Multimode surface wave tomography for the Australian region using a three-stage approach incorporating finite frequency effects

K. Yoshizawa[1] and B. L. N. Kennett

Research School of Earth Sciences, Australian National University, Canberra, Australia

Received 19 October 2002; revised 6 November 2003; accepted 4 December 2003; published 24 February 2004.

[1] A three-stage inversion technique for surface wave tomography is applied to the Australian region. The inversion procedure consists of three independent processes. In the first stage, path-specific one-dimensional (1-D) shear velocity profiles are derived from multimode waveform inversion to provide dispersion information. The information from all paths is then combined to produce multimode phase speed maps as a function of frequency. The first version of these phase speed maps is derived from linearized inversion based on the assumption of surface wave propagation along great circle paths. Subsequently, the 2-D phase speed maps are updated by including ray tracing and finite frequency effects through the influence zone around the surface wave paths over which the phase is coherent. Finally, in the third stage the 3-D shear wave speed distribution is reconstructed from the set of updated multimode phase speed maps. This three-stage inversion of surface waves has significant benefits because it is possible to incorporate multimode dispersion, off-great circle propagation, and finite frequency effects for surface waves in a common framework. The final 3-D model, which includes the effects of ray bending and finite frequency, shows improvement in the definition of the model in regions with high gradients in shear velocity, such as near tectonic boundaries, especially in eastern Australia. Despite the natural smoothing imposed by considering the influence zone around the surface wave paths, the final models still require rapid change in shear wave properties in the neighborhood of the edge of the craton. *INDEX TERMS:* 7218 Seismology: Lithosphere and upper mantle; 7255 Seismology: Surface waves and free oscillations; 8120 Tectonophysics: Dynamics of lithosphere and mantle—general; 8180 Tectonophysics: Tomography; 7260 Seismology: Theory and modeling; *KEYWORDS:* surface waves, tomography, upper mantle

Citation: 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.

1. Introduction

[2] Most existing methods for surface wave tomography are based on multistage processes using either the measurement of fundamental mode dispersion [e.g., *Ekström et al.*, 1997; *Ritzwoller and Levshin*, 1998] or multimode waveform inversion for a path-specific one-dimensional (1-D) model [e.g., *Cara and Lévêque*, 1987; *Nolet*, 1990]. The derivation of 3-D shear wave speeds from global surface wave studies has mostly employed the intermediary of phase speed dispersion maps for the fundamental mode [e.g., *Nataf et al.*, 1986; *Montagner and Tanimoto*, 1990, 1991; *Trampert and Woodhouse*, 1995, 1996; *Laske and Masters*, 1996; *Zhang and Lay*, 1996; *Ekström et al.*, 1997] and, to a lesser extent, higher modes [*Stutzmann and Montagner*, 1993; *van Heijst and Woodhouse*, 1997, 1999].

[3] In contrast, regional surface wave tomography has been dominated by a two-stage approach in which path-specific 1-D models are derived by nonlinear inversion of the waveforms of surface waves [*Nolet et al.*, 1986; *Cara and Lévêque*, 1987], and this path information is employed to construct a 3-D shear wave speed model [e.g., *Zielhuis and Nolet*, 1994; *van der Lee and Nolet*, 1997; *Simons et al.*, 1999; *Debayle and Kennett*, 2000a, 2000b]. In this partitioned waveform approach the 1-D models are interpreted as path averages of the 3-D structure, and a linear inversion is carried out to find a 3-D wave speed model compatible with the various averages.

[4] Improvements in such tomography models have been sought by enlarging the number of paths, so that more detailed structure can be recovered with dense path coverage. *Simons et al.* [1999] and *Debayle and Kennett* [2000a] have used around 2000 paths for the Australian region in inversions using Rayleigh waves. With such path densities it is possible to extend the linearized inversion in the second

[1]Now at Division of Earth and Planetary Sciences, Hokkaido University, Sapporo, Japan.

0148-0227/04/2002JB002254$09.00

stage to try to extract azimuthal anisotropy [*Debayle and Kennett*, 2000a].

[5] The next level of complexity is to include scattering concepts with surface wave tomography. There are only a few studies working with dispersion maps [e.g., *Yomogida and Aki*, 1987; *Alsina et al.*, 1996]. Recently, *Ritzwoller et al.* [2002] used simplified sensitivity kernels, which are equivalent to the Born/Rytov kernels based on the definition of *Spetzler et al.* [2001], to obtain global tomography model using group dispersion of fundamental mode Love and Rayleigh waves. Their sensitivity kernels have a width very close to, but slightly narrower than, the first Fresnel zone.

[6] The levels of heterogeneity and heterogeneity gradient in recent tomographic studies of the upper mantle are probably too large for the path-average approximation to be applied directly to 1-D models for surface waves under the assumption of propagation along the great circle between source and receiver. However, for the frequency range in which modal coupling can be neglected, the path-average assumption can be used for the phase of individual mode contributions. The 1-D model is then to be interpreted as a representation of the character of multimode dispersion along the source-receiver path. Recently, *Yoshizawa and Kennett* [2002a] investigated fully nonlinear inversion for surface wave trains and demonstrated the possibility of extracting different styles of 1-D models with a comparable fit to data. Although these models differ significantly, the dispersion of the first few modes over the relevant frequency range cannot be distinguished.

[7] The existence of large velocity perturbations in recent tomography models (e.g., the short-period phase speed models (<40 s) of *Ekström et al.* [1997] and the upper 200 km of the model of *Debayle and Kennett* [2000a]) also warns us of the need to rethink the great circle approximation for surface wave paths [*Spetzler et al.*, 2001]. In addition, we should take account of finite frequency effects on surface wave propagation, rather than assuming sensitivity just on the ray path. *Yoshizawa and Kennett* [2002b] have studied the approximate zone of influence around surface wave paths with careful investigation of a stationary phase field around a path. By examining the coherence of the phase behavior we have shown that the approximate influence zone can be represented as roughly one third of the width of the first Fresnel zone. The idea of the influence zone leads us to an alternative viewpoint with area-average phase speeds rather than the conventional path average. Such an approach allows us to incorporate the finite frequency effects of wave propagation as well as off-great circle propagation in tomographic inversion. The influence zone is somewhat smaller than that employed by *Ritzwoller et al.* [2002] because attention is concentrated on the coherent portion around the propagation path where the relative time shift is less than one-eighteenth of the period. At the limits of the zone employed by *Ritzwoller et al.* [2002] the time shifts approach half the period, and if such an enlarged zone is to be used, a higher level of approximation may be needed as in the work of *Friederich* [2003].

[8] Utilizing these new concepts for surface wave analysis, *Kennett and Yoshizawa* [2002] reformulated the process of surface wave tomography, especially at the regional scale, into a three-stage process working with multimode dispersion. The stages consist of (1) the extraction of path-specific information by waveform fitting, (2) construction of multimode phase-speed maps as a function of frequency, and (3) a final inversion for local shear wave speed properties.

[9] Such an approach has been exploited in earlier studies [e.g., *Nataf et al.*, 1986] based on observations of fundamental mode surface waves. However, our new scheme, which we call a three-stage inversion, offers the advantage of allowing the incorporation of various styles of information such as multimode dispersion, off-great circle propagation, and finite frequency effects within a single formulation. By working directly with phase speed we can readily incorporate the deviation of paths from the great circle using ray tracing for individual modes and taking account of the extended influence zone around each ray path. This approach can be applied not only to regional studies but also to global studies and therefore will be useful for reconciling surface wave tomography at different scales.

[10] In this paper, we apply the three-stage inversion scheme of *Kennett and Yoshizawa* [2002], for the first time, to the Australian region. The major objective of this paper is to present practical formulations for the three independent stages, that is, the method of extracting multimode dispersion for each path (the first stage), the iterative linearized inversion for multimode phase speed maps incorporating the effects of finite frequency and off-great circle propagation (the second stage), and the construction of the 3-D shear wave speed model combining improved multimode phase speed maps (the third stage). The method is then applied to extract a new Australian upper mantle model to illustrate the utility of the three-stage approach.

[11] In this first application of the three-stage approach we use a set of Rayleigh wave phase speeds for the fundamental and the first three higher modes for particular frequency ranges to obtain an isotropic shear wave speed model. No influence of anisotropy is considered in the present work; this topic is left to future studies on the development of the three-stage method.

2. Data Set

[12] The first step of the three-stage inversion is to measure multimode dispersion from observations. At this stage, we can employ any convenient method for estimating surface wave dispersion [*Kennett and Yoshizawa*, 2002]. Here we use 2000 path-specific 1-D shear wave speed profiles from *Debayle and Kennett* [2003], which have been derived from waveform inversion for the frequency range between 50 and 160 s using secondary observables based on the cross correlograms calculated from observed and synthetic seismograms [*Cara and Lévêque*, 1987], an approach similar to the isolation filter technique of *Gee and Jordan* [1992]. This data set employs paths solely within the Australian Plate to avoid possible complex effects from major structural boundaries in the Philippine Sea region and so provides a good basis from which to investigate the upper mantle structure beneath the Australian continent.

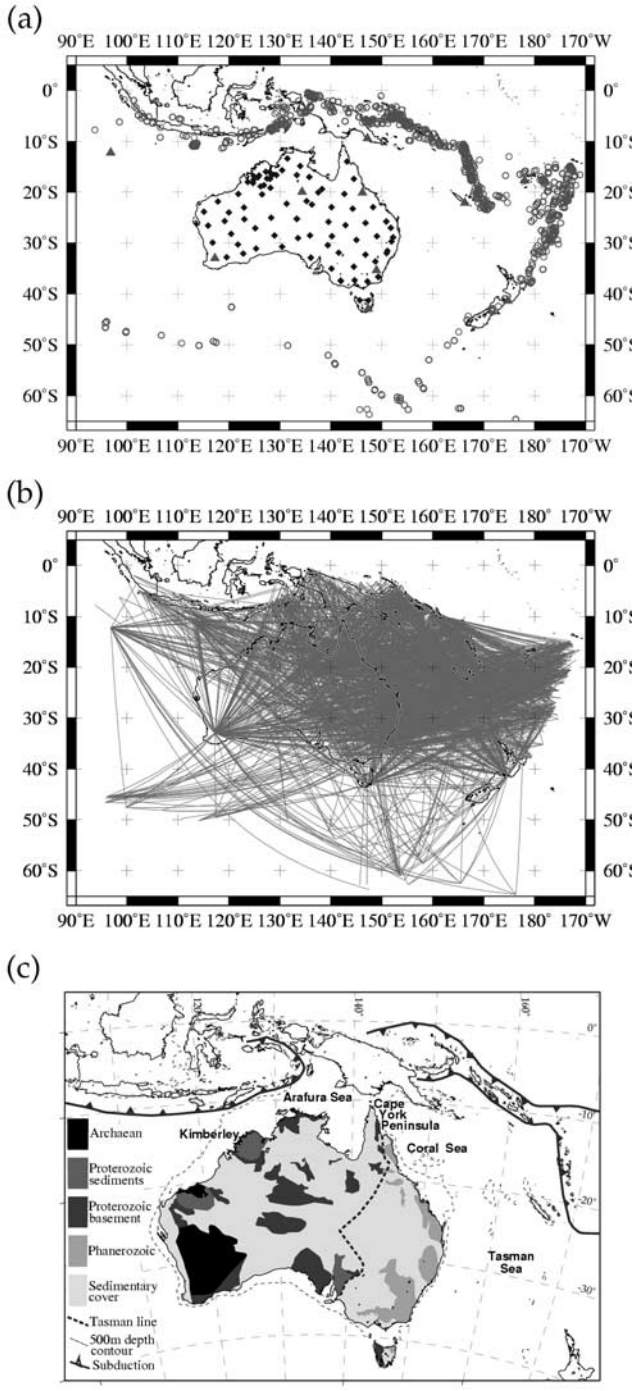

Figure 1. (a) Distribution of global permanent stations (triangles), portable stations (diamonds), and events (circles), (b) 2000 Rayleigh wave paths of *Debayle and Kennett* [2003], and (c) major geological blocks at the surface of the Australian region.

[13] The 1-D models are derived from the vertical component of Rayleigh waves recorded at the IRIS and GEOSCOPE stations as well as at portable broadband seismic stations of the SKIPPY and KIMBA experiments undertaken by the seismology group at the Australian National University from 1993 to 1998 (Figure 1a). We have corrected for the crustal structure encountered along each path using the 3SMAC model [*Nataf and Ricard*, 1996] to improve the calculation of phase speeds at shorter period. We have estimated phase speeds for the fundamental mode and three higher modes from the path-specific 1-D models for the period ranges shown in Table 1. With the crustal corrections we can estimate multimode phase speeds to periods of 40 s. The longest period used for the phase speed measurements depends on the particular Rayleigh mode (Table 1). Phase speed models are obtained over these period ranges with an increment of 10 s period.

[14] It should be noted that in the three-stage approach we assume that each surface wave mode propagates independently, without any effects of coupling between mode branches during the propagation. For periods much shorter than used in this study we would not be able to ignore the effects of mode conversion caused by strong heterogeneity in the crust and uppermost mantle [*Kennett and Nolet*, 1990], and thus the assumption of the independent mode propagation would be violated.

