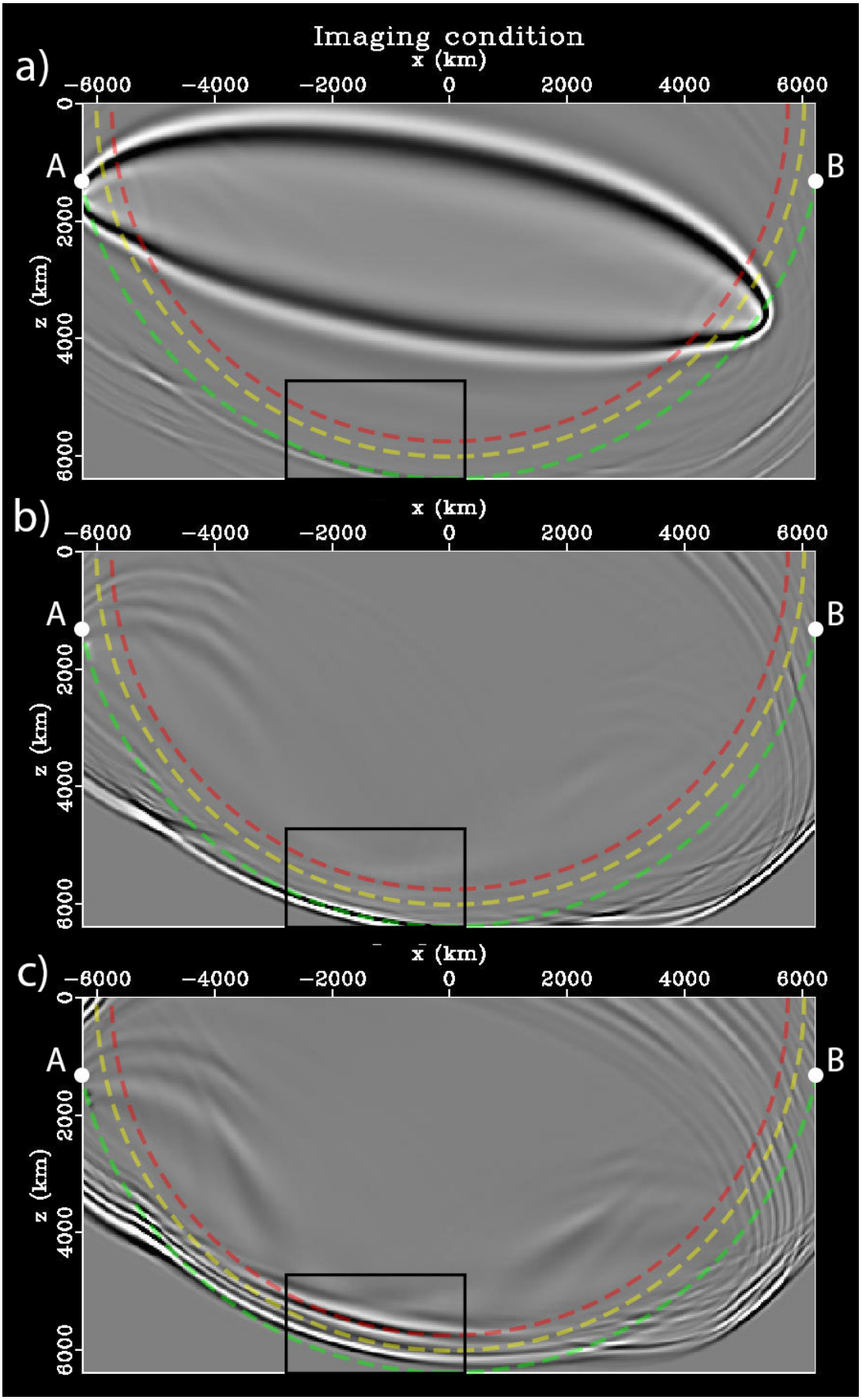

Figure 18: Imaging conditions with source wavelet peak frequency 0.02 Hz and migrated data of a) data set 1 (raw data); b) data set 2 (SS plus precursor window); c) data set 3 (filtered precursor window). Dashed lines represent the Earth's surface (green), the 410-discontinuity (yellow) and 660-discontinuity (red). A and B are the ends of the section line drawn in Figure 14. Black boxes outline the part in which the SS reflector aligns with the green dashed line.

