Upper mantle compositional variations and discontinuity topography imaged beneath Australia from Bayesian inversion of surface-wave phase velocities and thermochemical modeling

A. Khan,[1,2] A. Zunino,[3] and F. Deschamps[4]

[1]   Here we discuss the nature of velocity heterogeneities seen in seismic tomography images of Earth's mantle whose origins and relation to thermochemical variations are yet to be understood. We illustrate this by inverting fundamental-mode and higher-order surface-wave phase velocities for radial models of the thermochemical and anisotropic structure of the mantle to 450 km depth. Dispersion data are linked to thermochemical parameters through a thermodynamic formalism for computing mantle mineral phase equilibria and physical properties. The inverse problem is solved using a probabilistic inference approach whereby robust uncertainty estimates are obtained. We find that both compositional and thermal anomalies are required if observations are to be satisfied. Mantle thermochemical variations extend to 250 km depth beneath western and central Australia and are characterized by increased Mg/Fe and Mg/Si values relative to surrounding mantle. Correlated herewith are thermal variations that closely follow surface tectonics. We also observe a strong contribution to lateral variations in structure and topography across the "410 km" seismic discontinuity from thermochemically induced phase transformations that appear much stronger than lateral variations immediately above and below. Inside the transition zone, a general decoupling of structure relative to that of the upper mantle occurs driven by a reversal in Mg/Si, while thermal anomalies are smoothed out. Comparison of presently derived shear-wave tomography models with other regional models is encouraging. Radial anisotropy is strongest at 150/200 km depth beneath oceanic/continental areas, respectively, and appears weak and homogeneous below. Finally, geoid anomalies are computed for a subset of sampled model and compared to observations.

1. Introduction

[2]   The ubiquitous velocity anomalies seen in seismic tomography images of the upper mantle at global and regional scales are observed to correlate strongly with surface tectonic features that extend well into the upper mantle linking tectonically old and stable lithosphere with fast seismic velocities and seismically slow parts with regions that have undergone relatively recent tectonic activity. There is ample evidence that the processes responsible for these features involve thermochemical variations, but, because separating thermal and chemical effects from seismic wave speeds alone is complicated, a lot of details even at the long-wavelength scale remain perfunctory. Ultimately, and in spite of a mineral physics database that has steadfastly improved over the years in quantity and quality, the use of seismic tomography images for the purpose of understanding the underlying thermochemical variations is nonunique.

[3]   Several studies have tried to explain the observed variations in seismic wave speeds purely in terms of thermal variations [e.g., *Röhm et al.*, 2000; *Shapiro and Ritzwoller*, 2004; *Goes et al.*, 2005; *McKenzie et al.*, 2005; *Cammarano and Romanowicz*, 2007; *Cammarano et al.*, 2009; *Schuberth et al.*, 2009, 2012]. However, ability to fit data by thermal variations alone does not exclude the presence of compositional variations. While these studies have given preference to thermal variations based on observations of the limited range of seismic velocities of plausible lithologies, the latter derives from shallow mantle sources ($\sim$150 km depth) [e.g., *Schutt and Lesher*, 2006; *James et al.*, 2004; *Afonso et al.*, 2010; *Hieronymous and Goes*, 2010] that may not be representative of the mantle in general. Thus, stepping beyond the "xenolith hypothesis" in order to quantify the role played

[1]Institute of Geophysics, ETH Zürich, Zürich, Switzerland.
[2]Institute of Geochemistry and Petrology, ETH Zürich, Zürich, Switzerland.
[3]Mathematical and Computational Geoscience, DTU Space, Technical University of Denmark, Lyngby, Denmark.
[4]Institute of Earth Sciences, Academica Sinica, Taipei, Taiwan.

Corresponding author: A. Khan, Institute of Geophysics, Swiss Federal Institute of Technology, Sonneggstrasse 5, CH-8092 Zürich, Switzerland. (amir.khan@erdw.ethz.ch)
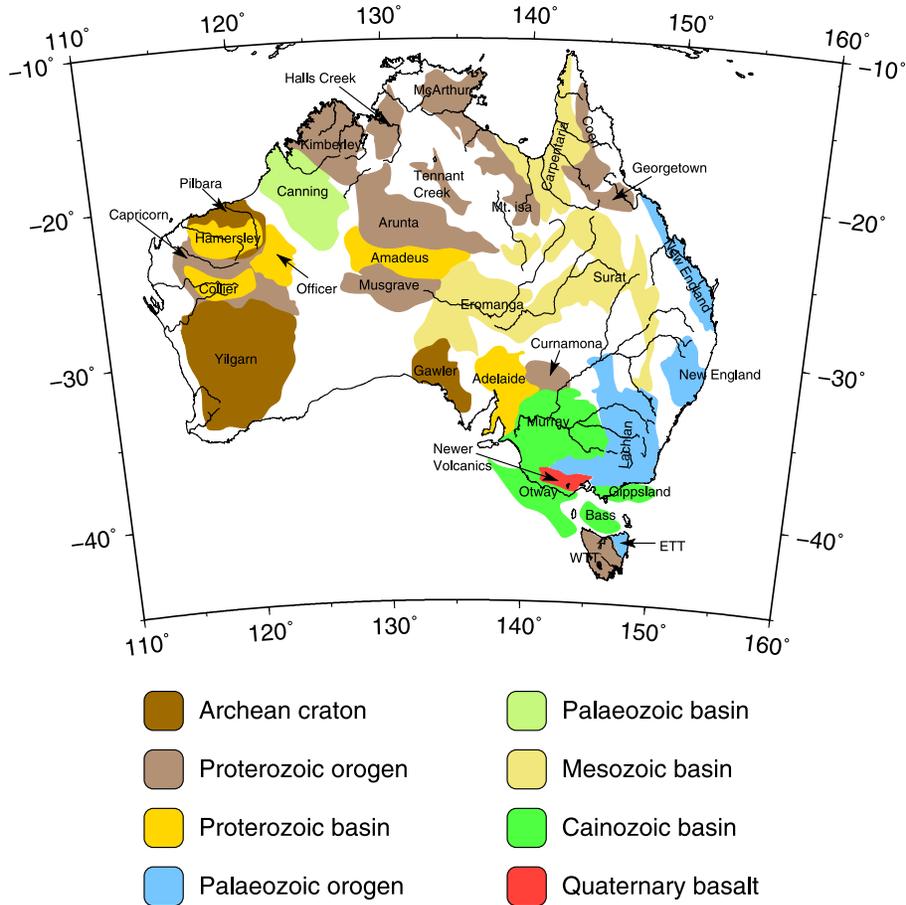
Figure 1. Simplified tectonic map of Australia.

by chemical heterogeneity may be important for understanding the structure of the lower part of the upper mantle and transition zone.

[4] In an attempt to overcome the problems associated with model construction followed by model interpretation, we present a model of the thermochemical structure of the mantle to 450 km depth beneath Australia, including its radially anisotropic shear-wave velocity structure derived directly from geophysical data. The link between geophysical observables and thermochemical state, which governs material and physical properties, is provided by thermodynamics. This ensures that temperature, composition, physical properties, and discontinuities associated with mineral phase transformations are anchored in laboratory-measured data of mantle minerals while permitting the use of inverse methods to sample a range of profiles of physical properties matching geophysical data. Relying, therefore, on a self-consistent thermodynamic formalism for computing mantle mineral phase equilibria and physical properties [e.g., *Stixrude and Lithgow-Bertelloni*, 2005; *Connolly*, 2005], quantitative inferences about the underlying processes that produce the observed variations in seismic wave speeds can be drawn.

[5] Tectonically, the Australian continent can be divided into several main provinces (Figure 1) with a westward age progression from Phanerozoic (<500 Ma) over Proterozoic (500–2500 Ma) to Archean (>2500 Ma). The Proterozoic shield regions located in the center and to the west, which

are divided into three main cratons to the west, north, and south, are separated by either orogenic belts or Phanerozoic sedimentary cover [e.g., *Betts et al.*, 2002]. The relation of these features to seismic velocity variations have been documented in several earlier studies of the region [e.g., *Simons et al.*, 1999; *Debayle and Kennett*, 2003; *Yoshizawa and Kennett*, 2004; *Fishwick et al.*, 2005, 2008; *Fichtner et al.*, 2010; *Arroucau et al.*, 2010; *Fishwick and Rawlinson*, 2012; *Kennett et al.*, 2013]. Although the models qualitatively agree, with western and central parts of the continent showing up as fast seismic velocity regions and the eastern and oceanic areas as seismically slow regions, the patterns of isotropic and anisotropic velocity anomalies and average velocities obtained in these studies nonetheless differ significantly from one another. As a result, and to the detriment of the interpreter, any geophysical implications to be drawn from these models will also differ. Model discrepancies result from differences in data, modeling aspects (forward and inverse problem), parameterization, regularization, and choice of reference model.

[6] In order to investigate the relative importance of thermochemical variations, we invert Love- and Rayleigh-wave phase velocities directly for regional models of upper mantle and transition zone thermochemistry and radial anisotropy to a depth of 450 km beneath the Australian continent following our previous approach [e.g., *Khan et al.*, 2011a]. In line with our earlier analysis, we employ a Bayesian probabilistic inference analysis [e.g., *Mosegaard and Tarantola*,

1995] to invert the regional surface-wave dispersion data following the methodology of *Shapiro and Ritzwoller* [2002] in their global-scale statistical tomographic study. In spite of advantages of Bayesian and Markov chain Monte Carlo (McMC) methods for solving inverse problems of which rigorous uncertainty, resolution and accuracy analysis have to be cited as the single most important set of items, they are, mainly because of the large computational demand involved, yet to become a mainstay of seismic tomography and currently applications are few [e.g., *Bosch*, 1999; *Shapiro and Ritzwoller*, 2002; *Bodin et al.*, 2012a; *Khan et al.*, 2011a, 2011b]. Other geophysical applications of McMC methods include the studies of *Voss et al.* [2006]; *Verhoeven et al.* [2009]; *Fullea et al.* [2012]; and *Afonso et al.* [2013a, 2013b], to name but some. Emphasis here is not on providing an estimator (e.g., mean, median, or maximum likelihood), i.e., a new tomographic model for purposes of interpretation but to provide uncertainty estimates of the relevant parameters. To assess results, we compare our thermodynamically derived seismic shear-wave tomography models with purely geophysically derived regional tomography models and find good agreement providing evidence in support of the approach taken here.

[7] Mapping anisotropy (radial and azimuthal) has been a sought-after attribute of fundamental interest because of the constraints that it provides on mantle flow. Anisotropy is believed to reflect present-day mantle strain field or past deformation frozen in the lithosphere providing a link between large-scale dynamical evolution and seismic tomogrophy [e.g., *Becker et al.*, 2008]. While regional and global anisotropic tomographic models based on surface-wave observations have detected what appears to be large-scale layering of azimuthal anisotropy within continental lithosphere and asthenosphere [e.g., *Simons and van der Hilst*, 2002; *Debayle et al.*, 2005; *Yuan and Romanowicz*, 2010] consensus is yet to emerge. The difficulty of uniquely assessing location and amplitude of anisotropic features is exemplified in several regional studies of the North American and Australian continents that appear to share few similarities [e.g., *Nettles and Dziewonski*, 2008; *Yuan et al.*, 2011; *Khan et al.*, 2011a; *Fichtner et al.*, 2010; *Yoshizawa and Kennett*, 2004].

[8] Finally, as thermal and chemical state of Earth's interior directly influence processes that determine the dynamical behavior and evolution of our planet, we employ gravity anomalies as a further means of refinement. The basic assumption being that mantle flow is driven by density anomalies that are determined by variations in the fundamental properties of interest [e.g., *Perry et al.*, 2003; *Nakagawa et al.*, 2009; *Cammarano et al.*, 2011; *Khan et al.*, 2011a].

2. Data

[9] Our data set consists of the azimuthally averaged isotropic part of the global azimuthally anisotropic phase-velocity maps of fundamental and higher-mode Love (to fifth overtone) and Rayleigh (to sixth overtone) waves of *Visser et al.* [2008] quoted at a lateral resolution of $5° \times 5°$. In an approach similar to the one employed by *Shapiro and Ritzwoller* [2002] in their McMC inversion of fundamental-mode surface-wave data for a global-scale

shear-wave velocity model, we extract from Visser et al.'s global maps dispersion curves at the center of each lateral grid node (henceforth abbreviated pixel) for the study area of interest, i.e., the Australian continent and surrounding ocean (see Figure 2). In total, this amounts to 13 dispersion curves consisting of a total of 149 distinct Love- and Rayleigh-wave phase velocities as a function of frequency for each $5° \times 5°$ pixel, which are inverted for radial profiles of composition, temperature, and radially anisotropic structure. Examples of dispersion curves, including frequency ranges studied, are shown later (see section 5.1, Figure 4). As previously, only a regional-scale study is undertaken due to the high computational requirements of McMC methods.

[10] There are inherent limitations in using surface-wave phase-velocity maps as data in place of the phase-velocity data from which the former are constructed. The range of models that will be mapped out here will be controlled by the phase-velocity maps and their nominal uncertainties. Whether these uncertainties are equivalent or similar to the uncertainties of the original data is difficult to ascertain. However, it has to be pointed out that the phase-velocity maps of *Visser et al.* [2008] were constructed from phase-velocity measurements obtained using a model-space search technique, which provides uncertainties on phase-velocity measurements as well as on phase-velocity maps. This means that the present inversion is not a tomographic inversion in the narrow sense, but an inversion for a set of (local) radial profiles of thermochemical and physical structure that result in a range of tomographic models that are consistent with Visser et al.'s phase-velocity maps and uncertainties. Regarding the latter, uncertainties are dominated by the smoothness-constraints employed in the construction of the maps, and therefore, only an average uncertainty per mode is given in Visser et al. Note that uncertainty varies with mode and frequency. For further details, we refer the reader to the work of *Visser et al.* [2008].

