Geophys. J. Int. (2003) **154,** 559–583

Simultaneous inversion for mantle shear velocity and topography of transition zone discontinuities

Yu J. Gu,* Adam M. Dziewoński and Göran Ekström

Department of Earth & Planetary Sciences, Harvard University, 20 *Oxford Street, Cambridge* MA 02138*, USA*

Accepted 2003 February 15. Received 2003 February 15; in original form 2002 April 10

SUMMARY

Until now, modelling of three-dimensional (3-D) velocity variations in the mantle and topography of the transition zone discontinuities have been considered separately. Velocity models were obtained assuming that the radii of the discontinuities are constant. Then, the travel time data sensitive to the topography, such as the *SS* precursors, were corrected for the effect of 3-D seismic structure and inverted for depth variations of a discontinuity. Such a procedure is unsatisfactory, as it may introduce artefacts that could significantly affect the topographic results; the opposite trade-off is less likely to introduce conceptually important changes in the velocity distribution but should also be considered. In this study we bring together the same set of *S*-velocity sensitive data as used by Gu *et al.* and combine it with a large set of differential travel times of *SS-S400S*, *S-S670S*, and direct measurements of *S400S–S670S*. We formulate the inverse problem in terms of the volumetric (3-D) and topographic (2-D) perturbations for both the 400- and 670-km discontinuities. The best-fitting model of the joint inversion significantly improves the variance reduction of *SS-S400S* and *SS-S670S* residuals. The velocity distribution in the resulting model, TOPOS362D1, is very similar to that in model S362D1 (with correlation coefficients >0.9 throughout the mantle), which indicates that lateral variations of discontinuity depths have only minor influence on global modelling of velocity. Important changes, however, have been made to the topography of the 400- and 670-km discontinuities with respect to those obtained earlier assuming an existing velocity model. The overall undulation of the 400-km discontinuity is considerably less than that reported by earlier global studies; in TOPOS362D1 its maximum variation does not exceed 12 km. The strong degree-1 component before the joint inversion has decreased, such that the correlation between the velocities above the discontinuity and the shape of the discontinuity itself has substantially diminished. Spatially, this result means a significant diminution in the strength of the earlier reported depression of the 400-km discontinuity under the Pacific. The power spectrum of the topography of the 670-km discontinuity has been enriched in long wavelength component, especially in degree 2. The range of depth variations is ± 18 km and its shape correlates well with the radially averaged velocity perturbations in the transition zone. At wavelengths greater than 1000 km, there is little correlation between the depth perturbations of the 400- and 670-km discontinuities. The topography of the 400-km discontinuity does not appear to be strongly influenced by thermal structures potentially associated with subduction processes and plumes. This implies that thermal influence on the olivine α- to β-phase transformation may not fully account for the observed depth variations; dynamical effects and potential variations in composition may be important near the top of the transition zone.

Key words: discontinuities, inversion, phase transitions, subduction.

1 INTRODUCTION

Efforts from seismology using reflected and transmitted waves are particularly instrumental in two areas: (1) development of local and global tomography of the Earth's velocity structure, and (2) study of depth variations of upper mantle discontinuities. The first area received much attention in recent years, as models of *P* and *S* speeds provided first-order evidence for the lateral variations in temperature, from narrow features such as the descent of lithospheric slabs and rise of mantle plumes to anomalies with horizontal dimensions exceeding 2000–3000 km. These results led to diverse

*Now at: Lamont Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA. E-mail: gu@ldeo.columbia.edu

interpretations of the mantle convection as a system, from whole-mantle to layered or partially layered system (e.g. Zhou & Clayton 1990; van der Hilst *et al.* 1991; Fukao *et al.* 1992). The behaviour of slabs near transition zone depths appears to be complex, suggesting that large-scale convection patterns may experience significant disruption in the top 900 km of the mantle (e.g. Gu *et al.* 2001b; Fukao *et al.* 2001).

Progress has been made during the last decade in the analysis of secondary reflected and converted waves. The improved imaging of discontinuity depths provided additional constraints on mantle temperature and dynamics. For example, Revenaugh & Jordan (1991) observed 20 km peak-to-peak topography on the 400- and 670-km discontinuities using multiple *SH* reverberations. Shearer & Masters (1992) analysed underside *SH*-component reflections from upper mantle discontinuities (*SS* precursors), and found a correlation between the regional depression of the 670-km discontinuity and the location of subduction zones. They attributed this correlation to possible flattening and accumulation of cold, subducted oceanic lithosphere near the widely believed post-spinel phase transformation (Navrotsky 1980; Ito & Takahashi 1989). Benz & Vidale (1993) analysed high-frequency precursors to $P'P'$ and suggested that a significant fraction of the jump in the physical properties associated with the 400-km discontinuity could occur within a relatively narrow zone (<4-km thick); this result enables an effective regional estimate for the Clapeyron slope of the pressure-induced olivine α- to β-phase transformation. More recent studies have improved the constraints on the lateral topographic variations of the transition zone discontinuities, both on a regional scale, e.g. near subduction zones (e.g. Kawakatsu & Niu 1994; Castle & Creager 1998; Gurrola & Minster 1998; Li *et al.* 2000b) and hotspots (e.g. Dueker & Sheehan 1997; Li *et al.* 2000a), and on a global scale (e.g. Gossler & Kind 1996; Gu *et al.* 1998; Flanagan & Shearer 1998; Chevrot *et al.* 1999).

Some issues regarding the origin and extent of the topography of transition zone discontinuities are still being debated. One of these is the correlation, or the lack of, between the depth perturbations of the 400- and 670-km discontinuities. It is generally believed that the Clapeyron slope of olivine α- to β-phase transformation is positive (Katsura & Ito 1989), which implies that the presence of a low- (or high-) temperature anomaly will cause a local uplift (or depression) of the 400-km discontinuity. The depth of the 670-km discontinuity is expected to behave in the opposite manner, owning to a negative Clapeyron slope of the post-spinel transformation (Ito & Takahashi 1989; Bina & Helffrich 1994). Hence the topography of these two discontinuities should be anticorrelated, assuming (1) thermal anomalies extend through the transition zone without changing sign, and (2) the effect of chemical heterogeneity is small. The result of Revenaugh & Jordan (1991) is consistent with a global anticorrelation. This observation is supported by some regional studies (e.g. Vidale & Benz 1992), but results from global studies of long-period waves have suggested a positive correlation (Stammler *et al.* 1992) or a de-correlation (Shearer 1993; Gu *et al.* 1998; Flanagan & Shearer 1998). A more recent debate centers on the range of the boundary perturbations on a global scale: studies using *SS* precursors (Gu *et al.* 1998; Flanagan & Shearer 1998; Gu & Dziewoński 2002) observe systematically greater variations in transition zone thickness ($\pm$16 km) than those found by Chevrot *et al.* (1999) using *P*-to-*SV* converted waves (<10 km). In addition to signals resulted from the inherent differences in data sensitivity, Fresnel volume and location, the effect of mantle velocity, particularly that of the transition zone, may be important.

Another key issue is the trade-off between velocity and topography of transition zone discontinuities. So far these two parameters have been mostly modelled independently: the velocity structure is obtained by assuming perfectly 'flat' (1-D) boundaries near 400 km and 670 km, and the topography is obtained by adopting *a priori* *P* or *S* velocity models to correct for mantle structure. These approaches are potentially flawed because (1) variations of $\pm$20 km in boundary topography (e.g. Shearer 1993) could influence the velocity structure, and (2) flaws in the velocity model due to trade-offs between velocity and discontinuity depths can, in the form of travel time corrections, bias the measurements of the latter.

This study proposes a solution to this problem by inverting for 3-D shear velocity and depths of transition zone discontinuities simultaneously. This approach effectively accounts for potential trade-offs between them and thereby improves both solutions. The resulting models allow us to examine the relationship between the depths of the 400- and 670-km discontinuities, as well as their association with large-scale thermal variations in the mantle. The resulting velocity variations also allow for a re-examination of the radial continuity of structures between the upper and lower mantle.

For convenience, we will refer to the two major transition zone discontinuities as the 400 and the 670. The depths of 400 and 670 km are chosen mainly because the Preliminary Reference Earth Model (PREM) (Dziewoński & Anderson 1981) is used as the reference model in our experiments. We will refer to Gu *et al.* (1998) and Gu *et al.* (2001b) as GDA98 and GDSE01, respectively.

2 DATA AND METHOD

Our first data set contains absolute travel times of *S*, *SS*, *ScS*, *SKS* and *SKKS*, and differential travel times of *SS-S*, *ScS-S*, *S-SKS* and *SKKS-SKS*. This data set is similar to that used by GDSE01 and provides important constraints for the structure of the lower mantle. The differential travel times are particularly reliable since they are much less sensitive to effects related to the earthquake sources and stations. The second data set contains nearly 30 000 of long-period mantle and body wave waveforms. These waveforms are sensitive to the structure at upper- and mid-mantle depths. The third data set consists of surface wave dispersion measurements of Ekström *et al.* (1997), which improves the resolution in the lithosphere and asthenosphere appreciably. The fourth set, and the most important data set for this study, is the differential travel times of *SS-SdS* and *S400S–S670S* (Gu & Dziewoński 2002). This data set is sensitive to the topography of the transition zone discontinuities, as well as the large-scale velocity variations in the top 2000 km of the mantle.

We use 362 normalized spherical B-splines to represent values on a 2-D surface. The spline nodes (Fig. 1a) are nearly equally spaced. The radial variations are represented by two sets of cubic B-splines (Fig. 1b) with a discontinuity at 670 km. This parametrization for the 3-D velocity structure is the same as that used by GDSE01 and allows for a direct comparison with results of that study. Two sets of spherical B-splines, with 362 nodes each, are used to represent the topography of the 400 and the 670. Since the depth perturbations of these discontinuities are likely dominated by spherical harmonic degrees $l < 10$ (e.g. GDA98), the density of the splines (with a lateral nominal resolution of $\sim$1000 km) should be sufficient to characterize the large-scale variations.

2.1 Theory

The travel time inversion approach for the 3-D velocity of the Earth is based on the theoretical formulations of Dziewoński (1984). We express the travel time t of *P* and *S* waves

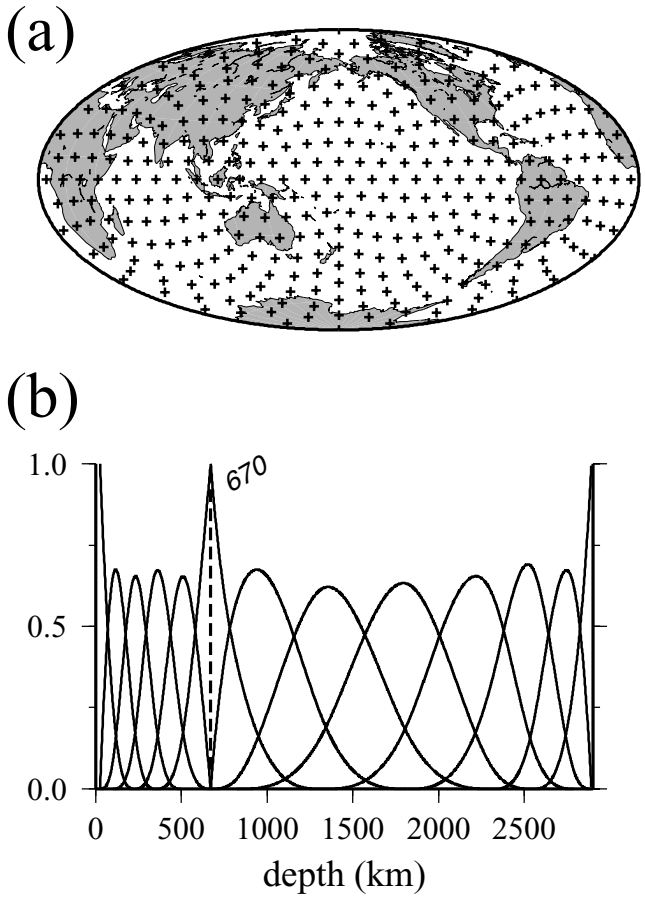

Figure 1. Basis functions (a) 362 spherical B-splines used in the lateral expansion of shear velocity and discontinuity topography, and (b) 14 cubic B-splines used in the radial expansion of shear velocity.

as

$$t = \int_{\text{path}} \frac{ds}{v}, \tag{1}$$

where ds is a segment of the path length along the ray and v is the wave speed. If we consider the effect of a depth perturbation δh to a mantle discontinuity, a perturbation in travel time δt can be expressed as the sum of the contributions from perturbations in shear velocity (δv) and depth of (δh_n) over the appropriate number of boundary interactions, that is

$$\delta t = -\int_{\text{path}} \frac{\delta v}{v^2} ds + \sum_n K_n \delta h_n, \tag{2}$$

where K_n is the travel time sensitivity of the ray to the depth perturbation to the nth discontinuity. If we expand the velocity and depth perturbations in B-splines, then

$$\delta t = \sum_k \sum_i C_{ki} \alpha_{ki} + \sum_n \sum_j D_{nj} \beta_{nj}, \tag{3}$$

where C_{ki} and D_{nj} are the unknown coefficients to spline expansions of velocity and discontinuity depth perturbations, respectively. The 3-D kernel α_{ki} can be calculated from

$$\alpha_{ki} = \int_{(\theta_0,\phi_0)}^{(\theta_1,\phi_1)} G[r(s)] B_k[r(s)] S_i[\theta(s), \phi(s)]\, ds, \tag{4}$$

where the integration path is from the source to the receiver. S_i represents the ith horizontal B-spline function, B_k is the radial function defined at the kth node, and $G[r(s)]$ is

$$G[r(s)] = -\frac{1}{v_0[r(s)]}, \tag{5}$$

where $v_0[r(s)]$ is the starting model velocity. For the second term in eq. (3), we consider kernels both for the reflected ($\beta_{\text{nj}}^{\text{Refl}}$) and for the transmitted ($\beta_{\text{nj}}^{\text{Trans}}$) waves at the transition zone discontinuities (Dziewoński & Gilbert 1976),

$$\beta_{\text{nj}}^{\text{Refl}} = -\sum_n \frac{2\left(\eta_{n-}^2 - p^2\right)}{r_n} S_j[\theta(s), \phi(s)],$$

$$\beta_{\text{nj}}^{\text{Trans}} = \sum_n \frac{\left[\left(\eta_{n+}^2 - p^2\right)^{\frac{1}{2}} - \left(\eta_{n-}^2 - p^2\right)^{\frac{1}{2}}\right]}{r_n} S_j[\theta(s), \phi(s)]. \tag{6}$$

S_j is the jth surface B-spline at a discontinuity n, r_n is the radius at the nth discontinuity, and the subscripts '+' and '−' represent the top and bottom of the discontinuity, respectively. Fig. 2 shows the travel time sensitivity of a shear wave, calculated from eq. (10) and (11) of Dziewoński & Gilbert (1976), to depth perturbations of the transition zone discontinuities. The difference in the slopes shows that a reflected wave is ten times more sensitive to a boundary perturbation than a transmitted wave; a major reason that the *SS* precursors are particularly useful in resolving boundary perturbations. The opposite signs indicate that, for example, a depression of a boundary will cause a reflected (transmitted) wave at this boundary to arrive earlier (later).

Eq. (3) leads to a convenient formulation for a simultaneous inversion of mantle velocity and topography of mantle discontinuities. The inverse problem can be written as:

$$\begin{pmatrix} \mathbf{A}_{3D} & \mathbf{B}_{400}^T & \mathbf{B}_{670}^T \\ \mathbf{A}_{3D} & \mathbf{B}_{400}^R & \mathbf{B}_{670}^T \\ \mathbf{A}_{3D} & \mathbf{B}_{400}^T & \mathbf{B}_{670}^R \end{pmatrix} \begin{pmatrix} \mathbf{m}_T \\ \mathbf{m}_{400} \\ \mathbf{m}_{670} \end{pmatrix} = \begin{pmatrix} \mathbf{d}_T \\ \mathbf{d}_{400} \\ \mathbf{d}_{670} \end{pmatrix}, \tag{7}$$

where $\mathbf{d}_T$ represents the non-precursor travel time data; $\mathbf{d}_{400}$ and $\mathbf{d}_{670}$ are travel time residuals of *SS-S400S* and *SS-S670S*, respectively. We correct for the effect of crustal thickness using CRUST5.1 (Mooney

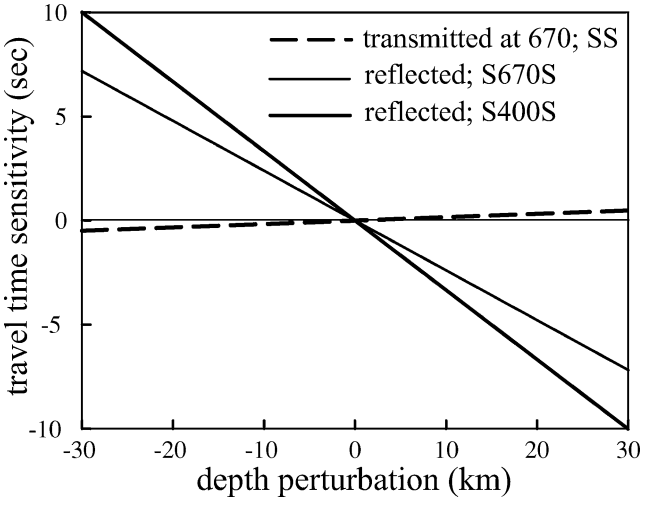

Figure 2. Travel time sensitivity of a shear wave to depth perturbations of the transition zone discontinuities. The dashed line shows the travel time sensitivity (130°) of a transmitted wave to depth perturbations of the 670. The thin and thick solid lines show the sensitivities of *S400S* and *S670S* to depth perturbations of the 400 and the 670, respectively, at the reflection point. The zero level is marked by a thin solid line. The travel time sensitivities of reflected and transmitted waves have opposite signs. Reflected waves are ten times more sensitive than transmitted waves to a depth perturbation.