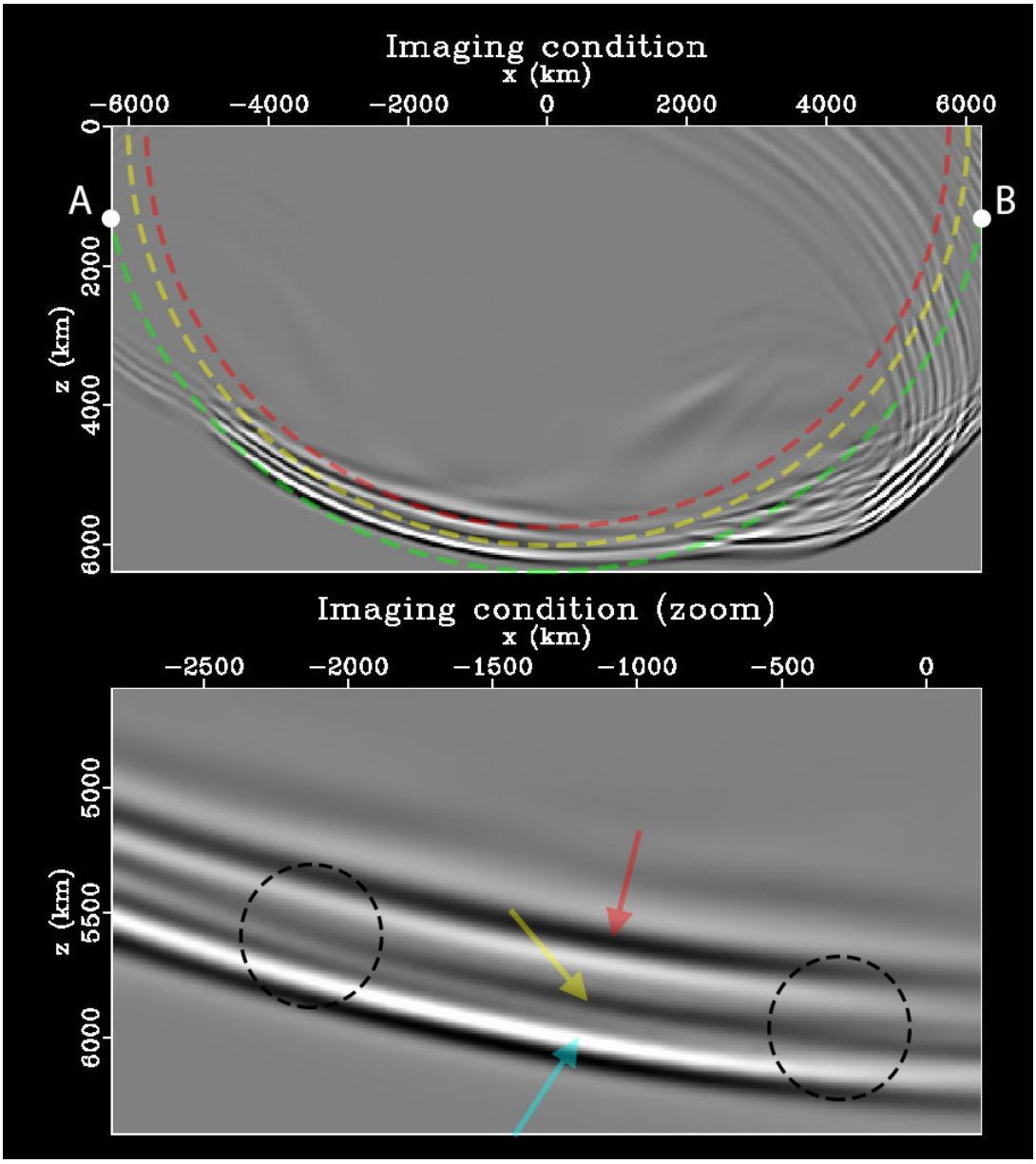

Figure 19: a) Imaging condition for migrated data set 4 (normalized, filtered, polarity reversed precursor window) and source peak frequency 0.02 Hz, divided by source illumination. b) Zoomed in on the black box of a). Arrows indicate the presumed peaks of the reflections, with red for the 660-reflection, yellow for the 410-reflection and blue for the reflection at ca. 200 km depth. Black dashed circles highlight changes in the 410-discontinuity depth.

successfully yet, whereas source illumination normalization has been implemented. The source-normalized image of the migrated precursor window has subdued reflections near the source location and wherever the power of the source wave field is high. This is easily spotted in a comparison between Figure 18c and Figure 19a. The reflectors that are not physical (mostly outside of the Earth near point A) and had high amplitudes purely for the source wave field having high power there, have faded away in Figure 19.

Figure 19b shows a zoomed in portion of the source-normalized imaging condition. The two reflections representing the transition zone boundaries at roughly 410 and 660 km depth have a negative (black) peak, whereas the most shallow reflection at roughly 200 km depth

seems to have a positive (white) peak. Peaks of the reflections are indicated with arrows. The 410-reflector seems to show some topography; it appears to move closer to the 660-discontinuity in the left black circle in Figure 19b, and move farther away from the 660-discontinuity in the circle towards B.

3.3 Results: synthetic modelling experiments

Several synthetic modelling experiments are set up for resolution tests and to explain features in the imaging conditions gathered from migrating real data. For these experiments, source wave fields are propagated through a background velocity model with and without added artificial reflectors. The difference between the synthetic data generated at the receivers in both models is the scattered data caused by these artificial reflectors. The steps that follow are equal to the steps taken for migrating real data. The scattered data is reversed in time and propagated through the background model from the receiver side, creating a receiver wave field. Source wave fields are also propagated through the background medium. The imaging condition is extracted in the usual way, by cross-correlating source and receiver wave fields at zero lag and summing over the sources.

To test resolution, point scatterers are added to the background velocity model (masked model, number 3 in list of models). The scatterers are placed at the boundaries of the transition zone, with four at 410 km and four at 660 km depth (Figure 20a). They are smoothed to keep the solution from becoming numerically unstable. In the resulting imaging condition for the real receiver array, the individual scatterers are highly distorted (Figure 20b). Therefore, images obtained for real data will be blurred as well. It implies that the amplitude at a point in the imaging condition is a function of the amplitudes at multiple surrounding points. The blurred point scatterers merge into a reflector appearing continuous, because of the source-reflector-receiver geometry. Even only two point scatterers will form a continuous reflector. To test whether poor resolution results from the limited extent of the receiver array, receivers are placed along the Earth's surface in a synthetic experiment. If receivers are placed everywhere at the Earth's surface, i.e. along a half-circle in the model, individual point scatterers are distinguishable (Figure 20c). When receivers are placed in a full enclosing circle around the scatterers, they would appear as points. Due to finite frequency, the point scatterers would still be a bit blurred and not be perfect points.

Another test that can be done with synthetic modelling, is adding a curved reflector to the background model (Figure 21a) and apply migration. If the curvature is well-resolved in the imaging condition, it could be possible to image curved reflectors for real data. Unfortunately, the curved reflector does not appear curved in the imaging condition (Figure 21b). So if there would be any large scale undulations in the discontinuities, they would not be well-resolved. At least one of the reasons is limited aperture, because when receivers are placed at the entire Earth's surface, most of the reflector's curvature is resolved in the imaging condition (Figure 21c).

To make sense of the features observed in the data imaging condition, two artifical strong discontinuities at 410 and 660 km depth are added to the background model in a final synthetic modelling experiment (Figure 22a). As we have seen, in the imaging condition for a migrated precursor window (data sets 3 and 4), three reflectors light up (Figure 19b). In the imaging condition for synthetic modelling with two discontinuities, three reflectors light up as well (Figure 22b). When the two images are compared, the three reflection depths at roughly 200, 410 and 660 km correspond to each other, with a slight phase shift. This argument makes it even more plausible to say that the deepest reflector (top reflector in the image) is the 660-discontinuity and the middle reflector is the 410-discontinuity. The 410-reflector is somewhat weaker than

