Seismic evidence for a plume-modified oceanic lithosphere–asthenosphere system beneath Cape Verde

Xin Liu1,2 and Dapeng Zhao3

1Frontiers Science Center for Deep Ocean Multispheres and Earth System, Key Lab of Submarine Geosciences and Prospecting Techniques, MOE and College of Marine Geosciences, Ocean University of China, Qingdao 266100, China. E-mail: liuxin@ouc.edu.cn
2Laboratory for Marine Mineral Resources, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266100, China
3Department of Geophysics, Graduate School of Science, Tohoku University, Sendai 980–8578, Japan. E-mail: zhao@tohoku.ac.jp

SUMMARY
We determine a new 3-D shear wave velocity ($V_s$) model down to 400 km depth beneath the Cape Verde hotspot that is far from plate boundaries. This $V_s$ model is obtained by using a new method of jointly inverting $P$- and $S$-wave receiver functions, Rayleigh-wave phase-velocity data and $S$-wave arrival times of teleseismic events. Two $V_s$ discontinuities at $\sim$15 and $\sim$60 km depths are revealed beneath volcanic islands, which are interpreted as the Moho discontinuity and the Gutenberg (G) discontinuity. Between the north and south islands, obvious high-$V_s$ anomalies exist in the uppermost mantle down to a depth of $\sim$100–150 km beneath the Atlantic Ocean, whereas obvious low-$V_s$ anomalies exist in the uppermost mantle beneath the volcanic islands including the active Fogo volcano. These low-$V_s$ anomalies merge into a significant column-like low-$V_s$ zone at depths of $\sim$150–400 km beneath the Cape Verde swell. We propose that these features in the upper mantle reflect a plume-modified oceanic lithosphere–asthenosphere system beneath the Cape Verde hotspot.

Key words: Joint inversion; Body waves; Seismic tomography; Surface waves and free oscillations; Dynamics of lithosphere and mantle; Hotspots.

1 INTRODUCTION
Oceanic lithosphere is born at the mid-ocean ridge and normally cools, becoming dense and thick with age towards subduction. Before its subduction, oceanic lithosphere may be rejuvenated if it passes over a plume rising from the deep mantle (Detrick & Crough 1978; Li et al. 2004). It is theoretically expected that a mantle plume can significantly alter the thermal structure, volatile and melt contents of the overlying oceanic lithosphere–asthenosphere (OLA) system (e.g. Nolet et al. 2006; Hirschmann 2010; Karato 2012). However, details about how they interact with each other remain poorly constrained, because there are few lines of evidence for the deep processes from high-resolution geophysical images without the influence of plate boundary processes.

The Cape Verde hotspot is a typical intraplate hotspot, about 2000 km away from the nearest plate boundary (Fig. 1). This hotspot rests on a broad bathymetric swell lying on a mature (110–150 Ma) oceanic lithosphere beneath the Atlantic Ocean, with a high geoid anomaly ($\sim$8 m; Monnereau & Cazenave 1990) and high surface heat flow ($\sim$61 mW m−2; Courtney & White 1986). It is composed of a group of Late Cenozoic oceanic islands and seamounts with long-term volcanism (Holm et al. 2008). The oldest exposed volcanic rocks ($\sim$28 Ma) are generally located in the eastern islands with a thicker oceanic crust, whereas recent volcanic eruptions and seismic activity mainly occurred in and around the western islands and seamounts with a thinner oceanic crust (e.g. Lodge & Helffrich 2006; Holm et al. 2008; Grevemeyer et al. 2010; Vales et al. 2014; Leva et al. 2019; Samrock et al. 2019; Fig. 1). The last known volcanic eruption occurred at the very active Pico do Fogo volcano during November 2014 to February 2015, after only $\sim$20 yr of quiescence (González et al. 2015; Mata et al. 2017).

The origin of the Cape Verde hotspot is generally attributed to a deep mantle plume (Zhao et al. 2013; Liu & Zhao 2014; French & Romanowicz 2015; Saki et al. 2015; Lodhia et al. 2018; Carvalho et al. 2019), rather than edge-driven convection (King & Ritsema 2000). In this hotspot region, previous studies found high 3He/4He ratios in carbonatites and alkaline silicate rocks (Christensen et al. 2001; Doucelance et al. 2003; Mata et al. 2010; Mouriño et al. 2012a), plume-like features in high-resolution seismic images such as a distinct low-$V_s$ zone in the lower mantle down to the core–mantle boundary (French & Romanowicz 2015), significantly depressed 410-km and shallower 660-km discontinuities (Saki et al. 2015) and a clear column-like low-velocity anomaly in the upper mantle (Liu & Zhao 2014). The potential deep mantle plume is expected to influence the overlying mature OLA system beneath Cape Verde. However, such interaction processes have not been clearly...
Figure 1. Tectonic settings of the Cape Verde hotspot and distribution of seismic stations used in this study (after Liu & Zhao 2014). (a) Distribution of seismic stations. Five seismic stations of the YW network (green squares), 39 seismic stations of the CVPLUME network (blue diamonds) and a GSN permanent seismic station (blue circle). The red triangle denotes the active Fogo volcano. Red numbers show volcanic evolution periods in Ma (Holm et al. 2008; Samrock et al. 2019). Green numbers show the Moho depths (in km) beneath the individual islands (Lodge & Helffrich 2006). (b) Map of the central Atlantic Ocean. The red box denotes the present study region. The white dashed lines denote plate boundaries. The seafloor ages (colours) are from a global age model (Müller et al. 2008), whose scale is shown on the right.

caught by previous local seismic images (e.g. Lodge & Helffrich 2006; Helffrich et al. 2010; Vinnik et al. 2012; Liu & Zhao 2014; Carvalho et al. 2019).

To clarify interaction processes between a mantle plume and a mature OLA system, here we use a new method of joint seismic inversion to image the upper mantle $V_s$ structure beneath the Cape Verde hotspot with multiple seismic data analysis techniques. Our results reveal unprecedented structural details of a plume-modified OLA system, a discovery that sheds new light on the plume dynamics and nature of the oceanic lithosphere–asthenosphere boundary (LAB).

2 DATA AND METHODS

2.1 Seismic stations

We use 45 seismic stations installed on the Cape Verde archipelago (Fig. 1), including 5 temporary broad-band stations of the YW network deployed from July 2002 to September 2004 (Lodge & Helffrich 2006; Helffrich et al. 2010), 39 temporary broad-band stations of the CVPLUME network deployed from November 2007 to September 2008 (Vinnik et al. 2012) and 1 GSN (global seismic network) permanent station (SACV). We collect three sets of data from teleseismic waveforms recorded at these stations. Hypocentral parameters of the teleseismic events used in this study are obtained from Bulletins of the International Seismological Center (ISC; Fig. 2).

2.2 $P$- and $S$-wave receiver function data

The first data set contains $P$-to-$S$ and $S$-to-$P$ receiver functions (PRFs and SRFs) newly measured from seismograms of teleseismic events (Figs 2a and b). For the PRF estimation, we select teleseismic events ($M \geq 5.8$) with epicentral distances of $\sim 30^\circ$–$90^\circ$, whereas for the SRF estimation, teleseismic events ($M \geq 5.8$) with epicentral distances of $\sim 55^\circ$–$90^\circ$ are selected, slightly longer than that ($55^\circ$–$85^\circ$) proposed by Yuan et al. (2006). To reduce the complexity of estimated receiver functions, we filter the vertical- and radial-component seismograms in a frequency band of 0.02–0.50 Hz, after correcting for the instrument response. For the PRF estimation, each filtered seismogram is truncated using a time window from 60 s before to 100 s after the theoretical direct $P$-wave arrival. For the SRF estimation, we use a time window from 100 s before to 60 s after the theoretical direct $S$-wave arrival. Then, for the PRF estimation, we calculate the deconvolution of the vertical component from the radial component ($R_{PRF}$) and the deconvolution of the vertical component from itself ($Z_{PRF}$) (Burdick & Langston 1977; Vinnik 1977). For the SRF estimation, the deconvolution of the radial component from the vertical component ($Z_{SRF}$) and the deconvolution of the radial component from itself ($R_{SRF}$) are calculated (Farra & Vinnik 2000; Zhou et al. 2000). These processes are performed in time domain with different Gaussian parameters of 0.5, 0.7 and 1.0 (Liggins & Ammon 1999). Then we calculate apparent incidence angles (AIAs) of $P$-wave, $\theta_P = \arctan[R_{PRF}(t=0)/Z_{PRF}(t=0)]$, and $S$-wave, $\theta_S = -\arctan[Z_{SRF}(t=0)/R_{SRF}(t=0)]$, following the method of Svenningsen & Jacobsen (2007). With the AIAs, the vertical- and radial-component seismograms can be rotated into the LQ system around $\theta_P$ for the PRF estimation or $\theta_S$ for the SRF estimation (e.g. Kind et al. 2012; Hu et al. 2015). In practice, we remove the data with unusual AIAs. In addition, to reduce the complexity of receiver functions, we rotate the vertical- and radial-component seismograms around various AIAs in a range of $\theta_P \sim 10^\circ$ to $10^\circ$, where as $\theta_S \sim 10^\circ$ to $10^\circ$. Then, we calculate the corresponding PRFs by deconvolution of the L component from the Q component or SRFs by deconvolution of the Q component from the L component. The obtained PRFs or SRFs are then stacked. The stacking is linear and unweighted. Then we reverse the amplitude of the stacked SRF, and remove the stacked traces with unusually large amplitudes.

