Geophys. J. Int. (1999) **137,** 601–620

Global high-resolution phase velocity distributions of overtone and fundamental-mode surface waves determined by mode branch stripping

Hendrik Jan van Heijst* and John Woodhouse
Department of Earth Sciences, University of Oxford, Parks Road, Oxford, OX1 3PR, *UK*

Accepted 1998 October 14. Received 1998 August 27; in original form 1998 January 26

SUMMARY

We apply the mode branch stripping (MBS) technique of van Heijst & Woodhouse (1997) to approximately 110 000 three-component seismograms. We select high-quality data with the reliability estimate of van Heijst & Woodhouse (1997). We assess the influence of different selection criteria and remove outliers using smooth degree-12 phase velocity maps. We present final results in terms of Rayleigh- and Love-wave fundamental-mode and overtone global high-resolution ($l=40$) phase velocity maps. We determine the optimum damping of the high-resolution maps with a method based on cross-validation.

Our fundamental-mode phase velocity maps are generally in good agreement with previous studies, especially with Ekström *et al.* (1997). They do, however, contain more short-wavelength structure than previous studies as we apply relatively little damping.

The Rayleigh wave overtone phase velocity measurements made with MBS are of high quality in broad frequency ranges. The measurements are generally well explained by phase velocity maps, and variance reductions for some modes, after rejecting outliers, are as high as 85 per cent. We compare our global phase velocity distributions to the previous results of Stutzmann & Montagner (1994) and to phase velocity maps predicted by 3-D tomographic mantle models. Agreement of the model predictions with our Rayleigh wave phase velocity maps, in terms of both amplitude and observed structures, is good.

For Love waves, the quality of the measurements made with MBS is not as high as for Rayleigh waves. The variance reductions achieved are lower and the agreement between model predictions and our phase velocity maps is less, especially in the frequency ranges where interference between different mode branches is strong.

Finally, as an additional check on the quality of our overtone measurements, we present a comparison of the fundamental-mode Rayleigh phase velocity distribution at 40 s and our fourth Rayleigh wave overtone phase velocity distribution at 62 s. These two modes are similarly sensitive to velocity anomalies in the top of the upper mantle. We show that the two phase velocity maps are in very close agreement.

Key words: global seismology, Love waves, Rayleigh waves, surface waves, tomography, waveform analysis.

1 INTRODUCTION

Mapping the structure of the transition zone between the upper and lower mantle is very important for the understanding of mantle dynamics. Detailed global mapping of the seismic structure in the transition zone has proven to be very difficult using the types of bulk data currently available. Body wave traveltime data as provided by the ISC, although capable of resolving the transition zone structure in specific regions (Spakman 1991; Fukao *et al.* 1992; Widiyantoro & van der Hilst 1996), do not provide global coverage of the transition zone (Robertson & Woodhouse 1995; van der Hilst *et al.*

Now at: Seismological Laboratory, California Institute of Technology, MS 252-21, Pasadena, CA 91125, USA. E-mail: hendrik@gps.caltech.edu

1997). Fundamental-mode surface wave dispersion data [e.g. Trampert & Woodhouse (1995, 1996); Ekström *et al.* (1997); Laske & Masters (1996); Zhang & Lay (1996) (TW95, TW96, ETL97, LM96 and ZL96)], on the other hand, provide very good global coverage but suffer from limited depth resolution below 400 km (Der *et al.* 1970). Moreover, fundamental-mode surface wave dispersion data alone has a limited resolving power when a low-velocity zone is present (van Heijst *et al.* 1994).

Surface wave overtones provide potentially important constraints on the structure of the deeper parts of the upper mantle and the shallower parts of the lower mantle. This has long been recognized and several studies have been devoted specifically to the measurement and use of surface wave overtone dispersion (Nolet 1975, 1977; Cara 1978, 1979; Mitchell 1980; Okal & Jo 1987; Stutzmann & Montagner 1993, 1994). Others have sought to include overtone dispersion information by using full-waveform fitting techniques (Tanimoto 1990; Woodhouse & Dziewonski 1986, 1989; Nolet 1990).

So far, it has not been possible to achieve good global resolution with the techniques aimed at directly measuring overtone dispersion characteristics. The array techniques (Nolet 1975, 1977) provide very limited coverage as they can only be used to measure the average overtone phase velocities under a station array. The hybrid waveform fitting technique of Stutzmann & Montagner (1993) aims to retrieve a path-averaged velocity structure and path-averaged overtone phase velocities at the same time. This technique was used to produce the first estimates of global overtone phase velocity distributions (Stutzmann & Montagner 1994). Although the technique provides global coverage of overtone phase velocities, resolution is limited due to the requirement that for each path several waveforms excited by an 'array' of sources distributed over a large depth range must be available.

Full-waveform techniques have been used widely in global (Woodhouse & Dziewonski 1984, 1986, 1989; Tanimoto 1990; Su *et al.* 1994) and regional (Nolet 1990; Zielhuis & Nolet 1994) tomographic studies. Full-waveform methods use recorded waveforms as direct constraints on mantle heterogeneity without the intermediate stage of estimating the dispersion properties of individual surface wave mode branches. In global studies, waveform techniques are usually applied to so-called mantle waves (predominantly long-period fundamental-mode surface waves) and long-period body wave waveforms. The body wave phases are modelled by a superposition of dispersed surface wave modes or normal modes. It has been noted, however (Li & Tanimoto 1993), that the path-averaged approximation does not describe the sensitivities of the body wave phases to Earth structure accurately. Recently, more accurate modelling of body wave phases, taking the coupling of surface wave or normal modes into account, has been introduced (Marquering & Snieder 1995; Li & Romanowicz 1995) and used in tomographic experiments (Li & Romanowicz 1996; Marquering & Snieder 1996).

Although the full-waveform techniques aim to include some surface wave overtone dispersion information, they usually do not involve extensive modelling of the low overtone signal that arrives just before the fundamental mode. Moreover, the weighting of the different contributions to the waveform as constraints on Earth structure is very difficult.

An alternative way of using overtone information using a waveform inversion technique was introduced by Lerner-Lam & Jordan (1983, 1987) and later extensively used and expanded by others (Cara & Lévêque 1987; Debayle & Lévêque 1997). In this technique a recorded waveform is not used as a direct constraint on the path-averaged velocity structure. Instead, new objective functions, called branch cross-correlation functions (bccfs), are constructed to increase sensitivity to the overtone signal. The bccfs are cross-correlations of the recorded waveforms with synthetic seismograms for individual mode branches that aim to enhance the sensitivity of the objective function to particular mode branches.

Recently, we introduced a method that makes it possible to estimate surface wave overtone phase velocities for single source–receiver combinations [van Heijst & Woodhouse (1997) (VHW97)]. Our mode branch stripping technique (MBS) makes use of bccfs as well. In our technique, however, the bccfs are not used as constraints on path-averaged structure, but as constraints on mode branch phase velocities. This has the potential advantage that the phase measurements can be interpreted in the framework of surface wave ray theory. This means that it is, for instance, possible to account for mode branch- and frequency-dependent off-great-circle path propagation. Moreover, MBS allows for mode branch- and frequency-dependent amplitude perturbations that may result in apparent phase shifts when a single amplitude normalization is applied to the whole multimode waveform.

In this paper, we present results of the application of MBS to approximately 110 000 three-component seismograms. We convert the measurements to phase velocity maps which we compare with phase velocity distributions derived in earlier studies and with phase velocity distributions predicted by 3-D models of the seismic structure of the mantle. In addition to providing constraints on the 3-D structure of the Earth, the construction of phase velocity maps allows us to evaluate the coherence and quality of the overtone phase velocity measurements as well.

2 MEASUREMENT TECHNIQUE

We use the mode branch stripping technique of VHW97 to measure the phase velocities of the fundamental-mode and overtone surface waves. This technique relies on the construction of mode bccfs (Lerner-Lam & Jordan 1983) for narrow frequency bands. In each frequency band the difference between a synthetic bccf and a real data bccf is minimized in terms of phase velocity and amplitude perturbations with respect to a starting model. In this paper, we use the technique to try to retrieve the dispersion properties for the lowest eight mode branches, including the fundamental mode, for both Rayleigh and Love waves.

We use the anisotropic PREM (Dziewonski & Anderson 1981) as the reference model. All synthetic seismograms are corrected for the effects of the Earth's rotation and ellipticity. We parametrize the eigenfrequency and amplitude perturbations as smooth functions of frequency in the period range between 1000 and 30 s using cubic spline basis functions with 29 spline nodes $\bar{\omega}_i$ (Fig. 1). We use the same nodal frequencies for all mode branches. In the measurement process we use a standard statistical non-linear optimization algorithm (Tarantola & Valette 1982a,b) which is stabilized by a smoothness constraint and *a priori* correlations between neighbouring amplitude and eigenfrequency nodes.

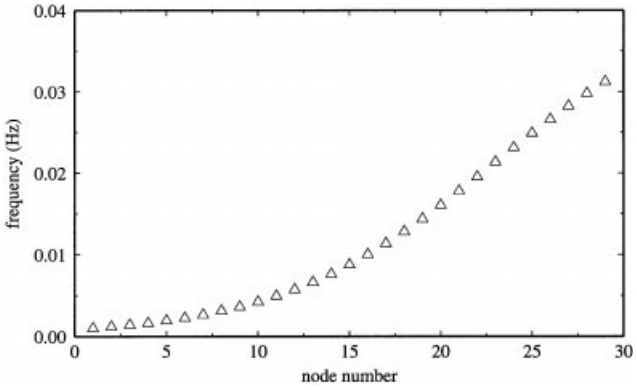

Figure 1. Nodal frequencies of the 29 spline nodes used to parametrize the phase and amplitude perturbations as a function of frequency.

3 DATA RELIABILITY ESTIMATE

Estimating the quality of overtone phase velocity and amplitude measurements is very difficult. In fundamental-mode studies one can usually identify and isolate the fundamental-mode branch waveform from the real data. By comparing predicted and measured waveforms one can then roughly estimate the quality of the dispersion measurements (see for instance TW95 and ETL97). In overtone studies, however, it is generally not possible to isolate individual mode branch seismograms from the data due to the interference of different mode branches. Therefore, the 'waveform fit' for individual overtone mode branches cannot be determined.

In order to assign a measure of quality to individual dispersion measurements we use the reliability estimate of VHW97. The reliability $r^n_{\bar{\omega}_i}$ associated with the phase velocity perturbation measurement for mode branch n at the frequency node $\bar{\omega}_i$ is defined as the inner-product of two vectors:

$$r^n_{\bar{\omega}_i} = \mathbf{p}^n_{\bar{\omega}_i} \cdot \mathbf{f}_{\bar{\omega}_i} . \tag{1}$$

In this expression, $\mathbf{p}^n_{\bar{\omega}_i}$ represents the strength of mode branch n with respect to the other mode branches for a narrow frequency band around $\bar{\omega}_i$, as calculated from the synthetics after fitting a waveform. $\mathbf{f}_{\bar{\omega}_i}$ represents the fit of the full synthetic (i.e. the sum of all synthetic mode branches after adjusting phase velocities and amplitudes) to the recorded seismogram. $\mathbf{p}^n_{\bar{\omega}_i}$ and $\mathbf{f}_{\bar{\omega}_i}$ are represented as n-D vectors as the seismogram is split into n group velocity windows for which the measures of fit and relative mode strength are calculated. (See VHW97 for more details on the reliability estimate.)

In our estimate, the reliability of a phase velocity measurement for a given mode branch at a certain frequency is high if that mode branch contributes significantly to the waveform in one or several group velocity windows and if the recorded waveform in those group velocity windows is explained well by the full synthetic seismogram. We determine the reliability of our measurements at all nodes of the parametrization for all mode branches for which we aim to retrieve the dispersion properties.

Clearly, there is no formal quantitative relation between the reliability estimate and the measurement error. However, examples based on real and synthetic data in VHW97 show that the reliability estimate and the accuracy of measurements are related. The results presented in this paper confirm this. Most importantly, the reliability estimate is a good tool for selecting high-quality measurements. Therefore, it is of great importance to this study as it allows us to automate the data selection for our very large data set.

4 MAKING PHASE VELOCITY MAPS

In surface wave ray theory, the phase delay and ray path of a surface wave for a certain mode branch at a given period are completely determined by the phase velocity distribution of the medium. The average phase velocity along a path between a source and a receiver is given by

$$\frac{\delta c}{c_0} = \frac{1}{\Delta} \int_{\text{ray}} \frac{\delta c(\theta, \phi)}{c_0} \, ds , \tag{2}$$

where $\delta c / c_0$ is the phase velocity perturbation with respect to the phase velocity c_0 in the reference model, Δ is the path length and $\delta c(\theta, \phi)$ represents the phase velocity distribution. The integration in (2) is along the ray path, which itself is dependent on the phase velocity distribution $\delta c(\theta, \phi)$.

In most global surface wave studies to date, the great-circle approximation is used. In this approximation, the integral between source and receiver is not carried out over the actual ray path, which is not known *a priori*, but instead is taken over the corresponding great circle (see e.g. ZL96; TW95; TW96; ETL97). The great-circle approximation has been very successful in explaining large fundamental-mode surface wave dispersion data sets, even for periods as short as 35 s (ETL97). Some studies account for off-great-circle path propagation in the construction of phase velocity maps (Wong 1989; Pollitz 1994; Wang *et al.* 1998). Although these studies show differences between maps that are constructed accounting for the off-great-circle propagation and maps that are constructed using the great-circle approximation, overall agreement between the two types of maps is strong.

The global phase velocity maps that we present here are also derived using the great-circle path approximation. These maps can be considered a first iteration in an iterative modelling procedure that accounts for off-great-circle path propagation [see Woodhouse & Wong (1986) or Pollitz (1994) for such a scheme].

The importance of off-great-circle path propagation effects increases with the size of the anomalies sampled by the surface waves. Surface wave overtones are sensitive to deeper, and therefore probably smaller, heterogeneities than the fundamental-mode surface wave (see e.g. Su *et al.* 1994). Therefore, we expect that the great-circle approximation is sufficiently accurate to construct overtone phase velocity maps.

We expand our phase velocity maps in spherical harmonics $Y_{lm}(\theta,\phi)$. A global phase velocity distribution is then given by:

$$\frac{\delta c(\theta, \phi)}{c_0} = \sum_{l=0}^{L} \sum_{m=-l}^{l} c_{lm} Y_{lm}(\theta, \phi) , \tag{3}$$

where $\delta c(\theta, \phi)$ represents the phase velocity distribution and c_0 the phase velocity in the reference model.