3. Computational Aspects

3.1. Parameterization

[11] Lateral and radial model parameterization is shown in Figure 2. Surface-wave dispersion data are strongly affected by the velocity structure and thickness of the crust [e.g., *Boschi and Ekström*, 2002; *Bozdag and Trampert*, 2008]. As *Nettles and Dziewonski* [2008], we use global crustal model CRUST-2.0 [*Bassin et al.*, 2000] as initial model to derive a regional crustal model that has many features in common with recent models obtained from crustal-specific studies [e.g., *Kennett et al.*, 2011; *Bodin et al.*, 2012b]. The crust is parameterized using purely physical properties within 10 layers: P, S wave velocity (V_P, V_S), density (ρ), and depth to crust-mantle interface (d_{Moho}). As in our previous work, we employ the crustal model CRUST2.0 (http://mahi.ucsd.edu/Gabi/rem.html) as initial model for the crust and change properties in top and bottom layers only while interpolating linearly for the remainder. Mantle layers (10) are described in terms of composition c, temperature T, anisotropy parameters ξ, ϕ, η (to be defined below), and seismic wave attenuation Q. As only the mantle is of interest here, this results in a total of 97 parameters for each pixel. Parameters and prior information are summarized further in Table 1.
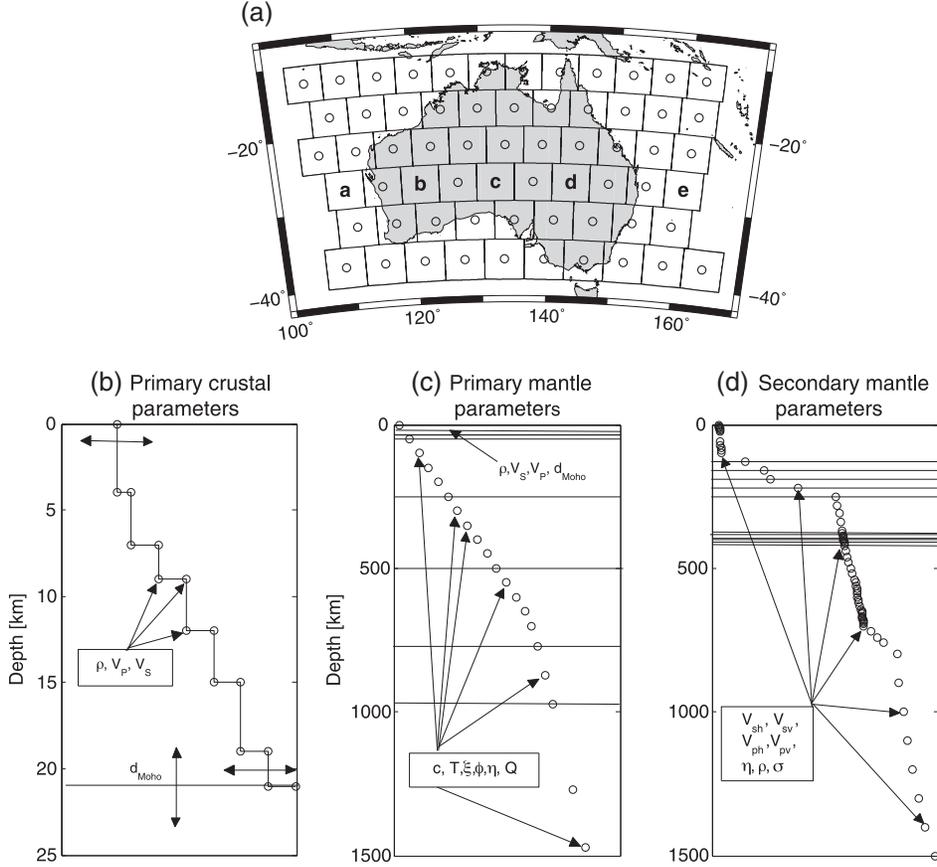
Figure 2. (a) Lateral and (b–d) radial model parameterization. (Figure 2a) Dots at the center of each pixel denote the locations at which properties are defined laterally with a grid spacing of $5° \times 5°$. Letters a–e refer to locations for which radial shear-wave velocity profiles are displayed in Figure 11. Radially (Figures 2b–2d), the model is parameterized in terms of layers. The primary (fundamental) parameters are as follows: (Figure 2b) crust is delineated by density (ρ), P and S wave velocity (V_p, V_s), and depth to crust-mantle interface (d_{Moho}) while mantle layers (Figure 2c) are described by the main parameters: composition (c), temperature (T), anisotropy (ξ, ϕ, η), and attenuation (Q). (Figure 2d) Secondary parameters (physical properties: V_{SH}, V_{SV}, V_{PH}, V_{PV}, η, and ρ) are computed on the basis of the primary parameters on a grid, which is finely spaced around locations of major mantle mineral phase transitions. Circles in both plots denote location of a set of prefixed depth nodes, except in the crust where depth to crust-mantle interface is variable.

3.2. Forward Problem

[12] The forward problem is summarized in Figure 3 and will be briefly discussed here.

[13] Step 1: $\{c, T\} \xrightarrow{g_1} \{M\} \xrightarrow{g_2} \{V_p, V_s, \rho\}$. Equilibrium mineralogy M and isotropic V_p and V_s, and ρ beneath each pixel are computed using the thermodynamic code Perple_X developed by *Connolly* [2005]. Assuming thermodynamic equilibrium Perple_X computes mantle mineralogy by Gibbs free energy minimization as a function of composition, pressure, and temperature. The thermodynamic formalism of *Stixrude and Lithgow-Bertelloni* [2005] as parameterized by *Xu et al.* [2008] for mantle minerals in the model chemical system $Na_2O\text{-}CaO\text{-}FeO\text{-}MgO\text{-}Al_2O_3\text{-}SiO_2$ (abbreviated NCFMAS) is employed. The Gibbs energy minimization procedure yields the amounts, compositions, and physical properties, including elastic moduli, of the stable minerals in the model chemical system, while bulk rock elastic moduli are estimated by Voigt-Reuss-Hill averaging.

The pressure profile is obtained by integrating the load from the surface (boundary condition $p = 10^5$ Pa).

[14] As an alternative to our assumption of thermodynamic equilibrium, *Xu et al.* [2008] have proposed that the mantle consists of a nonequilibrium mechanical mixture. In this model, the mantle has undergone complete differentiation into end-members, and bulk properties are computed by averaging the properties of the minerals in these end-members. However, from the point of view of geophysics, there is no argument for or against either model as discussed in *Khan et al.* [2009], whereas from a petrological viewpoint, it can be argued that while the mechanical mixture model plausibly depicts the influence of chemical segregation on the equilibrium model, it cannot be argued to be a more realistic end-member for the Earth's mantle because it is inconsistent with mid-ocean ridge volcanism.

[15] Step 2: $\{V_p, V_s, \rho, Q\} \xrightarrow{g_3} \{v_p, v_s, \rho\}$. Anelastic wave speeds due to attenuation Q and dispersion are estimated

Table 1. Prior Model Parameter Information[a]

Description	Layers	Parameter	Value	Reference/Comment
Crust	2	V_P	CRUST2.0$\pm$10%	Top and bottom crustal layer
	2	V_S	CRUST2.0$\pm$10%	Top and bottom crustal layer
	2	ρ	CRUST2.0$\pm$10%	Top and bottom crustal layer
		d_{Moho}	CRUST2.0$\pm$20 km	Beneath continents
		d_{Moho}	CRUST2.0$\pm$5 km	Beneath oceans
Composition (in wt %)	10	c_1 (CaO)	2.32–3.88	LK07
		c_2 (FeO)	5–10	
		c_3 (MgO)	35–41.6	LK07
		c_4 (Al$_2$O$_3$)	2.92–4.87	LK07
		c_5 (Na$_2$O)	0.157–0.439	LK07
		c_6 (SiO$_2$)	$100 - \sum_{j=1}^{5} c_j^i$	All components in any given layer i should sum to 100 wt %.
Temperature	10	T	$T_{i-1} \leq T_i \leq T_{i+1}$ $T_{surface}=0°C$	Temperature cannot decrease with depth.
Anisotropy	10	ξ	0.97–1.15	PR06,KU08
		ϕ	0.97–1.15	PR06,KU08
		η	0.97–1.15	PR06,KU08
Attenuation		Q	E_a=500 kJ/mol	SB96
			V_a=2.5·10^5 cm^3/mol	SB96
			α=0.2	JA02
			Q_0=1	CB08

[a]Index i refers to the radial direction, e.g., T_i is the temperature in layer i. PR06 and KU08 refer to the whole-mantle anisotropic tomography models of *Panning and Romanowicz* [2006] and *Kustowski et al.* [2008], respectively, while SB96, JA02, and CB08 refer to the studies of *Sobolev et al.* [1996], *Jackson et al.* [2002], and *Cobden et al.* [2008], and LK07 refers to Table 2 in the study of *Lyubetskaya and Korenaga* [2007].

from the expression [e.g., *Anderson*, 1989]

$$v_{s/p}(p,T,c) = V_{s/p}(p,T,c)\left[1 - \frac{2Q_{s/p}^{-1}}{\tan(\alpha\pi/2)}\right], \qquad (1)$$

while shear-wave attenuation Q_S is computed from

$$Q_s = Q_0 \exp\left[\frac{\alpha(E_a + pV_a)}{RT}\right], \qquad (2)$$

where Q_0 is a constant, E_a the activation energy, V_a the activation volume, p pressure, T temperature, R the gas constant, and α an exponent, which has been determined experimentally to be between 0.15 and 0.25 [e.g., *Jackson et al.*, 2002]. Parameter values are tabulated in Table 1. The compressional quality factor is obtained from $Q_p^{-1} = (4V_s^2/3V_p^2)Q_s^{-1}$. We should note that the anelastic correction of seismic wave speeds (equation (1)) only applies for an anelastic absorption band within which attenuation has a power law dependence upon period. Any dissipation peak, associated with elastically accommodated grain-boundary sliding or melt squirt, results in dispersion that is more strongly localized in frequency [e.g., *Jackson and Faul*, 2010].

[16] Step 3: $\{v_p, v_s, \rho, \xi, \phi, \eta\} \xrightarrow{g_4} \{V_{SV}, V_{SH}, V_{PV}, V_{PH}, \eta, \rho\}$. As in our previous work and elsewhere [e.g., *Panning and Romanowicz*, 2006; *Kustowski et al.*, 2008], we assume

transverse isotropy (vertical symmetry axis) and $\eta \sim 1$, i.e., small anisotropy, in which case anisotropic Voigt-averaged velocities can be computed from the above isotropic (anelastically corrected) P and S wave velocities [*Babuska and Cara*, 1991] using

$$V_{SV} = \sqrt{\frac{3v_s^2}{2\xi}}, \quad V_{SH} = \sqrt{\frac{3v_s^2}{2}}, \quad V_{PH} = \sqrt{\frac{5v_p^2}{4\phi}}, \quad V_{PV} = \sqrt{\frac{5v_p^2}{4}}, \qquad (3)$$

where $\xi = V_{SH}^2/V_{SV}^2$ and $\phi = V_{PV}^2/V_{PH}^2$ quantify S and P wave anisotropy, respectively. V_{SV}, V_{SH}, V_{PV}, and V_{PH} are the velocities of vertically (V) and horizontally (H) polarized S waves and vertically and horizontally propagating P waves, respectively, and $\eta = F/\rho(V_{PH}^2 - 2V_{SV}^2)$, where F is a function of velocity at intermediate incidence angles [*Dziewonski and Anderson*, 1981].

[17] Step 4: $\{V_{SV}, V_{SH}, V_{PV}, V_{PH}, \eta, \rho\} \xrightarrow{g_5} \{C_R(\omega), C_L(\omega)\}$. The final step, which involves computing fundamental-mode and higher-order Rayleigh- and Love-wave dispersion curves as a function of frequency ω, is performed using the minor-based code of *Nolet* [2008]. Finally, we would like to recall that relative sensitivity of higher modes to compressional velocity (for Rayleigh waves) and density grows with increasing overtone number

$$\{c, T\} \xrightarrow{g_1} \{M\} \xrightarrow{g_2} \{V_p, V_s, \rho\} \quad\longrightarrow\quad \{V_{SV}, V_{SH}, V_{PV}, V_{PH}, \rho, \eta\} \xrightarrow{g_5} \{C_R, C_L\}$$

with $\{\xi, \phi, \eta\}$ connected via g_4 and $\{Q\}$ connected via g_3.
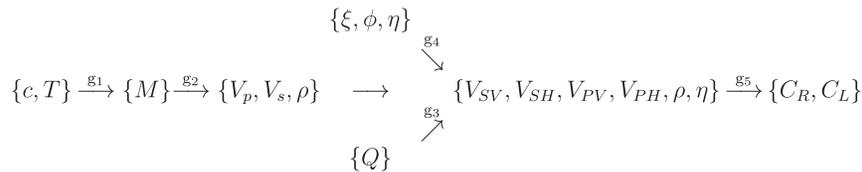
Figure 3. A schematic illustration of the forward problem and the different model parameters $(c, T, \xi, \phi, \dots)$, physical theories $(g_1, g_2, \dots)$, and data (C_R, C_L) used to describe it. Symbols are described in section 3.2.

[see *Anderson and Dziewonski*, 1982]. Note also the complete absence of preassigned scaling factors between V_p, V_s, and ρ in contrast to common practice in surface-wave and waveform tomography [e.g., *Nettles and Dziewonski*, 2008; *Yuan et al.*, 2011; *Fichtner et al.*, 2010].

4. Inverse Problem

[18] We employ the probabilistic approach of *Tarantola and Valette* [1982] to solve the nonlinear inverse problem. A detailed treatment is given in, e.g., *Khan et al.* [2007] and what follows is a very brief outline. Within a Bayesian framework, the solution to the general inverse problem **d** = g(**m**), where **d** is a data vector containing observations and g is a typically nonlinear operator that maps a model parameter vector **m** into data, is given by [e.g., *Tarantola and Valette*, 1982; *Mosegaard and Tarantola*, 1995]

$$\Omega(\mathbf{m}) = kf(\mathbf{m})\mathcal{L}(\mathbf{m}), \qquad (4)$$

where k is a normalization constant, $f(\mathbf{m})$ is the prior probability distribution on model parameters (see Table 1), i.e., information about model parameters obtained independently of the data under consideration, $\mathcal{L}(\mathbf{m})$ is the likelihood function, which can be interpreted as a measure of misfit between the observations and the predictions from model **m**, and $\Omega(\mathbf{m})$ is the posterior model parameter distribution containing the solution to the inverse problem. The particular form of $\mathcal{L}(\mathbf{m})$ is determined by the observations, their uncertainties, and how these are employed to model data noise (to be enunciated below).

[19] To sample the posterior distribution (equation (4)) in the model space, we employ the Metropolis algorithm. Although this algorithm is based on random sampling of the model space, only models that result in a good data fit and are consistent with prior information are frequently sampled (importance sampling). The Metropolis algorithm is capable of sampling the model space with a sampling density proportional to the target posterior probability density without excessively sampling low-probability areas. This is particularly important in the case of high-dimensional model spaces in which a large proportion of the volume may have near-zero probability density.

4.1. Sampling the Posterior Distribution

[20] We assume that data noise can be modeled using a Gaussian distribution and that observational uncertainties and calculation errors between Rayleigh and Love waves are independent as a result of which the likelihood function can be written as

$$\mathcal{L}(\mathbf{m}) \propto \exp\left(-\sum_{\text{mode}}\sum_{\text{frequency}}\frac{[d_{\text{obs}}^{R} - d_{\text{cal}}^{R}(\mathbf{m})]^2}{2\sigma_R^2}\right.$$
$$\left. -\sum_{\text{mode}}\sum_{\text{frequency}}\frac{[d_{\text{obs}}^{L} - d_{\text{cal}}^{L}(\mathbf{m})]^2}{2\sigma_L^2}\right), \qquad (5)$$

where d_{obs} denotes the observed data, $d_{\text{cal}}(\mathbf{m})$ denotes the calculated data, superscripts R and L are the shorthand notations for Rayleigh and Love waves, respectively, and σ uncertainty on either of these.

[21] Of importance when running Monte Carlo simulations are questions of convergence and sample independence, in addition to how many samples are required to adequately approximate the posterior distribution. The number of iterations needed for the algorithm to converge to a high-probability region (burn-in time), i.e., the number of iterations before samples can be retained from the posterior distribution, was found to be around $5\cdot10^3$. In total, one million models were sampled with an overall acceptance rate of about 40%. To ensure convergence of the McMC algorithm in practice, time series of all output parameters were checked for stationarity throughout the entire sampling stage. Adequate sampling of the model space was ensured by sampling until no palpable changes to the characteristics of the posterior probability distributions were observed, in addition to recommencing the algorithm at a number of different random places in the model space. To further ensure near-independent models, only every hundredth sample was retained. This was found by analyzing the autocorrelation function of the fluctuations of the likelihood function.

[22] As mentioned earlier, the main purpose of using a probabilistic inference approach to solving nonlinear inverse problems is not to present a single realization from the posterior distribution, but rather the information gathered from a whole series of sampled models. This implies that only the statistical nature of model features is of interest. To quantitatively summarize information from such a necessarily complex probability density function (PDF) standard resolution measures involving means and covariances are inadequate, and we have to resort to a more general approach, which will typically depend on the questions that we are trying to address. There are several different strategies that can be employed to analyze the posterior distribution, e.g., through the use of marginal posterior distributions, resolution measures, Bayesian hypothesis testing, computation of credible or highest posterior density regions, and not least the movie strategy [e.g., *Box and Tiao*, 1992; *Bernardo and Smith*, 1994; *Mosegaard and Tarantola*, 2002; *Tarantola*, 2005]. Here we will concentrate on computing highest probability density intervals in order to provide the reader with a quantitative notion of model parameter uncertainties.