As a result, we obtain 1924 PRFs from 186 teleseismic events and 493 SRFs from 99 teleseismic events (Figs 3a–c). These PRFs and SRFs are moveout-corrected to a reference slowness of 0.06 and 0.10 s km$^{-1}$, respectively, following the method of Duiker & Sheehan (1997). The corrected receiver functions are isolated in time windows with 5 s before to 80 s after the direct $P$-wave arrival and 60 s before to 5 s after the direct $S$-wave arrival. For each seismic station, we stack all the corrected PRF or SRF traces for each of the Gaussian parameters. The stacked PRFs and SRFs at each station are further stacked, respectively, and the results are shown in Figs 3(d) and (e). According to such stacked traces (Figs 3d and e), we check all the moveout-corrected traces and only keep the traces with clear $Pms$ ($P$-to-$S$ conversion at the Moho discontinuity) arrivals, or clear $Smp$ ($S$-to-$P$ conversion at the Moho discontinuity) arrivals similar to those in the stacked traces (Figs 3d and e). As a result, we
Figure 2. Distribution of the teleseismic events finally used in this study. (a) Events (green dots) used to calculate P-to-S receiver functions. (b) Events (blue dots) used to calculate S-to-P receiver functions. (c) Events (red dots) used to calculate S-wave relative traveltime residuals. (d) Events (orange dots) used to calculate fundamental mode Rayleigh-wave phase velocities. The brown dashed lines in (a)–(d) denote plate boundaries.

Figure 3. The PRF and SRF data. (a–e) The initial stacked PRFs and SRFs with the data distribution. (f–j) The final stacked PRFs and SRFs with the data distribution. (a) 186 and (f) 143 events (green dots) used to calculate PRFs. (b) 99 and (g) 44 events (blue dots) used to calculate SRFs. (c and h) Distribution of the piercing points at 15 and 60 km depths, calculated with the optimal 1-D $V_s$ model obtained by this study (Fig. 7a). P15s is P-to-S conversion at a depth of 15 km. S15p is S-to-P conversion at a depth of 15 km. S60p is S-to-P conversion at a depth of 60 km. (d and i) Stacked SRFs and (e and j) stacked PRFs with the data shown in (c) and (h), respectively, for Gaussian parameters of 0.5, 0.7 and 1.0. Smp and Sgp are S-to-P conversions at the Moho and Gutenberg (G) discontinuities, respectively. Pms and Pgs are P-to-S conversions at the Moho and G discontinuities, respectively. Note that the amplitudes of SRFs in (d) are enlarged 2 times relative to those of the PRFs in (e), whereas they are equal in (i) and (j).

obtain 1097 PRFs from 143 teleseismic events and 186 SRFs from 44 teleseismic events (Figs 3f–h). Then we stack the kept traces again. The stacked SRFs at all stations (Fig. 3i) do not show a clear S410p (S-to-P conversion at the 410-km discontinuity) arrival, but the stacked PRFs (Fig. 3j) show visible P410s (P-to-S conversion at the 410-km discontinuity) and P660s (P-to-S conversion at the 660-km discontinuity) arrivals. These stacked PRFs and SRFs at all stations (Figs 3i and j) are used to conduct subsequent inversions for a 1-D $V_s$ model, whereas the stacked PRFs and SRFs at each station (green traces in Supporting Information Figs S1 and S2) are
Figure 4. An example of measuring $S$-wave relative traveltime residuals. (a) The blue square and red star denote the study region and a teleseismic event ($M_{\text{L}} \geq 6.9$, focal depth $\sim 30$ km), respectively. (b) $S$-wave relative traveltime residuals of the teleseismic event shown in (a). The colour squares show the seismic stations used. The red and blue colours denote delayed and early arrivals, respectively. The residual scale is shown at the bottom. (c) Seismograms after alignment with the relative traveltime residuals. The traces are sorted by epicentral distances.

2.3 Teleseismic $S$-wave traveltime data

Our second data set includes newly measured 742 $S$-wave relative traveltime residuals from 69 teleseismic events (Fig. 2c), using the multichannel cross-correlation technique (VanDecar & Crosson 1990; Liu & Zhao 2016). The teleseismic events used are $M_{\text{L}} \geq 5.8$, and their epicentral distances are in a range of $\sim 30^\circ - 90^\circ$ (Fig. 2c). The azimuthal coverage of these events is very good in all directions, and most of the events occurred in plate boundary regions. The $S$-wave relative residuals are measured from the horizontal-component seismograms filtered in a frequency band of 0.02–0.1 Hz, after correcting for the instrument response. An example of this data set is shown in Fig. 4.

2.4 Teleseismic Rayleigh-wave data

Our third data set contains 26 754 amplitude and phase data pairs of fundamental mode Rayleigh waves in a period range of 20–150 s, which are newly measured from the vertical-component seismograms of 238 teleseismic events (focal depths $<100$ km; $M_{\text{L}} \geq 5.8$) with epicentral distances of $\sim 30^\circ - 150^\circ$ (Fig. 2d). An example of the data processing is shown in Fig. 5, following the approach of Liu & Zhao (2016). At first, the Rayleigh waves are isolated by applying adaptive time windows to broad-band vertical-component seismograms after correcting for the instrument response (Fig. 5a). The time windows are selected in a group velocity range of 2.5 to 5.0 km/s. The windowed broad-band seismogram is then filtered into 27 different period bands from 20 to 150 s, with an interval of 5 s using a period-dependent Gaussian bandpass filter (Fig. 5b). Then envelopes of the Gaussian-filtered Rayleigh waves are calculated, and the amplitudes of the obtained envelopes are normalized for each period. We select the optimal group velocity at each period between 20 and 150 s, with a clear and large enough normalized envelope amplitude peak ($>0.7$) (Fig. 5c). With the obtained optimal group velocity at each period, we re-isolate the Gaussian-filtered Rayleigh wave with a time window of 450 s before and 450 s after the group arrival time. Then using the method of Forsyth & Li (2005), we determine amplitudes and phases of the re-isolated Gaussian-filtered Rayleigh waves by Fourier analysis. The obtained amplitudes for each event are normalized by the root-mean-square (RMS) amplitude of each event. The phase data are corrected with a reference station phase for each event to reduce the effect of hypocentral uncertainties of the teleseismic events. To obtain a robust result, we adopt only those events that were recorded at more than four seismic stations in the study region at each period. As a result, the number of amplitude and phase data pairs measured at each period ranges from 828 to 3678.

2.5 Joint inversion method

To better constrain the 3-D $V_s$ structure beneath the Cape Verde hotspot, in this study we adopt a new joint inversion method. First, we calculate the 1-D fundamental mode Rayleigh-wave phase velocity (FMRWPV) in the study region (Fig. 6), using the Rayleigh-wave data set. Second, we determine the optimal 1-D $V_s$ model of the study region (Fig. 7) by jointly inverting the obtained 1-D FMRWPV and the stacked PRFs and SRFs at all stations. Then using the updated 1-D FMRWPV model for the 1-D $V_s$ model, we determine 2-D FMRWPV models at various periods (Fig. 8). Finally, we obtain a 3-D $V_s$ model beneath the study region (Figs 9 and 10) by conducting a joint inversion of the teleseismic $S$-wave relative traveltime (TS-RTT) residuals, the 2-D FMRWPV data, and the receiver functions.

Note that this new joint inversion approach is based on three well-established seismic inversion methods, which are the receiver-function linearized inversion method (Ammon et al. 1990), the Rayleigh-wave two-plane-wave (TPW) method (Forsyth & Li 2005), and the body-wave velocity tomographic method (Zhao et al. 1994). In addition, our approach is also an updated version of the joint inversion method by Liu & Zhao (2016). The joint inversion may bridge resolution gaps associated with each individual data set (Julià et al. 2000). However, a critical issue of the joint inversion of different kinds of data is to find the optimal relative weights for them used in the inversion. This issue exists in all studies dealing with a joint inversion of different kinds of geophysical data (e.g. Julià et al. 2000; Gallardo & Meju 2004, 2011; Maceira & Ammon 2009; Moorkamp et al. 2010, 2011; Villagòmez et al. 2014; Zhang et al. 2014; Syracuse et al. 2015, 2016; Fang et al. 2016; Liu & Zhao 2016; Gao & Zhang 2018; Liu et al. 2018). In this work we have conducted many tomographic inversions with
Figure 5. An example of measuring Rayleigh-wave data. (a) Broad-band three-component seismograms of a teleseismic event (red star in Fig. 4a) recorded at a seismic station (9A.CVST2). Z, vertical component; T, transverse component; R, radial component. (b) Gaussian-filtered Rayleigh waves in a period range of 20–150 s. (c) Each yellow diamond symbol denotes group velocity of the picked fundamental mode Rayleigh wave at each period with a clear and large enough normalized amplitude peak for the envelope of a Gaussian-filtered Rayleigh wave that is shown as a blue line in (b). The amplitudes and phases of the Gaussian-filtered Rayleigh waves are then determined around the group arrival times for different periods by Fourier analysis.