We solve for the coefficients $c_{lm} = \mathbf{m}$ using a damped least-squares technique. Instead of damping towards the smallest model, we damp towards the smoothest model in a first-derivative sense. The optimum model is the model that simultaneously minimizes the data misfit and the squared

surface derivative:

$$|\mathbf{d}-\mathbf{A}\mathbf{m}|^2+\epsilon\mathbf{m}^{\mathrm{T}}D\mathbf{m}\,, \tag{4}$$

where $\mathbf{A}$ is the matrix containing the partial derivatives of the data with respect to the model $\mathbf{m}$, ϵ is the damping parameter that controls the trade-off between data fit and smoothness and D represents the first-derivative operator. We calculate the model that minimizes (4) using an exact inverse technique based on eigenvalue decomposition.

5 CROSS-VALIDATION

A linear inverse problem, like the one described in the previous section, can be damped arbitrarily. In principle, each damping (in our case defined by ϵ) is perfectly valid as long as the results are interpreted together with appropriate measures of error and resolution. Unfortunately, it is virtually impossible to present a comprehensive resolution or error analysis in a paper such as this one. For the reader to be able to test hypotheses, for instance, the full resolution and covariance matrices should be made available. As phase velocity maps are usually interpreted without this information, we seek to damp the inverse problem so that the resulting map approximates the actual global phase velocity distribution as closely as possible.

We use a cross-validation technique (Hastie & Tibshirani 1990) to find an optimum damping parameter ϵ. The optimum ϵ, as determined from cross-validation, approximately minimizes the 'average mean squared error' in the data space. The average mean squared error $\mu(\epsilon)$ is a measure of the difference between the data predicted by model $\mathbf{m}(\epsilon)$ and the data (without measurement errors) corresponding to the 'true' model. The error in the data predicted by $\mathbf{m}(\epsilon)$ is due to the propagation of measurement errors from the data in the modelling procedure and to the imperfect model resolution determined by the damping and the parametrization of the inverse problem:

$$\mu(\epsilon)=\frac{1}{N}\sum_{i=1,N}E(\hat{d}_i(\epsilon)-\bar{d}_i)^2\,, \tag{5}$$

where $\hat{d}_i(\epsilon)$ is the data prediction for datum i from a model constructed with damping ϵ and the sum is over N data. In our case, the index i corresponds to a source–receiver pair. E denotes the 'expectation value' and $\bar{d}_i$ is the data value for source–receiver pair i associated with the 'true' model $\bar{\mathbf{m}}$:

$$d_i=\bar{d}_i+\sigma_i\,, \tag{6}$$

where d_i represents the actual measurement and σ_i the associated measurement error. When the mapping between model and data is linear, the model that minimizes (5) is the closest model to the true model in a least-squares sense, as it is the model that predicts data that is closest to the true data vector $\bar{d}_i$. As we use the great-circle approximation, the relation between phase velocity data and a phase velocity map is linear. Therefore, we expect that the models that we derive using cross-validation are very close to the best model in the average mean squared error sense.

One can estimate the ϵ that minimizes $\mu(\epsilon)$ by constructing the cross-validation sum of squares:

$$c(\epsilon)=\frac{1}{N}\sum_{i=1,N}(d_i-\hat{d}_\epsilon^{-i})^2\,, \tag{7}$$

where d_i is a measurement and $\hat{d}_\epsilon^{-i}$ is the data prediction for data point d_i from the model that is constrained using all data except for d_i and damped with damping parameter ϵ. For normally distributed measurement errors σ_i, it can be shown that the ϵ that minimizes the cross-validation sum (7) is equivalent to the value that minimizes the average mean squared error (5) (Hastie & Tibshirani 1990).

For N data points, constructing the full cross-validation sum of squares requires solving the inverse problem N times, leaving one data point out for each of the N runs. In addition, for each of these N runs, a model needs to be determined for a range of damping parameters ϵ to determine the dependence of c upon ϵ. Because of the large number of data that we consider in this study and the computational time required for solving the inverse problem for a high-resolution phase velocity map, it is impracticable to construct the full cross-validation sum. Therefore, we subdivide each phase velocity data set into M (typically 20) random subsets. We then treat each of these subsets as one 'data point' i for the purpose of calculating the cross-validation sum. This means that we group the data for $M-1$ subsets together, solve the inverse problem as explained in the previous section for a range of damping parameters ϵ and study how the derived model explains the data for the Mth subset that is not included in the data set used to construct the model. Analogous to (7) we write for the approximate cross-validation sum

$$c'(\epsilon)=\frac{1}{M}\sum_{j=1,M}(\mathbf{d}_j-\hat{\mathbf{d}}_\epsilon^{-j})^2\,, \tag{8}$$

where the sum is now over M data subsets given by data vectors $\mathbf{d}_j$.

In Fig. 2(a), we show the residual variance curves $(\mathbf{d}_j-\hat{\mathbf{d}}_\epsilon^{-j})^2$ for the $M=20$ data subsets for our 200 s Rayleigh wave phase velocity distribution. The dependence on the damping parameter ϵ is represented in terms of the effective number of unknowns modelled as given by the trace of the resolution matrix. For this map, we use approximately 30 000 measurements that we expand in spherical harmonics up to degree 40 (1681 unknowns). The behaviour observed in Fig. 2(a) is typical for cross-validation: as the damping is reduced and the number of unknowns modelled increases, the residual variance for the data subset that is not included in the modelling initially decreases. At some point, however, this trend reverses and the residual variance for the subset increases with decreasing damping. This is where the model is fitting data in the the $M-1$ subsets that does not agree with the data in the Mth subset that is not modelled.

To construct the approximate cross-validation sum of squares (8) shown in Fig. 2(b), we sum the 20 curves of Fig. 2(a). The sum has a well-defined minimum when approximately 750 unknowns are modelled. This defines the optimum damping for the inverse problem. If we were to reduce the damping any further we would over-interpret the data and model incoherent signal in the data.

Cross-validation is limited by the approximations that we use in our modelling, such as the neglect of off-great-circle path propagation effects, of azimuthal anisotropy and of source and receiver effects. These effects may result in coherent noise which cross-validation cannot distinguish from coherent signal. Nevertheless, cross-validation establishes the minimum damping that one should apply.

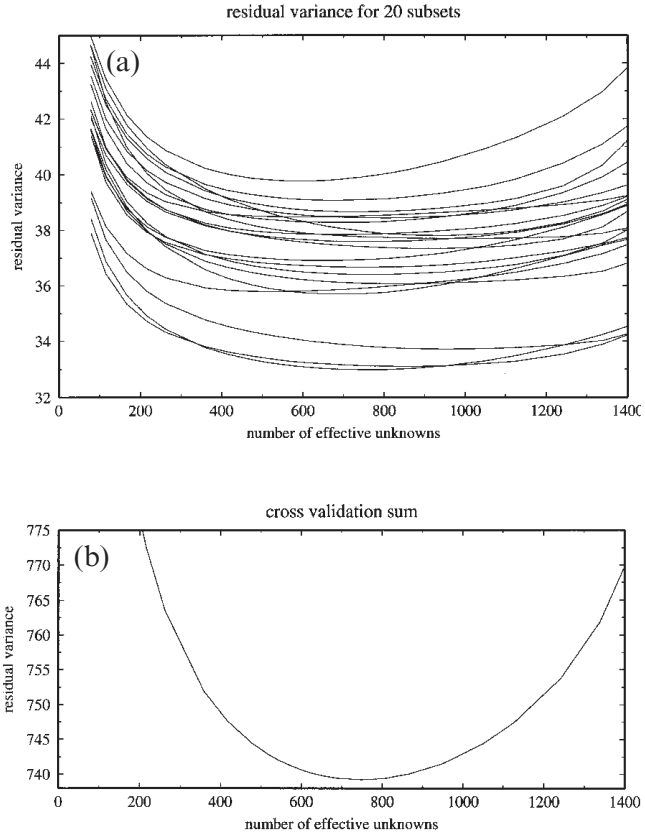

Figure 2. Residual variance curves (a) and the sum of residual squares (b) as a function of the number of effective unknowns calculated for 20 randomly selected subsets in a cross-validation experiment. The data used are 200 s fundamental-mode Rayleigh wave phase measurements. Note the well-defined minimum in cross-validation sum that corresponds to the optimum model in a least-squares sense. See text for details.

6 DATA

We have applied MBS to approximately 110 000 three-component minor arc seismograms that were recorded between 1980 and 1996 on the GDSN and GEOSCOPE (only between 1986–1991) networks. We resample all data traces (group velocity window between 8 and 3 km s^{-1}) at 4 s intervals, detrend them and remove their mean. We rotate the horizontal components to a longitudinal and a transverse trace.

In most surface wave studies, only vertical-component and transverse-component seismograms are processed. In this study, we also use the longitudinal-component seismograms to make Rayleigh wave phase velocity measurements. These measurements are made independently from the vertical component. The main reason for analysing the longitudinal component in addition to the vertical component is that some overtones have a stronger amplitude relative to the noise level on the longitudinal component than on the vertical component. Moreover, the longitudinal component Rayleigh seismograms contribute additional independent fundamental-mode phase velocity measurements. Of these measurements a small fraction, of the order of 5 per cent, is really additional; that is, the phase velocities are measured for source–receiver combinations that have not rendered a reliable measurement from the vertical component.

In Fig. 3, we show the distribution of events and stations used in this study. We use the source mechanisms, times and locations as published in the Harvard CMT catalogue. We use data for all events in the Harvard CMT catalogue (approximately 1700) between 1980 and 1996 with a seismic moment greater than 1×10^{25} dyne-cm. In addition, we have processed 2000 smaller events (M_0 between 3×10^{24} and 1×10^{25} dyne-cm) from regions with relatively sparse data coverage, especially the Southern Hemisphere. Although these smaller events are less likely to contribute high-quality phase velocity measurements than big events, they contribute a significant number of valuable phase data.

The great majority of events that we use are shallow (depth $<$50 km). Overtones are usually not as well excited by shallow earthquakes as by deep earthquakes. For many shallow events, however, some overtone dispersion properties can be measured when the signal-to-noise ratio is high. Of course, including shallow events improves the global data coverage significantly.

In this paper, we consider only minor arc surface waves recorded at stations at epicentral distances greater than 30° and smaller than 160°. We find that it can be difficult to make accurate overtone dispersion measurements for paths shorter than 30° . This is due to both the strong influence of the initial source phase and the very strong interference of the different mode branches in the time domain, even for mode branches with distinct group velocities. We do not process seismograms for paths longer than 160° to avoid the interference of waves at the antipode.

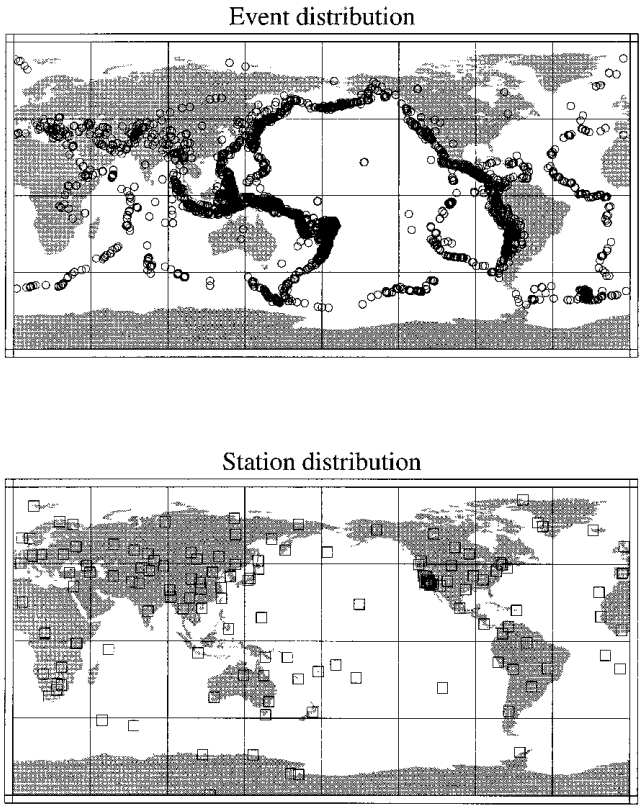

Figure 3. Sources (3952) and stations (164) used in this study. Only a few stations were operational through the whole time period covered by this study (1980–1996).

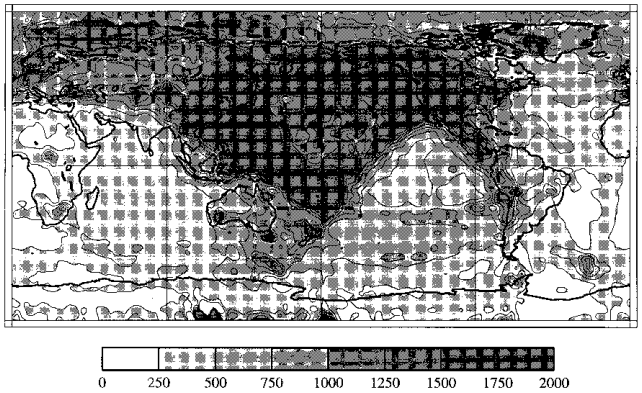

Figure 4. Ray coverage of all the seismograms processed. The scale represents the number of great-circle rays going through each 333×333 km surface area. For most modes, many of the processed seismograms will not yield acceptable measurements. For the best sampled mode (fundamental-mode Rayleigh wave at 100 s) the coverage is approximately half as dense as shown in this figure. Note the very strong bias towards the Northern Hemisphere as only minor-arc seismograms are used in this study.

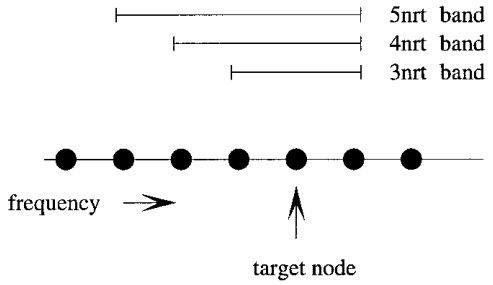

Figure 5. Schematic representation of reliability threshold bands (rtbs). The measurement reliability (van Heijst & Woodhouse 1997) has to exceed the reliability threshold at all nodes within a rtb to be accepted. Reliability threshold bands of three (3nrtb) and five (5nrtb) nodes are used in this study. See text for details.

