Fine-scale structure of the mid-mantle characterised by global stacks of PP precursors

H.L.M. Bentham [a,*], S. Rost [a], M.S. Thorne [b]

[a] *School of Earth & Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK*
[b] *Department of Geology and Geophysics, University of Utah, Salt Lake City, UT 84112, USA*

A B S T R A C T

Subduction zones are likely a major source of compositional heterogeneities in the mantle, which may preserve a record of the subduction history and mantle convection processes. The fine-scale structure associated with mantle heterogeneities can be studied using the scattered seismic wavefield that arrives as coda to or as energy preceding many body wave arrivals. In this study we analyse precursors to PP by creating stacks recorded at globally distributed stations. We create stacks aligned on the PP arrival in $5°$ distance bins (with range 70–120°) from 600 earthquakes recorded at 193 stations stacking a total of 7320 seismic records. As the energy trailing the direct P arrival, the P coda, interferes with the PP precursors, we suppress the P coda by subtracting a best fitting exponential curve to this energy. The resultant stacks show that PP precursors related to scattering from heterogeneities in the mantle are present for all distances. Lateral variations are explored by producing two regional stacks across the Atlantic and Pacific hemispheres, but we find only negligible differences in the precursory signature between these two regions. The similarity of these two regions suggests that well mixed subducted material can survive at upper and mid-mantle depth. To describe the scattered wavefield in the mantle, we compare the global stacks to synthetic seismograms generated using a Monte Carlo phonon scattering technique. We propose a best-fitting layered heterogeneity model, BRT2017, characterised by a three layer mantle with a background heterogeneity strength ($\epsilon = 0.8\%$) and a depth-interval of increased heterogeneity strength ($\epsilon = 1\%$) between 1000 km and 1800 km. The scalelength of heterogeneity is found to be 8 km throughout the mantle. Since mantle heterogeneity of 8 km scale may be linked to subducted oceanic crust, the detection of increased heterogeneity at mid-mantle depths could be associated with stalled slabs due to increases in viscosity, supporting recent observations of mantle viscosity increases due to the iron spin transition at depths of ∼1000 km.

© 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Mantle convection is the process that drives the interaction of tectonic plates and recycles oceanic lithosphere introduced into the mantle at subduction zones. Most of our knowledge of the present day convective system comes from seismic tomography (e.g. Ritsema et al., 2011; Van der Hilst et al., 1997), which inverts for seismic travel times and waveforms, and has revealed large seismic velocity variations at both the upper and lower mantle boundaries. In particular, tomographic studies have imaged fast velocity features associated with subducted slabs, some of which are continuous from the surface to the core–mantle boundary (CMB) (e.g. Van der Hilst et al., 1997). Such deep subduction is evidence

for the whole mantle interacting in one convective system rather than in separate multi-layered cells. In contrast, some studies have also imaged flat lying fast velocity features in the transition zone and mid-mantle (e.g. Sigloch and Mihalynuk, 2013), suggesting there may be some barriers to these downwellings such as an increase in viscosity just below the transition zone (e.g. Forte and Mitrovica, 2001) or as deep as 2000 km (Justo et al., 2015) due to the iron spin transition.

Subducted slabs partly consist of basaltic crust, which has a different composition to that of mantle peridotite. The crust deforms slowly over time in reaction to convection related stresses (Stixrude and Lithgow-Bertelloni, 2012) leading to long-lived heterogeneity that preserves the path taken by subduction. These heterogeneities have scale lengths on the order of ∼10 km and are below the current resolution levels of global tomography models (>100 km). Thus other seismic methods, such as analysis of the high frequency scattered seismic wavefield (Sato and Fehler, 1998),

have to be employed to map these heterogeneities and analyse their properties (e.g. Rost et al., 2006). Since subducting plates are sensitive to barriers to flow, we may expect higher density of heterogeneities at boundaries of convection cells. These barriers might only influence thermal convection and the uniformity of mixing, and therefore may not be obvious in tomographic images. To provide additional temporal constraints on subduction processes, the crustal component of slabs can be traced using seismic scattering approaches.

Most previous seismic scattering studies have tried to characterise heterogeneity in the deep or whole mantle and different, often contradicting, distributions of heterogeneity in the mantle have been suggested. Some evidence has been put forth in support of heterogeneities distributed evenly throughout the mantle (Earle and Shearer, 2001; Hedlin et al., 1997; Hedlin and Shearer, 2000; Mancinelli and Shearer, 2013; Margerin and Nolet, 2003a; Shearer and Earle, 2004). In contrast, models with radially varying heterogeneity have also been proposed. Many of the models prefer strong heterogeneities concentrated in the lowermost mantle (Bataille and Flatté, 1988; Cleary and Haddon, 1972; Doornbos, 1978; Niu and Wen, 2001; Tono and Yomogida, 1996) but others suggest a reduction in heterogeneity in the lowermost mantle is necessary (Shearer and Earle, 2004). Furthermore, Hedlin and Shearer (2000) found that models are not well constrained and multiple models can explain the same data. The lack of agreement may result from modelling the whole mantle simultaneously without independent constraints on the upper and mid-mantle heterogeneity structure.

In this study we use precursors to the PP arrival to characterise global averages of heterogeneity since we feel PP precursors provide unique insight into mid-mantle structure that cannot be achieved with other seismic probes (e.g. Bentham and Rost, 2014; Rost et al., 2008; Shearer and Flanagan, 1999). We create global and regional seismic stacks and build on the modelling procedure established by Shearer and Earle (2004) by quantifying the model misfit and systematically searching for the best fitting heterogeneity model. We consider models with constant heterogeneity in the mantle and also increase model complexity by varying heterogeneity with depth to gain insight into the resultant effect of different scattering distributions on the PP precursory wavefield.

We characterise heterogeneity from the lithosphere to the mid-mantle using global signatures of the high frequency PP precursory wavefield. Through stacking of a large global dataset we identify radial changes in heterogeneities, and established whether lateral variations of heterogeneities exist through grouping data into two hemispheres. The heterogeneity structure in the lithosphere and mantle is characterised through forward modelling of the scattered wavefield using a 1D Monte Carlo phonon method (Shearer and Earle, 2004) and the resulting synthetic envelopes are compared to the observed global stacks. The PP precursory wavefield is analysed and modelled for a range of distances, attempting to resolve the depth dependence of the scattering giving rise to the precursory energy. We generate more than 150 models of radial scattering heterogeneity and show that the heterogeneity parameters can be constrained when systematically varied. To limit the size of the parameter space examined in the forward modelling we only explore contributions from four different lithospheric layers within a fixed thickness of 100 km. Additionally, we focus on the contribution from the mantle by considering radial variations in scattering scale length and Root Mean Square (RMS) velocity perturbations, and find the best fitting model overall has depth varying velocity perturbation.

2. Global stacking of PP

2.1. Data

We create stacks of the seismic wavefield recorded by 193 seismometers (Fig. 1). Most of the stations used are part of the Global Seismic Network (GSN) (Albuquerque Seismological Laboratory (ASL)/USGS, 1988; Scripps Institution of Oceanography, 1986) with supplementary stations from the USArray Transportable Array (TA) (IRIS Transportable Array, 2003); Canadian National network (CN), POLARIS network (POL) and Canadian Northwest Experiment (XN) added to improve coverage. These networks are selected due to their appropriate distance range from regions of high earthquake activity. Vertical broadband data for 600 earthquakes (Fig. 1) are obtained from the Incorporated Research Institutions for Seismology (IRIS) database. Earthquakes are selected from 2003 to 2012 with depths from 0 to 100 km, magnitudes (Mw) larger than 5.8, and epicentral distance between source and station from 70° to 120°. The chosen distance range extends the range used in previous studies of scattered PP precursors (Bentham and Rost, 2014; Rost et al., 2008; Wright, 1972). The extended distance range analysed in this work allows us to detect heterogeneities within a larger depth range of the mantle. For an epicentral distance of 70°, PP turns at about 850 km depth and for 120° PP turns at 1550 km depth (see Suppl. Material, Fig. S1) allowing good sampling of the uppermost lower and mid-mantle. Therefore variations in the character of PP in stacks across this distance range should be linked to properties of the mantle between 850 km and 1550 km. Furthermore PP precursors arrive with slowness similar to the PP arrival (Rost et al., 2006), it is likely that the precursors travel to similar depths as PP (see Suppl. Material, Fig. S1). However some scattered arrivals may have scattered multiple times (Shearer, 2007) and/or have travelled out of the great circle plane, thus travelling with deeper or shallower depths than PP.