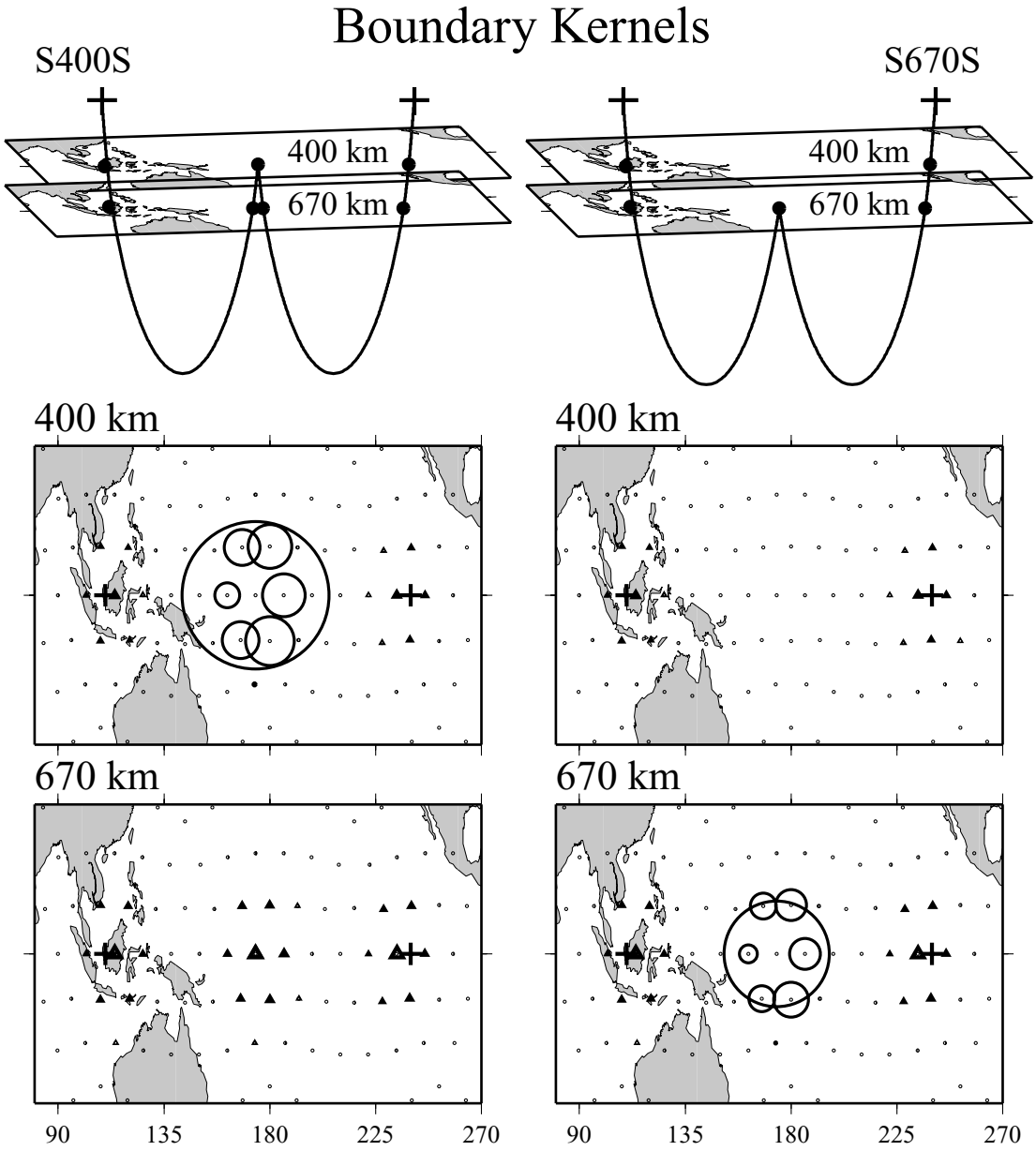

Figure 3. Travel time sensitivity of *S400S* and *S670S* In the top two panels the solid circles represent the spatial locations where rays intersect the transition zone discontinuities. The earthquake and station locations are marked by plus signs in all six panels. The bottom four panels show (in map view) the travel time sensitivity kernels due to reflections (circles) and transmissions (triangles); The S400S and S670S sensitivities are shown on the left and right, respectively. The small circles on the background show the locations of the B-spline nodes.

et al. 1998) and free surface topography using the ETOPO5 database. The unknown coefficients of shear velocities are represented by $\mathbf{m}_T$, and the coefficients of discontinuity depths are $\mathbf{m}_{400}$ and $\mathbf{m}_{670}$ for the 400 and the 670, respectively. $\mathbf{B}_d^T$ and $\mathbf{B}_d^R$ (d for a given discontinuity) are computed from kernels of the transmitted and reflected waves at d, respectively. Waveforms and surface waves are primarily used to constrain the 3-D variations in the mantle.

The travel time kernels can be readily computed from Dziewoński & Gilbert (1976). For waves that are not reflected by the upper mantle discontinuities, for example, *S*, the kernels include 3-D contributions from the mantle portion of the ray path (large) and 2-D contributions from the piercing points at these discontinuities (small). The primary contribution to an *SS*-*SdS* residual comes from the reflection points at a given discontinuity d, although other interactions of the ray with the discontinuities also need to be considered. Fig. 3 shows the travel time kernels for *S400S* and *S670S* due to reflections and transmissions at the 400 and 670. The circles represent the reflection kernels near the center of the ray paths. The sensitivity is low beneath the source and receiver locations (shown by triangles) where the waves transmitted through the discontinuities nearly vertically. It is necessary, however, to include these smaller contributions in our modelling since (1) they improve the global coverage considerably, and (2) their overall amplitudes become significant after summing all the contributions from transmitted waves such as *S*, *SS*, *ScS*, and *SKS*.

The kernels for the differential travel times, for example, *SS*-*S400S* and *ScS*-*S*, are computed by subtracting the kernels for the respective absolute residuals. The kernels for the differential travel

time of *S400S–S670S* are calculated by subtracting the *SS-S670S* and *SS-S400S* kernels.

2.2 The *SS* precursor data sets

The first step of our analysis is to re-examine the differential travel times of *SS-S400S* and *SS-S670S*. This is mainly motivated by the significantly improved data coverage obtained during the last five years. Our long-period *SS* precursor data set includes recordings collected by GSN, GEOSCOPE, MEDNET and other seismic networks, for earthquakes occurring between 1989 and 2000. The data selection criteria are the same as in GDA98. Fig. 4 compares the *SS* reflection points of GDA98 (Fig. 4a) with those of this study (Fig. 4b). Significantly denser coverage is evident in the previously under-sampled regions such as Africa, South America, and the southern Atlantic ocean.

Different phase equalization approaches have been introduced in earlier studies of *SS* precursors. A common technique for identifying the travel time and waveform of secondary phases is to use a well known phase that is orders of magnitude larger, e.g. *SS*, as a reference phase. In analysing the *SS* precursors, one could stack the original waveforms or the deconvolution/correlation functions that rely on the consistency between *SdS* and *SS*. We examine both approaches in this analysis to evaluate the robustness of our *SS-dS* measurements. In the first approach, we use filtered *SH*-component acceleration seismograms to form stacks at each cap location (Fig. 5a); this approach has the advantage of preserving the original signals in the seismograms. In the second approach (Fig. 5b) we correlate the waveform of *SS* with the rest of the seismogram for each record. A good overall correlation is expected between their waveforms due to the similarities between the underside reflections from the Earth's surface (*SS*) and those from mantle discontinuities (*SdS*), i.e. reflections from mantle discontinuities are expected to produce peaks in the correlation functions.

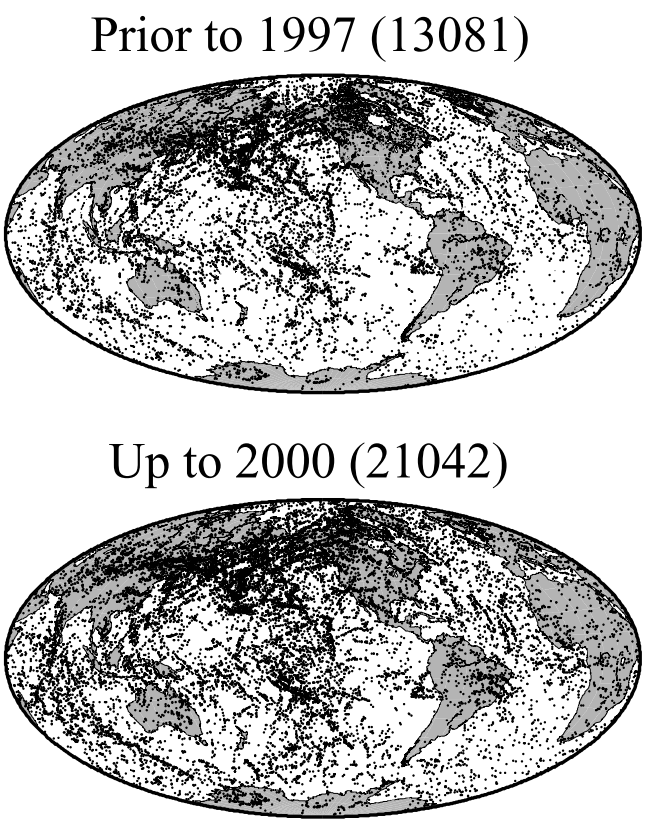

Figure 4. Improvement in data coverage. (a) Ray theoretical reflection points of 13 080 high quality *SS* precursors for data recorded between 1989 and 1997 (used by GDA98). (b) Ray theoretical reflection points of 21 042 high-quality *SS* precursors used by this study. The latter data set significantly improves the coverage in previously under-sampled regions such as the southern Atlantic and Africa.

A high correlation is observed between the resulting *SS-S400S* (Fig. 5c) and *SS-S670S* (Fig. 5d) residuals obtained by these two approaches. Both approaches obtain predominantly negative residuals for *SS-S400S* and positive residuals for *SS-S670S*, which imply a deeper 400 and a shallower 670 than PREM. In addition to corrections for the crust and surface topography, we have also accounted for mantle structure using S12 (Su *et al.* 1994), the starting model for the joint inversion procedure. Residuals obtained from the first approach show slightly larger topography on the 400 than those obtained from the second approach; this reflects greater complexities in the original waveforms. The overall agreement between these two approaches demonstrates that our measurements are generally robust. Fig. 6(a) and (b) show five correlation stacks for the final measurements of *SS-S670S* and *SS-S400S*, respectively. Each stack is formed by over 100 records and the highlighted phase is best aligned and ready to be measured; clear *S400S* and *S670S* are present in all stacks. Variations up to several seconds can be observed between the data and PREM synthetic arrivals, which suggest absolute depth variations exceeding ± 10 km. In addition to the 400 and the 670 reflections, there are peaks that correspond to possible reflectors between 230 km and 270 km in the sample caps in Indonesia and Eurasia (Ryberg *et al.* 1995; Gu *et al.* 2001a), and between the 400 and the 670 in the Pacific ocean (e.g. Shearer 1993; Gossler & Kind 1996; GDA98; Deuss & Woodhouse 2001).

The large-scale features of our topography maps, as obtained directly from corrected *SS-dS* times, are generally consistent with those of Flanagan & Shearer (1998). For the 400 both studies show a long-wavelength depression of the boundary in many parts of the Pacific ocean (Fig. 7a and b). This feature differs strongly from the Indian and Atlantic oceans where the boundary is either elevated or nearly unperturbed. Our new measurements of the 400 show slightly less variations than Flanagan & Shearer (1998). For the 670, both studies obtain a large-scale depression in the western Pacific, eastern Eurasia and South America (Fig. 7c and d). The amplitude (>20 km) and the lateral scales (3000–5000 km) of the perturbations are generally consistent between these two studies, e.g. the uplifts of the 670 in the Central Pacific and North Atlantic. The spectra of the topography maps are dominated by low degree spherical harmonics. Although little discussion was offered by Shearer (1993) and Flanagan & Shearer (1998) in the context of the long-wavelength degree signature, a dominating degree 1 (Fig. 7c) and a strong degree 2 (Fig. 7d) exist in their maps of the 400 and 670, respectively. The overall agreement between the results of Flanagan & Shearer (1998) and those of this study reflects the stability of the large-scale features of discontinuity topography obtained from the *SdS* data without using a joint inversion approach.

2.3 Additional assumptions

It is well known that due to the low amplitude of the *SS* precursors, travel times of *SS-dS* can only be obtained by stacking multiple records in a cap-averaging scheme (Shearer 1993; GDA98). For the purpose of a joint inversion, however, it is advantageous to

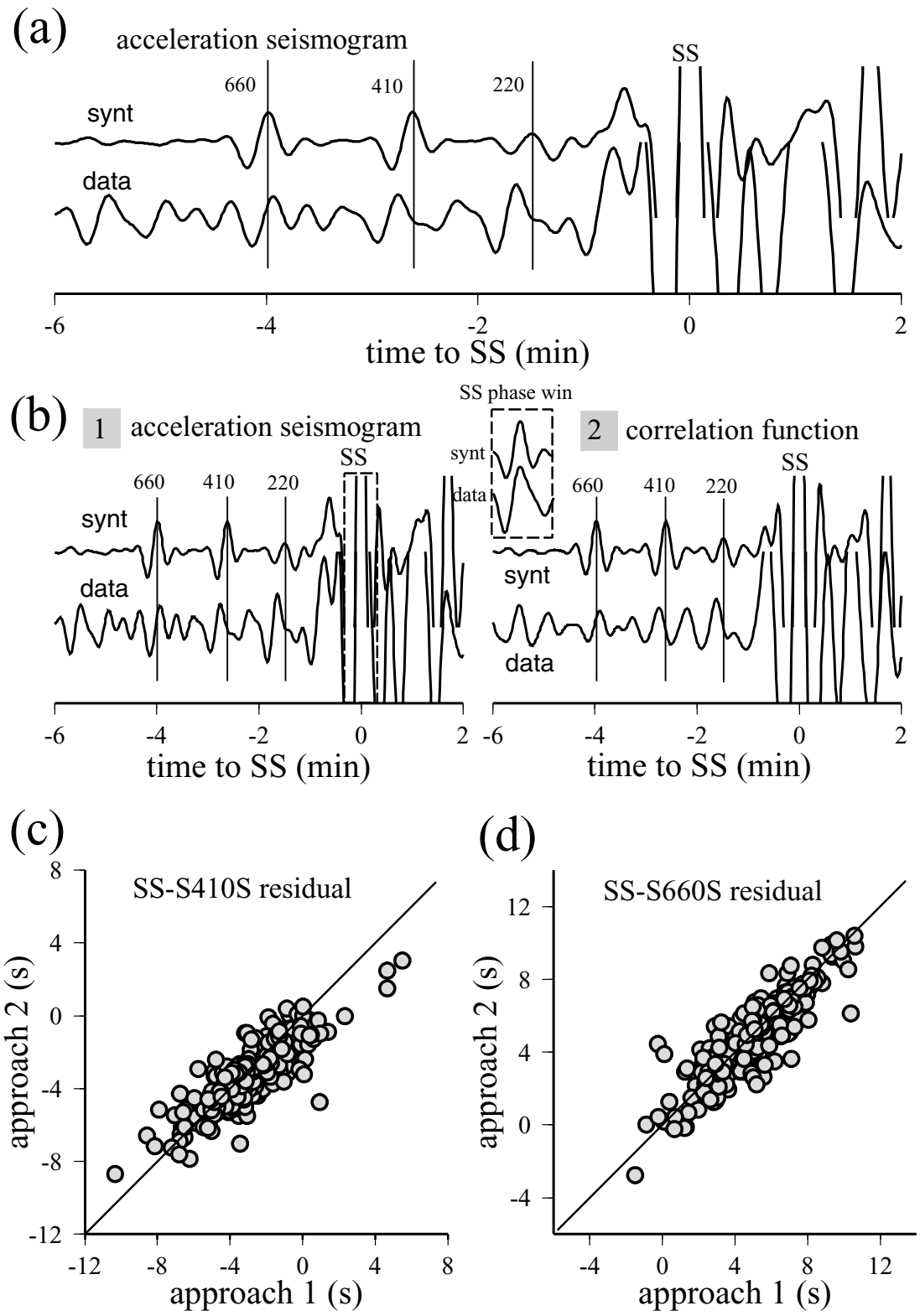

Figure 5. Stacking approaches. (a) A typical SH-component, band-pass filtered acceleration seismogram used in the stacking procedure. Similar seismograms are used to form stacks after appropriate travel time corrections; we refer to this approach as 'approach 1.' (b) Steps taken in obtaining the correlation functions for the second approach. In this approach, we select a phase window enclosing the *SS* waveforms and correlate with the respective seismograms. The correlation functions, shifted to by the length of the *SS* phase window, are shown in the right panel. The measurement approach which uses these correlation functions is referred to as 'approach 2.' (c) Correlation of the *SS-S400S* residuals (relative to those predicted by PREM) obtained by approach 1 and by approach 2. The straight line represents the case of a perfect correlation. (d) Similar to panel (c), but for the *SS-S670S* residuals. The consistency of the measurements obtained by these two approaches shows the robustness of the measurements.