Figure 6. The 1-D Rayleigh-wave phase velocity. Red and blue colours denote low and high RMS values of the two-plane-wave real and imaginary component residuals for various periods and phase velocities. The white line denotes the initial 1-D phase velocity used in 2-D Rayleigh-wave tomography. Blue diamond symbols denote the 1-D phase velocity after five iterations, which are used to determine the 1-D $V_p$ model.

2.6 Determining 1-D $V_p$ model and 2-D Rayleigh-wave phase-velocity models

We determine a 1-D $V_p$ model beneath the Cape Verde hotspot by jointly inverting the stacked PRFs at 0–20 s after the direct
approach of Zhao et al. (1992, 1994). In this 3-D grid, however, and periods, we obtain an image as shown in Fig. 6. On the basis of real and imaginary component residuals in various phase-velocities 20–150 s. Finally, using the obtained series of RMS values of the arrival and SRFs at 0–20 s before the direct

The optimal 1-D model in a period range of 20–150 s (blue diamond symbols in Fig. 6). The 1-D FMR WPV model is obtained by using the teleseismic fundamental Rayleigh-wave amplitude and phase data in our third data set and applying the TPW interference technique (Forsyth & Li 2005) to account for multipath interference effects caused by heterogeneities outside the modelling space. The incoming Rayleigh wave at each period from a teleseismic event can be regarded as the sum of two horizontally propagating plane waves that are described by their respective amplitude, reference phase and direction (Forsyth & Li 2005). The TPW method adopts residuals of the real and imaginary components rather than the observed amplitude and phase data. In this study, with a given phase-velocity at a given period, we first determine the best-fitting TPW parameters for all events, in a least-squares sense using a simulated annealing method (Press et al. 1992), and then calculate the RMS values of the real and imaginary component residuals for all data. This operation is repeated in a phase-velocity range of 3.0 to 5.0 km s$^{-1}$ and in a period range of 20–150 s. Finally, using the obtained series of RMS values of the real and imaginary component residuals in various phase-velocities and periods, we obtain an image as shown in Fig. 6. On the basis of this image (Fig. 6), we set an initial 1-D FMR WPV of 4.0 km/s for all periods from 20 to 150 s (white line in Fig. 6).

With the initial 1-D FMR WPV model, we conduct a joint inversion for 2-D FMR WPV distribution at periods of 20–150 s in the study region, using the teleseismic fundamental mode Rayleigh-wave amplitude and phase data in our third data set, and applying the TPW method (Forsyth & Li 2005) that takes the finite-frequency effects into account (Yang & Forsyth 2006). We set up a 3-D grid with a lateral grid interval of 0.1° in the study region following the approach of Zhao et al. (1992, 1994). In this 3-D grid, however, each grid node is defined by latitude, longitude and period. The nodes in each 2-D grid mesh (for latitude and longitude) have the same period. The 2-D grid meshes are arranged at periods of 20 to 150 s with an interval of 5 s. The amplitude, reference phase and direction of each two-plane wave, and FMR WPV perturbations at every grid nodes from the initial 1-D FMR WPV model at periods of 20–150 s, are taken as unknown parameters. The phase-velocity perturbation at any point in the modelling space is computed by linearly interpolating the phase-velocity perturbations at the four grid nodes surrounding that point at each period. The LSM algorithm (Paige & Saunders 1982) is adopted to solve the system of observation equations relating the observed real and imaginary component residuals to the unknown phase-velocity perturbations and parameters of two-plane waves. Smoothing and damping regularization (Liu & Zhao 2016) are adopted to suppress the dramatic short-scale phase-velocity variations at each period and adjacent periods. In practice, we conducted many tomographic inversions using different values of the damping and smoothing parameters, to seek their optimal values considering the balance between the reduction of residuals and the smoothness of the resulting velocity model. A 2-D tomographic image at each period is obtained after three iterations. Then we calculate the average phase-velocity perturbation at each period, which is used to perturb the initial 1-D FMR WPV model at each period. The perturbed 1-D FMR WPV model is then used to conduct a new tomographic inversion. After five iterations, we obtain a new 1-D FMR WPV model (blue diamond symbols in Fig. 6), which is used to determine the optimal 1-D $V_s$ model for the study region with the stacked PRFs and SRFs. Such an iterative process for determining the optimal 1-D FMR WPV model is similar to the process used to determine the optimal 1-D Q model of the Japan subduction zone (Liu & Zhao 2015; Wang & Zhao 2019).

To determine a 1-D $V_s$ model, we arrange a 1-D grid in the modelling space with a vertical grid interval of 5 km at depths of 5–20 km, 10 km interval at depths of 20–100 km, and 25 km interval at depths of 100–700 km. $V_s$ perturbations at every grid node from a starting 1-D $V_s$ model are taken to be unknown parameters following the method of Zhao et al. (1994, 2012). At depths <100 km, the $V_s$ (or $V_s$ perturbation) at any point in the modelling space is equal to the $V_s$ (or $V_s$ perturbation) at the deeper one of the two grid nodes adjacent to that point. Below 100 km depth, however, the $V_s$ (or $V_s$ perturbation) at any point in the modelling space is computed by linearly interpolating the $V_s$ values (or $V_s$ perturbations) at the two grid nodes adjacent to that point. This scheme allows significant variations of $V_s$ at depths <100 km. The starting 1-D $V_s$ model is derived from the AK135 model (Kennett et al. 1995) without the...
Figure 8. Rayleigh-wave phase-velocity tomography beneath Cape Verde. The red and blue colours denote low and high phase velocities, respectively, whose perturbation (in per cent) scale is shown at the bottom. The period is shown at the lower-right corner of each map. The red triangle in each map denotes the active Fogo volcano. The 1-D FMRWPV model used in this tomography is shown as the red line in Fig. 7(d).

crust (grey line in Fig. 7a). A 1-D $V_p$ model is calculated by assuming $V_p/V_s = 1.75$, and a 1-D density ($\rho$) model is calculated using Birch’s law, $V_p = 3.125\rho - 2.4$ (Birch 1961). The stacked PRFs and SRFs for Gaussian parameters of 0.5, 0.7 and 1.0 (blue lines in Figs 7b and c), and the 1-D FMRWPV values at periods of 20–150 s (blue diamond symbols in Fig. 7d) form a set of observed data. During the inversion, sensitivity kernels and predictions of phase velocities are calculated using the code DISPER80 (Saito 1988), whereas those of receiver functions are calculated using a code we modified from an original code in the Computer Programs in Seismology (CPS; Herrmann 2013). Note that our modified code is used to calculate predictions of PRFs and SRFs from the LQ components rather than from the vertical and radial components in the original CPS code. The LSQR algorithm is adopted to solve the system of observation equations relating the observed data to the unknown $V_s$ parameters. We use smoothing and damping regularizations to suppress the dramatic short-scale $V_s$ variations (Zhao et al. 2012; Liu & Zhao 2016). The final 1-D $V_s$ model is obtained after fifty iterations (red line in Fig. 7a). The predicted FMRWPV values at periods of 20–150 s (red line in Fig. 7d) are adopted as the optimal 1-D FMRWPV model for the study region, which is used to conduct a joint inversion for 2-D FMRWPV distribution at periods of 20–150 s in the study region with the above-mentioned method. The RMS values of residuals for all periods are reduced significantly after the 2-D surface wave tomographic inversion (Supporting Information Fig. S5). The obtained 2-D FMRWPV distributions at periods of 20–150 s (Fig. 8) are then used to determine a 3-D $V_s$ model as described in the following.

2.7 Determining 3-D $V_s$ model

To constrain the detailed 3-D $V_s$ structure of the Cape Verde hotspot, we update the joint inversion method of Liu & Zhao (2016) that is
any point in the modelling space is computed by linearly interpolating the values (or \( V_s \) perturbations) at any point \( \delta V_s / V_s \) along the ray path between the intersection and the receiver as

\[
\delta V_s / V_s = \int_{\text{ray path}} K^{i}_{rs} \frac{\delta V_s}{V_s} \, dz \tag{1}
\]

where \( K^{i}_{rs} = (\partial T / \partial V_p)V_s \) is the sensitivity kernel at ray segment \( z \) relating the TS-RTT residual \( r \) to \( V_s \) anomaly \( \delta V_s / V_s \), and \( \partial T / \partial V_s \) is partial derivative of traveltime \( T \) with respect to \( V_s \).

The Rayleigh-wave phase-velocity perturbation at each period at a point in the study region is related to the 1-D \( V_s \), \( V_p \), and \( \rho \) structures beneath that point (e.g. Zhang & Tanimoto 1993; Yoshizawa et al. 2010). The Rayleigh-wave phase velocity is much more sensitive to \( V_s \) than to \( V_p \) and \( \rho \) (e.g. Saito 1988). Following Liu & Zhao (2016), we assume that the Rayleigh-wave phase-velocity perturbation \( \delta c / c \) at position \( x \) and period \( k \) mainly reflects \( V_s \) anomalies \( \delta V_s / V_s \) along a vertical ray in the crust and upper mantle beneath position \( x \) as

\[
\frac{\delta c}{c} = \int_{\text{ray path}} K^{i}_{rs} \frac{\delta V_s}{V_s} \, dz \tag{2}
\]

where \( K^{i}_{rs} \) is the sensitivity kernel at ray segment \( z \) and period \( k \). \( V_p \) and \( \rho \) are calculated with \( V_p / V_s = 1.75 \) and Birch’s law, respectively, which are the same as those used to determine the 1-D \( V_s \) model. The sensitivity kernel \( K^{i}_{rs} \) and the predicted phase velocity are calculated using the code DISPER80 (Saito 1988). This scheme is similar to that dealing with the TS-RTT residuals as mentioned above (Zhao et al. 1994).