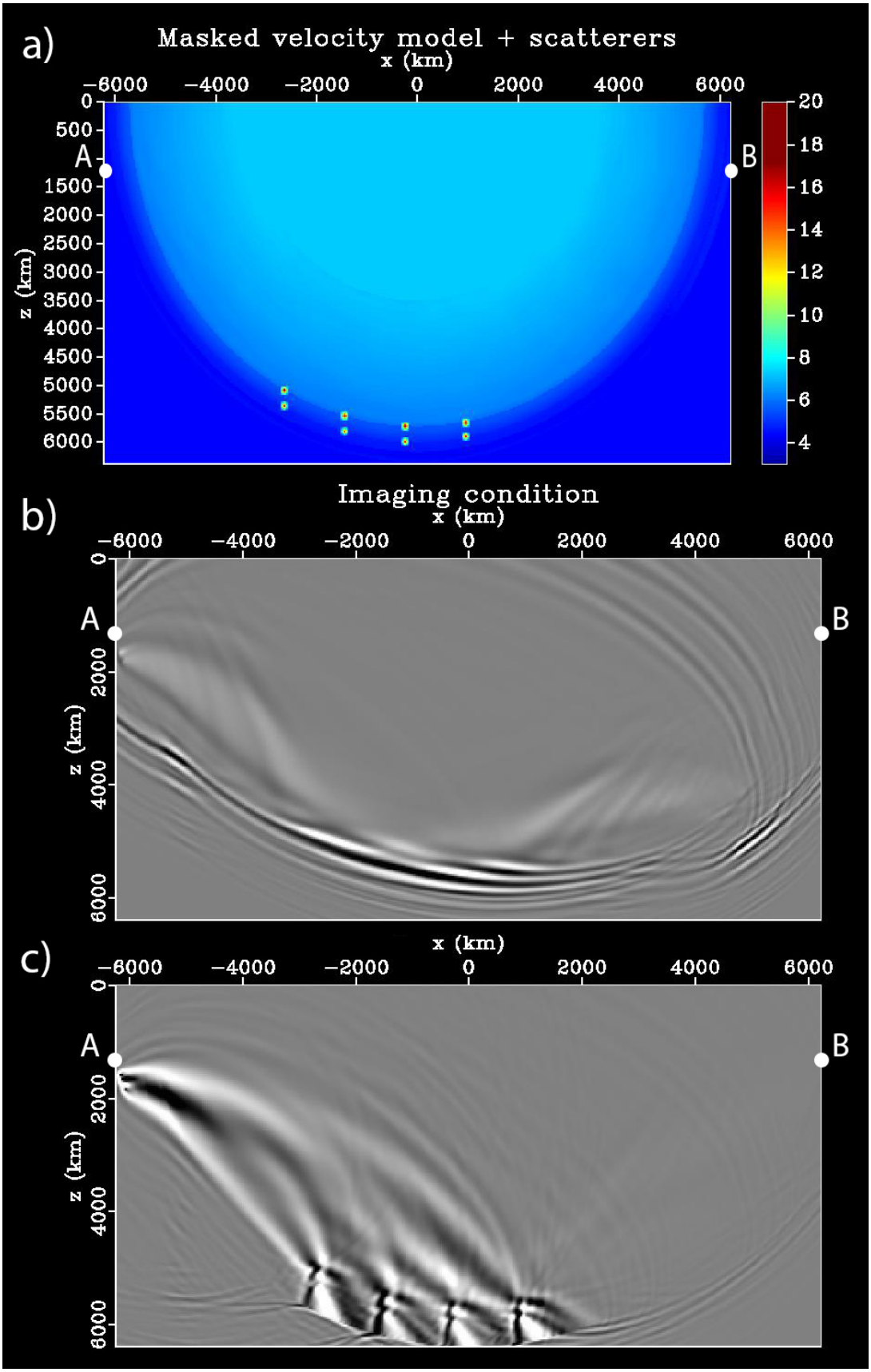

Figure 20: a) Masked velocity model with artificial point scatterers at 410 and 660 km depth. Velocities are in km/s. It is used for synthetic modelling, yielding the imaging conditions for b) the real receiver geometry; c) receivers placed along the entire Earth's surface. The point scatterers are distinguishable in c), in contrast to the limited aperture of the real receiver array in b). The point scatterers do not appear as points in c), because the receivers form a half-circle around the scatterers, rather than a full circle.

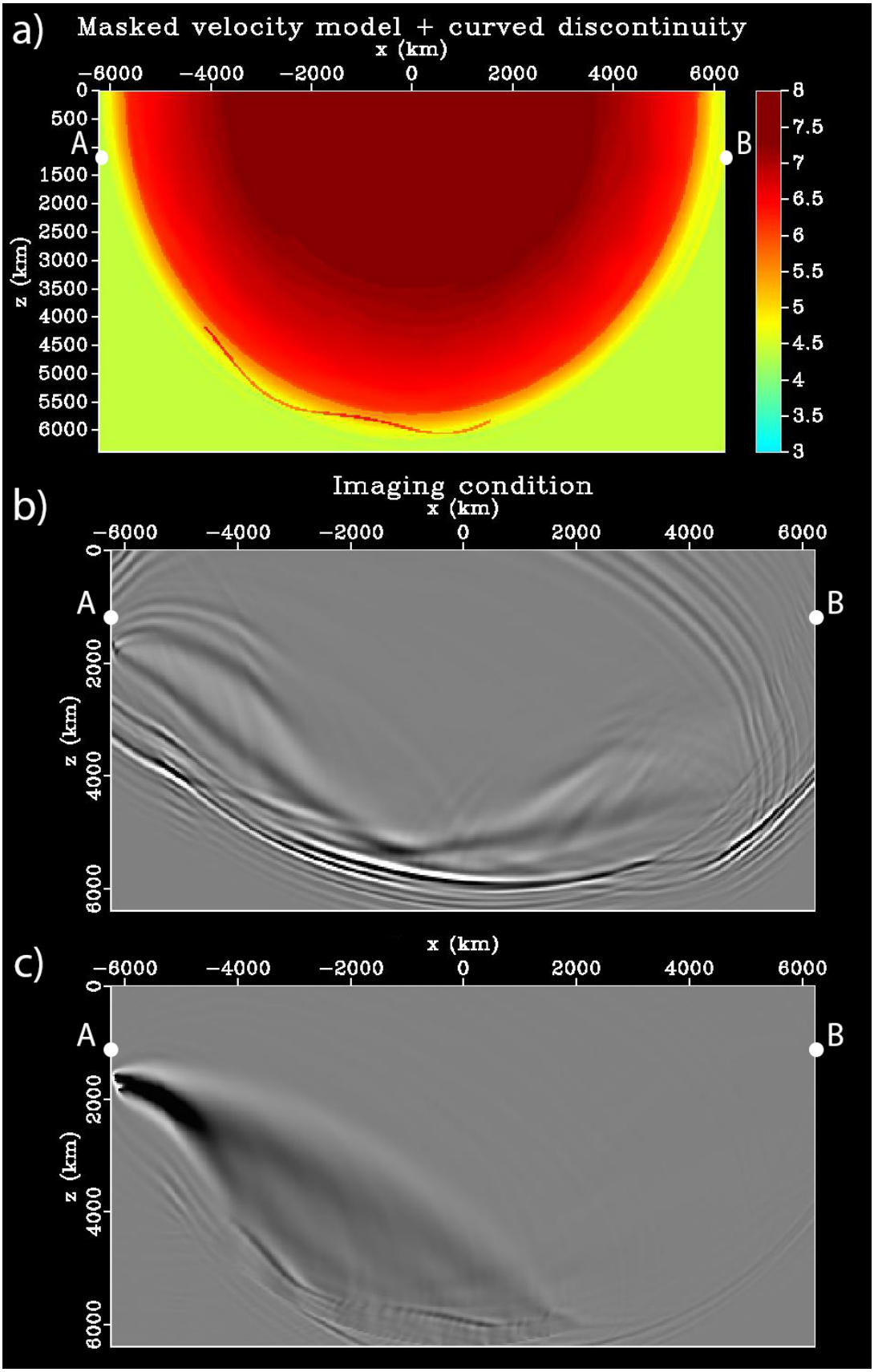

Figure 21: a) Masked velocity model with a strongly curved artificial discontinuity at a mean depth of 410 km. Velocities are in km/s. It is used for synthetic modelling, yielding the imaging conditions for b) the real receiver geometry; c) receivers placed along the entire Earth's surface. The curvature of the discontinuity is recovered in the imaging condition in c).

