Monte Carlo search techniques applied to the measurement of higher mode phase velocities and anisotropic surface wave tomography

Karin Visser

Monte Carlo search techniques applied to the measurement of higher mode phase velocities and anisotropic surface wave tomography

Monte Carlo zoektechnieken toegepast op het meten van fasesnelheden van boventonen en tomografie van oppervlaktegolven

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR
AAN DE UNIVERSITEIT UTRECHT
OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF. DR. J.C. STOOF,
INGEVOLGE HET BESLUIT VAN HET COLLEGE VOOR PROMOTIES
IN HET OPENBAAR TE VERDEDIGEN
OP VRIJDAG 29 FEBRUARI 2008 DES MIDDAGS TE 12.45 UUR

DOOR

Karin Visser

geboren op 8 mei 1979, te Biddinghuizen

Promotoren: Prof. dr. J. Trampert
                Prof. dr. B.L.N. Kennett

The Earth:
"We thought it was a very stable place - that if you could go back to the time when the oceans and continents first separated out, they'd have been in the same places they are now And we thought that the interior of the Earth was pretty simple...."

"We were wrong...."

Terry Prachett: "The Science of Discworld"

"Nothing has to be true forever.
Just for long enough, to tell you the truth."

Terry Prachett: "The Truth"

Contents

Chapter 1

Introduction

1.1 Seismic Tomography

Seismic tomography was developed in the mid 1970s (Aki et al., 1977; Dziewonski et al., 1977; Dziewonski, 1984; Woodhouse and Dziewonski, 1984) as a way to infer information about the Earth's three-dimensional structure from seismic waves. Since then, body waves have been used to map the Earth globally and regionally (van der Hilst et al., 1997; Kennett et al., 1998; Widiyantoro et al., 1999; Bijwaard and Spakman, 2000; Masters et al., 2000; Kárason and van der Hilst, 2001; Amaru, 2007) while surface waves have mainly been used to map the crustal and upper mantle structure (Ekström et al., 1997; Ritzwoller and Levshin, 1998; Debayle and Kennett, 2000; Shapiro and Ritzwoller, 2002; Romanowicz, 2003; Trampert and Woodhouse, 2003). Over the years, tomographic maps have improved due to the increase in data and computing power as well as improvements in the theory. In the 1970s, global models were only able to resolve the largest wavelengths (∼3000-5000km, corresponding to spherical harmonic degree 6-8) while now lateral resolutions may be obtained corresponding to wavelengths of less than 1000km and in some places less than 500km. Tomography has provided many sharp images of fast velocity slabs, or subducting plates (van der Hilst et al., 1997; Widiyantoro et al., 1999; Bijwaard and Spakman, 2000), and showed that some slabs penetrate the lower mantle while others appear to lie flat (e.g. Fukao, 1992 ,van der Hilst et al, 1997). There is also evidence of plumes (regions of low velocity), for example underneath Iceland and Hawaii (Bijwaard and Spakman, 1999; Zhao, 2001; Montelli et al., 2004). Seismic tomography has also been extended to anisotropy (Tanimoto and Anderson, 1985; Montagner and Tanimoto, 1990; Ekström and Dziewonski, 1998; Trampert and van Heijst, 2002;

Panning and Romanowicz, 2006), attenuation (Durek et al., 1993; Romanowicz, 1995; Selby and Woodhouse, 2002) and density (Ishii and Tromp, 1999, 2001). Furthermore, tomographic images have been used to obtain information about the temperature and composition of the Earth (van der Hilst and Kárason, 1999; Deschamps et al., 2002; Godey et al., 2004).

1.2   Limitations of seismic tomography

The two types of tomography, body wave and surface wave tomography, are more or less complementary. This is because they see different parts of the Earth at different scales of resolution. Body wave tomography lacks resolution in the oceanic parts of the Earth due to the uneven geographical coverage of sources and stations. It is however well suited to study subducting slabs in great detail (due to the many sources in the subducting slab). Surface wave tomography, on the other hand, has good coverage in oceanic areas but suffers from limited depth resolution due to the use of mainly fundamental mode surface waves.

Another issue is anisotropy. Most tomography assumes an isotropic Earth for theoretical and computational convenience. By now however, there is considerable evidence for both radial and azimuthal anisotropy at a variety of depths in the Earth. Radial anisotropy, the discrepancy between Rayleigh and Love waves, was first observed by Anderson (1961); Aki and Kaminuma (1963); McEvilly (1964) and was later included in the upper 200km of the global reference Earth model PREM (Dziewonski and Anderson, 1981). Azimuthal anisotropy, the azimuthal variation of phase velocity, was first observed by Forsyth (1975). Radial and azimuthal anisotropy are different expressions of the underlying general anisotropy of the Earth's interior. The source of anisotropy in the mantle is usually assumed to be the alignment (lattice preferred orientation or LPO) of intrinsically anisotropic minerals under strain in the mantle (Karato, 1998a; Montagner, 1998). Nataf et al. (1986) first inverted radial anisotropy and related the results to mantle flow. They found evidence for veritical flow beneath ridges and subduction zones and horizontal flow beneath lithospheric plates. Tanimoto and Anderson (1985) first inverted for azimuthal anisotropy at different periods and found that the fast directions correlate with plate motions at the lowest wavelengths. These studies were follwed by joint inversions for radial and azimuthal anisotropy (Montagner and Nataf, 1986; Montagner, 1986; Nishimura and Forsyth, 1989). Although anisotropy was first found to be limited to the upper mantle. Montagner and Kennett (1996), using normal modes, and Vinnik and P. (1996), using receiver functions, first found evidence for anisotropy in the transition zone. Finally the
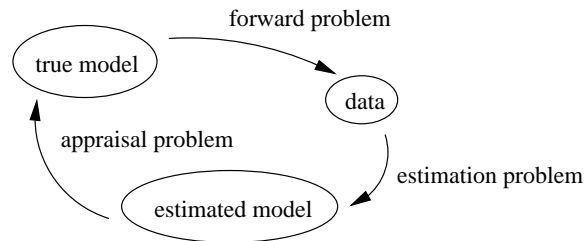
Figure 1.1: *Definition of the forward and inverse problem viewed as a combination of the estimation problem and the appraisal problem from Snieder and Trampert (2000).*

use of higher mode phase velocities allowed Trampert and van Heijst (2002) to construct the first global model of azimuthal anisotropy in the transition zone.

Seismic tomography is based on solving an inverse problem (figure 1.1). The data are observables obtained from seismograms such as travel times, phase velocities and splitting measurements and the objective is to find the Earth's structure ('true model'). Due to the limited number of data and the many degrees of freedom necessary to reconstruct an Earth model, the inverse problem is not unique. There are many models that explain the data equally well. The model obtained through inversion is therefore not equal to the true model but represents some estimated model. The appraisal problem tries to relate the estimated model to the true model. There are two reasons why the estimated model differs from the true model. One is the non-uniqueness mentioned above and the other is the propagation of uncertainties present in the data. In the appraisal problem, therefore, we need to find which properties of the true model are retrieved and what uncertainties are attached to them. An example of non-uniqueness in tomography is shown in figure 1.2, taken from Deal et al. (1999). Deal et al. (1999) first obtained an initial tomographic model by inverting delay time data. They then defined a theoretical slab model by finding velocity profiles that were obtained from an assumed slab temperature model. The initial tomographic model satisfies the delay time data, the theoretical slab model does not. They then projected the difference between the initial model and the theoretical slab model onto the null space (the area of the model space not constrained by the data of their initial problem). The resulting enhanced tomographic image now has a comparable fit to the delay time data as the initial tomographic image but contains null space components that were introduced by the theoretical slab model. Even though the initial model and the enhanced model fit the delay time data equally well by construction, the models themselves are quite different. In the initial model (figure 1.2), the slab seems to