[23] Bayesian highest probability density intervals of marginal distributions were computed with the method of *Chen and Shao* [1999], which provides lower and upper bound given a confidence level (here 95%). As estimators, we use the maximum likelihood model and the mean, i.e., the simple average of all samples. While the maximum likelihood model provides the best fit to data in probabilistic terms, it should be clear that this particular model is not more credible than any other model taken randomly from the posterior PDF. This results directly from the observation that posterior models, although different (e.g., with and without upper mantle low-velocity zones, shallow and deep-lying "410 km" seismic discontinuities, etc.), are nonetheless models with high likelihood values that predict observed data within uncertainties (see Figure 4). Moreover, it is also the reason why the mean of such a collection of models, which itself is necessarily smooth, has a lower posterior probability of being sampled as it is less likely to fit data to the extent required. That said, if the PDF is Gaussian-like one often encounters instances where the mean resembles single samples such as the maximum likelihood model; where this is not the case the PDF is non-Gaussian and any single estimator will have to be interpreted carefully.
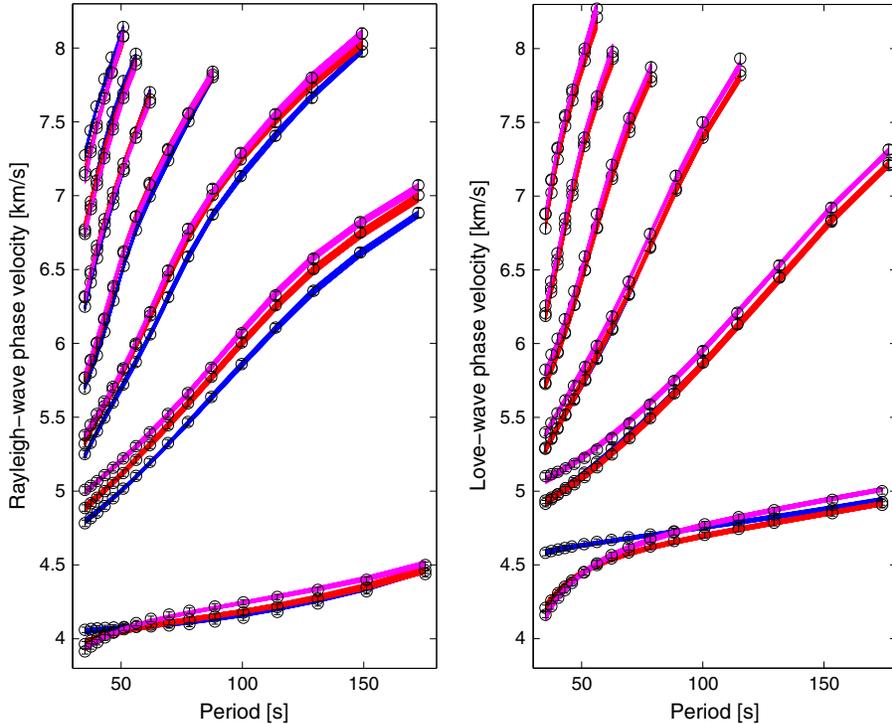
Figure 4. Comparison of calculated (red, blue, and magenta lines) and observed Rayleigh- and Love-wave phase velocities (circles), including uncertainties (error bars) at an oceanic (blue lines) and two continental (red and magenta lines) locations. Colored lines show calculated dispersion curves for all sampled models. Geographic location of sites is shown in Figure 2 as pixels **a** (blue), **c** (magenta), and **d** (red), respectively.

[24] Finally, and on a more technical note, our models are parametrized on an unstructured grid subdivided in cells of approximately $5° × 5°$ of latitude and longitude at a set of fixed depth nodes. Values at selected depth locations of 100, 150, 200, 250, 300, 350, 400, and 450 km have been obtained by linear interpolation along the vertical axis for each pixel. In order to plot our results, we constructed a series of depth slices of our models interpolating in longitude and latitude on a finer grid of $120 × 60$ nodes. The interpolation has been carried out employing radial basis functions (cubic) for unstructured grids as provided by the Python "Matplotlib" package [*Hunter*, 2007]. We utilize this method for all plots with the only exception of *Bodin et al.*'s [2012a] Moho depth model (see Figure 5), where a two-dimensional spline interpolation method was employed. The limiting polygon shown in the plots represents the outer border of our $5° × 5°$ grid.

4.2. Modeling Uncertainties

[25] The dominant source of uncertainty in the calculations performed here is related to the paucity of experimental constraints on the parameters of relevance for employing the thermodynamic formalism of *Stixrude and Lithgow-Bertelloni* [2005] and parametrized by *Xu et al.* [2008]. Assessment of the error is further complicated by the correlations existing between the various parameters such as bulk modulus and its pressure derivative, which are known to trade off strongly. However, as the covariance matrix is not available for the published thermodynamic parameters

employed here, we have to abstain from modeling uncertainties. Where data are available, *Bass* [1995] has reported, based on current experimental precision, minimum errors of 1–2% in computed velocities. As a consequence, estimates on sampled thermochemical parameters appear to be better constrained than they are in reality.

[26] There is currently only limited information on attenuation structure, which arises from an absence of experimental data and large uncertainties associated with attenuation-related parameters (see equation (2)) and anelastic effects. To circumvent this several tests were conducted, where relevant parameters were varied by as much as 10–20% to ascertain sensitivity of data to attenuation structure and to ensure that results did not significantly change as a result.

[27] Of additional importance for the computations performed here are prior constraints. If based on physical theory, experimental evidence, or other observational arguments, prior information ($f(\mathbf{m})$ in equation (4)) can provide strong constraints on the main parameters of interest. Given the essentially unknown thermochemical state of Earth's mantle, we employ very wide prior parameter probability distributions as indicated in Table 1. This will be shown in the following.

[28] Finally, we disregard potentially important contributions arising from the presence of melt and/or water [e.g., *Karato and Jung*, 1998; *Chen et al.*, 2002; *Smyth and Frost*, 2002]. Although there is geophysical evidence for the presence of either component or phase [e.g., *Toffelmier and Tyburczy*, 2007; *Tauzin et al.*, 2010; *Khan and Shankland*,
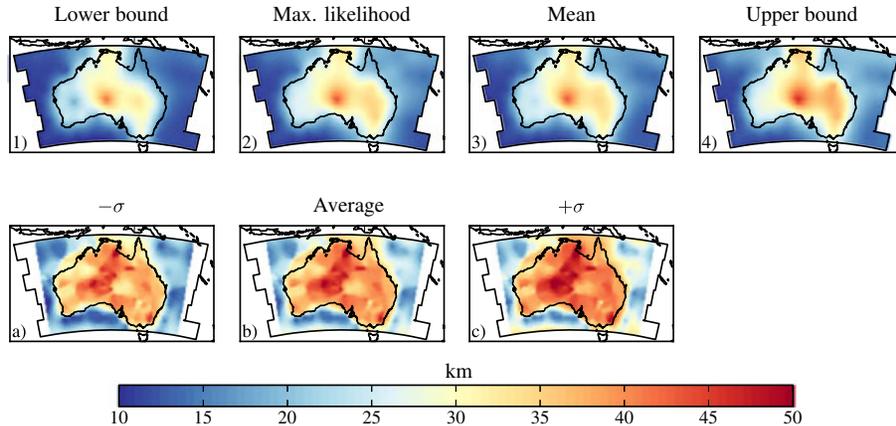
Figure 5. Maps of Moho depth (in kilometers) showing (plot 2) maximum likelihood model, (plot 3) mean model, and (plot 4) upper and (plot 1) lower bounds in the form of 95% highest probability density regions as measures of overall uncertainty on Moho depth obtained in the present study. Note that the maximum likelihood model is but one among many models that fit data. Plots in the bottom row display mean Moho depth (b) ± one standard deviation (σ) (plots a and c) obtained by *Bodin et al.* [2012a] in their study of crustal structure based on a large diverse set of seismic data.

2012], quantitative thermodynamical modeling in the current context is not feasible given present-day lack of relevant data.

5. Results and Discussion

5.1. Data Fit

[29] Differences in data between the various locations are shown in Figure 4 displaying significant tectonic variations as expected. For oceanic/continental differences, higher-order Rayleigh-wave data are clearly distinct, while for fundamental-mode data, this is most strongly the case for Love waves and short-period Rayleigh waves. In the case of local continent/continent differences, data are distinguishable across most of the branches. Also shown in Figure 4 is our fit to data where sampled models are seen to fit observations within uncertainties.

5.2. Crustal Structure

[30] To exemplify our inversion for crustal structure, we focus on Moho depth. Maps of Moho depth obtained here are shown in Figure 5. Also included here, for comparison, is Moho depth across the Australian continent from the study of *Bodin et al.* [2012a]. The latter is obtained from probabilistic surface reconstruction that incorporates data from different active and passive seismic source surveys. Setting the comparatively low resolution of our model aside, both maps generally show the same pattern, with the thickest crust appearing beneath central Australia and eastern highlands reaching Moho depths between 40 and 50 km. Regions showing shallow crustal thicknesses include almost the entire western edge as well as the easternmost parts with Moho depths in the range 20–30 km. Relatively thin crust is also observed to underlie the Pilbara craton. We find uncertainties in Moho depth to be about 10 km beneath most of continental Australia. These estimates are higher than Bodin et al.'s, but reflect the low-resolution nature of our model as well as the fact that the uncertainty measure (one standard deviation) quoted for Bodin et al. is generally narrower than

the 95% credible interval used here. Given the relatively low sensitivity of our surface-wave data to crustal structure in comparison to the high-resolution data of Bodin et al. that were acquisitioned for the purpose of crustal structure determination, more quantitative comparison is not warranted here. The level of agreement between the models shown here as well as others, e.g., Australian seismological reference model (AusREM) [*Kennett et al.*, 2011] is nonetheless encouraging.

5.3. Upper Mantle Thermochemical Structure

[31] As indicators of chemical variations in the mantle beneath the Australian continent and surrounding oceans, we limit ourselves to showing the following elemental ratios, Mg/Fe and Mg/Si, as these are the best resolved compositional parameters. Maps of Mg/Fe, Mg/Si, and temperature variations in the mantle are shown in Figures 6–8. A striking feature of these maps is a clear correlation between surface tectonics and thermochemical structure that seems to persist, within the uncertainties of the maps (upper and lower bounds in Figures 6–8), to a depth of ∼250 km. In particular, we find that the old central and western Archean and Proterozoic regions are relatively "cold" and characterized by high Mg/Fe and Mg/Si values relative to the eastern Phanerozoic and oceanic regions. The observed compositional variations are in general agreement with what is expected for depleted peridotite. While diminishing in amplitude and areal extent (becoming centered on the Australian continent), the thermochemical signature appears robust to a depth of 200–250 km. Comparison of these maps indicates that mantle temperature is better resolved than composition as the tectonic division, for example, is clearly resolved in both upper and lower bounds. Although this appears less distinct in the compositional maps, it should be noted that high Mg/Si values across the western and central parts appear resolvable to ∼200 km depth. Also, the maps do strongly suggest a geographically resolvable structure in the retrieved compositional pattern inasmuch as areas with relative high/low values in the lower bound
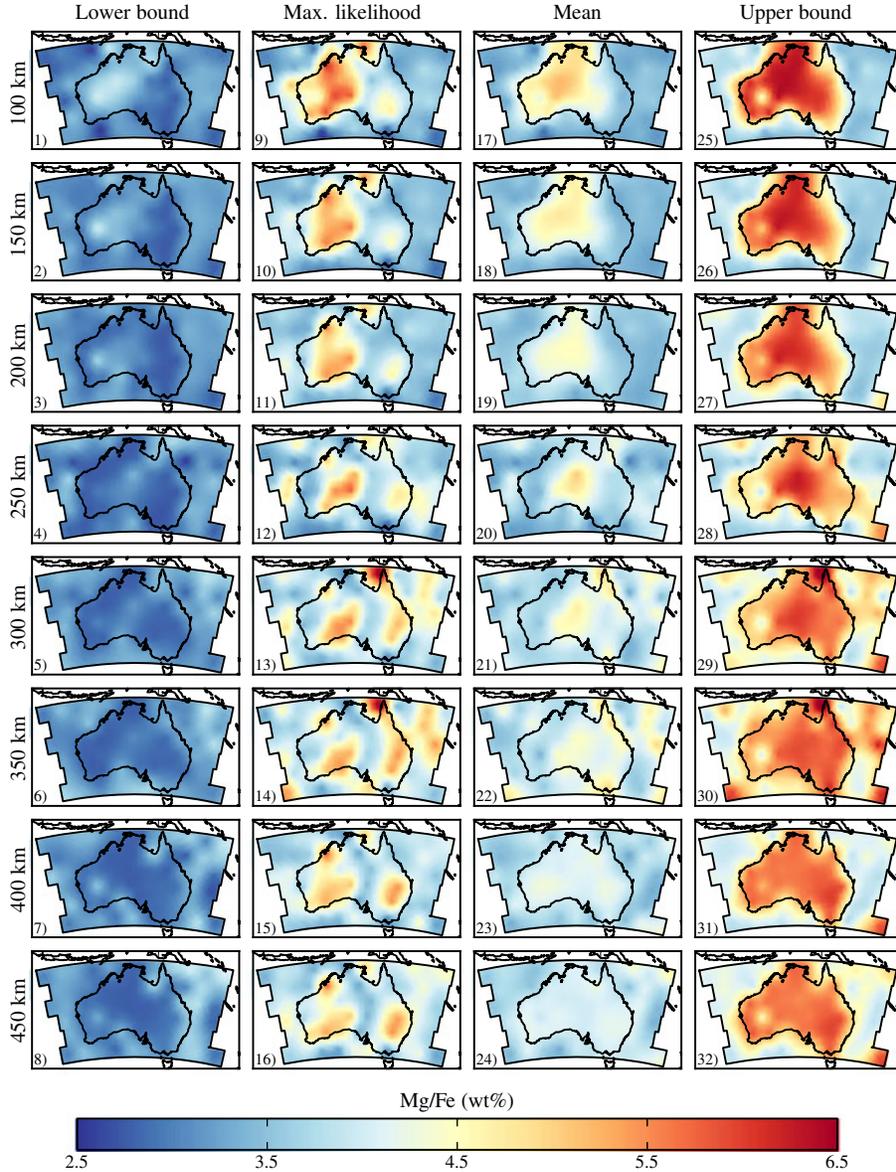
Figure 6. Maps of mantle composition showing variations in elemental Mg/Fe (wt %) at various depths underneath the Australian continent and surrounding oceans. (plots 9–16) Maximum likelihood model, (plots 17–24) mean model, (plots 1–8) lower bounds, and (plots 25–32) upper bounds, respectively, in the form of 95% highest probability density regions.

reappear with the same relative high/low values in the upper bound. While some of these features have been observed elsewhere based on purely thermal parameterizations (to be discussed in the following section), the important point here is that both compositional and temperature variations are required for the surface-wave data to be fit as illustrated in the following.