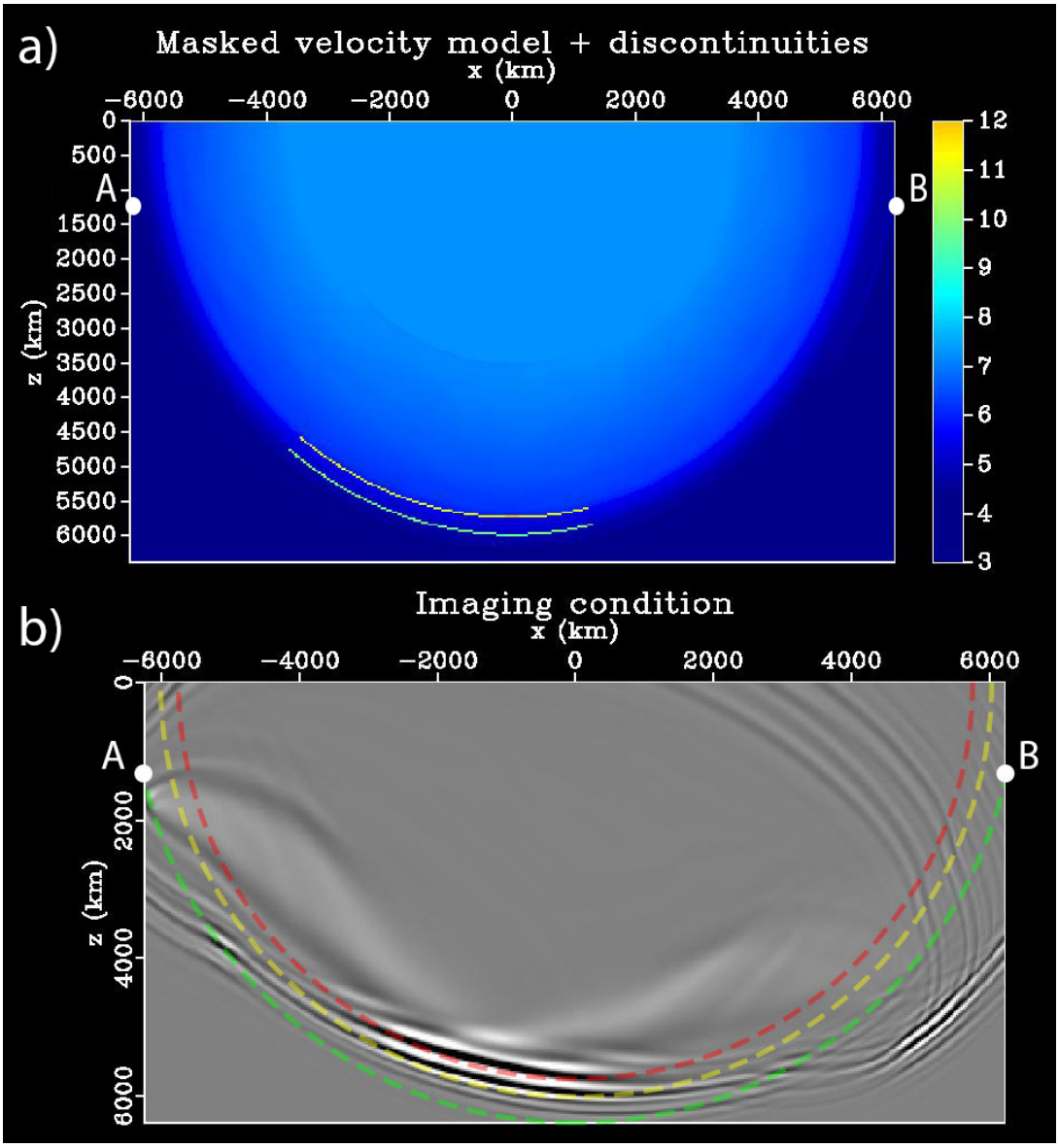

Figure 22: a) Masked velocity model with artificial discontinuities at 410 and 660 km depth. It is used for synthetic modelling, to yield b) the imaging condition, with source wavelet peak frequency 0.02 Hz. An extra reflector appears vaguely in between the surface (green dashed line) and 410-reflector (yellow dashed line).

the 660-reflector. It is probably caused by the data rather than blurring. The point scatterer resolution test showed that blurring affects features at 660 km depth as badly as at 410 km depth. Furthermore, the S410S phase has on average a smaller amplitude than S660S in the stacked data of Figure 8c.

The most shallow reflector at ca. 200 km depth in Figure 22b (lowest in the image) is less prominent than the other two in the synthetic image and appears to have a negative polarity, instead of positive in Figure 19b. The amplitude of the 200-reflector is one order of magnitude smaller than the amplitudes of the transition zone discontinuities. Since only scattered data by the artificially added reflectors are migrated for the synthetic image, this reflector must be related to scattering from the artificial reflectors. It is most likely a multiple following the path depicted in Figure 23, which is projected to a position closer to the Earth's surface by

migration, because it arrives later in the data. Cutting the real data from -300 to -110 seconds with respect to the theoretical SS arrival, instead of -300 to -60 seconds, gets rid of the most shallow reflector. This proves that it is a coherent phase that arrives later in the data than the two transition zone SS precursors, S410S and S660S. However, a multiple would be expected to have a smaller amplitude than the underside reflections that have only one bounce point. The data image does not show this, contrary to the synthetic image. Furthermore, a multiple would have a negative polarity, as in the synthetic image. The reflection has a positive polarity in the real data image. The origin of this extra reflector in the data will be discussed in the Discussion.