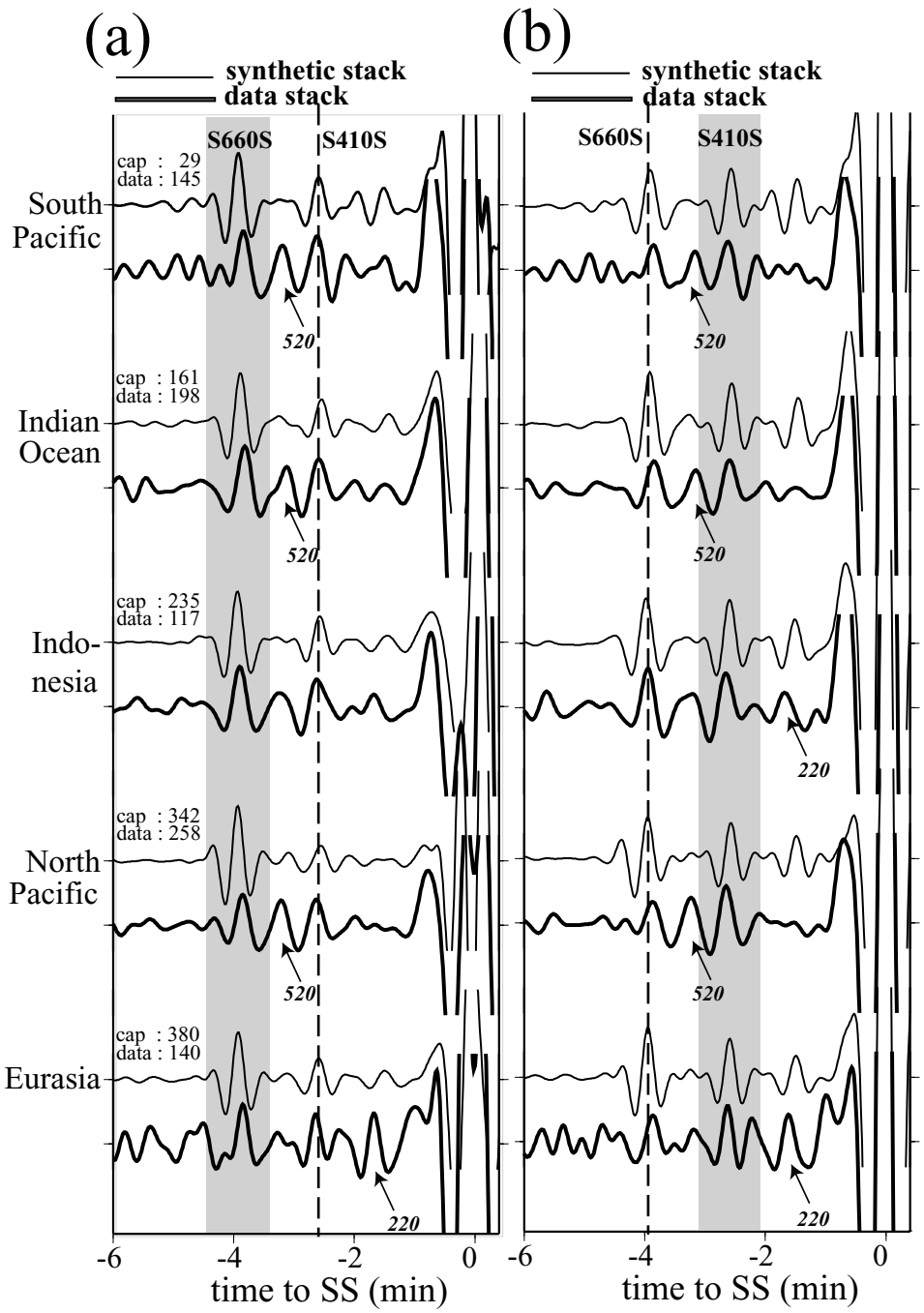

Figure 6. Sample stacks of data and synthetic seismograms. (a) Stacks of synthetic (thin line) and data (thick line) seismograms for five measurements of *SS-S670S*. The seismograms which form these stacks have been shifted in time using the Fourier transform to enhance the 670 reflection (shaded region). The dashed line shows the consistent arrivals of *S400S*. Small reflections due to the 220 and the 520 are as indicated. (b) Similar to (a), but for the *SS-S400S* measurements. The cap locations are the same as those shown in (a). Consistent arrivals of *S400S* and *S670S* make it possible to obtain global measurements of *SS-S400S* and *SS-S670S*.

have a travel time residual for each *SS* and *SdS* pair to extract the information of boundary perturbation and mantle heterogeneity along the path. This is achieved by weighing the residuals of adjacent caps to the reflection point of the ray: we determine the weight of a neighboring cap by the value of the spherical B-spline function (see fig. 1b of GDSE01) that is inversely related to its distance from the reflection point. The expanded data sets (not shown) are consistent with the original cap residuals, i.e. the pattern in the cap residuals is nearly perfectly recovered.

Furthermore, we scale the residual of each *SS* precursor and its corresponding kernels by the square root of the data count within the neighboring caps. This procedure accounts for the uneven distribution of the data. It should also be noted that the quarter-period travel time Fresnel zone associated with a 20 s *SS* precursor has an approximate width of 15–20° (e.g. Shearer 1993). Such large 'footprint' of the waves is insensitive to small-scale variations in velocity and discontinuity topography. It effectively applies an automatic averaging that is necessary for the study of the long-wavelength signals in

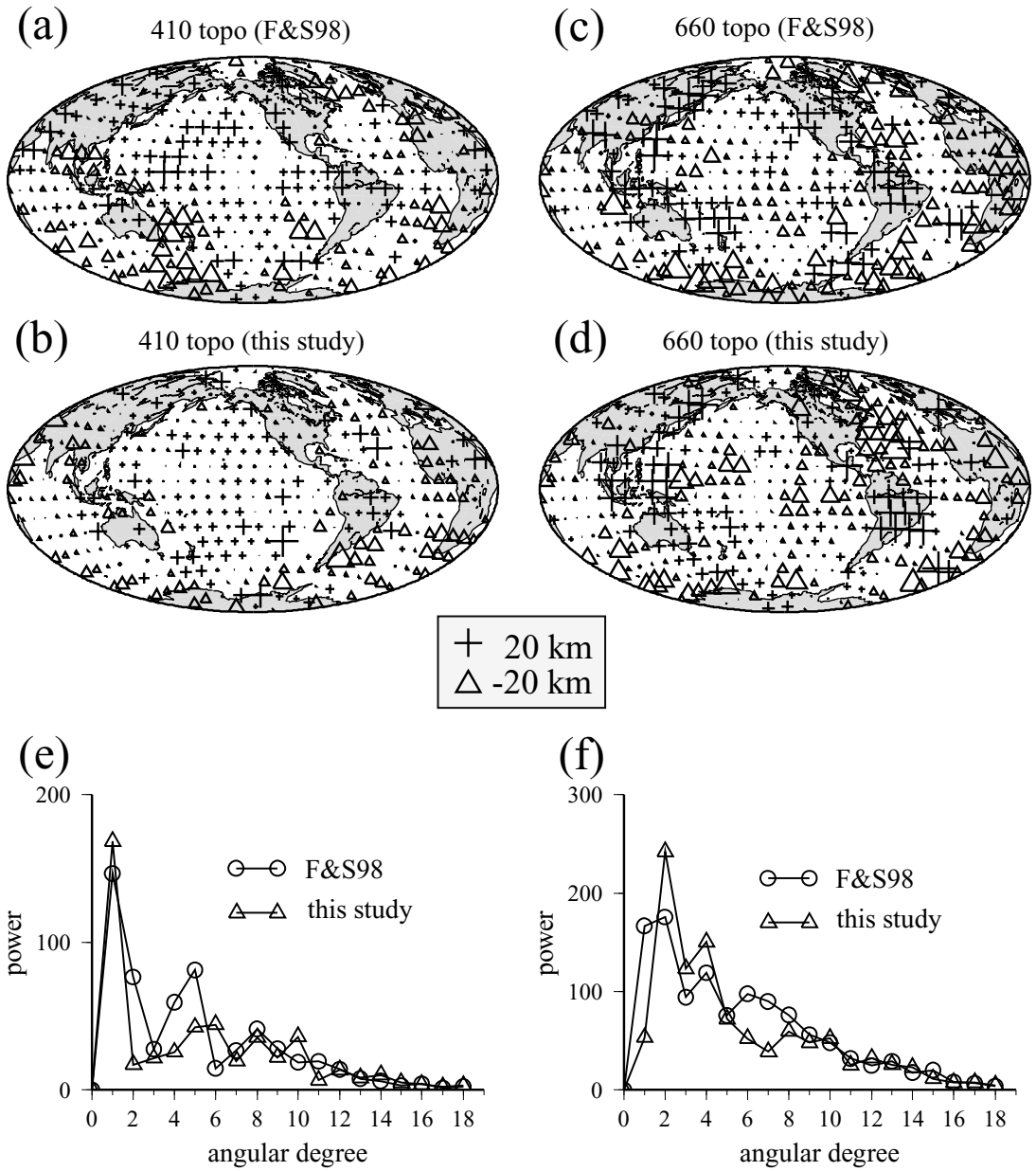

Figure 7. Topography of transition zone discontinuities after the joint inversion. The depths are obtained from the measurements of *SS-S400S* and *SS-S670S* residuals without jointly modelling mantle velocity. In each map triangles represent elevations of a boundary and crosses represent depressions of a boundary. (a) Topography of the 400 in Flanagan & Shearer (1998). (b) Topography of the 400 from this study; S12/WM13 (Su *et al.* 1994) is used to correct for mantle structure. (c) Topography of the 670 in Flanagan & Shearer (1998). (d) Topography of the 670 in this study. (e) Power spectra of the 400 which show a dominating degree-1 spherical harmonic. (f) Power spectra of the 670 which show strong signals at degrees 2 and 4. The spectrum of the 670 in Flanagan & Shearer (1998) contains a strong degree 1 which is not present in our map. Note: the residuals of Flanagan & Shearer (1998) have been plotted on the cap locations of this study.

the mantle. We also apply corrections for average depths of 410 km and 650 km to all the travel time residuals prior to the inversions; this accounts for the effect of the non-zero averages to the depths of discontinuities in the reference model.

2.4 Inversion and damping

The set of equations for the inversion can be written in the following form:

$$\mathbf{d} = \mathbf{A} \cdot \mathbf{x} + \mathbf{e}, \tag{8}$$

where $\mathbf{d}$ is a vector of observations, $\mathbf{x}$ is the model vector, and $\mathbf{e}$ is the vector of errors. Solutions can only be obtained from the formulation below (see Su *et al.* 1994, for detailed discussion):

$$\delta\mathbf{d}_0 = \mathbf{A}_0 \cdot \delta\mathbf{x}, \tag{9}$$

where $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$, $\mathbf{x}_0$ is the starting model, $\delta\mathbf{d}_0$ is the difference between the observation and the starting model prediction, and $\mathbf{A}_0$ is the respective kernel matrix. We impose horizontal and radial damping to the $(\mathbf{A}^T \cdot \mathbf{A})$ matrix of the 3-D velocities by minimizing the integrated squared gradients of the model (see GDSE01).

For the 2-D portion of the model, i.e. the topography of the 400 and 670, we define a depth perturbation δh as

$$\delta h = \sum_j D_j B_j(\Delta), \tag{10}$$

where $B_j(\Delta)$ is the jth horizontal B-spline and D_j is the unknown coefficient to this spline. The smoothness constraint is imposed as

$$\int_\Omega \|\nabla \delta h\|^2 \, d\Omega = \int_\Omega \left(\frac{\partial \delta h}{\partial \theta}\right)^2 + \left(\frac{1}{\sin(\theta)}\frac{\partial \delta h}{\partial \phi}\right)^2 d\Omega = \min, \tag{11}$$

where Ω represents the surface of a sphere (see GDSE01 for details).

3 RESOLUTION TESTS

One of the key components in assessing the reliability of the results from seismic tomography is resolution analysis. The solution to our inverse problem is strongly dependent on the data coverage, the weight of each data set, and the smoothing criteria. For the subsequent resolution tests we compute the synthetic travel time for each datum by tracing rays through the input velocity model with hypothetical perturbations to the transition zone discontinuities. The inversion algorithm, the data weights and smoothing are the same as those used in the inversion of the actual data sets.

In the first experiment we use a simple velocity model (3-D) that contains fast transition zone velocities in the western Pacific and South America, and slow velocities in the northern Pacific and Atlantic oceans (see Fig. 8a and b; left). These anomalies have similar dimensions and spatial locations as those frequently observed in the transition zone, and their amplitudes decay quickly at depths away from the transition zone. The input topography for the 400 (Fig. 8c; left) mainly consists of a long-wavelength depression under the Pacific; this is one of the key features of the observed topography prior to the joint inversion. We use a pattern similar to the transition zone velocity perturbations as the input topography for the 670 (Fig. 8d; left); this is motivated by the observed depressions in regions of fast velocities. The right hand column of Fig. 8 shows the recovered velocity and discontinuity topography. At 400 km, the input anomalies in the western and central Pacific are well recovered, both in lateral scale and amplitude, by the joint inversion. Only modest high-frequency artefacts are present, for example, in South America and the southern Atlantic; these artefacts are likely due to the coarser data coverage in the southern hemisphere. At 550 km, the overall resolution improves and the joint inversion nearly recovers 95 per cent of the input amplitude. The large amplitude of the input signal at this depth (5 times larger than that at 400 km) also contributes to the excellent recovery of the patterns. The topography of the 400 (Fig. 8c; right) contains most of the large-scale patterns in the input model, for example, the depression under the Pacific and the elevation under Africa. Only minor differences are observed between the input and output models, e.g. beneath South America and the Polar regions. The recovered topography of the 670 (Fig. 8d; right) also correlates well with the input model. Hence the data and the joint inversion approach can simultaneously resolve, with confidence, most of the large-scale velocity structure and boundary perturbations.

In the second test we introduce a non-zero 3-D structure that contains fast transition zone velocities under the western Pacific and South America, and slow velocities in the northern Pacific and northern Atlantic; we assume zero topography on the 400 and the 670. The goal is to examine the influence of velocity structure on boundary undulations. Fig. 9a (left panel) shows the input velocity map at 550 km. The anomalies are spatially limited to the transition zone and the amplitude drops to zero at 700 km (Fig. 9b; left). The blank maps shown by the left-hand-side panels of Fig. 9(c) and (d) indicate the absence of topography on the 400 and 670. The recovered 3-D velocity pattern at 550 km (Fig. 9a; right) shows a reasonably good correlation with the main velocity features of the input model. The location and lateral extent of the anomalies are well resolved except in South America and the northern Atlantic ocean. This is partly related to the low density of rays in the southern hemisphere and, in part, to lateral and radial smoothing of the velocities in the inversion procedure. The effect of radial smoothing is evident in the right panel of Fig. 9(b) where the velocity becomes nonzero at 700 km. The resulting depth variations vary between -1 and 1 km for the 400 and 670 (Fig. 9c and d; right), solely due to trade-offs between velocity and topography. The shape of the 400 is similar to the input velocity pattern at 550 km (see Fig. 9c; right), and yet the output topography of the 670 (Fig. 9d; right) is clearly anti-correlated with the input velocity pattern. The nonzero topography in the output model can be understood with the help of the schematic diagrams in Fig. 9(e); for brevity we only use a high velocity anomaly as an illustration. Only *S400S* and *S670S* are shown since they have the highest sensitivity, and thus contribution, to the velocity and boundary topography near the transition zone. Because of lateral and radial smoothing, the resulting transition zone velocity structure of the resolution test cannot fully recover that of the input model near 400 km, as demonstrated in Fig. 9a (right). This will decrease the speed at which *S400S* travels inside the transition zone. To compensate for the effective loss of speed, a joint inversion tends to depress the boundary near the reflection point; this allows *S400S* to reflect at a deeper depth and, consequently, arrive at an earlier time (see Fig. 9e; right). The opposite mechanism takes place at depths below the transition zone. The radial smoothing causes the velocity structure to extend beyond the 670, which leads to a faster *S670S* travel time than that expected from a zero model. To offset the net gain in speed, the inversion procedure tends to uplift the 670 and effectively slows down *S670S*.

The experiment shown in Fig. 9 provides a validity check on the expected relationship between velocity and discontinuity topography during the joint inversion procedure. It indicates that the undesired trade-offs between the boundary topography and velocity structure are relatively small: a 1 per cent velocity perturbation only introduces errors on the order of ± 1 km in the topography of transition zone discontinuities. The size of such errors, induced by smoothing and other modelling effects, is an order of magnitude smaller than those resulted from the joint inversion using the 'true' data (± 10 to ± 20 km). More importantly, as will be shown in the next section, the observed depth variations of the transition zone discontinuities are generally associated with the velocity perturbations in the opposite manners as those shown by Fig. 9(e).