A similar scheme is also used to deal with the receiver functions. If we assume that fixed-thickness horizontal layers exist in the study.
Figure 10. Vertical cross-sections of $V_s$ tomography beneath Cape Verde. This tomographic model is determined by a joint inversion of the teleseismic S-wave traveltime data, the Rayleigh-wave phase-velocity data and the receiver functions. Locations of the vertical cross-sections (a and b) are shown in blue lines on the inset map. The surface topography along each profile is shown atop each cross-section. The red and blue colours denote low- and high-$V_s$ perturbations, respectively, whose scale (in per cent) is shown above the inset map. The dashed lines in panels (a) and (b) denotes the 410-km discontinuity. The red crosses on the inset map and in panels (a) and (b) denote the piercing points at 15 km depth of P15s and S15p as shown in Fig. 3h. The purple circles on the inset map and in panels (a) and (b) denote the piercing points at 60 km depth of $S60p$ as shown in Fig. 3h. The piercing points (red crosses and purple circles) in panels (a) and (b) are within a 25-km width of each profile. Red line and green lines in (c) show the 1-D $V_s$ model used and the obtained 3-D $V_s$ model, respectively. Three grey dashed lines in (c) show synthetic $V_s$ models obtained with a function of melt fraction 0.1, 0.5 and 1.0 per cent, respectively (see the text for details). Panels (d) and (e) show 3-D views of the $V_s$ tomography, and the iso-surface of velocity anomalies is rendered where $V_s$ perturbation is $-1$ per cent.

The TS-RTT residuals, the 2-D FMRWPV perturbations, and the PRF and SRF residuals form a set of observed data with a column vector $d$ of dimension $N$. Corrections to all the unknown parameters form a column vector $\Delta m$, which can be arranged as: $\Delta m^T = (\delta V_{s1}, \delta V_{s2}, \ldots, \delta V_{sJ})$, where $J$ is the number of unknown parameters. The vectors for the data and the unknown parameters are related in the following system of observation equations:

$$d = G \Delta m + e,$$

where $e$ is an error vector and $G$ is a matrix whose $N \times J$ elements are the sensitivity kernels as mentioned above. The LSQR algorithm is then adopted to solve the system of observation equations (5) relating the observed data to the unknown $V_s$ parameters. To suppress the dramatic short-scale $V_s$ variations, we adopt smoothing and damping regularizations (Zhao et al. 2012; Liu & Zhao 2016). The final tomographic model is obtained after twenty iterations (Supporting Information Fig. S6). The RMS values of various volume, then the receiver function is more sensitive to $V_s$ than to $V_p$ and $\rho$ (Owens et al. 1984; Ammon et al. 1990). In this study, at a seismic station, we assume that the stacked PRF residual $b_p$ and the stacked SRF residual $b_s$ at time $t$ mainly reflect $V_s$ anomalies $\delta V_s / V_s$ along a vertical ray in the crust and upper mantle beneath the station as

$$b_p = \int_{ray \ path} K^{bp} \cdot V_s \cdot \frac{\delta V_s}{V_s} \ dz,$$

$$b_s = \int_{ray \ path} K^{bs} \cdot V_s \cdot \frac{\delta V_s}{V_s} \ dz,$$

where $K^{bp} \cdot V_s$ is the sensitivity kernel at ray segment $z$ and time $t$. Parts of the sensitivity kernel, $K^{bp}$, and the predicted receiver function are calculated using the above-mentioned code modified from the CPS (Herrmann 2013).
3 ANALYSIS AND RESULTS

To evaluate the resolution of our inversion results, we have conducted extensive checkerboard resolution tests (CRTs) and restoring resolution tests (RRTs) (Zhao et al. 1992; Liu & Zhao 2016). The test results (Supporting Information Figs S7–S14) show that our seismic images obtained (Figs 7–10) are quite reliable. Note that these tests rely on the assumption that the theory and techniques used to compute synthetic traveltimes, dispersion data and receiver function waveforms are error-free.

3.1 1-D $V_s$ model

To determine the 1-D $V_s$ model of the study region, we conducted a joint inversion of the stacked receiver function data and the 1-D FMRWPV data. The final results are shown in Fig. 7. Many inversions are performed to find reasonable values of the damping, smoothing, relative weights and iteration parameters. The finally adopted parameters are shown in Supporting Information Table S1. The fit of the obtained model to the stacked receiver functions and the 1-D FMRWPV data is improved significantly after the joint inversion (Fig. 7).

We conducted an RRT (Zhao et al. 1992) to confirm the obtained 1-D $V_s$ model (Fig. 7). To perform an RRT, the obtained 1-D $V_s$ model is adopted as the input model for which we calculate synthetic receiver functions with various Gaussian parameters and 1-D FMRWPV values at periods of 20–150 s. Then we inverted the synthetic data to see whether the input model could be recovered or not. The RRT results (Supporting Information Figs S7a–h) show that main features of the optimal 1-D $V_s$ model are well recovered, suggesting that the obtained 1-D $V_s$ model (Fig. 7) is reliable. In addition, we only inverted the synthetic receiver functions or the synthetic 1-D FMRWPV data, but the input model is not well recovered (Supporting Information Figs S7i–p). These test results indicate that the joint inversion is necessary to get a reliable 1-D $V_s$ model.

We further conducted many synthetic tests to estimate the vertical resolution of the obtained 1-D $V_s$ model. The test results (Supporting Information Fig. S8) show that the 1-D $V_s$ model has a vertical resolution of $\sim$10 km at depths of 0–60 km, $\sim$20 km at depths of 60–100 km and $\sim$50–200 km at depths of $\sim$100–400 km.

3.2 2-D Rayleigh-wave phase-velocity model

To evaluate the spatial resolution of the Rayleigh-wave phase-velocity (RWPV) tomography, we conducted many CRTs to assess the adequacy of the ray coverage. To perform a CRT, we first assign positive and negative RWPV perturbations of 6 per cent to all the 2-D grid nodes, then calculate synthetic real and imaginary component residuals for the checkerboard model with the same numbers of seismic stations, teleseismic events and parameters of the two-plane waves as those in the real data set. The synthetic data are then inverted to investigate whether the assigned RWPV anomalies could be recovered or not. To simulate errors contained in the observed data, random noise with a normal distribution having a standard deviation of 1 per cent anomaly was added to the synthetic real and imaginary component residuals before the tomographic inversion. The test results for the RWPV tomography at 9 different periods are shown in Supporting Information Fig. S9, which indicate that the obtained 2-D RWPV models have a lateral resolution of $\sim$50–130 km in the study region. The final results of the RWPV tomography are shown in Fig. 8. The main features of the RWPV tomography are confirmed by the corresponding RRT (Supporting Information Fig. S10).

3.3 3-D $V_s$ model

We conducted many CRTs to assess the adequacy of the ray coverage and to evaluate the robustness of the joint inversion of multiple seismic data. Supporting Information Figs S11–S13 show comparisons of four test results obtained by inverting the three data sets (i.e. the TS-RTT residuals, the FMRWPV data, and the receiver functions). The Rayleigh-wave data and the receiver functions can constrain the shallow structure well down to $\sim$100–200 km depth, whereas the TS-RTT data mainly constrain the deeper structure of the study region. A joint inversion of the three kinds of seismic data can result in a much more robust 3-D $V_s$ model with a lateral resolution of $\sim$50–100 km beneath the Cape Verde hotspot (Supporting Information Figs S11–S13).

We performed many inversions to find reasonable values of the damping, smoothing, relative weights and iteration parameters (Table S1). The final optimal tomographic results are shown in Figs 9 and 10. The RMS values of residuals of the receiver functions, the FMRWPV data as well as the TS-RTT data are all reduced significantly after the joint inversion (Supporting Information Fig. S4).

To confirm our tomographic results, we determined four 3-D $V_s$ models by inverting the three data sets (Supporting Information Figs S14 and S15). These $V_s$ models are confirmed by the corresponding RRT results (Supporting Information Fig. S14). Note that in these RRTs, random noises with a normal distribution having a standard deviation of 0.1 s, 0.01 km/s and 10 per cent anomalies are added to the synthetic TS-RTT, FMRWPV and receiver function data, respectively, before the tomographic inversion. The model obtained by the joint inversion of all the three data sets is much improved, as compared with the other models (Supporting Information Figs S14 and S15), indicating that our joint inversion method is effective.

To investigate the effect of the initial $V_p/V_s$ ratio on the final 3-D $V_s$ model, we conducted tomographic inversions with three different $V_p/V_s$ values (1.70, 1.75 and 1.80). The patterns of $V_s$ images thus obtained are very similar to each other (Supporting Information Fig. S16), suggesting that the 3-D $V_s$ model obtained by this study is not very sensitive to the initial $V_p/V_s$ ratio adopted.