[32] As in our previous studies, we verified the robustness of the lateral compositional variations by inverting the data set using a single fixed composition throughout. As compositional model, we employed the pyrolite composition of, e.g., *Ringwood* [1962]. Performing the same inversion, but only varying mantle temperature, in addition to crustal parameters, anisotropy, and attenuation, we found that data could not be fit. In particular, we observed that throughout

most of the overtone branches were fit, while fundamental-mode and short-period part of the first overtone branch were consistently "misfit." An example of such a "misfit" at three locations is displayed in Figure 9. The figure shows how both fundamental-mode Rayleigh and Love waves, which are sensitive to upper mantle structure (to ~250–300 km depth), do not fit observations within uncertainties (compare with data fit shown in Figure 4). Only fundamental-mode branches are shown here as these are most severely affected. Also affected, albeit to a lesser extent, are the first overtone branches at short periods (not shown for clarity). Although fundamental-mode branches were observed to fit data within uncertainties in some locations, this was not generally the case. No systematic effects could be observed though.
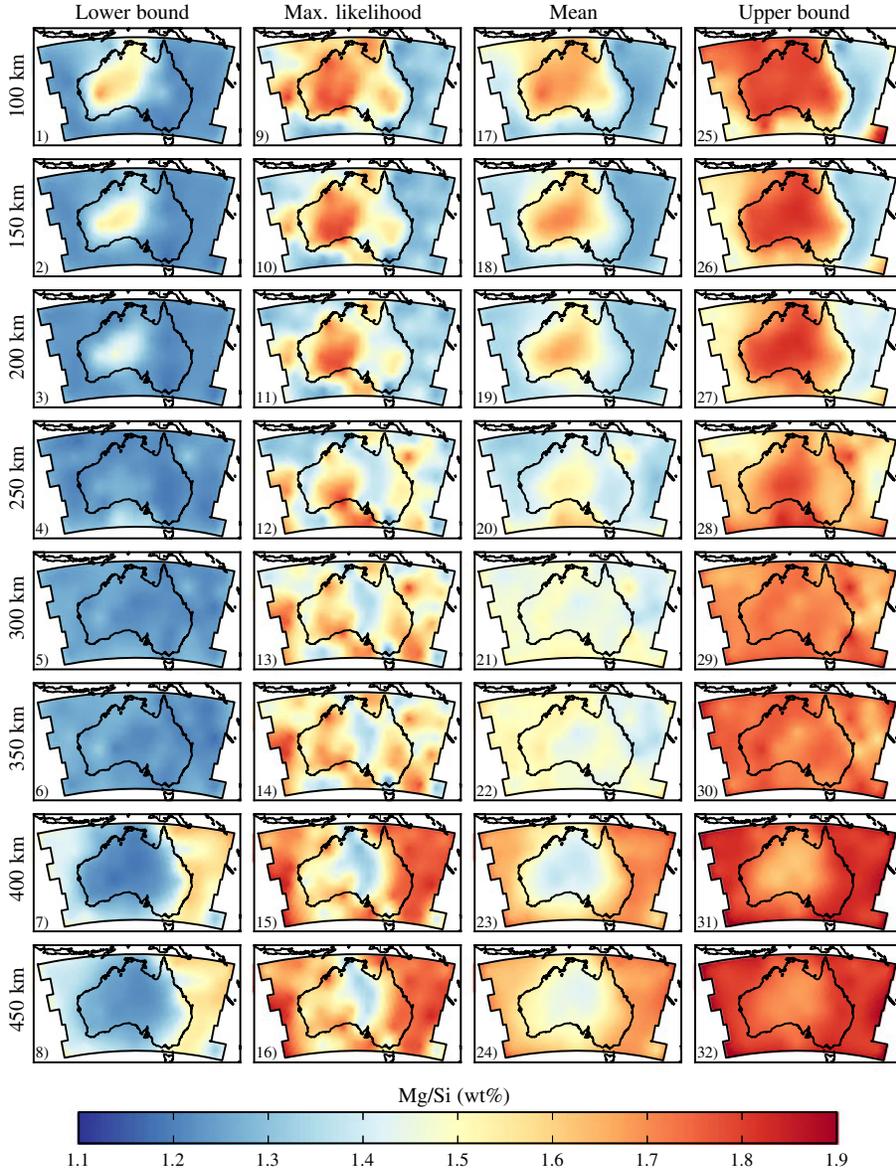
Figure 7. Maps of mantle composition showing variations in elemental Mg/Si (wt %) at various depths underneath the Australian continent and surrounding oceans. (plots 9–16) Maximum likelihood model, (plots 17–24) mean model, (plots 1–8) lower bounds, and (plots 25–32) upper bounds, respectively, in the form of 95% highest probability density regions.

[33] Overall, this test serves to highlight the relative importance of chemical variations in the upper mantle. Since overtone branches (of order >1) are generally fit within uncertainties, this suggests that chemical variations in the deeper mantle are likely to be smaller in line with the general observation that lateral velocity variations below the transition zone (>660 km depth) appear to be relatively homogeneous [e.g., *Becker and Boschi*, 2002; *Panning and Romanowicz*, 2006; *Kustowski et al.*, 2008]. Thus, chemical variations appear to be strongest in the uppermost mantle and affecting those branches most sensitive to its structure, i.e., fundamental-modes. As a final test, we also investigated the correlation between the fundamental thermochemical parameters inverted for here. As in our previous work, no large trade-offs were found signaling that temperature and composition are independently resolvable.

5.4. Comparison With Miscellaneous Studies

[34] In support of our results, we compare briefly with results from field (xenolith), regional seismic tomography, mineral physics-related studies, and experimental observations. The latter include temperatures derived from petrological experiments on mineral phase transitions in the system $(Mg,Fe)_2SiO_4$. *Ito and Takahashi* [1989], for example, found that the olivine→wadsleyite reaction (the "410 km" discontinuity) occurred at a temperature of ∼1750 ± 100 K, while in similar experiments undertaken by *Katsura et al.* [2004] temperatures of 1760 ± 45 K were obtained. These observations are in agreement with temperatures determined here that lie in the range 1723 ± 100 K across the entire continent.

[35] Relying on mineral physics and a parameterized phase-diagram approach, data thermal anomaly maps of the
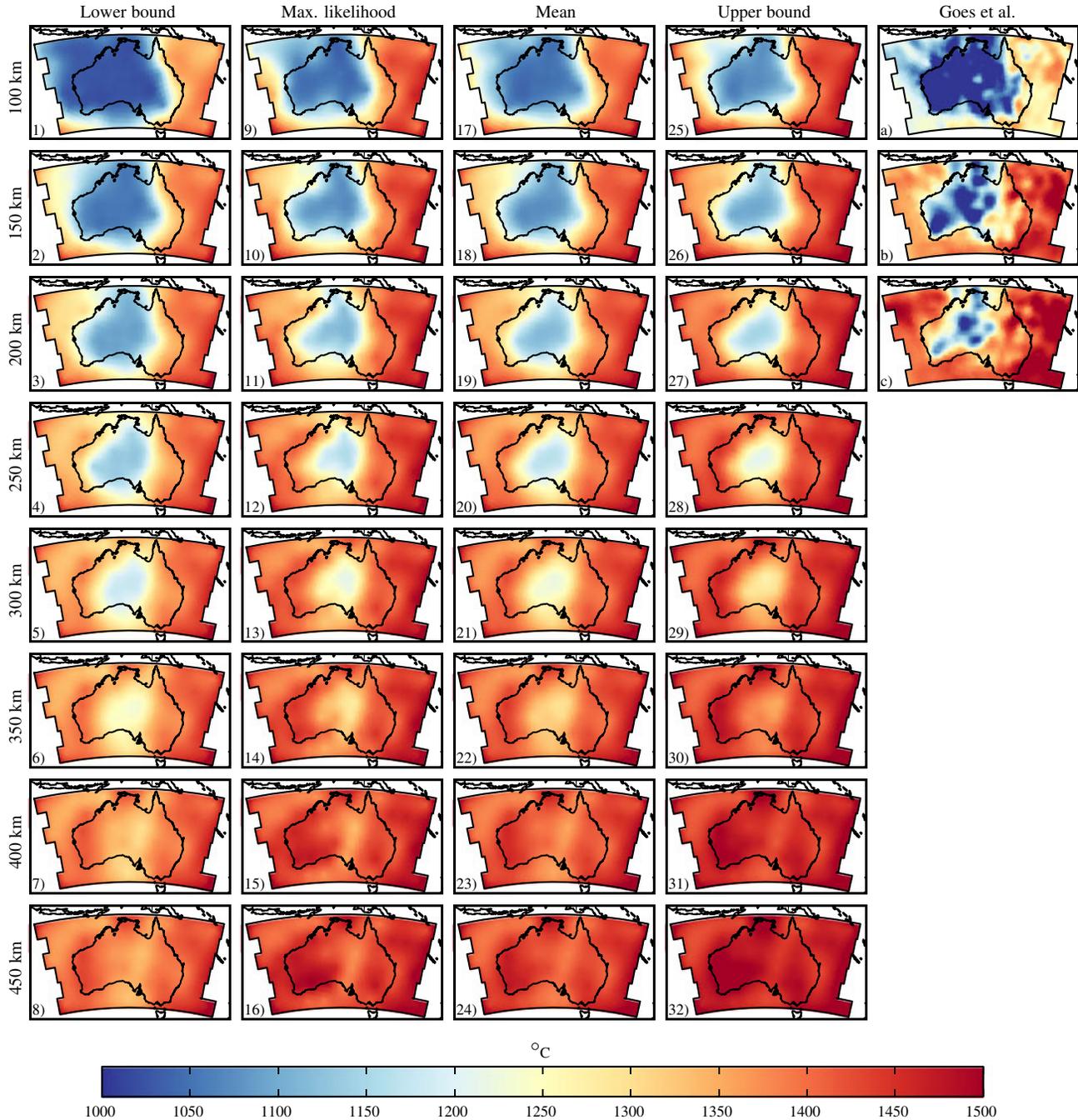
Figure 8. Maps of mantle temperature showing variations in temperature (in °C) at various depths underneath the Australian continent and surrounding oceans. (plots 9–16) Maximum likelihood model, (plots 17–24) mean model, (plots 1–8) lower bounds, and (plots 25–32) upper bounds, respectively, in the form of 95% highest probability density regions. Plots a–c show thermal maps derived by *Goes et al.* [2005]. See text for details.

upper mantle of Australia were obtained by *Goes et al.* [2005] from thermal mapping (see *Goes et al.* [2000] for details) of the regional shear-wave tomography model of *Yoshizawa and Kennett* [2004]. Their thermal maps (see Figure 8) qualitatively agree with ours with temperatures in the ranges 800°C–1100°C and 1300°C–1500°C for "cold" and "hot" provinces, respectively. In line with what is observed here, Goes et al. also find a thermal division that closely follows surface-tectonic provinces and that

qualitatively agrees with the regional surface heat flow measurements compiled by *Pollack et al.* [1993]. The latter indicates elevated heat flow underneath the western and eastern Phanerozoic provinces, and diminished heat flow beneath the western Archean and Proterozoic regions [*Pearson et al.*, 1991]. An additional model of the thermal structure of the continental lithosphere (TC1) has been developed by *Artemieva* [2006] based on analysis of various thermal, seismic, and petrologic data. In TC1, temperatures
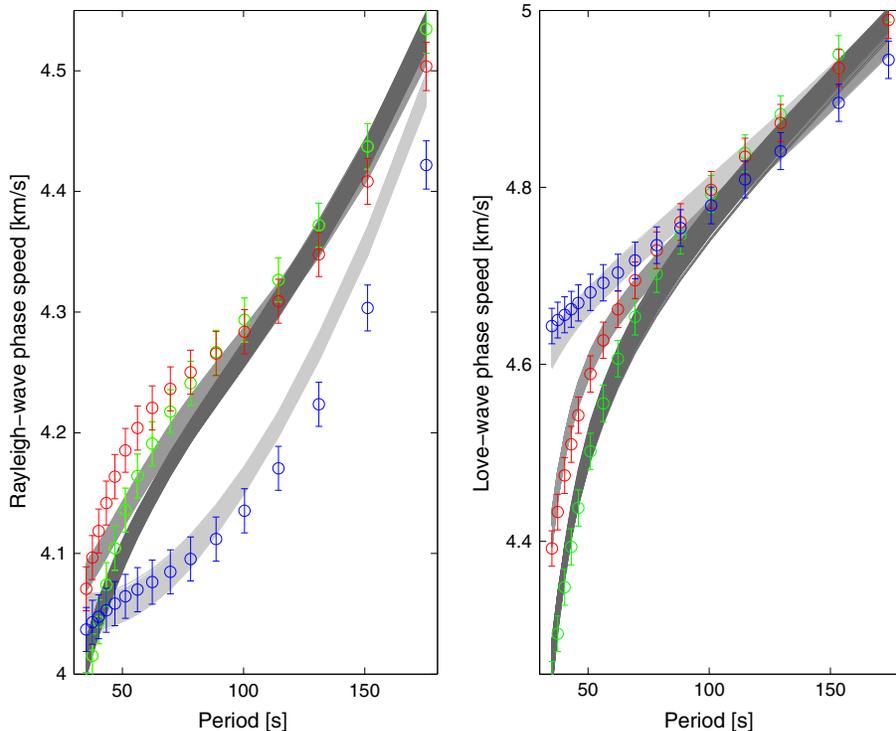
Figure 9. Comparison of calculated (gray lines) and observed Rayleigh- and Love-wave phase velocities (circles in red, blue, and green), including uncertainties (error bars in red, blue, and green) at an oceanic location (light gray/blue) and two continental locations (gray/red and dark gray/green) for the case of inversion with fixed mantle composition. Only fundamental-mode (left) Rayleigh- and (right) Love-wave branches are shown as these are consistently "misfit." For details, see main text.

around 1000°C–1100°C are found for the major cratons at a depth of 150 km (with uncertainties estimated to be around 150°C). Additional information on thermal conditions of the lithosphere derive from geothermobarometry data at a couple of locations. The temperature profiles derived from these data indicate that the lithosphere underneath Proterozoic domains is relatively cool, while non-steady state thermal conditions beneath southeastern Australia are associated with Cenozoic volcanism (e.g., *O'Reilly et al.*, 1997; *Pearson et al.*, 1991).

[36] The agreement between Goes et al. and the present study are purely qualitative given the fundamental difference in approach: thermal mapping of a seismic tomographic model versus inversion of dispersion data directly for thermochemistry. Although temperature appears relatively more well determined than composition providing an explanation for why "thermal-only" studies to a large extent can explain the observed velocity variations and obtain larger thermal variations than obtained here. Nonetheless, our results show that compositional contributions are separable and cannot be disregarded if dispersion data are to be fit (cf., Figures 4 and 9).

[37] Resolving temperatures and lithospheric depths within the various tectonic terranes is more difficult, although it appears, within model resolution, that Phanerozoic terranes are "hotter" than the older Australian Archean and Proterozoic cratonic blocks.

[38] Compositional parameters derived from garnet xenocrysts, where available, show a decreasing degree of depletion in the lithospheric mantle from the Proterozoic cratonic sections to Phanerozoic sections [e.g., *Gaul et al.*, 2000, 2003]. Moreover, the geochemical parameters further indicate a decrease in magnesium number (Mg#) with depth within the cratonic lithospheric sections from relatively depleted values (92–93) to values approaching those of the asthenosphere (89–90). This variation of Mg# with depth has also been observed beneath the Kimberley area in South Africa [*Griffin et al.*, 2008]. This change in composition with depth is not observed in the Phanerozoic sections as these are generally characterized by lower Mg#s overall. Also observed in these field studies is the lithospheric thickening of the continent from east to west, increasing from ∼100 to 150 km depth beneath Phanerozoic localities in the east to ∼150–200 and ∼200 km underneath Proterozoic regions and margins, respectively, to ∼250 km depth in the west beneath the Archean craton.