4 RESULTS OF THE JOINT INVERSION

4.1 Velocity perturbations

We obtain the best-fitting model TOPO362D1 from the joint inversion of mantle velocity and the topography of transition zone discontinuities. The average depths of the 400 and the 670 are 409 km and 649 km, respectively; both are $\sim$0.5 km shallower than the global averages obtained without using the joint inversion approach. Fig. 10 compares the 3-D portion of TOPO362D1 (Fig. 10a) with S362D1 (Fig. 10b); the latter is a model obtained by an

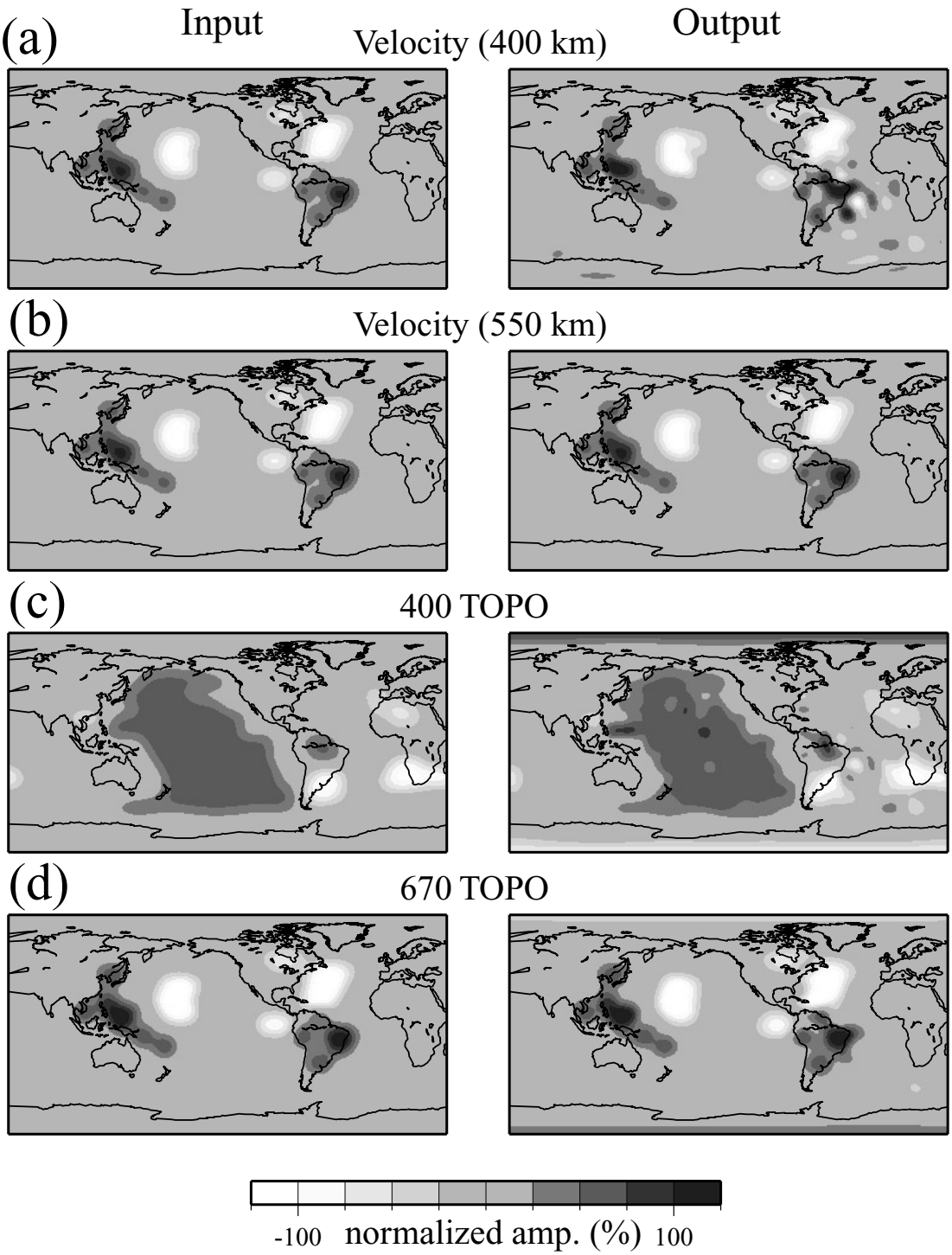

Figure 8. A resolution test with non-zero input velocity and topography. Light colours represent low velocities or elevations of a discontinuity and dark colours represent high velocities or depressions of a discontinuity. The left column shows the input models and the right column shows the resulting models of the resolution test. The maps are: (a) velocity structure at 400 km, (b) velocity structure at 550 km, (c) topography of the 400, and (d) topography of the 670. The input patterns are generally well resolved by our data sets and modelling approach.

independent shear velocity inversion using similar data sets excluding the *SS* precursors (GDSE01). Both models show significant high velocities under continents and low velocities under mid-ocean ridges at 100 km, indicating strong thermal variations in the shallow mantle. The only visible difference between these two models is the amplitude of the high-velocity anomalies under the northern Pacific ocean. At 350 km, TOPO362D1 still shows traces of imprints of surface tectonics, though the fast velocities under continents are visibly less pronounced than those of S362D1; this could, in part, result from differences in the weighting of the data sets due to the inclusion of *SS*-*dS*. At 550 km, TOPO362D1 shows a weaker velocity anomaly in South America than S362D1, indicating that a fraction of the

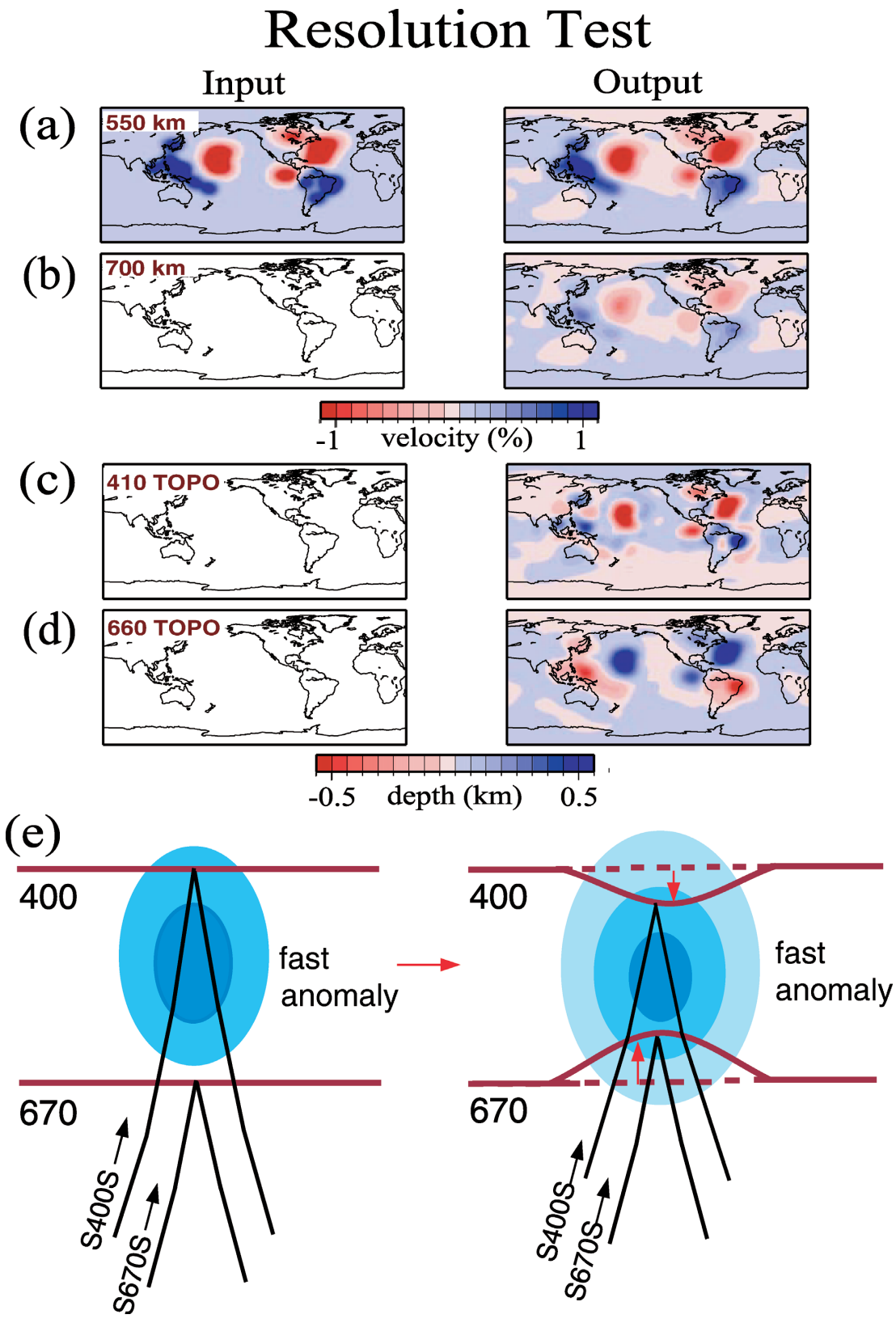

Figure 9. A resolution test with zero input topography. (a) A resolution test with input (left) and recovered (right) velocities at 550 km. The blue colours show fast velocities and the red colours show slow velocities. (b) The input and recovered velocities at 700 km. The blank map indicates a zero input velocity model. (c) Topography of the 400. The input model contains no assumed topography on the 400. (d) Topography of the 670. (e) A schematic diagram demonstrating the relationship between the input and recovered models. For simplicity a fast velocity anomaly is used.

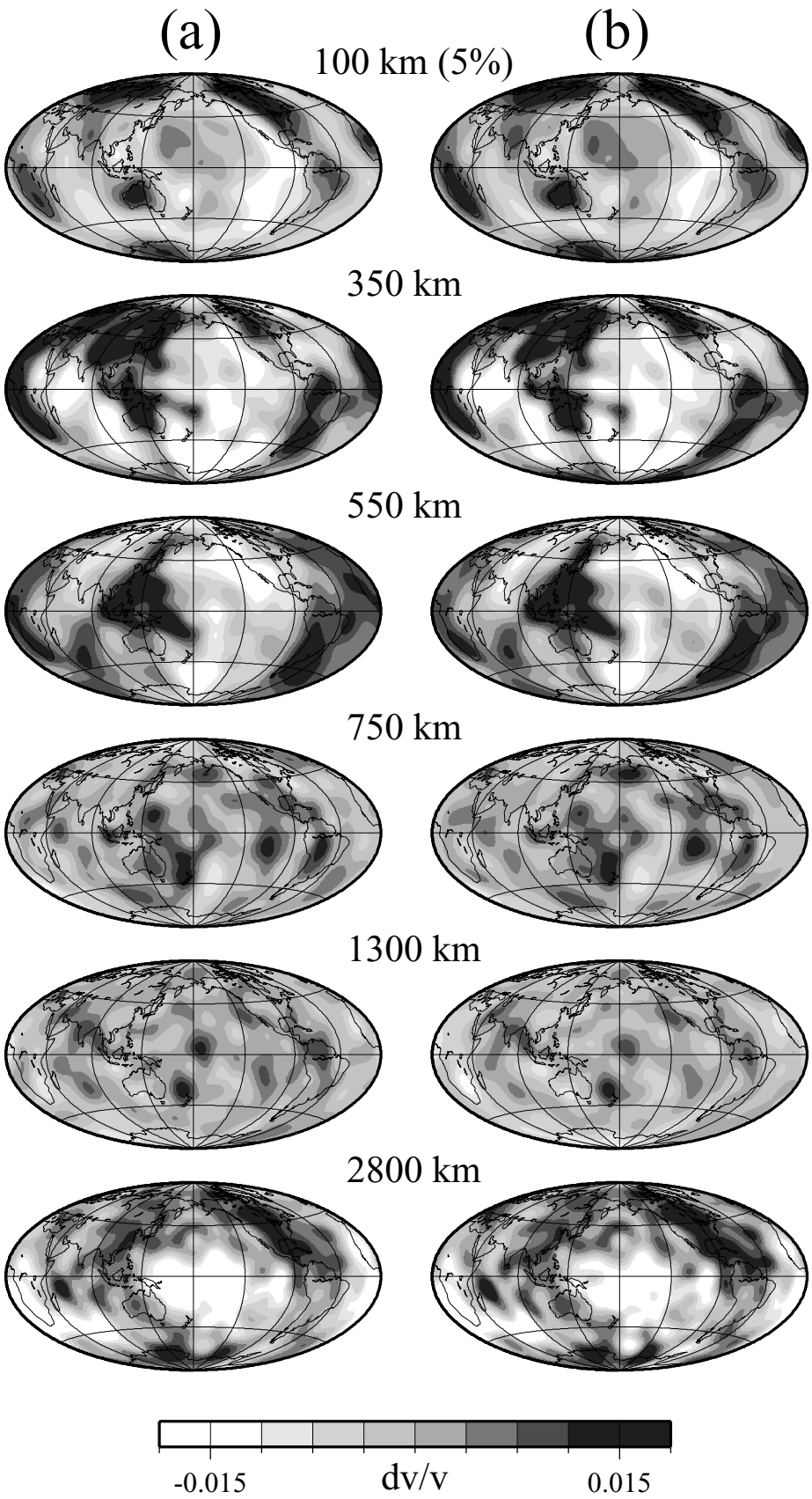

Figure 10. Shear velocities in the mantle. The velocities are obtained by (a) TOPO362D1 (joint inversion of this study), and (b) S362D1 of GDSE01. The positive values represent faster than average velocities. A good overall correlation is observed between these two models, which suggests that discontinuity topography has only second order effects on the velocity structure.

velocity anomaly could be explained more effectively by topographic variations on the discontinuities. The mid-mantle velocity perturbations of TOPO362D1 are much weaker at 750 km and 1300 km than at the upper mantle depths. The decrease in amplitude has been associated with a possible disruption to the large-scale flow pattern near the 670 (e.g. Liu 1997, GDSE01). Both models show significant lateral heterogeneities in the lowermost mantle, e.g. 2000–3000 km wide fast anomalies in the Pacific and Africa and a fast velocity ring around the circum-Pacific (e.g. Su *et al.* 1994; Li & Romanowicz 1996; Masters *et al.* 2000).

The velocity variations of TOPO362D1 are consistent with those obtained by modelling velocity alone. For further comparisons we invert for 3-D velocity variations alone using the same data sets (including the *SS* precursor travel times) and smoothing criteria. The resulting model shows nearly perfect correlation with the 3-D part of TOPO362D1 (>0.95) throughout the mantle. The lowest correlation is in the middle of the transition zone; this is expected since transition zone velocity structure has the highest trade-offs with the topography of the 400 and the 670. In general the topography of discontinuities has only minor overall influence on the 3-D structure of the mantle.

4.2 Topography of the 400 and the 670

Fig. 11 compares the topography of the transition zone discontinuities, obtained from a simple spherical harmonic expansion of the residuals in Fig. 7(a) and (b) (hereafter 'topography before inversion', Fig. 11a and e), with those obtained by two different inversion approaches. Fig. 11(b) shows the results from the first inversion approach – an independent inversion for the topography of the 400 using only the *SS-S400S* times; this is useful as a benchmark for evaluating how well the *SS-S400S* residuals can be explained by the spline parametrization without considering potential influences from mantle structure, depth of the 670, and variations in transition zone thickness. This model shows a long-wavelength depression of this boundary under most of the Pacific and uplifts under Asia and South Africa. These features are in excellent agreement with the 400 topography before inversion, which demonstrates that the model parametrization and data sampling are adequate in recovering the main topographic features of the 400. Results from the joint inversion (Fig. 11c) show a considerable decline in the range of depth variations, most notably in the northwest and central Pacific. Coupled with significant elevation of the 400 under the East Pacific Rise, the long-wavelength depression of the Pacific (the most notable feature before inversion) becomes less significant after the joint inversion. This is likely a robust feature, having known that long-wavelength topography of the 400 in the Pacific ocean can be sufficiently resolved by our data sets and inversion approach (see resolution test in Fig. 8c). The diminished topography implies that a considerable fraction of the Pacific depression (e.g. GDA98; Flanagan & Shearer 1998) can be either explained by perturbations to the velocity structure or, possibly, by trade-offs with the 670 topography. The spectrum of the 400 topography after the joint inversion (Fig. 11d; shown in red) shows a notable decrease in the power of degree-1 spherical harmonic. Although degree 1 is still the strongest signal in the spectrum, its dominance over other degrees, particularly degrees 4 and 5, is not as pronounced as that observed in earlier studies (e.g. GDA98).

A near-perfect recovery of the initial 670 topography (Fig. 11e) is achieved by an inversion without 3-D structure and the topography of the 400 (Fig. 11f). In comparison, results from the joint inversion of velocity and topography (Fig. 11g) show a slight increase in the lateral scales of the main anomalies. This is evident in the western Pacific and South America where the boundary is depressed, and in the northern Pacific and the Atlantic oceans where the boundary is elevated. Smaller-scale depressions in the Pacific ocean (see Fig. 11e and f), with lateral dimensions of $\sim$1000 km, are generally replaced by low-amplitude elevations after the joint inversion (see Fig. 11g). The power spectrum of the 670 from TOPO362D1 shows a stronger degree 2 (Fig. 11h; shown in red) in comparison with the other maps. Despite the slight decrease in amplitude, the large-scale depressions (e.g. the western Pacific) appear to be robust features that cannot be explained by trade-offs with velocity nor with the topography of the 400.