(a) Initial   (b) Theoretical slab   (c) Enhanced
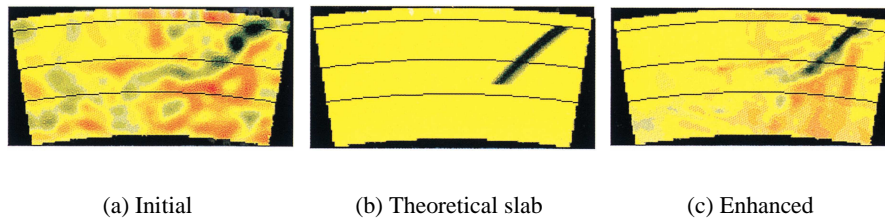
Figure 1.2: *Non-uniqueness in tomography taken from Deal et. al. 1999 . The enhanced tomography result is obtain by putting the difference between the initial model and the theoretical slab into the null space. The final image has been strongly biased towards the theoretical slab model and only deviates from it as required by the seismic data.*

lie flat and may even penetrate the 670km discontinuity while it clearly does not in the enhanced tomography result, illustrating the problem of non-uniqueness.

To illuminate the importance of model uncertainty we show in figure 1.3 two phase velocity models for the fundamental mode Rayleigh at 150 seconds. Although the long wavelength features are very similar the models display differences. Differences in tomographic maps are caused by choices in the inversion algorithm (exact, iterative), assumptions in the theory (great circle arc propagation, 'fat' rays), regularisation (norm damping, gradient damping), parameterisation (spherical harmonics, latitude-longitude grid) and type of data (body waves, surface waves, normal modes) (Boschi and Dziewonski, 1999; Romanowicz, 2003). Without appropriate uncertainties we cannot determine whether tomographic models are similar or different.

1.3   Dealing with the limitations

As already mentioned, one of the limitations of surface wave tomography is the depth resolution due to the use of mainly fundamental mode surface waves. To increase the resolution of surface wave tomography at depth we need to use higher mode surface waves (Nolet, 1975; Cara, 1979; Stutzmann and Montagner, 1993; Van Heijst, 1997; Yoshizawa and Kennett, 2002; Beucler et al., 2003). For similar periods, higher modes are sensitive to larger depths. This is illustrated in figures 1.4 and 1.5 which show several spheroidal (Rayleigh wave) and toroidal (Love wave) shear wave sensitivity curves. For the period range of 35s to 200s, the sensitivities for higher modes extend well into the lower mantle which should

Visser et al. (2007)

Trampert & Woodhouse (2003)

Percent
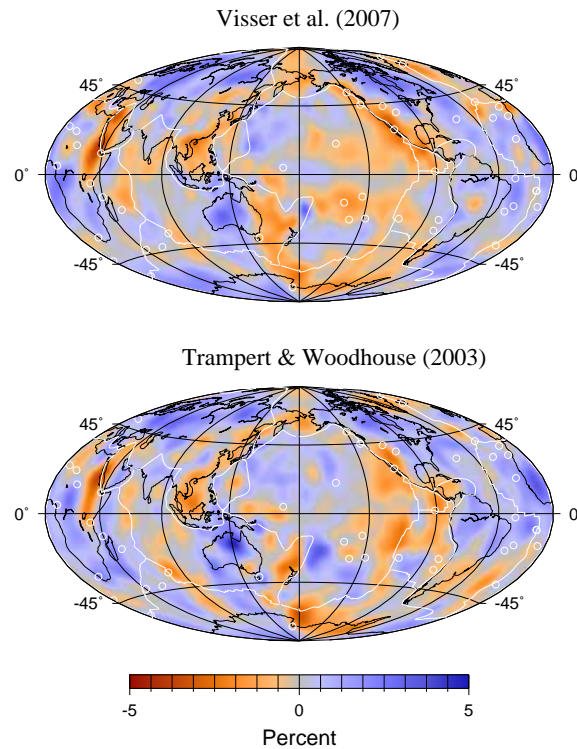
Figure 1.3: *Rayleigh wave phase velocity models at a period of 150s for Trampert & Woodhouse (2003) and Visser et. al. (2007).*

significantly improve the depth resolution of surface wave mantle models.

Waveform inversion techniques inherently use higher mode information whether by direct use of the seismograms (Nolet et al., 1986), or by the use of secondary observables based on cross-correlations between the observed and synthetic seismograms (Cara and Leveque, 1987). The relation between the seismograms and the underlying velocity model is highly non-linear, which is the reason that waveform inversion is often linearized. In most waveform inversion approaches, the one-dimensional velocity model obtained by waveform fitting is interpreted as the average structure along the source-receiver path. This assumption breaks down for high frequency higher modes (Marquering and Snieder, 1996) since they represent body waves with a sensitivity along the body wave path. This may be avoided by introducing mode coupling, however at large computational costs (Li and Romanowicz, 1996; Capdeville et al., 2000). This traditional two

fundamental mode   first higer mode   fourth higer mode
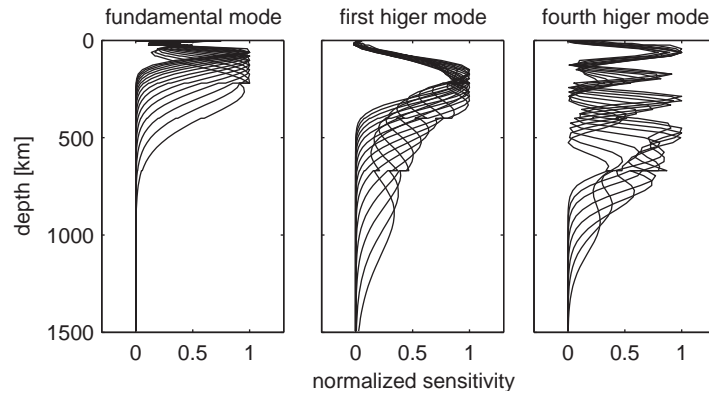
Figure 1.4: *Spheroidal shear wave sensitivities for the fundamental mode from a period of 35s to 175s, first higher mode from 35s to 175s and fourth higher mode from 35s to 65s. The sensitivities are normalized.*

fundamental mode   first higer mode   fourth higer mode
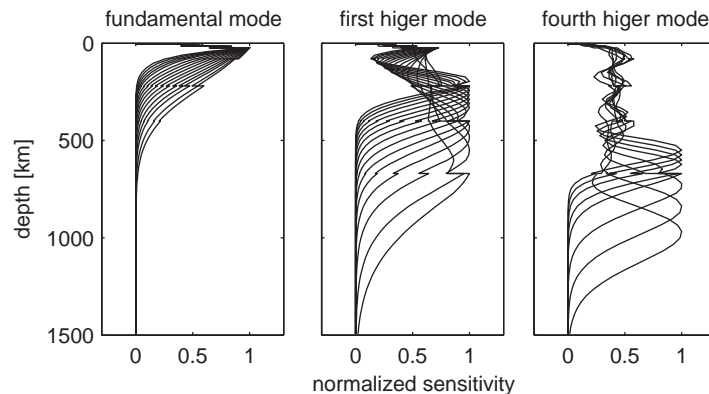
Figure 1.5: *Toroidal shear wave sensitivities for the fundamental mode from a period of 35s to 175s, first higher mode from 35s to 175s and fourth higher mode from 35s to 65s. The sensitivities are normalized.*