3.4 Seismic velocity structure beneath Cape Verde

The optimal 1-D $V_s$ model for the Cape Verde hotspot is obtained by a joint inversion of the stacked receiver functions and the 1-D FMRWPV data, which is confirmed by extensive resolution tests (Supporting Information Figs S7 and S8). This model (red line in Fig. 7a) is characterized by two $V_s$ discontinuities at $\sim$15 ± 10 km depth and $\sim$60 ± 10 km depth beneath Cape Verde, which are interpreted as the Moho discontinuity and the Gutenberg (G) discontinuity, respectively. Note that the piercing points of the P15s and S15p converted phases at 15 km depth are mainly located beneath the individual islands (green and blue cross symbols in Fig. 3h), whereas the piercing points of the S60p converted phases at 60 km depth are generally located around the volcanic islands (red circle symbols...
in Fig. 3h). Hence, we can only constrain the average Moho depth beneath the individual islands from the Pms and Smp converted phases in the stacked PRFs and SRFs (Figs 3i and j), and the average depth of the G discontinuity around the volcanic islands from the Sgo converted phases in the stacked SRFs (Fig. 3i).

The optimal 3-D Vp model with a lateral spatial resolution of ~50–100 km is obtained by a joint inversion of the receiver functions, the 2-D FMRWPV data, and the TS-RTT data, which is confirmed by extensive resolution tests (Supporting Information Figs S11–S14). The tomographic results (Figs 9 and 10) show that, between the north and south islands, obvious high-Vp anomalies exist in the uppermost mantle down to a depth of ~100–150 km beneath the Atlantic Ocean, whereas obvious low-Vp anomalies exist in the uppermost mantle in and around the volcanic islands including the active Fogo volcano. Note that clear S60p converted phases mainly sample the areas with low-Vp anomalies in the uppermost mantle in and around the volcanic islands (Supporting Information Fig. S17). These low-Vp anomalies merge into a significant column-like low-Vp zone at a depth greater than ~150 km beneath the Cape Verde swell. The column-like low-Vp zone is visible down to a depth of ~400 km (Fig. 10).

4 DISCUSSION

4.1 Magma chambers in the oceanic lithosphere

Wide-angle seismic data from a marine geophysical survey revealed a normal oceanic crust with a thickness of ~7–13 km beneath the Cape Verde swell (Pim et al. 2008; Wilson et al. 2010, 2013), whereas the crust seems to be thicker (10–22 km) under the individual volcanic islands revealed by an analysis of PRFs (Lodge & Helffrich 2006) (Fig. 1). A thicker crust (20–30 km) including a ~10–20 km thick layer of magmatic underplate has been revealed by a joint inversion of PRFs and SRFs (Vinnik et al. 2012). But flexure modelling with a subcrustal load shows the Moho discontinuity at a depth of ~20 km (Wilson et al. 2013). Our present 1-D Vp model (red line in Fig. 7a) indicates a Moho depth of ~15 ± 10 km beneath the individual islands.

An early PRF study revealed a high-velocity and low-density layer overlying a low-Vs zone starting at ~80 km depth beneath Cape Verde (Lodge & Helffrich 2006). A subsequent simultaneous inversion of PRFs and SRFs detected a Vs discontinuity at ~65–70 km depth (Vinnik et al. 2012). In addition, the compensation depth for the Cape Verde swell may be also at ~70 km depth revealed by analysis of the long-wavelength gravity-topography slope and geoid-topography ratio (Wilson et al. 2013). Our present 1-D Vp model (red line in Fig. 7a) shows a Vs discontinuity at ~60 km depth beneath the volcanic islands. This discontinuity is not sharper than the Moho discontinuity at ~15 km depth. We interpret the discontinuity at ~60 km depth as the G discontinuity beneath oceanic regions (Rychert et al. 2019, 2020). The G discontinuity is often related to the oceanic lithosphere–asthenosphere boundary (LAB), which is a transition from the upper rigid conductively cooling plate to the underlying ductile convecting mantle (Fischer et al. 2010; Geissler et al. 2017; Kawakatsu & Utada 2017; Karato & Park 2019).

Abundant low-frequency earthquakes occur beneath the Fogo volcano, most of which are located directly below the volcano crater (Faria & Fonseca 2014), whereas the regular earthquake activity beneath the Fogo volcano appears to be not high (Greinemeyer et al. 2010; Vales et al. 2014; Leva et al. 2019). The CVPLUME network deployed from November 2007 to September 2008 did not detect significant activity at the Fogo volcano (Vales et al. 2014). During September 2011 to May 2013, another seismic network detected only 57 volcano-tectonic events, all of which occurred in the crust (~7 km depth) (Faria & Fonseca 2014). In August 2016, a swarm of deep regular events with focal depths of ~38–44 km was detected in the uppermost mantle beneath the Fogo volcano, which was related to fracturing induced by magmatic injection (Leva et al. 2019).

Our 3-D Vp model shows obvious low-Vp anomalies in the uppermost mantle beneath most of the volcanic islands (Figs 9 and 10), which may represent hot and/or melt-rich zones possibly relating to mantle magma chambers feeding the active volcanoes. Thermo-barometric and geochemical studies of magmas erupted in 1995 and 2014 at the Fogo volcano also suggested the existence of several magma chambers mainly located in the uppermost mantle at depths of ~15–25 km (Hildner et al. 2011; Mata et al. 2017). The existence of potential mantle magma chambers was also suggested by geodetic modelling that did not reveal deformation precursors to eruptions in 2014 at the Fogo volcano (González et al. 2015). The existence of magma chambers in the uppermost mantle seems to be common beneath some ocean islands with low magma supply rates (Klügel et al. 2015).

4.2 Volatile-rich asthenosphere

An early PRF study (Lodge & Helffrich 2006) revealed a low-Vs zone (~4.1 km/s) in the Cape Verde asthenosphere. A subsequent simultaneous inversion of PRFs and SRFs (Vinnik et al. 2012) detected a low-Vs zone (down to ~3.8 km/s at ~70–100 km depths) with a normal Vp/Vs (~1.8) in the asthenosphere. A recent Rayleigh-wave study by cross-correlating teleseismic data between pairs of stations also shows a low-Vs zone (down to ~4.2 km/s at 170 km depth) in the asthenosphere (Carvalho et al. 2019). Our present results reveal obvious low-Vs anomalies in the asthenosphere beneath the Cape Verde hotspot (Figs 9 and 10).

The obvious low-Vs zone in the asthenosphere beneath the Cape Verde hotspot may result from high-temperature, mantle volatiles, partial melting and oxygen fugacity (Chantel et al. 2016; Cline et al. 2018; Karato & Park 2019). The mantle potential temperature beneath Cape Verde (~1511 °C, Putirka 2008; ~1400–1420 °C, Carvalho et al. 2019) is considered to be higher than that below oceanic ridges (~1300–1400 °C, Herzberg et al. 2007; Stixrude & Lithgow-Bertelloni 2007; Katsura et al. 2010), which may be underestimated by ~60 °C (Sarafian et al. 2017). The H2O content of the primitive Cape Verde magmas is estimated to be 1.78 wt per cent (Putirka 2008) or ~0.5–2.5 wt per cent (Moussofam et al. 2019), and the CO2 content is estimated to be 1000–4500 ppm (Moussofam et al. 2019). In addition, ferric iron to total iron ratios (Fe3+/ΣFe) are up to ~0.2–0.3 for melt inclusions from lava samples of the active Fogo volcano (Moussofam et al. 2019). Such a hot volatile-rich oxidized condition beneath Cape Verde may incite wet carbonated silicate melting at a depth of much greater than ~300 km (Dasgupta et al. 2013; Moussollem et al. 2019). Some magnesium carbonatites from Cape Verde are characterized by Mg# of ~80–90 (Mourão et al. 2012b), similar to those obtained experimentally by melting of carbonated peridotites (Foley et al. 2009). Oceanic carbonatites, which are found only in the Cape Verde and Canary Islands among oceanic islands, exhibit a high 3He/4He ratio, favouring a deep origin (Mata et al. 2010), though the deep source may be from recycled marine carbonates via subduction, not from primordial carbon in the D∗ layer (Doucelance et al. 2014).
To constrain the melt fraction in the asthenosphere beneath Cape Verde, we calculate the $V_s$ decrease due to the presence of melt, from in situ ultrasonic velocity measurements on a series of partially molten samples at 2.5 GPa and up to 1623 K (Chantel et al. 2016). Here $V_s$ is a function of melt fraction $\phi$, $V_s = 0.065\phi^2 - 0.5565\phi + V_{so}$, where $V_{so}$ is the $V_s$ with a melt fraction of 0 per cent (Chantel et al. 2016). The results (Fig. 10c) show that the obvious low-$V_s$ anomalies in our present model could be matched with a melt fraction between ~0.5 per cent and ~1.0 per cent, relative to 0 per cent melt for the AK135 model. The estimated melt fraction between ~0.5 per cent and ~1.0 per cent in the asthenosphere beneath the Cape Verde hotspot is significantly higher than that (~0.1–0.2 per cent) beneath the normal oceanic lithosphere (Chantel et al. 2016; Selway & O’Donnell 2019; Yang et al. 2020).