It is worth noting that the depths of the 400 and the 670 after the joint inversion do not appear to be strongly influenced by the mechanism shown in Fig. 9(e). This is apparent in the large-scale depression of the 670 in regions of fast shear velocity: results from the joint inversion generally show a depressed, rather than an elevated, 670 (see Fig. 9). Hence artefacts resulted from our modelling approach and limited data distribution are generally small.

4.3 Model evaluation

Model TOPO362D1 reduces the data variance of the transmitted waves by 30 per cent or more for each data set (Fig. 12a). It shows marginal improvements over S362D1 in fitting the *SKS-S* and *SKS-SKKS* travel times, while both B-spline models significantly improve the fit to all the included data sets over S12. This improvement results from a combination of factors including differences in the number of free model parameters, data coverage and quality, and the fact that S12 was obtained by an inversion using different data sets. By considering both velocity and discontinuity topography, TOPO362D1 explains more than 60 per cent of the data variance in the travel times of the *SS* precursors (Fig. 12b); the fit to the *SS-S400S* and *SS-S670S* data sets is considerably higher than that achievable by S362D1 or S12. Considering the modest increase in free model parameters (724) and the considerable increase in the number of data ($>54\,000$ measurements), a high reduction of data variance demonstrates the robustness of the joint inversion as well as a great deal of consistency among the various travel time data sets. These are the important reasons that the joint inversion approach is more desirable than separate modelling of velocity and boundary topography.

The spectrum of TOPO362D1 (Fig. 12c) shows high amplitudes at low degrees (1–3) in the upper and lowermost mantle; this is consistent with results from earlier studies of normal modes and shear velocities (e.g. Masters *et al.* 1982; Dziewoński 1984; Su *et al.* 1994; Li & Romanowicz 1996; Masters *et al.* 2000, GDSE01). However, the amplitudes of most harmonics, particularly degree 2, drop substantially in the upper mantle transition zone. The root-mean-squared velocities (RMS; Fig. 12d) of the lateral heterogeneities are generally consistent between TOPO362D1 and S362D1. Modest differences exist between TOPO362D1 and S12 in the upper and lowermost mantle. Fig. 12(e) shows a high correlation between TOPO362D1 and S362D1 (0.9 or better) at all mantle depths, but the correlation between TOPO362D1 and S12 is much lower, particularly at mid- and lower-mantle depths.

A key issue to be investigated is the sensitivity of our modelling result to the choice of reference depths of the 400 and 670. Since PREM fixes the discontinuities at 400 km and 670 km, biases could potentially be introduced when we solve for perturbations to these

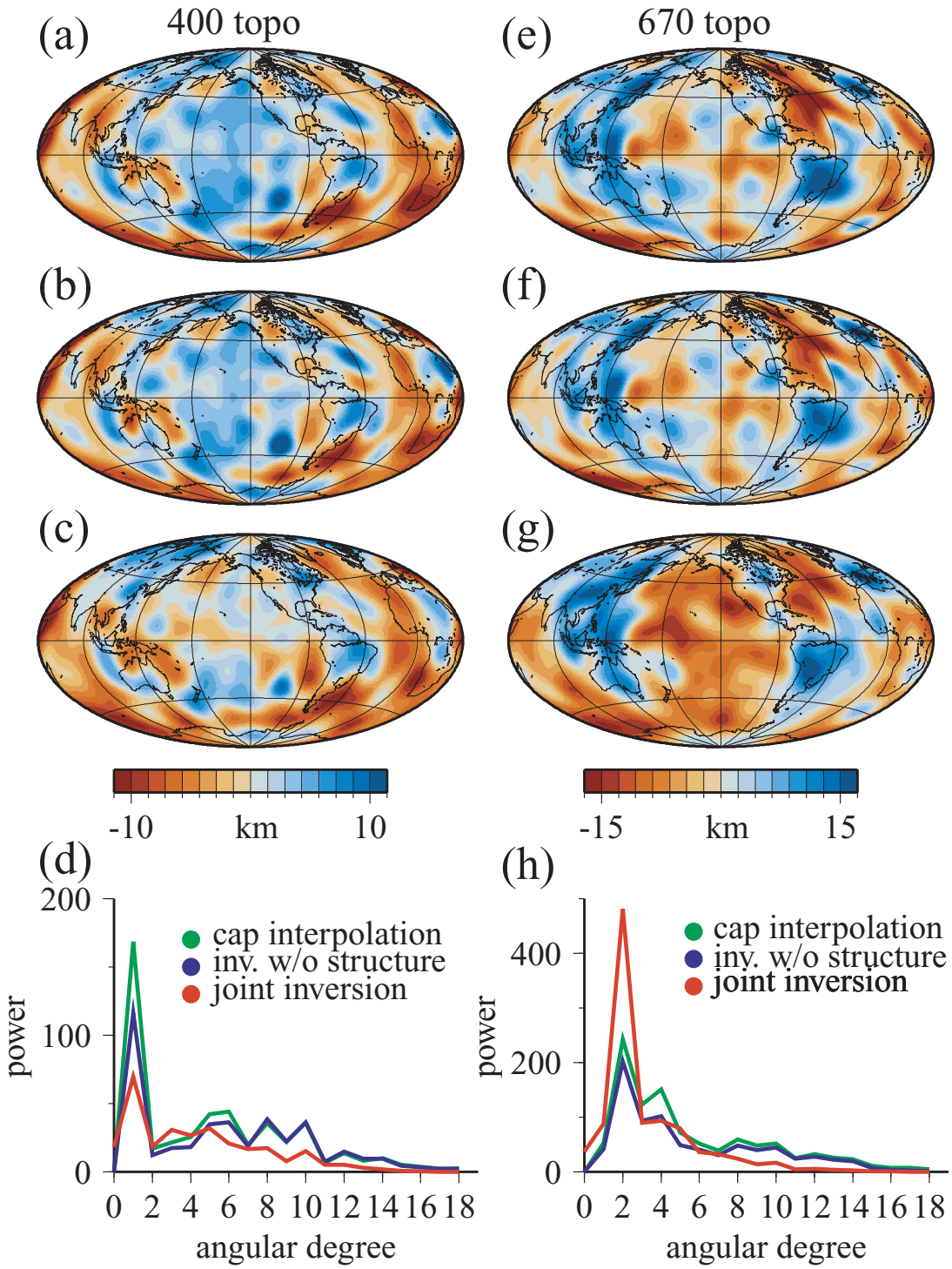

Figure 11. Discontinuity topography and power spectra. (a) Topography of the 400 obtained by expanding the residuals in Fig. 7(a) up to the degree 18 spherical harmonic. (b) Inversion for the 400 topography using the *SS–S400S* data set only. The velocity structure and the undulations on the 670 are ignored in this inversion. (c) Resulting topography of the 400 from the joint inversion of velocity and discontinuity depths. (d) A comparison of the power spectra of maps shown in (a)–(c). (e)–(h) Similar to figures (a)–(d), but for the topography of the 670.

depths. To address this issue we performed a series of inversions with reference discontinuity depths ranging from −10 km to +20 km from the PREM values. The correlations between TOPO362D1 and models obtained by assuming different starting depths are extremely high: In the case of the 400 the correlation exceeds 0.98, both for the structure and for the topography of the discontinuities; marginally lower correlations (>0.94) are achieved when we modify the reference depth of the 670. The excellent correlations suggest that the modelling result of this study is nearly insensitive to the choice of starting discontinuity depths.

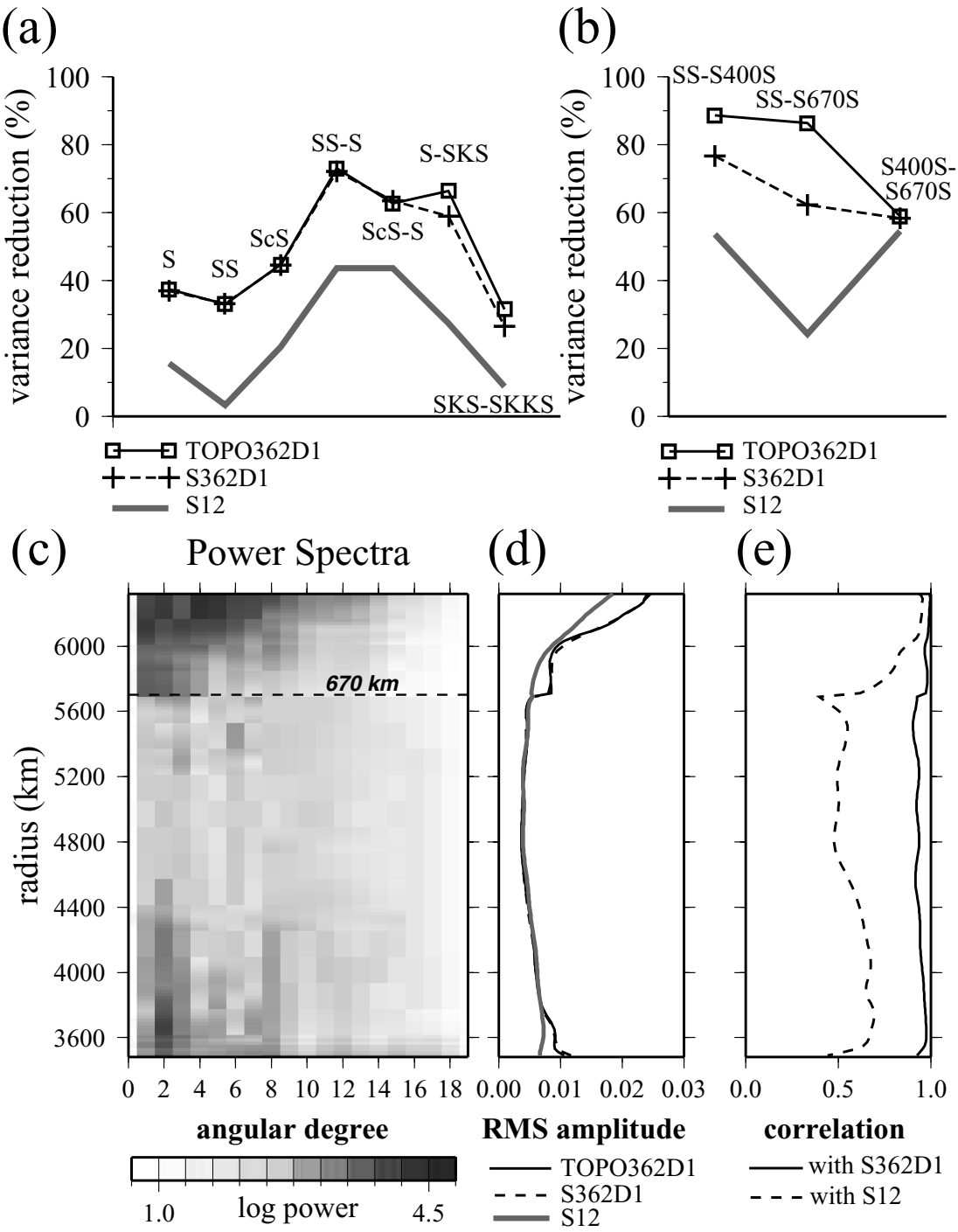

Figure 12. Variance reduction and model analysis. (a) Reduction of data variance of the transmitted waves through the transition zone discontinuities. The models are S12 (Su *et al.* 1994), S362D1 (GDSE01), and TOPO362D1 (joint inversion of this study). (b) Reduction of data variance to the *SS* precursor data sets. TOPO362D1 improves the fit to *SS-S400S* and *SS-S670S* significantly. (c) Logarithmic power spectra of model TOPO362D1. (d) A comparison of the root-mean-squared (RMS) amplitude of the lateral heterogeneities. (e) Correlations of TOPO362D1 with S362D1 (solid line) and with S12 (dotted line).

4.4 Cross-sections

To examine the regional features of TOPO362D1 we resort to cross-sections that focus on the subduction zones. For each cross-section, the velocity model is plotted on the background and the depths of discontinuities are represented by black lines; the amplitude of the topography is magnified by a factor of 3. The white lines show the globally averaged discontinuity depths.

Fig. 13 shows two sample cross-sections of TOPO362D1 through the northwest Pacific. The left panel shows well-defined fast velocity anomalies beneath the northern Japan. The dip of the anomalies closely follows that of the seismicity in this region. Fast velocities near the top of the transition zone appear to extend down to the 670, where a local boundary depression is observed. The topography of the 400 is generally small; only a slight regional elevation is visible. In southern Japan the fast anomalies broaden within the transition

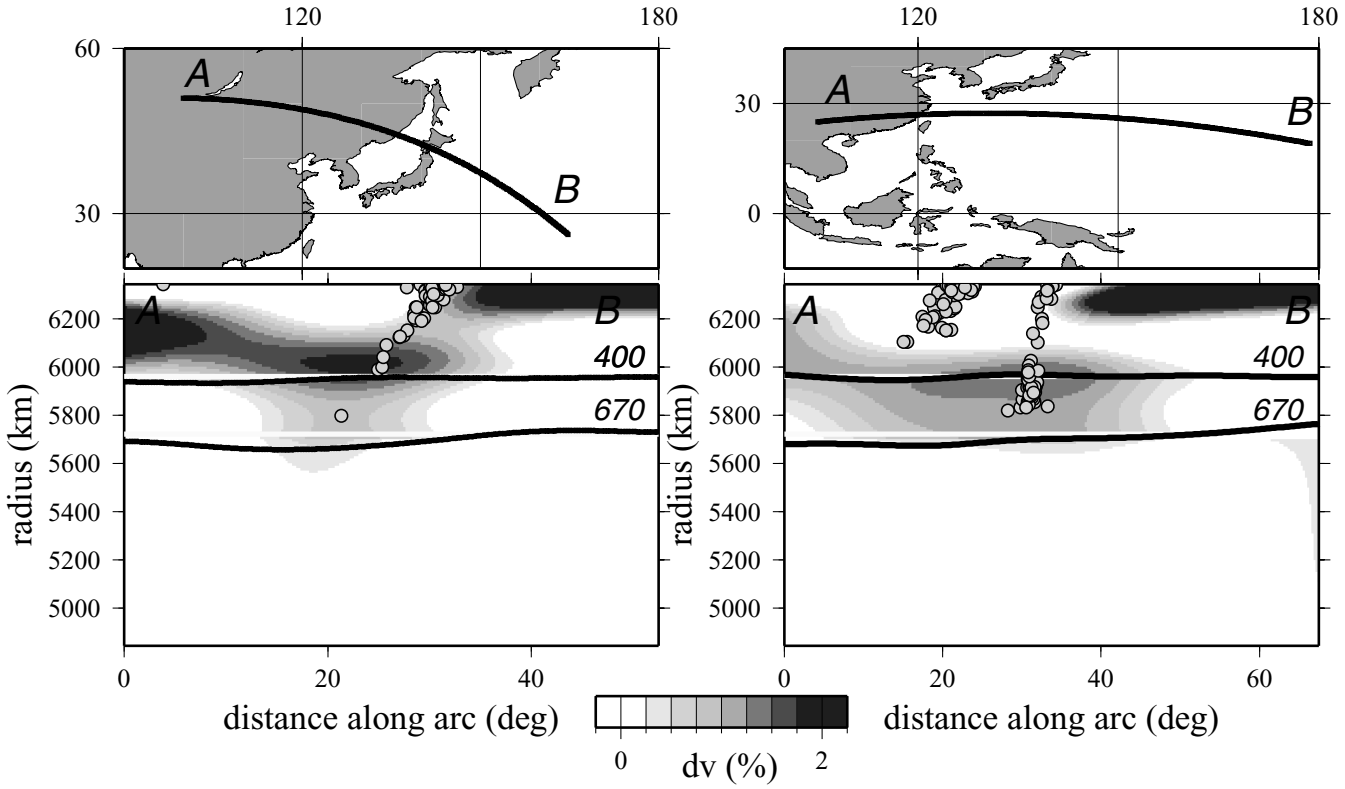

Figure 13. Regional cross-sections of TOPO362D1 through Japan. The background contours show the shear velocity variations predicted by TOPO362D1. The thick black lines represent the depth perturbations of the 400 and 670 along the profile. The white lines indicate the globally averaged depths of 409 km and 649 km, respectively. The depth perturbations have been exaggerated by a factor of 3. The circles represent the locations of earthquakes (taken from the Harvard CMT-catalog) within a distance of 5° on either side of the cross-section.

zone (right panel), and the amplitude increases appreciably. The location and lateral dimension of this structure are supported by a 2000–3000 km wide depression of the 670. If the fast anomalies are indeed associated with the subduction process, their lateral extent may reflect some horizontal flattening and accumulation of subducted lithospheric material on top of the lower mantle (e.g. Zhou & Clayton 1990; van der Hilst *et al.* 1991; Fukao *et al.* 1992; Shearer & Masters 1992; Castle & Creager 1998). The main fast velocity anomaly does not appear to extend below the 670.