In Fig. 4, we show the data coverage of all paths that we have processed. Fig. 4 shows the number of great circle rays that hit a 333×333 km ($3 \times 3^\circ$ on the equator) surface area. Due to the distribution of sources and stations, the distribution is strongly biased towards the Northern Hemisphere. Obviously, we do not actually succeed in making reliable measurements for all the paths in Fig. 4, but for many modes the pattern of the data coverage of reliable measurements is similar to Fig. 4. For some overtones, however, the bias towards the Northern Hemisphere may be stronger than in Fig. 4, as they are only excited strongly by deep earthquakes.

7 DATA SELECTION AND QUALITY

Selecting good-quality data is a very important step in automated phase velocity determinations. We select high-quality phase velocity measurements using a reliability threshold based on the reliability estimate discussed in Section 3.

Normally, the reliability spectrum for a given mode branch is a smooth function of frequency. Occasionally, however, peaks arise at single-frequency nodes in the reliability spectrum. A reliability peak at a single node is usually associated with noise, and the corresponding phase measurement should not be accepted. Therefore, for a phase measurement at a certain frequency node to be accepted, we require the reliability to exceed the threshold value not only at the node considered but also in a frequency band around the node.

It is usually sufficient to require the reliability at the node itself and at its two neighbouring nodes (one higher frequency and one lower frequency) to exceed the threshold. We call this a three-node reliability threshold band (=3nrtb). We may also insist that the reliability should exceed the threshold by one more lower-frequency node than in the 3nrtb (=4nrtb), or even two more lower-frequency nodes (=5nrtb), as is schematically shown in Fig. 5.

Occasionally, especially at very long periods and for very short paths where the signal-to-noise ratio is low, phase velocity measurements selected based on reliability can be biased towards zero perturbation. This is due to the *a priori* covariances that we use to regularize the phase velocity retrieval algorithm (eq. 24 in VHW97): in the phase velocity measurement process there is a trade-off between waveform fit and deviation of the phase velocities from the reference model. At the longest periods considered during waveform fitting, the regularization term may dominate this trade-off and hence bias the measurements towards the reference model. Although the waveform fit in this case will not be perfect, the reliability may be reasonably high, especially when the fit is acceptable in a wide group velocity window. To make sure that we do not include measurements that may suffer from this bias, we require that the phase velocity perturbation at a given node has been inverted for in at least two frequency steps (paragraph 4.1, VHW97). This condition is usually sufficient to remove potentially biased measurements.

After the initial selection of phase velocity measurements using an appropriate reliability threshold band, we construct damped phase velocity maps up to spherical harmonic degree 12 solving for a small effective number of unknowns (typically 80). We then reject those paths for which the misfit after modelling is greater than twice the standard deviation of the misfit distribution. This ensures that outliers (mostly 2π phase shifts at the shorter periods) are removed from the phase velocity perturbation data set. The remaining data are of high quality and we use them to resolve global phase velocity distributions with high lateral resolution.

To illustrate the data selection procedure outlined above, we use our vertical-component measurements for the first Rayleigh wave overtone at 100 s. Using the measurements selected on reliability, we construct a phase velocity map up to degree 12 using approximately 80 effective unknowns. In Table 1, we present variance reductions, in terms of both phase velocity anomalies and total phase anomalies accumulated along the path, for data that are selected using different reliability thresholds and nrtbs.

The variance reductions increase when a higher-reliability threshold is imposed. Similarly, the variance reductions increase by using a 5nrtb instead of a 3 nrtb. These increases in variance reductions can partly be attributed to the fact that fewer data points are modelled using the same number of unknowns for the stricter reliability thresholds. Inversions of data subsets randomly selected from the initial 3nrtb, rt=0.5 data set, however, show a much smaller increase in variance reduction than inversions of the subsets that are selected based

Table 1. Number of data selected using six different reliability threshold bands for the first overtone at 100 s. The variance reductions quoted are for smooth degree-12 maps with approximately 80 effective unknowns. Both the variance reduction for the total phase anomalies and the variance reduction for the phase velocity anomalies are shown.

rej. criteria	nr of paths	var. red. $\frac{\delta c}{c_0}$ (%)	var. red. tot. phase (%)
3nrtb, $rt=.25$	15101	50	61
5nrtb, $rt=.25$	13770	51	63
3nrtb, $rt=.5$	13186	53	65
5nrtb, $rt=.5$	11484	55	68
3nrtb, $rt=1$	10634	59	70
5nrtb, $rt=1$	8521	61	74

on reliability and nrtb. This implies that the increased variance reductions are indeed due to the selection of higher-quality data.

In this example, we invert the measured phase velocity anomalies to obtain the phase velocity maps. As shown in Table 1, the variance reductions for the total phase anomalies are significantly higher than those for the phase velocity anomalies. This reflects the fact that the phase velocity anomalies of longer paths are usually better explained by phase velocity maps than the phase velocity anomalies for shorter paths. This effect is commonly observed in phase velocity studies and it can largely be attributed to initial source phase anomalies. Source phase anomalies—due to mislocation and misorientation of the estimated source—have a relatively larger effect on the observed phase velocity anomalies for short paths than for long paths. The influence of an initial source phase error on the total observed phase anomaly, however, is not dependent on epicentral distance. This implies that in order to obtain data of roughly uniform variance it is appropriate to invert total phase anomalies instead of phase velocity anomalies.

Interestingly, experiments show that the actual differences between maps constructed by inverting phase velocity and maps constructed inverting total phase are very small. For most modes the correlations between the two types of maps up to degree 12 are of the order of 0.98. We attribute this to the high data quality and the large number of data that are used to solve for relatively few unknowns.

In Table 2, we present the variance reductions for the same reliability thresholds and nrtbs as in Table 1, but after rejecting outliers. We define outliers as those measurements that do not fit the initial smooth degree-12 phase velocity map to within a specified misfit. In this case, we have rejected of the order of 7 per cent of the original data sets as outliers. Importantly, the outliers are geographically similarly distributed to the remainder of the data set. As expected, the variance reductions improve significantly by rejecting data, rendering a high-quality data set.

Table 2. Same as Table1, but after rejecting outliers.

rej. criteria	nr of paths	var. red. $\frac{\delta c}{c_0}$ (%)	var. red. tot. phase (%)
3nrtb, $rt=.25$	14078	70	76
5nrtb, $rt=.25$	12845	71	77
3nrtb, $rt=.5$	12340	74	79
5nrtb, $rt=.5$	10758	75	80
3nrtb, $rt=1$	9964	76	81
5nrtb, $rt=1$	7978	78	83

Of course, achieving a high variance reduction is not an aim in itself. Instead, we seek to determine global phase velocity distributions as accurately as possible. When we reject noisy measurements we may improve the quality of the constraints on the phase velocity distribution, but at the same time we reduce the number of data constraints. The trade-off between these two effects makes careful selection of the rejection criteria very important.

8 FUNDAMENTAL-MODE PHASE VELOCITY MAPS

Using the mode branch stripping technique we can estimate fundamental-mode phase velocities from long periods (300 s) to short periods (35 s). As the stripping technique accounts for overtone signal, we are not limited to studying the fundamental mode from shallow earthquakes as is the case with most methods used to date. In this section, we present phase velocity maps for the fundamental-mode Love and Rayleigh waves at selected periods and compare them with maps obtained in other studies.

In Fig. 6, we show the number of Love and Rayleigh wave fundamental-mode measurements that we retrieve for different reliability thresholds as a function of frequency. Increasing the reliability threshold and the width of the reliability threshold band both reduce the number of measurements that are accepted.

The largest number of reliable measurements is retrieved for both Rayleigh and Love waves at periods around 100 s. At longer periods, the number of reliable measurements is smaller as the signal-to-noise ratio deteriorates. At shorter periods, the number of reliable measurements is smaller due to insufficient excitation of short-period fundamental modes by deep events. At shorter periods the waveforms are probably also affected by wave propagation effects such as multipathing and scattering that cannot be modelled by phase velocity and amplitude perturbations that are smooth as a function of frequency.

Horizontal-component seismograms are usually significantly noisier than vertical-component seismograms. Therefore, only half as many reliable Rayleigh wave phase velocity measurements can be made from the longitudinal component as from the vertical component.

As outlined above, the first step in our modelling procedure is the construction of smooth degree-12 phase velocity maps. For these maps, we model the total phase anomalies using approximately 80 effective unknowns.

In Fig. 7(a), we show the variance reductions achieved with the smooth degree-12 maps for the Rayleigh wave fundamental mode for data selected using a 3nrtb of 5 and a 3nrtb of 10 before rejecting outliers. We present the variance reduction before rejecting outliers because it is a more honest measure of the quality and internal consistency of our measurements than the variance reduction after rejecting outliers. As expected, the variance reduction achieved for the $r=10$ data is higher than for the $r=5$ data. The same is observed for the Love wave data in Fig. 7(b).

To make high-resolution phase velocity maps, we expand the $r=5$, 3nrtb data, after rejecting outliers, in spherical harmonics up to degree 40. We determine the optimum damping parameters using cross-validation as outlined in Section 5.

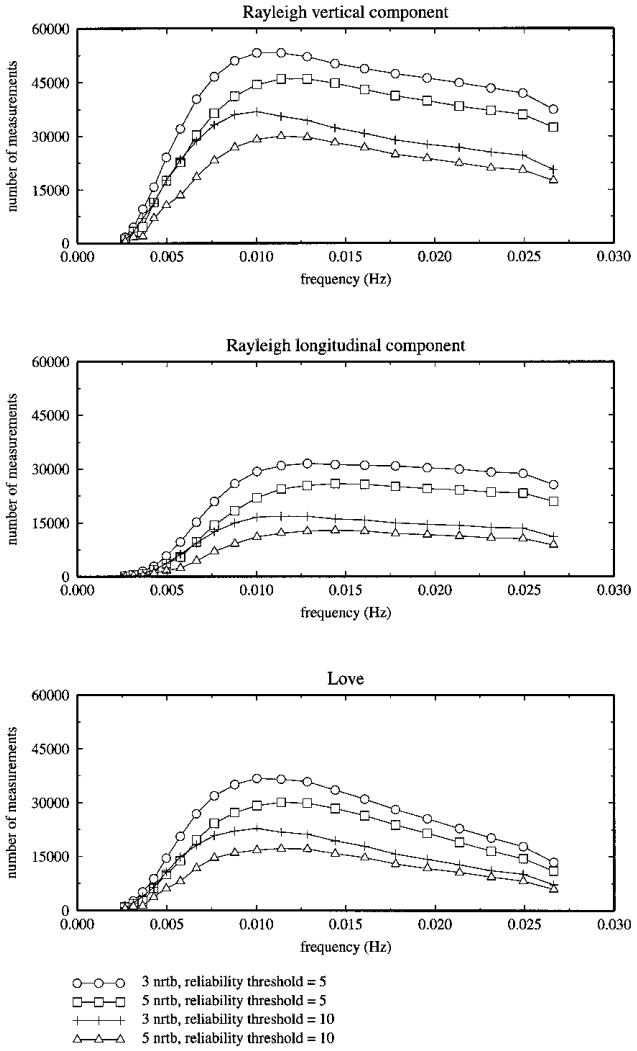

Figure 6. Number of accepted measurements as a function of frequency selected using different rejection criteria for the three different components. Increasing the reliability threshold and widening the reliability threshold band decrease the number of data that are accepted.

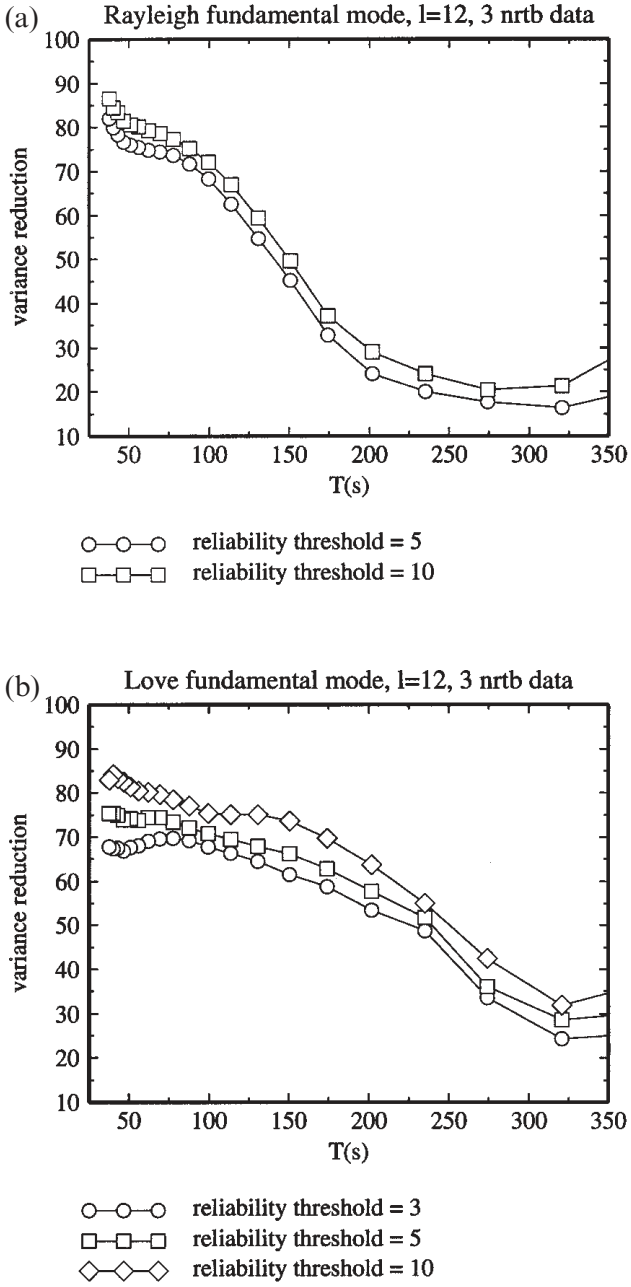

Figure 7. Variance reductions as a function of period for the fundamental-mode phase velocity measurements. The data are expanded in spherical harmonics up to degree 12 and approximately 80 effective unknowns are modelled. For the Rayleigh waves, the vertical-component and longitudinal-component data sets are combined. Note that the variance reductions increase with increasing reliability threshold. Also note the strong increase in variance reduction with decreasing period. See text for details.