Such a relatively high melt fraction in the asthenosphere may originate from a deeper onset of melting in a volatile-rich mantle beneath the Cape Verde hotspot (Hirschmann 2010). Global topographic models (Zhao et al. 2013; French & Romanowicz 2015) generally show low seismic velocity anomalies in the lower mantle under Cape Verde, which may indicate a deep mantle plume.

Although an early PRF study revealed a normal thickness of the mantle transition zone beneath Cape Verde and suggested that the Cape Verde hotspot does not require a vertical thermal anomaly from the lower mantle (Helffrich et al. 2010), the subsequent analysis of PRFs and SRFs argued that the mantle transition zone is up to ~30 km thinner than that of the ambient mantle and suggested that a deep mantle plume exists under this hotspot (Vinnik et al. 2012). A result of analysing teleseismic PP and SS precursors also favours a thin mantle transition zone beneath the Cape Verde hotspot, due to a deep mantle plume (Saki et al. 2015). However, our present PRF results (Fig. 3j) show a standard differential time of 24 s between P410s and P660s arrivals, consisting with the results of Helffrich et al. (2010). Such an issue may call for more detailed studies in the near future.

5 CONCLUSIONS

To study the influence of a mantle plume on the oceanic lithosphere–asthenosphere system, we investigate the detailed 3-D $V_s$ structure of the upper mantle beneath the Cape Verde hotspot by jointly inverting teleseismic S-wave travel times, fundamental mode Rayleigh-wave data, and PRFs and SRFs, recorded by several seismic networks on the Cape Verde archipelago. Our results reveal an unprecedented detailed 3-D $V_s$ structure of the upper mantle beneath the hotspot. Two distinct $V_s$ discontinuities are detected at ~15 km and ~60 km depths beneath the volcanic islands, which are interpreted as the Moho and G discontinuities, respectively. Significant low-$V_s$ anomalies exist in the oceanic lithospheric mantle beneath the volcanic islands, which may be related to hot and/or melt-rich zones feeding the active volcanoes. A prominent column-like low-$V_s$ zone appears in the asthenosphere down to ~400 km depth, which may reflect a hot, volatile- and melt-rich plume beneath Cape Verde.

ACKNOWLEDGEMENTS

We thank the IRIS and GEOFON data centres for providing the waveform data used in this study. We are very grateful to Prof. Martin Schimmel (the Editor), Prof. Wolfram Geissler and three anonymous reviewers for their thoughtful review comments and suggestions which have greatly improved this paper. The free software packages GMT (Wessel & Smith 1998), TauP (Crotwell et al. 1999), SAC (Goldstein et al. 2003), CPS (Herrmann 2013) and ParaView (www.paraview.org) are used in this study. We appreciate helpful discussions with Prof. Shengyao Yu, as well as Drs Chuanxu Chen, Xiyao Li and Shujuan Zhao. This work was supported by grants from the Natural National Science Foundation of China (41972211 and 41602207) to XL and a grant from Japan Society for the Promotion of Science (19H01996) to DZ.

REFERENCES

Ammon, C., Randall, G. & Zandt, G., 1990. On the nonuniqueness of receiver function inversions, J geophys. Res., 95, 15 303–15318.

Birch, F., 1961. The velocity of compressional waves in rocks to 10 kilobars: 2., J geophys. Res., 66, 2199–2224.

Burdick, L. & Langston, C., 1977. Modeling crustal structure through the use of converted phases in teleseismic body-wave forms, Bull. seism. Soc. Am., 67, 677–691.

Carvalho, J., Bonadio, R., Silveira, G., Lebedev, S., Mata, J., Arroucau, P., Meier, T. & Celli, N., 2019. Evidence for high temperature in the upper mantle beneath Cape Verde archipelago from Rayleigh-wave phase-velocity measurements, Tectonophysics, 770, 228225, doi:10.1016/j.tecto.2019.228225.

Chantel, J., Manthilake, G., Andraud, D., Novella, D., Yu, T. & Wang, Y., 2016. Experimental evidence supports mantle partial melting in the asthenosphere, Sci. Adv., 2, e1600246, doi:10.1126/sciadv.1600246.

Christensen, B., Holm, P. & Wilson, J., 2001. Helium, argon and lead isotopic composition of volcanics from Santo Antão and Fogo, Cape Verde Islands, Chem. Geol., 178, 127–142.

Cline, C., H. Faul, U., David, E., Berry, A. & Jackson, I., 2018. Redox-influenced seismic properties of upper-mantle olivine, Nature, 555, 355–358.

Courtney, R. & White, R., 1986. Anomalous heat flow and geoid across the Cape Verde Rise: evidence for dynamic support from a thermal plume in the mantle, Geophys. J. Int., 87, 815–867.

Crotwell, H., Owens, T. & Ritsema, J., 1999. The TauP toolkit: flexible seismic travel-time and ray-path utilities, Seismol. Res. Lett., 70, 154–160.

Dasgupta, R., Mallik, A., Tsuno, K., Withers, A., Hirth, G. & Hirschmann, M., 2013. Carbon-dioxide-rich silicate melt in the Earth’s upper mantle, Nature, 493, 211–215.

Detrick, R. & Crough, S., 1978. Island subsidence, hot spots, and lithospheric thinning, J geophys. Res., 83, 1236–1244.

Doucclanze, R., Eserig, S., Moreira, M., Garijep, C. & Kurz, M., 2003. Pb-Sr-He isotope and trace element geochemistry of the Cape Verde Archipelago, Geochim. Cosmochim. Acta, 67, 3717–3733.

Doucclanze, R., Bellot, N., Boyet, M., Hammouda, T. & Bosq, C., 2014. What coupled cerium and neodymium isotopes tell us about the deep source of oceanic carbonatites, Earth planet. Sci. Lett., 407, 175–186.

Ducker, K. & Sheehan, A., 1997. Mantle discontinuity structure from mid-point stacks of converted P to S waves across the Yellowstone hotspot track, J geophys. Res., 102, 8313–8327.

Fang, H., Zhang, H., Yao, H., Allam, A., Zigone, D., Ben-Zion, Y., Thubler, C. & van der Hilst, R.D., 2016. A new algorithm for three-dimensional joint inversion of body wave and surface wave data and its application to the Southern California plate boundary region, J geophys. Res., 121, 3557–3569.

Faria, B. & Fonseca, J., 2014. Investigating volcanic hazard in Cape Verde Islands through geophysical monitoring: network description and first results, Nat. Hazards Earth Syst. Sci., 14, 485–499.

Farra, V. & Vinnik, L., 2000. Upper mantle stratification by P and S receiver functions, Geophys. J. Int., 141, 699–712.

Fischer, K., Ford, H., Abt, D. & Rybeck, C., 2010. The lithosphere–asthenosphere boundary, Annu. Rev. Earth planet. Sci., 38, 551–577.

Foley, S., Yaxley, G., Rosenthal, A., Buhre, S., Kiseeva, E., Rapp, R. & Jacob, D., 2009. The composition of near-solidus melts of peridotite in...
the presence of CO₂ and H₂O between 40 and 60 kbar, Lithos, 112, 274–283.

Forstyth, D. & Li, A., 2005. Array analysis of two-dimensional variations in surface wave phase velocity and azimuthal anisotropy in the presence of multipathing interference, in Seismic Earth. Array Analysis of Broad-band Seisograms, Vol. 157, pp. 81–97, eds Levander, A. & Nolet, G., Geophysical Monograph, AGU.

French, S. & Romanowicz, B., 2015. Broad plumes rooted at the base of the Earth’s mantle beneath major hotspots, Nature, 525, 95–99.

Gallardo, L. & Meju, M., 2011. Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints, J. geophys. Res., 116, B03311, doi:10.1029/2010JB007216.

Gallardo, L. & Meju, M., 2011. Structure-coupled multiphysics imaging in geophysical sciences, Rev. Geophys., 49, RG1003, doi:10.1029/2010RG000330.

Gao, J. & Zhang, H., 2018. An efficient sequential strategy for realizing cross-gradient joint inversion: method and its application to 2-D cross borehole seismic traveltime and DC resistivity tomography, Geophys. J. Int., 213, 1044–1055.

Geissler, W., Jokat, W., Jegen, M. & Baba, K., 2017. Thickness of the oceanic crust, the lithosphere, and the mantle transition zone in the vicinity of the Tristan da Cunha hot spot estimated from ocean-bottom and ocean-island seismometer receiver functions, Tectonophysics, 716, 33–51.

Goldstein, P., Dodge, D., Firpo, M. & Minner, L., 2003. SAC2000: signal processing and analysis tools for seismologists and engineers, in The IASPEI International Handbook of Earthquake and Engineering Seismology, pp. 1613–1641, eds Lee, W., Kanamori, H., Jennings, P. & Kisslinger, C., Academic Press.

González, P., Bagnardi, M., Hooper, A., Larsen, V., Marinovic, P., Samsonov, S. & Wright, T., 2015. The 2014–2015 eruption of Fogo volcano: geodetic modeling of Sentinel-1 TOPS interferometry, Geophys. Res. Lett., 42, 9239–9246.