Further south, the deep seismicity clearly suggests the presence of subducted oceanic lithosphere beneath the Mariana Islands, New Guinea, and Indonesia (Fig. 14). In profile A–A1 the high velocity anomalies seem to be mostly confined within the transition zone. The main anomaly extends nearly horizontally to the west; only a small portion of it appears to sink below 700 km. Profiles B–B1 and C–C1 show a high velocity anomaly in the top 300–400 km – potentially associated with the deep continental root beneath western Australia. Note that although profile C–C1 shows fast velocities well below 700 km, the scale of the anomaly in the transition zone is significantly larger than that in the lower mantle. The strong amplitude difference implies that an unimpeded slab penetration is unlikely in these regions. The extensive fast velocities within the transition zone also appear to cause a notable depression of the 670. The topography of the 400 is, however, much more difficult to interpret – only modest regional elevations are observed.

Two cross-sections through the Tonga-Kermadec islands are shown in Fig. 15(a). Profile A–A1 shows significant fast velocities below the Tonga trench and the structure dips westward. This anomaly appears to extend into the lower mantle, possibly down to 1000–1200 km (e.g. Grand *et al.* 1997; van der Hilst 1997). Both the 400 and the 670 experience a slight depression in the region of fast velocity. The small topography of the 670 in this region is consistent with slab penetration rather than subhorizontal deflection/accumulation in the transition zone since the latter should cause stronger and wider depressions on this discontinuity. The A–A1 profile through South America (Fig. 15b) shows a strong anomaly under the South American continent at depths above 200 km. In the transition zone the deep earthquakes fall nicely within a contour of fast velocity; this anomaly appears to be confined within the top 800 km. The apparent depression of the 670 corroborates, in view of an endothermic phase transition, with the presence of transition zone fast velocities. Similar velocity and depth perturbations are also observed beneath northern part of South America, where a considerable fast velocities seem to reach the top of the lower mantle. This result is consistent with low- (Su *et al.* 1994) and high- (e.g. van der Hilst 1997) resolution tomography of the mantle. Its presence has been associated with potential remains of the subducted Faralon plate (van der Hilst & Kárason 1999).

It is evident that the shear velocity perturbations examined above have greater lateral dimensions than expected from the Wadati–Benioff zones. The scales of these structures can be partly accounted for by the inherent low-pass filtration applied to the inverse problem and, partly, by the wide aperture of the long-period waves. Nonetheless, the main velocity structures shown by the cross-sections are generally consistent with regional seismicity and tectonics. The suggestion is that thermal variations generally, but not aways, have a significant impact on the topography of the 670. In regions such as Northwest Pacific and South America, the depth variation of the 670

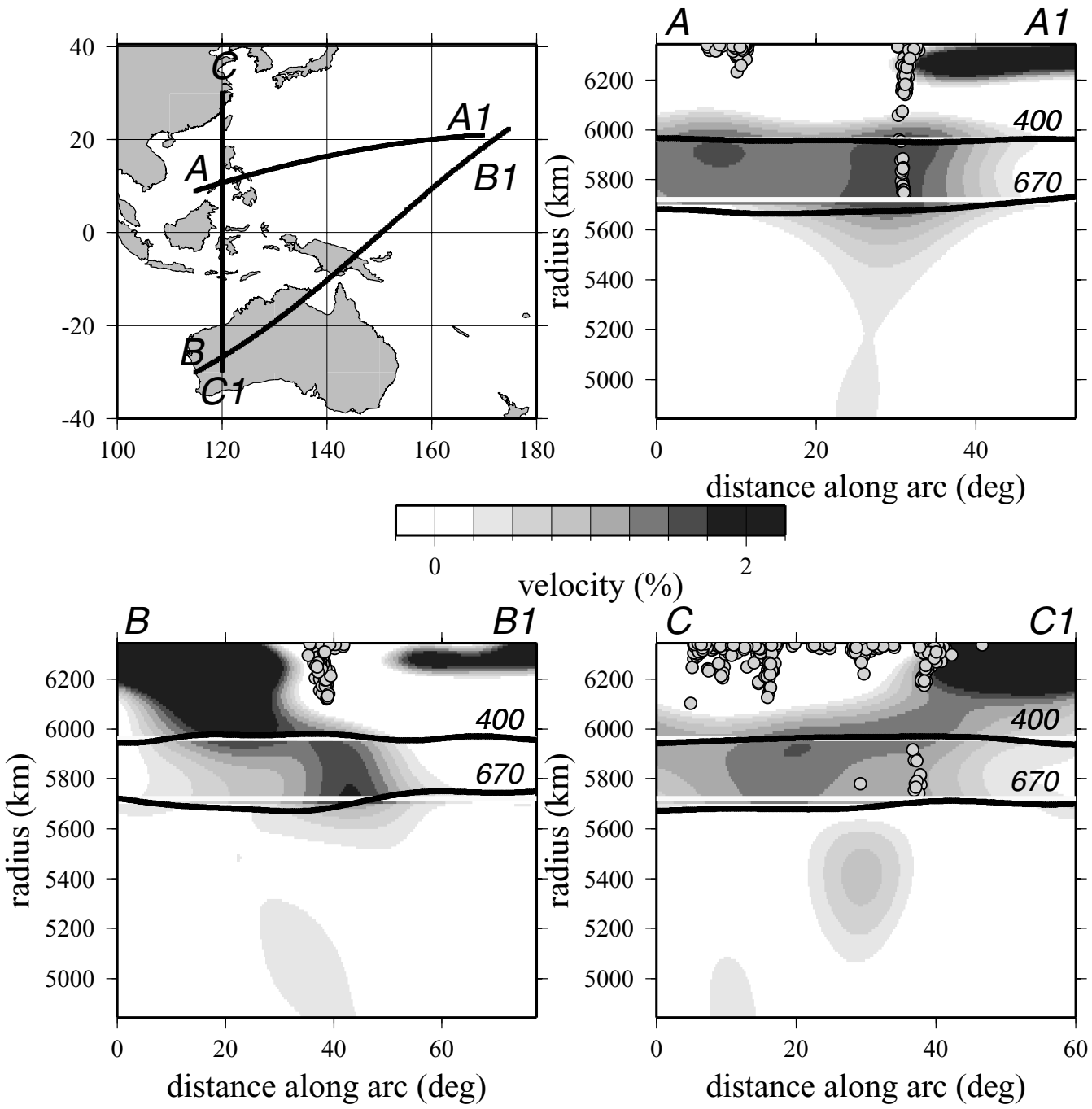

Figure 14. Regional cross-sections of TOPO362D1 through the western Pacific. Regions of fast velocities take up a significant portion of the upper mantle transition zone. The bulk of the anomalies, however, terminates in the top 800 km of the mantle.

is consistent with the velocity (and likely thermal) variations. The topography of the 400, however, does not correlate well with either the shear velocities or the regional tectonics.

5 DISCUSSION

5.1 Inference for mantle convection

One of the often-debated issue is whether the style of convection is whole-mantle or, at least, partially layered. Regional cross-sections through TOPO362D1 reveal that not all slab-like structures are continuous across the 670. To examine this issue further we compute the RMS amplitude of velocity anomalies near well documented subduction zones (Fig. 16). The sampled regions (Fig. 16a) follow those of Fukao *et al.* (2001) for a better comparison. Although differences exist at most mantle depths between the selected models, all four models (Fig. 16b) show a relative maximum of RMS in the depth range of 350–800 km (shade region); relative minimums are observed near 350 km and 850 km. The best examples of this feature are shown in the western Pacific and South America. The secondary RMS peak in the transition zone is consistent with that of Fukao *et al.* (2001), which have been associated with significant

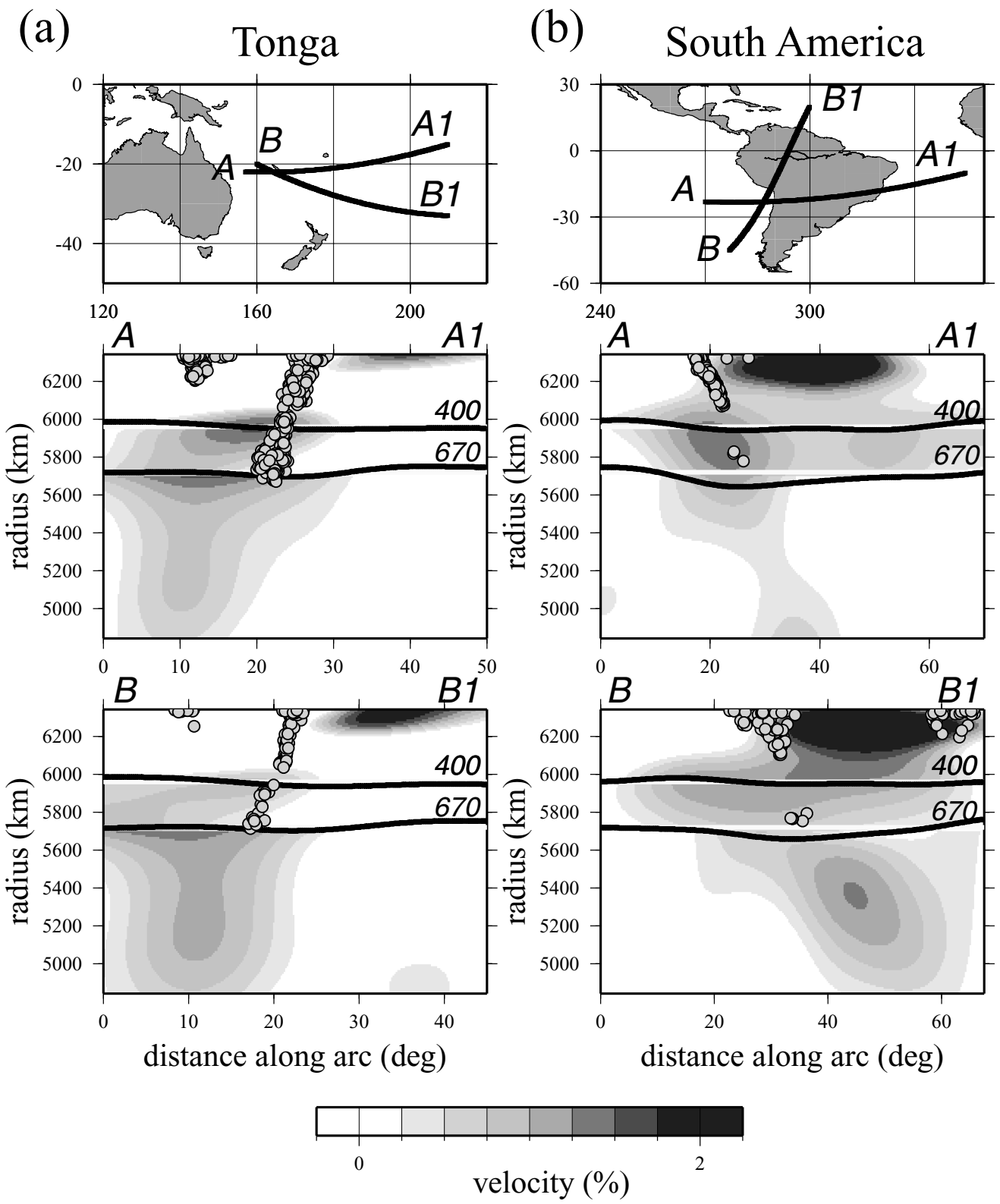

Figure 15. Regional cross-sections of TOPO362D1 through Tonga–Fiji Islands and South America. Cross-sections through (a) Tonga–Fiji islands, and (b) South America. Some of the fast velocity anomalies appear to extend down to 800–1000 km depths.

ponding or flattening of subducted slab material. The steep falloff of the RMS in the lower mantle suggests that much of the cold material is confined to the upper mantle, and that the pattern of heterogeneity below 800 km is considerably different from that above it. This upper-lower mantle difference is also evident on the global scale (Fig. 17). Above the 670, the pattern shows clear fast anomalies in the western Pacific and South American subduction zones (Fig. 17a). These anomalies span 3000–5000 km laterally and the

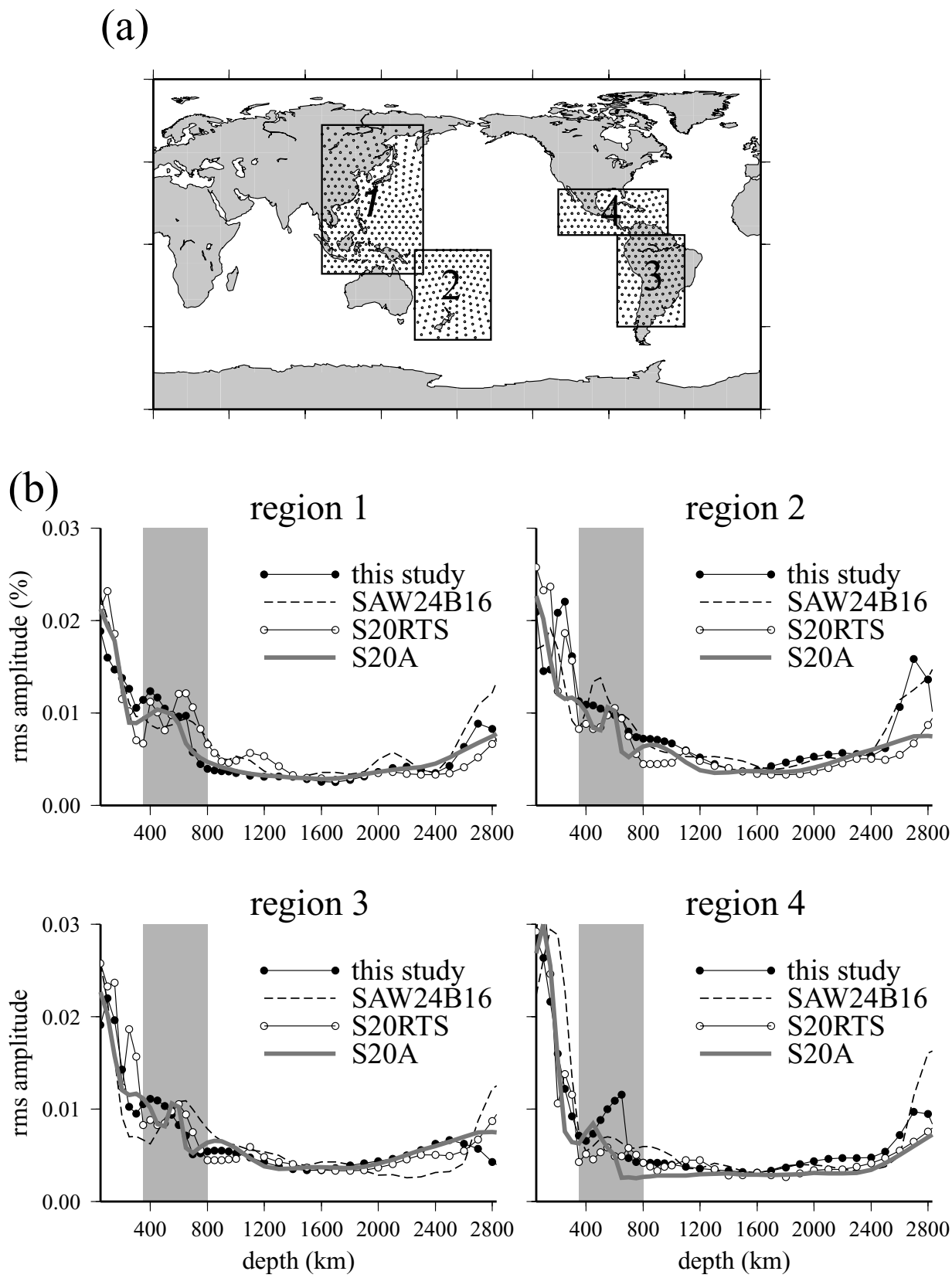

Figure 16. RMS values of velocity in four well-documented subduction zones The sampling areas are selected based on those shown by Fukao *et al.* (2001). The black dots in each region represent the locations of the sampling points used in the averaging. (b) RMS of the regions shown in (a). The shaded areas mark the depth range where the RMS changes from high (upper mantle, more variable) to low (lower mantle, less variable), as suggested by all four models. A secondary peak is present in these models in the transition zone, which likely results from the fast velocities associated with the accumulation of subducted lithospheric material.