[39] Joint inversions of regional gravity data and a seismic tomographic model have also shown that the upper mantle beneath Proterozoic domains is chemically different from that beneath Phanerozoic domains [*van Gerven et al.*, 2004]. Based on the relative density-to-shear-wave velocity scaling (ζ) as a diagnostic for the presence of chemical variations [*Deschamps et al.*, 2001] (thermal effects produce positive values of ζ, whereas depletion in dense elements result in a decrease of amplitude of density anomalies, implying that small and negative values of ζ are indicative of compositional effects), *van Gerven et al.*
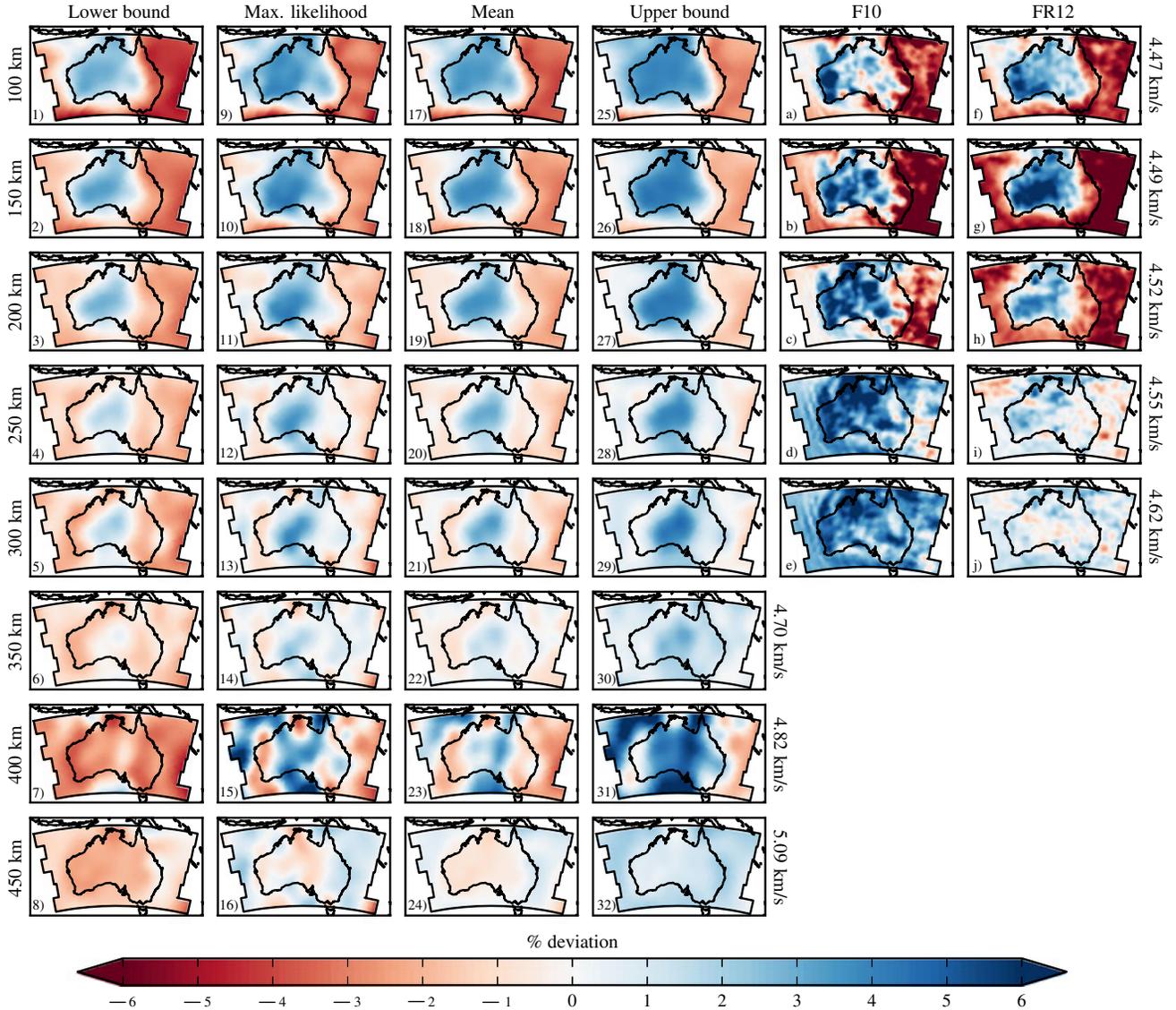
Figure 10. Maps of mantle isotropic shear-wave velocities showing variations in shear-wave speeds at various depths underneath the Australian continent and surrounding oceans. Shear-wave speeds are given in percent deviations from the mean shear-wave velocities computed here (see profile labeled AUS in Figure 11) and are indicated on the right side of the figure. Maximum likelihood model (maps 9–16), mean model (maps 17–24), lower bounds (maps 1–8), and upper bounds (maps 25–32), respectively, in the form of 95% highest probability density regions. Plots a–e depict the regional seismic full waveform tomography model of *Fichtner et al.* [2010], whereas plots f–j show the surface-wave tomography model of *Fishwick and Rawlinson* [2012].

[2004] concluded that the continental roots beneath the mainly Proterozoic regions of central Australia appeared chemically distinct from those of the Phanerozoic regions. While radial variations in ζ beneath the latter regions could be explained without having to invoke significant compositional variations in the lithosphere, an observed sharp decrease in ζ beneath the former was only compatible with a significant compositional contribution [see also *Forte and Perry*, 2000; and *Deschamps et al.*, 2001]. No robust conclusion could be reached for the mantle beneath Archean regions due to lack of resolution [see also *Kaban et al.*, 2003].

5.5. Comparison With Regional Seismic Tomography Models

[40] Isotropic shear-wave velocity variations obtained here are shown in Figure 10. Shear-wave velocity maps reported elsewhere (the Australian upper mantle seismic tomography models of, e.g., *Simmons et al.* [2009], *Debayle and Kennett* [2003], *Yoshizawa and Kennett* [2004], *Fishwick et al.* [2005, 2008], *Fichtner et al.* [2010], *Rawlinson and Fishwick* [2012], and *Fishwick and Rawlinson* [2012] to name but a few) show much the same features as our models. Here we have singled out two recent models for comparison, which include the
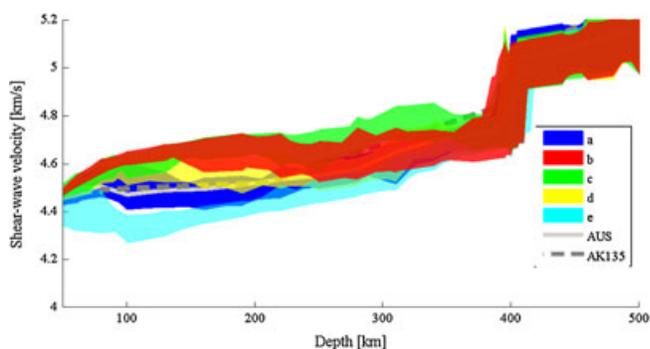
Figure 11. Shear-wave velocity profiles beneath different tectonic settings in the upper mantle and transition zone: Oceanic (pixels a and e) and continental (pixels b–d, with decreasing age progression from pixels b to d). Envelopes encompass all sampled models. Refer to Figure 2 for geographic location of pixels a–e. For reference, two mean radial models are also shown: AK135 [*Kennett et al.*, 1995] and the average profile (AUS) of the Australian continent computed as the mean of all sampled models across all pixels.

fundamental-mode surface-wave tomography model of *Fishwick and Rawlinson* [2012, hereinafter FR12] and the full waveform radially anisotropic tomography model of *Fichtner et al.* [2010, hereinafter F10]. These models are sensitive to about 300 km depth and are also shown in Figure 10. All three shear-wave models are plotted relative to an average shear-wave velocity profile computed as the mean of all sampled models across all pixels (here referred to as AUS in Figure 11).

[41] The tectonic division of slow (corresponding to "hot" and relatively low Mg/Fe and Mg/Si values) and fast (corresponding to "cold" and relatively high Mg/Fe and Mg/Si values) regimes extending to 300 km depth is distinct in the isotropic shear-wave velocity models and appears to be relatively robust inasmuch as they are apparent on both upper and lower bounds. In comparison, the tectonic division in models F10 and FR12 is less prevalent for depths >200–250 km. This is likely due to loss of resolution in the latter models at 300 km depth, although *Fichtner et al.* [2010] reported the tectonic division as extending to ~250–300 km depth. Also, in model F10 at 250–300 km depth velocity anomalies appear overall to be faster, which is related to F10 employing AK135 [*Kennett et al.*, 1995] as reference model. This is evident from Figure 11: on average AK135 is faster than AUS (the mean model derived here) below 200 km depth. A slight "ringing" in F10 can be observed off the western coast of Australia, which is believed to be an artifact due to reduced data coverage in the region as the computational costs of including events from the distant 90° East Ridge were too high (A. Fichtner, personal communication, 2012). Differences in lateral resolution notwithstanding, the various regional tomographic models shown in Figure 10, are seen to vary in detail as a result of differences in methodology, data, and parametrization. However, the overall level of agreement with present and previous regional models is encouraging.

[42] Lithospheric thickness (based on assessment of velocity gradients—not shown for brevity) beneath the

eastern Phanerozoic and oceanic regions appears to be ≤100 km, whereas the western Archean craton is seen to extend to at least 200 km depth, while the central Proterozoic regions have roots that are still observable at 300 km depth implying a conductive lithosphere that extends well into the upper mantle. There is considerable agreement between this and estimates of the depth to the base of the lithosphere derived from a variety of seismological data based on the seismic wave speed distribution in the AusREM mantle model [*Kennett et al.*, 2013], interstation-derived surface-wave dispersion data across the Australian continent [*Lebedev et al.*, 2009], and more integrated approaches reconciling thermal, seismic, and petrologic data for continental lithosphere [e.g., *Priestley and McKenzie*, 2006; *Priestley et al.*, 2006; *Kuskov et al.*, 2006, 2011; *Artemieva*, 2009]. As additional support, Figure 11 shows radial shear-wave velocity profiles for a number of locations (see Figure 2). Shear-wave velocities beneath continental locations are faster than underneath oceans as expected, with velocity differences persisting almost to the base of the upper mantle. This is in qualitative agreement with predictions of the model of *Faul and Jackson* [2005] based on experimental observations of frequency-, temperature-, and grain size-dependence of shear modulus and attenuation for olivine.

5.6. Lower Upper Mantle and Transition Zone Structure

[43] Toward the base of the upper mantle (regions not sampled by F10 and FR12), we generally observe a decrease in amplitude of shear-wave anomalies followed by a relatively strong increase at and around 400 km depth. The sudden increase in shear-wave amplitudes coincides with the mineral phase transformation of olivine to wadsleyite corresponding to the "410 km" seismic discontinuity marking the boundary between upper mantle and transition zone [e.g., *Ito and Takahashi*, 1989; *Katsura et al.*, 2004]. Lateral velocity variations are clearly larger than those immediately above and below the transition, which serves to highlight the importance of considering phase-induced variations in addition to variations arising purely from thermochemical changes that determine the structure away from mineral phase transitions [*Anderson*, 1987]. This will be discussed further in section 5.7 below.

[44] In the transition zone at 450 km depth, a reversal is seen to occur such that subcontinental mantle has become slower than suboceanic mantle. Although this feature appears less well resolved, it is nonetheless present in both lower and upper bound; the continental regions consistently appear slower than the surrounding oceanic regions. This reversal between the base of the upper mantle and the transition zone is also observed in global seismic tomography models [e.g., *Lebedev and van der Hilst*, 2008; *Schaeffer and Lebedev*, 2013]. The interesting thing to note is that the reversal in mantle shear-wave velocity structure does not appear to be related to a similar change in temperature and Mg/Fe. Thus, the apparent correlation that we observed between Mg/Fe and Mg/Si in the upper mantle is no longer extant in the transition zone. What appears to control shear-wave velocity anomalies in the absence of strong temperature gradients are variations in Mg/Si as the latter determines the relative amount of majorite (Mj) garnet
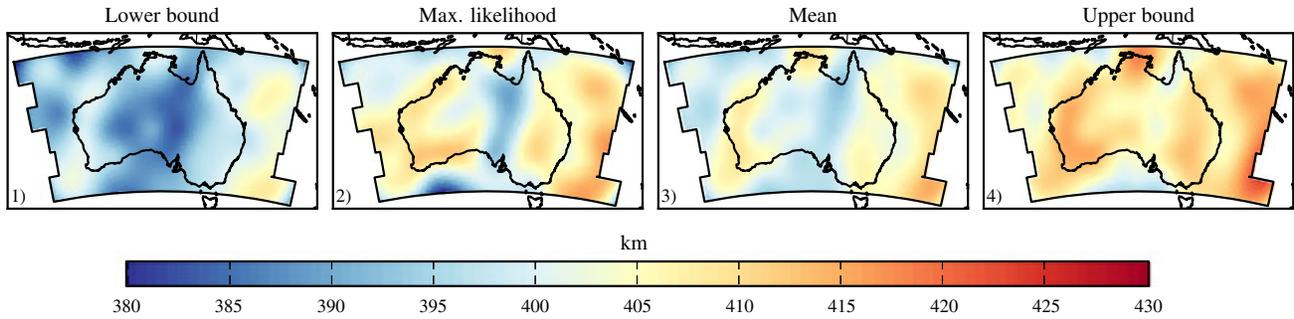
Figure 12. Depth to "410" km seismic discontinuity marking the transition between upper mantle and transition zone as defined by the mineral phase transition olivine→wadsleyite. Maximum likelihood model (plot 2), mean model (plot 3), and lower and upper bounds (plots 1 and 4) in the form of 95% highest probability density regions.

present. Mj garnet is slower than the other transition zone minerals wadsleyite and ringwoodite, and a higher proportion of this mineral (with increased Si) will thus tend to lower bulk velocities producing the observed negative shear-wave velocity anomalies. While our results point to the relative importance of Mg/Si over temperature and Mg/Fe in determining shear-wave velocity structure in the transition zone, we do also observe what appears to be a general decoupling of structure relative to that of the upper mantle apparently driven by a reversal in Mg/Si with depth in line with earlier suggestions of a chemically distinct transition zone [e.g., *Anderson and Bass*, 1986; *Ita and Stixrude*, 1992; *Anderson*, 2002].

5.7. Transition-Zone Topography

[45] Another interesting feature is the lateral variation in transition zone topography, i.e., depth of the boundary between upper mantle and transition zone (the so-called "410 km" seismic discontinuity), which is shown in Figure 12. The topography of this discontinuity appears to reveal a distinct thermal signature in that regions that are "cold" are associated with elevated "410s", whereas relatively "hot" regions have depressed "410s" in agreement with expectations across a subduction zone setting [e.g., *Bina and Helffrich*, 1994]. Generally, we find that the central part of the image from top to bottom, in addition to the region beneath the Indian Ocean (the NW of the image), shows elevated "410s" ranging from 380 to 400 km depth and decreased temperatures. Underneath much of the western and eastern coasts, and Pacific Ocean, the "410" appears depressed (~400–425 km depth) where temperatures are elevated. Effects on 410-topography arising from compositional variations, while present, are of more subtle nature and are discussed further below.

[46] Comparison with results from global receiver function studies, e.g., *Tauzin et al.* [2008], shows reasonable agreement, although it should be borne in mind that depth to interface can trade off with velocity structure immediately above rendering the interpretation of travel time residuals in terms of transition zone depth nonunique. Tauzin et al. considered four stations across the Australian continent: two stations situated on the eastern coast over Paleozoic regions, one to the north situated over a Proterozoic locality, and a station to the southwest overlying the Archean

craton. For the easternmost stations, Tauzin et al. find travel times of the phase converted at the "410" ($P_{410}s$) to be delayed relative to those predicted by a reference velocity model suggesting a depressed "410," whereas the station over the Proterozoic locality shows a faster-than-average $P_{410}s$ travel time indicating elevation of the "410." The station overlying the Archean craton is anomalous in that a slower-than-average $P_{410}s$ travel time is observed, contrary to expectations. Curiously enough, results here are compatible with this, inasmuch as most models show a depressed "410." We leave it for a future study to consider the nature of this anomaly in more detail.