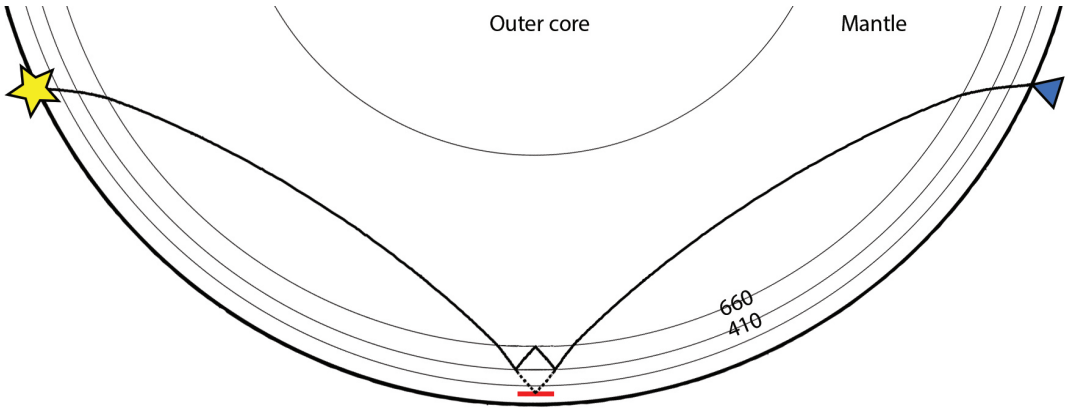

Figure 23: Ray path (solid black line) of the multiple reverberating once between the 410- and 660-discontinuities. The phase is projected according to the dashed lines to the location of the red line in migration, which explains the vague reflector in between the surface and 410-discontinuity in synthetic imaging.

4 Discussion

4.1 Interpretation of stacking results

In part 1, S410S and S660S have been robustly resolved using conventional stacking analysis in the slowness-time domain, according to an error analysis (Figure 12). Although this error analysis (wrongly) assumes a normal distribution and does not fully include several assumptions made in the slant-stacking approach, which will be discussed later, it gives an indication of robustness of the peaks. The presence and non-splitting of S520S in the most southern bins in the study area is consistent with observations from Deuss & Woodhouse (2001). They did not find this mid-transition zone discontinuity near the equator in the Pacific, but did observe it further away from the equator.

Transition zone thickness and depths of the 410- and 660-discontinuity have been inferred from slant-stacking. Of these three variables, transition zone thickness is most reliable. When calculating transition zone thickness, velocity corrections largely cancel out, because S410S and S660S follow similar paths. It is mainly the velocity structure in the bounce point region that affects S410S and S660S differently (Lawrence & Shearer, 2008). Absolute discontinuity depth depends more on the velocity model used to go from time to depth. Near the three hotspots, the transition zone is slightly thinner (1-5 km) than the global average (Figure 9). A thin transition zone would correspond to an intermediate to deep hotspot origin in an olivine-dominated mantle composition. However, this interpretation can change when uncertainties are accounted for, to be discussed below.

A regional study of the same three hotspots (Marqueses, Tahiti and Pitcairn) following a similar stacking procedure, found a thickness of 213 ± 5 km for the transition zone beneath Tahiti (Niu et al., 2002). This is much thinner than the transition zone thickness of the bin closest to Tahiti (bin 10), which has a thickness of 241 km. In the study by Niu et al. (2002) it is mainly the result of a deep 410-discontinuity at 448 ± 4 km depth. For the other hotspots, the transition zone thickness did not show any clear anomalies. The most logical explanation put forward by the authors is a low velocity anomaly caused by a mantle plume that has not been accounted for in the time-to-depth conversion. This velocity anomaly would lead to 10 km more shallow discontinuities and the situation of Figure 1B for olivine-dominated discontinuities. Combined with the observation of Montelli et al. (2006) of a well-resolved low velocity conduit in tomographic images below Tahiti, it is plausible that Tahiti is fed by a mantle plume. Discrepancies between this study and that of Niu et al. (2002) could come from differences in the stacking method and/or variations in data coverage and azimuths of source-receiver pairs, since data coverage here is much more sparse around the hotspot locations and azimuths are more aligned than in their study.

There are some assumptions made in the slant-stacking approach, that are relaxed in the migration approach. First of all, the slant-stacking method does not account for waveform propagation effects as function of frequency. These propagation effects could have shifted the position of the precursor peaks, thereby shifting the discontinuity depths.

Secondly, the velocity model used for inferring discontinuity depths is an assumption that affects the resulting discontinuity depths. In this study the 1-D velocity model PREM was used. Because using PREM in converting from time to depth does not account for 3-D and crustal velocity structure, resulting discontinuity topography may be the result of anomalous velocity structure. On the ray paths from source to receivers, waves encounter high-velocity zones in the form of subducted slabs near New Zealand and the Pacific coast of North-America, and the cold thick North-American craton. However, most of the mantle velocities in the Pacific Ocean are lower than the 1-D average, according to S40RTS (Ritsema et al., 2011). In a PP and SS precursor study of the transition zone beneath India (Lessing et al., 2014), travel time

corrections for crust and mantle heterogeneities resulted in a systematic depth error estimation of ±8 km in discontinuity depth. In a receiver functions study of mantle discontinuities beneath Iceland (Jenkins et al., 2016), accounting for the crust and 3-D variations in velocity made the depressions of both mantle discontinuities found with PREM less extreme (ca. 30 km less deep), while transition zone thickness remained relatively invariant to the choice of velocity model. It should be noted that low velocity anomalies are up to -11% in the Icelandic region (Rickers et al., 2013), so for the Pacific region the effect of velocity model will likely be smaller.

Absolute depths should nevertheless be interpreted with care, even in the studies that have used a 3-D velocity model to go from time to depth. For SS precursors, the time residual between measured travel time and predicted travel time by ray theory through a 1-D background model is the combined result of 3-D heterogeneities and discontinuity topography. A complication for observing discontinuity depths is that the travel time residual cannot be linearly decomposed into a topography part and a 3-D velocity structure part (Koroni & Trampert, 2016). To properly infer discontinuity topography, one needs to use non-linear techniques, such as full waveform inversion, to obtain both topography and 3-D velocity structure (Koroni & Trampert, 2016).

Other assumptions are made in dividing the study area into bounce point bins. In the high frequency approximation, i.e. ray theory, the rays bounce midway at a single point and their sensitivity is limited to that point. Seismograms are grouped according to their ray theoretical bounce point in the stacking procedure, whereas in reality, the bounce points are not single points. Wave energy interpreted as precursor arrivals recorded at the receiver may not have originated from the bounce point location. Rather, it may have originated from a scatterer far away in the X-shaped Fresnel zone. Excluding seismograms from stacking that have their ray theoretical bounce point just outside of a bin/cap does not take all the energy in the bin into account. An improvement on this aspect could be weighting the data by distance from bounce point to the stack location (Shearer et al., 1999; Niu et al., 2002) or including the sensitivity kernels of each trace in stacking (Lawrence & Shearer, 2008).