The optimum models found using cross-validation tend to be significantly less damped than models damped according to more subjective criteria. There are two main reasons for this. The first and probably most important reason is that cross-validation cannot distinguish between coherent signal and coherent noise, as mentioned in Section 5. Therefore, if possible, coherent noise will be fitted with model structure. Coherent noise is undoubtedly present in our data sets, especially due to event mislocations and misrepresentation of the local structure in source and receiver regions. Moreover, unmodelled wave propagation effects may result in coherent noise. The second, rather trivial, reason is that less damping usually gives rise to higher-amplitude small-scale heterogeneity. This looks unusual and 'blobby' on a global scale so that researchers may have tended to over-damp global phase velocity maps.

The high-resolution maps that we present in this paper are damped using the cross-validation result as a reference. We choose our models to have approximately 75 per cent of the number of effective unknowns that should be modelled according to the cross-validation criterion. Unfortunately, this step makes selecting damping somewhat subjective again. By choosing a slightly stronger damping than the damping suggested by cross-validation we aim to decrease the influence of coherent noise on the model. This does not, however, completely remove the effects of coherent noise. This can only be achieved with more sophisticated modelling techniques.

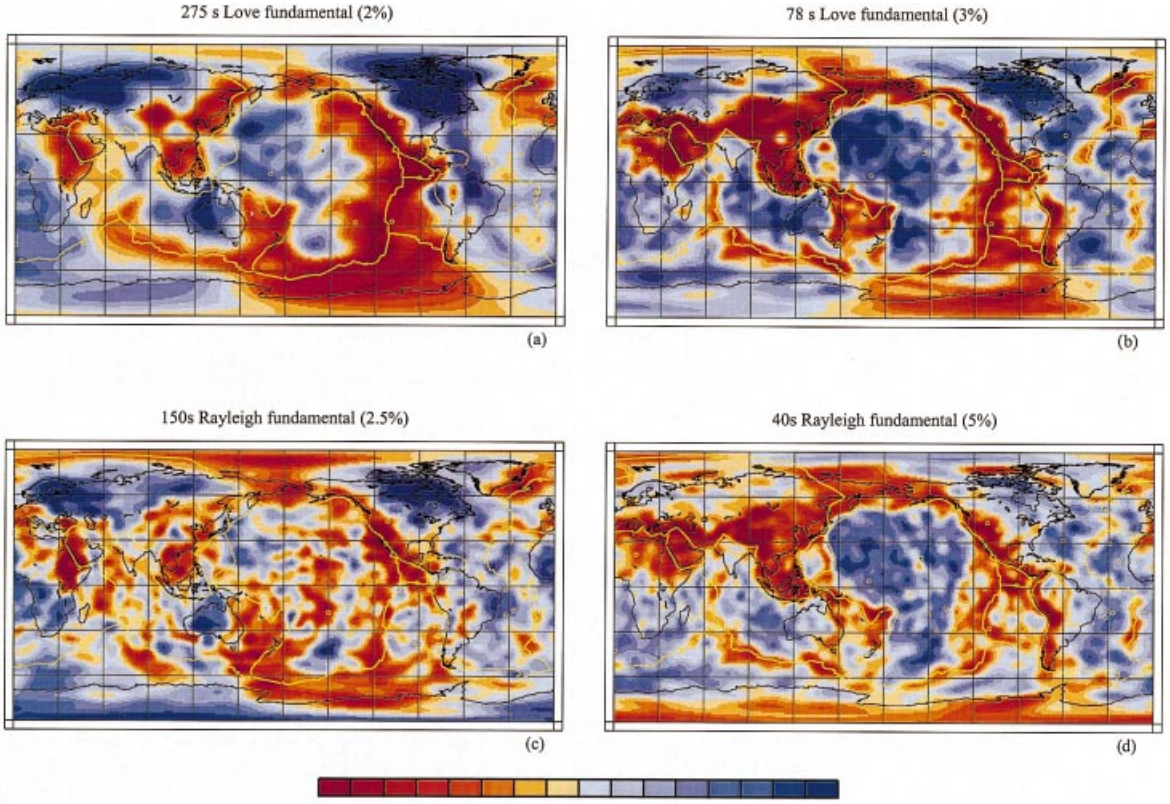

Figure 8. Fundamental-mode phase velocity maps determined in this study. The number in brackets at the top of each figure corresponds to the maximum and minimum used in the colourscale that is shown at the bottom of the plate. See text for further details.

In Fig. 8, we present four representative fundamental-mode Love and Rayleigh wave phase velocity maps. The Love wave maps are for periods of 275 s and 78 s. The Rayleigh wave phase velocity maps are for periods of 150 s and 40 s. All maps are expanded in spherical harmonics up to degree 40, except for the 275 s Love wave map. This map is expanded up to degree 25 as there are relatively few reliable phase measurements for this mode available (in this paper we may refer to a given mode branch at a given period as a 'mode'). In Table 3, we list the number of data used to constrain each map, the number of unknowns modelled and the variance reductions achieved. For reference, we also list a few modes that are not shown in Fig. 8.

In Fig. 9, we show the amplitude spectra of our maps and those of W89, LM96, TW96 and ETL97. We also include degree-by-degree correlations between our maps and the phase velocity distributions obtained in the other studies.

Table 3. Number of measurements, number of effective unknowns modelled and total phase variance reduction for the fundamental-mode high resolution ($l=40$) phase velocity maps.

branch	period (s)	nr of data	nr unkn. mod.	var. red. (%)
0S	200	20480	574	55
0S	150	53217	975	63
0S	100	78027	1119	88
0S	40	65376	1068	92
0T	275	4874	246	65
0T	150	25682	671	86
0T	78	33411	826	90
0T	40	16387	666	91

The 275 s Love wave phase velocities agree very well with the work of LM96 and W89, especially up to degree 6. This confirms that our fundamental-mode phase measurements from minor-arc waveforms are reliable at long periods. For degrees higher than 8, however, agreement between the models is poor. This may be due to the small number of measurements in our work at this period, although the correlation between LM96 and W89 is poor at degrees higher than 8 as well (not shown). The amplitude spectra of our map and the other two maps are very similar, except between degrees 8 to 12, where LM96 and W89 have an oscillating amplitude spectrum.

Our Love wave phase velocity map at 78 s agrees extremely well with results from ETL97 at 75 s. The overall correlation of the two maps is as high as 94 per cent and degree-by-degree correlations are significant throughout the spectrum (Fig. 9b). Correlations with LM96 (83 s) and TW96 (80 s) are lower (73 and 81 per cent overall), but nevertheless very high.

The two Rayleigh wave phase velocity distributions in Fig. 8 appear to be quite different from the distributions obtained in previous studies. This apparent difference is most striking at 150 s. When analysed more carefully, however, our maps correlate very well with those of W89, LM96 and TW96 (Figs 9c and d). The correlations with ETL97 are especially high: 85 per cent overall at 150 s and 92 per cent overall at 40 s. What makes our maps appear so different from ETL97 and LM96 is that they contain more power at shorter wavelengths. Note, for instance, that the amplitude spectrum of our map at 150 s is almost flat beyond degree 20. The difference between our amplitude spectrum and that of ETL97 is primarily due to the fact that we apply less damping; the similarity of the amplitude spectra of our model at 150 s and that of ETL97

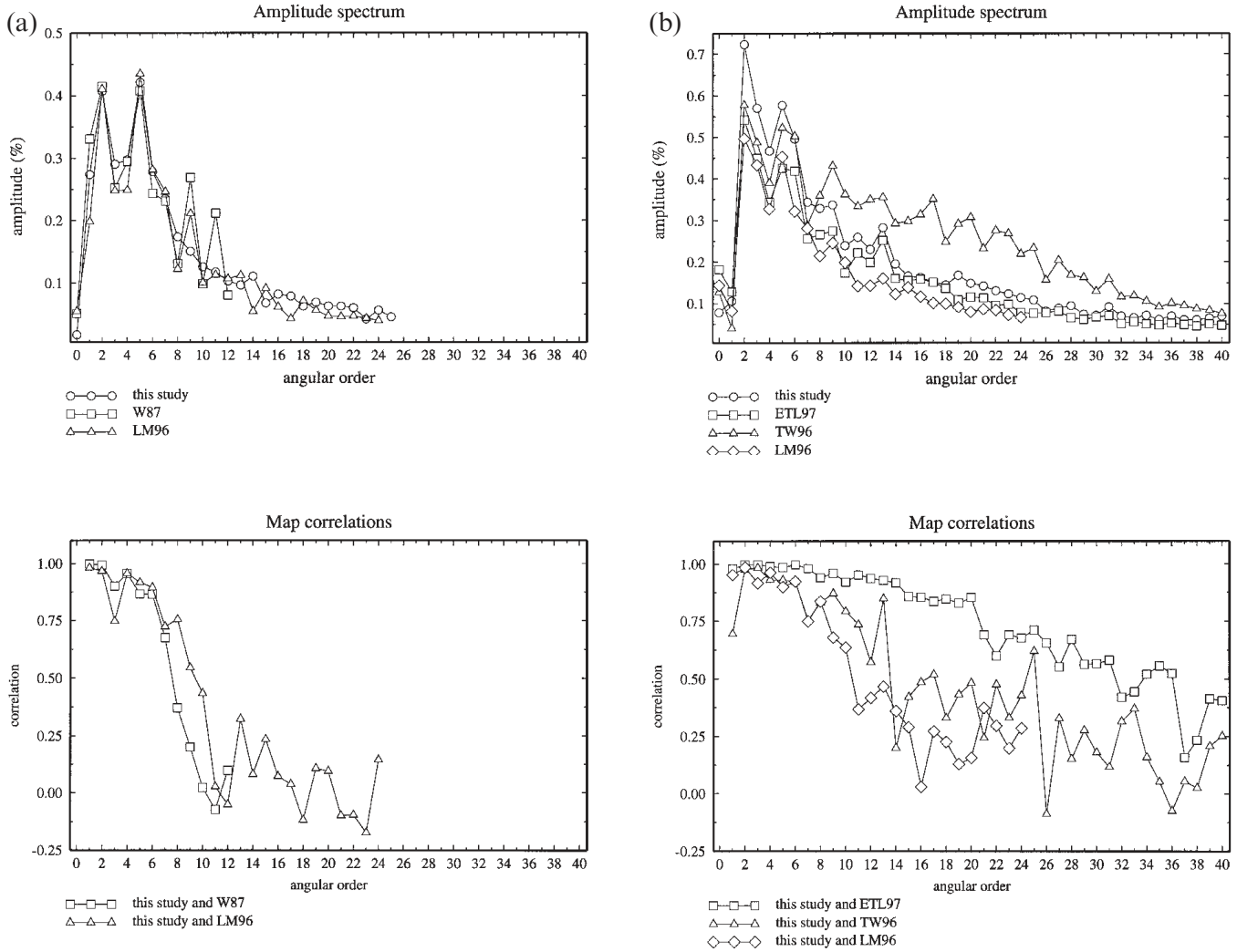

Figure 9. (a) Top: comparison of the amplitude spectrum of our fundamental-mode Love wave phase velocity map at 275 s with those of W89 and LM96. Bottom: Correlations between our work and the other two studies calculated for individual spherical harmonic degrees. (b) As (a), but for the Love fundamental mode at 75 s. Comparison is with ETL97, TW96 and LM96. Note the extremely high correlations with ETL97 and the almost flat amplitude spectrum of our study at high angular orders. (c) As (b), but for the Rayleigh fundamental mode at 150 s. (d) As (b), but for the Rayleigh fundamental mode at 40 s.

increases significantly when we model fewer unknowns. For a model with 585 effective unknowns instead of 975, for instance, our amplitude spectrum is very similar to that of ETL97. Moreover, by increasing the damping the overall correlation with ETL97 increases from 85 to 91 per cent.

It may be noted that the spectral amplitudes of TW96 are significantly higher than ours between degrees 10 and 30. This is due to the different damping strategy used in TW96: TW96 minimizes the second spatial derivative of the phase velocity distribution. This imposes a stronger spectral fall-off than a first-derivative damping as used in this study and in ETL97, as seen in Figs 9c and d.

As the large-wavelength patterns of our models agree very well with previous studies, we will not discuss them in detail (see e.g. ZL96 for an extensive discussion). Of the many interesting short-wavelength features in our maps, we want to point out two here. In the 150 s Rayleigh map there are remarkable linear features running east–west in the Pacific, west of Hawaii. These linear features may corresponds to so-called Richter rolls (Richter & Parsons 1975). Richter rolls are linear convection cells in the mantle that line up with the direction of movement of the overlying oceanic lithosphere. The presence of such rolls in the Pacific has previously been suggested by Katzman *et al.* (1998). The other interesting small-scale feature that we would like to point out is the low-velocity structure in the middle of the Pacific south of Hawaii. This structure is also observed in ETL97, although at lower amplitude. There is no obvious interpretation of this structure, although it is probably related to very low S_v velocities observed in the Pacific around 150 km depth (Ekström & Dziewonski 1998).