stage inversion consists of obtaining one-dimensional velocity models through waveform fitting and inverting them, using the path average assumption, for a three-dimensional velocity model. Kennett and Yoshizawa (2002) proposed a three stage approach in stead of the regular two stage approach which we adopt in this work. The three stage approach consists of obtaining one-dimensional veloc-

ity models through waveform fitting, building multimode phase velocity models as a function of frequency using the path average assumption and a final inversion for local wave speed properties to obtain a three-dimensional velocity model. The one-dimensional velocity model in the first stage is now regarded as a representation of the character of multimode dispersion along the source-receiver path. This three stage approach is flexible due to the ability to use different methods: waveform inversion (Nolet, 1990; Lebedev et al., 2005), mode stripping (Van Heijst and Woodhouse, 1999), roller-coaster technique (Beucler et al., 2003) and different types of information (Rayleigh waves, Love waves, polarisation information) to obtain the one-dimensional velocity model in the first stage. It is also a robust method. Yoshizawa and Kennett (2002) showed that multiple one-dimensional shear wave velocity models obtained through waveform fitting with a slight difference in misfit share the same dispersion characteristics indicating that the one-dimensional velocity model in the first stage may be regarded as a representation of the multimode dispersion characteristics along the source-receiver path.

Rather than assuming isotropy we invert for azimuthal and radial anisotropy as well. Radial and azimuthal anisotropy are both observed by surface waves which is why these are well suited to study anisotropy. The alignment (lattice preferred orientation or LPO) of intrinsically anisotropic minerals under strain in the mantle is assumed to be the major cause of upper mantle anisotropy (Montagner, 1998; Karato, 1998a). Anisotropy is thus an indicator of mantle deformation and flow. Adding higher mode information, also in the azimuthally and radially anisotropic tomographic maps, will increase our understanding of the dynamics of the mantle, especially in the transition zone and lower mantle.

A very important aspect of any inversion are the uncertainties (the appraisal problem). There have been many attempts to assign uncertainties to tomographic models. For this study, we focus on uncertainties of fundamental and higher mode phase velocity measurements and the uncertainties of the final three-dimensional shear wave speed models. In general, the standard deviations of the phase velocity measurements (and therefore also of the phase velocity models) and the shear wave velocity models are not known. The focus is more on the resolution of shear wave velocity models which is often assessed by synthetic tests (for example, checkerboard tests) (Bijwaard et al., 1998) or more recently by calculating the model resolution matrix (Boschi, 2003; Soldati and Boschi, 2005). The synthetic tests check whether an input model will be retrieved using the same source station geometry as for the real data. This is dependent on the input model and the source station geometry (Lévêque et al., 1993). The calculation of the model resolution matrix, on the other hand, is computationally very demanding. Un-

certainties for the (fundamental mode) phase velocity measurements have mostly been obtained using cluster analysis (Ritzwoller and Levshin, 1998; Trampert and Woodhouse, 2001; Shapiro and Ritzwoller, 2002), where similar paths are used to obtain the standard deviations. Trampert and Woodhouse (2001) showed that phase velocity standard deviations obtained with cluster analysis are in agreement with uncertainties obtained for comparing model predictions to real seismograms. Higher mode phase velocity uncertainties have been impossible to obtain using cluster analysis since higher modes are very difficult to measure using traditional techniques and there are too few measurements to reliably use cluster analysis. Van Heijst (1997) assigned a reliability measure for his higher mode measurements based on the amount of higher mode information in the seismogram and the fit between the seismogram and the synthetic prediction of the seismogram. Yoshizawa and Kennett (2002) measured fundamental and higher mode phase velocities with a model space search and took the 1000 best dispersion models to calculated approximate standard deviations.

In this study, we follow the three stage approach of Kennett and Yoshizawa (2002) and measure phase velocities using a model space search approach, as Yoshizawa and Kennett (2002). In a model space search each point in the model space represents a set of parameters describing some physical property of the Earth, for example velocity. The objective of the search is to find the (set of) models that minimizes an objective function, which in geophysical inverse problem is usually a measure of misfit between the observations and theoretical predictions. Contrary to Yoshizawa and Kennett (2002), we use the whole ensemble of models obtained with the model space search to obtain Bayesian statistical inferences about the ensemble. In this way we obtain not only the best model but also consistent and reliable uncertainties. These uncertainties are used in the second stage to invert the phase velocity measurements for global isotropic and azimuthally anisotropic phase velocity maps. After inversion, we obtain the posterior uncertainties of the phase velocity models. In the final inversion for the radially anisotropic shear wave velocity model, we use the phase velocity model uncertainties to obtain consistent posterior uncertainties on the anisotropic model.

1.4   Overview

In chapter 2 we present the fully automated waveform inversion method to obtain fundamental and higher mode phase velocity measurements. We follow Yoshizawa and Kennett (2002), and measure phase velocities using a model space search approach which enables us to obtain consistent standard deviations for the

phase velocity measurements. Chapter 2 also presents the tests we have performed to validate the method (convergence of the model space search, tuning parameters, parameterisation, approximations). In chapter 3 we present the Love wave higher mode measurements specifically since they are more difficult to measure than Rayleigh wave higher modes due to their overlap with the fundamental mode and the, in general, noisier seismograms for Love waves. Chapter 4 presents global isotropic and azimuthally anisotropic phase velocity models for the fundamental up to the sixth higher mode for Rayleigh and up to the fifth higher mode for Love waves. Finally, the isotropic phase velocity models are inverted for a global depth dependent radially anisotropic shear wave velocity model in chapter 5. We obtain the radially anisotropic shear wave velocity model using a model space search approach, which takes the full non-linearity of the problem into account and allows us to obtain consistent uncertainties and probabilities of radial anisotropy. Finally, in chapter 6 we present the general conclusion.

Chapter 2

Measuring phase velocities for fundamental and higher mode surface waves

In this chapter we explain the method we used to obtain fundamental and higher mode phase velocity measurements and corresponding uncertainties for Rayleigh and Love waves and discuss all the tests we performed to validate the automated procedure.

2.1   Separation of higher modes

One of the problems of surface wave tomography is the lack of depth resolution due to the use of mainly fundamental mode surface waves. For commonly used periods of 50s to 200s, the fundamental mode surface waves are sensitive to the upper 400km of the Earth. Adding higher mode surface waves extends the depth resolution to the transition zone and lower mantle.

Figure 2.1 shows the PREM (Dziewonski and Anderson, 1981) phase (c) and group (u) velocities for the fundamental and higher mode Rayleigh and Love waves. For the spheroidal modes at a phase velocity close to 8 km/s (figure 2.1, top), the Stonely wave at the CMB cuts across the sequence of branches. Especially at short periods, the group velocities for the higher modes only slightly differ from each other indicating that these higher modes will arrive almost simultaneously in the seismogram. At longer periods (>100s) for Rayleigh waves, the fundamental mode group velocities differ significantly from the first higher mode
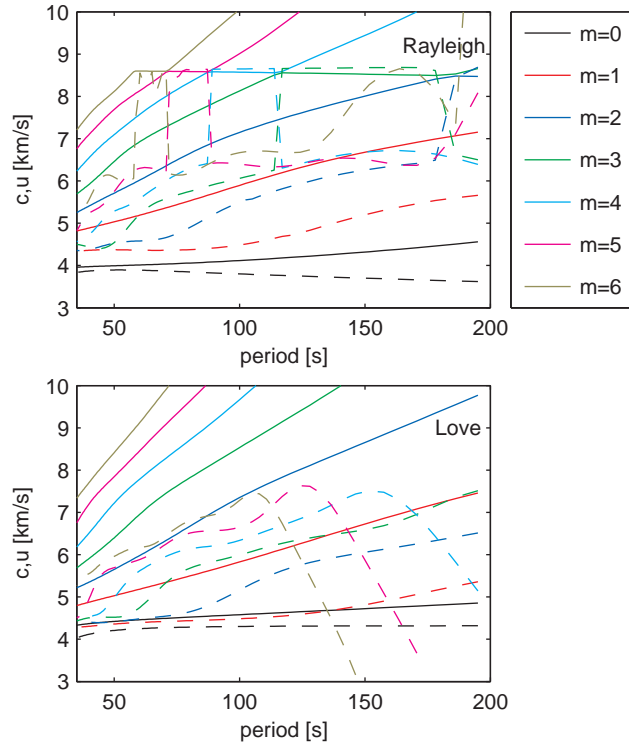
Figure 2.1: *Rayleigh and Love fundamental and higher mode dispersion curves for phase (c, solid lines) and group (u, dotted lines) velocity.*