4.2 Interpretation of migration results

In part 2, migration was applied to the same data set. In selecting an appropriate time window, i.e. the precursor window, the 410- and 660-discontinuities can be imaged. The resulting imaging condition in the migration approach (Figure 19b) shows topography of the 410-discontinuity in the black circles. Projection of the locations of the two black circles onto the section line gives locations on either side just outside of the bounce point covered region of Figure 14. This raises the question whether the observed topography is due to true discontinuity topography or due to reaching the limits of the region covered sufficiently by the data. The discontinuities in the synthetic image of Figure 22b do not show topography in the locations of the black circles, which is an argument in favour of real discontinuity topography. However, whether the migrated image of the data is actually showing discontinuity topography or an imaging artefact, requires further investigation.

An extra reflector at approximately 200 km depth can be observed in imaging conditions both for real and synthetic data. In synthetic imaging, it is a multiple, since it is an order of magnitude smaller than the synthetic 410- and 660-discontinuities and has the same polarity. In real data, it could be a S220S precursor, reflected off the Lehmann discontinuity, which has been observed in the study area (Deuss & Woodhouse, 2004). However, the polarity of the extra reflector seems to be opposite to those of the 410- and 660-discontinuities, and its amplitude is 1.5 times larger than of the 410- and 660-discontinuities, which makes the S220S precursor hypothesis illogical. The answer can be found in the stacking results, where a large negative peak

occurs around -100 seconds relative to SS, especially visible in Figures 12 and 7, corresponding to ca. 200 km depth. It has an opposite polarity to the S410S and S660S precursors and a larger amplitude. The peak is most likely a consequence of stacking, since it is not distinguishable in individual seismograms. It is probably part of the long-period down swing before the SS arrival (Schmerr & Garnero, 2006).

Poor resolution inferred from scatterer tests and other synthetic modelling experiments illustrates the resolution challenge, since the amplitude of one image point depends on the surrounding image points as well. The high degree of blurring in this study is a consequence of migration with limited aperture and finite frequency approximation and without the integral blurring operator. Wave energy is not focussed in an infinitely narrow beam but rather it is spread over an area. Furthermore, migration allows for waveform propagation effects as function of frequency and at finite frequency ranges. Methods based on ray theory cannot adequately model wave propagation effects in complex media (Rickett, 2000). That is the reasoning for using a wave equation migration approach. However, depth of the reflectors is still dependent on the background velocity model. The migration background model (PREM + S40RTS) is more laterally heterogeneous and arguably more realistic than PREM alone, which is used in slant-stacking.

Another assumption in migration is that the medium is acoustic, because we have used the scalar wave equation as starting point of the theoretical derivation. It implies that the waves are purely compressional and the shear modulus is equal to zero. Since the real data contain mostly S-wave energy, imaging condition amplitudes should not be interpreted quantitatively. The positions of scatterers can still be interpreted qualitatively. How can we justify using a scalar wave equation for S-waves? The homogeneous wave equation for elastic wave propagation in an infinite homogeneous isotropic medium is given by (Aki & Richards, 2002):

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + 2\mu) \nabla (\nabla \cdot \mathbf{u}) - \mu \nabla \times (\nabla \times \mathbf{u}) \tag{49}$$

where μ and λ are the Lamé parameters and ρ the density. The total vector wave field $\mathbf{u}$ can be decomposed into a scalar (ϕ) and vector ($\overline{\psi}$) Helmholtz potential:

$$\mathbf{u} = \nabla \phi + \nabla \times \overline{\psi} \tag{50}$$

in which $\nabla \phi$ and $\nabla \times \overline{\psi}$ are the P-wave and S-wave components of $\mathbf{u}$ (Aki & Richards, 2002). Solving for the vector potential (S-wave component) results in the following differential equation for S-wave potential, with $\beta = \sqrt{\mu/\rho}$ as S-wave velocity.

$$\left(\nabla^2 - \frac{1}{\beta^2} \frac{\partial^2}{\partial t^2} \right) \overline{\psi} = 0 \tag{51}$$

Note the similarity to equation 3 in the frequency domain. Equation 51 has vector potential $\overline{\psi}$ instead of a scalar potential, as in equation 3 used in migration. The solution to equation 51 is a three-component vector. Reducing the vector wave equation to a scalar wave equation for S-waves is equivalent to only taking one of the three components into account. For isotropic media, this approximation is justified.

4.3 Future research

Two methods have been used to image transition zone discontinuities. The first method, slant-stacking, is a tried and tested approach (e.g. Lawrence & Shearer, 2008; Deuss, 2009; Saki et al., 2015), whereas reverse time migration for global seismology is a novel approach. This study

raises opportunities for further research into these methods. Assumptions have been made in both methods. However, their effects have not been quantified. It would, for example, be useful to do the time-to-depth conversion for the stacked traces with different velocity models, accounting for various degrees of heterogeneity, and see how much the outcome changes.

As of yet, it is hard to truly compare the outcomes of the two methods, because they live in different coordinate frames and because of different implicit assumptions regarding wave propagation regimes. To make the results more comparable, discontinuity depths should be extracted from the migrated images by making radial cross-sections through the migrated images. To make the two methods more similar, migration could be done in 3-D for the same 15 bounce point bins employed in stacking. Migration was done in 2-D, which required 3-D to 2-D conversions and geometrical spreading corrections. Now that we know the discontinuities can be imaged, migration can be done directly in 3-D to see topographic changes in the entire study area, if resolution permits it. Resolution of the images can be improved further by techniques such as deconvolution imaging, as explained in the section on improving the imaging condition.

Finally, a more extensive receiver array would improve the resolution of the images, as synthetic modelling experiments have shown. The current source-receiver configuration has too little coverage. Acquisition of a wider receiver array is limited by the preferred epicentral distance range of 100-160 degrees for SS precursors. Fortunately, migration can be used to include different phases that are more suitable for smaller epicentral distances, such as ScS reverberations.

5 Conclusion

The transition zone discontinuities beneath three hotspots in the Pacific Ocean have been successfully imaged with SS precursors in slant-stacking and reverse time migration. Data of three earthquakes near New Zealand were divided into spherical caps of 10° width according to their geometrical bounce point and subsequently stacked with slowness corrections. Stacking results show that the transition zone is 1-5 km thinner in bounce point bins covering the three hotspots (Marqueses, Tahiti and Pitcairn), possibly indicating a hot olivine-dominated transition zone and mantle plume origin. The 410-discontinuity is 1-4 km deeper than the estimated global average for Tahiti and Pitcairn, whereas the thinner transition zone for Marqueses is caused by a 7 km more shallow 660-discontinuity than average. These results may change with the velocity model that is used to convert time of the precursors to discontinuity depth.