[15] The waveform inversion procedure used by *Debayle and Kennett* [2000a] used the *Cara and Lévêque* [1987] approach for the first four modes. We have retained this four-mode treatment since we can then be confident of independent mode propagation. We can achieve good resolution of the continental lithosphere beneath Australia since most of the surface wave energy in the frequency band that we have used is confined to the first four modes. Some implementations of partitioned waveform inversion [e.g., *Lebedev*, 2000] utilize much higher modes up to 20–30, which in principal enables resolution of structure in the transition zone, but mode coupling induced by heterogeneity gradients is likely to be a problem for the highest modes.

[16] Some examples of the 1-D models and estimated phase dispersion curves of the fundamental and the first three modes for different paths to the NWAO station in the southwest Australia are shown in Figure 2. These 1-D models are smoothed over the 400- and 670-km discontinuities. A path passing through the Indian Ocean (Figure 2a) shows a clear slower shear wave speed anomaly around 150 km depth, whereas a continental path (Figure 2b) passing mainly through the Proterozoic and Archaean blocks in the central and western Australia (Figure 1c) shows noticeably higher wave speed anomalies in the top 200 km. Another example in Figure 2c shows a path which traverses both oceanic and continental regions. The corresponding path-average 1-D model shows the average features of the oceanic and continental structures with no remarkable anomalies in the upper mantle.

[17] The use of the sets of 1-D models of *Debayle and Kennett* [2003] enables us to assess the model from the three-stage approach compared with that derived previously from the two-stage approach. The number of higher modes, which can be reliably extracted from the observations,

Table 1. Minimum and Maximum Period Range[a]

	Mode Branch			
	0	1	2	3
Minimum period, s	40	40	40	40
Maximum period, s	150	140	100	60

[a]Rayleigh wave phase speeds are estimated from the 1-D shear wave speed profiles of *Debayle and Kennett* [2003].

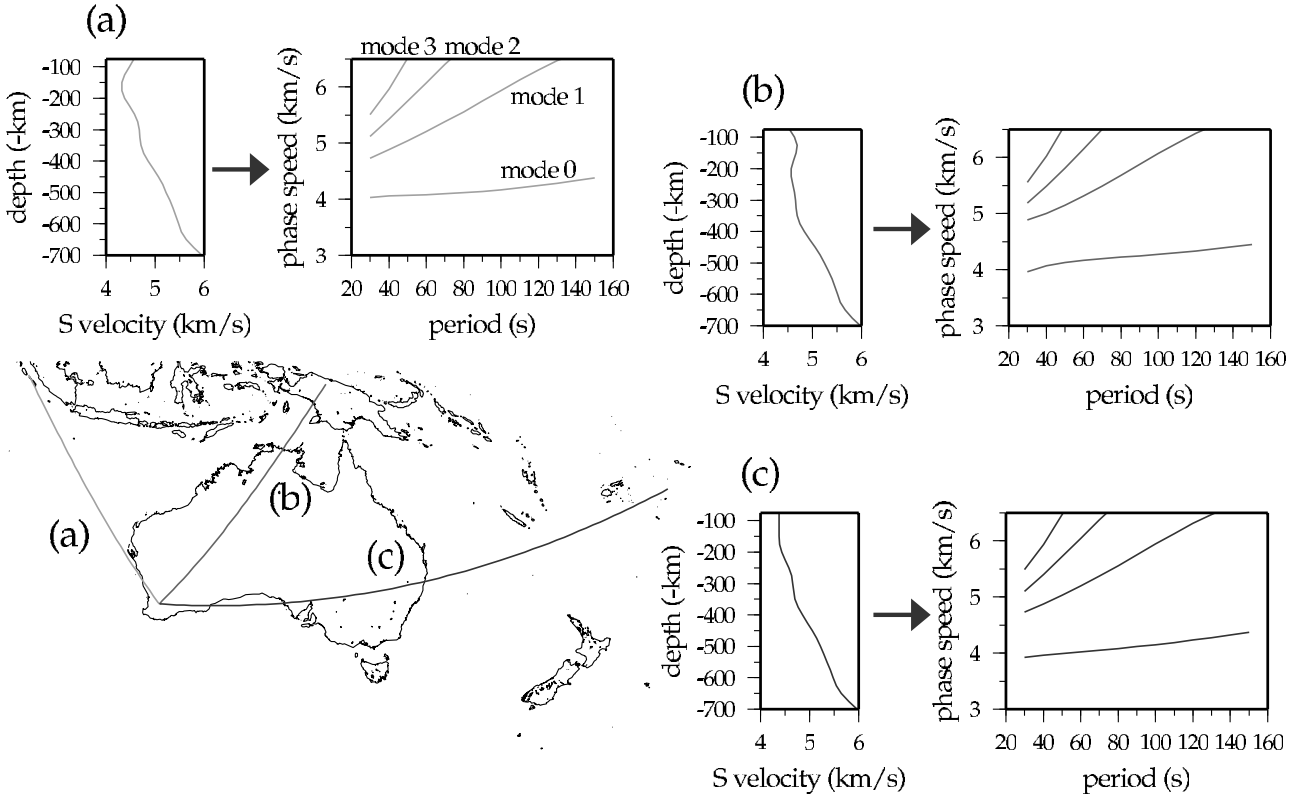

Figure 2. Examples of the path-average 1-D profiles of *Debayle and Kennett* [2003] and multimode phase dispersion curves estimated from these 1-D models. The three paths are chosen so that they sample (a) mainly the oceanic region, (b) mainly the continental region, and (c) both the oceanic and continental regions.

depends strongly on the excitation of the modes at the source. Some of the 1-D models are primarily constrained by the fundamental mode for shallow events. For testing the development of our three-stage method we have used the 1-D models to generate higher-mode dispersion, even where the model is almost entirely constrained by the fundamental mode. Thus for very shallow sources, the higher-mode dispersion will not represent independent information.

[18] The reliability of the measured phase speeds from such 1-D models can be taken into account by using a posteriori errors [e.g., *Nataf et al.*, 1986] in the 1-D shear wave speed profiles, which represent how well the model is constrained via waveform fitting [*Debayle and Kennett*, 2000a]. Thus for 1-D models that have been constrained only by the fundamental mode the estimated errors in shear wave speed become large at depth where there is little sensitivity. We use this information to estimate the errors in the estimated phase speeds, which are subsequently used as weights on the phase speed data in inversions for phase speed maps, as a function of frequency, in the second stage.

3. Inversion for Multimode Phase Speed Maps

3.1. Formulation of Inversion

[19] In the second step of the three-stage inversion the ensembles of path-average phase speeds for each mode are inverted to produce mode-dependent phase speed maps as a function of frequency. The linear relationship between perturbation of phase of seismograms, $\delta\psi$, and phase speeds for the jth mode, δc^j, can be represented as [e.g., *Woodhouse and Wong*, 1986]

$$\delta\psi_j(\omega) \simeq -k_j(\omega) \int_{\mathrm{ray}_j} ds \frac{\delta c^j(s,\omega)}{c_0^j(\omega)}, \tag{1}$$

where k_j is the wave number ($= \omega/c_0^j$) for a reference model. Hereafter we omit the dependency on frequency ω and mode number j. The contribution to the phase of the observed seismogram for each mode is thus represented by the average phase speed perturbation along either the great circle path or the ray path to the station with propagation distance Δ:

$$\delta\psi^{\mathrm{obs}} \simeq -k \frac{\langle \delta c \rangle^{\mathrm{obs}}}{c_0} \Delta, \qquad \frac{\langle \delta c \rangle^{\mathrm{obs}}}{c_0} = \frac{\langle c \rangle^{\mathrm{obs}} - c_0}{c_0}, \tag{2}$$

where $\langle c \rangle^{\mathrm{obs}}$ is the average phase speed along the path.

3.1.1. Linear Relation for Geometrical Rays

[20] The path-average phase speeds can be represented through the linear relationship derived from equations (1) and (2):

$$\frac{\langle \delta c \rangle^{\mathrm{obs}}}{c_0} = \frac{1}{\Delta} \int_{\mathrm{path}} ds \frac{\delta c(s)}{c_0}, \tag{3}$$

where the integration is taken along either the great circle path or the actual ray path between the source and receiver. Using relationship (3), frequency- and mode-dependent phase speed maps are obtained on the basis of geometrical ray theory.

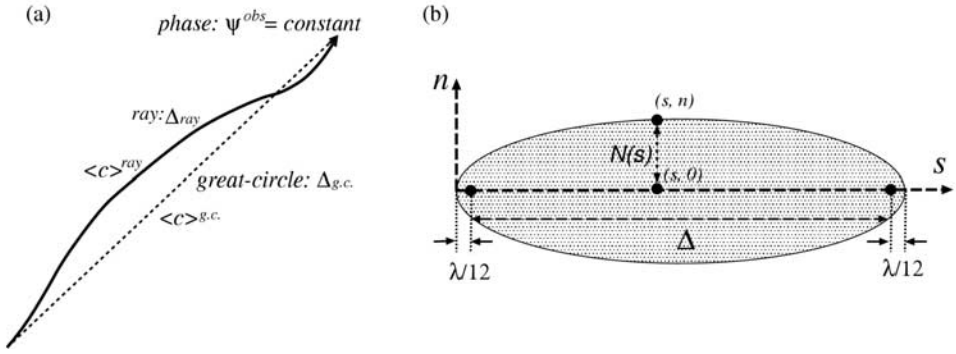

Figure 3. (a) Illustration of a great circle with epicentral distance Δ_{gc} and the actual ray path with length Δ_{ray}. Since the phase of an observed seismogram is constant, the average phase speeds $\langle c \rangle^{gc}$ measured along the great circle and $\langle c \rangle^{ray}$ along the ray have slight differences associated with the differences in the distance traveled. (b) Schematic illustration of the influence zone in a ray-centered coordinate system (s, n). After *Yoshizawa and Kennett* [2002b].

[21] The observed phase speed $\langle c \rangle^{obs}$ is usually measured along the great circle path. However, the actual ray path will be longer than the epicentral distance Δ_{gc} along the geodesic great circle. When we take account of off-great circle propagation in constructing phase speed, we need to make a correction to the observed phase speed so that it represents an average along the actual ray path. Since the phase contribution to a seismogram does not depend on the ray path (Figure 3a), the average phase speeds $\langle c \rangle^{ray}$ along the ray path with length, $\Delta_{ray} = \int_{ray} ds$, can thus be represented by using $\langle c \rangle^{gc}$ along the great circle: $\langle c \rangle^{ray} = \langle c \rangle^{gc} \Delta_{ray}/\Delta_{gc}$. This corrected average phase speed is used when we invert for phase speed maps employing ray tracing in a heterogeneous reference model.

3.1.2. Linear Relation for Finite Width Paths

[22] Phase speed maps can be improved not only via ray tracing but also by including the influence zone around the surface wave paths due to finite frequency effects. With careful investigation of a stationary-phase field, *Yoshizawa and Kennett* [2002b] have identified that the influence zone of surface waves is nearly one third of the width of the first Fresnel zone. Since the influence zone has been defined as the finite area over which surface waves are coherent in phase, we can regard the observed phase speeds as an average within the influence zone rather than just as an average along the path.

[23] Using the ray centered coordinate system (s, n) used by *Yoshizawa and Kennett* [2002b], we first consider the average phase speed variation perpendicular to the ray at a particular point s on the path:

$$\frac{\langle \delta c(s) \rangle}{c_0} = \frac{1}{2\mathcal{N}(s)} \int_{width} dn \frac{\delta c(s,n)}{c_0}, \tag{4}$$

where $\mathcal{N}(s) = \int_0^{\mathcal{N}}(s) dn$ is the half width of the influence zone at a point s on the ray path (Figure 3b). Equation (4) can then be integrated along a path to give the average phase speed variations within the influence zone:

$$\frac{\langle \delta c \rangle}{c_0} = \frac{1}{\Delta} \int_{path} ds \frac{1}{2\mathcal{N}(s)} \int_{width} dn \frac{\delta c(s,n)}{c_0}. \tag{5}$$

The path integration can be calculated along either the great circle path or the actual ray path. If we take into account the influence zone around the source and receiver locations as discussed by *Yoshizawa and Kennett* [2002b, appendix], the integration along the path should be undertaken between the two edges of the influence zone on the ray trajectory (Figure 3b), so that the total zone should be slightly longer than the ray path length ($\sim\lambda/6$, where λ is the wavelength).

[24] When we integrate the phase speed perturbation across the zone in equation (4), we apply a weight function to smooth the edges of the influence zone where the assumption of the phase coherency tends to be violated [*Yoshizawa and Kennett*, 2002b]. Equation (5) is then modified to

$$\frac{\langle \delta c \rangle}{c_0} = \frac{1}{\Delta} \int_{path} ds \frac{1}{2\mathcal{N}_w(s)} \int_{width} dn \mathcal{W}(s,n) \frac{\delta c(s,n)}{c_0}, \tag{6}$$

where $\mathcal{N}_w(s) = \int dn W(s, n)$. Since the variations in the phase contributions within the influence zone are very small, we adopt a cosine taper as the weight function, rather than a Gaussian:

$$\mathcal{W}(s,n) = \cos\left[\frac{\pi}{2}\left(\frac{n}{\mathcal{N}(s)}\right)^2\right]. \tag{7}$$

[25] The process of double integration in equation (5) or (6) allows us to provide an expression of the two-dimensional distribution of sensitivity to surface wave phase:

$$\delta\psi = \int_{path} ds \int_{width} dn \, \mathcal{K}_\psi(s,n) \, \frac{\delta c(s,n)}{c_0}, \tag{8}$$

where $\mathcal{K}_\psi$ is the 2-D sensitivity kernel for phase speed structures. Examples of the weighted sensitivity kernels are displayed in Figure 4. The influence zone becomes wider at longer periods because the absolute phase speed is faster. It is apparent that the weighted surface wave sensitivity to phase speed structure varies along the path but is nearly constant over the width of the influence zone. The highest sensitivity is concentrated near the source and receiver, as expected from the sensitivity kernels evaluated with first-order scattering theory [e.g., *Yomogida*, 1992; *Marquering et al.*, 1998, 1999; *Dahlen et al.*, 2000; *Hung et al.*, 2000; *Zhao et al.*, 2000]. *Vasco et al.* [1995] have shown that the use of such kernels provides very similar results to the use of more rigorous sensitivity kernels derived from wave theory.