[47] Another aspect of the "410" discontinuity, which has received attention concerns its nature, i.e., size and sharpness [e.g., *Duffy et al.*, 1995; *Frost*, 2003; *Helffrich and Wood*, 1996; *Helffrich*, 2000; *Katsura et al.*, 2004]. The strength of the discontinuity is essentially determined by the amount of olivine present and thus Mg/Si, whereas sharpness is controlled by size of the coexistence loop between olivine and wadsleyite. In order to match a preliminary reference Earth model-like "410," a composition with a higher basaltic component is required to produce an olivine-depleted mantle and thus smaller jump [*Duffy et al.*, 1995; *Khan et al.*, 2009]. A purely basaltic composition, for example, completely lacks the olivine→wadsleyite transition. These features are clearly apparent in the maps at 400 km depth (see Figure 10) where the strength in shear-wave velocity has increased relative to shear-wave amplitudes at 300 km depth. A purely thermal origin of lateral structure would not be compatible with this. The extent to which these amplitudes are required will, however, have to be examined using data that are more sensitive to velocity discontinuities [e.g., *Helffrich*, 2000; *Schmerr and Garnero*, 2007]. The important point to note here is that much of the structure visible in the images at the base of the upper mantle derives from the lateral variations in phase transitions associated with the olivine→wadsleyite transformation. The corresponding lateral variations in velocity are clearly observed to be much larger than those associated with purely lateral variations in temperature and composition immediately above and below the transition. Neglect of this contribution to lateral structure will likely result in lateral variations in temperature that vary too rapidly and probably also too strongly as apparent in some studies [e.g., *Goes et al.*, 2005; *Trampert et al.*, 2004].
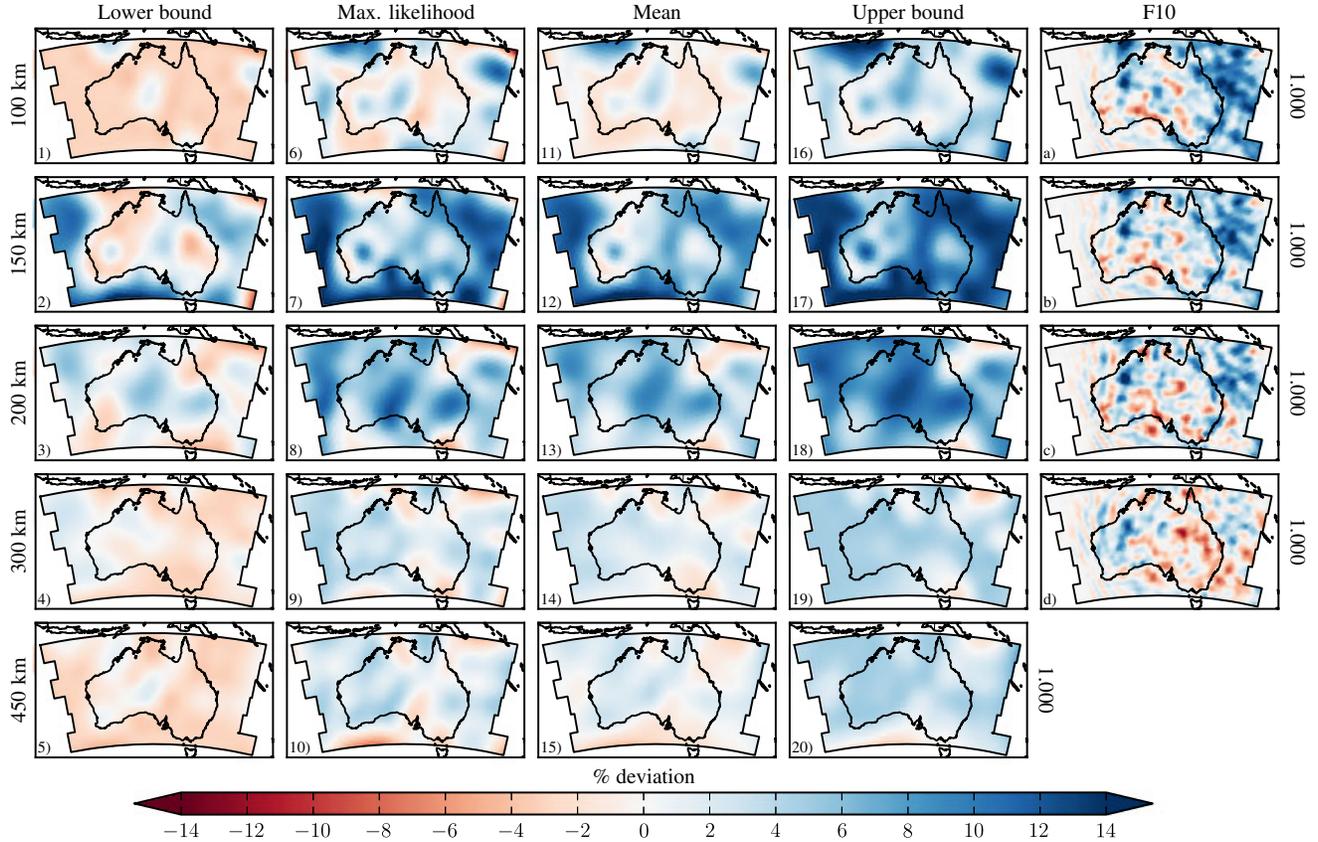
Figure 13. Maps of mantle shear-wave anisotropy variations (ξ) at various depths beneath the Australian continent and surrounding oceans. Shear-wave anisotropy is given in percent deviations from the mean (here $\xi=1$). Maximum likelihood model (plots 6-10) mean model (plots 11-15) lower bounds (plots 1-5), and upper bounds (plots 16-20), respectively, in the form of 95% highest probability density regions. Plots a–d depict shear-wave anisotropy variations in the full waveform tomography model of *Fichtner et al.* [2010].

5.8. Anisotropic Shear-Wave Velocity Variations

[48] Changes in sign of anisotropy corresponding to $\xi < 1$ ($V_{SV} > V_{SH}$) or $\xi > 1$ ($V_{SH} > V_{SV}$), can be interpreted as indicating changes from vertical to horizontal flow ($\xi < 1$) under the assumption that anisotropy is the result of a preferred orientation of the crystal lattice of the anisotropic mantle minerals as these are subjected to strains due to mantle flow. For depths above 200 km, this is interpreted as a record of horizontal flow in the lithosphere dating from the time when the cratons where undergoing deformation, i.e., prior to their stabilization. This is termed "frozen-in" anisotropy in contrast to what is observed for depths below 200 km, which is typically associated with recent or present-day flow within the asthenosphere [*Deschamps et al.*, 2008].

[49] Variations in the shear-wave anisotropy parameter ξ are shown in Figure 13. For comparison, we are also showing the radial anisotropy variations found in the model of *Fichtner et al.* [2010]. From our models, we observe the following general features:

[50] 1. The strength of anisotropy follows the general trend of previous studies in that ξ increases from shallow depths to a maximum at around 150–200 km beneath which it steadily decreases.

[51] 2. The strongest signal at 100 and 150 km depth emanates from beneath oceanic regions with predominantly positive shear-wave anisotropy, i.e., $\xi > 1$ ($V_{SH} > V_{SV}$). Underneath continental regions, we observe anisotropy to peak at 200 km depth, being dominantly characterized by $\xi > 1$. This pattern has also been observed beneath the North American continent by *Nettles and Dziewonski* [2008]. Looking at evidence provided by other global- and regional-scale studies shows that in the upper mantle to 200 km depth, radial anisotropy beneath continental areas appears to be preferentially aligned such that $\xi > 1$, implying horizontally oriented anisotropic fabric [e.g., *Gung et al.*, 2003; *Sebai et al.*, 2006; *Pedersen et al.*, 2006; *Lebedev et al.*, 2009].

[52] 3. Radial anisotropy below 200 km depth is weak but seems to be smooth overall with little apparent heterogeneity as can be discerned from both upper and lower bound. Within the uncertainties of the maps, ξ can be either weakly positive or negative.

[53] 4. Agreement between the model reported here and F10 is fair but limited to 150 km depth, implying that the ocean-continent anisotropic division reported as extending beyond 150 km depth in model F10 is likely more ambiguous. Note also that there is a considerable amount of "ringing" present in the entire eastern part of model
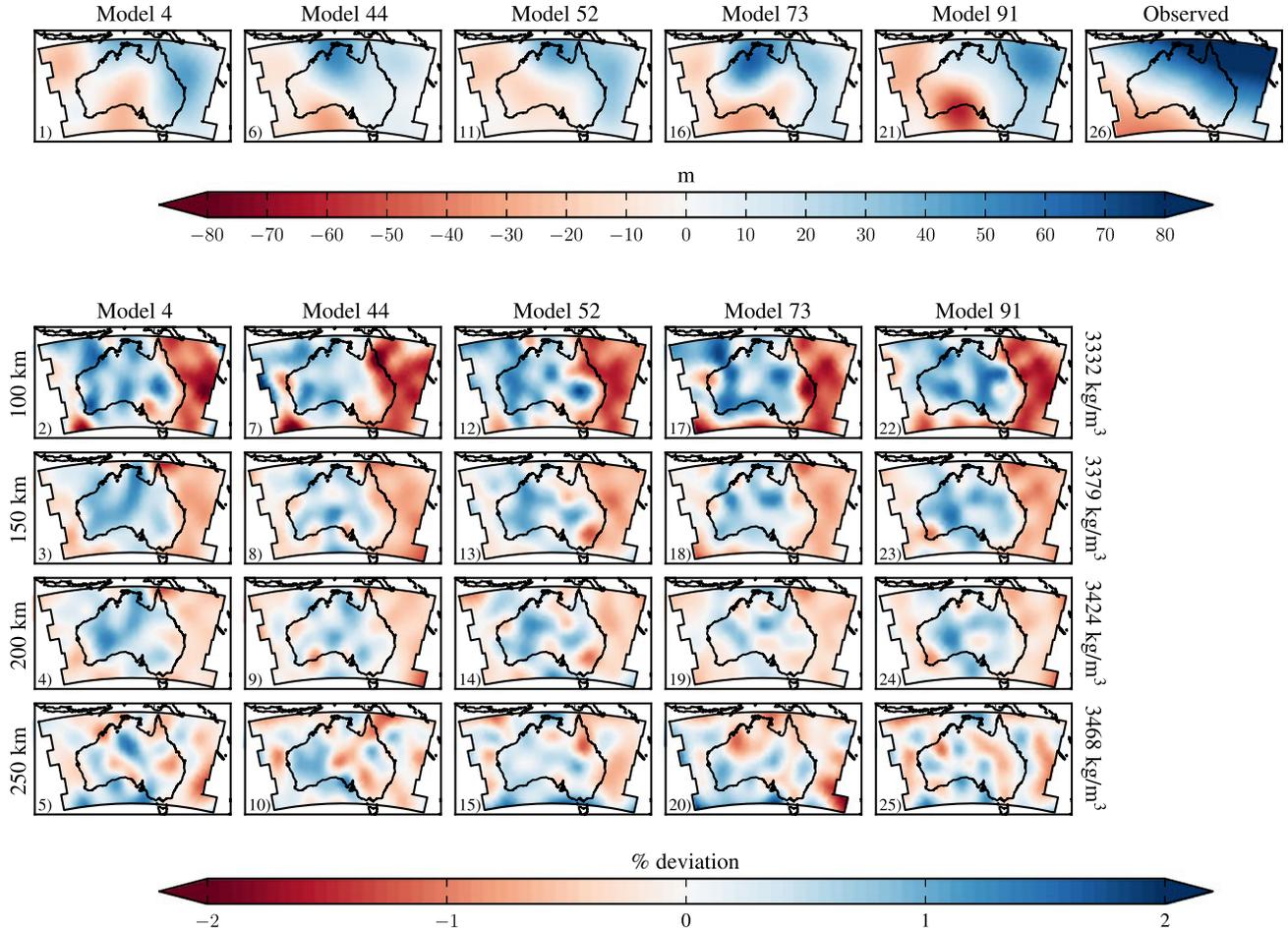
Figure 14. A selection of "best-fitting" (plots labeled "Model 4, 44, 52, 73, and 91") and corresponding models of density anomalies in the upper mantle at depths of 100, 150, 200, and 250 km depth. For comparison, observed geoid anomaly is also shown (labeled "Observed"). Data fit is measured in terms of variance reduction and correlation to observed geoid anomalies (see Table 2 and main text for details).

F10 that results from lack of data and thus sensitivity in this region, which prevents Fichtner et al. from assessing a complete model. In the context of discrepancies between anisotropic models, we find that radial anisotropy can trade-off with isotropic structure such that a large part of the data signal can be explained by isotropic variations alone as also noted by, e.g., *Fichtner et al.* [2010] and *Nettles and Dziewonski* [2008].

5.9. Posterior Filtering of Tomographic Models Using Geoid Anomalies

[54] To further refine the collection of sampled tomographic models, we propose to use additional geophysical data in the form of geoid anomalies as these are related directly to inverted density structure. For this purpose, we compute geoid anomalies for a set of sampled density distributions corresponding to a subset of thermochemical models discussed in the previous sections and compare these anomalies to the observed geoid anomalies obtained from Gravity Recovery and Climate Experiment (GRACE) [*Tapley et al.*, 2005].

[55] Geoid modeling is performed in a manner similar to our previous study [*Khan et al.*, 2011a]. Briefly, geoid

anomalies are computed by depth-integrating modeled density anomalies appropriately weighted by geoid kernels. Geoid kernels describe the response of the geoid to a density anomaly located at a given depth and take into account the dynamic contribution (the part related to mantle flow) to the geoid [e.g., *Ricard et al.*, 1984; *Hager and Richards*, 1989]. Here we use local geoid kernels, which we calculate following the method described in *van Gerven et al.* [2004] by summing up radial kernels over a prescribed spherical harmonic filter fixed in the range 2°–20°.

[56] Geoid kernels are sensitive to the viscosity structure. Here we impose a different radial viscosity profile for each tectonic province. In all provinces, we assume a viscosity jump of 30 between upper and lower mantle. The crust and lithospheric mantle are more viscous than the upper mantle by 3 and 2 orders of magnitude, respectively, and crustal and lithospheric thickness is modeled as varying with geological domain. The boundaries between crust and lithosphere, and lithosphere and upper mantle, respectively, are located at 35 and 80 km depth beneath oceanic regions, and 80 and 280 km depth beneath the Archean regions, respectively.

[57] We would like to point out that the Australian cratons appear slightly denser than the surrounding areas.

Table 2. Summary of Data Fits for the Five "Best-Fitting" Computed Geoid Maps[a]

Model #	dN^{obs} (m)	dN^{obs}_{rms} (m)	dN^m (m)	dN^m_{rms} (m)	Variance Reduction	Correlation
4	30.452	50.079	6.739	21.371	0.466	0.753
44	30.452	50.079	7.146	19.615	0.449	0.740
52	30.452	50.079	6.955	20.360	0.477	0.781
73	30.452	50.079	7.927	24.481	0.465	0.697
91	30.452	50.079	1.980	28.103	0.459	0.816

[a]The geoid anomaly maps are shown in Figure 14. dN^{obs}, dN^{obs}_{rms}, dN^m, and dN^m_{rms} are the average observed geoid anomaly, the RMS observed geoid anomaly, the average modeled geoid anomaly, and the RMS modeled geoid anomaly, respectively. See main text for further details.

The relatively low temperature observed in these regions (Figure 8) is not fully compensated by the depletion in iron (Figure 6). Overall, this results in positive density anomalies (Figure 14), which may induce negative buoyancy and contradict the tectosphere hypothesis [*Jordan*, 1975]. Stability analysis of continental roots [*Lenardic and Moresi*, 1999; *Deschamps et al.*, 2002; *Jaupart et al.*, 2007], however, indicates that negative buoyancy induced by cooling of a root does not need to be fully compensated by chemical density anomalies for the root to remain stable. Conditions for stability can be estimated in terms of a critical buoyancy ratio (B_c), the root being stable if the observed buoyancy ratio is larger than B_c. *Lenardic and Moresi* [1999] showed that this buoyancy ratio depends on the dimension of the root. In particular, the stability of the root increases (B_c decreases) with increasing thickness and length of the chemical root. Furthermore, larger viscosity and lateral variations in viscosity (related to temperature anomalies for instance) strongly stabilize the root [*Deschamps et al.*, 2002]. Practically, the critical buoyancy ratio may be rather low, 0.5 or less, i.e., the root may remain stable even if the thermal density contrast is twice the chemical density contrast or more.

[58] Computed geoids for the five "best-fitting" models including the density anomaly maps employed are displayed in Figure 14. For comparison, the observed geoid is also shown. On average, sampled density models explain the observed anomalies with a variance reduction in the range 0.2–0.3, whereas for the "best-fitting" models, correlations and variance reductions >0.6 and >0.4 (up to 0.75 and 0.46), respectively, are achieved (see Table 2). Overall, the modeled geoid is smaller in amplitude than the observed geoid. For the spherical harmonic filter employed, RMS of the observed geoid is around 50 m, while that of the modeled geoids varies between 10 and 40 m, depending on the particular density model considered (note that RMS for the best-fitting models is within 20–30 m; see Table 2). This discrepancy may be due to the fact that the observed geoid contains signal that originates from deep within the mantle, which is not imaged in our inversion. Other possible sources of error include the nature of the computed geoid kernels, which depend on the particular viscosity profile employed. Furthermore, our approach does not take into account possible lateral variations of viscosity within each tectonic province, which may substantially influence the geoid kernels.

[59] Thus, while *a posteriori* filtering using geoid anomalies presents an important means for positively identifying, among the set of all sampled (density) models those

that are consistent with gravity data, the most important limitation that we presently face is the influence of the (unknown) three-dimensional (3-D) viscosity model. Future efforts incorporating more realistic viscosity estimates will, however, most probably have to await improvements in experimental determinations [e.g., *Korenaga and Karato*, 2008], in addition to the use of 3-D geodynamical methods based on finite-difference to model the geoid.

6. Summary and Conclusion

[60] In this study, we have employed a combined geophysical-thermodynamical approach to estimating mantle composition (major element chemistry) and thermal state beneath the Australian continent by inverting fundamental-mode and higher-order surface-wave phase velocities for a regional tomographic model of the upper mantle. Geophysical data are linked to physical rock properties (density, P, and S wave speed) and thermochemistry through the use of thermodynamics providing a scheme for computing mantle mineral phase equilibria that depend only on composition, temperature, and pressure. The advantage of this approach lies in the combination of geophysical, mineral physics, and petrological information to constrain, for example, location, size, and sharpness of mantle discontinuities that would otherwise not be resolvable. With this method, we thus obtain models of mantle conditions that simultaneously combine features of both laboratory and geophysical data. Moreover, and unlike more traditional approaches, the statistical nature of the inverse method allows us to compute rigorous uncertainty estimates on all inverted parameters. Emphasis here is on drawing inferences from an assembly of models rather than a single image.