Reverse time migration has been applied to the same data. It relaxes some of the assumptions made in slant-stacking, because it is a finite frequency method that takes the Fresnel zone and waveform propagation effects into account. Although they might show some topography of the 410-discontinuity, the resulting migrated images are blurred, because of limited aperture and finite frequency wave propagation. Synthetic modelling experiments show that the migration method itself will probably be more successful for a wider receiver array than the current set-up.

6 Acknowledgements

I would like to thank my supervisors Arwen Deuss and Ivan Pires de Vasconcelos for their guidance and enthusiasm throughout this project. Thanks to the “lunchgroepje” for the pleasant distraction every day, especially to Bente de Vries and Laura Coumou. Figures 5, 7-14, 23 were made with GMT mapping software (Wessel & Smith, 1991). Figures 3 and 6 were made with ObsPy.

7 References

Aki, K., & Richards, P. G. (2002). Quantitative seismology (Vol. 1).

Bai, Y., Bao, G., Cao, J.J., Cheng, H., Dai, Y.H., Dong, C.Z., Dulikravich, G.S., Egorov, I.N., Fu, C.L., Gelius, L.J. & Hao, L.L. (2012). Computational methods for applied inverse problems (Vol. 56). *Walter de Gruyter*.

Cao, Q., Wang, P., Van der Hilst, R. D., De Hoop, M. V., & Shim, S. H. (2010). Imaging the upper mantle transition zone with a generalized Radon transform of SS precursors. *Physics of the Earth and Planetary Interiors*, 180(1), 80-91.

Chambers, K., & Woodhouse, J. H. (2006). Investigating the lowermost mantle using migrations of long-period S-ScS data. *Geophysical Journal International*, 166(2), 667-678.

Chattopadhyay, S., & McMechan, G. A. (2008). Imaging conditions for prestack reverse-time migration. *Geophysics*, 73(3), S81-S89.

Claerbout, J. F. (1971). Toward a unified theory of reflector mapping. *Geophysics*, 36(3), 467-481.

Claerbout, J. F., & Fomel, S. (2008). Image estimation by example: Geophysical soundings image construction: multidimensional autoregression. *Stanford University*.

Clouard, V., & Bonneville, A. (2001). How many Pacific hotspots are fed by deep-mantle plumes?. *Geology*, 29(8), 695-698.

Courtillot, V., Davaille, A., Besse, J., & Stock, J. (2003). Three distinct types of hotspots in the Earth's mantle. *Earth and Planetary Science Letters*, 205(3), 295-308.

Crotwell, H.P., Owens, T.J., & Ritsema, J. (1999). The TauP Toolkit: Flexible seismic travel-time and ray-path utilities, *Seismological Research Letters*, 70, 154-160.

Dai, W., Boonyasiriwat, C., & Schuster, G. T. (2010). 3D multi-source least-squares reverse time migration. In *SEG Technical Program Expanded Abstracts 2010* (pp. 3120-3124). Society of Exploration Geophysicists.

Dai, W., Fowler, P., & Schuster, G. T. (2012). Multi-source least-squares reverse time migration. *Geophysical Prospecting*, 60(4), 681-695.

Davies, G.F. (2000). Dynamic Earth: Plates, Plumes and Mantle Convection, *Cambridge University Press, New York*.

Deuss, A. (2007). Seismic observations of transition-zone discontinuities beneath hotspot locations. *Geological Society of America Special Papers*, 430, 121-136.

Deuss, A. (2009). Global observations of mantle discontinuities using SS and PP precursors. *Surveys in geophysics*, 30(4-5), 301-326.

Deuss, A., & Woodhouse, J. (2001). Seismic observations of splitting of the mid-transition zone discontinuity in Earth's mantle. Science, 294(5541), 354-357.

Deuss, A., & Woodhouse, J. H. (2004). The nature of the Lehmann discontinuity from its seismological Clapeyron slopes. *Earth and Planetary Science Letters*, 225(3), 295-304.

Deuss, A., Andrews, J., & Day, E. (2013). Seismic observations of mantle discontinuities and their mineralogical and dynamical interpretation. *Physics and Chemistry of the Deep Earth*, 295-323.

Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4), 297-356.

Dziewonski, A. M., Chou, T. A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research: Solid Earth*, 86(B4), 2825-2852.

Ekström, G., Nettles, M., & Dziewonski, A. M. (2012). The global CMT project 2004-2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200, 1-9.

Hirose, K. (2002). Phase transitions in pyrolitic mantle around 670-km depth: Implications for upwelling of plumes from the lower mantle. *Journal of Geophysical Research: Solid Earth*, 107(B4).

Hu, J., Schuster, G. T., & Valasek, P. A. (2001). Poststack migration deconvolution. *Geophysics*, 66(3), 939-952.

Ita, J. & Stixrude, L. (1992). Petrology, elasticity and composition of the mantle transition zone, *Journal of Geophysical Research*, 97(B5), 6849-6866.

Jenkins, J., Cottaar, S., White, R. S., & Deuss, A. (2016). Depressed mantle discontinuities beneath Iceland: Evidence of a garnet controlled 660 km discontinuity?. *Earth and Planetary Science Letters*, 433, 159-168.

Karato, S. (2011). Water distribution across the mantle transition zone and its implications for global material circulation. *Earth and Planetary Science Letters*, 301(3-4): 413-423.

Kennett, B. L. N., & Engdahl, E. R. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2), 429-465.

Kennett, B. L. N., Engdahl, E. R., & Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1), 108-124.

Koroni, M., & Trampert, J. (2016). The effect of topography of upper-mantle discontinuities on SS precursors. *Geophysical Journal International*, 204(1), 667-681.

Lawrence, J. F., & Shearer, P. M. (2008). Imaging mantle transition zone thickness with SdS-SS finite-frequency sensitivity kernels. *Geophysical Journal International*, 174(1), 143-158.

Lessing, S., Thomas, C., Rost, S., Cobden, L., & Dobson, D. P. (2014). Mantle transition zone structure beneath India and Western China from migration of PP and SS precursors. *Geophysical Journal International*, 197(1), 396-413.

Miller, D., Oristaglio, M., & Beylkin, G. (1987). A new slant on seismic imaging: Migration and integral geometry. *Geophysics*, 52(7), 943-964.

Montelli, R., Nolet, G., Dahlen, F. A., & Masters, G. (2006). A catalogue of deep mantle plumes: New results from finite-frequency tomography. *Geochemistry, Geophysics, Geosystems*, 7(11).

Neele, F., de Regt, H., & Van Decar, J. (1997). Gross errors in upper-mantle discontinuity topography from underside reflection data. *Geophysical Journal International*, 129(1), 194-204.

Nemeth, T., Wu, C., & Schuster, G. T. (1999). Least-squares migration of incomplete reflection data. *Geophysics*, 64(1), 208-221.