For the longer-period fundamental-mode Love and Rayleigh waves, our results agree well with previous studies, especially with W89. At shorter periods, the very good agreement between our results and those of TW96 and especially ETL97 are very encouraging given the differences in the phase velocity measurement techniques. One should bear in mind, however, that ETL97 and this study use the same type of data

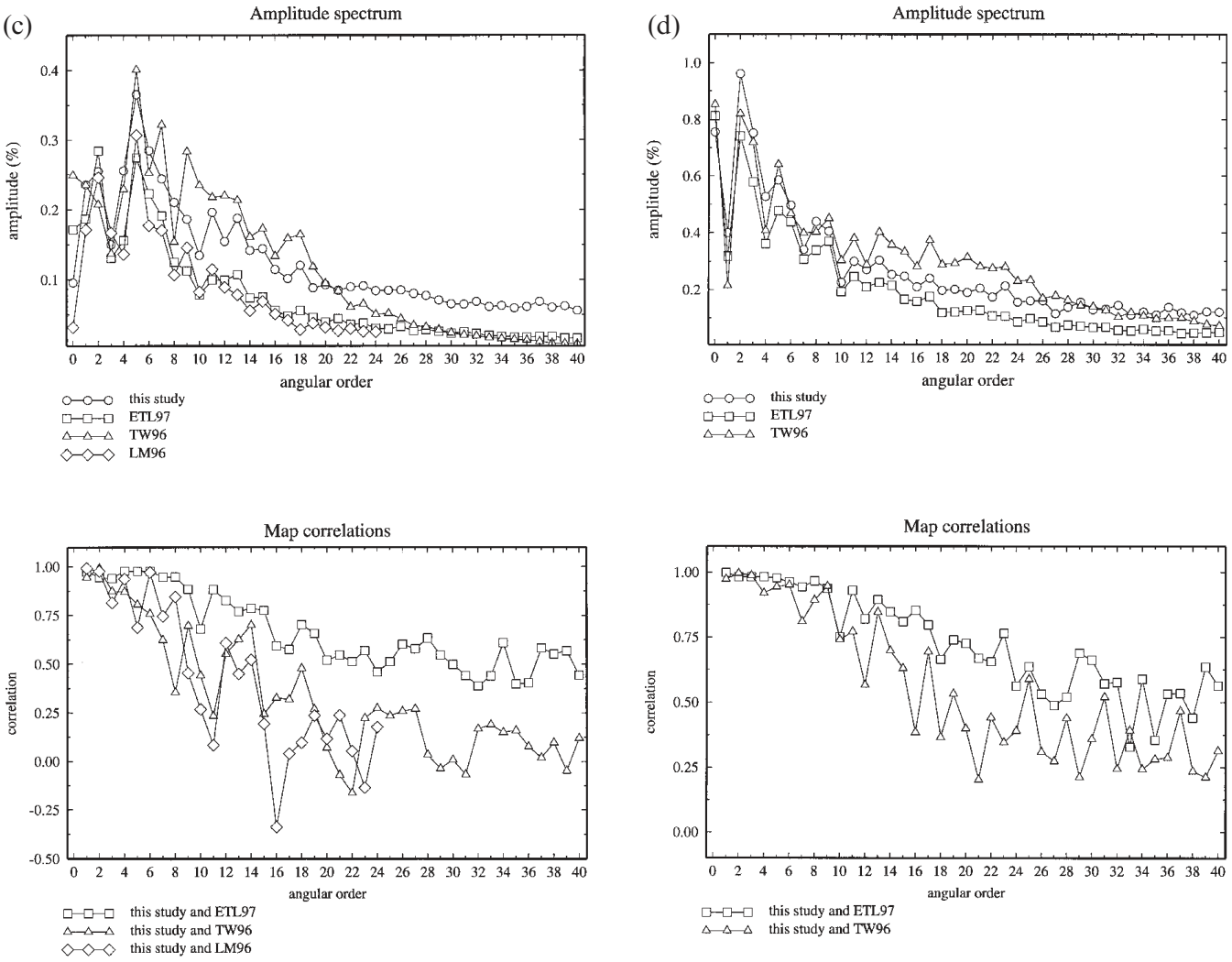

Figure 9. (*Continued.*)

(only minor-arc phase velocity measurements) and method of analysis (great-circle approximation, first-derivative damping). Therefore, the studies will be sensitive to similar modelling bias and errors.

It may be noted that the variance reductions that we achieve before rejecting outliers (ETL97 does not present variance reductions after rejecting outliers) are slightly lower than those reported in ETL97. There are two possible explanations for this. First, we use a greater number of frequency nodes to parametrize the phase velocity perturbations as a function of frequency than ETL97. This will, following a standard variance versus resolution trade-off argument, increase the uncertainty of our phase measurement. Second, and probably more importantly, ETL97 uses an aspherical starting model. This helps to stabilize the measurement process as it prevents 2π cycle skips. Moreover, the variance reductions in ETL97 are reported with respect to PREM. This means that the signal introduced by the aspherical reference model is included in the variance calculation. By definition, this signal can be accounted for exactly in the modelling and is not representative of the measurement process.

A final note on the long-period phase velocity maps concerns the applicability of the great-circle path approximation. The wavelength of a 150 s fundamental-mode Rayleigh wave is of the order of 650 km, which is only half the wavelength of a degree 40 spherical harmonic. Therefore, strictly speaking, ray theory is not applicable. The failure of ray theory when the wavelength of waves approaches the wavelength of heterogeneity is caused by it not predicting wavefield smoothing effects correctly; ray theory will produce erratic ray behaviour and it will predict phase and especially amplitude anomalies (Wang & Dahlen 1995) that are larger than observed.

Although our models have more power at short wavelengths than many previous studies, most of the power is still contained in the longer wavelengths. In a synthetic experiment using a model with a similar spectral structure to ours, Wang & Dahlen (1995) show that the phase, unlike the amplitude, of a 150 s minor-arc Rayleigh wave is well predicted by WKBJ theory. It is not directly obvious that this result can be extrapolated to our results as we use the great-circle approximation instead of WKBJ theory. However, the great-circle approximation yields models that are very similar to models derived with WKBJ theory, as discussed in Section 4. Moreover, we speculate that the great-circle approximation may still be useful when ray theory fails: when ray theory predicts unphysical ray behaviour, the great-circle approximation

will be sensitive only to the average properties of the medium. This is probably more accurate than ray theory, especially when the wavelength of heterogeneity is significantly smaller than the wavelength of the wave. This conjecture can and should be investigated in more detail using accurate synthetic seismograms for a 3-D earth model.

9 OVERTONE PHASE VELOCITY MAPS

In this section, we present the results for our overtone phase velocity measurements. As for the fundamental modes, we select reliable data using the reliability estimate. In the first column of Fig. 10(a), we show the number of reliable overtone phase velocity measurements for the vertical, longitudinal and transverse components selected using a 3nrtb of 0.5. The number of reliable measurements for each mode branch is strongly dependent on the period range considered. Moreover, for Rayleigh waves the number of reliable measurements is strongly dependent on whether the longitudinal- or vertical-component traces are analysed. Especially for the high overtones, the longitudinal-component seismograms render reliable data at lower frequencies than the vertical-component seismograms. Therefore, longitudinal-component traces provide valuable additional constraints on the structure of the mantle.

In the second column we show the variance reductions in terms of total phase anomalies achieved for these data with the damped degree-12 models before rejecting outliers. For the low reliability threshold of 0.5, good variance reductions are achieved, although lower than for the fundamental mode at similar periods. There are a few reasons that the variance reductions for the overtones are lower than for the fundamental mode. First, measuring overtone phase velocities is a noisier process than measuring fundamental-mode phase velocities due to the interference of overtone mode branches. Second, wrongly estimated source parameters have a larger effect on overtone phase velocity estimates than on fundamental-mode phase velocities. At the same period, overtones are sensitive to deeper and hence probably smaller velocity anomalies than the fundamental mode. Therefore, the phase perturbations due to anomalous structure are probably smaller for overtones than for the fundamental mode, whereas the phase perturbations due to mislocation and orientation of the source estimate are of similar size for the overtones and the fundamental mode. Moreover, at the same period overtones have higher phase velocities than the fundamental mode. Therefore, the total phase at the station will be smaller for the overtones, so that the relative influence of initial phase errors is larger.

In Fig. 10(b), we show the number of data and the variance reductions achieved with the smooth degree-12 maps for the overtones measured from the vertical component selected with a 3nrtb of 1. The doubling of the reliability threshold reduces the number of accepted measurements by approximately one-third. The variance reductions are significantly higher, which is consistent with the selection of higher-quality measurements, as discussed in Section 7.

The variance reductions, both before and after rejecting outliers, increase with increasing frequency only up to a given point. This is especially evident for the vertical-component measurements. We attribute the decrease in variance reduction to the stronger interference of neighbouring mode branches at shorter periods. Comparison of the variance reductions in Fig. 10 and the group velocities calculated for PREM (Fig. 11) shows that this may indeed be the case. At periods shorter than 50 s, for instance, the interference of the first, second and third overtone Rayleigh waves may be strong if two or more of the three branches are excited significantly. If this is the case, measurements are likely to be noisier and lower variance reductions will be obtained. [Note that the group velocities plotted in Fig. 11(a) are for reordered Rayleigh mode branches (VHW97, paragraph 2), omitting core and Stoneley modes.]

The interference of different mode branches is especially strong for Love waves. Fig. 11(b) shows that the Love wave overtone group velocities are almost identical to the group velocity of the fundamental mode in a wide period range. This makes the quality of the measurements, even at high reliabilities, questionable in those period ranges. At longer periods, however, the Love wave overtone group velocities are different enough from the fundamental-mode group velocity so that MBS should yield accurate measurements. Unfortunately, only a small number of reliable measurements can be made at these periods for most modes, as is evident from Fig. 10(a). This means that the use of the Love wave overtone phase velocity measurements from minor-arc waveforms as constraints on Earth structure is restricted to relatively few measurements in limited frequency ranges.

9.1 High-resolution overtone phase velocity maps

After removing outliers, we construct high-resolution phase velocity maps up to angular order 40 for Rayleigh wave overtones and angular order 25 for Love wave overtones. For the maps presented here, we use data selected with a reliability threshold of 0.5 and a 3 nrtb. In Fig. 12, we show four representative phase velocity maps: the first Rayleigh overtone at 150 s, the second Rayleigh overtone at 100 s, the fourth Rayleigh overtone at 78 s and the third Love wave overtone at 70 seconds. Note that we do not include the spherical average phase velocity perturbation in these maps.

We calculate the overtone phase velocity maps in exactly the same way as we calculate the fundamental-mode phase velocity maps. We choose the damping so that we model 75 per cent of the number of effective unknowns suggested by cross-validation. The numbers of effective unknowns, the number of phase measurements used and the variance reductions for the four maps in Fig. 12 are listed in Tables 4 (Rayleigh waves) and 5 (Love waves). For reference, we include the same information for several modes that are not shown in Fig. 12.

In Fig. 13, we plot sensitivity kernels for the maps in Fig. 12 with respect to perturbations in density and isotropic P and S velocities calculated for PREM. The sensitivities are the derivatives of the eigenfrequency with respect to $\delta\rho/\rho$, $\delta V_p/V_p$ and $\delta V_s/V_s$ for the eigenoscillation with the closest frequency to the nodal frequencies at which the maps in Fig. 12 are plotted. Note the increasing complexity of the kernels with increasing overtone number. This makes it difficult to relate phase velocity perturbations to seismic anomalies in a certain depth range for the high overtones.

The four modes plotted here have significant sensitivity to changes in the S velocity in the lower mantle. The sensitivity to changes in the P velocity is small for the first overtone at 150 s and the second overtone at 100 s. For the fourth overtone, however, P-velocity anomalies may contribute significantly

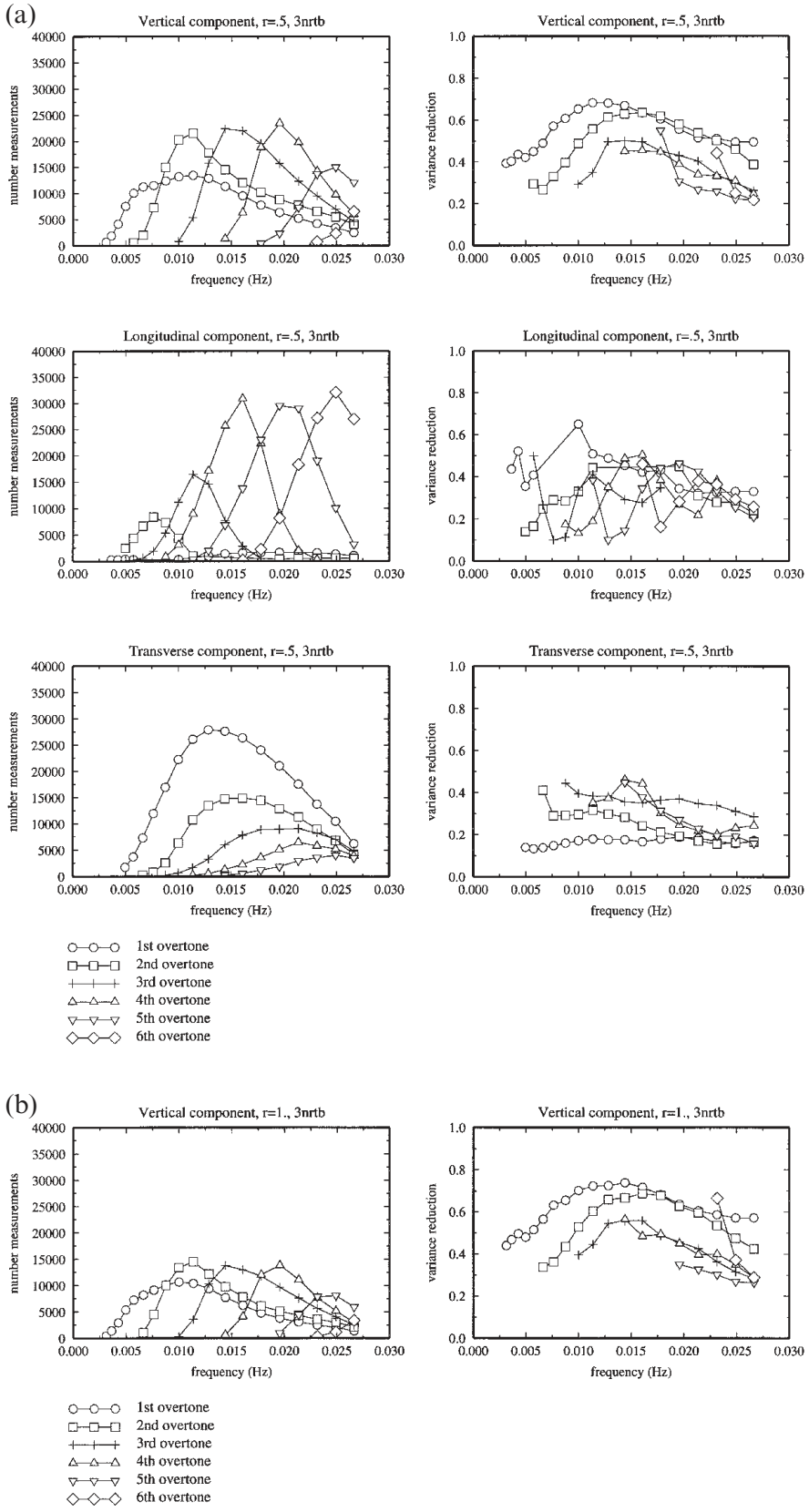

Figure 10. (a) Top left: Number of vertical component Rayleigh wave phase velocity measurements selected for the first six overtone mode branches using a three-node reliability threshold of 0.5. Top right: Variance reduction for the selected data obtained with a smooth degree-12 phase velocity map. Middle: Same as above, but for the longitudinal component. Bottom: Same, but for Love waves. (b) Same as (a) (Top), but for data selected using a 3nrtb of 1. The higher reliability threshold increases the variance reductions significantly but at the same time it reduces the number of accepted measurements.