[61] Main findings and conclusions revealed by our study are as follows:

[62] 1. To fit geophysical data, we find that both compositional and thermal anomalies appear to be required in the upper mantle and upper part of the transition zone to 450 km depth,

[63] 2. Tests have shown that models relying solely on thermal variations appear not to fit fundamental-mode Rayleigh- and Love-wave data within uncertainties that are sensitive to upper mantle structure.

[64] 3. The thermochemical structure of the upper mantle follows the surface-tectonic age division closely to depths of 200–250 km. In particular, we observe the older western and central parts of the continent to be relatively cold, with high Mg/Fe and Mg/Si values, whereas the younger eastern and oceanic regions are comparatively hot with low Mg/Fe and Mg/Si.

[65] 4. Across the "410 km" seismic discontinuity, we observed a strong contribution to lateral variations in structure from thermochemically induced phase transformations that appear to be much stronger than those arising from lateral variations in composition and temperature only.

[66] 5. In the transition zone, we observe a general decoupling of the structure relative to that of the upper mantle apparently driven by an increase in Mg/Si with depth, which appears uncorrelated to Mg/Fe.

[67] 6. The strongest anisotropic signal occurs in the shallow upper mantle at 150 km depth. It is strongest beneath oceanic regions being preferentially horizontally oriented,

whereas underneath continental regions radial anisotropy peaks at 200 km depth.

[68] 7. Radial anisotropy below 200 km depth is less resolved but appears to be weak and laterally uniform across the Australian region.

[69] 8. The geoid anomalies computed on the basis of the obtained density models are largely compatible with the observed geoid anomalies of GRACE.

[70] 9. Finally, comparison of our laboratory-based tomographic shear-wave speed models with purely geophysically derived models shows good agreement and supports the methodology adopted here.

[71] We consider the present approach a step toward resolving the seismic tomography problem using fully non-linear inverse strategies. In this respect, it is far superior to the standard techniques, centered on linearized methods, that are employed in present-day seismic tomography because of its ability to provide rigorous estimates of uncertainty and resolution. However, in exchange for this, sampling-based strategies suffer from the limitations imposed by their immense computational requirements. To overcome these limitations and to render the problem tractable modeling simplifications are required. In the present case, this limitation is reflected in the use of surface-wave phase-velocity maps rather than the original phase-velocity data from which the former are constructed. Future application of the present strategy should consider the original phase-velocity data in order that the fundamental tomographic problem be solved from the outset. As an alternative, stochastic waveform inversion as initiated by *Käufl et al.* [2013] presents another interesting avenue. This, however, will likely necessitate, even on a regional-scale level, diminishing lateral resolution further. However, it should be emphasized that with the use of sampling-based strategies comes the ability to perform uncertainty analysis, which, ultimately, determines the reliability of results from inverse calculations and thereby our confidence in these.

References

Afonso, J. C., G. Ranalli, M. Fernandez, W. L. Griffin, S. Y. O'Reilly, and U. Faul (2010), On the Vp/Vs - Mg# correlation in mantle peridotites: Implications for the identification of thermal and compositional anomalies in the upper mantle, *Earth Planet. Sci. Lett.*, *289*, 606–618, doi:10.1016/j.epsl.2009.12.005.

Afonso, J. C., J. Fullea, W. L. Griffin, Y. Yang, A. G. Jones, J. A. D. Connolly, and S. Y. O'Reilly (2013a), 3D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle I: a priori information and geophysical observables, *J. Geophys. Res. Solid Earth*, *118*, 2586–2617, doi:10.1002/jgrb.50124.

Afonso, J. C., J. Fullea, Y. Yang, J. A. D. Connolly, and A. G. Jones (2013b), 3D multi-observable probabilistic inversion for the compositional and thermal structure of the lithosphere and upper mantle II: General methodology and resolution analysis, *J. Geophys. Res. Solid Earth*, *118*, 1650–1676, doi:10.1002/jgrb.50123.

Anderson, D. L. (1987), Thermally induced phase changes, lateral heterogeneity of the mantle, continental roots, and deep slab anomalies, *J. Geophys. Res.*, *92*, 13,968–13,980.

Anderson, D. L. (1989), *Theory of the Earth*, Blackwell, Oxford.

Anderson, D. L. (2002), The case for irreversible chemical stratification of the mantle, *Int. Geol. Rev.*, *44*, 97–116.

Anderson, D. L., and A. M. Dziewonski (1982), Upper mantle anisotropy: Evidence from free oscillations, *Geophys. J. R. Astron. Soc.*, *69*, 383.

Anderson, D. L., and J. D. Bass (1986), Transition region of the Earth's upper mantle, *Nature*, *320*, 321–328.

Arroucau, P., N. Rawlinson, and M. Sambridge (2010), New insight into Cainozoic sedimentary basins and Palaeozoic suture zones in southeast Australia from ambient noise surface wave tomography, *Geophys. Res. Lett.*, *37*, L07303, doi:10.1029/2009GL041974.

Artemieva, I. M. (2006), Global 1°× 1 ° thermal model TC1 for the continental lithosphere: Implications for lithosphere secular evolution, *Tectonophysics*, *416*, 245–277.

Artemieva, I. M. (2009), The continental lithosphere: Reconciling thermal, seismic, and petrologic data, *Lithos*, *109*(1-2), 23–46.

Babuska, V., and M. Cara (1991), *Seismic Anisotropy in the Earth*, Kluwer Academic Press, Boston.

Bass, J. D. (1995), Elastic properties of minerals, melts, and glasses, in *Handbook of Physical Constants*, edited by T. J. Ahrens, pp. 45–63, AGU, Washington, D. C.

Bassin, C., G. Laske, and G. Masters (2000), The current limits of resolution for surface-wave tomography in North America, *Eos Trans. AGU*, *81*(48), Fall Meet. Suppl., Abstract S12A-03.

Becker, T. W., and L. Boschi (2002), A comparison of tomographic and geodynamic mantle models, *Geochem. Geophys. Geosyst.*, *3*(1), 1003, doi:10.1029/2001GC000168.

Becker, T. W., B. Kustowski, and G. Ekström (2008), Radial seismic anisotropy as a constraint for upper mantle rheology, *Earth Planet. Sci. Lett.*, *267*, 213–227.

Bernardo, J. M., and A. F. M. Smith (1994), *Bayesian Theory*, John Wiley, Chichester, U. K.

Betts, P. G., D. Giles, G. S. Lister, and L. R. Frick (2002), Evolution of the Australian lithosphere, *Austral. J. Earth Sci.*, *49*, 661.

Bina, C., and G. R. Helffrich (1994), Phase transition Clapeyron slopes and transition zone seismic discontinuity topography, *J. Geophys. Res.*, *99*(B8), 15,853–15,860.

Bodin, T., and M. Sambridge (2009), Seismic tomography with the reversible jump algorithm, *Geophys. J. Int.*, *178*, 1411–1436, doi:10.1111/j.1365-246X.2009.04226.x.

Bodin, T., M. Sambridge, N. Rawlinson, and P. Arroucau (2012a), Transdimensional tomography with unknown data noise, *Geophys. J. Int.*, *189*, 1536–1556, doi:10.1111/j.1365-246X.2012.05414.x.

Bodin, T., M. Salmon, B. L. N. Kennett, and M. Sambridge (2012b), Probabilistic surface reconstruction from multiple data sets: An example for the Australian Moho, *J. Geophys. Res.*, *117*, B10307, doi:10.1029/2012JB009547.

Bosch, M. (1999), Lithologic tomography: From plural geophysical data to lithology estimation, *J. Geophys. Res.*, *B1*, 749–766.

Boschi, L., and G. Ekström (2002), New images of the Earth's upper mantle from measurements of surface wave phase velocity anomalies, *J. Geophys. Res.*, *107*(B4), 2059, doi:10.1029/2000JB000059.

Box, G., and G. Tiao (1992), *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading, Mass.

Bozdag, E., and J. Trampert (2008), On crustal corrections in surface wave tomography, *Geophys. J. Int.*, *172*, 1066–1082, doi:10.1111/j.1365-246X.2007.03690.x.

Cammarano, F., and B. Romanowicz (2007), Insights into the nature of the transition zone from physically constrained inversion of long-period seismic data, *Proc. Natl. Acad. Sci. U.S.A.*, *104*, 9139–9144.

Cammarano, F., B. Romanowicz, L. Stixrude, C. Lithgow-Bertelloni, and W. Xu (2009), Inferring the thermochemical structure of the upper mantle from seismic data, *Geophys. J. Int.*, *179*, 1169–1185.

Cammarano, F., P. Tackley, and L. Boschi (2011), Seismic, petrological and geodynamical constraints on thermal and compositional structure of the upper mantle: Global thermo-chemical models, *Geophys. J. Int.*, *187*, 1301–1318.

Chen, J., T. Inoue, H. Yurimoto, and D. J. Weidner (2002), Effect of water on olivine-wadsleyite phase boundary in the (Mg, Fe)$_2$SiO$_4$ system, *Geophys. Res. Lett.*, *29*(18), 1875, doi:10.1029/2001GL014429.

Chen, M.-H., and Q.-M. Shao (1999), Monte Carlo estimation of Bayesian credible and HPD intervals, *J. Comput. Graph. Stat.*, *8*, 69–92.

Cobden, L., S. Goes, F. Cammarano, and J. A. D. Connolly (2008), Thermochemical interpretation of one-dimensional seismic reference models for the upper mantle: Evidence for bias due to heterogeneity, *Geophys. J. Int.*, *175*, 417–830, doi:10.1111/j.1365-246X.2008.03903.x.

Connolly, J. A. D. (2005), Computation of phase equilibria by linear programming: A tool for geodynamic modeling and an application to subduction zone decarbonation, *Earth Planet. Sci. Lett.*, *236*, 524–541.

Debayle, E., and B. L. N. Kennett (2003), Surface-wave studies of the Australian region, *Soc. Australia Spec. Publ. 22, and Geol. Soc. America Spec. Pap.*, *372*, 25–40.

Debayle, E., B. N. L. Kennett, and K. Priestley (2005), Global azimuthal seismic anisotropy and the unique plate-motion deformation of Australia, *Nature*, *433*, 509–512.

Deschamps, F., R. Snieder, and J. Trampert (2001), The relative density-to-shear velocity scaling in the uppermost mantle, *Phys. Earth Planet. Inter.*, *124*, 193–211.

Deschamps, F., J. Trampert, and R. Snieder (2002), Anomalies of temperature and iron in the uppermost mantle inferred from gravity data and tomographic models, *Phys. Earth Planet. Inter.*, *129*, 245–264.

Deschamps, F., S. Lebedev, T. Meier, and J. Trampert (2008), Stratified seismic anisotropy reveals past and present deformation beneath the East-central United States, *Earth Planet. Sci. Lett.*, *274*, 489–498, doi:10.1016/j.epsl.2008.07.058.

Duffy, T. S., C.-S. Zha, R. T. Downs, M. Ho-kwang, and J. H. Russell (1995), Elasticity of forsterite to 16 GPa and the composition of the upper-mantle, *Nature*, *378*, 170–173.

Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Phys. Earth Planet. Inter.*, *25*, 297–356.

Faul, U. H., and I. Jackson (2005), The seismological signature of temperature and grain size variations in the upper mantle, *Earth Planet. Sci. Lett.*, *234*, 119–134, doi:10.1016/j.epsl.2005.02.008.

Fichtner, A., B. L. N. Kennett, H. Igel, and H.-P. Bunge (2010), Full waveform tomography for radially anisotropic structure: New insights into present and past states of the Australasian upper mantle, *Earth Planet. Sci. Lett.*, *290*, 270–280, doi:10.1016/j.epsl.2009.12.003.

Fishwick, S., and N. Rawlinson (2012), Structure of the Australian lithosphere: Evolving tomographic models from evolving seismic datasets, *Aust. J. Earth Sci.*, *59*, 809–826.

Fishwick, S., B. L. N. Kennett, and A. M. Reading (2005), Contrasts in lithospheric structure within the Australian craton—Insights from surface wave tomography, *Earth Planet. Sci. Lett.*, *231*, 163–176, doi:10.1016/j.epsl.2005.01.009.

Fishwick, S., M. Heintz, B. L. N. Kennett, A. M. Reading, and K. Yoshizawa (2008), Steps in lithospheric thickness within eastern Australia, evidence from surface wave tomography, *Tectonics*, *27*, doi:10.1029/2007TC002116..

Forte, A. M. (2000), Seismic-geodynamic constraints on mantle flow: Implications for layered convection, mantle viscosity, and seismic anisotropy in the deep mantle, in *Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale, Geophys. Monogr. Ser.* vol. 117, edited by S.-I. Karato et al., pp. 3–36, AGU, Washington, D. C., doi:10.1029/GM117p0003.

Forte, A. M., and C. A. Perry (2000), Seismic-geodynamic evidence for a chemically depleted continental tectosphere, *Science*, *290*, 1940–1944.

Frost, D. J. (2003), The structure and sharpness of $(Mg,Fe)_2SiO_4$ phase transformations in the transition zone, *Earth Planet. Sci. Lett.*, *216*, 313–328.

Fullea, J., S. Lebedev, M. R. Agius, A. G. Jones, and J. C. Afonso (2012), Lithospheric structure in the Baikal-central Mongolia region from integrated geophysical-petrological inversion of surface-wave data and topographic elevation, *Geochem. Geophys. Geosyst.*, *13*, Q0AK09, doi:10.1029/2012GC004138.

Gaul, O. F., W. L. Griffin, S. Y. O'Reilly, and N. J. Pearson (2000), Mapping olivine composition in the lithospheric mantle, *Earth Planet. Sci. Lett.*, *182*, 223–235.

Gaul, O. F., S. Y. O'Reilly, and W. L. Griffin (2003), Lithosphere structure and evolution in southeastern Australia, *Geol. Soc. Australia Spec. Publ. 22, and Geol. Soc. America Spec. Pap.*, *372*, 185–202.

Goes, S., R. Govers, and P. Vacher (2000), Shallow mantle temperatures under Europe from *P* and *S* wave tomography, *J. Geophys. Res.*, *105*, 11,153–11,169.

Goes, S., F. J. Simons, and K. Yoshizawa (2005), Seismic constraints on temperature of the Australian uppermost mantle, *Earth Planet. Sci. Lett.*, *236*, 227–237.

Griffin, W. L., S. Y. O'Reilly, J. C. Afonso, and G. C. Begg (2008), The composition and evolution of lithospheric mantle: A re-evaluation and its tectonic implications, *J. Petrol.*, *50*, 1185–1204, doi:10.1093/petrology/egn033.

Gung, Y., M. Panning, and B. Romanowicz (2003), Global anisotropy and the thickness of continents, *Nature*, *442*, 707–710.

Hager, B. H., and M. A. Richards (1989), Long-wavelength variations in Earth's geoid: Physical models and dynamical implications, *Philos. Trans. R. Soc. London, Ser. A*, *328*, 309–327.

Helffrich, G. R., and B. J. Wood (1996), 410 km discontinuity sharpness and the form of the olivine α–β phase diagram: Resolution of apparent seismic contradictions, *Geophys. J. Int.*, *126*, F7–F12.

Helffrich, G. R. (2000), Topography of the transition zone seismic discontinuities, *Rev. Geophys.*, *38*, 141–158.

Hieronymous, C. F., and S. Goes (2010), Complex cratonic seismic structure from thermal models of the lithosphere: Effects of variations in deep radiogenic heating, *Geophys. J. Int.*, *180*, 999–1012, doi:10.1111/j.1365-246X.2009.04478.x.

Hunter, J. D. (2007), Matplotlib: A 2D graphics environment, *Comp. Sci. Eng.*, *9*(3), 90–95.

Ita, J., and L. Stixrude (1992), Petrology, elasticity, and composition of the mantle transition zone, *J. Geophys. Res.*, *97*(B5), 6849–6866.