group velocities such that the fundamental mode can be separated from the higher modes. For Love waves at all periods, the group velocity of the fundamental mode and the first higher mode do not differ a lot, which makes it challenging to separate the fundamental mode from the higher modes. This is also illustrated in figure 2.2, where we show the contribution per higher mode (m) to a synthetic seismogram for Rayleigh and for Love waves. For Rayleigh waves, the fundamental mode waveform (as shown with $m = 0$ in figure 2.2) does not change, the higher modes only add information in time before the fundamental mode, due to their higher group velocities. The fundamental mode is completely separated in time from the higher modes. For Love waves, the fundamental mode waveform ($m = 0$) changes if you add higher modes to the seismogram, illustrated by the changes around 2000s when adding higher modes (for example: compare
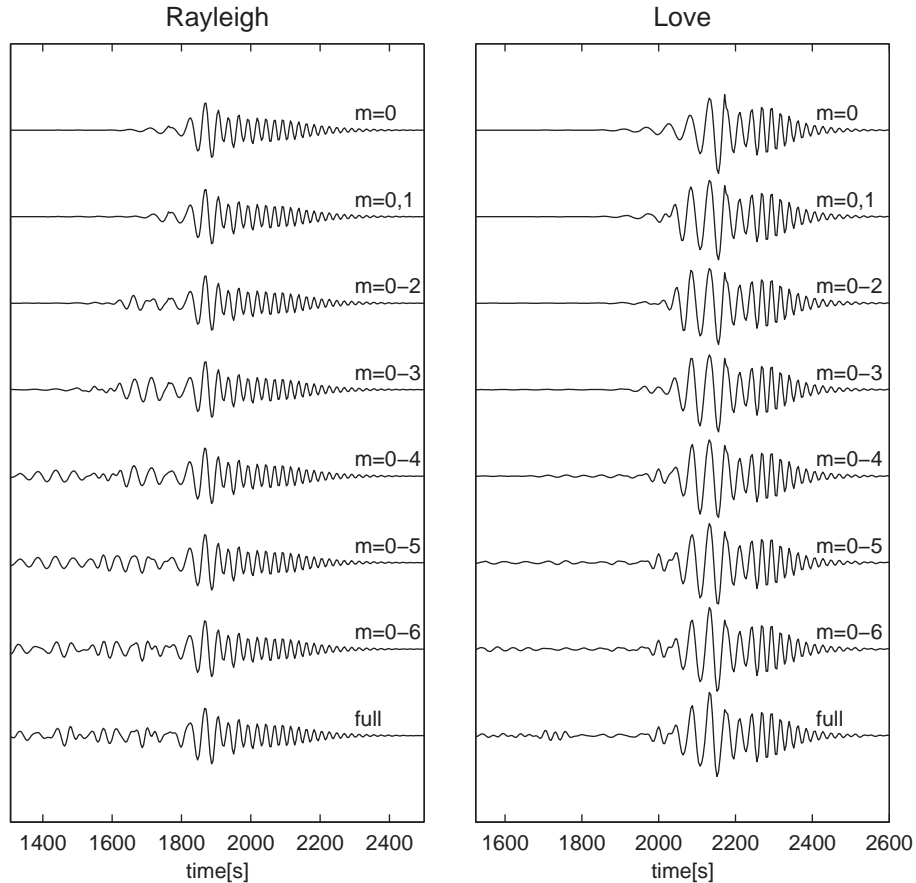
Figure 2.2: *Rayleigh and Love synthetic seismogram (filtered between 20 and 50 seconds) build with only the fundamental mode (m=0), the fundamental and first higher mode (m=0,1) up to full synthetic seismogram (full). For the Rayleigh seismogram the epicentral distance is 7584 km and the depth is 33.8 km. For the Love seismogram the epicentral distance is 9332 km and the depth is 15.0 km.*

$m = 0$ with $m = 0 - 3$ around 2000s). The fundamental mode and higher modes arrive almost simultaneously, which makes it challenging to separate the fundamental mode from the higher modes. In spite of this, most methods that measure higher modes are based on the separation of modes in the seismogram. Overtones were separated in the frequency-wavenumber (ω, k) domain, first by Nolet (1975, 1977) and later by Cara (1978), Mitchel (1980) and Okal and Jo (1987) us-

ing an array of seismometers. Cluster based techniques (Nolet, 1975; Cara, 1979; Stutzmann and Montagner, 1993; Beucler et al., 2003) use either a clustering of stations (Nolet, 1975; Cara, 1979) or of events (Stutzmann and Montagner, 1993; Beucler et al., 2003), which reduces the number of paths to measure phase velocities and the geographical ray coverage. The geographical ray coverage will be high in areas with lots of events, for example the Pacific subduction zone, or lots of stations, for example the continents. The models are generally valid only in regions of good ray coverage. The advantage of such an approach is that the clusters will have a better signal to noise ratio and phase velocities will be easier to measure.

Van Heijst and Woodhouse (1999) and Yoshizawa and Kennett (2002) used single seismograms to measure phase velocities, thus obtaining more measurements with better geographical coverage. Van Heijst and Woodhouse (1999) used a mode branch stripping technique which measures the phase velocities of the fundamental and first few higher modes. It relies on the separation of modes in the seismogram and it thus can only measure seismograms with sufficiently long paths where the modes are well separated. Yoshizawa and Kennett (2002) used a fully non-linear waveform inversion to obtain path specific multimode dispersion measurements for Rayleigh and Love waves. The method searches the model space for the shear wave velocity model that explains the seismogram best. This shear wave velocity model is interpreted as providing implicit information on the multimode dispersion of the specific source receiver path, but it is not meant as a direct representation of the Earth. A direct interpretation of path specific shear wave velocity models is only valid for small lateral perturbations along the path, whereas less constrictive conditions apply to the multimode dispersion curves (Kennett and Yoshizawa, 2002). We follow Yoshizawa and Kennett (2002) and perform a waveform inversion using a model space search approach. So our estimation problem is to obtain velocity models that explain the data (single observed seismograms) after which we need to solve the appraisal problem to make inferences about the uncertainties. For the appraisal problem, Yoshizawa and Kennett (2002) used the best 1000 dispersion models and calculated approximate standard deviations. In contrast, we use the whole ensemble of models obtained with the model space search to obtain Bayesian statistical inferences about the ensemble. In this way we obtain not only the best model but consistent uncertainties as well.