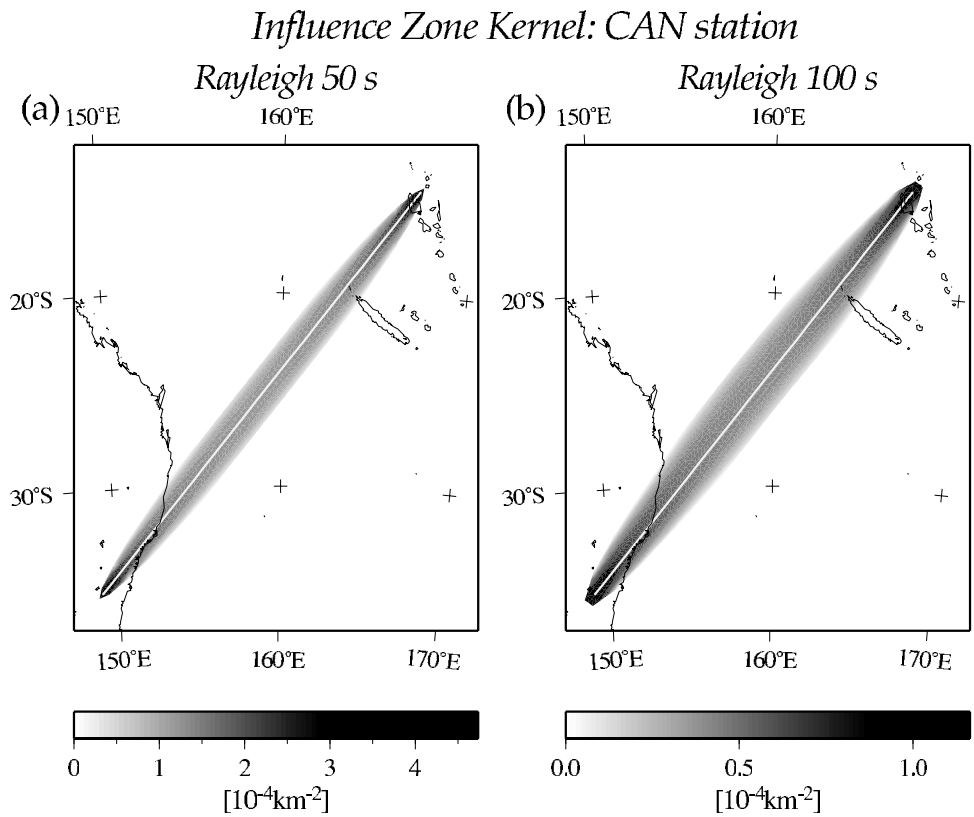

Figure 4. Representation of the weighted sensitivity kernel $\mathcal{K}_\psi$ for a path between an event in Fiji and the CAN station in southeastern Australia. The influence zones of the fundamental mode Rayleigh waves at (a) 50 and (b) 100 s are displayed. The PREM model is used as the reference model to calculate all the kernels. The white line is the great circle.

[26] To take account of the wider effects of scattering outside the influence zone, we can envisage the use of the Born approximation with more rigorous calculations of the sensitivity kernels, as have been used in previous studies [e.g., *Yomogida and Aki*, 1987; *Meier et al.*, 1997; *Dahlen et al.*, 2000]. Tackling these more complicated problems is beyond the present scope; such forms of sensitivity kernels will be discussed in a future publication.

3.2. Least Squares Inversion for 2-D Phase Speed Maps

[27] The linearized equations (3) and (6) can be written in the generalized form, $\mathbf{d} = \mathbf{Gm}$, where the data vector $\mathbf{d}$ consists of the observed phase speed variations $\langle \delta c/c \rangle_i (i = 1, 2, \ldots, M)$ and M is the total number of paths (about 2000); $\mathbf{m}$ is a vector of model parameters m_j ($j = 1, 2, \ldots, N$), and $\mathbf{G}$ is the kernel matrix.

[28] We use a spherical B spline function $\mathcal{F}$ (θ, ϕ) [e.g., *Lancaster and Salkauskas*, 1986; *Wang and Dahlen*, 1995] defined at the center of a geographic cell as a basis function to expand the phase speed perturbation:

$$\frac{\delta c(\theta, \phi)}{c_0} = \sum_{j=1}^{N} m_j \mathcal{F}_j(\theta, \phi), \tag{9}$$

where the model parameter m_j is the coefficient of the jth basis function $\mathcal{F}_j$. The model space is parameterized using the B spline basis functions with an interval of 2°, which provides Gaussian type smoothing around the center of cells [*Yoshizawa*, 2002]. The total number of model parameters N is about 1200.

[29] Using the spherical B splines, the components of the kernel matrix $\mathbf{G}$ for the ith path and the jth model parameter can be written as

Average along the path

$$G_{ij} = \frac{1}{\Delta_i} \int_0^{\Delta_i} ds \mathcal{F}_j \tag{10a}$$

Average within the influence zone

$$G_{ij} = \frac{1}{\Delta_i} \int_0^{\Delta_i} ds \frac{1}{2\mathcal{N}_w(s)} \int_{\text{width}} dn \mathcal{W}(s, n) \mathcal{F}_j, \tag{10b}$$

where the epicentral distance Δ_i is measured along either the great circle or the ray path.

[30] The linearized inversion equation is solved with a damped least squares scheme, minimizing the objective function

$$\Phi(\mathbf{m}) = (\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{Gm}) + \lambda^2 \mathbf{m}^T \mathbf{m}, \tag{11}$$

where $\mathbf{C}_d^{-1}$ is the inverse data covariance matrix controlling the relative contribution of the individual data misfits and λ is an arbitrary damping parameter that controls the trade-off between the model variance and resolution, which subse-

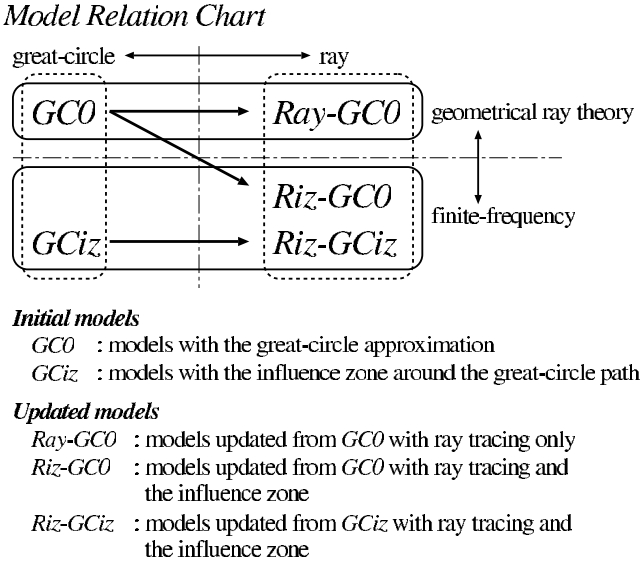

Figure 5. Chart representing the relations between the five different sets of phase speed models. Model sets GC0 and GCiz are derived from the great circle approximation, and these models are used as initial heterogeneous models in the subsequent inversions for the updated models (Ray-GC0, Riz-GC0, and Riz-GCiz).

quently affects the maximum amplitude and smoothness of the model.

[31] Assuming that measured phase speeds for different paths are uncorrelated (although there may be some correlation through the source location) and their variances are different for each path, the covariance matrix can be represented as $\mathbf{C}_d = \sigma_{d_i}^2\mathbf{I}$, where σ_{d_i} is the measurement error for the ith datum and $\mathbf{I}$ is the identity matrix. It is convenient to introduce a rescaled data vector $d_i' = d_i/\sigma_{d_i}$, and kernel matrix $G_{ij}' = G_{ij}/\sigma_{d_i}$ and thereby to enhance the importance of data with relatively small measurement error. Thus we can rewrite the objective function as $\Phi(\mathbf{m}) = |\mathbf{d}' - \mathbf{G}'\mathbf{m}|^2 + \lambda^2|\mathbf{m}|^2$. The inverse problem is then to solve the linear equation system

$$\begin{bmatrix} \mathbf{G}' \\ \lambda\mathbf{I} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d}' \\ \mathbf{0} \end{bmatrix}. \tag{12}$$

for which we have used the iterative LSQR algorithm [*Paige and Saunders*, 1982].

4. Multimode Phase Speed Models

4.1. Five Sets of Phase Speed Models

[32] We can obtain different types of multimode phase speed models by using various types of inversion kernels (equation (10)) working with the great circle approximation, with ray tracing or with allowance for the influence zone around the assumed path. In order to distinguish the sets of phase speed models derived from the different assumptions, we use a naming scheme based on the inversion method; this convention and the relation of the model sets are summarized in Figure 5. The PREM model [*Dziewonski and Anderson*, 1981] is used as a reference for all the tomography models shown in this paper.

[33] The first set of models GC0 uses the great circle approximation without any consideration of the zone of influence; so that the procedure is similar to conventional inversion for phase speed maps. A second set of models GCiz is derived incorporating the influence zone around the great circle paths.

[34] The phase speed models GC0 can be updated by allowing for the deviations of the surface wave rays from the great circle in the construction of the phase speed maps (models Ray-GC0). We also use GC0 to obtain an improved model set Riz-GC0 considering both off-great circle propagation and the influence zone for each path.

[35] Further updates to the model set GCiz incorporating ray tracing and the influence zone produce the final model set Riz-GCiz. This set would be expected to be the best models in this study, in that we take full account of the possible effects of off-great circle propagation and finite frequency effects.

[36] We refer to the various updating processes for the sets of models as global iteration to distinguish them from the other class of iterative processes for each phase speed inversion using the LSQR algorithm (local iteration). Although the new sets of phase speed models (i.e., Ray-GC0, Riz-GC0, and Riz-GCiz) could be further updated by global iteration, there is no need to repeat this process more than once because in most cases, the second global iteration does not alter (or improve) the models significantly.

[37] The most time-consuming process in the second stage is the computation of the inversion kernels (equation (10)). In particular, when we incorporate ray tracing for all the paths in phase speed maps, the computation time is almost doubled compared to the great circle approximation. When we include the influence zone, the total amount of computation for the spatial integrations depends on the mode and frequency because the width of the influence zone to be integrated varies significantly with these factors. In the examples in this paper, the computation for the models GCiz is more than 5 times than that for GC0. The construction of Riz-GC0 and Riz-GCiz requires nearly 10 times the computation time of GC0.

4.2. Assessment of Models

4.2.1. Trade-Off

[38] Figure 6a shows the behavior of the misfit and model norms as a function of the local iterations of the LSQR algorithm, with varying damping parameters (λ = 0.4, 1.0, and 1.6) for the phase speed model GC0 for the fundamental mode at 100 s. The number of local iterations that give adequate convergence depends on the choice of damping parameter. For a relatively large damping, λ = 1.6, both the model and misfit norms converge by the tenth iteration with a suppressed model norm and slightly higher misfit. With a smaller damping of λ = 0.4 the model norm grows rapidly, and models do not converge within 10 iterations.

[39] With damping λ = 1.0, the model reaches a satisfactory level of convergence within 10 iterations with a compromise in the trade-off between the misfit and the model norm. The trade-off curve is displayed in Figure 6b, with the misfit represented by $(|\mathbf{d} - \mathbf{Gm}|^2/|\mathbf{d}|^2) \times 100(\%)$.