2.2. Pre-processing

Data are re-sampled to 100 samples per second and filtered using a two way bandpass filter with corner frequencies at 0.5 Hz and 2.5 Hz. This frequency band is chosen as it is sensitive to the small-scale (10 km) spatial wavelengths in lower mantle (Shearer and Earle, 2004) and beneficially, the high frequency noise is removed. The envelope time function of each trace is calculated to obtain phase independent amplitudes for stacking. The stacking of noise can be reduced by ensuring the mean of the noise is zero. An estimate of the mean noise level is found using pre-signal noise calculated from 60 s to 35 s before P (or P_{diff}), and this level is then subtracted from the complete time series before stacking.

We chose to pick observed PP traveltime as the absolute maximum amplitude of the PP waveform using theoretical times as calculated through IASP91 (Kennett and Engdahl, 1991) as a guide. The signal to noise ratio (SNR) of PP amplitude to the amplitude of the noise before P (or P_{diff}) is calculated and traces with SNR less than 5 are removed. The data are organised into 5° distance bins, centred at 72.5°, 77.5°, 82.5°, 87.5°, 92.5°, 97.5°, 102.5°, 107.5°, 112.5° and 117.5°.

2.3. Stacking of PP

Data are aligned on PP and stacked (summed) within each distance bin, normalised by the number of traces and the absolute amplitude of PP (Fig. 2). The number of traces in each distance bin generally decreases with distance, ranging from a minimum of ∼500 traces for 117.5° to a maximum of >1000 traces for 72.5° (Fig. 2). The final stacked dataset provides good global coverage
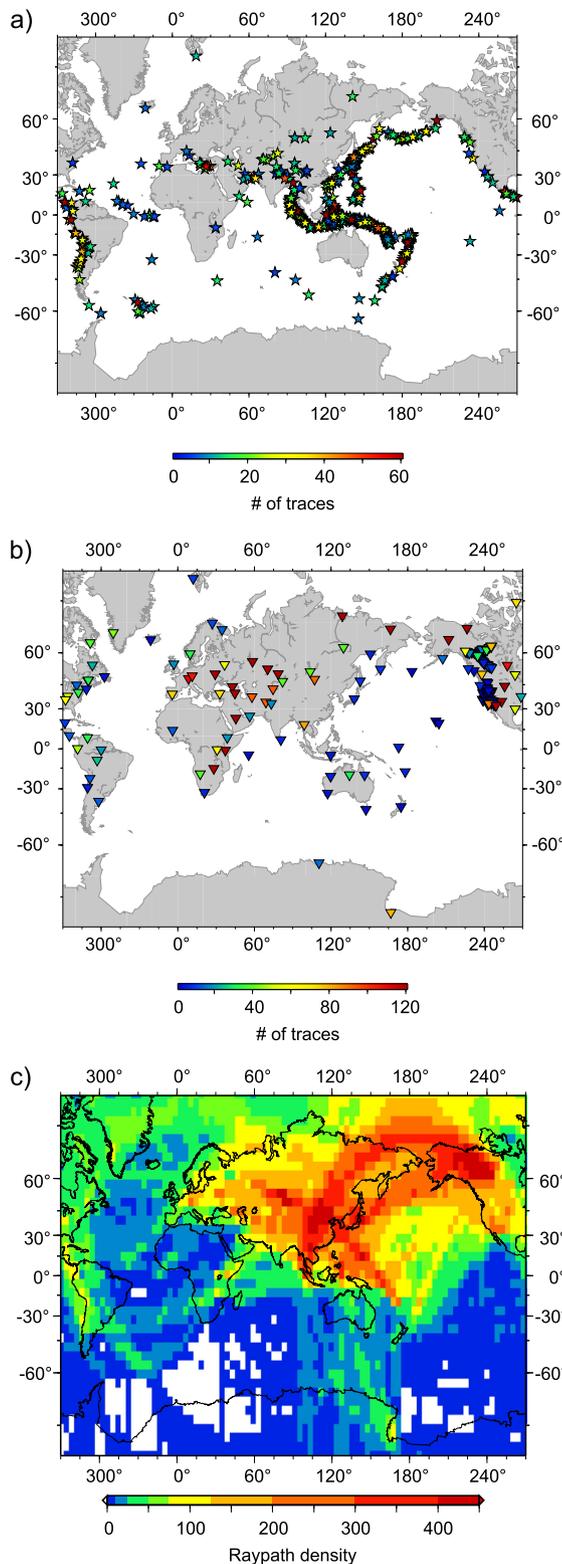
Fig. 1. a) Earthquake locations used in global stacks with symbol (stars) colour showing the number of traces used from each earthquake in the stacks (max. 64 traces). b) Stations used for global stacking of PP precursors (triangles), with symbol colour showing the number of traces used from each station in the stacks. c) Spatial sampling of the seismic traces used in the global stacks. Raypath density in 4° bins is found by summing the source–receiver great circle paths from every trace used in stacking. Most of the globe is sampled, with the greatest density observed in west North America, Japan and south–east Asia. There is poor sampling for Antarctica both due to limited station coverage and earthquakes occurring in this region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
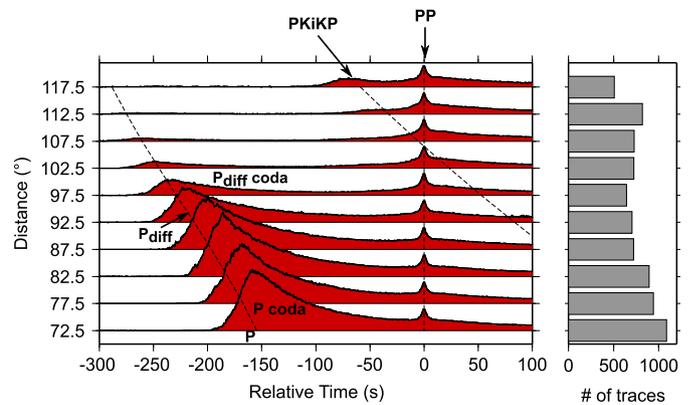
Fig. 2. Global envelope-function stacks of 7905 traces stacked in 5° distance bins. Time is shown relative to the PP arrival as indicated. P, P_{diff} and PKiKP travel times from IASP91 (Kennett and Engdahl, 1991) are marked as thin, black, dashed lines. P and P_{diff} are non-impulsive and the arrivals gradually increases with time due to aligning on PP and stacking in 5° distance bins. The number of seismograms stacked in each bin are shown in the right panel. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

with the best sampled region beneath the Pacific Ocean and poorest sampling beneath Antarctica (Fig. 1).

The wavelet shape of PP is consistent in all distance stacks, and the precursory energy before PP is also clearly visible (Fig. 2). However despite aligning on PP before stacking, P coda and P_{diff} coda amplitudes are also large in the same time window (∼100 s before PP). For shorter distances, 97.5° and less, P coda is dominant and decays slowly. For greater distances (>97.5°), the P coda energy is not prevalent. The transition in P coda dominance is likely caused by the direct P wave energy starting to diffract along the core–mantle boundary (i.e. P becoming P_{diff}) at ∼98°.

As the presence of P coda causes difficulty in examining the PP precursors for distances between 70° to 95° (Fig. 2) we remove the P coda interference before stacking. As the P coda is predominantly produced by scattering and intrinsic attenuation in the lithosphere for both source- and receiver-sides, and thus decays exponentially (Sato and Fehler, 1998), we remove the energy by fitting and subtracting an exponential trend to the coda for each seismic trace. Other studies avoid this issue by using deep earthquakes (Shearer and Earle, 2004) but in this study we find that the depth phase pP would arrive in the PP precursor time window adding complexity to the precursory wavefield. After removing the P coda, the data are stacked within each distance bin by summing the traces together and normalising by the stacked amplitude of PP (Fig. 3).

After removing the P coda, PP precursors and coda energy are clearly visible for all distances (Fig. 3 and 4). The PP precursory energy starts within 100 s prior to the onset of PP and increases in strength with time (Figs. 3 and 4). The onset of the precursors is difficult to determine since the energy builds gradually over time and the amplitude of the base noise level is unclear. The PP coda amplitude also increases with distance (from 30% to 50% normalised amplitude) and decays slowly still retaining ∼20% of PP amplitude 100 s after the PP arrival (Fig. 3 and 4).