Figure 2.3: *flowchart of the method. Red block correspond to main programs, green to input blocks and blue blocks correspond to measurements.*
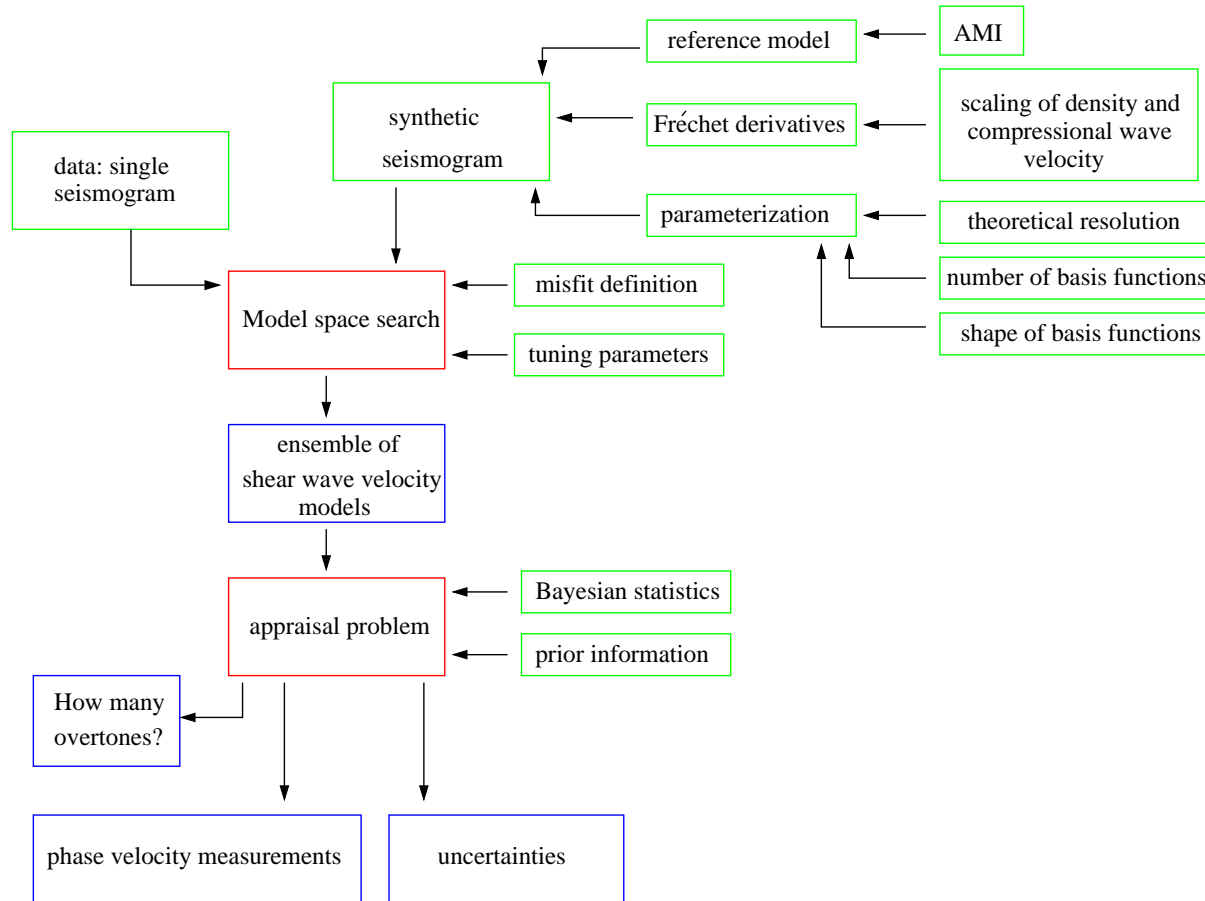
2.2   Overview: method and tests

Figure 2.3 gives an overview of all the ingredients of the method to obtain the phase velocity measurements. We will first briefly discuss all the ingredients and later discuss them in detail.

In a waveform inversion, the objective is to find the velocity model that gives a synthetic seismogram explaining the data best. The relation between the velocity model and the synthetic seismogram is highly non-linear. Therefore, we calculate the synthetic seismogram using linearized mode summation. For this linearisation, we need a reference velocity model and Fréchet derivatives of the reference model. The Fréchet derivatives are the derivatives of the phase velocity of the reference model with respect to the compressional wave velocity, shear wave velocity and density of the reference model. A change in the velocity model (in terms of a change in compressional, shear wave velocity and density) can then be expressed, using the Fréchet derivatives, as a change in phase velocity. This procedure relies heavily on the reference model chosen. Therefore, we use the best possible reference velocity model for the specific seismogram obtained with AMI (Lebedev et al., 2005). We perform waveform inversion using a model space search approach. The dimensions of the model space are determined by the parameterisation of the velocity model. We find the best possible parameterisation by calculating the resolution kernels for different parameterisations and determine the number and shape of the basis functions. Since we use a model space search approach, we need to define a search interval for every parameter. We define each search interval as a certain percentage of change with respect to the reference model. The model space search is guided to areas where the sampled velocity models give synthetic seismograms that fit the data well by a misfit criteria between the data and the synthetics. Finally, the tuning parameters of the model space search determine whether we perform a slow but extensive search or a fast converging search. Since we want to apply this method to thousands of seismograms we need to find the tuning parameters that give a fast, but reliable, ensemble of velocity models. At this stage, we have solved the estimation problem.

For the appraisal problem, we resample the ensemble of velocity models to obtain the posterior probability density function of the ensemble of velocity models and, at the same time, we obtain the posterior probability density function for specific surface wave mode phase velocities. This is possible since the relation between the velocity models and the phase velocities is known. Phase velocities are obtained by integrating the velocity model using sensitivity kernels for specific surface wave modes. We obtain one-dimensional marginals for specific phase velocity parameters by integrating the posterior probability density function over all

other phase velocity parameter using Bayesian statistics. The posterior probability density function depends on the prior information (parameterisation, search boundaries, forward theory) and the likelihood (misfit). Therefore, we examine the effect of the prior information on the inversion. Finally, we decide on the number of higher modes that are constrained by the seismogram. The one-dimensional marginals for the phase velocity parameters are Gaussian shaped which enables us to represent the phase velocity measurement by a mean and a standard deviation. In the end, we have obtained phase velocity measurements for the fundamental mode as well as a number of higher modes including inferences about the uncertainties of the phase velocity measurements.

2.3   Theory

Waveform inversion is a highly non-linear problem due to the complicated relation between the synthetic seismogram and the Earth model. A synthetic seismogram (s) can be calculated as a sum of modes (m) using the JWKB approximation (Dahlen and Tromp, 1998) in the frequency domain (ω):

$$s(\omega) = \sum_m A_m(\omega) exp[i\omega\Delta/c_m(\omega)], \tag{2.1}$$

where $A_m(\omega)$ is the complex amplitude of the modes, Δ is the source-receiver distance and $c_m(\omega)$ is the phase velocity of mode m. A part of the non-linearity of equation 2.1 is in the calculation of the phase velocity given an earth model. This can be linearized using an expansion for the phase velocity

$$c_m(\omega) = c_m^0(\omega) + \delta c_m(\omega) + \frac{1}{2}\delta c_m^2(\omega) + ... \tag{2.2}$$

Ignoring the second order terms and higher, equation 2.1 becomes

$$s(\omega) = \sum_m A_m(\omega) exp[i\omega\Delta/(c_m^0(\omega) + \delta c_m(\omega))], \tag{2.3}$$