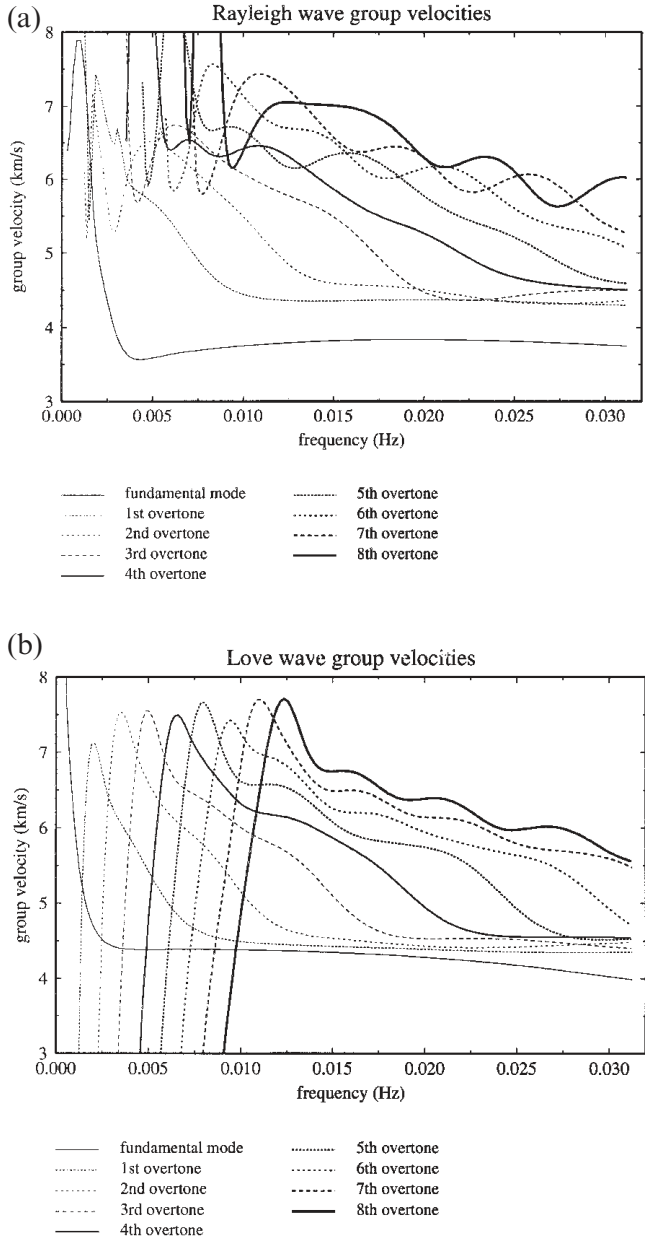

Figure 11. (a) Rayleigh wave group velocities calculated for PREM. The group velocities are for reordered mode branches, omitting core and Stonely modes. (b) Love wave group velocities calculated for PREM.

to the observed phase velocity perturbations as the strongest velocity anomalies are probably found in the top 250 km of the Earth (see e.g. Su *et al.* 1994). Assuming that *P*-velocity anomalies are approximately half the size of *S*-velocity anomalies, up to a third of the phase velocity anomalies observed for this mode may be attributed to changes in *P* velocity rather than changes in *S* velocity. A similar strong sensitivity to *P* velocity is usually observed for the modes that have high amplitudes on the longitudinal component.

The sensitivity of the Love wave third overtone at 70 s is typical for Love wave overtones. Love wave overtones usually have a strong peak in the *S*-velocity sensitivity around the greatest depth that they sample. This peak shifts to shallower depths for shorter periods.

Table 4. Number of measurements, number of effective unknowns modelled and total phase variance reduction for the Rayleigh wave overtone high resolution ($l=40$) phase velocity maps. All data are selected using a 3nrtb of 0.5 and outliers have been rejected.

branch	period (s)	nr of data	nr unkn. mod.	var. red. (%)
1S	200	6998	287	61
1S	150	10524	406	69
1S	100	12404	430	83
1S	70	11989	450	84
2S	150	8059	183	35
2S	100	22749	541	72
2S	70	13684	481	80
3S	100	10415	389	53
3S	70	28218	637	65
3S	56	18974	590	71
4S	78	16110	424	52
4S	56	39028	792	64
4S	43	13742	473	61
5S	62	12978	446	54
5S	47	34041	779	70
6S	51	7634	323	46
6S	40	31589	742	59

Table 5. As Table 4, but for Love waves. The Love wave overtone maps are expanded up to degree 25 only as relatively few reliable measurements are available.

branch	period (s)	nr of data	nr unkn. mod.	var. red. (%)
2T	100	5963	171	46
3T	70	5596	184	52
4T	62	2107	134	65

The large-wavelength anomalies observed in the Rayleigh wave phase velocity maps in Fig. 12 can be interpreted in terms of relatively shallow velocity anomalies (< 300 km). All three maps have a strong contribution from the ocean–continent function. The first and second overtone modes sample the deeper parts of the continents where continents are seismically faster than oceans. For the fourth overtone at 78 s in Fig. 12(c), however, the continents are slower than the oceans. At this period, the fourth overtone has a strong sensitivity to very shallow structure. Therefore, the low continental phase velocities are related to high topography and the thick crust on continents.

Most of the high velocities in Fig. 12(a) are related to the cratonic areas of the continents. In the regions of Borneo, the Himalayas and also western South America, however, the high velocities are probably associated with the subduction of cold oceanic lithosphere. Among the other interesting features observed in Fig. 12(a) are a possible signal of the Walvis ridge, an old hotspot track (Vandecar *et al.* 1995), low velocities around the East African Rift, and linear features in the West Pacific that are similar to those found for the fundamental mode at 150 s, as discussed in Section 8. These linear features are even more prominent for the first overtone at 100 s (not shown here). Importantly, the highs and lows of the linear structures observed in the first overtone phase velocity maps correlate well with those observed in our fundamental-mode maps.

Owing to the very oscillatory sensitivity kernels of the fourth overtone, Fig. 12(c) is much harder to interpret than Figs 12(a)

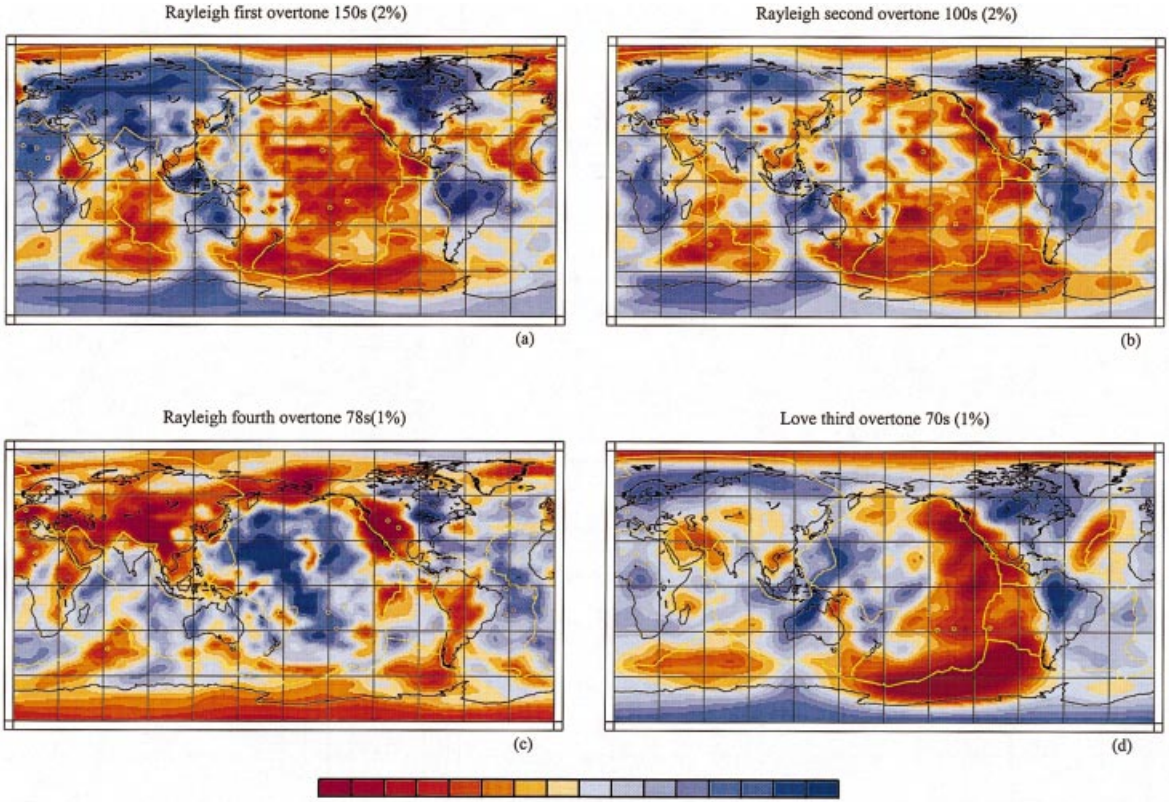

Figure 12. Overtone phase velocity maps for the four modes indicated. The number in brackets at the top of each figure corresponds to the maximum and minimum used in the colourscale that is shown at the bottom of the plate. See text for further details.

and (b). Nevertheless, some familiar structures are observed: low velocities around the Himalayas and the Tibetan Plateau, low velocities at the East Pacific Rise, very low velocities in the western Unites States and fast old oceanic lithosphere in the western Pacific.

It is interesting to compare the three Rayleigh wave maps together with their kernels around Hawaii. Both Figs 12(b) and (c) show anomalously low velocities around Hawaii, whereas the phase velocities around Hawaii do not appear to be different from the surrounding ocean in Fig. 12(a). The kernels in Fig. 13 show that there are two possible explanations for this:

(1) There is a region of anomalously low velocities in the transition zone under Hawaii. The modes shown in Figs 12(b) and (c) are especially sensitive to velocity anomalies between 400 and 600 km depth, whereas Fig. 12(a) has a reduced sensitivity in that depth range. Hence, this could account for the observations.

(2) The region around Hawaii is anomalously slow at very shallow depths. The first overtone has very little sensitivity to very shallow structure compared to the other two modes. This suggests a possible shallow origin for the observed difference.

Of course, it should be possible to establish which of these two explanations is the best using joint 3-D modelling of fundamental and overtone phase velocities.

The Love wave overtone shown in Fig. 12(d) samples the upper mantle evenly. We find high velocities associated with cratonic areas and low velocities associated with oceans. The high velocities in the western Pacific are probably related to subduction of oceanic lithosphere. Note that the resolution in Fig. 12(d) is considerably lower than in the Rayleigh wave phase velocity maps in Fig. 12. This is due to the smaller number of reliable measurements and the lower quality of the Love wave overtone phase measurements.

9.2 Comparison with Stutzmann & Montagner (1994)

We compare our overtone phase velocity maps with the first overtone phase velocity distributions derived by Stutzmann & Montagner (1994) (SM94). SM94 use an iterative waveform fitting technique (Stutzmann & Montagner 1993) that is hybrid in the sense that in each step phase velocity measurements are made that are used to construct a path-averaged velocity model. The intermediate modelling of depth-dependent path-averaged velocity anomalies is used to stabilize the inversion process.

At very long wavelengths, our results are broadly consistent with the phase velocity maps of SM94. In Table 6, we show degree-by-degree correlation between our maps and those of SM94 for the lowest four mode branches at 100 s. We compare the maps only up to angular order 4, as few paths (300) are used to constrain the maps in SM94. For these low angular orders, the overtone maps of SM94 correlate well with our maps at degree 1 and reasonably well at degree 2. The fundamental-mode map correlates well for all 4 degrees.

The correlations between our work and SM94 are encouraging. However, we feel that more detailed comparison is not warranted due to the large difference in lateral resolution between the two studies.

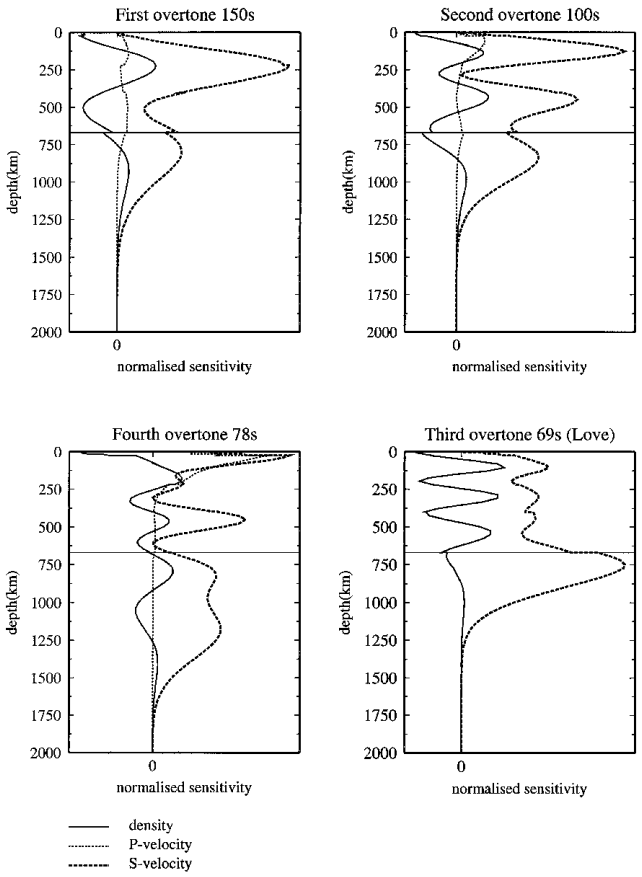

Figure 13. Normalized sensitivities to density, *S* velocity and *P* velocity for the four modes shown in Fig. 12.