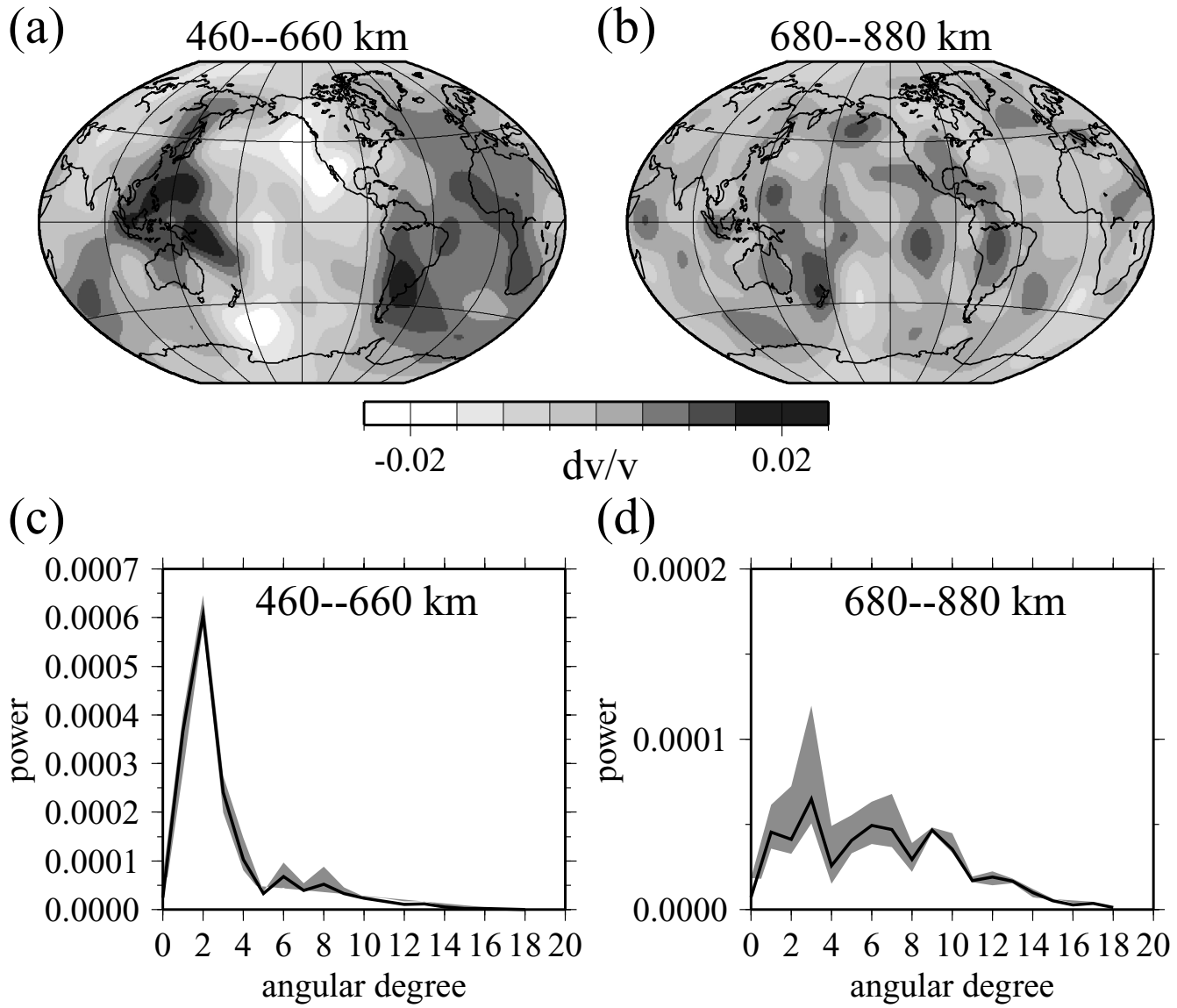

Figure 17. Average velocity and power spectra above and below the 670. (a) The average shear velocity of TOPO362D1 between 460 km and 660 km. This map is obtained by averaging 11 equally-spaced shear velocity maps in this depth range. (b) Similar to (a), but for the average structure between 680 km and 880 km. (c) Power spectra of the velocity patterns above the 670. The shaded region shows the range of variations within 460 km and 660 km for each degree; the thick black line shows the average spectrum for this depth range. (d) Similar to (c), but for the depth range of 680–880 km. The average structures above and below the 670 are significantly different.

magnitude of the perturbations is nearly 2 per cent relative to PREM. The slow anomalies are well defined within the central Pacific and Eurasia. In the lower mantle (particularly below 800 km), however, the pattern of heterogeneity shifts to smaller scales (Fig. 17b): fast anomalies in Indonesia, Tonga-Kermadec, and South America become weaker (typically <1.2 per cent) and more localized spatially. There are regions of fast velocities, e.g. in the central Pacific, that do not correlate with plate subduction of the last 200 Myr. The spectra of velocities above the 670 (Fig. 17c) show a dominant degree 2 and little variation within the transition zone, but the spectra below the 670 (Fig. 17d) contain less than 30 per cent of the peak spectral amplitude observed in the upper mantle. This suggests that the global pattern of heterogeneity undergoes significant modifications near the 670.

A number of existing models of shear velocity also show marked changes in the organization of heterogeneities from within to below the upper mantle transition region. Fig. 18 shows a comparison of radial correlations (see Jordan *et al.* 1993) between mid-transition zone velocity and the rest of the mantle. The maximum of the correlation is 1 at 540 km – the reference depth. Four of the five models show a significant decrease in the correlation near 300 km and 800 km (shaded region), particularly for TOPO362D1, S20A (Ekström & Dziewoński 1998), and SAW24B16 (Mégnin & Romanowicz 2000) where the correlation becomes zero or negative below 800 km. A similar drop-off in the global correlation between the structure in the transition zone and those in the lower mantle was documented by GDSE01 in the context of a potential disruption of mantle flow near the 670.

Although the presence and mechanism of the slab deflection are still being debated, many people now believe that the potential presence of a negative Clapeyron slope of the 670 (Ito & Takahashi 1989) and pronounced viscosity contrast (e.g. Bunge *et al.* 1996), rather

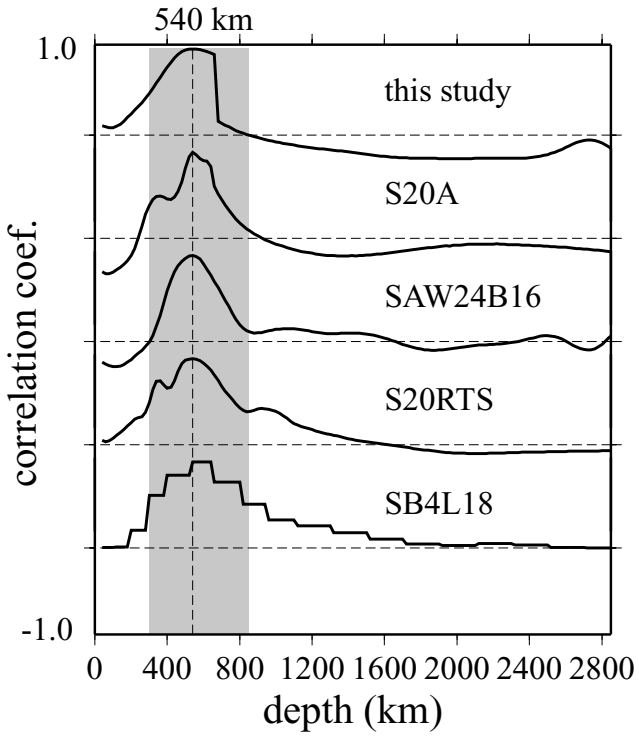

Figure 18. Radial correlations (see Jordan *et al.* 1993) with the velocity structure at 540 km. The horizontal dotted lines represent the zero correlation in each of the five subplots. The top four selected models show a significant decrease in correlation between the structure at 540 km and those below 850 km, suggesting considerable differences in the velocity patterns. There are other models, for example, SB4L18 of Masters *et al.* (2000), that do not contain this feature.

than compositional density contrast, can effectively slow down and resist the direct penetration of subducted oceanic lithosphere. Tackley *et al.* (1993, 1994) characterize the accumulation of subducted material in the transition zone using numerical modelling and raise the possibility of sudden avalanches when the cold material at the 670 becomes gravitationally unstable. A recent inversion result from seismic and geoid observations (Forte & Woodward 1997) reveals a model, among a family of 3-D models, that satisfies the observations while maintaining a layered mantle that is resistant to vertical flow near the 670. Overall, the large-scale heterogeneity obtained from the joint inversion approach is consistent with a partially layered mantle with a major reorganization of the flow pattern between 600–800 km. This does not necessarily mean that the slabs and plumes do not cross the 670, but the extent of mass and heat flow across the 670 is episodic and varies significantly among subduction zones.

5.2 Depth variability and correlation of the 400 and the 670

The range of depth variations of the 400 and 670 is a useful measure of the lateral variations in temperature. Past studies of *SS* precursors (e.g. Shearer 1991, 1993; GDA98; Flanagan & Shearer 1998) have shown that the topography of the 670 is significantly greater than that of the 400. This was considered to be one of the most robust features of mantle discontinuity observations (see Shearer 2000). Fig. 19(a) and (b) compare the distributions of the 400 depth variations before and after the joint inversion. Although the peak location (average) is nearly unchanged, the approximate half-width of the distribution decreases considerably after the joint inversion. A significantly higher percentage of the caps have variations on the order of ± 10 km (after inversion) rather than ± 15 km (prior to the inversion). The reduced topography suggests that part of the amplitude variations before inversion may have resulted from trade-offs with the velocity model and, hence, corrections for structure.

The topography of the 670 (Fig. 19c and d) shows a moderate decrease in amplitude after the joint inversion, particularly in the negative (shallower) anomalies. Most of the depth perturbations fall between ± 15 km, while the largest positive (deeper) anomalies are $\sim$18 km; this range of variations is significantly larger than that for the 400 (e.g. Shearer 1991, 1993). If the Clapeyron slope near 400 km is comparable to or greater than that near 670 km, this result would imply that the lateral temperature variations are considerably greater at the bottom of the transition zone than near the top.

Models obtained by the joint inversion also allow us to examine the relationship between the depths of the 400 and 670. As mentioned in the introduction, in the absence of chemical variations, the topography of the 400 and the 670 should anticorrelate based on the pressure-induced phase transitions of olivine. Results from *SS* precursors, however, have generally shown a decorrelation on the global scale (Shearer 1993, GDA98). The topography before the joint inversion (Fig. 19e) provides further evidence for this decorrelation; if anything, there is a slight positive correlation (see also Stammler *et al.* 1992). Flanagan & Shearer (1998) attributes the lack of a negative correlation to potential inaccuracies in the travel time corrections computed from *a priori* shear velocity model. The result from the joint inversion (Fig. 19f) shows nearly zero correlation, despite having effectively accounted for potential trade-offs between velocity and discontinuity topography; this result strongly suggests that the decorrelation is a robust global feature. It is questionable, however, that this global decorrelation of the 400 and 670 can be extrapolated to small scales (Shearer 2000), although several recent regional-scale studies have also reported uncorrelated depths between their depths (e.g. Dueker & Sheehan 1997; Li *et al.* 1998; Owens *et al.* 2000; Simmons & Gurrola 2000; Gilbert *et al.* 2003).

The relationship between velocity and topography of the 400 appears to be complex. Fig. 20(a) compares the spectrum of shear velocity at 350 km (thick solid line) with the spectra of the topography of the 400 before (labeled as 'old'; dotted line) and after (labeled as 'new'; thin solid line) the joint inversion. The amplitudes have been normalized by the largest amplitude in each spectrum for a better visual comparison. We use 350 km to represent velocity variations near the 400, though the result will not be quantitatively different if the depth is selected at 450 km instead. Both the velocity and the topography patterns have a strong degree 1, but the topography has much diminished power relative to velocity pattern at degrees 2 and 3. The modest correlation observed between a continent–ocean signature and the topography of the 400 (e.g. GDA98; Gossler & Kind 1996) becomes less significant, primarily as a result of the diminution of the Pacific depression (see Fig. 11). A degree-by-degree and cumulative correlations plot (Fig. 20b) shows a negative correlation between the velocity and the topography prior to the joint inversion (dotted curves), especially at the degree-1 spherical harmonic. A less significant correlation (-0.1) is observed between shear velocity and the topography after inversion. The lack of a better anticorrelation could indicate the presence of small but complex thermal/chemical variations near 400 km. In part, the modest overall topography of the 400 may reflect the transient nature of subducted lithosphere at depths close to 400 km; thermal

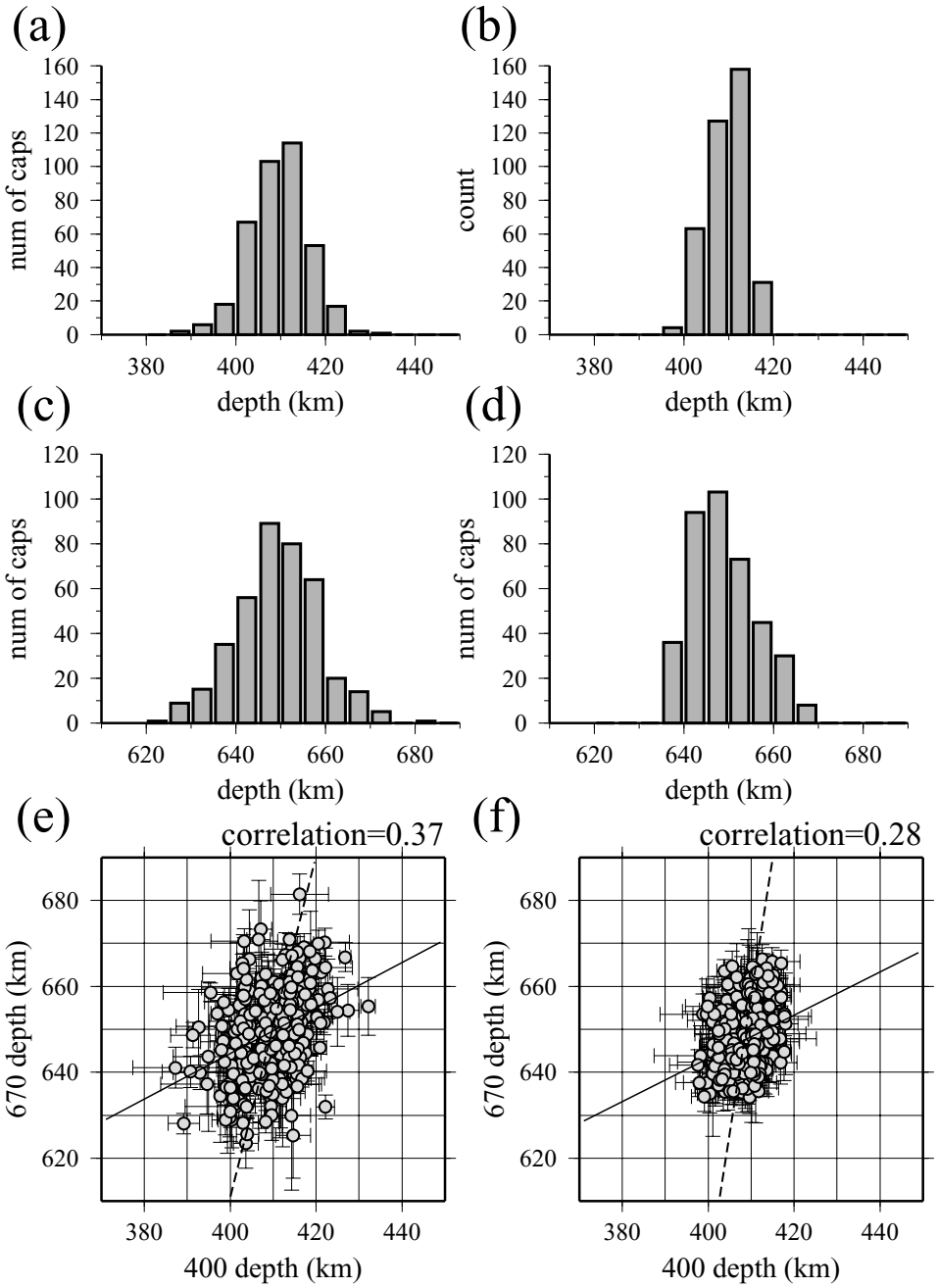

Figure 19. Topography distributions and correlations. (a) Topography distribution of the 400, obtained from *SS-S400S* before the joint inversion. (b) Topography distribution of the 400 after the joint inversion. (c)–(d) Similar to (a) and (b), but for the 670. A notable decrease in the range of variations is evident after the inversion, which suggests that some of the topographic variations can be accounted for by variations in the mantle shear velocity. (e) Correlation between the depth perturbations of the 400 and the 670 before the inversion. The solid and dotted lines show the least-squares regressions obtained by assuming the 400 and the 670 topography to be the independent variables, respectively. (f) Similar to (e), but for the resulting topography after the joint inversion. Little correlation is observed between these two discontinuities, indicating complexities in the structure and dynamics.

anomalies associated with these slabs generally have smaller amplitudes and lateral dimensions than the anomalies at the base of the upper mantle. Trade-offs between thermal and the kinetic effects associated with the positive Clapeyron slope of α-olivine to β-phase transformation (e.g. Yuen *et al.* 1994; Thoraval & Machetel 2000) may also contribute to the small topography on the 400 and a poor correlation with shear velocities at comparable depths.