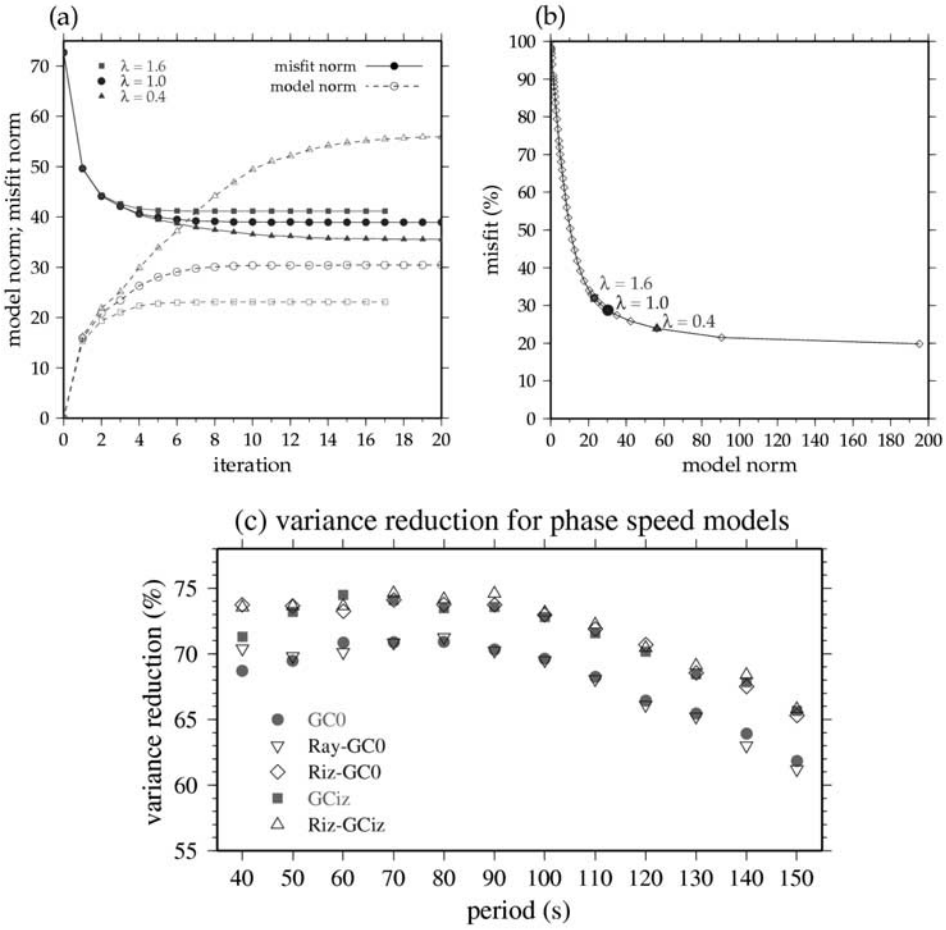

Figure 6. (a) Model and misfit norms as a function of local iterations of the LSQR algorithm for a phase speed model GC0 at 100 s. (b) Trade-off curves for varying damping parameter λ. We choose λ = 1.0 as a preferred damping for this example. (c) Variance reductions estimated by using synthetic data calculated from the linearized forward modeling employing equation (6) for five sets of fundamental mode phase speed models.

In this example, we chose λ = 1.0 as the appropriate damping for this model.

[40] The other classes of phase speed models (Figure 5) for different frequencies and different modes show a similar behavior to Figure 6b with slight differences in the values of the damping parameter. As suitable values of the damping parameter λ with a good compromise in the trade-offs for ray theoretical models (i.e., GC0 and Ray-GC0), we use from 0.8 to 1.2 for the fundamental modes and from 0.6 to 0.8 for the higher modes, depending on frequency. For finite frequency models (i.e., GCiz, Riz-GC0, and Riz-GCiz), we can choose somewhat smaller damping than for ray theoretical models to achieve the comparable levels of misfit.

4.2.2. Variance Reduction

[41] We estimate the variance reductions for the five types of phase speed maps using $(1 - \sigma_1^2/\sigma_0^2) \times 100(\%)$, where σ_0 is a data variance for a spherical Earth model (i.e., PREM) and σ_1^2 is a variance for an obtained phase speed map. We estimate σ_1^2 by using synthetic data calculated from the linearized forward modeling employing equation (6) considering the effects of both the ray path bending and the influence zone.

[42] The patterns of variance reduction in the phase speed maps for the fundamental mode are displayed in Figure 6c. All the models achieve quite good variance reductions of around 70% for the fundamental mode and over 80% for higher modes (not shown). There is a clear tendency in Figure 6c that the phase speed models including the effects of finite frequency (GCiz, Riz-GC0, and Riz-GCiz) consistently provide better variance reductions than the ray theoretical models (GC0 and Ray-GC0). This behavior indicates that the theory with the influence zone is superior to conventional geometrical ray theory. We can also see that the inclusion of ray path bending does not bring significant improvements in the phase speed models for most of the current period range. An exception can be seen in the phase speed models at 40 s for which the off-great circle propagation becomes conspicuous compared to the longer periods, resulting in somewhat higher variance reduction for updated models including ray path bending.

4.2.3. Resolution

[43] The most widely used procedure to assess the resolution of tomographic models is the use of a checkerboard resolution test. However, it is not trivial to visualize the realistic resolution of tomography models taking account of

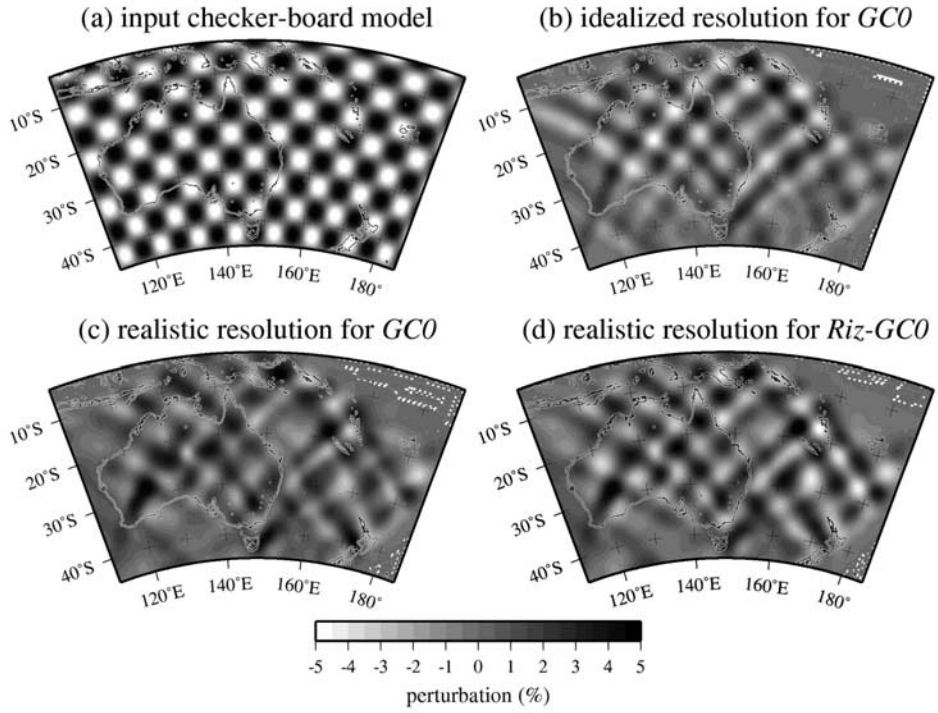

Figure 7. (a) Input checkerboard models with 5° cells. (b) A retrieved model based on the geometrical ray theory employing synthetic data calculated from the ray theory for fundamental mode Rayleigh wave at 40 s, representing ''idealized'' resolution for a ray theoretical model (GC0). Models retrieved by using (c) the geometrical ray theory (corresponding to GC0) and (d) the influence zone (corresponding to Riz-GC0) employing synthetic data which include the effects of finite frequency and off-great circle propagation, representing ''realistic'' resolution of phase speed models.

the nonlinear effects in the observed data. If we just work with linearized ray theory for both forward and inverse problems, the test only provides us with an ''idealized'' apparent resolution which just reflects the ray path coverage as modulated by the employed damping and smoothing in the inversion.

[44] Such an example is shown in Figure 7b, where we calculated synthetic data (fundamental mode Rayleigh waves at 40 s) for an input checkerboard model with 5° cells (Figure 7a) employing equation (3) based on the geometrical ray theory and then performed inversion with the same linear equation. For regions with sufficient path coverage we can achieve a fairly good recovery of the original checkerboard pattern and amplitude with the damping used for real data inversion. We refer to this type of resolution as an idealized resolution because everything is linear for both forward and inverse problems.

[45] To provide a more realistic estimate of the attainable resolution, we perform a further class of checkerboard tests using synthetic data calculated from equation (6). These synthetics include the effects of ray path bending as well as the finite frequency of the surface waves. The synthetic data are still based on linear approximation but should be closer to observed data that reflect the off-path sensitivity of surface waves propagating in a 3-D structure.

[46] We perform inversions with these synthetic data for the five types of models (Figure 5). Examples of ''realistic'' resolutions for phase speed models GC0 and Riz-GC0 are shown in Figures 7c and 7d, respectively. It is apparent that the resolution for GC0 with the new data set is somewhat suppressed compared to that for Riz-GC0. This is because that the synthetic data employed here include the effects of finite frequency which cannot be fully accounted for in a framework of geometrical ray theory. Therefore Figure 7c represents a realistic assessment of the attainable resolution for ray theoretical models. From the comparison of Figures 7b and 7c it is clear that the limited approximation of geometrical ray theory tends to overestimate the apparent resolution of the derived models and such results therefore need to be treated with caution.

[47] Figure 7d includes an allowance for the influence zone and is updated from Figure 7c in a similar way to real data inversions. The retrieved model (Figure 7d) provides a better recovery of the original checkerboard pattern and amplitude, even though the reference velocity model in Figure 7c contains quite erroneous features. Also, we can see a uniform spatial resolution of the finite frequency model in Figure 7d, since we can achieve more uniform path coverage with the inclusion of spatial integration about a path.

[48] In Figure 8 we display some more examples of realistic checkerboard models for the fundamental mode with different cell sizes: 12° (Figures 8a–8c) and 8° (Figures 8d–8f) at 80 s period and 4° (Figures 8g–8i) at 40 s period. The resolution patterns for GC0 (Figures 8b, 8e, and 8h) show some erroneous patch-like features and somewhat reduced amplitudes. For Riz-GC0 (Figures 8c, 8f, and 8i), which is updated from GC0, such artificial

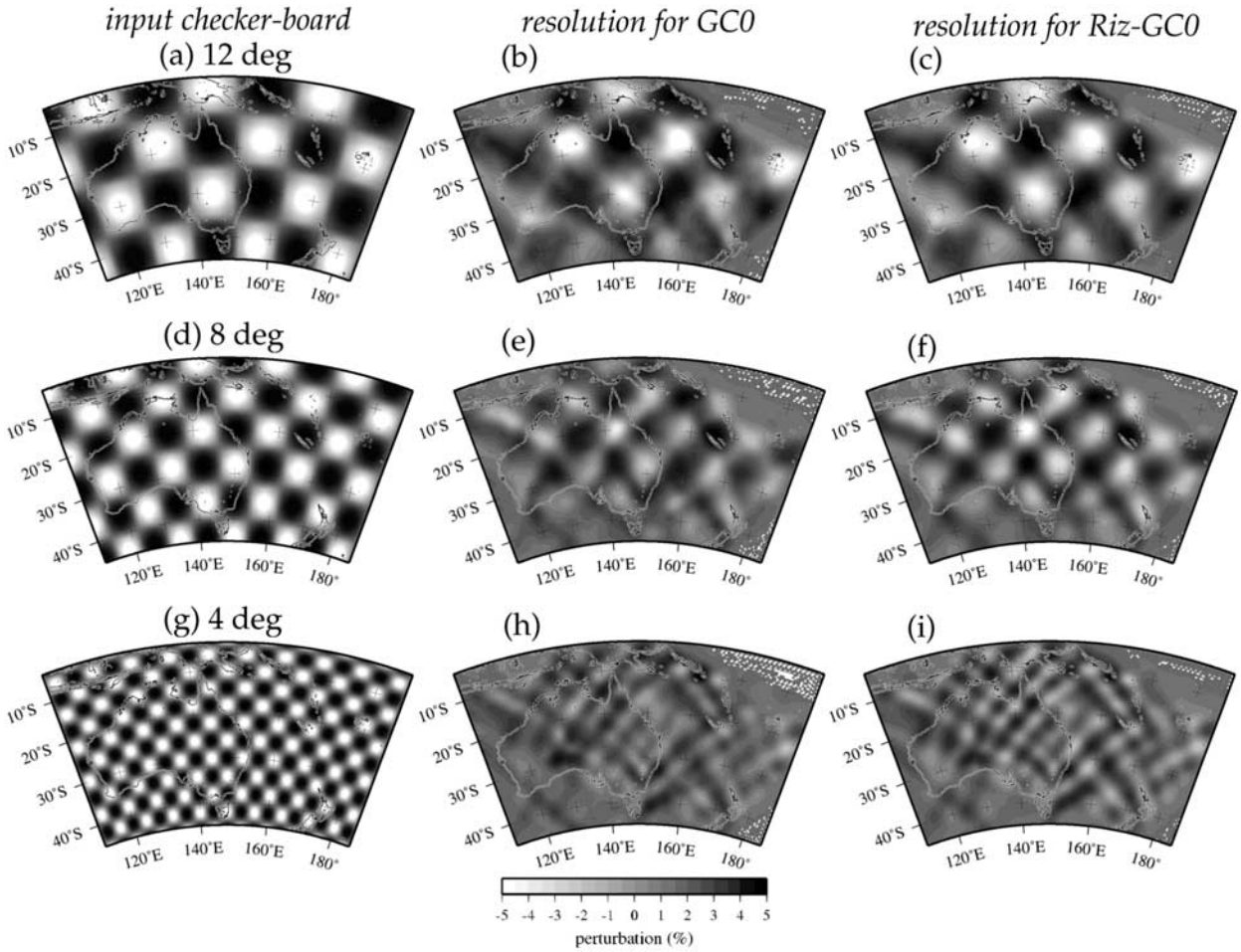

Figure 8. Realistic resolution tests for the fundamental mode with different cell sizes: (a–c) 12° and (d–f) 8° at 80 s period and (g–i) 4° at 40 s period. Input models are displayed in Figures 8a, 8d, and 8g, the retrieved checkerboard is displayed for ray theoretical models (GC0) in Figures 8b, 8e, and 8h and for the finite frequency models (Riz-GC0) in Figures 8c, 8f, and 8i.

structures are smoothed out, and the original patterns and amplitude of heterogeneity can be recovered quite well. The natural smoothing introduced by the inclusion of off-path sensitivity is also seen in the group dispersion maps of *Ritzwoller et al.* [2002], who employed much wider sensitivity kernels based on the Born/Rytov approximation [*Spetzler et al.*, 2002].