Grevemeyer, I., Helffrich, G., Faria, B., Booth-Rea, G., Schnabel, M. & Weinrebe, W., 2010. Seismic activity at Câdamento seamount near Fogo Island, Cape Verdes—formation of a new ocean island? Geophys. J. Int., 180, 552–558.

Helffrich, G., Faria, B., Fonseca, J., Lodge, A. & Kaneshima, S., 2010. Transition zone structure under a stationary hot spot: Cape Verde, Earth planet. Sci. Lett., 289, 156–161.

Herrmann, R., 2013. Computer programs in seismology: an evolving tool for instruction and research, Seismol. Res. Lett., 84, 1081–1088.

Hertzberg, C., Asimow, P., Arndt, N., Niu, Y., Lesher, C., Filton, J., Cheadle, M. & Saunders, A., 2007. Temperatures in ambient mantle and plumes: constraints from basalts, picrites, and komatiites, Geochim. Geophys. Geosyst., 8, Q00B06, doi:10.1029/2006GC001390.

Hildner, E., Klugel, A. & Hauff, F., 2011. Magma storage and asthenospheric evolution in the upper mantle beneath cratons, Science, 329, 1137–1140.

King, S. & Ritsema, J., 2000. African hot spot volcanism: small-scale convection in the upper mantle beneath cratons, Earth planet. Sci. Lett., 180, 85–99.

Kempf, R., Engdahl, E. & Buland, R., 2015. Constraints on seismic velocities in the Earth from travel times, Geophys. J. Int., 196, 108–124.

Kawakatsu, H. & Utada, H., 2017. Seismic and electrical signatures of the lithosphere–asthenosphere system of the normal oceanic mantle, Annu. Rev. Earth planet. Sci., 45, 139–167.

Kawasaki, H. & Utada, H., 2017. Seismic and electrical signatures of the lithosphere–asthenosphere system of the normal oceanic mantle, Annu. Rev. Earth planet. Sci., 45, 139–167.

Kennett, B., Engdahl, E. & Buland, R., 1995. Constraints on seismic velocities in the Earth from travel times, Geophys. J. Int., 122, 108–124.

Kind, R., Yuan, X. & Kumar, P., 2012. Seismic receiver functions and the lithosphere–asthenosphere boundary, Tectonophysics, 536–537, 25–43.

Klingel, A., Longpré, M., García-Cañada, L. & Stix, J., 2015. Deep intrusions, lateral magma transport and related uplift at ocean island volcanoes, Earth planet. Sci. Lett., 431, 140–149.

Leva, C., Rümpker, G., Link, F. & Wöllbern, I., 2019. Mantle earthquakes beneath Fogo volcano, Cape Verde: evidence for subcrustal fracturing induced by magmatic injection, J. Volc. Geotherm. Res., 386, 106672.

Li, X., Kind, R., Yuan, X., Wöllbern, I. & Hanka, W., 2004. Rejuvenation of the lithosphere by the Hawaiian plume, Nature, 427, 827–829.

Ligorría, J. & Ammon, C., 1999. Iterative deconvolution and receiver-function estimation, Bull. seism. Soc. Am., 89, 1395–1400.

Liu, H., Chen, F., Leng, W., Zhang, H. & Xu, Y., 2018. Crustal footprint of the Hainan plume beneath southeast China, J. geophys. Res., 123, 3065–3079.

Liu, X. & Zhao, D., 2014. Seismic evidence for a mantle plume beneath the Cape Verde hotspot, Int. Geol. Rev., 56, 1213–1225.

Liu, X. & Zhao, D., 2015. Seismic attenuation tomography of the Southwest Japan arc: new insight into subduction dynamics, Geophys. J. Int., 201, 135–156.

Liu, X. & Zhao, D., 2016. P and S wave tomography of Japan subduction zone from joint inversions of local and teleseismic travel times and surface-wave data, Phys. Earth planet. Int., 252, 1–22.

Lodge, A. & Helffrich, G., 2006. Depleted swell root beneath Cape Verde Islands, Geology, 34, 449–452.

Lodhia, B., Roberts, G., Fraser, A., Fishwick, S., Goes, S. & Jarvis, J., 2018. Continental margin subsidence from shallow mantle convection: example from West Africa, Earth planet. Sci. Lett., 481, 350–361.

Maceira, M. & Ammon, C., 2009. Joint inversion of surface wave velocity and gravity observations and its application to central Asian basins shear velocity structure, J. geophys. Res., 114, B02314, doi:10.1029/2007JB005157.

Mata, J., Moreira, M., Doucelance, R., Ader, M. & Silva, L., 2010. Noble gas and carbon isotopic signatures of Cape Verde oceanic carbonatites: implications for carbon provenance, Earth planet. Sci. Lett., 291, 70–83.

Mata, J. et al., 2017. The 2014–15 eruption and the short-term geochemical evolution of the Fogo volcano (Cape Verde): evidence for small-scale mantle heterogeneity, Lithos, 288, 91–107.

Monnerault, J. & Cazenave, A., 1990. Depth and geoid anomalies over oceanic hotspot swells: a global survey, J. geophys. Res., 95, 15429–15431.

Moorkamp, M., Jones, A. & Fishwick, S., 2010. Joint inversion of receiver functions, surface wave dispersion, and magnetotelluric data, J. geophys. Res., 115, B04318, doi:10.1029/2009JB006369.

Morán-Menéndez, A., Begg, D., Maceira, M., Schwendeling, M., Roberts, A. & Hobbs, R., 2011. A framework for 3-D joint inversion of MT, gravity and seismic refraction data, Geophys. J. Int., 184, 477–493.

Mourão, C., Moreira, M., Mata, J., Raquin, A. & Madeira, J., 2012a. Primary and secondary processes constraining the noble gas isotopic signatures of carbonatites and silicate rocks from Brava Island: evidence for a lower mantle origin of the Cape Verde plume, Contrib. Mineral. Petrol., 163, 995–1009.

Mourão, C., Mata, J., Doucelance, R., Madeira, J., Millet, M. & Moreira, M., 2012b. Geochemical temporal evolution of Brava Island magma: constraints on the variability of Cape Verde mantle sources and on carbonatite–silicate magma link, Chem. Geol., 334, 44–61.
Vinnik, L., 1977. Detection of waves converted from P to SV in the mantle, *Earth planet. Sci. Lett.* 248, 685–699.

Owens, T., Zandt, G. & Taylor, S., 1984. Seismic evidence for an ancient rift beneath the Cumberland Plateau, Tennessee: a detailed analysis of broadband teleseismic P waveforms, *J. geophys. Res.* 89, 7783–7795.

Paige, C. & Saunders, M., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.* 8, 43–71.

Pim, J., Peirce, C., Watts, A., Grevermeyer, I. & Krabbenhoft, A., 2008. Crustal structure and origin of the Cape Verde Rise, *Earth planet. Sci. Lett.* 272, 422–428.

Press, W., Teukolsky, S., Vetterling, W. & Flannery, B., 1992. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd edn, Cambridge Univ. Press.

Putirka, K., 2008. Excess temperatures at ocean islands: implications for mantle layering and convection, *Geology* 36, 283–286.

Rychert, C., Harmon, N. & Thirumala, S., 2019. Seismic imaging of the base of the ocean plates, in *Lithospheric Discontinuities*, Vol. 239, pp. 71–87, eds Yuan, H. & Romanowicz, B., Geophysical Monograph, AGU.

Rychert, C., Harmon, N., Connable, S. & Wang, S., 2020. The nature of the lithosphere–asthenosphere boundary, *J. geophys. Res.* 125, e2018JB016463, doi:10.1029/2018JB016463.

Saito, M., 1988. DISPER80: a subroutine package for the calculation of seismic normal mode solutions, in *Seismological Algorithms: Computational Methods and Computer Programs*, pp. 293–319, ed. Dooomsb, D., Academic Press.

Saki, M., Thomas, C., Nippres, S. & Lessing, S., 2017. Topography of upper mantle seismic discontinuities beneath the North Atlantic: the Azores, Canary and Cape Verde plumes, *Earth planet. Sci. Lett.* 409, 193–202.

Samrock, L., Wartho, J. & Hansteen, T., 2019. 40Ar–39Ar geochronology of the Canary and Cape Verde plumes, *Academic Press.*

Sakuraba, L. & Lithgow-Bertelloni, C., 2007. Influence of lateral heterogeneity and dynamics in Earth’s mantle, *Earth planet. Sci. Lett.* 517, 117–124.

Stixrude, L. & Lithgow-Bertelloni, C., 2007. Influence of phase transformations on lateral heterogeneity and dynamics in Earth’s mantle, *Earth planet. Sci. Lett.* 263, 45–55.

Svensen, L. & Jacobsen, B., 2007. Absolute S-velocity estimation from receiver functions, *Geophys. J. Int.* 170, 1089–1094.

Syracuse, E., Maceira, M., Zhang, H. & Thurber, C., 2015. Seismicity and seafloor spreading beneath the central Pacific from surface wave analysis of OBS records, *Earth planet. Sci. Lett.* 534, 116086.

Wang, Y. & Forsyth, D., 2006. Regional tomographic inversion of the amplitude and phase of Rayleigh waves with 2-D sensitvity kernels, *Geophys. J. Int.* 166, 1148–1160.