On the other hand, the joint inversion significantly increases the amplitude of the degree-2 spherical harmonic (see Fig. 20c) in the topography of the 670. The relative strengths in degrees 3 and 4, in particular, are less important than they were before the joint inversion. This result agrees well with the spectral characteristics of mid-transition zone velocities (thick gray line). A high correlation ($\sim$0.6) is observed between the transition zone velocity pattern and the 670 topography after the joint inversion (Fig. 20d); $\sim$0.4 is observed before the inversion. This apparent correlation shows thermal, rather than compositional, variations in the presence of γ-spinel to perovskite and magnesiowüstite are most likely responsible for

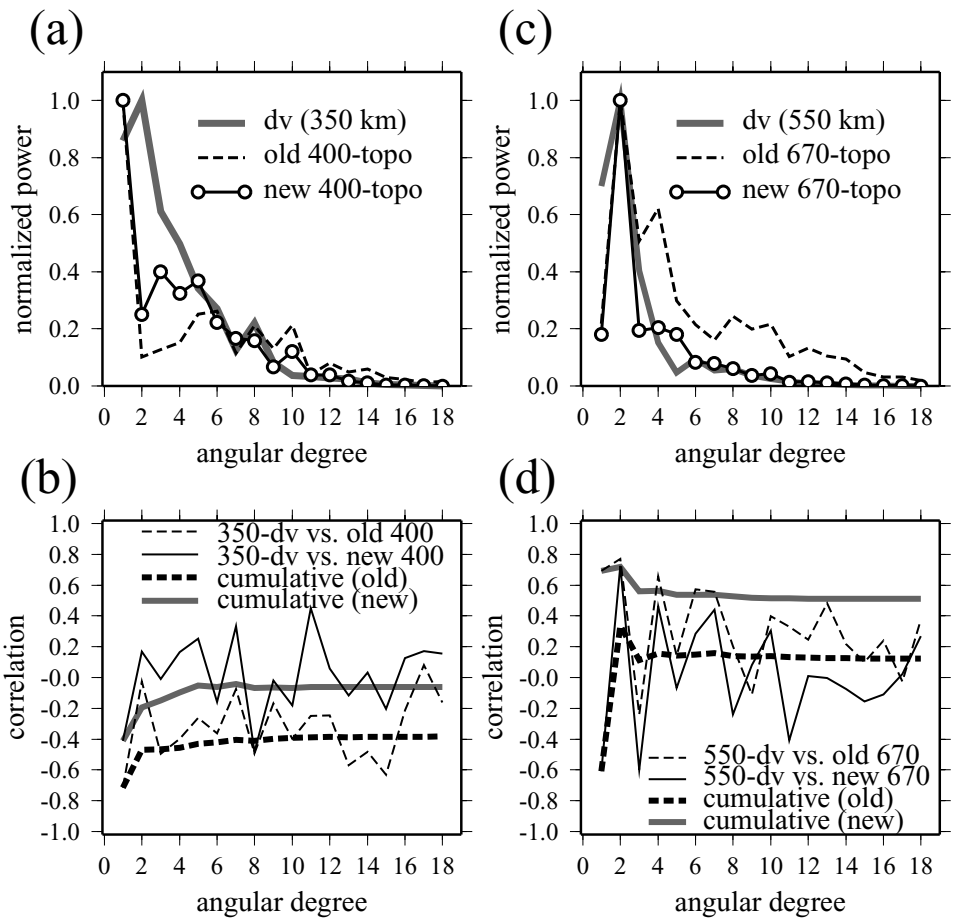

Figure 20. Correlations between velocity and topography. (a) The thick gray line shows the power spectrum of the shear velocity perturbations of TOPO362D1 at 350 km. The thin dotted and solid lines show the power spectra of the 400 before (referred to as 'old') and after (referred to as 'new') the joint inversion, respectively. (b) Degree-by-degree and cumulative correlations (in spherical harmonics) between the velocity map at 350 km and the depth variations of the 400. The topography after the inversion shows a significantly higher correlation at degree 1. (c) Similar to (a), but for mid transition zone velocities (at 550 km) and the topography of the 670. (d) Similar to (b), but for velocity perturbations at 550 km and the topography of the 670. An overall increase in the correlations is seen at the low degrees after the inversion.

a significant portion of the signal on the 670. However, it does not mean that thermal influence alone can fully account for the amplitude and scale of the observed topography. For example, if the main cause is thermal, we would expect comparable thermal contributions to the topography from velocity structures that are both above and below the 670. But a visual examination of Fig. 17 suggests that only structures above the 670 are well correlated with its topography; the pattern in the lower mantle, especially below 800 km, is appreciably different from that in the upper mantle. This suggests sources other than the endothermic phase transition, for example, the dynamical effect of cold (and therefore intrinsically dense) slab material pressing against a strongly viscous lower mantle, that may contribute to the observed topography of the 670.

6 CONCLUSIONS

This study presents a novel approach to jointly invert for mantle shear velocity and topography of transition zone discontinuities. The resulting model provides a good explanation for the data variance in the *SS-dS* travel time residuals. The velocity model obtained from the joint inversion is very similar to S362D1, but important changes have been made to the topography of the 400 and the 670. The depth variations of the 670 are generally correlated with shear velocity, and thus thermal, variations in the transition zone. In particular, the large-scale depressions of the western Pacific and South America remain after the joint inversion. There are exceptions, for example, in Tonga, where the topography does not appear to be associated with shear velocities. The topography of the 400, however, does not clearly correlate with velocity, which suggests complexities in thermal, and possibly chemical, variations within the upper mantle. The large-scale depression of the 400 in the Pacific, documented by earlier studies, has diminished significantly after the joint inversion; this implies that a considerable fraction of the previously observed pattern potentially resulted from trade-offs between velocity and discontinuity topography.

Results from the joint inversion show significantly larger topography on the 670 than on the 400. The depth variations of the 400 and 670 are characterized by a mild degree 1 (considerably decreased after the joint inversion) and a dominating degree 2 spherical harmonic, respectively. We also observe small overall variations in discontinuity depths after the joint inversion in comparison with those prior to the inversion. This suggests that some of the topographic variations obtained earlier can be effectively explained by variations in shear velocity; the main reason that a joint inversion is desirable. This observation is consistent with the recent results of Reif *et al.*

(2001) from a similar modelling approach. Future efforts could extend our analysis to the modelling of regional *S* (or *P*) velocities and smaller-scale topography of upper mantle discontinuities. We expect that the improved database of secondary *P*-to-*S* and *S*-to-*P* converted waves, with a lateral resolution of 200–500 km, will lead to a better understanding of small-scale velocity structures and regional depths of discontinuities.

ACKNOWLEDGMENTS

We thank Weijia Su, Karen Felzer, and Christine Reif for scientific discussions and comments. We also thank Peter Shearer and Steven Ward for constructive reviews and suggestions for improvements. This research is supported by grant EAR98-05172 from the National Science Foundation.

REFERENCES

Benz, H.M. & Vidale, J.E., 1993. Sharpness of upper-mantle discontinuities determined from high-frequency reflections, *Nature,* **365,** 147–150.

Bina, C. & Helffrich, G., 1994. Phase transition Clapeyron slopes and transition zone seismic discontinuity topography, *J. geophys. Res.,* **99,** 15 853–15 860.

Bunge, H.P., Richards, R.M. & Baumgardner, J.R., 1996. Effect of depth-dependent viscosity on the planform of mantle convection, *Nature,* **379,** 436–438.

Castle, J.C. & Creager, K.C., 1998. Topography of the 660-km seismic discontinuity beneath Izu-Bonin: Implications for tectonic history and slab deformation, *J. geophys. Res.,* **103,** 12 511–12 527.

Chevrot, S., Vinnik, L. & Montagner, J.-P., 1999. Global-scale analysis of the mantle *Pds* phases, *J. geophys. Res.,* **104,** 20 203–20 219.

Deuss, A. & Woodhouse, J., 2001. Seismic observations of splitting of the mid-transition zone discontinuity in Earth's mantle, *Nature,* **294,** 354–357.

Dueker, K.G. & Sheehan, A.F., 1997. Mantle discontinuity structure from midpoint stacks of converted *P* to *S* waves across the Yellowstone hotspot track, *J. geophys. Res.,* **102,** 8313–8327.

Dziewoński, A.M., 1984. Mapping the lower mantle: Determination of lateral heterogeneity in *P* velocity up to degree and order 6, *J. geophys. Res.,* **89,** 5929–5952.

Dziewoński, A.M. & Anderson, D.L., 1981. Preliminary reference Earth model, *Phys. Earth planet. Inter.,* **25,** 297–356.

Dziewoński, A.M. & Gilbert, F., 1976. The effect of small, aspherical perturbations on travel times and a re-examination of the corrections for ellipticity, *Geophys. J. R. astr. Soc.,* **44,** 7–18.

Ekström, G. & Dziewoński, A.M., 1998. The unique anisotropy of the Pacific upper mantle, *Nature,* **394,** 168–172.

Ekström, G., Tromp, J. & Larson, E.W.F., 1997. Measurements and global models of surface wave propagation, *J. geophys. Res.,* **102,** 8137–8157.

Flanagan, M.F. & Shearer, P.M., 1998. Global mapping of topography on transition zone velocity discontinuities by stacking *SS* precursors, *J. geophys. Res.,* **103,** 2673–2692.

Forte, A.M. & Woodward, R.L., 1997. Seismic-geodynamic constraints on 3-D, vertical flow, and heat transfer in the mantle, *J. geophys. Res.,* **102,** 17 981–17 994.

Fukao, Y., Obayashi, M., Inoue, H. & Nenbai, M., 1992. Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.,* **97,** 4809–4822.

Fukao, Y., Widiyantoro, S. & Obayashi, M., 2001. Stagnant slabs in the upper and lower mantle transition region, *Rev. Geophys.,* **39,** 291–323.

Gilbert, H.J., Sheehan, A.F., Dueker, K.G. & Molnar, P., 2003. Receiver functions in the western United States, with implications for upper mantle structure and dynamics, *J. geophys. Res.,* in press.

Gossler, J. & Kind, R., 1996. Seismic evidence for very deep roots of continents, *Earth planet. Sci. Lett.,* **138,** 1–13.

Grand, S.P., van der Hilst, R.D. & Widiyantoro, S., 1997. Global seismic tomography: a snapshot of convection in the Earth, *GSA Today,* **7,** 1–7.

Gu, Y.J. & Dziewoński, A.M., 2002. Global variability of transition zone thickness, *J. Geophys. res.,* **107,** 10.1029/2001JB000489.

Gu, Y., Dziewoński, A.M. & Agee, C.B., 1998. Global de-correlation of the topography of transition zone discontinuities, *Earth planet. Sci. Lett.,* **157,** 57–67.

Gu, Y.J., Dziewoński, A.M. & Ekström, G., 2001a. Preferential detection of the Lehmann discontinuity under continents, *Geophys. Res. Lett.,* **28,** 4655–4658.

Gu, Y.J., Dziewoński, A.M., Su, W.-J. & Ekström, G., 2001b. Models of the mantle shear velocity and discontinuities in the pattern of lateral heterogeneities, *J. geophys. Res.,***106,** 11 169-11 199.

Gurrola, H. & Minster, J.B., 1998. Thickness estimates of the upper-mantle transition zone from bootstrapped velocity spectrum stacks of receiver functions, *Geophys. J. Int.,* **133,** 31–43.

Ito, E. & Takahashi, E., 1989. Postspinel transformations in the system Mg_2SiO_4–Fe_2SiO_4 and some geophysical implications, *J. geophys. Res.,* **94,** 10 637–10 646.

Jordan, T.H., Puster, P., Glatzmaier, G.A. & Tackley, P.J., 1993. Comparison between seismic Earth structures and mantle flow models based on radial correlation functions, *Science,* **261,** 1427–1431.

Katsura, T. & Ito, E., 1989. The system Mg_2SiO_4–Fe_2SiO_4 at high pressures and temperatures; precise determination of stabilities of olivine, modified spinel, and spinel, *J. geophys. Res.,* **94,** 15 663–15 670.

Kawakatsu, H. & Niu, F., 1994. Seismic evidence for a 920-km discontinuity in the mantle, *Nature,* **371,** 301–305.

Li, X.-D. & Romanowicz, B., 1996. Comparison of global waveform inversions with and without considering cross-branch model coupling, *Geophys. J. Int.,* **121,** 695–709.

Li, A., Fischer, K.M., Wysession, M.E. & Clarke, T.J., 1998. Mantle discontinuities and temperature under the North American continental keel, *Nature,* **395,** 160–163.

Li, X., Kind, R., Priestley, K., Sobolev, S.V., Tilmann, F., Yuan, X. & Weber, M., 2000a. Mapping the Hawaiian plume with converted seismic waves, *Nature,* **405,** 938–941.

Li, X., Sobolev, S.V., Kind, R., Yuan, X. & Estabrook, C.H., 2000b. A detailed receiver function image of the upper mantle discontinuities in the Japan subduction zone, *Earth planet. Sci. Lett.,* **183,** 527–541.

Liu, X.-F., 1997. The 3-D shear-wave velocity structure of the Earth's lower mantle, *thesis,* Harvard Univ., Cambridge, Mass, p. 202.

Masters, G., Jordan, T.H., Silver, P.G. & Gilbert, F., 1982. Aspherical Earth structure from fundamental spheroidal-mode data, *Nature,* **298,** 609–613.

Masters, G., Laske, G., Bolton, H. & Dziewoński, A., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure, in *Earth's deep interior: mineral physics and tomography from the atomic to the global scale*, Vol. 117, pp. 63–87, eds Karato, S., Forte, A.M., Libermann, R.C., Masters, G. & Stixrude, L., *Geophys. Monogr. Ser.,* AGU, Washington, DC.

Mégnin, C. & Romanowicz, B., 2000. The 3-D shear velocity structure of the mantle from the inversion of body, surface, and higher mode waveforms, *Geophys. J. Int.,* **143,** 709–728.

Mooney, W., Laske, G. & Masters, G., 1998. A new global crustal model at 5 × 5 degrees: CRUST-5.1, *J. geophys. Res.,* **103,** 727–747.

Navrotsky, A., 1980. Lower mantle phase transitions may generally have negative pressure-temperature slopes, *Geophys. Res. Lett.,* **7,** 709–711.

Owens, T.J., Nyblade, A.A., Gurrola, H. & Langston, C.A., 2000. Mantle transition structure beneath Tanzania, East Africa, *Geophys. Res. Lett.,* **27,** 827–830.

Reif, C., Masters, G., Laske, G., Shearer, P.M. & Flanagan, M.F., 2001. Joint inversions for velocity and discontinuity structure in the mantle, *EOS, Trans. Am. geophys. Un.,* **82,** Fall Meet. Suppl., Abstract S41B-03.

Revenaugh, J. & Jordan, T.H., 1991. Mantle layering from ScS reverberation: 2. the transition zone, *J. geophys. Res.,* **96,** 19 763–19 780.

Ryberg, T., Fuchs, K., Egorkin, A.V. & Solodilov, L., 1995. Observation of high-frequency teleseismic P_n on the long-range Quartz profile across northern Eurasia, *J. geophys. Res.,* **100,** 18 151–18 163.

Shearer, P.M., 1991. Imaging global body wave phases by stacking long-period seismograms, *J. geophys. Res.,* **96,** 20 353–20 364.

Shearer, P.M., 1993. Global mapping of upper mantle reflectors from long-period SS precursors, *Geophys. J. Int.,* **115,** 878–904.

Shearer, P.M., 2000. Upper mantle seismic discontinuities, in *Earth's deep interior: mineral physics and tomography from the atomic to the global scale*, Vol. 117, pp.115–131, *Geophys. Monogr. Ser.,* AGU, Washington, DC.

Shearer, P.M. & Masters, T.G., 1992. Global Mapping of topography on the 670-km discontinuity, *Nature,* **355,** 791–796.

Simmons, N.A. & Gurrola, H., 2000. Multiple seismic discontinuities near the base of the transition zone in the Earth's mantle, *Nature,* **405,** 559–562.

Stammler, K., Kind, R., Petersen, N., Kosarev, G., Vinnik, L. & Liu, Q.-Y., 1992. The upper mantle discontinuities: correlated or anticorrelated, *Geophys. Res. Lett.,* **19,** 1563–1566.

Su, W.-J., Woodward, R.L. & Dziewoński, A.M., 1994. Degree-12 model of shear velocity heterogeneity in the mantle, *J. geophys. Res.,* **99,** 6945–6980.

Tackley, P.J., Stevenson, D.J., Glatzmaier, G. & Schubert, G., 1993. Effects of an endothermic phase transition at 670-km in a spherical model of convection in the Earth's mantle, *Nature,* **361,** 699–704.

Tackley, P.J., Stevenson, D.J., Glatzmaier, G.A. & Schubert, G., 1994. Effects of multiple phase transitions in a 3-D spherical model of convection in the Earth's mantle, *J. geophys. Res.,* **99,** 15 877–15 901.

Thoraval, C. & Machetel, P., 2000. Accounting for phase changes and their kinetics within geodynamic model for the geoid, *EOS, Trans. Am. geophys. Un.,* **81,** F1240.

van der Hilst, R. & Kárason, H., 1999. Compositional heterogeneity in the bottom 1000 kilometers of Earth's mantle: toward a hybrid convection model, *Science,* **283,** 1885–1888.

van der Hilst, R., Engdahl, R., Spakman, W. & Nolet, G., 1991. Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature,* **353,** 37–43.

van der Hilst, R., Widiyantoro, S. & Engdahl, R., 1997. Evidence for deep mantle circulation from global tomography, *Nature,* **386,** 578–584.

Vidale, J.E. & Benz, H.M., 1992. Upper-mantle seismic discontinuities and the thermal structure of subduction zones, *Nature,* **356,** 678–683.

Yuen, D.A., Reuteler, D.M., Balachandar, S., Steinbach, V., Malevsky, A.V. & Smedsmo, J.J., 1994. Various influences on 3-D mantle convection with phase transitions, *Phys. Earth planet. Inter.,* **86,** 185–203.

Zhou, H.-W. & Clayton, R.W., 1990. *P* and *S* wave travel time inversions for subducting slab under the island arcs of the northwest Pacific, *J. geophys. Res.,* **95,** 6829–6851.