[49] With the current data set we can achieve fairly good recovery of both patterns and amplitudes of input checkerboard down to 5° cells. The patterns of smaller-scale heterogeneities (<4°) can also be recovered quite well for regions with good path coverage, but the amplitude is slightly suppressed. This is mainly due to somewhat larger damping applied to observed data for shorter-period Rayleigh waves. In the present study, the minimum scale length of the lateral heterogeneity that we can resolve is a few hundred kilometers for fundamental mode phase speed maps at 40 s.

4.3. Multimode Phase Speed Maps for Rayleigh Waves

[50] The five types of phase speed maps for the fundamental mode Rayleigh wave at 100 s are displayed in Figure 9. All the models show a fair degree of consistency at large scale. The dominant features are fast velocity anomalies in the Archaean and Proterozoic blocks in central and western Australia and slow velocity anomalies in the eastern Phanerozoic region of the eastern Australia and in the Coral and Tasman Sea (Figure 1c).

[51] The models with the inclusion of finite frequency effects represent a better approximation to the forward modeling for the surface waves as shown in Figure 6c. The spatial resolution is more uniform with the inclusion of the spatial integration about the paths. A further factor is enhanced sensitivity to the source and receiver points (in common to all improved modeling procedures). Checkerboard tests including nonlinear effects in Figures 7 and 8 suggest that finite frequency models would provide better realistic resolution than ray theoretical models.

[52] The influence of the different styles of inversions change the character of the phase speed distributions in Figure 9, so that some regions with strong local heterogeneity in the ray theoretical models (GC0 and Ray-GC0) are smoothed out in the finite frequency models (GCiz, Riz-GC0, and Riz-GCiz). This effect is most noticable in the Tasman Sea and in the Proterozoic blocks of central Australia. Also, we can see more enhanced image of subduction in the New Hebrides and Tonga-Kermadec trenches. The

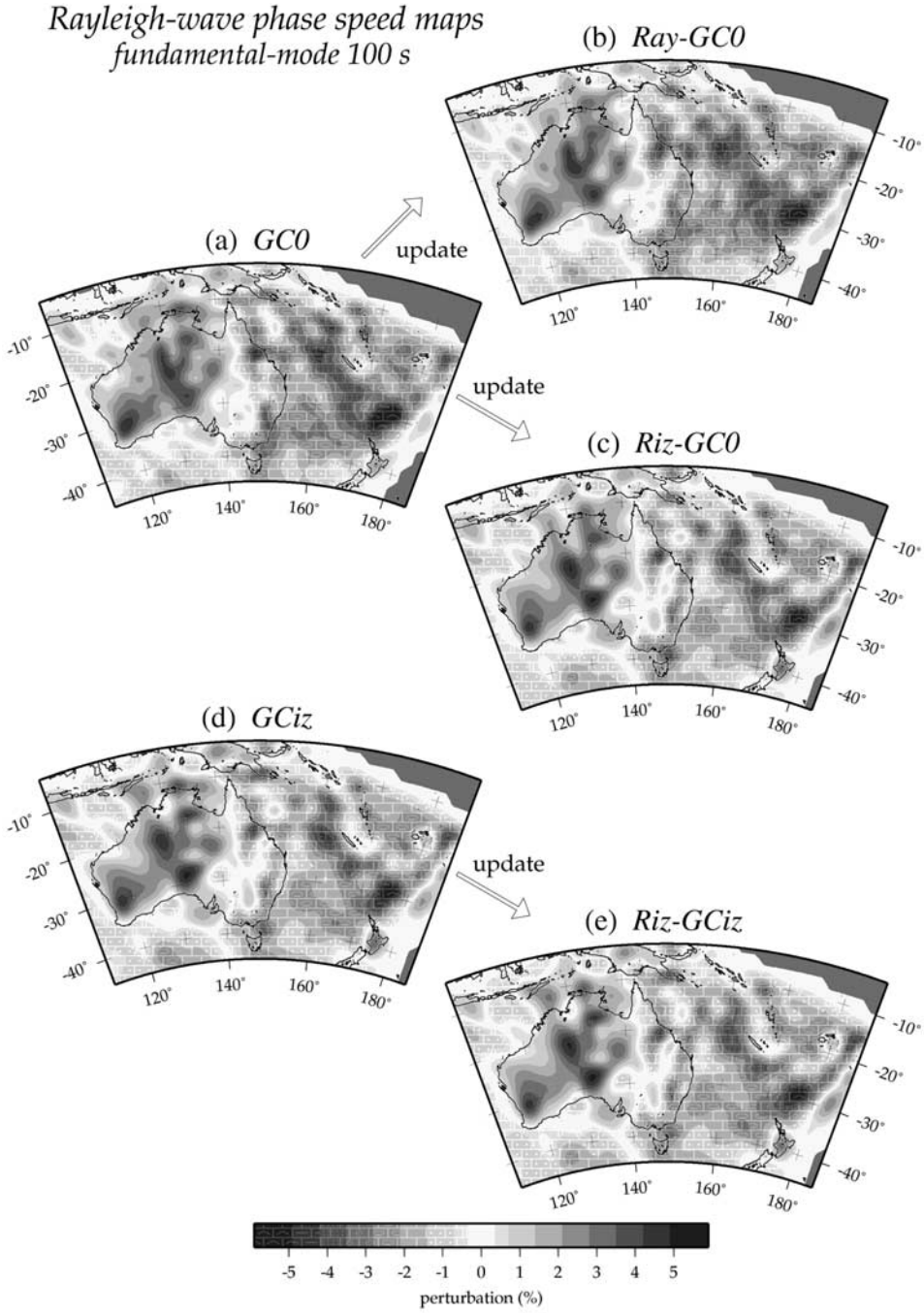

Figure 9. Five types of phase speed maps for the fundamental mode Rayleigh waves at 100 s. All the models are shown as a perturbation from PREM. (a) Model GC0 derived from the great circle approximation. (b) Updated model Ray-GC0 with surface wave ray tracing in the model in Figure 9a. (c) Updated model Riz-GC0 with ray tracing as well as the influence zone about the path. (d) Model GCiz including the influence zone around the great circle paths. (e) Updated model Riz-GCiz from Figure 9d using ray tracing and the influence zone.

updated models including finite frequency effects, Riz-GC0 and Riz-GCiz, show excellent spatial correlation at the 0.99 level. The correlations between the other models are around 0.95.

[53] We extend the illustration of the phase speed models to the higher modes in Figure 10 for GC0 and Riz-GC0 at 60 s. The smoothing effects from the inclusion of finite frequency effects are more pronounced for the higher modes because phase speeds are higher leading to a wider influence zone.

[54] The phase speed perturbations for the higher modes are usually smaller than the fundamental mode model, which indicates that the velocity perturbations tend to decrease in the deeper part of the mantle. Thus the effect of off-great circle propagation is not critical for the higher modes. However, for the fundamental mode at 60 s, there is a clear

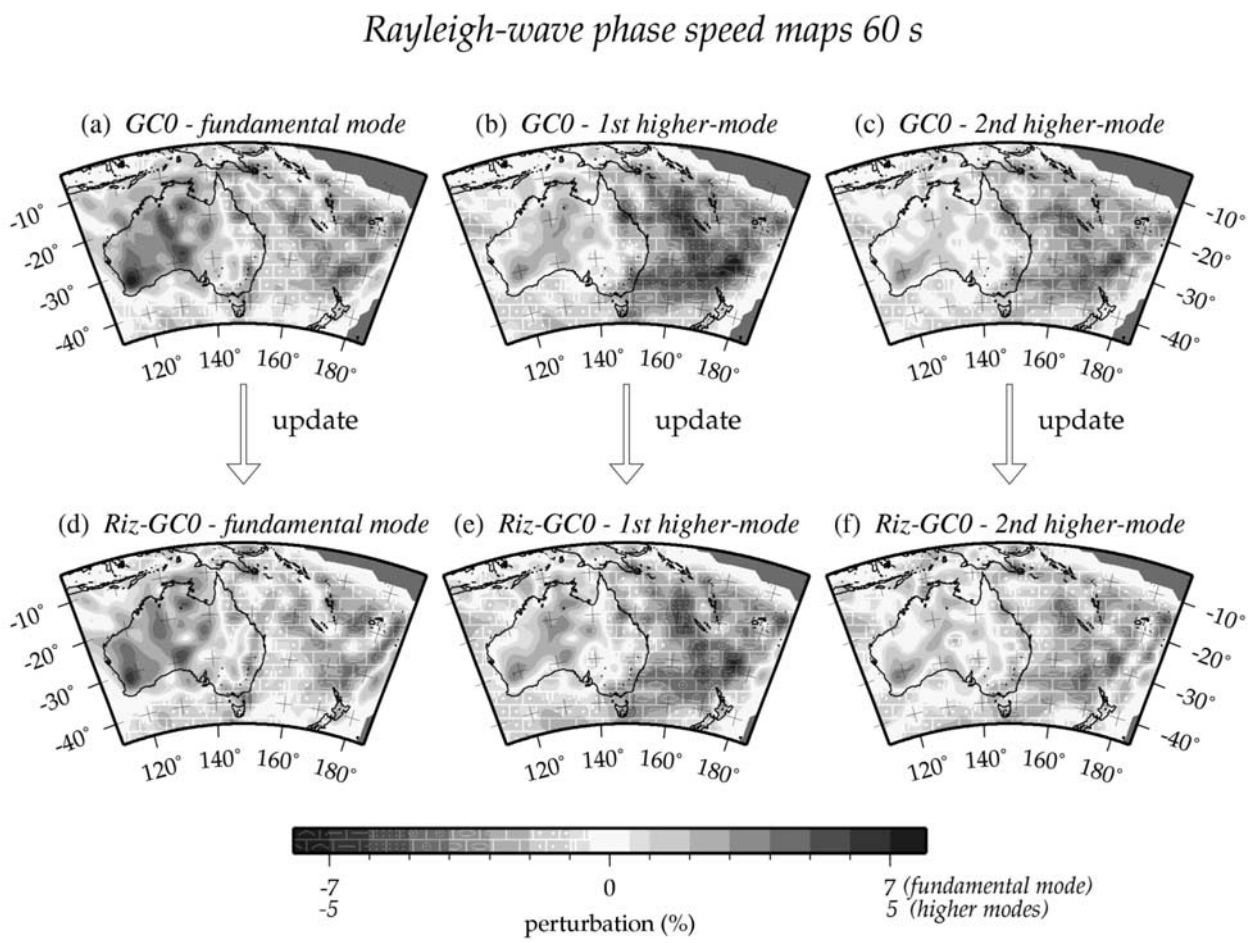

Figure 10. Two sets of phase speed maps for the first three modes of Rayleigh waves at 60 s. All the models are shown as a perturbation from PREM. GC0 models for (a) the fundamental mode, (b) the first higher mode, and (c) the second higher mode. Riz-GC0 models for (d) the fundamental mode, (e) the first higher mode, and (f) the second higher mode, updated from GC0.

difference in the boundaries between the lower and faster phase speed anomalies in the eastern Australia, when ray tracing is employed.

5. Nonlinear Inversion for Local Shear Wave Speed Models

5.1. Multimode Phase Speed Maps and 3-D Shear Wave Speed Structures

[55] The third stage of the three-stage process is to invert for a set of local shear wave speed models using local multimode dispersion information assembled from a set of phase speed maps derived in the second stage. The 3-D wave speed model is then constructed from the full suite of local models.

[56] The relation of multimode phase dispersion and local one-dimensional structure can be represented by a linearized relation as [e.g., *Takeuchi and Saito*, 1972; *Dahlen and Tromp*, 1998]

$$\frac{\delta c(\omega)}{c} = \int_0^R \left\{ K_\rho(\omega, z)\frac{\delta\rho(z)}{\rho} + K_\alpha(\omega, z)\frac{\delta\alpha(z)}{\alpha} + K_\beta(\omega, z)\frac{\delta\beta(z)}{\beta} \right\} dz, \tag{13}$$

where δc is the perturbation of the phase speed; $\delta\rho$, $\delta\alpha$, and $\delta\beta$ are the perturbations of density, P wave speed, and shear wave speed, respectively; and R is the radius of the Earth. K_ρ, K_α, and K_β are the sensitivity kernels which represent the partial derivatives of phase speed with respect to each of the model parameters.

[57] In general, the effects of density and P wave speed on a Rayleigh wave phase speed perturbation are not very significant compared with the influence of shear wave speed, particularly in the intermediate frequency range used in this study. Thus we fix the density and P wave speed structures, and only a perturbation of shear wave speed is considered in the inversion for a 1-D model.

[58] We need to emphasize that the 1-D models in the first stage are path-specific and are employed to estimate multimode dispersion, whereas in this third stage we are interested in an ensemble of local 1-D wave speed models to assemble the 3-D model. The local 1-D models are derived from a nonlinear inversion of local dispersion for a number of modes. The results are sensitive to the local shallow structure, and therefore we need good representations for the crustal structure across the region.