Where $c_m^0(\omega)$ is the phase velocity in the reference model and $\delta c_m(\omega)$ is the phase velocity perturbation defined as

$$\delta c_m(\omega) = \int_0^a \left\{ \frac{\partial c_m^0(\omega)}{\partial V_P(r)}\delta V_P(r) + \frac{\partial c_m^0(\omega)}{\partial V_S(r)}\delta V_S(r) \right.$$
$$\left. + \frac{\partial c_m^0(\omega)}{\partial \rho(r)}\delta\rho(r) \right\} dr. \tag{2.4}$$

$\partial c_m^0(\omega)/\partial V_P(r)$, $\partial c_m^0(\omega)/\partial V_S(r)$ and $\partial c_m^0(\omega)/\partial \rho(r)$ are the Fréchet derivatives which relate the change in compressional, shear wave velocity and density from the reference model to changes in the phase velocity and a is the radius of the Earth. We normalise the amplitude of the synthetic seismogram $(A_m(\omega))$ to the real seismogram and solve mainly for the phase of the seismogram. The only unknown in equation 2.3 is now the phase velocity perturbation $(\delta c_m(\omega))$ and the waveform inversion is now reduced to finding the phase velocity perturbation given an Earth model. In this equation we recognise the definition of a linear inverse problem $\mathbf{d} = \mathbf{Gm}$, where the data $(\mathbf{d})$ is given by the phase veloctity perturbations $\delta c_m(\omega)$ as a function of mode m, the model $(\mathbf{m})$ is given by the perturbations in V_P, V_S and ρ $(\delta V_P, \delta V_S, \delta \rho)$ as a function of depth (r) and $\mathbf{G}$ the Fréchet derivatives. We use a model space search to solve this linearized inverse problem. Usually a model space search is used to solve non-linear problems, we use it to map the nullspace of a linear problem instead. For each sampled Earth model, we calculate the corresponding perturbations in phase velocity. The perturbations in phase velocity are used to compute the synthetic seismogram which is then compared to the observed seismogram. In the end, the model space search provides us with an ensemble of Earth models and their corresponding fit to the observed seismogram.

2.4   Isotropic reference model

We have assumed an isotropic reference model and isotropic perturbations to this reference model mainly for computational reasons. We use transversely polarized seismograms to estimate Love wave phase velocities and vertically recorded seismograms to estimate Rayleigh wave phase velocities. This is, in theory, only correct for an isotropic medium. We therefore examined whether we would obtain the same phase velocity measurements assuming a transversely isotropic medium. We first calculated Love and Rayleigh wave phase velocities for the anisotropic PREM model (Dziewonski and Anderson, 1981). We then separated the anisotropic PREM model in a horizontally polarized model (V_{PH}, V_{SH}) and a vertically polarized model (V_{PV}, V_{SV}) and calculated the Love and Rayleigh wave phase velocities separately assuming isotropy. Since anisotropic PREM contains only shallow anisotropy, we performed the same test for a radially anisotropic model taken from chapter 5, located in the Baltic Shield (figure 2.5). Figure 2.4 shows the differences in phase velocity assuming isotropy and radial anisotropy. For both tests the differences in the phase velocities calculated assuming isotropy or anisotropy are small. Therefore we conclude that as long as the standard devi-
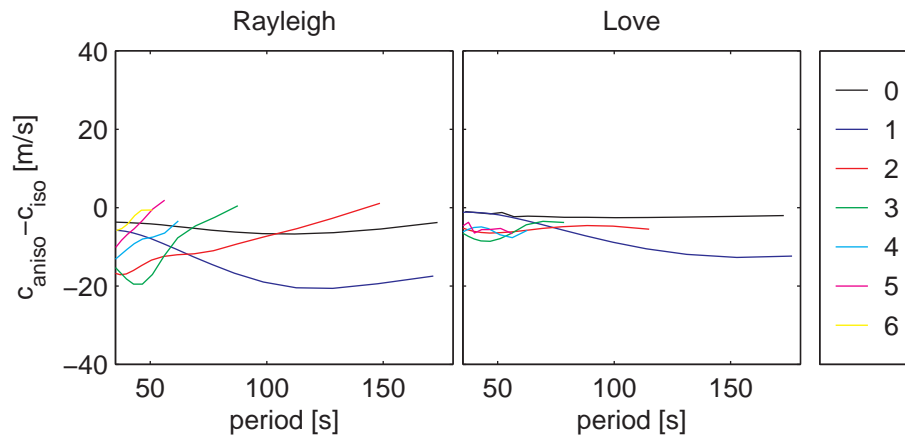
Figure 2.4: *The difference between the Love and Rayleigh fundamental and higher mode phase velocities calculated assuming anisotropic and isotropic profiles (see figure 2.5). For Love, the isotropic model has V_S and V_P equal to V_{SH} and V_{PH} of the anisotropic model, for Rayleigh the isotropic model has V_S and V_P equal to V_{SV} and V_{PV} of the anisotropic model.*
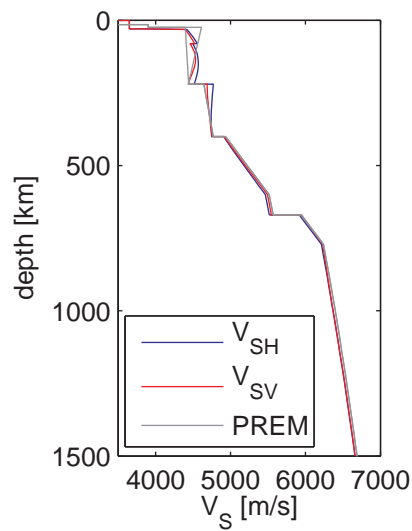
Figure 2.5: *The radially anisotropic model taken from chapter five at the Baltic shield location plotted with the anisotropic PREM model.*

ations on the phase velocity measurements are larger than these differences it is justified to assume isotropy.

2.5   Why do we need to linearise the phase velocity?

In principle, we could include the full non-linearity of the forward problem and solve equation 2.1. In practise, this is a very time-consuming process since we would have to recalculate the eigenfunctions for each model we sample in the model space search. Each model space search typically samples thousands of models per seismogram and we want to perform the model space search on hundreds of thousands of seismograms, which makes it impossible to include the full non-linearity of the problem. Therefore we chose to linearise the inversion by using the Fréchet derivatives. This enables us to calculate the phase velocity perturbations at each point in the model space using a best reference model and the perturbation to the reference model. It does introduce a dependence on the reference model. We will only be able to obtain a global minimum if our final model is close to the reference model. Therefore we choose the reference model to be the best shear wave velocity model that results from the Automated Multimode Inversion (AMI, Lebedev et al. , 2005) method. AMI is a non-linear waveform inversion method that uses carefully chosen time and frequency windows to obtain a shear wave velocity model that explains both the fundamental as well as the higher modes in the seismogram. The resulting shear wave velocity model acts as our reference model but is in fact already the best shear wave velocity model which we would find using the model space search. Our best shear wave velocity model is (close to) the AMI best shear wave velocity model. The small differences between the models can be explained by different parameterisations and windowing. AMI also acts as a data quality control and rejects seismograms that are too noisy, are invalid according to the JWKB approximation, or are close to the nodes of the fundamental or higher mode radiation pattern.