Other phases are expected to align sufficiently to be visible in the stacks. The core phase PKiKP is present in the 112.5° and 117.5° stacks in Fig. 4. However, PKiKP is not visible for distances shorter than the crossover distance of 105° where PKiKP becomes post-critical. In contrast, the underside reflections from the 410 km and 660 km discontinuities are noticeably absent for all distances studied (Fig. 3 and 4) indicating that they are not coherent in this global stack, possibly due to depth variations of the responsible phase transitions.
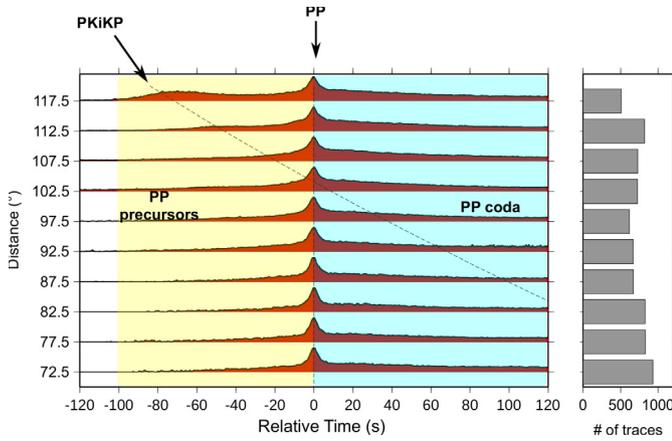
Fig. 3. Envelope-function stacks for all data after processing and P coda removal. The numbers of seismograms summed within each 5 distance bin are shown on the right panel, for a total of 7320 traces. Time is shown relative to the PP arrival as indicated. PP, PP precursors (yellow box), PP coda (cyan box) and PKiKP are also identified. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
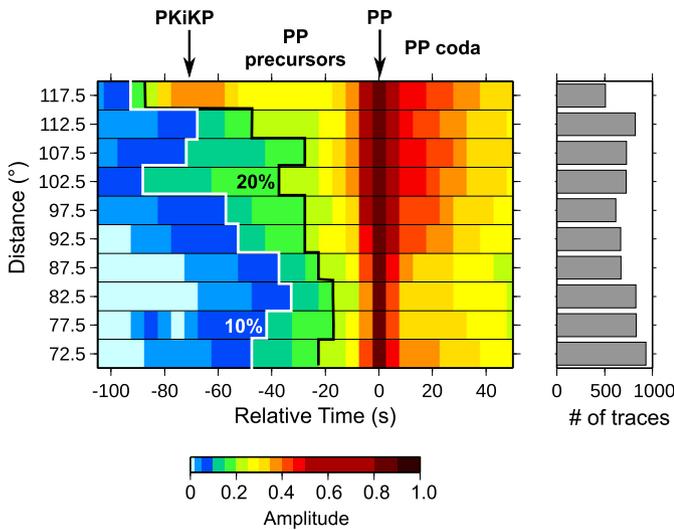
Fig. 4. Power density representation of the stacks for all data after processing and P coda removal. The numbers of seismograms summed within each 5 distance bin are shown on the right panel, for a total of 7320 traces. Seismograms are aligned on PP and time is relative to the PP. The onset of PP precursors is emergent and difficult to pick. Variations in precursor strength can be found by comparing constant PP amplitude ratios, specifically contours of 10% (white) and 20% (black) of amplitude of PP. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

3. Regional differences in PP stacks

We create two regional stacks of our data to look for lateral variations in the PP precursory wavefield that are sensitive to heterogeneities in the upper and mid-mantle. The stacks are generated for two quasi-hemispheric regions: the Pacific and the Atlantic (Fig. 5) and these regions are primarily selected as they likely have experienced different tectonic and mantle processes. In particular the subduction history of the Atlantic includes older, ceased subduction across most of north America, active subduction in South America and no recent subduction beneath the eastern Atlantic, although some subduction in the western Atlantic has been postulated (Domeier et al., 2016). Likely the subducted material in the upper mantle from subduction at the western shores of the Americas is located too deep to be detected with the data of our quasi Atlantic hemisphere. In contrast, the Pacific has had long
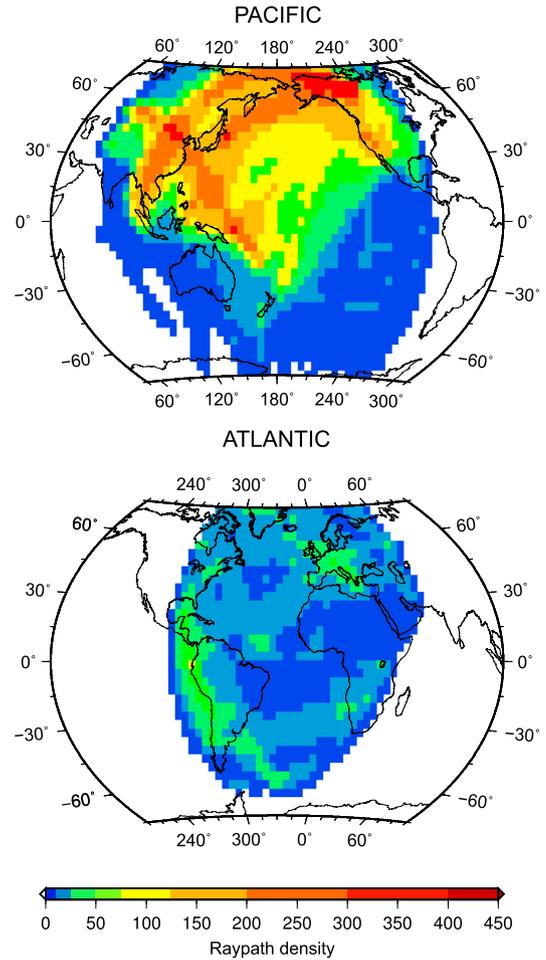

Fig. 5. Map showing PP raypath coverage for two hemispherical regions in this study. The Pacific region (left) uses earthquakes and stations with longitudes in the range of 150°W to 60°E, whereas the Atlantic region (right) has longitudinal limits of 150°W and 60°E. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

lasting subduction over the last 120 Myr (Domeier et al., 2016). We select data for source-receiver paths between longitudes 60°E and 60°W for the Pacific region with a total of 4166 traces contributing to the stack and longitudes 150°W and 60°E for the Atlantic region (a total of 896 traces). Both datasets produce smooth PP precursor stacks. Fig. 5 shows the PP raypath coverage for each region.

Following the same procedure of the global stacks (Section 2.3) the P coda is removed, the time series are aligned on PP and traces are summed (Fig. 6). In general, the PP precursors and the PP coda are almost identical in both envelope and amplitude for both regions. Because the data are normalised on PP amplitude we would expect that changes in heterogeneity structure would lead to amplitude variations or differences in the shape of the precursor envelope. This is not observed for most distances, except minor differences for the distance bin from 85° to 90°. We expected differences between the regions as they have different tectonic histories, but instead we find that tectonics have not left significant imprints on the fine scale structure of the upper & mid-mantle. Smaller scale regional differences might exist but further subdividing the dataset might lead to unrepresentative and rough stacks.

4. Monte Carlo phonon scattering modelling

We employ a forward modelling approach to find a 1D heterogeneity model in agreement with the observed data. Due to the
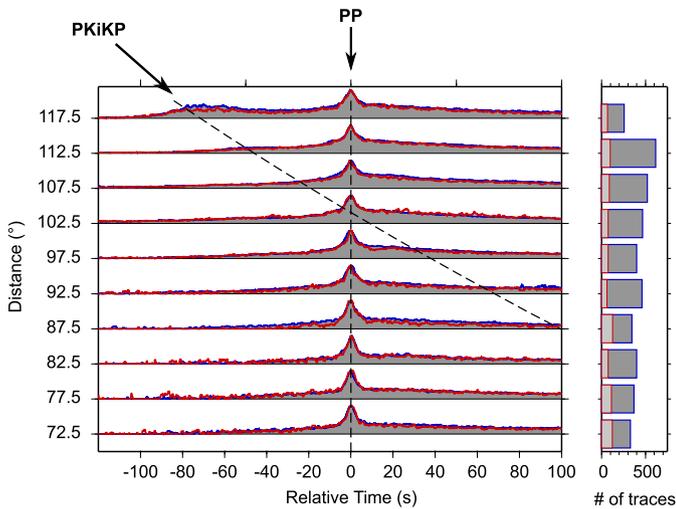
Fig. 6. Comparison of the Pacific and Atlantic regional envelope-function stacks with P coda removed. Time is relative to the PP arrival as indicated. P, P$_{diff}$ and PKiKP are marked as thin, black, dashed lines. Number of traces in each distance bin (right) are: 4166 traces for Pacific stack (blue outline, dark grey fill) and 896 for Atlantic stack (red outline, no fill). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lack of regional differences on a hemispheric scale, we aim to fit the global averages representing the global average of small-scale structure. For modelling the PP precursory wavefield in this study, a Monte Carlo phonon based method by Shearer and Earle (2004) is used. The approach uses radiative transfer theory to model energy transport of phonons through whole Earth scattering models, assuming random (velocity and density) perturbations represented by an exponential autocorrelation function, and taking into account the energy reduction in the primary wavefield and permitting multiple scattering. This method is chosen to model the PP precursors because it can simulate high frequency multiple scattering that is depth dependent. Multiple scattering has been shown to contribute a substantial amount of energy to the P coda and therefore modelling of multiple scattering in the lithosphere, and potentially deeper in the mantle, is necessary (Shearer, 2007). Additionally, intrinsic attenuation is incorporated into the modelling and can vary with depth. The method assumes 1D velocity structure and is limited by ray-theoretical assumptions not allowing us to model diffracted arrivals or lateral heterogeneity structure. Since we do not seek to model for diffracted waves, nor non-radial variations in heterogeneity, both limitations do not restrict our modelling approach.