[59] Throughout this study, we use the crustal structure from the 3SMAC model [*Nataf and Ricard*, 1996], which

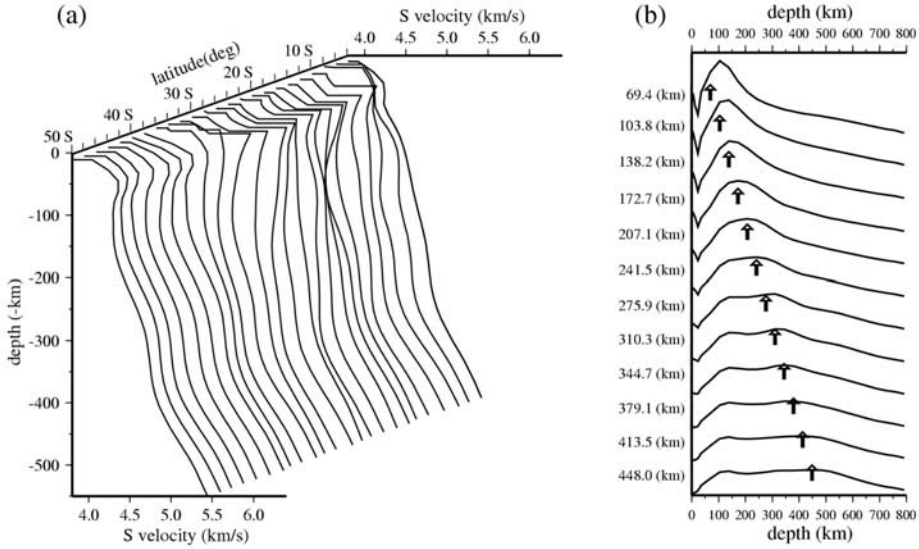

Figure 11. (a) Local shear wave speed models along longitude 130°E with varying latitude. (b) Resolution kernels for shear wave speed structure beneath the location (28°S, 130°E).

defines local wave speed structures at 2° × 2° cells. The use of a global crustal model such as 3SMAC is convenient, but even with corrections it may be an inadequate representation of the true structure in the Australian region. A systematic development of a detailed model of crustal structure is being made, but the 3SMAC model is retained here for comparison with previous studies since the emphasis is on the influence of different approximations in carrying out the inversion for 3-D structure.

5.2. Method of Inversion

[60] The iterative least squares inversion scheme of *Tarantola and Valette* [1982] is used to invert multimode dispersion data for a local 1-D shear wave speed model. This nonlinear inversion procedure has been widely adopted in a number of surface wave studies [e.g., *Montagner and Jobert*, 1981; *Nataf et al.*, 1986; *Cara and Lévêque*, 1987; *Nishimura and Forsyth*, 1989], and the details of the method are explained in these works. Here we explain only the practical application of the method, especially focusing on the choice of a priori smoothing parameters.

[61] In this iterative method, the smoothness of the model is controlled by a priori information, which is used in the linearized inversion as a model covariance. Previous studies have shown that the shear wave speeds in the upper mantle have a peak to peak variation of about 0.5 km/s over a 100-km interval in depth [e.g., *Nishimura and Forsyth*, 1989]. We prefer models which vary smoothly with depth but allow some degree of rapid variation so that we can treat rapid changes with large wave speed perturbation. We therefore employ a correlation length L = 20 km and a standard deviation σ = 0.1 km/s over the depth range below the Moho to construct the model covariance. For the shallower structure above the Moho the correlation length is chosen to be L = 5 km so that more rapid variation is allowed.

[62] We start the local inversion from the PREM model [*Dziewonski and Anderson*, 1981] for the oceanic regions and in continental regions modify the crust and uppermost mantle to allow for the thicker crust (PREMC). Both the PREM and PREMC models are adapted to have smooth variation across the 220-, 400-, and 670-km boundaries. Prior to the inversion, the local crustal structure of the reference model is corrected by using the 3SMAC model [*Nataf and Ricard*, 1996].

5.3. Local Shear Wave Speed Models

[63] The local phase dispersion curves are assembled from the multimode dispersion maps at 2° grid points in both longitude and latitude and then are inverted for local shear wave speed structures. We display the 1-D shear wave speed models for Riz-GCiz along a N-S line at 130°E longitude in Figure 11a and note the high-velocity anomalies near the center of the Australian continent in the upper 300 km. Along this longitude, the continent-ocean boundary occurs around 34°S latitude, and there are slower velocities to the south of this latitude and faster anomalies to the north below 80 km depth, indicating a rather rapid change at the boundary.

[64] The resolution kernels for a typical 1-D shear wave speed profile are shown in Figure 11b. It is apparent that the shallower parts of the 1-D structure from 100 to 300 km are comparatively well constrained. However, the deeper parts of the upper mantle are not well resolved even though we include information up to the third higher mode. The lower resolution at depth is mainly due to the smaller sensitivity of the higher modes to the shear wave speed structure compared with the fundamental mode. There is still some sensitivity around 450 km depth which could not be achieved with just the fundamental Rayleigh mode with our maximum period of 150 s.

6. The 3-D Models in the Australian Region

[65] The final 3-D models are obtained by repeating the inversions for the local shear wave speed profile across the whole region using the dispersion information for each 2° × 2° cell. Five types of 3-D shear wave speed models are

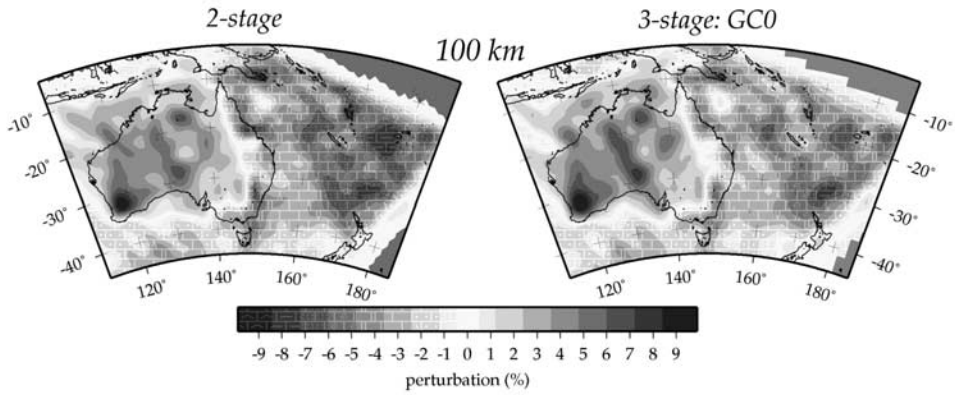

Figure 12. The 3-D shear wave speed models in the Australian region at 100 km depth. (left) Model derived from the two-stage approach and (right) GCO three-stage model derived from the great circle approximation. The reference wave speed is 4.41 km/s.

derived from the corresponding sets of phase speed models in Figure 5.

6.1. Comparison of 3-D Models

6.1.1. Two-Stage and Three-Stage Approaches

[66] The process of three-stage approach is somewhat indirect compared to the conventional two-stage approach, although it provides us with a number of benefits for improving 3-D models. A comparison of 3-D models derived from both two-stage and three-stage inversions is therefore helpful for assessing how the differences in the inversion processes affect the final 3-D models.

[67] We employ a two-stage inversion with the direct use of the same path-specific 1-D profiles of *Debayle and Kennett* [2003] as have been used to estimate the multimode phase speed maps. The 1-D profiles are treated as averages of the 3-D model along the great circle between source and receiver [*Debayle and Kennett*, 2000a] and the inversion of the full set of path information is performed using the LSQR algorithm [*Paige and Saunders*, 1982] with a damped least squares formulation.

[68] The 3-D model derived from the two-stage inversion scheme shows a similar behavior in the trade-off between the misfit and model norm to the phase speed situation. For the region above 200 km, more than 70% variance reductions are achieved through the direct use of the path-average 1-D models, and below 250 km, about 40% variance reductions can be achieved. Nevertheless, the model explains the data quite well with respect to the estimated errors.

[69] A comparison of the 3-D models derived from the two-stage and three-stage approaches GC0 at 100 km depth is shown in Figure 12 under the same assumptions of great circle propagation with no allowance for finite frequency effects. Despite the intrinsic differences in the inversion processes, the final 3-D models are extremely well correlated. The spatial correlation coefficients of these velocity structures exceed 0.95 at all depths. The idealized resolution map for the two-stage model is similar to that shown in Figure 7b. However, it is not at all simple to represent the realistic resolution for a two stage model because unlike the phase speed maps for different frequencies, it is not straightforward to relate the path-average information (i.e., 1-D models in this case) to the effects of finite frequency.

[70] The similarity of the models in Figure 12 indicates that the use of the intermediate step of phase speed maps in the three-stage approach does not bring in any deleterious effects in the final 3-D model for the intermediate frequency range we have used. We can therefore expect to gain from use of the three-stage inversion when we include improved propagation models with frequency dependent off-great circle propagation and allowance for the influence zone around the paths.

6.1.2. Three-Stage Models

[71] Using the five sets of phase speed models in section 4, we obtain five types of corresponding 3-D shear wave speed models. In this section, we mainly focus on the comparison of these 3-D models, especially how the models are improved by considering the influence zone of surface wave paths.

6.1.2.1. Differential Maps at 120 km Depth

[72] The shear wave speed models at 120 km depth for the different types of 3-D models are shown in Figures 13a–13e. The major features of these models are quite similar to those of the corresponding phase speed maps in Figure 9, sharing the similar patterns of the fast and slow wave speed anomalies.

[73] We emphasize the differences between the models by using maps of the differential velocity perturbations $\Delta\beta_{ab}/\beta_0$ between two models a, b compared to the reference profile (PREM). Examples of these differential maps for several combinations of five models are shown in Figures 13f–13j. Figures 13f, 13g, and 13h show differential maps between updated models (including ray tracing) and the corresponding starting models, whereas Figures 13i and 13j show differential maps between two updated models which include the effects of ray bending and/or the influence zone.

[74] It is apparent that the differences between Riz-GC0 and GC0 (Figure 13g) are the most conspicuous among all the differential maps; while there is almost no difference (less than ±0.5%) between the models Riz-GCiz and Riz-GC0 with ray tracing and influence zone effects included (Figure 13j), even though they are derived from different heterogeneous models.

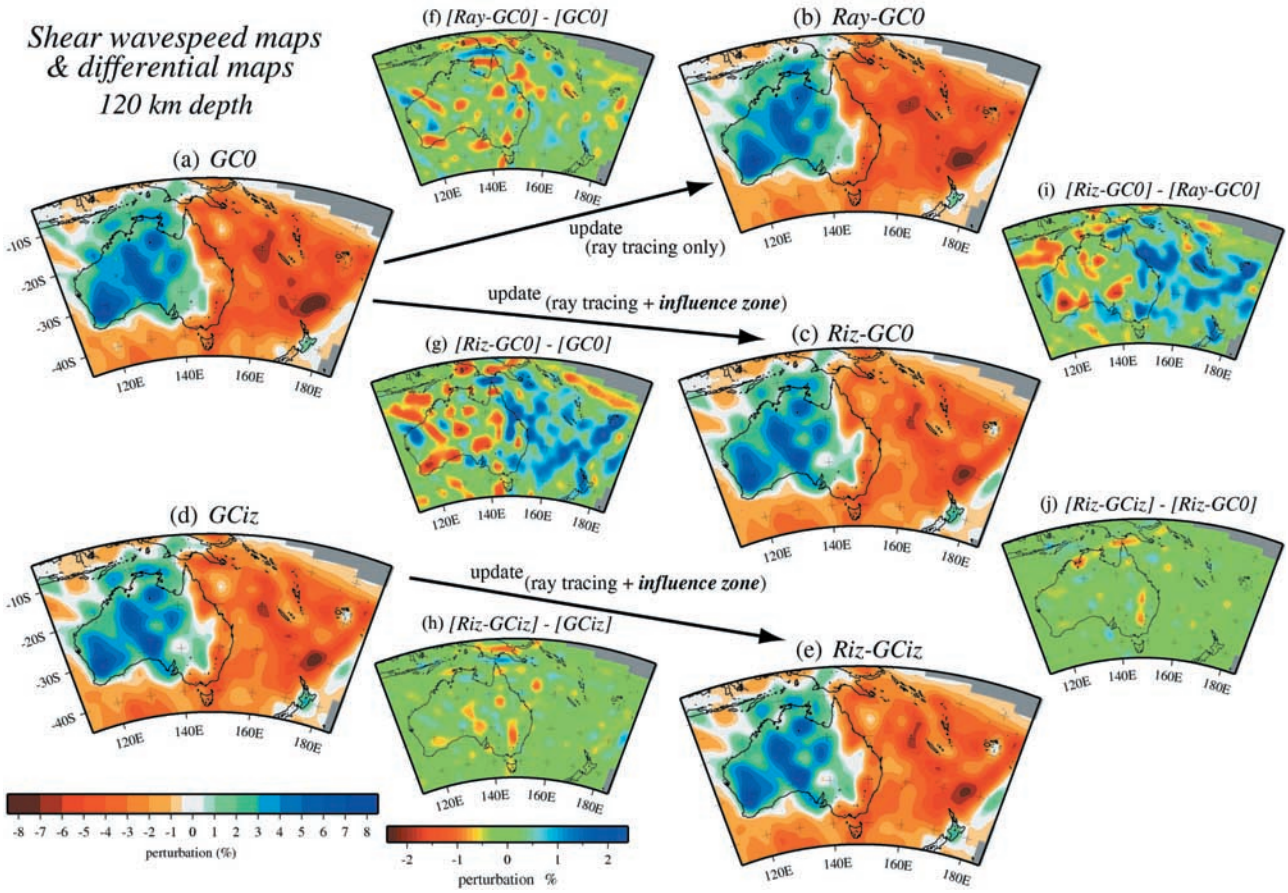

Figure 13. (a–e) Five types of shear wave speed models at 120 km derived from the corresponding phase speed maps in Figure 9 derived by the three-stage approach. Differential maps at 120 km between shear wave speed models in Figures 13a–13e: (f) Ray-GC0 and GC0, (g) Riz-GC0 and GC0, (h) Riz-GCiz and GCiz, (i) Riz-GC0 and Ray-GC0, and (j) Riz-GCiz and Riz-GC0.