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

Jackson, I., J. D. Fitz Gerald, U. H. Faul, and B. H. Tan (2002), Grain-size-sensitive seismic wave attenuation in polycrystalline olivine, *J. Geophys. Res.*, *107*(B12), 2360, doi:10.1029/2001JB001225.

Jackson, I., and U. H. Faul (2010), Grainsize-sensitive viscoelastic relaxation in olivine: Towards a robust laboratory-based model for seismological application, *Phys. Earth Planet. Inter.*, *183*, 151–163, doi:10.1016/j.pepi.2010.09.005.

James, D. E., F. R. Boyd, D. Schutt, D. R. Bell, and R. W. Carlson (2004), Xenolith constraints on seismic velocities in the upper mantle beneath southern Africa, *Geochem. Geophys. Geosyst.*, *5*, Q01002, doi:10.1029/2003GC000551.

Jaupart, C., P. Molnar, and E. Cottrell (2007), Instability of a chemically dense layer heated from below and overlain by a deep less viscous fluid, *J. Fluid Mech.*, *572*, 433–469.

Jordan, T. H. (1975), The continental tectosphere, *Rev. Geophys. Space Phys.*, *13*, 1–12.

Kaban, M. K., P. Schwintzer, I. M. Artemieva, and W. D. Mooney (2003), Density of continental roots: Compositional and thermal contributions, *Earth Planet. Sci. Lett.*, *209*, 53–69.

Karato, S.-I., and H. Jung (1998), Water, partial melting and the origin of seismic low velocity and high attenuation zone in the upper mantle, *Earth Planet. Sci. Lett.*, *157*, 193–207.

Katsura, T., et al. (2004), Olivine-wadsleyite transition in the system $(Mg,Fe)_2SiO_4$, *J. Geophys. Res.*, *109*, B02209, doi:10.1029/2003JB002438.

Käufl, P., A. Fichtner, and H. Igel (2013), Probabilistic full waveform inversion based on tectonic regionalization—Development and application to the Australian upper mantle, *Geophys. J. Int.*, *193*, 437–451, doi:10.1093/gji/ggs131.

Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995), Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, *122*, 108–124.

Kennett, B. L. N., and I. Jackson (2009), Optimal equations of state for mantle minerals from simultaneous non-linear inversion of multiple datasets, *Phys. Earth Planet. Inter.*, *176*, 98–108, doi:10.1016/j.pepi.2009.04.005.

Kennett, B. L. N., M. Salmon, E. Saygin, and AusMoho Working Group (2011), AusMoho: The variation of Moho depth in Australia, *Geophys. J. Int.*, *187*, 946–958, doi:10.1111/j.1365-246X.2011.05194.x.

Kennett, B. L. N., A. Fichtner, S. Fishwick, and K. Yoshizawa (2013), Australian Seismological Reference Model (AuSREM): Mantle component, *Geophys. J. Int.*, *192*(2), 871–887, doi:10.1093/gji/ggs065.

Khan, A., J. A. D. Connolly, J. Maclennan, and K. Mosegaard (2007), Joint inversion of seismic and gravity data for lunar composition and thermal state, *Geophys. J. Int.*, *168*, 243–258, doi:10.1111/j.1365-246X.2006.03200.x.

Khan, A., L. Boschi, and J. A. D. Connolly (2009), On mantle chemical and thermal heterogeneities and anisotropy as mapped by inversion of global surface wave data, *J. Geophys. Res.*, *114*, B09305, doi:10.1029/2009JB006399.

Khan, A., A. Zunino, and F. Deschamps (2011a), The thermo-chemical and physical structure beneath the North American continent from Bayesian inversion of surface-wave phase velocities, *J. Geophys. Res.*, *116*, B09304, doi:10.1029/2011JB008380.

Khan, A., L. Boschi, and J. A. D. Connolly (2011b), Mapping the Earth's thermochemical and anisotropic structure using global surface wave data, *J. Geophys. Res.*, *116*, B01301, doi:10.1029/2010JB007828.

Khan, A., and T. J. Shankland (2012), A geophysical perspective on mantle water content and melting: Inverting electromagnetic sounding data using laboratory-based electrical conductivity profiles, *Earth Planet. Sci. Lett.*, *317–318*, 27–43, doi:10.1016/j.epsl.2011.11.031.

Korenaga, J., and S.-I. Karato (2008), A new analysis of experimental data on olivine rheology, *J. Geophys. Res.*, *113*, B02403, doi:10.1029/2007JB005100.

Kuskov, O. L., V. A. Kronrod, and H. Annersten (2006), Inferring upper-mantle temperatures from seismic and geochemical constraints: Implications for Kaapvaal craton, *Earth Planet. Sci. Lett.*, *244*, 133–154, doi:10.1016/j.epsl.2006.02.016.

Kuskov, O. L., V. A. Kronrod, and A. A. Prokof'ev (2011), Thermal structure and thickness of the lithospheric mantle underlying the Siberian Craton from the kraton and kimberlit superlong seismic profiles, *Phys. Solid Earth*, *47*, 155–175, doi:10.1134/S1069351310111011.

Kustowski, B., G. Ekström, and A. M. Dziewonski (2008), Anisotropic shear-wave velocity structure of the Earth's mantle: A global model, *J. Geophys. Res.*, *113*, B06306, doi:10.1029/2007JB005169.

Lebedev, S., and R. D. van der Hilst (2008), Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms, *Geophys. J. Int.*, *173*, 505–518, doi:10.1111/j.1365-246X.2008.03721.x.

Lebedev, S., J. Boonen, and J. Trampert (2009), Seismic structure of Precambrian lithosphere: New constraints from broadband surface-wave dispersion, *Lithos*, *109*, 96–111, doi:10.1016/j.lithos.2008.06.010.

Lenardic, A., and L.-N. Moresi (1999), Some thoughts on the stability of cratonic lithosphere: Effects of buoyancy and viscosity, *J. Geophys. Res.*, *104*, 12,747–12,748.

Lyubetskaya, T., and J. Korenaga (2007), Chemical composition of Earth's primitive mantle and its variance: 1. Method and results, *J. Geophys. Res.*, *112*, B03211, doi:10.1029/2005JB004223.

McKenzie, D. P., J. A. Jackson, and K. F. Priestley (2005), Thermal structure of oceanic and continental lithosphere, *Earth Planet. Sci. Lett.*, *233*, 337–349.

Mosegaard, K., and A. Tarantola (1995), Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.*, *100*, 12,431–12,447.

Mosegaard, K., and A. Tarantola (2002), Probabilistic approach to inverse problems, in *International Handbook of Earthquake and Engineering Seismology*, edited by W. H.K. Lee, P. Jennings, C. Kisslinger, and H. Kanamori, pp. 237–265, Academic Press, International Association of Seismology and Physics of the Earth Interior (IASPEI).

Nakagawa, T., P. J. Tackley, R. Deschamps, and J. A. D. Connolly (2009), Incorporating self-consistently calculated mineral physics into thermochemical mantle convection simulations in a 3-D spherical shell and its influence on seismic anomalies in Earth's mantle, *Geochem. Geophys. Geosyst.*, *10*, Q03004, doi:10.1029/2008GC002280.

Nettles, M., and A. M. Dziewonski (2008), Radially anisotropic shear velocity structure of the upper mantle globally and beneath North America, *J. Geophys. Res.*, *113*, B02303, doi:10.1029/2006JB004819.

Nolet, G. (2008), *A Breviary of Seismic Tomography: Imaging the Interior of the Earth and Sun*, Cambridge Univ. Press, New York.

O'Reilly, S. Y., W. L. Griffin, and O. Gaul (1997), Paleogeothermal gradients in Australia: Key to 4-D lithospheric mapping, *AGSO J. Aust. Geol. Geophys.*, *17*, 63–72.

Panning, M., and B. Romanowicz (2006), A three-dimensional radially anisotropic model of shear velocity in the whole mantle, *Geophys. J. Int.*, *167*, 361–379, doi:10.1111/j.1365-246X.2006.03100.x.

Pearson, N. J., S. Y. O'Reilly, and W. L. Griffin (1991), Heterogeneity in the thermal state of the lower crust and upper mantle beneath eastern Australia, *Explor. Geophys.*, *22*, 295–298.

Pedersen, H. A., M. Bruneton, V. Maupin, and SVEKALAPKO Seismic Tomography Working Group (2006), Lithospheric and sublithospheric anisotropy beneath the Baltic shield from surface-wave array analysis, *Earth Planet. Sci. Lett.*, *244*, 590–605.

Perry, H. K. C., A. M. Forte, and D. W. S. Eaton (2003), Upper-mantle thermochemical structure beneath North America from seismic-geodynamic flow models, *Geophys. J. Int.*, *154*, 279–299.

Priestley, K. F., and D. P. McKenzie (2006), The thermal structure of the lithosphere from shear wave velocities, *Earth Planet. Sci. Lett.*, *244*, 285–301.

Priestley, K. F., D. P. McKenzie, and E. Debayle (2006), The state of the upper mantle beneath southern Africa, *Tectonophysics*, *416*, 101–112.

Pollack, H. N., S. J. Hurter, and R. J. Johnson (1993), Heat flow from the Earth's interior: Analysis of the global data set, *Rev. Geophys.*, *31*, 267–280.

Rawlinson, N., and S. Fishwick (2012), Seismic structure of the southeast Australian lithosphere from surface and body wave tomography, *Tectonophysics*, *572*, 111–122, doi:10.1016/j.tecto.2011.11.016.

Ricard, Y., L. Fleitout, and C. Froidevaux (1984), Geoid heights and lithospheric stresses for a dynamic Earth, *Ann. Geophys.*, *2*, 267–285.

Ringwood, A. E. (1962), A model for the upper mantle, *J. Geophys. Res.*, *67*, 857–867.

Röhm, A. H. E., R. Snieder, S. Goes, and J. Trampert (2000), Thermal structure of continental upper mantle inferred from S-wave velocity and surface heat flow, *Earth Planet. Sci. Lett.*, *181*, 395–407.

Schaeffer, A. J., and S. Lebedev (2013), Global shear-speed structure of the upper mantle and transition zone, *Geophys. J. Int.*, *194*, 417–449, doi:10.1093/gji/ggt095.

Schmerr, N., and E. Garnero (2007), Upper mantle discontinuity topography from thermal and chemical heterogeneity, *Science*, *318*, 623–626.

Schutt, D. L., and C. E. Lesher (2006), Effects of melt depletion on the density and seismic velocity of garnet and spinel lherzolite, *J. Geophys. Res.*, *111*, B05401, doi:10.1029/2003JB002950.

Schuberth, B. S. A., C. Zaroli, and G. Nolet (2012), Synthetic seismograms for a synthetic Earth: Long-period P- and S-wave traveltime variations can be explained by temperature alone, *Geophys. J. Int.*, *188*, 1393–1412, doi:10.1111/j.1365-246X.2011.05333.x.

Schuberth, B. S. A., H.-P. Bunge, and J. Ritsema (2009), Tomographic filtering of high-resolution mantle circulation models: Can seismic heterogeneity be explained by temperature alone?, *Geochem. Geophys. Geosyst.*, *10*, Q05W03, doi:10.1029/2009GC002401.

Sebai, A., E. Stutzmann, and J. P. Montagner (2006), Anisotropic structure of the African upper mantle structure from Rayleigh and Love wave tomography, *Phys. Earth Planet. Inter.*, *155*, 48–62.

Simons, F. J., A. Zielhuis, and R. D. van der Hilst (1999), The deep structure of the Australian continent from surface-wave tomography, *Lithos*, *48*, 17–43.

Simons, F. J., and R. D. van der Hilst (2002), Age-dependent seismic thickness and mechanical strength of the Australian lithosphere, *Geophys. Res. Lett.*, *29*(11), 1529, doi:10.1029/2002GL014962.

Simmons, N. A., A. M. Forte, and S. P. Grand (2009), Joint seismic, geodynamic and mineral physical constraints on three-dimensional mantle heterogeneity: Implications for the relative importance of thermal versus compositional heterogeneity, *Geophys. J. Int.*, *177*, 1284–1304.

Shapiro, N. M., and M. H. Ritzwoller (2002), Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle, *Geophys. J. Int.*, *151*, 88–105.

Shapiro, N. M., and M. H. Ritzwoller (2004), Thermodynamic constraints on seismic inversions, *Geophys. J. Int.*, *157*, 1175–1188, doi:10.1111/j.1365-246X.2004.02254.x.

Smyth, J. R., and D. J. Frost (2002), The effect of water on the 410-km discontinuity: An experimental study, *Geophys. Res. Lett.*, *29*, 1485, doi:10.1029/2001GL014418.

Sobolev, S. V., H. Zeyen, G. Stoll, F. Werling, R. Altherr, and K. Fuchs (1996), Upper mantle temperatures from teleseismic tomography of French Massif Central including effects of composition, mineral reactions, anharmonicity, anelasticity and partial melt, *Earth Planet. Sci. Lett.*, *139*, 147–163.

Stixrude, L., and C. Lithgow-Bertelloni (2005), Mineralogy and elasticity of the oceanic upper mantle: Origin of the low-velocity zone, *J. Geophys. Res.*, *110*, B03204, doi:10.1029/2004JB002965.

Tapley, B., et al. (2005), GGM02—An improved Earth gravity field model from GRACE, *J. Geod.*, *79*, 467–478.

Tarantola, A., and B. Valette (1982), Inverse problems: Quest for information, *J. Geophys.*, *50*, 159–170.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.

Tauzin, B., E. Debayle, and G. Wittlinger (2008), The mantle transition zone as seen by global *Pds* phases: No clear evidence for a thin transition zone beneath hotspots, *J. Geophys. Res.*, *113*, B08309, doi:10.1029/2007JB005364.

Tauzin, B., E. Debayle, and G. Wittlinger (2010), Seismic evidence for a global low velocity layer in the Earth's upper mantle, *Nat. Geosci.*, *3*, 718–721, doi:10.1038/ngeo969.

Toffelmier, D. A., and J. A. Tyburczy (2007), Electromagnetic detection of a 410-km-deep melt layer in the Southwestern United States, *Nature*, *447*, 991–994, doi:10.1038/nature05922.

Trampert, J., F. Deschamps, J. Resovsky, and D. Yuen (2004), Chemical heterogeneities throughout the lower mantle, *Science*, *306*, 853–856.

van Gerven, L., F. Deschamps, and R. D. van der Hilst (2004), Geophysical evidence for chemical variations in the Australian Continental Mantle, *Geophys. Res. Lett.*, *31*, L17607, doi:10.1029/2004GL020307.

Verhoeven, O., A. Mocquet, P. Vacher, A. Rivoldini, M. Menvielle, P.-A. Arrial, G. Choblet, P. Tarits, V. Dehant, and T. Van Hoolst (2009), Constraints on thermal state and composition of the Earth's lower mantle from electromagnetic impedances and seismic data, *J. Geophys. Res.*, *114*, B03302, doi:10.1029/2008JB005678.

Visser, K., J. Trampert, and B. L. N. Kennett (2008), Global anisotropic phase-velocity maps for higher mode Love and Rayleigh waves, *Geophys. J. Int.*, *172*, 1016–1032.

Voss, P., K. Mosegaard, S. Gregersen, and TOR Working Group (2006), The Tornquist Zone, a north east inclining lithospheric transition at the south western margin of the Baltic Shield: Revealed through a nonlinear teleseismic tomographic inversion, *Tectonophysics*, *416*, 151–166, doi:10.1016/j.tecto.2005.11.016.

Xu, W., C. Lithgow-Bertelloni, L. Stixrude, and J. Ritsema (2008), The effect of bulk composition and temperature on mantle seismic structure, *Earth Planet. Sci. Lett.*, *275*, 70–79, doi:10.1016/j.epsl.2008.08.012.

Yoshizawa, K., and B. L. N. Kennett (2004), Multimode surface-wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects, *J. Geophys. Res.*, *109*, B02310, doi:10.1029/2002JB002254.

Yuan, H., and B. Romanowicz (2010), Lithospheric layering in the North American Craton, *Nature*, *466*, 1063–1068.

Yuan, H., B. Romanowicz, K. Fisher, and D. Abt (2011), 3-D shear wave radially and azimuthally anisotropic velocity model of the North American upper mantle, *Geophys. J. Int.*, *184*, 1237–1260, doi:10.1111/j.1365-246X.2010.04901.x.