The approach defines the exponential heterogeneity structure in radial layers through the RMS fractional velocity fluctuation, ϵ, and the heterogeneity correlation scale length, a. These two parameters, as well as the radial shell locations, are changed to find a best fitting model to the data. The following parameters are fixed for all models:

1. Intrinsic attenuation – crust and mantle attenuation values are selected for 1 Hz (Warren and Shearer, 2000). For P waves: $Q = 227$ from 0 to 100 km and $Q = 138$ from 220 to 2889 km. For S waves, attenuation is $\frac{4}{9}$ of P wave attenuation assuming a Poisson solid and that all attenuation is in shear. Inner core attenuation of $Q_{core} = 360$ (Bhattacharyya et al., 1993) is applied to both the inner and outer core to prevent strong PKiKP at distances >110° in the modelling.
2. The P/S energy ratio is chosen to be a constant value of 23.4 following Shearer and Earle (2004) assuming a double-couple source.

3. P and S velocities have the same fractional velocity variation and the density/velocity fluctuation scaling factor is set to 0.8.
4. The background velocity profile are taken from IASP91 (Kennett and Engdahl, 1991) and the density structure from PREM (Dziewonski and Anderson, 1981).

In order to find the most appropriate mantle heterogeneity model to fit the globally stacked data, we varied the number of scattering layers (N) as well as the thickness (h_n), scattering correlation length (a_n) and velocity fluctuation (ϵ_n) of each scattering layer. The modelled synthetic seismograms are converted to envelopes and the P coda is removed in the same way as for the data before comparing PP and PP precursors.

A Root Mean Square (RMS) misfit of each model (Δ_d) is calculated between the observed stacked data (D) and the synthetic modelled data (M) for each distance bin (d). The misfit is calculated at every sample (i) with a sample interval of 1 s within a 40 s time window (starting at 40 s before PP). This time window contains the clearest precursor information for all distances. The RMS misfit for one distance stack is calculated as:

$$\Delta_d = \sqrt{\frac{\sum_{i=1}^{S} (D_i - M_i)^2}{S}}$$

Therefore the mean misfit ($\overline{\Delta}$) over all 10 distance bins (bin centres from 72.5° to 117.5°) is defined as:

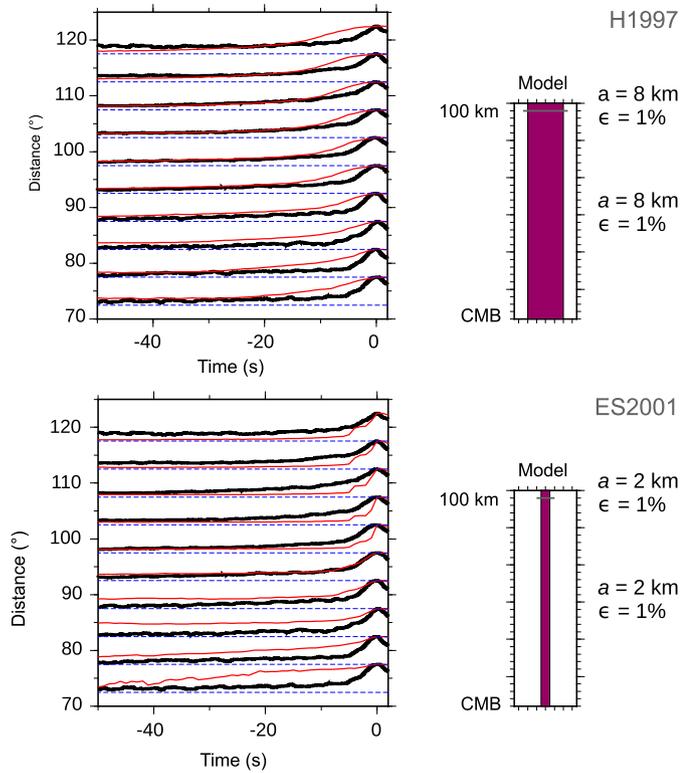$$\overline{\Delta} = \frac{1}{10} \sum_{d=1}^{10} \Delta_d$$

Calculating the misfit in this way allows us to determine how well a specific model fits the data over all distance ranges that are roughly equivalent to different sampling depths. Additionally, as the model misfit is determined on the sample level, we can describe the temporal evolution of the fit of the scattering envelope.

4.1. Existing heterogeneity models

As a starting point we compare how well existing heterogeneity models fit the PP precursor data from this study. Due to the varying sensitivity of different scattering probes to the heterogeneity structure of the mantle, differences in fit between the models can be expected. We use the heterogeneity parameters reported by three studies: Hedlin et al. (1997) [H1997], Earle and Shearer (2001) [ES2001] and Shearer and Earle (2004) [SE2004]. For each model synthetic stacks are produced as described above (see Suppl. Material, Table S1). Models H1997 and ES2001 are quite similar as they have depth independent heterogeneity (through lithosphere and mantle) with 1% velocity variation but differing correlations of 8 km and 2 km respectively. In contrast SE2004 has depth varying heterogeneity with correlation length a, ranging from 4 to 8 km and velocity variation ϵ, ranging from 0.5 to 4% (Table S1).

For these models the P coda is predicted to be stronger than observed in the global stacked data. The effect is most noticeable for model SE2004. In this model, the P coda is extended in time and dominates the PP precursory window whilst completely masking PP. In contrast for H1997 and ES2001, there is clear separation between the P-coda and PP precursors for most distances (Fig. 7). In particular the shape of precursory energy in the data stacks is best fit by model H1997 for distances up to 97.5°. However, neither H1997 nor ES2001 provide a good match for distances larger than 97.5°, indicating that depth variation of heterogeneity may be required.

In summary, none of the previously published models explored here fit the PP stacked data, however, H1997 and ES2001 models have a reasonable fit for some distances indicating that a more

Fig. 8. Computed RMS misfit for systematic search of whole mantle heterogeneity models. Misfit is shown for 196 models across four lithospheric models of correlation length a_L & velocity variation ϵ_L: 2 km/1% (bottom left), 2 km/2% (top left), 8 km/1% (bottom right) and 8 km/2% (top right). Within each lithospheric model group there are 48 models with a_M ranging from 2 km to 20 km and ϵ_M ranging from 0.1% to 2%. The lowest RMS misfit in each group is marked with a white circle except for model with the lowest misfit over all groups (top left): model $a_L/\epsilon_L/a_M/\epsilon_M$ of 2 km/2%/8 km/0.8% marked with a white star. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)
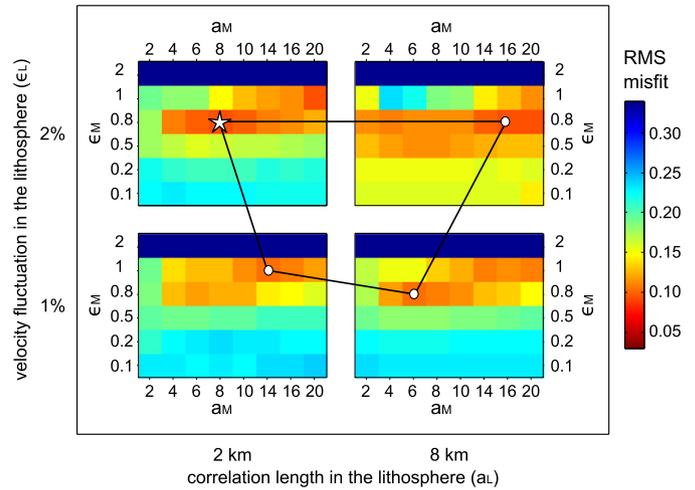
Fig. 7. Comparison of global stacks (black) to synthetic seismograms (red) from published models: H1997 (Hedlin et al., 1997) and ES2001 (Earle and Shearer, 2001) for 5° distances bins from 70°–120°. The P coda has been removed from both the global and modelled stacks. The parameters used in the modelling are shown on the right. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

complicated radial heterogeneity structure may be present than is captured by these models (Fig. 7). We use models H1997 and ES2001 as base models in constructing our model space including models with constant (Section 4.2) and depth varying (Section 4.3) mantle heterogeneity.