[75] The significant differences between Riz-GC0 and GC0 (more than ±1% variation) in Figure 13g suggest that the great circle model GC0, based only on geometrical ray theory, contains a number of erroneous features that are revealed as relatively large differences in the differential velocity perturbation (±1 ∼ 2%). Such erroneous structures for ray theoretical models can be expected from the results of realistic resolution tests in Figures 7 and 8. The inclusion of off-great circle propagation alone (Ray-GC0) is not sufficient to eliminate such errors (Figures 13f, 13g, and 13i). The similarity in patterns of model differences shown in Figures 13g and 13i) indicates that the effect of ray path bending is not sufficient to improve the tomography models at period range in this study.

[76] On the other hand, models derived including the influence zone around the surface wave paths (Figures 13c, 13d, and 13e) share rather similar features of somewhat smoothed heterogeneities, regardless of the inclusion of off-great circle propagation. Figure 13h is a differential map between the best updated model Riz-GCiz and a finite frequency great circle model GCiz. The differences are less than ±1.0% in most areas and are barely noticable, suggesting that the finite frequency model GCiz is rather similar to the model Riz-GCiz even without corrections from ray tracing. For the intermediate frequencies that we have used, the inclusion of an allowance for finite frequency effects using the influence zone makes a more significant contribution to the improvement of tomographic models than ray tracing.

[77] The main differences between the models in Figures 13g and 13i indicate that updating processes including the influence zone make the wave speed in the oceanic region faster and the continental region slower, resulting in smoother variations across the whole region. Velocity differences within the continent are mainly found in those regions with relatively large gradients in velocity, such as near the western and eastern margins of the continent and around the craton in the central Australia. In the oceanic region we can see more pronounced features of subduction in New Hebrides and Tonga-Kermadec with faster wave speeds relative to the surroundings.

6.1.2.2. Effects of the Influence Zone and Structural Implications

[78] The major features of Figure 13, particularly the differences between Riz-GC0 and GC0, are common to other slices through the 3-D models at different depths. We note that in Figures 13a and 13c some small patch-like features in GC0 are smoothed out in Riz-GC0; the effect is

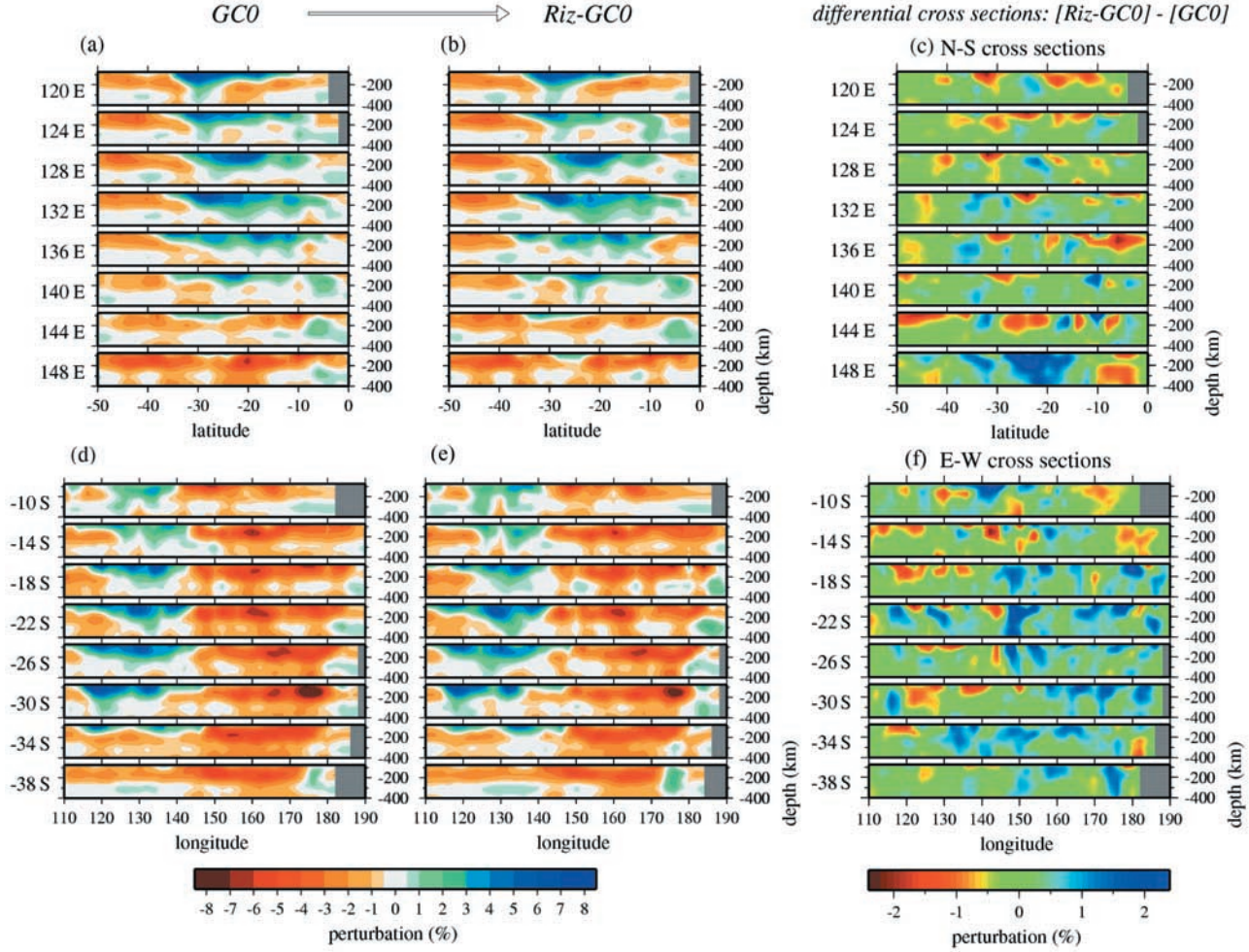

Figure 14. Cross sections of 3-D shear wave speed models for GC0 and Riz-GC0. N-S cross sections through varying longitudes for (a) GC0 and (b) Riz-GC0. E-W cross sections through various latitudes for the model (d) GC0 and (e) Riz-GC0. Differential cross sections between shear wave speed models Riz-GC0 and GC0 along (c) the N-S direction and (f) the E-W direction.

noticable near the faster velocity anomalies beneath the Proterozoic blocks in the center of the Australian Continent and also in the offshore regions with slower velocity in the Coral Sea and Tasman Sea to the east. The off-path smoothing effects in the finite frequency model Riz-GC0 become even clearer in the deeper parts of the mantle below 150 km (not shown). The structures at these greater depths are constrained mainly by the higher modes and longer-period fundamental modes whose influence zones are wider than those for the shorter-period fundamental mode. The effect of the off-path smoothing is preserved in the cross sections (Figure 14) for the same regions where there is noticable smoothing in the heterogeneity in the map views (Figures 13a and 13c).

[79] Differential cross sections between GC0 and Riz-GC0 are shown in Figures 14c and 14f. The most striking feature is that differences in the east of Australia (around 148°E and 20° $\sim$ 30°S) exist down to 400 km depth, whereas in the cratonic region in the central and western Australia the main velocity differences are confined to the top 200 km. This may in part reflect greater sampling by paths with strong higher-mode excitation.

[80] A striking feature of previous models for the structure of Australia derived from two-stage inversion is the rapid transition from fast to slow wave speed at the edge of the craton (around 140°E). This feature is preserved in the three-stage models but with slightly reduced gradients in Riz-GC0 compared with GC0. The zone of change is comparable to the achievable spatial resolution and this suggest that the actual transition must be quite sharp.

[81] In Figure 15 we show a full set of map slices for our final 3-D model Riz-GCiz of the Australian region, which shares an almost identical wave speed distribution with Riz-GC0. We can track the faster wave speed anomaly in the center of the continent in these images down to 300 km depth.

[82] The depth of the root of such continental lithosphere can be estimated from the largest velocity gradient in the wave speed profiles. We can see higher-velocity anomalies down to depths of 200–250 km beneath the middle and western parts of Australia, corresponding to the continental lithosphere of the Australian Continent. In a region just beneath the Proterozoic blocks in the central Australia (around 20°S and 132°) the continental lithosphere seems

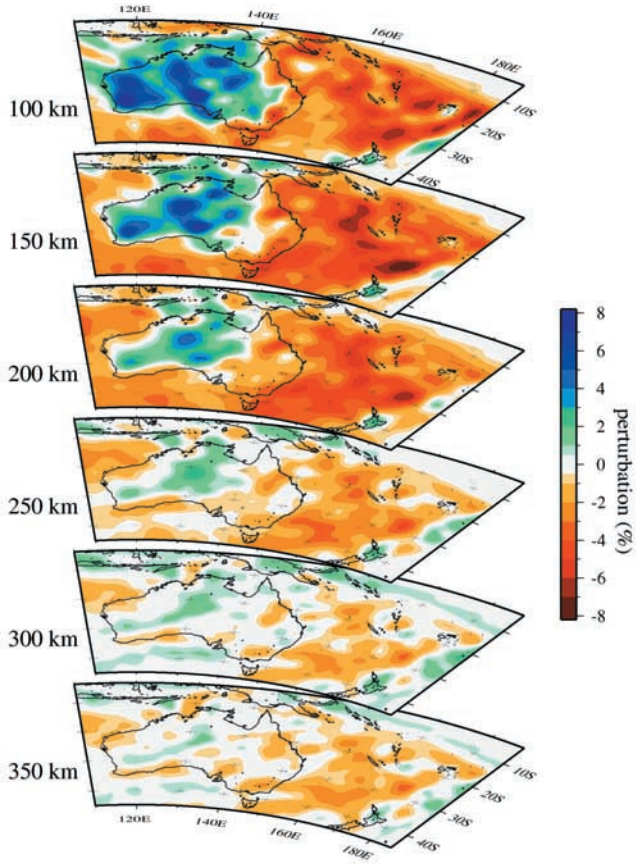

Figure 15. Shear wave speed model Riz-GCiz in the upper mantle. Reference velocities are 4.41 km/s at 100 km, 4.43 km/s at 150 km, 4.51 km/s at 200 km, 4.61 km/s at 250 km, 4.70 km/s at 300 km, and 4.75 km/s at 350 km.

to reach 300 km, in agreement with the results of *Simons et al.* [1999].

[83] The neighboring subduction zones are imaged as zones of relatively fast wave speeds, under New Zealand and to the north of Australia (around 130°E and 0° ∼ 10°S) from the Indonesian subduction zone. However, the details of the subducting plate cannot be resolved with the long-wavelength surface waves.

[84] One of the major difference from the models of *Simons et al.* [1999] is the absence of high wave speed anomalies in eastern Australia below 250 km in the zone 140° ∼ 150°E and 20° ∼ 25°S. At this depth, the wave speed perturbations from the PREM become very small beneath the Australian continent, and the maximum perturbations do not exceed ±2% in our models (Figure 15). The model differences can be attributed to the nature of the data analysis, with different waveform fitting procedures. Our data set of path-specific 1-D models of *Debayle and Kennett* [2003] has been derived from fitting cross correlograms as secondary observables, whereas *Simons et al.* [1999] have used a procedure of fitting the multimode waveforms directly for path-specific 1-D models [e.g., *Nolet*, 1990], which is more sensitive to the choice of initial models to start the inversion [*Hiyoshi*, 2001].

[85] Higher wave speed anomalies in western Australia just beneath the NWAO station, corresponding to the root of the Archaean craton, seem to get fainter below a depth of 250 km in Figures 15 and 14e at 30°S. Earlier studies [*Simons et al.*, 1999; *Debayle and Kennett*, 2000a] have shown high wave speed anomalies at this depth. The weakened higher wave speed anomaly in this region in Riz-GCiz comes from the improved allowance for structure, but the sparse data coverage still gives limited resolution; so we need to be careful about discussing the structure in the western blocks of the Australian continent. Even so, the ray-tracing experiments shown by *Yoshizawa* [2002] suggest that conspicuous ray path deviations from the great circle are likely to appear in the paths from the Tonga-Kermadec region to the NWAO station. Therefore the use of ray tracing in the estimation of the phase speed maps should play a role in suppressing some undesirable effects caused by large ray path deviations to NWAO.

7. Discussion

[86] In this paper, we have completed the development of the three-stage inversion scheme of *Kennett and Yoshizawa* [2002] by obtaining a suite of 3-D shear wave speed models for the Australian region based on the sets of multimode phase speed maps incorporating the effects of finite frequency. In particular, emphasis has been placed on the explicit formulation of the inversion scheme for phase speed maps in the second step, which enables us to incorporate off-great circle propagation and the influence zone for surface wave paths in the frequency and modal domains.