Niu, F., Solomon, S. C., Silver, P. G., Suetsugu, D., & Inoue, H. (2002). Mantle transition-zone structure beneath the South Pacific Superswell and evidence for a mantle plume underlying the Society hotspot. *Earth and Planetary Science Letters*, 198(3), 371-380.

Nolet, G., Allen, R., & Zhao, D. (2007). Mantle plume tomography. *Chemical Geology*, 241(3), 248-263.

Rickers, F., Fichtner, A., & Trampert, J. (2013). The Iceland-Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion. *Earth and Planetary Science Letters*, 367, 39-51.

Rickett, J. (2000). Traveltime sensitivity kernels: Banana-doughnuts or just plain bananas?.

Ringwood, A.E. (1975). Composition and Petrology of the Earth's Mantle, *McGraw-Hill, New York*.

Ritsema, J., van Heijst, H. J., & Woodhouse, J. H. (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science*, 286(5446), 1925-1928.

Ritsema, J., Deuss, A., Van Heijst, H.J., & Woodhouse, J.H. (2011). S40RTS: a degree-40 shear-velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltime and normal-mode splitting function measurements. *Geophysical Journal International*, 184(3), 1223-1236.

Rost, S., & Thomas, C. (2009). Improving seismic resolution through array processing techniques. *Surveys in Geophysics*, 30(4-5), 271-299.

Saki, M., Thomas, C., Nippress, S. E., & Lessing, S. (2015). Topography of upper mantle seismic discontinuities beneath the North Atlantic: the Azores, Canary and Cape Verde plumes. *Earth and Planetary Science Letters*, 409, 193-202.

Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1), 10-21.

Shearer, P. M., Flanagan, M. P., & Hedlin, M. A. (1999). Experiments in migration processing of SS precursor data to image upper mantle discontinuity structure. *Journal of Geophysical Research: Solid Earth*, 104(B4), 7229-7242.

Squires, G. L. (2001). Practical physics. *Cambridge university press*.

Tromp, J., Tape, C., & Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1), 195-216.

Wessel, P., & Smith, W. H. F. (1991). Free software helps map and display data, *Eos Trans. AGU*, 72, 441, 445-446.

Yu, Y. G., Wentzcovitch, R. M., Vinograd, V. L., & Angel, R. J. (2011). Thermodynamic properties of MgSiO3 majorite and phase transitions near 660 km depth in MgSiO3 and Mg2SiO4: A first principles study. *Journal of Geophysical Research: Solid Earth*, 116(B2).

Zheng, Z., & Romanowicz, B. (2012). Do double 'SS precursors' mean double discontinuities?. *Geophysical Journal International*, 191(3), 1361-1373.

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. *Geophysical Journal International*, 202(1), 175-189.

Appendices

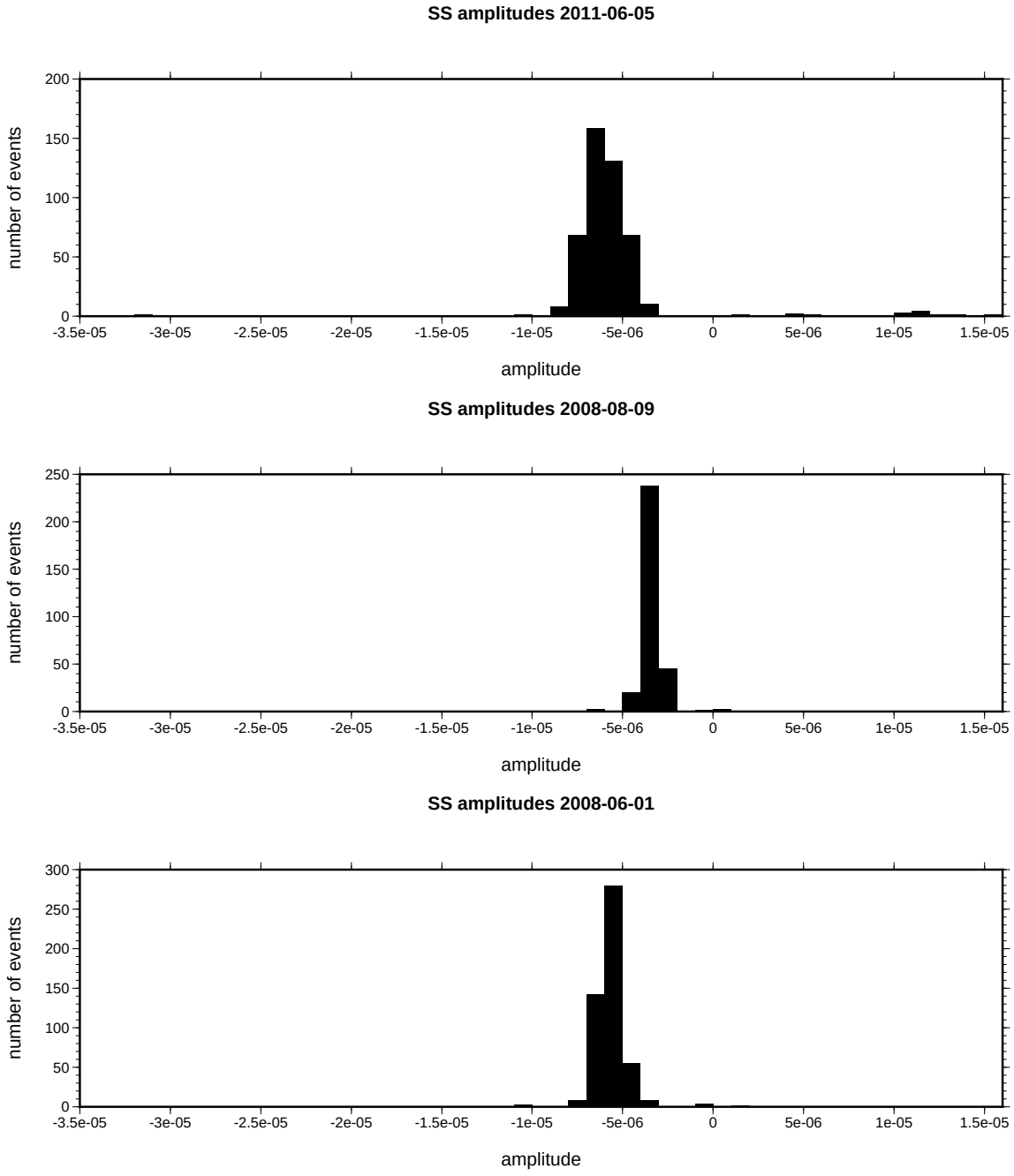

Figure A.1: Amplitudes of SS picks before normalization for all three events.