4.2. Modelling with constant mantle heterogeneity

Using models H1997 and ES2001 (Section 4.1) we build and test models with homogeneous velocity variation and correlation length throughout the mantle (i.e. depth range 100–2891 km for all models). We start by constructing a lithospheric layer (0–100 km depth) using four model types based on H1997 and ES2001. These four lithospheric model types considered are (a/ϵ): i) 2 km/1%; ii) 2 km/2%; iii) 8 km/1%; iv) 8 km/2%. Next, we vary the correlation length and RMS velocity variation in the mantle for a range of parameters. For each lithospheric model type, 48 models with various correlation length (2 km, 4 km, 6 km, 8 km, 10 km, 14 km, 16 km and 20 km) and velocity variation (0.1%, 0.2%, 0.5%, 0.8%, 1.0% and 2.0%) were assessed, resulting in a combined total of 192 models.

The misfits between model and data are shown in Fig. 8. The total misfit ranges from 0.094 to 0.838 (nearly an order of magnitude), with a mean of 0.243 and standard deviation of 0.180. All models with mantle ϵ of 2% are characterised by a high misfit, regardless of mantle correlation length and lithospheric heterogeneity. The best fitting models in each group have ϵ of 0.8 to 1% but there is a wide range of correlation lengths amongst these models that fit these data equally well. As the optimum models have similar mantle a_M and ϵ_M, this may suggest that lithospheric heterogeneity is not the main contributor to the PP precursory wavefield. To test this we compare the data fit of the four models

with $a_M = 8$ km and $\epsilon_M = 0.8\%$ and another model without any lithospheric structure. We find that all model misfits are <0.15, similar misfit to the best fitting models including lithospheric heterogeneity (see Suppl. Material, Fig. S2). Additionally, there seems to be a trade-off between all parameters (a_L, ϵ_L, a_M and ϵ_M) collectively, rather than trade-off between the layers (i.e. correlation length a_L vs a_M and velocity variation ϵ_L vs ϵ_M).

Though lithospheric structure only has minor influence on model misfit, the best model overall (with a misfit of 0.0939) has lithospheric parameters of $a_L = 2$ km and $\epsilon_L = 2\%$ and mantle parameters of $a_M = 8$ km and $\epsilon_M = 0.8\%$. Systematic search of these parameters provides strong evidence that average RMS mantle velocity variations are likely with the range from 0.8 to 1%. Though the best mantle correlation length is 8 km, it is less well defined and ranges between 6 km to 16 km since there is a trade-off between mantle and lithospheric heterogeneity parameters. The correlation length range might indicate that the mantle heterogeneity contains a variety of scale lengths.

4.3. Modelling with depth varying mantle heterogeneity

4.3.1. Modelling with depth varying mantle heterogeneity I – two layer case

For the single layer models, the misfits between data and model vary with distance and the lowest misfits are observed in the distance range from 90° to 100°. This suggests that the small-scale structure may deviate from the single layer model for depth intervals corresponding to the varying turning depths of PP. Therefore, we attempt to improve the overall model misfit by introducing depth varying mantle heterogeneity to the best fitting single layer model. Initially, we explore two cases, each with two layers of heterogeneity in the mantle:

- Model type 2A – increase velocity variation in the mantle from 0.8% (upper layer) to 1% (lower layer) and vary the depth of the transition from 200 km to 2800 km (every 200 km);
- Model type 2B – decrease velocity variation in the mantle from 1% (upper layer) to 0.8% (lower layer) and vary the depth of the transition from 200 km to 2800 km (every 200 km).
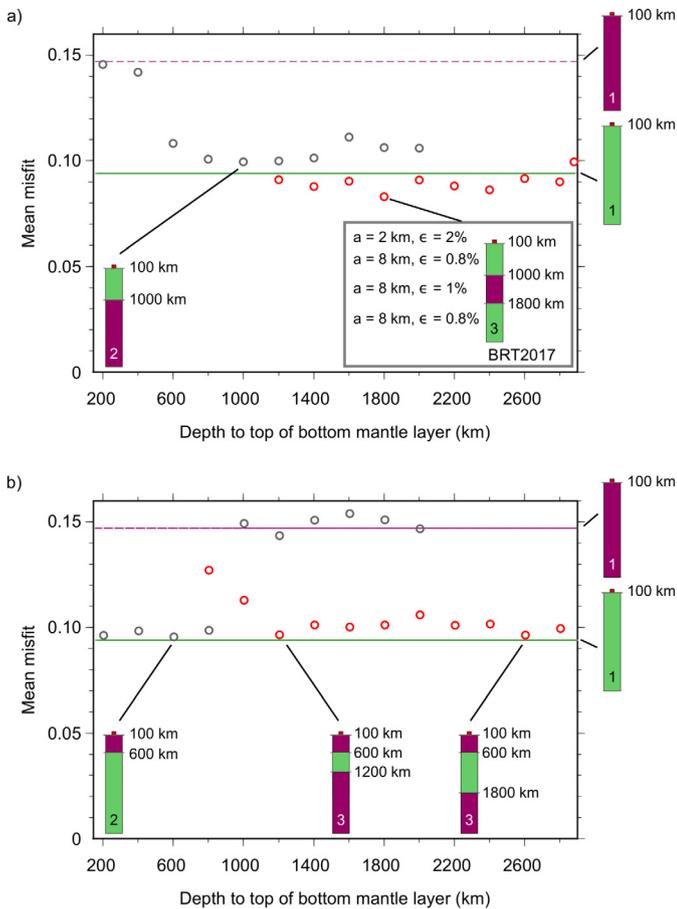
Fig. 9. Variation of mean misfit for models with two layers (grey circles) and three layers (red circles) of heterogeneity in the mantle, with depth of the top of the deepest layer (layer 2 and 3 respectively). The models with the smallest misfit in each category are indicated. For comparison, the misfit for two single layer mantle models are shown: $\epsilon = 1\%$ (purple dash line) and $\epsilon = 0.8\%$ (green line). a) Lithospheric $\epsilon = 2\%$ and correlation length = 2 km, and the mantle correlation length is 8 km for both model types for all layers. The two layer mantle models (Model type 2A) have $\epsilon = 0.8\%$ in layer 1 and $\epsilon = 1\%$ in layer 2, with the depth of layer 2 varying from 200 km to 2000 km. The three layer mantle models (Model type 3A) have $\epsilon = 0.8\%$ in layer 1, $\epsilon = 1\%$ in layer 2 and $\epsilon = 0.8\%$ in layer 3, with the depth of layer 3 varying from 1200 km to 2880 km. Best-fitting layered heterogeneity model, BRT2017 is shown in the box. b) Lithospheric $\epsilon = 2\%$ and correlation length = 2 km, and the mantle correlation length is 8 km for both model types for all layers. The two layer mantle models (Model type 2B) have $\epsilon = 1\%$ in layer 1 and $\epsilon = 0.8\%$ in layer 2, with the depth of layer 2 varying from 200 km to 2000 km. The three layer mantle models (Model type 3B) have $\epsilon = 1\%$ in layer 1, $\epsilon = 0.8\%$ in layer 2 and $\epsilon = 1\%$ in layer 3, with the depth of layer 3 varying from 800 km to 2800 km. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

We analyse the misfit for Model type 2A and 2B (Fig. 9). For type 2A, we find that by increasing the depth of the top of the second layer (in the mantle) we observe a decrease in mean misfit to a minimum of $\overline{\Delta} = 0.0995$ at 1000 km (Fig. 9a). When this boundary is deeper than 1000 km, the misfit increases by up to 15%. In contrast for type 2B, the mean misfit is very similar to the misfit of the single layer model with $\epsilon_M = 0.8\%$ ($\overline{\Delta} = 0.0939$) when the top of the second layer is no deeper than 800 km (Fig. 9b). When the boundary is 1000 km or deeper, the misfit is close to the misfit of the single layer model with $\epsilon_M = 1\%$. The minimum mean misfit for type 2B models is $\overline{\Delta} = 0.0955$ with the boundary at a depth of 600 km.