[87] The advantage of the three-stage approach is that various types of information can be combined in a common framework to form a final 3-D model. Polarization anomalies due to ray path bending [*Laske and Masters*, 1996; *Yoshizawa et al.*, 1999], and the influence zone [*Yoshizawa and Kennett*, 2002b] that takes account of finite frequency effects of surface waves can be brought together efficiently by working with multimode phase speed maps at each frequency.

[88] Regional studies cannot recover long-wavelength structure which dominates most global models. However, in the three-stage approach we can employ portions of the phase speed maps from global studies as starting models for the regional tomographic inversion of phase speeds, which enables us to combine tomography models at different scales.

[89] Moreover, we can use any convenient technique in the first stage to make reliable measurements of multimode phase speeds, and thus different data sets derived from different techniques can be combined in producing the tomographic models. All models are viewed as representations of multimode surface wave dispersion and can therefore be brought together in the phase speed maps as a function of frequency for the different modes. With a two-stage approach all the 1-D models for the paths need to have been generated in a common representation to provide comparable constraints on the final 3-D model.

[90] The three-stage inversion scheme requires the computation of phase speed maps for a number of modes followed by inversions for local shear wave speed models. Therefore, in total, the three-stage approach requires more computation than a two-stage process such as partitioned waveform inversion but is still efficient enough to treat the

complicated phenomena of off-great circle propagation and finite frequency effects.

[91] The inclusion of finite frequency effects provides superior results compared with conventional ray theory for both the estimates of variance reduction and the synthetic experiments for determining realistic resolution for phase speed models. The 2-D phase speed models, which are obtained using the influence zone about the propagation path, are naturally smoothed through the incorporation of finite frequency effects and share similar features in the variations of phase speed irrespective of the inclusion of the off-great circle propagation. The effects of this smoothing are retained in the final 3-D shear models. The dominant influence comes from the inclusion of the off-path sampling associated with the finite frequency of the surface waves, although the inclusion of ray tracing is beneficial.

[92] The three-stage approach enables us to include slightly higher frequencies of surface waves than in a two-stage scheme, since the effects of ray bending caused by moderate lateral heterogeneity can be treated through the updating process of phase speed maps. Such effects are particularly important for the fundamental mode, which samples the shallower layers of the Earth with stronger heterogeneity.

[93] The frequency range that we have employed is such that we can still be confident of the use of the approximation of independent mode propagation for the four modes. In order to go to higher-frequency ranges a careful treatment of the effects of the mode branch coupling as well as of crustal structure will be required to avoid undesirable effects on the tomography models [*Kennett and Nolet*, 1990; *Kennett*, 1995]. For such complex structures with strong lateral heterogeneity the assumption of independent mode propagation will be violated, and coupling between mode branches cannot be ignored. The treatment of mode branch coupling in a full 3-D structure is still too complex for the practical use because different directions of propagation for all the scattered waves in a 3-D structure must be considered [*Kennett*, 1998]. With the neglect of backscattering, *Friederich* [1999, 2003] has developed a scheme including mode coupling on a spherical Earth. However, the sharp gradients in shear wave speed found in the Australian region are such that the forward scattering hypothesis may be too restrictive.

[94] The three-stage inversion scheme has considerable potential for further development to include the recovery of azimuthal anisotropy, the combined use of Love and Rayleigh waves and the simultaneous use of group and phase speed maps at the second stage.

[95] **Acknowledgments.** We are grateful to Eric Debayle for the provision of his 1-D velocity models for the Australian region and for numerous useful discussions. We also wish to thank the members of RSES at the Australian National University for their invaluable support for collecting the SKIPPY and KIMBA data and Rob van der Hilst for his efforts to create the data management system. We thank two anonymous reviewers and an associate editor for their useful and constructive comments. K.Y. was supported by a Ringwood Scholarship and an International Postgraduate Research Scholarship at the Australian National University.

References

Alsina, D., R. L. Woodward, and R. K. Snieder (1996), Shear wave velocity structure in North America from large-scale waveform inversions of surface waves, *J. Geophys. Res.*, *101*, 15,969–15,986.

Cara, M., and J. J. Lévêque (1987), Waveform inversion using secondary observables, *Geophys. Res. Lett.*, *14*, 1046–1049.

Dahlen, F. A., and J. Tromp (1998), *Theoretical Global Seismology*, Princeton Univ. Press, Princeton, N. J.

Dahlen, F. A., S.-H. Hung, and G. Nolet (2000), Fréchet kernels for finite-frequency travel times, I. Theory, *Geophys. J. Int.*, *141*, 157–174.

Debayle, E., and B. L. N. Kennett (2000a), The Australian continental upper mantle: Structure and deformation inferred from surface waves, *J. Geophys Res.*, *105*, 25,243–25,450.

Debayle, E., and B. L. N. Kennett (2000b), Anisotropy in the Australian upper mantle from Love and Rayleigh waveform inversion, *Earth Planet. Sci. Lett.*, *184*, 339–351.

Debayle, E., and B. L. N. Kennett (2003), Surface wave studies of the Australian region, in *Evolution and Dynamics of the Australian Plate*, edited by R. R. Hillis and R. D. Müller, *Spec. Pap. Geol. Soc. Am.*, *372*, 25–40.

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

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

Friedrich, W. (1999), Propagation of seismic shear and surface waves in a laterally heterogeneous mantle by multiple forward scattering, *Geophys. J. Int.*, *136*, 180–204.

Friedrich, W. (2003), The *S*-velocity strucutre of the East Asian mantle from inversion of shear and surface waveform, *Geophys. J. Int.*, *153*, 88–102.

Gee, L. S., and T. H. Jordan (1992), Generalized seismological data functionals, *Geophys. J. Int.*, *111*, 363–390.

Hiyoshi, Y. (2001), Regional surface waveform inversion for Australian paths, Ph.D. thesis, Aust. Natl. Univ., Canberra.

Hung, S.-H., F. A. Dahlen, and G. Nolet (2000), Fréchet kernels for finite-frequency travel times, II. Examples, *Geophys. J. Int.*, *141*, 175–203.

Kennett, B. L. N. (1995), Approximations for surface-wave propagation in laterally varying media, *Geophys. J. Int.*, *122*, 470–478.

Kennett, B. L. N. (1998), Guided waves in three-dimensional structures, *Geophys. J. Int.*, *133*, 159–174.

Kennett, B. L. N., and G. Nolet (1990), The interaction of the S-wavefield with upper mantle heterogeneity, *Geophys. J. Int.*, *101*, 751–762.

Kennett, B. L. N., and K. Yoshizawa (2002), A reappraisal of regional surface wave tomography, *Geophys. J. Int.*, *150*, 37–44.

Lancaster, P., and K. Salkauskas (1986), *Curve and Surface Fitting*, Academic, San Diego, Calif.

Laske, G., and G. Masters (1996), Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, *101*, 16,059–16,075.

Lebedev, S. (2000), The upper mantle beneath the western Pacific and southeast Asia, Ph.D. thesis, Princeton Univ., Princeton, N. J.

Marquering, H., G. Nolet, and F. A. Dahlen (1998), Three-dimensional waveform sensitivity kernels, *Geophys. J. Int.*, *132*, 521–534.

Marquering, H., F. A. Dahlen, and G. Nolet (1999), Three-dimensional sensitivity kernels for finite-frequency travel times: The banana-doughnut paradox, *Geophys. J. Int.*, *137*, 805–815.

Meier, T., S. Lebedev, G. Nolet, and F. A. Dahlen (1997), Diffraction tomography using multimode surface waves, *J. Geophys. Res.*, *102*, 8255–8267.

Montagner, J.-P., and J. Jobert (1981), Investigation of upper mantle structure under young regions of the southeast Pacific using long-period Rayleigh waves, *Phys. Earth Planet. Inter.*, *27*, 206–222.

Montagner, J.-P., and T. Tanimoto (1990), Global anisotropy in the upper mantle inferred from the regionalization of phase velocities, *J. Geophys. Res.*, *95*, 4797–4819.

Montagner, J.-P., and T. Tanimoto (1991), Global upper mantle tomography of seismic velocity and anisotropies, *J. Geophys. Res.*, *96*, 20,337–20,351.

Nataf, H. C., and T. Ricard (1996), 3SMAC: An a priori tomographic model of the upper mantle based on geophysical modeling, *Phys. Earth Planet. Inter.*, *95*, 101–122.

Nataf, H. C., I. Nakanishi, and D. L. Anderson (1986), Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy: 3. Inversion, *J. Geophys. Res.*, *91*, 7261–7307.

Nishimura, C. E., and D. W. Forsyth (1989), The anisotropic structure of the upper mantle in the Pacific, *Geophys. J.*, *96*, 203–229.

Nolet, G. (1990), Partitioned waveform inversion and two-dimensional structure under the network of autonomously recording seismographs, *J. Geophys. Res.*, *95*, 8499–8512.

Nolet, G., J. van Trier, and R. Huisman (1986), A formalism for nonlinear inversion of seismic surface waves, *Geophys. Res. Lett.*, *13*, 26–29.

Paige, C. C., and M. A. Saunders (1982), LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Software*, *8*, 43–71.

Ritzwoller, M. H., and A. L. Levshin (1998), Eurasian surface wave tomography: Group velocities, *J. Geophys. Res.*, *103*, 4839–4878.

Ritzwoller, M. H., N. M. Shapiro, M. P. Barmin, and A. L. Levshin (2002), Global surface wave diffraction tomography, *J. Geophys. Res.*, *107*(B12), 2335, doi:10.1029/2002JB001777.

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.

Spetzler, J., J. Trampert, and R. Snieder (2001), Are we exceeding the limits of the great circle approximation in global surface wave tomography?, *Geophys. Res. Lett.*, *28*, 2341–2344.

Stutzmann, E., and J.-P. Montagner (1993), An inverse technique for retrieving higher mode phase velocity and mantle structure, *Geophys. J. Int.*, *113*, 669–683.

Takeuchi, H., and M. Saito (1972), Seismic surface waves, in *Seismology: Surface Waves and Free Oscillations*, *Methods Comput. Phys.*, vol. 11, edited by B. A. Bolt, pp. 217–295, Academic, San Diego, Calif.

Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problems solved using the least-squares criterion, *Rev. Geophys.*, *20*, 219–232.

Trampert, J., and J. H. Woodhouse (1995), Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.*, *122*, 675–690.

Trampert, J., and J. H. Woodhouse (1996), High resolution global phase velocity distributions, *Geophys. Res. Lett.*, *23*, 21–24.

van der Lee, S., and G. Nolet (1997), Upper mantle *S* velocity structure of north America, *J. Geophys. Res.*, *102*, 22,815–22,838.

van Heijst, H. J., and J. H. Woodhouse (1997), Measuring surface-wave overtone phase velocities using a mode-branch stripping technique, *Geophys. J. Int.*, *131*, 209–230.

van Heijst, H. J., and J. H. Woodhouse (1999), Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode branch stripping, *Geophys. J. Int.*, *137*, 601–620.

Vasco, D. W., J. E. Peterson, and E. L. Majer (1995), Beyond ray tomography: Wavepath and Fresnel volumes, *Geophysics*, *60*, 1790–1804.

Wang, Z., and F. A. Dahlen (1995), Spherical-spline parameterization of three-dimensional Earth models, *Geophys. Res. Lett.*, *22*, 3099–3102.

Woodhouse, J. H., and Y. K. Wong (1986), Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. Astron. Soc.*, *87*, 753–773.

Yomogida, K. (1992), Fresnel zone inversion for lateral heterogeneities in the Earth, *Pure Appl. Geophys.*, *138*, 391–406.

Yomogida, K., and K. Aki (1987), Amplitude and phase data inversion for phase velocity anomalies in the Pacific Ocean basin, *Geophys. J. R. Astron. Soc.*, *88*, 161–204.

Yoshizawa, K. (2002), Development and application of new techniques for surface wave tomography, Ph.D. thesis, Aust. Natl. Univ., Canberra.

Yoshizawa, K., and B. L. N. Kennett (2002a), Non-linear waveform inversion for surface waves with a neighbourhood algorithm - application to multimode dispersion measurements, *Geophys. J. Int.*, *149*, 118–133.

Yoshizawa, K., and B. L. N. Kennett (2002b), Determination of the influence zone for surface wave paths, *Geophys. J. Int.*, *149*, 440–453.

Yoshizawa, K., K. Yomogida, and S. Tsuboi (1999), Resolving power of surface wave polarization data for higher-order heterogeneities, *Geophys. J. Int.*, *138*, 205–220.

Zhang, Y.-S., and T. Lay (1996), Global surface wave phase velocity variations, *J. Geophys. Res.*, *101*, 8415–8436.

Zhao, L., T. H. Jordan, and C. H. Chapman (2000), Three-dimensional Fréchet differential kernels for seismic delay times, *Geophys. J. Int.*, *141*, 558–576.

Zielhuis, A., and G. Nolet (1994), Shear-wave velocity variations in the upper mantle beneath central Europe, *Geophys. J. Int.*, *117*, 695–715.

B. L. N. Kennett, Research School of Earth Sciences, Australian National University, Canberra, ACT 0200, Australia. (brian@rses.anu.edu.au)

K. Yoshizawa, Division of Earth and Planetary Sciences, Hokkaido University, Sapporo 060-0810, Japan. (kazu@ep.sci.hokudai.ac.jp)