Yang, X., Luo, Y., Xu, H. & Zhao, K., 2020. Shear wave velocity and radial anisotropy structures beneath the central Pacific from surface wave analysis of OBS records, *Earth planet. Sci. Lett.* 534, 116086.

Wang, Z.W. & Zhao, D., 2019. Updated attenuation tomography of Japan subduction zone, *Geophys. J. Int.* 219, 1679–1697.

Wessel, P. & Smith, W., 1998. New improved version of the Generic Mapping Tools released, *EOS, Trans. Am. geophys. Un.* 79, 579.

Wilson, D., Peirce, C., Watts, A., Grevermeyer, I. & Krabbenhoft, A., 2010. Uplift at lithospheric swells—I: seismic and gravity constraints on the crust and uppermost mantle structure of the Cape Verde mid-plate swell, *Geophys. J. Int.* 182, 531–550.

Wilson, D., Peirce, C., Watts, A. & Grevermeyer, I., 2013. Uplift at lithospheric swells—II: is the Cape Verde mid-plate swell supported by a lithosphere of varying mechanical strength? *Geophys. J. Int.* 193, 798–819.

SUPPORTING INFORMATION

Supplementary data are available at GJI online.

**Figure S1.** Stacked PRFs and SRFs. Blue lines show the stacked (a) SRFs and (b) PRFs at all seismic stations with Gaussian parameters of 0.5, 0.7 and 1.0, which are the same as the receiver functions shown in Figs 3(i) and (j). Each green line denotes the stacked (a) SRF or (b) PRF at each seismic station with the assigned Gaussian parameter.

**Figure S2.** PRFs and SRFs at every seismic stations. Each green line denotes the stacked PRF or SRF at each seismic station with Gaussian parameters of 0.5, 0.7 and 1.0. These green lines are the SRF or PRF at each seismic station with the assigned Gaussian parameter.

**Figure S3.** Depth migration of stacked receiver functions. (a) The 1-D $V_p$ model obtained by this study. Depth migration of stacked (b) SRFs at 0–20 s before the direct $S$ arrival and (c) PRFs at 0–20 s after the direct $P$ arrival with Gaussian parameters of 0.5, 0.7 and 1.0.
1.0. The depth migration is obtained with the 1-D $V_s$ model (a) and $V_p/V_s = 1.75$, following the method of Dueker & Sheehan (1997).

(d) Distribution of the piercing points at 180 and 160 km depths, calculated for the 1-D $V_s$ model obtained by this study (a). $P180s$ is $P$-to-$S$ conversion at a depth of 180 km. $S160p$ is $S$-to-$P$ conversion at a depth of 160 km.

**Figure S4.** Residual distributions before and after tomographic inversions. (a, c, e) TS-RTT residuals, (b, d, f) Rayleigh-wave phase-velocity residuals and (g) receiver function residuals before (grey lines) and after (red lines) the tomographic inversions. T: only inverting the TS-RTT residuals; R: only inverting the Rayleigh-wave phase-velocity data; TR: jointly inverting the TS-RTT residuals and the Rayleigh-wave phase-velocity data; TRF: jointly inverting the TS-RTT residuals, the Rayleigh-wave phase-velocity data, and the receiver function residuals. The corresponding RMS values of residuals before (grey numerals) and after (red numerals) the tomographic inversions are shown in each panel.

**Figure S5.** Distributions of the Rayleigh-wave real and imaginary component residuals at nine periods before (grey lines) and after (red lines) tomographic inversions. The corresponding RMS values of residuals before (grey numerals) and after (red numerals) the inversions are shown in each panel. The period is shown at the upper-left corner of each panel.

**Figure S6.** Changes of the RMS of all residuals with the number of iterations. The tomographic model is obtained by jointly inverting the TS-RTT residuals, the Rayleigh-wave phase-velocity data and the receiver function residuals.

**Figure S7.** Results of four RRTs for the 1-D $V_s$ model. Blue lines and squares denote input synthetic data. Red lines denote output results with initial models (grey lines). (a–d) Results of the first test obtained by jointly inverting the PRFs and SRFs and the Rayleigh-wave phase-velocity data. (e–h) Results of the second test which is similar to the first one. In this test, random noise is added to the synthetic data before the tomographic inversion. (i–l) Results of the third test obtained by only inverting the Rayleigh-wave phase-velocity data. (m–p) Results of the fourth test obtained by only inverting the PRFs and SRFs. The other labelling is the same as that in Fig. 7.

**Figure S8.** Results of eight synthetic resolution tests for the 1-D $V_s$ model. Results of the first (a–d), second (e–h), third (i–l), fourth (m–p), fifth (a’–d’), sixth (e’–h’), seventh (i’–l’) and eighth (m’–p’) tests. Blue lines and squares denote input synthetic data. Red lines denote output results with initial models (grey lines). These tests are conducted by jointly inverting the PRFs and SRFs and the Rayleigh-wave phase-velocity data, except for the eighth test (m’–p’) only inverting the PRFs and SRFs. The other labelling is the same as that in Fig. 7.

**Figure S9.** Results of five CRTs for the 2-D $V_s$ model obtained at different periods. Results of the first (a and 1–9), second (b and 10–18), third (c and 19–27), fourth (d and 28–36) and fifth (e and 37–45) tests. The input models are shown in (a–e). The output models are shown in (1–45). The period is shown at the lower-right corner of each map.

**Figure S10.** Results of an RRT for the 2-D Rayleigh-wave phase-velocity tomography at different periods. The input model is the same as that shown in Fig. 8. To simulate errors contained in the observed data, random noise with a normal distribution having a standard deviation of the synthetic data of 1 per cent is added to the synthetic real and imaginary component residuals before the tomographic inversion. The red and blue colours denote low and high phase velocities, respectively. The other labelling is the same as that in Fig. 8.

**Figure S11.** Results of four CRTs for the 3-D $V_s$ model. (a–f) The input models at different depths. (1–24) The output images at different depths. (1–6) Results of the first test obtained by only inverting the TS-RTT residuals (T Output). (7–12) Results of the second test obtained by only inverting the FMRWPV data (R Output). (13–18) Results of the third test obtained by a joint inversion of the TS-RTT residuals and the FMRWPV data (TR Output). (19–24) Results of the fourth test obtained by a joint inversion of the TS-RTT residuals, the FMRWPV data and the receiver functions (TRF Output). In these tests, the lateral grid interval is 0.5°.

**Figure S12.** Results of four CRTs for the 3-D $V_s$ model. The same as Fig. S11 but for the results of four tests with a lateral grid interval of 1.0°.

**Figure S13.** Results of four CRTs for the 3-D $V_s$ model. The same as Fig. S11 but for the results of four tests with a lateral grid interval of 1.0°.

**Figure S14.** Vertical cross-sections of $V_s$ tomography and results of corresponding RRTs for Cape Verde with different data sets. The same as Fig. 10 but for models determined by inverting different data sets. (1–8) The obtained $V_s$ tomography. (9–16) Results of the corresponding RRTs. (T) Results obtained by only inverting the TS-RTT residuals. (R) Results obtained by only inverting the FMRWPV data. (TR) Results obtained by a joint inversion of the TS-RTT residuals and the FMRWPV data. (TRF) Results obtained by a joint inversion of the TS-RTT residuals, the FMRWPV data and the receiver functions. The other labelling is the same as that in Fig. 10.

**Figure S15.** Map views of $V_s$ tomography for Cape Verde obtained with different data sets. (1–6) A $V_s$ model determined by only inverting the TS-RTT residuals (T). (7–12) A $V_s$ model determined by only inverting the FMRWPV data (R). (13–18) A $V_s$ model determined by a joint inversion of the TS-RTT residuals and the FMRWPV data (TR). (19–24) A $V_s$ model determined by a joint inversion of the TS-RTT residuals, the FMRWPV data and the receiver functions (TRF). The other labelling is the same as that in Fig. 9.

**Figure S16.** Vertical cross-sections of 3-D $V_s$ tomography obtained with different initial $V_p/V_s$ values. The other labelling is the same as that in Fig. 10.

**Figure S17.** Stacked S-wave receiver functions at various areas. (a) Map view of $V_p$ tomography at 60 km depth. Blue stars, grey squares and red triangles denote the piercing points of $S60p$ at 60 km depth, which are the same as the red circles in Fig. 3(h), calculated with the 1-D $V_s$ model obtained by this study (Fig. 7a). Blue stars are located in areas with low-$V_s$ anomalies ($\leq -2$ per cent), grey squares in areas with relatively normal $V_s$ anomalies ($-2$ per cent $< V_s < 2$ per cent), and red triangles in areas with high-$V_s$ anomalies ($\geq 2$ per cent). (b) SRFs with Gaussian parameters of 0.5, 0.7 and 1.0 sampling various areas are stacked respectively. The stacked SRFs are migrated with the 1-D $V_s$ model obtained by this study (Fig. 7a) and $V_p/V_s = 1.75$, following the method of Dueker & Sheehan (1997). The $V_s$ discontinuity at $60$ km depth (the G-discontinuity) seems to be clearer in areas with low-$V_s$ anomalies beneath the volcanic islands.

**Table S1.** Parameters used in the tomographic inversions.

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