From introducing a second mantle layer, we are unable to find a model that has a better mean misfit than the single layer mantle model with $\epsilon_M = 0.8\%$. However, exploring the two model types 2A and 2B there is a suggestion that we require $\epsilon_M = 0.8\%$ at

depths of 600 km to 1000 km, though the velocity variation for the rest of the mantle is unclear. In the next section we seek to improve the model misfit by applying further radial variation and including a third and final mantle layer in our models.

4.3.2. Modelling with depth varying mantle heterogeneity II – three layer case

We explore a three mantle layer case using the type 2A and 2B models that have the smallest misfits as found in Section 4.2. The two layer mantle models show increases in misfit at depths larger than about 1000 km and 600 km (types 2A and 2B respectively) (Fig. 9). To limit the model space we use the information gained from the single layer and two layer models and thus only consider models with combinations of velocity variations equal to 0.8% or 1% in the three layers. Specifically, the two types of models we analyse are:

- Model type 3A – increase velocity variation from $\epsilon_M = 0.8\%$ (mantle layer 1 depths 100–1000 km) to 1% (mantle layer 2 from 1000 km) then decrease back to 0.8% (mantle layer 3). The depth of the boundary between mantle layer 2 and 3 is varied from 1200 km to 2800 km (every 200 km);
- Model type 3B – decrease velocity variation from $\epsilon_M = 1\%$ (mantle layer 1 depths 100–600 km) to 0.8% (mantle layer 2 from 600 km) then decrease back to 1% (mantle layer 3). The depth of the boundary between mantle layer 2 and 3 is varied from 800 km to 2800 km (every 200 km).

All the type 3A models have smaller misfits than all one and two layer mantle models (Fig. 9a). In contrast, the type 3B models have larger misfits than the best fitting one layer mantle model (with $\epsilon_M = 0.8\%$) (Fig. 9b). The best fitting model overall, BRT2017 (Fig. 9a) has an increase of velocity variation ($\epsilon = 1\%$) between depths of 1000–1800 km and a decrease of velocity heterogeneity at 1800 km depth ($\epsilon = 0.8\%$) (Fig. 9a for full parameters). The misfit of this model: $\overline{\Delta} = 0.0830$, is a decrease of 12% with respect to the single layer model misfit ($\overline{\Delta} = 0.0939$).

5. Discussion

5.1. PP precursor observations

From the global stacks shown in Figs. 3 and 4, it is clear that analysing incoherent PP precursors provides important insights into the heterogeneity structure of the Earth's upper and upper lower mantle that have not been observed previously. We have found that PP precursors are distinct from the P coda (mainly caused by heterogeneities in the lithosphere) suggesting the PP precursory energy is predominantly created by scattering at small-scale heterogeneities in the mantle. Compared to other scattering probes, PP precursors allow better study of heterogeneities in upper and mid-mantle due to the turning depths of PP. With the observations of scattered energy at different distances, it is now possible to quantify radial heterogeneity structure in more detail than with previous probes (Shearer, 2007).

The striking similarity in the distance stacks between the Pacific and the Atlantic suggests that the heterogeneities in the sampled depth range of the mantle do not vary significantly despite different tectonic settings, subduction histories and observations of mid-mantle heterogeneities (e.g. Bentham and Rost, 2014; Kaneshima and Helffrich, 1998; Vidale and Garcia-Gonzalez, 1988). The regional similarity is in contradiction to variations of lower mantle heterogeneities with stronger than average heterogeneities observed beneath central Africa (Hedlin and Shearer, 2000), North America (Hedlin and Shearer, 2000), north of Tonga (Vidale and Hedlin, 1998) and the southwest Pacific (Margerin and Nolet,

2003b). The difference of heterogeneity structure at the CMB compared to the mid-mantle could indicate changes in viscosity, convective vigour or configuration of mantle flow at these depths. Smaller scale variations in heterogeneity structure might exist, but we are limiting our analysis to hemispherical differences to retain enough traces per distance bin to ensure a smooth stack.

5.2. Monte Carlo modelling of heterogeneity

As described in Section 4.1, none of the three existing heterogeneity models considered (H1997, ES2001 and SE2004) match the PP precursory energy. Model SE2004, constructed to explain P & P_{diff} codas (Shearer and Earle, 2004), produces no PP arrivals and associated precursory energy despite covering the same time window and method used in this study. Models ES2001 and H1997 provide a better fit to data analysed in this study (Fig. 7) and have been useful in constructing our model space (Fig. 8). Recently a correction to the H1997 model has been presented reducing the velocity variation to 0.1% (Mancinelli and Shearer, 2013). However, applying the revised model by Mancinelli and Shearer (2013) degrades the model fit producing one of the highest misfits in our single layer mantle model space (Fig. 8).

We observe a difference in preferred heterogeneity correlation length between lithosphere and mantle (Fig. 8 and 9). As the data are filtered between 0.5 and 2.5 Hz, scattered arrivals can be generated by lithospheric structures with scalelengths on the order of 2–12 km. The best fitting lithospheric correlation length of 2 km falls within this range of resolvable wavelengths. Similarly for the mantle, the frequency range used will be sensitive to structures with scalelengths of ~4–23 km. The range of best whole mantle models (for different lithospheric heterogeneity) of 6–16 km lie in the middle of this group with the optimum correlation length of 8 km in agreement with expected scalelengths for subducted crust. Though the preferred correlation length of 8 km is not unexpected and the range of resolvable scalelengths is limited by the studied frequency range, we feel this value is well constrained due to the grid search approach (Fig. 8) and is in agreement with previous work (e.g. Hedlin et al., 1997). Since there is a trade-off between mantle correlation length and velocity variation however, our modelling result could be non-unique.

We have shown that varying heterogeneity parameters in the lithosphere has little effect on the optimum velocity variation in the mantle found through our modelling (ϵ_M: 0.8 to 1%) (Fig. 8). This demonstrates that lithospheric scattering is not the main scattering source for PP precursory wavefield. Nonetheless the heterogeneous lithosphere does contribute to the precursory wavefield mainly through the interaction of the PP wavefield with the lithosphere structure at the PP mid-path reflection point. With our assumed lithospheric thickness of 100 km we may overestimate the contribution from the lithosphere to our wavefield especially since many PP reflection points are in oceanic areas. In some preliminary tests we did find that changing the thickness of the lithosphere does have a small effect on the width and shape of PP within a narrow time window around PP, but shows a negligible effect on the overall PP precursor time window we model.

We show the model fit is improved by varying RMS velocity variation with depth within the mantle (Fig. 9). RMS velocity variations of 0.8% and 1% are selected based on the single mantle layer models that match the data consistently for a wide range of distances (Section 4.2). Due to our systematic modelling approach, we were able to investigate radial changes in parameters and found depth varying heterogeneity does significantly reduce the overall mean model misfit by 12% for our preferred model BRT2017. Depth varying mantle heterogeneity has been explored previously using P coda energy (Shearer and Earle, 2004) but stronger scattering in the upper mantle ($\epsilon = 3\%$, up to 600 km depth) and weaker
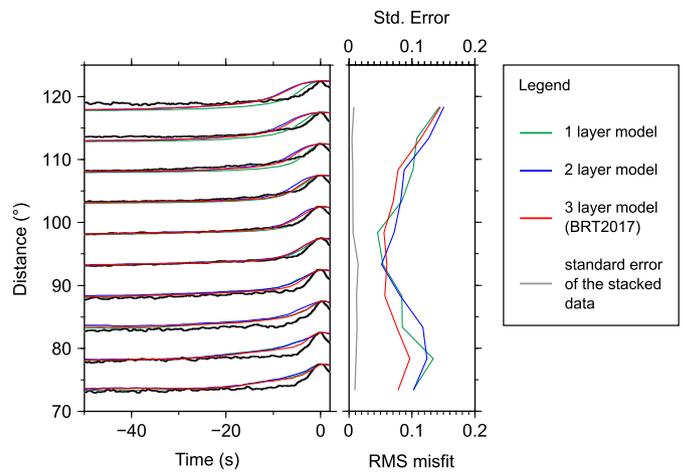
Fig. 10. Comparison of seismograms (left) and RMS misfit (right) for the best fitting 1 (green), 2 (blue) and 3 (red) layer mantle heterogeneity models with respect to global stacked data (black) for every distance bin. The standard error (grey) of the observed global stack data (black) is of the order of 10^{-2} and 5–10 times smaller than the RMS model misfits. The misfits vary with distance with the smallest misfits calculated for distances 90–100°. The best overall model with the lowest misfit, BRT2017, has three mantle layers (red) and improves the misfit for distances, especially 70–90°. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

scattering ($\epsilon = 0.5\%$) in the rest of the mantle was the preferred model (SE2004) in contrast to the BRT2017 model. As explained in Section 4.1, we explored the SE2004 model but found very strong scattering in the P coda, masking the PP arrival and PP precursors and thus variations of this model were excluded from further investigating.