2.6   Scaling the density and the compressional wave velocity

The phase velocity perturbation depends on the changes of the compressional wave velocity, shear wave velocity and density. For Rayleigh waves the sensitivity to shear wave velocity is highest, followed by the compressional wave velocity and density. For Love waves the shear wave velocity sensitivity is highest fol-

lowed by that of density. In principle, all variations (compressional wave velocity, shear wave velocity and density) could be allowed in the inversion. In practice, computational resources force us to limit the number of variations allowed in the inversion since the number of parameters would triple for Rayleigh if we take variations in compressional wave velocity, shear wave velocity and density into account. In general, a model space is empty and an increase in the number of dimensions severely increases the time needed for a thorough search of the model space. Since we will use the Neighbourhood Algorithm (Sambridge, 1999a,b), the computation time needed for increasing number of parameters would grow exponentially. The Neighbourhood Algorithm is efficient only for a small number of parameters ($<$30) and since we want to obtain a relatively fast model space search to apply to hundreds of thousands of seismograms, we need a small number of parameters. The density and compressional wave velocity are of secondary importance (due to the lower sensitivities compared to the shear wave velocity) and thus we decided to scale them to the shear wave velocity perturbations using the scaling relation of Ritsema and Van Heijst (2002) ($R = \partial V_s / \partial V_p$) for the compressional wave perturbations and Deschamps et al. (2001) ($\xi = \partial ln\rho / \partial ln V_s$) for the density perturbations. The scaling relation for the compressional wave velocity from Ritsema and Van Heijst (2002), increases from $R = 1.25$ at the surface to $R = 3.0$ at the core mantle boundary. This positive correlation between the compressional wave velocity and the shear wave velocity has been found earlier by other studies (Su and Dziewonski, 1997; Vasco and Johnson, 1998; Masters et al., 2000; Saltzer et al., 2001; Resovsky and Trampert, 2003). The increase with depth for the scaling relation is found up to 2000 km (Saltzer et al., 2001; Resovsky and Trampert, 2003). We are interested in the scaling relation up to a depth of 1500 km, since this is the limit of our one-dimensional velocity model, so the increasing scaling relation is appropriate. For the scaling of the density with the shear wave velocity, there is no consensus, but luckily surface waves are not very sensitive to density due to the strong oscillations in the density sensitivity curves with depth. This lack of consensus is illustrated by Resovsky and Trampert (2003), who used probabilistic seismic tomography to test mantle velocity-density relationships. While the relation between the compressional wave velocity and the shear wave velocity is clearly constrained and positive, the scaling relation between the density and shear wave velocity is not as well constrained with the probability of a positive or negative scaling relation non-zero at all depths.

2.7   The effect of scaling the density

Since there is no consensus on a scaling relation for density, we need to check the effect of the scaling relation on our results. We use the scaling relation of Deschamps et al. (2001) ($\xi = \partial ln\rho/\partial lnV_s$) which is derived from an inversion of gravity and S-wave velocities. They propose different scaling relations for sub-oceanic and sub-continental paths. Therefore, for each seismogram we extracted the ocean/continental information from CRUST2.1 (Mooney et al., 2004). An alternative density scaling relation is given by Kaban and Schwintzer (2001), who inverted a global shear wave velocity model and residual crust-free gravity anomalies to obtain a sub-oceanic scaling relation that extends into the lower mantle. There are significant differences between the two scaling relations under oceans. The scaling relation of Kaban and Schwintzer (2001) is larger in the top 60 km of the model and at depths greater than 250 km.

We applied the waveform inversion using no density scaling, the scaling relation of Deschamps et al. (2001) and the scaling relation of Kaban and Schwintzer (2001) and compared the phase velocity measurements. The phase velocity measurements are very similar and the difference is well within our standard deviations. In the case where we assume no density scaling, the number of modes we are able to measure is slightly smaller (see section 2.18).

This test suggests that it does not matter which density scaling relation we use as long as we use a density scaling relation.

2.8   Parameterisation

After scaling equation 2.5 simplifies to

$$\delta c_m(\omega) = \int_0^a \frac{\partial c_m^0(\omega)}{\partial V_S(r)} \delta V_S(r) dr. \tag{2.5}$$

We parameterise the shear wave velocity perturbations $\delta V_s(r)$ as a set of one-dimensional basis functions $h_i(r)$:

$$\delta V_S(r) = \sum_{i=1}^{N} \gamma_i h_i(r), \tag{2.6}$$

where the γ_i are the coefficients to be found in the waveform inversion and N the number of basis functions. Combining equations 2.5 and 2.6 we obtain

$$\delta c_m(\omega) = \int \frac{\partial c_m^0(\omega)}{\partial V_S(r)} \sum_{i=1}^{N} \gamma_i h_i(r) dr \tag{2.7}$$

$$\sum_{i=1}^{N} \gamma_i \int \frac{\partial c_m^0(\omega)}{\partial V_S(r)} h_i(r) dr. \tag{2.8}$$

Where we finally find the expression for the theory that links the data to the model in the inverse problem $\mathbf{G} = \int \frac{\partial c_m^0(\omega)}{\partial V_S(r)} h_i(r) dr$.

An overparameterisation (large N) will capture all details of the one-dimensional shear wave velocity models but will also significantly increase the computation time needed for the model space search due to the increase in the number of dimensions. On the other hand, an underparameterisation (small N) affects the depth resolution and causes a bias in our phase velocity measurements while providing us with a fast model space search. Since the resolution of surface waves decreases with depth we expect to be able to solve for thinner layers at the top of our model and for thicker layers as the depth increases. We performed a Backus-Gilbert resolution test to identify the optimal number and depths of the basis functions.
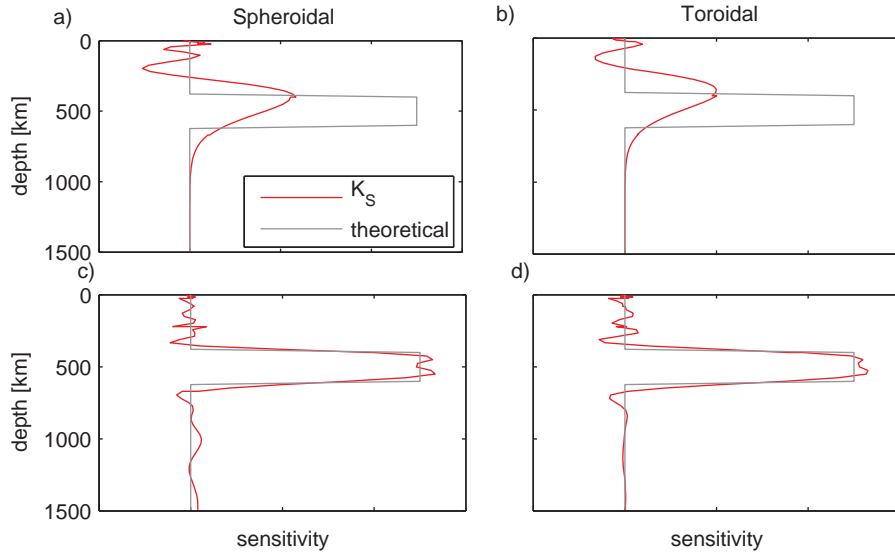
Figure 2.6: *Resolving a layer from 400 km to 600 km using only the fundamental (a,b) or the fundamental and higher modes (c,d) for the Spheroidal (a,c) and Toroidal (b,d) target sensitivity kernels. In grey the target sensitivity kernel and in red the best fit resolution kernel for the shear wave sensitivity.*
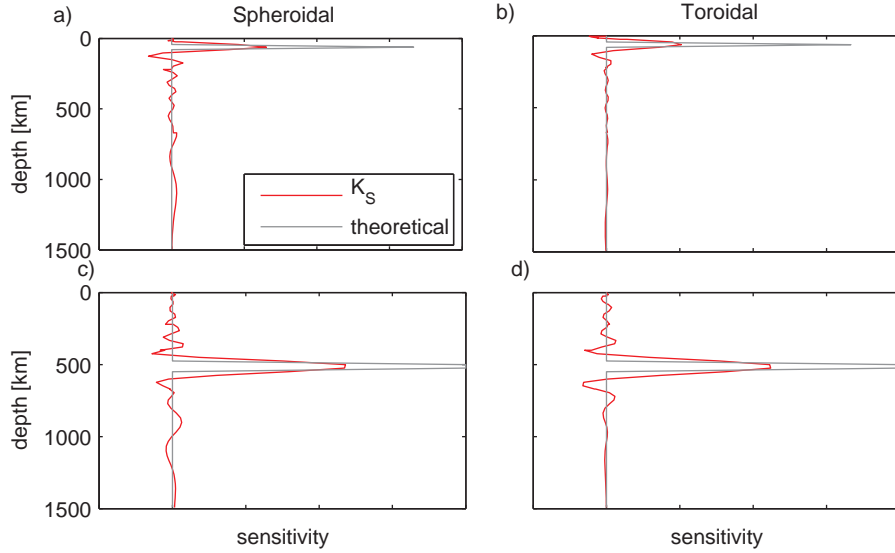
Figure 2.7: *Resolving thin layers for the Spheroidal (a,c) and Toroidal (b,d) target sensitivity kernels. In grey the target sensitivity kernel and in red the best fit resolution kernel for the shear wave sensitivity.*