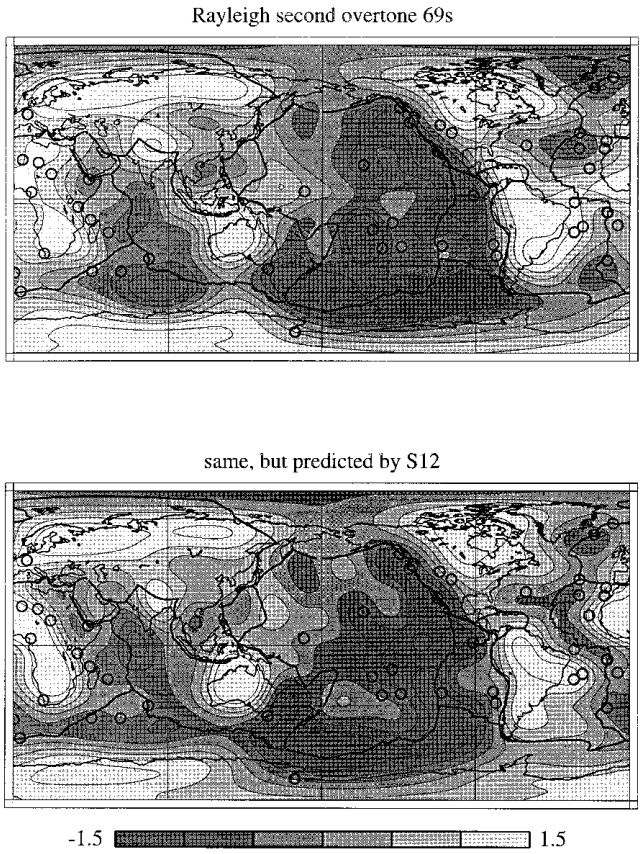

Figure 14. Comparison of our phase velocity map for the Rayleigh wave second overtone at 69 s with the phase velocity distribution for the same mode predicted by the whole-mantle model S12. For comparison, our high-resolution map is truncated at degree 12. The spherical averages are omitted.

9.3 Comparison with model predictions

In order to check our overtone measurements against previously established results in more detail, we compare our overtone phase velocity distributions with phase velocity maps predicted by 3-D tomographic models of the mantle.

In Fig. 14, we plot the phase velocity distribution for the second Rayleigh wave overtone at 70 s measured by us and the phase velocity distribution predicted by the mantle shear velocity model S12 (Su *et al.* 1994). We show the maps up to degree 12 and we omit degree zero. The observed and predicted phase velocity distributions are very similar in both pattern and amplitude. In our phase velocity map, however, the amplitudes are lower than those predicted by S12 in areas of poor data coverage such as South America and Africa. In these areas our maps are significantly more damped than in areas of good coverage such as the northern Pacific.

In the upper panel of Fig. 15, we show the amplitude spectrum of our phase velocity map for the first overtone at 70 s together with the amplitude spectrum of the phase velocity distributions predicted by three different models of shear velocity anomalies in the mantle: M84c (Woodhouse & Dziewonski 1984), S12 (Su *et al.* 1994) and S16 (Woodhouse & Trampert 1995). The amplitude spectrum of our map agrees well with the spectra predicted by the 3-D models. Note that the spectrum of our map is nearly flat beyond degree 12. This is typical for our overtone phase velocity maps.

In the lower panel of Fig. 15, we show the degree-by-degree correlations between our map and the predicted maps. The correlations between the measured phase velocity map and the predicted maps are very high up to degree 6. For higher degrees the correlation decreases with increasing angular order. We observe similar correlations between our overtone phase velocity maps and maps predicted by S12 and S16 for all Rayleigh wave overtone phase velocity maps.

In Fig. 16, we compare our phase velocity map for the fourth Love wave overtone at 62 s to the phase velocity distribution predicted by the *SH*-velocity whole-mantle model U84L85/SH (Woodhouse & Dziewonski 1989). At this period, the fourth overtone is outside the frequency range where the Love wave group velocities overlap (see Fig. 11(b)) and the variance

Table 6. Degree-by-degree correlations between our phase velocity maps and the phase velocity maps of Stutzmann & Montagner (1994) for the lowest four Rayleigh wave mode branches at 100 s period. Correlations are only presented up to degree 4, because of the relatively low lateral resolution in Stutzmann & Montagner (1994).

degree	fund. mode	first ovt.	2nd ovt.	3rd ovt.
1	.92	.95	.56	.17
2	.95	.53	.92	.86
3	.74	.13	.33	.39
4	.66	.66	−.37	.67

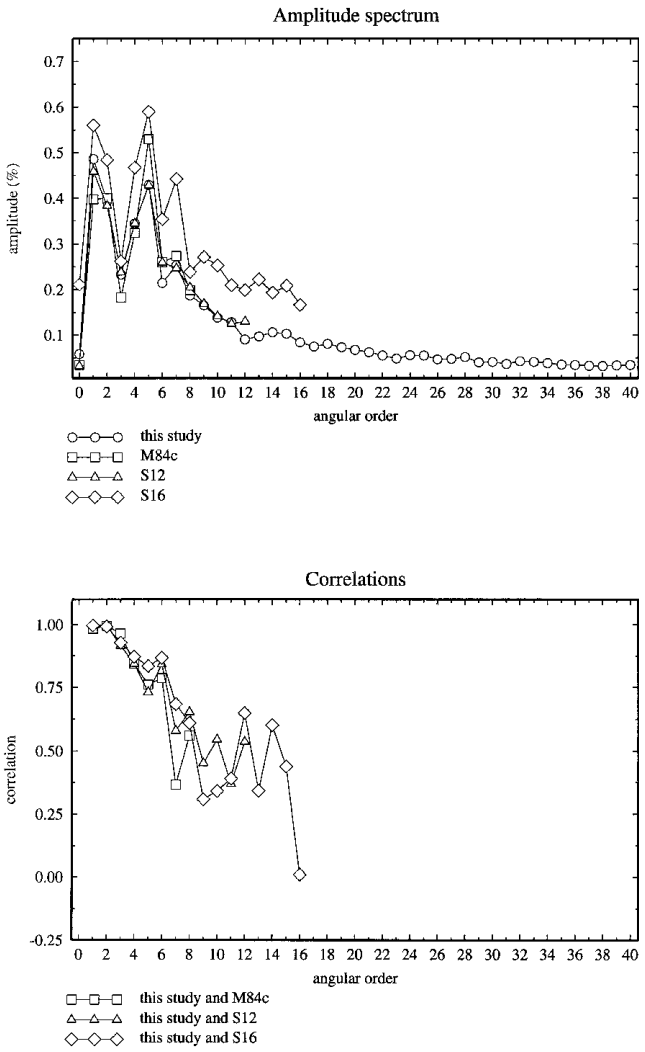

Figure 15. Top: Comparison of the amplitude spectrum of our phase velocity map for the second overtone at 69 s shown in Fig. 14 with phase velocity maps predicted by three mantle models: M84c (Woodhouse & Dziewonski 1984), S12, (Su *et al.* 1994) and S16 (Woodhouse & Trampert 1995). Bottom: Correlations of our phase velocity map with the maps predicted by the 3-D mantle models listed above.

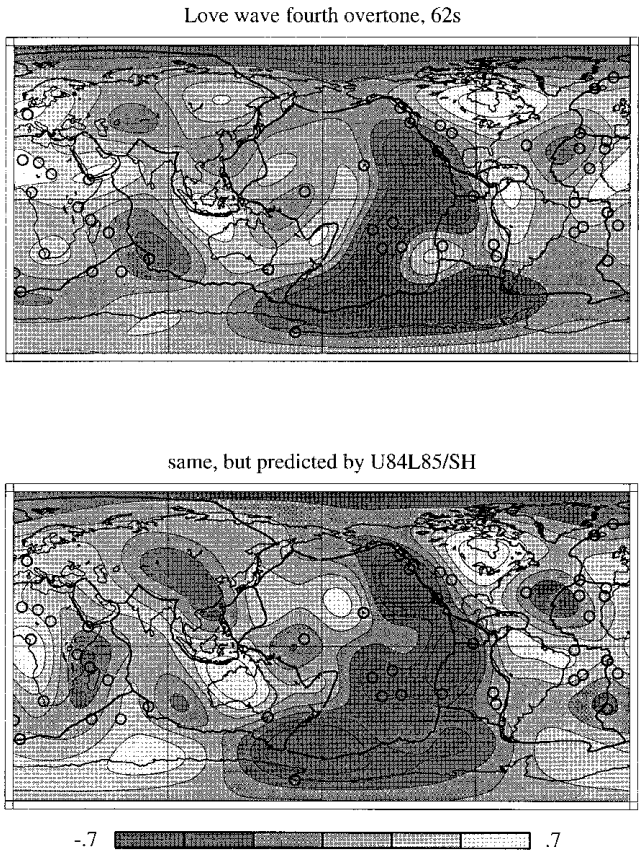

Figure 16. Comparison of phase velocities for the Love wave fourth overtone at 62 s measured in this study and those predicted by the whole-mantle (pure *SH* velocity) model U84L85/SH (Woodhouse & Dziewonski 1989). As U84L85/SH is a degree 8 model, we plot our high-resolution map up to degree 8. We omit the spherical average.

reduction achieved is high (65 per cent, after rejecting outliers) compared to the variance reductions achieved for most Love wave overtone modes.

The overall correlation of the two maps in Fig. 16 is 71 per cent. The amplitudes of our map, however, are somewhat lower than those predicted by U84L85/SH. This is also evident from the amplitude spectra shown in Fig. 17. In Fig. 17, we also show the amplitude spectrum of the phase velocity distribution predicted by S12. Interestingly, the spectral amplitudes predicted by S12 are significantly higher than those in our phase velocity map and those predicted by U84L85/SH. This may, of course, be due to the strong damping that we apply; for this map we model only 134 effective unknowns using few measurements (2107). It could, however, also be due to a more profound difference between the anomalies sensed by Love and Rayleigh waves, as S12 is mainly constrained by waves of *SV*-type.

As noted before, at shorter periods the interference of the Love wave mode branches increases. At these periods, our phase velocity maps have significantly smaller amplitudes than those predicted by the whole-mantle models, although

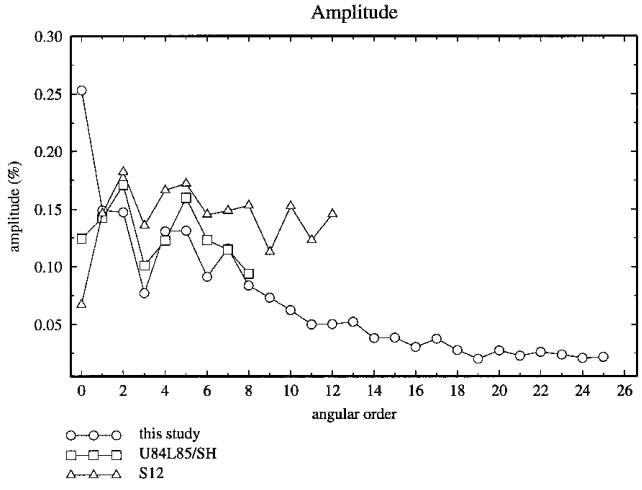

Figure 17. Comparison of the amplitude spectra of the fourth Love wave overtone phase velocity maps in Fig. 12 with the amplitude spectrum predicted by S12 and U84L85/SH.

the observed velocity patterns are similar. This behaviour is representative of our Love wave overtone phase measurements at frequencies where cross-branch interference is strong.

9.4 Comparison of fundamental mode and overtone phase velocity maps

As a final test of our Rayleigh wave overtone phase velocity measurements, we compare the fundamental-mode Rayleigh wave phase velocity map at 40 s with the phase velocity distribution determined for the fourth Rayleigh wave overtone at 62 s. We choose these two modes because they are sensitive to the structure of the top 150 km of the upper mantle in a very similar way, as shown in Fig. 18.

In Fig. 19, we show the phase velocity perturbations for both maps up to degree 12. Approximately the same number of data have been used for both maps in Fig. 19 and the damping is chosen to be as similar as possible for the two maps.

In order to be able to compare the amplitudes of the maps, the amplitude of the fourth overtone has been multiplied by a factor of 3.3. This factor accounts for the lower sensitivity of the fourth overtone in the top 200 km of the upper mantle. After this multiplication, an *S*-velocity anomaly in the top of the upper mantle will give rise to approximately the same perturbation in both maps. Therefore, the amplitudes in Fig. 19 can readily be compared.

The phase velocity distributions in Fig. 19 are very similar: low velocities associated with topography and tectonically active regions and high velocities associated with old oceanic lithosphere. In most places the amplitudes agree well but there are some interesting differences. The difference in amplitude in the region between Indonesia and Fiji, for instance, is remarkable. In this area the fourth overtone is considerably faster than the fundamental mode. This suggests that material that is faster than average must be present at greater depths. This can probably be attributed to the presence of subducting slabs in this region.

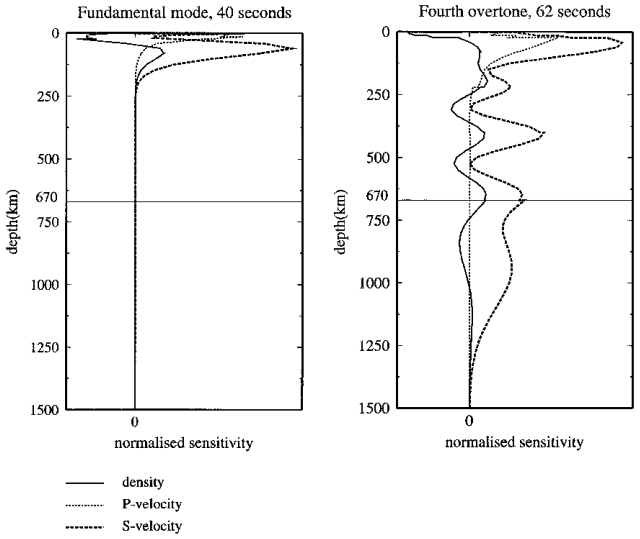

Figure 18. Normalized sensitivity to density, *S* velocity and *P* velocity for the modes shown in Fig. 19. Note that the two modes are similarly sensitive to velocity and density anomalies in the top 150 km of the mantle.