Though model BRT2017 (with a 1% layer from 1000 to 1800 km, see Table S1) has the lowest misfit, the model fits less well for shorter distances as it over-estimates the precursor energy for distances 70°–85° (up to 20 s before PP) (Fig. 10). Improvements to the shape of the precursor energy for BRT2017 for distances 110°–120° might be made by decreasing the velocity variation in the lower mantle (deeper than 1800 km) or by exploring depth variations in correlation length in the mantle. This would greatly increase the available parameter space leading to a large number of additional models to be simulated. We feel this is impractical given the long simulation runtimes for the approaches of scattered seismic wavefield modelling that are currently available.

5.3. Geodynamical significance

Many studies link the heterogeneity detected in the mantle from studying scattered seismic waves to subducted oceanic crust (e.g. Bentham and Rost, 2014; Castle and Creager, 1999; Kaneshima and Helffrich, 1998; Rost et al., 2008; Kito et al., 2008). Subducted crust is typically 8–10 km thick, a scale that correlates well with the heterogeneity correlation length for the better fitting models found in this study. Though subducted crust is likely deformed through stretching and folding mechanism during the mechanical mixing (Albarède, 2005; Olson, 1981), this process is likely to be very slow (due to the high mantle viscosity) and as such the crustal heterogeneity can keep its dominant scale length of ~8 km for more than 2.5 Ga (Stixrude and Lithgow-Bertelloni, 2012). With these timescales in mind, we expect that heterogeneity from most of Earth's 3 Ga subduction history (Tang et al., 2016) can still be detected seismically in the mantle in the present day. Furthermore, numerical models show that mantle convection is able to distribute the crustal heterogeneity throughout the mantle away from its source at subduction zones (Li and McNamara, 2013).

Mancinelli and Shearer (2013) discuss other causes of heterogeneity including primordial material which has been mixed into younger material (e.g. Albarède, 2005) as demonstrated by geodynamic models (Manga, 1996). However, this scenario seems more likely at the core mantle boundary where primordial material is thought to survive the age of Earth and it would be difficult to amass primordial material in the mid-mantle. Furthermore, differentiating between these competing heterogeneity sources would be challenging, particularly since velocity variations expected from primordial material are unknown.

By assuming the mantle heterogeneity is subducted basalt, the changes in seismic velocity variation with depth may provide a way of observing the seismic velocity of basaltic compositions in situ for comparison to computed velocities (e.g. Stixrude and Lithgow-Bertelloni, 2012). Using the HeFESTo model (Stixrude and Lithgow-Bertelloni, 2011), the contrast in the velocity of basalt/eclogite to the velocity of a harzburgite/pyrolite mechanical mixture is thought to greatly vary in the upper mantle and transition zone (Stixrude and Lithgow-Bertelloni, 2012). In addition, the ratio of these velocities (eclogite/mantle) increases at ∼1000 km, in agreement with our best-fitting model. However, to fully compare observed velocity variations to computed velocities, advanced modelling using a wide range of seismic probes and improved understanding of mantle & slab materials at these depths, are required.

An increase in velocity variation results in a rise in scattered energy for a random medium. Similarly an increase of the amount of heterogeneity (i.e. an increase in the number of scatterers per volume) will lead to higher scattered energy. Therefore the mid-mantle (1000–1800 km) could contain more crustal remnants than the upper and lower mantle. Tomographic models commonly resolve slab-like fast velocity features from the surface to depths of 1000 km as well as a few lateral fast velocity features at depths of 1000–1500 km (e.g. Becker and Boschi, 2002; Ritsema et al., 2011; Sigloch and Mihalynuk, 2013), suggesting subducted material can reach these depths and potentially remain there. The scattering power we have deduced in this study could be additional evidence that slabs and associated basaltic crust can remain or be stalled at depths of around 1000 km.

To stall slabs in the mid-mantle, there needs to be a change in properties at this depth and to date, no first order discontinuities have been detected. However, a radial increase in viscosity of the mantle can delay descending slabs (e.g. Torii and Yoshioka, 2007) and many mantle viscosity models show an increase in viscosity in the lower mantle (e.g. Rudolph et al., 2015). In fact, mantle viscosity structure is difficult to resolve and the detailed viscosity profile is still debated. Some models suggest that viscosity in the mid- and lower mantle increases steadily in the mid-mantle to depths of 2000 km (Mitrovica and Forte, 2004) but alternative models show sharp increases below the transition zone (10^{22}–10^{23} Pa s), or at 2000–2200 km (to 10^{23}–10^{24} Pa s) (Forte and Mitrovica, 2001; Steinberger and Calderwood, 2006). The viscosity might also decrease at 660 km, followed by an increase at 1000 km (Cserepes et al., 2000). Recent mineral physical studies have suggested that the strength of ferropericlase can increase by 2.3 orders of magnitude in high-strain regions for depths associated with the iron spin transition and will broaden and stagnate subducted slabs (Marquardt and Miyagi, 2015). Additionally, a recent study also suggests that undulations in mid-mantle viscosity (at 1300 and 2000 km) can be explained by the creep mechanism of ferropericlase undergoing a spin transition (Justo et al., 2015) proposing a different mechanism for viscosity changes at similar depths.

Regardless of where the jump in viscosity occurs, such an increase will also increase the lateral length scale of mantle flow compared to the upper mantle (Forte and Mitrovica, 2001) leading to a change of the radial distribution of heterogeneity between the mid-mantle (e.g. 1000 km) and the lower mantle (depths >2000 km). The radially varying small-scale mantle heterogeneity detected by PP precursors in this study is potentially a unique way to observe the effect of depth varying viscosity structure that has been proposed using a wide variety of different approaches (Forte and Mitrovica, 2001; Marquardt and Miyagi, 2015; Justo et al., 2015; Rudolph et al., 2015). Using the information contained in the short-period scattered seismic wavefield might deliver important complementary information to large-scale tomographic images allowing better insight into the dynamics and evolution of our planet.

Acknowledgements

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.epsl.2017.05.027.

References

Albarède, F., 2005. The survival of mantle geochemical heterogeneities. In: Earth's Deep Mantle: Structure, Composition, and Evolution, pp. 27–46.

Bataille, K., Flatté, S.M., 1988. Inhomogeneities near the core–mantle boundary inferred from short-period scattered PKP waves recorded at the global digital seismograph network. J. Geophys. Res., Solid Earth 93 (B12), 15057–15064. http://dx.doi.org/10.1029/JB093iB12p15057.

Becker, T.W., Boschi, L., 2002. A comparison of tomographic and geodynamic mantle models. Geochem. Geophys. Geosyst. 3 (1). http://dx.doi.org/10.1029/2001GC000168.

Bentham, H.L.M., Rost, S., 2014. Scattering beneath Western Pacific subduction zones: evidence for oceanic crust in the mid-mantle. Geophys. J. Int. 197 (3), 1627–1641. http://dx.doi.org/10.1093/gji/ggu043.

Bhattacharyya, J., Shearer, P., Masters, G., 1993. Inner core attenuation from short-period PKP (BC) versus PKP (DF) waveforms. Geophys. J. Int. 114 (1), 1–11. http://dx.doi.org/10.1111/j.1365-246X.1993.tb01461.x.

Castle, J.C., Creager, K.C., 1999. A steeply dipping discontinuity in the lower mantle beneath Izu–Bonin. J. Geophys. Res., Solid Earth 104 (B4), 7279–7292. http://dx.doi.org/10.1029/1999JB900011.

Cleary, J.R., Haddon, R.A.W., 1972. Seismic wave scattering near the core–mantle boundary: a new interpretation of precursors to PKP. Nature 240, 549–551. http://dx.doi.org/10.1038/240549a0.

Cserepes, L., Yuen, D.A., Schroeder, B.A., 2000. Effect of the mid-mantle viscosity and phase-transition structure on 3D mantle convection. Phys. Earth Planet. Inter. 118, 135–148. http://dx.doi.org/10.1016/S0031-9201(99)00140-5.

Domeier, M., Doubrovine, P.V., Torsvik, T.H., Spakman, W., Bull, A.L., 2016. Global correlation of lower mantle structure and past subduction. Geophys. Res. Lett. 43, 2016GL068827. http://dx.doi.org/10.1002/2016GL068827.

Doornbos, D.J., 1978. On seismic-wave scattering by a rough core–mantle boundary. Geophys. J. Int. 53 (3), 643–662. http://dx.doi.org/10.1111/j.1365-246X.1978.tb03765.x.

Dziewonski, A.M., Anderson, D.L., 1981. Preliminary reference Earth model. Phys. Earth Planet. Inter. 25 (4), 297–356. http://dx.doi.org/10.1016/0031-9201(81)90046-7.

Earle, P.S., Shearer, P.M., 2001. Distribution of fine-scale mantle heterogeneity from observations of P-diff coda. Bull. Seismol. Soc. Am. 91 (6), 1875–1881. http://dx.doi.org/10.1785/0120000285.

Forte, A.M., Mitrovica, J.X., 2001. Deep-mantle high-viscosity flow and thermochemical structure inferred from seismic and geodynamic data. Nature 410 (6832), 1049–1056. http://dx.doi.org/10.1038/35074000.

Hedlin, M.A.H., Shearer, P.M., 2000. An analysis of large-scale variations in small-scale mantle heterogeneity using global seismographic network recordings of precursors to PKP. J. Geophys. Res., Solid Earth 105 (B6), 13655–13673. http://dx.doi.org/10.1029/2000JB900019.

Hedlin, M.A.H., Shearer, P.M., Earle, P.S., 1997. Seismic evidence for small-scale heterogeneity throughout the Earth's mantle. Nature 387, 145–150. http://dx.doi.org/10.1038/387145a0.

Justo, J.F., Morra, G., Yuen, D.A., 2015. Viscosity undulations in the lower mantle: the dynamical role of iron spin transition. Earth Planet. Sci. Lett. 421, 20–26. http://dx.doi.org/10.1016/j.epsl.2015.03.013.

Kaneshima, S., Helffrich, G., 1998. Detection of lower mantle scatterers northeast of the Marianna subduction zone using short-period array data. J. Geophys. Res., Solid Earth 103 (B3), 4825–4838. http://dx.doi.org/10.1029/97JB02565.

Kennett, B.L.N., Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification. Geophys. J. Int. 105 (2), 429–465. http://dx.doi.org/10.1111/j.1365-246X.1991.tb06724.x.

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

Li, M., McNamara, A.L., 2013. The difficulty for subducted oceanic crust to accumulate at the Earth's core–mantle boundary. J. Geophys. Res., Solid Earth 118 (B4). http://dx.doi.org/10.1002/jgrb.50156.

Mancinelli, N.J., Shearer, P.M., 2013. Reconciling discrepancies among estimates of small-scale mantle heterogeneity from PKP precursors. Geophys. J. Int. 195, 1721–1729. http://dx.doi.org/10.1093/gji/ggt319.

Manga, M., 1996. Mixing of heterogeneities in the mantle: effect of viscosity differences. Geophys. Res. Lett. 23, 403–406. http://dx.doi.org/10.1029/96GL00242.

Margerin, L., Nolet, G., 2003a. Multiple scattering of high-frequency seismic waves in the deep Earth: PKP precursor analysis and inversion for mantle granularity. J. Geophys. Res., Solid Earth 108 (B11). http://dx.doi.org/10.1029/2003JB002455.

Margerin, L., Nolet, G., 2003b. Multiple scattering of high-frequency seismic waves in the deep Earth: and numerical examples. J. Geophys. Res., Solid Earth 108 (B5). http://dx.doi.org/10.1029/2002JB001974.

Marquardt, H., Miyagi, L., 2015. Slab stagnation in the shallow lower mantle linked to an increase in mantle viscosity. Nat. Geosci.. http://dx.doi.org/10.1038/NGEO2393.

Mitrovica, J.X., Forte, A.M., 2004. A new inference of mantle viscosity based upon joint inversion of convection and glacial isostatic adjustment data. Earth Planet. Sci. Lett. 225, 177–189. http://dx.doi.org/10.1016/j.epsl.2004.06.005.

Niu, F., Wen, L., 2001. Hemispherical variations in seismic velocity at the top of the Earth's inner core. Nature 410, 1081–1084. http://dx.doi.org/10.1038/35074073.

Olson, P., 1981. Mantle convection with spherical effects. J. Geophys. Res., Solid Earth 86 (B6), 4881–4890. http://dx.doi.org/10.1029/JB086iB06p04881.

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. Geophys. J. Int. 184, 1223–1236. http://dx.doi.org/10.1111/j.1365-246X.2010.04884.x.

Rost, S., Thorne, M.S., Garnero, E.J., 2006. Imaging global seismic phase arrivals by stacking array processed short-period data. Seismol. Res. Lett. 77 (6), 697–707. http://dx.doi.org/10.1785/gssrl.77.6.697.

Rost, S., Garnero, E.J., Williams, Q., 2008. Seismic array detection of subducted oceanic crust in the lower mantle. J. Geophys. Res., Solid Earth 113 (B6). http://dx.doi.org/10.1785/gssrl.77.6.697.

Rudolph, M., Lekic, V., Lithgow-Bertelloni, C., 2015. Viscosity jump in the Earth's mid mantle. Science 360 (6266), 1349–1352. http://dx.doi.org/10.1126/science.aad1929.

Sato, H., Fehler, M., 1998. Seismic Wave Propagation and Scattering in the Heterogeneous Earth. AIP Press/Springer, New York.

Shearer, P.M., 2007. Seismic scattering in the deep Earth. In: Treatise on Geophysics, vol. 1, pp. 695–730.

Shearer, P.M., Earle, P.S., 2004. The global short-period wavefield modelled with a Monte Carlo seismic phonon method. Geophys. J. Int. 158, 1103–1117. http://dx.doi.org/10.1111/j.1365-246X.2004.02378.x.

Shearer, P.M., Flanagan, M.P., 1999. Seismic velocity and density jumps across the 410- and 660-kilometer discontinuities. Science 285, 1545–1548. http://dx.doi.org/10.1126/science.285.5433.1545.

Sigloch, K., Mihalynuk, M.G., 2013. Intra-oceanic subduction shaped the assembly of Cordilleran North America. Nature 496, 50–56. http://dx.doi.org/10.1038/nature12019.

Steinberger, B., Calderwood, A.R., 2006. Models of large-scale viscous flow in the Earth's mantle with constraints from mineral physics and surface observations. Geophys. J. Int. 167, 1461–1481. http://dx.doi.org/10.1111/j.1365-246X.2006.03131.x.

Stixrude, L., Lithgow-Bertelloni, C., 2011. Thermodynamics of mantle minerals-II. Phase equilibria. Geophys. J. Int. 184 (3), 1180–1213. http://dx.doi.org/10.1111/j.1365-246X.2010.04890.x.

Stixrude, L., Lithgow-Bertelloni, C., 2012. Geophysics of chemical heterogeneity in the mantle. Annu. Rev. Earth Planet. Sci. 40, 569–595. http://dx.doi.org/10.1146/annurev.earth.36.031207.124244.

Tang, M., Chen, K., Rudnick, R.L., 2016. Archean upper crust transition from mafic to felsic marks the onset of plate tectonics. Science 351, 372–375. http://dx.doi.org/10.1126/science.aad5513.

Tono, Y., Yomogida, K., 1996. Complex scattering at the core–mantle boundary observed in short-period diffracted P-waves. J. Phys. Earth 44 (6), 729–744. http://dx.doi.org/10.4294/jpe1952.44.729.

Torii, Y., Yoshioka, S., 2007. Physical conditions producing slab stagnation: constraints of the Clapeyron slope, mantle viscosity, trench retreat, and dip angles. Tectonophysics 445, 200–209. http://dx.doi.org/10.1016/j.tecto.2007.08.003.

Van der Hilst, R.D., Widiyantoro, S., Engdahl, E.R., 1997. Evidence for deep mantle circulation from global tomography. Nature 386, 578–584. http://dx.doi.org/10.1038/386578a0.

Vidale, J.E., Garcia-Gonzalez, D., 1988. Seismic observation of a high-velocity slab 1200–1600 km in depth. Geophys. Res. Lett. 15 (4), 369–372. http://dx.doi.org/10.1029/GL015i004p00369.

Vidale, J.E., Hedlin, M.A.H., 1998. Evidence for partial melt at the core–mantle boundary north of Tonga from the strong scattering of seismic waves. Nature 391, 682–685. http://dx.doi.org/10.1038/35601.

Warren, L.M., Shearer, P.M., 2000. Investigating the frequency dependence of mantle Q by stacking P and PP spectra. J. Geophys. Res., Solid Earth 105 (B11), 25391–25402. http://dx.doi.org/10.1029/2000JB900283.

Wright, C., 1972. Array studies of seismic waves arriving between P and PP in the distance range 90° to 115°. Bull. Seismol. Soc. Am. 62, 385–400.