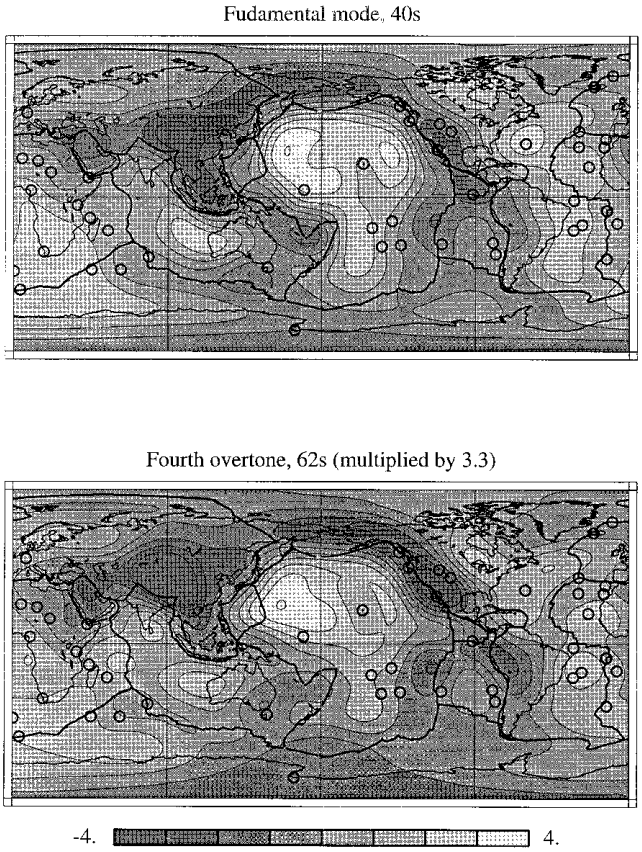

Figure 19. Comparison of the phase velocity distributions determined for the Rayleigh wave fundamental mode at 40 s and the fourth Rayleigh wave overtone at 62 s as determined in this study. (See Fig. 18 for the sensitivities of these modes.) The maps have very similar data coverage and similar numbers of unknowns are resolved in both maps. The amplitudes in the bottom map are multiplied by a factor of 3.3 to allow comparison of the amplitudes. To emphasize the similarity of the maps at long wavelengths, we truncate the maps at degree 12. See text for details.

In Fig. 20, we show the amplitude spectra for the two maps and the correlation between the two maps up to degree 36. For comparison, the amplitude spectrum for the fourth overtone has been multiplied by the same factor as the map in Fig. 19.

The amplitude spectra are very similar except at degree zero. This means either that higher spherical average velocities somewhere in the top 150 km are compensated by lower average velocities in the sensitivity range of the fourth overtone, or that the average velocities in the Earth are only significantly higher than PREM's in the top 50 km where the sensitivities of the two modes are somewhat different.

The correlations between the maps are very high, especially for the lower degrees. The correlations decrease rapidly with increasing angular order, which may, of course, reflect the fact that the high angular orders are more difficult to resolve and hence noisier. An alternative explanation is that the low angular orders are so well correlated because of the dominance of low-order spherical harmonic structure in the top of the upper mantle. If this dominance is restricted to relatively shallow depths sampled by both modes, the fourth overtone signal at low angular orders will be dominated by shallow structure giving rise to high correlations between the two modes. At greater depths the spectrum of heterogeneity may

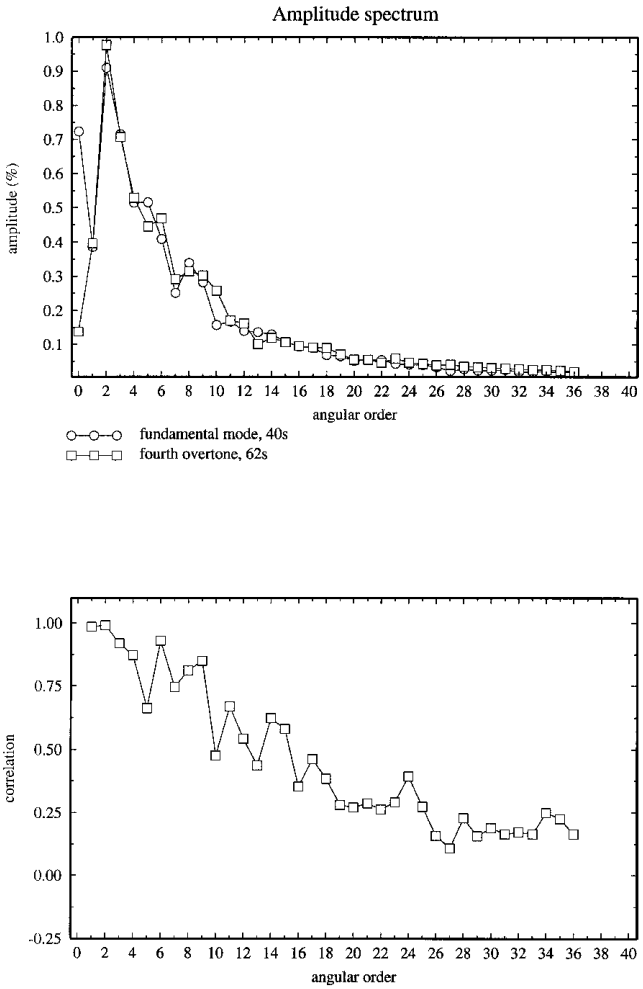

Figure 20. Amplitude spectra and correlations for the maps in Fig. 19.

be whiter and hence the contribution of deeper structure to the overtone map relatively stronger at short wavelengths. Assuming that the seismic anomalies at greater depths are uncorrelated with the seismic anomalies at shallow depths this would naturally decrease the correlations between the two maps at high angular orders. This would be consistent with our earlier observation that the spectrum for a 150 s fundamental-mode Rayleigh wave is significantly whiter than that at 40 s. Interestingly, the spectrum of heterogeneity as a function of depth for the mantle model S12 also shows a whitening of the spectrum with increasing depth (Su *et al.* 1994; Dziewonski 1996).

10 CONCLUSIONS

In this paper, we have shown that the mode branch stripping technique is a powerful method for measuring Rayleigh wave fundamental-mode and overtone phase velocities and fundamental-mode Love wave phase velocities. The results for the Love wave and Rayleigh wave fundamental modes agree very well with previous studies. Even at periods as long as 275 s, our measurements from minor arcs are compatible with earlier work that is largely constrained by major arcs and multiple orbits.

Our Rayleigh wave overtone phase velocity measurements are well explained by phase velocity maps. These maps are consistent with phase velocity distributions predicted by existing mantle models and phase velocity distributions obtained for the fundamental mode. Importantly, we have found similar structures in Rayleigh wave overtone phase velocity maps that are measured independently.

Interestingly, we have observed that longitudinal-component waveforms contribute valuable Rayleigh wave phase information for the high overtones that cannot be retrieved from vertical-component seismograms.

It is difficult to assess the reliability of our Love wave overtone phase measurements. They may not be trustworthy in some frequency ranges due to the strong interference of several mode branches including the fundamental mode. This may even be true for measurements that are associated with a high reliability as determined using our reliability estimate. We expect that 3-D modelling of Love wave fundamental mode and overtone dispersion will allow us to address this problem in more detail.

Finally, mode branch stripping method is well suited for automatic application to bulk data: the 110 000 three-component seismograms were processed in less than two months using a network of workstations.

The data set described in this paper consists of well over 500 000 individual reliable phase velocity measurements. Clearly, this data set provides extremely powerful constraints on the 3-D structure of the Earth. This should enable us to resolve the global structure of the transition zone with unprecedented accuracy.

ACKNOWLEDGMENTS

We would like to thank the operators of the GEOSCOPE, IRIS and USGS networks for providing high-quality waveform data. We also thank Goran Ekström and an anonymous reviewer for their constructive criticisms.

Spherical harmonic coefficients for the phase velocity maps presented in Figs 8 and 12 are available via ftp at ftp.gps.caltech.edu, in the directory /pub/hendrik.

REFERENCES

Cara, M., 1978. Regional variations of Rayleigh-mode velocities: a spatial filtering method, *Geophys. J. R. astr. Soc.,* **54,** 439–460.

Cara, M., 1979. Lateral variations of S-velocity in the upper mantle from higher Rayleigh modes, *Geophys. J. R. astr. Soc.,* **57,** 649–670.

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

Debayle, E. & Lévêque, J.J., 1997. Upper mantle heterogeneities in the Indian Ocean from waveform inversion, *Geophys. Res. Lett.,* **24,** 245–248.

Der, Z., Masse, R. & Landisman, M., 1970. Effects of observational errors on the resolution of surface waves at intermediate distances, *J. geophys. Res.,* **75,** 3399–3409.

Dziewonski, A.M., 1996. Earth's mantle in three dimensions, in *Seismic Modelling of Earth Structure,* pp. 507–572, eds Boschi, E., Ekström, G. & Morelli, A., Editrice Compositori, Bologna.

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

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

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

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

Hastie, T.J. & Tibshirani, R.J., 1990. *Generalized Additive Models,* Chapman & Hall, London.

Katzman, R., Zhao, L. & Jordan, T., 1998. High-resolution, 2-dimensional vertical tomography of the central pacific mantle using ScS reverberations and frequency-dependent traveltimes, *J. geophys. Res.,* **103,** 17 933–17 971.

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

Lerner-Lam, A.L. & Jordan, T.H., 1983. Earth structure from fundamental and higher-mode waveform analysis, *Geophys. J. R. astr. Soc.,* **75,** 759–797.

Lerner-Lam, A.L. & Jordan, T.H., 1987. How thick are the continents, *J. geophys. Res.,* **92,** 14 007–14 026.

Li, X.D. & Romanowicz, B., 1995. Comparison of global waveform inversions with and without considering cross-branch modal coupling, *Geophys. J. Int.,* **121,** 695–709.

Li, X.D. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. geophys. Res.,* **101,** 22 245–22 272.

Li, X.D. & Tanimoto, T., 1993. Wave-forms of long-period body waves in a slightly aspherical earth model, *Geophys. J. Int.,* **112,** 92–102.

Marquering, H. & Snieder, R., 1995. Surface-wave mode-coupling for efficient forward modeling and inversion of body-wave phases, *Geophys. J. Int.,* **120,** 186–208.

Marquering, H. & Snieder, R., 1996. Shear-wave velocity structure beneath Europe, the Northeastern Atlantic and Western Asia from wave-form inversions including surface-wave mode-coupling, *Geophys. J. Int.,* **127,** 283–304.

Mitchell, R.G., 1980. Array measurements of higher mode Rayleigh wave dispersion: an approach utilising source parameters, *Geophys. J. R. astr. Soc.,* **63,** 311–331.

Nolet, G., 1975. Higher Rayleigh modes in Western Europe, *Geophys. Res. Lett.,* **2,** 60–62.

Nolet, G., 1977. The upper mantle under Western Europe inferred from the dispersion of Rayleigh modes, *J. Geophys.,* **43,** 265–285.

Nolet, G., 1990. Partitioned wave-form inversion and 2-dimensional structure under the network of autonomously recording seismographs, *J. geophys. Res.,* **95,** 8499–8512.

Okal, E.A. & Jo, B.G., 1987. Stacking investigations of the dispersion of higher-order mantle Rayleigh-waves and normal-modes, *Phys. Earth planet. Inter.,* **47,** 188–204.

Pollitz, F.F., 1994. Global tomography from Rayleigh and Love wave-dispersion—effect of ray-path bending, *Geophys. J. Int.,* **118,** 730–758.

Richter, F.M. & Parsons, B., 1975. On the interaction of two scales of convection in the mantle, *J. geophys. Res.,* **80,** 2529–2541.

Robertson, G.S. & Woodhouse, J.H., 1995. Evidence for proportionality of *P* and *S* heterogeneity in the lower mantle, *Geophys. Res.,* **123,** 85–116.

Spakman, W., 1991. Delay-time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geophys. J. Int.,* **107,** 309–332.

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

Stutzmann, E. & Montagner, J.P., 1994. Tomography of the transition zone from the inversion of higher-mode surface-waves, *Phys. Earth planet. Inter.,* **86,** 99–115.

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

Tanimoto, T., 1990. Long-wavelength *S*-wave velocity structure throughout the mantle, *Geophys. J. Int.,* **100,** 327–336.

Tarantola, A. & Valette, B., 1982a. Generalized non-linear inverse problems solved using the least-squares criterion, *Rev. Geophys.,* **20,** 219–232.

Tarantola, A. & Valette, B., 1982b. Inverse problems = quest for information, *J. Geophys. Zeitschrift Fur Geophysik,* **50,** 159–170.

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

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

Vandecar, J.C., James, D.E. & Assumpcao, M., 1995. Seismic evidence for a fossil mantle plume beneath South America and implications for plate driving forces, *Nature,* **378,** 25–31.

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

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

van Heijst, H.J., Snieder, R. & Nowack, R., 1994. Resolving a low-velocity zone with surface-wave data, *Geophys. J. Int.,* **118,** 333–343.

Wang, Z. & Dahlen, F.A., 1995. Validity of surface-wave ray theory on a laterally heterogeneous earth, *Geophys. J. Int.,* **123,** 757–773.

Wang, Z., Tromp, J. & Ekström, G., 1998. Global and regional surface-wave inversions: a spherical-spline parameterization, *Geophys. Res. Lett.,* **25,** 207–210.

Widiyantoro, S. & van der Hilst, R., 1996. Structure and evolution of lithospheric slab beneath the Sunda arc, Indonesia, *Science,* **271,** 1566–1570.

Wong, Y.K., 1989. Upper mantle heterogeneity from phase and amplitude data of mantle waves, *PhD thesis,* Harvard University, Cambridge, MA.

Woodhouse, J.H. & Dziewonski, A.M., 1984. Mapping the upper mantle–3-dimensional modeling of earth structure by inversion of seismic waveforms, *J. geophys. Res.,* **89,** 5953–5986.

Woodhouse, J.H. & Dziewonski, A.M., 1986. Three dimensional mantle models based on mantle wave and long period body wave data, *EOS, Trans. Am. geophys. Un.,* **67,** 307.

Woodhouse, J.H. & Dziewonski, A.M., 1989. Seismic modeling of the earths large-scale 3-dimensional structure, *Phil. Trans. R. Soc. Lond.,* A **328,** 291–308.

Woodhouse, J.H. & Trampert, J., 1995. Global upper mantle structure inferred from surface wave and body wave data, *EOS, Trans. Am. geophys. Un.,* **76,** 422.

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

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

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