2.9   Backus-Gilbert

We tested how well we can resolve shear wave velocity for a target depth layer. We use ray theoretical toroidal and spheroidal sensitivity kernels up to the sixth higher mode in the period range between 35s and 200s. The kernels are calculated in PREM (Dziewonski and Anderson, 1981). We further remove surface wave modes with sensitivity in the core and the Stoneley mode. We start with equation 2.5 and define the change in phase velocity as

$$\delta ln c_m(\omega) \quad = \quad \int_o^a \left\{ \frac{\partial c_m^0(\omega)}{\partial V_P(r)} \delta ln V_P(r) + \frac{\partial c_m^0(\omega)}{\partial V_S(r)} \delta ln V_S(r) \right. \qquad (2.9)$$
$$\left. + \frac{\partial c_m^0(\omega)}{\partial \rho(r)} \delta ln \rho(r) \right\} dr.$$

The compressional wave velocity and density are again scaled to the shear wave velocity using the scaling relations of Ritsema and Van Heijst (2002) and De-

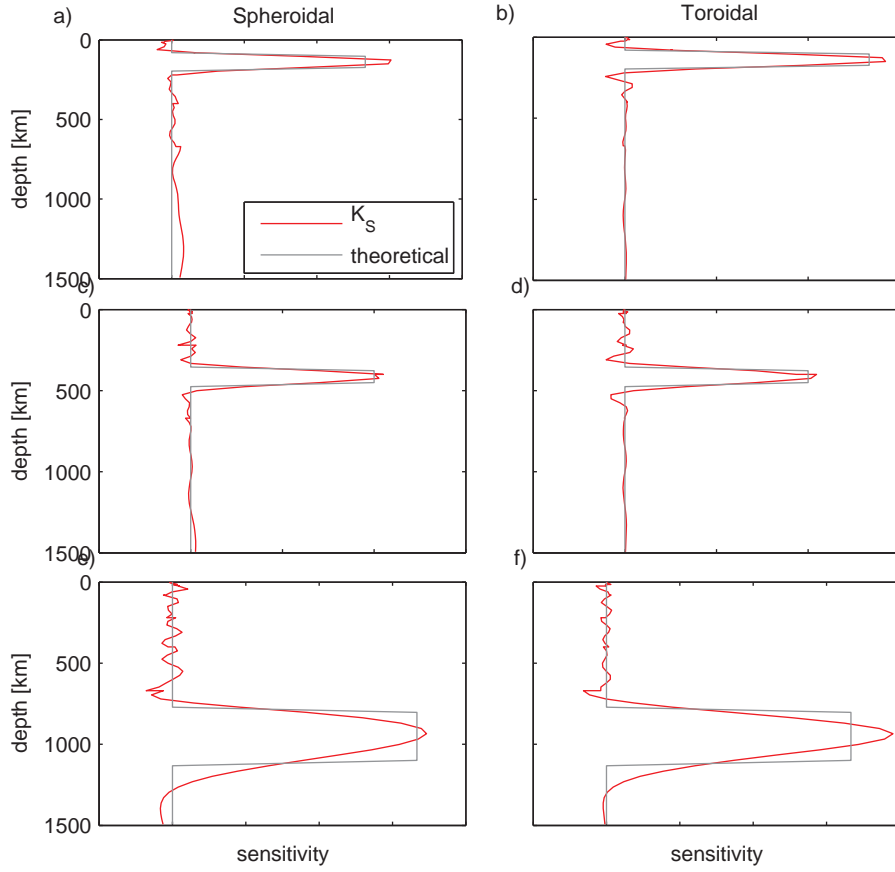Figure 2.8: *Best fit resolution kernels (red) for the Spheroidal (a,c) and Toroidal (b,d) target sensitivity kernels (grey).*

schamps et al. (2001). This reduced equation 2.10 to

$$\delta ln c_m(\omega) = \int_o^a \frac{\partial c_m^0(\omega)}{\partial V_S(r)} \delta ln V_S(r) dr. \tag{2.10}$$

For convenience, we write the shear wave sensitivity kernel as

$$K_S = \partial c_m^0(\omega)/\partial V_S(r). \tag{2.11}$$

A target shear wave velocity $\delta ln V_S(r^0)$ centered at radius r^0 can be found by

$$\delta ln V_S(r^0) \quad = \quad \sum_{i=1}^{N} \alpha_i \delta ln c_i(\omega) \tag{2.12}$$

$$= \quad \sum_{i=1}^{N} \alpha_i \int_0^a K_S^i \delta ln V_S(r) dr \tag{2.13}$$

$$= \quad \int_0^a (\sum \alpha_i K_S^i) \delta ln V_S(r) dr, \tag{2.14}$$

where N is the total number of surface wave modes, α_i are the coefficients and $\sum \alpha_i K_s^i$ is the resolution kernel. We need to find the coefficients (α_i) such that $dln V_s(r^0) - \sum \alpha_i K_s^i$ is minimized. This can be done by a classical least-squares optimization. The resolution kernel can be used to illustrate the fit to any target depth layer. First we checked the influence of higher modes on the resolution kernel for both the spheroidal as well as the toroidal modes. As expected, higher modes are indispensable if we want to resolve layers at larger depths (for example in the transition zone in figure 2.6). Also, it is impossible to resolve thin layers both at shallow as well as at larger depths (figure 2.7). The ideal parameterisation is able to fit small layers at shallow depths. At larger depths, the layers should increase in size, as illustrated in figure 2.8.

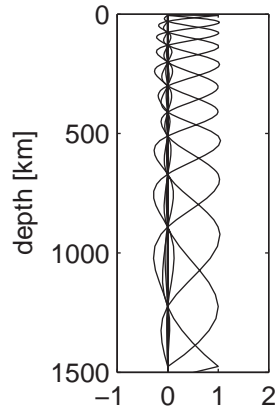| # Spline	Toroidal [%]	Spheroidal [%]
1	62	43
2	61	58
3	85	86
4	88	85
5	93	86
6	90	82
7	91	84
8	89	74
9	91	73
10	89	61
11	81	25
12	41	0

Figure 2.9: *Variance reduction in percent of recovery (left) for the target depth layer corresponding to the twelve spline parameterisation (right) using spheroidal and toroidal modes. The splines are numbered from the top (1) to the bottom (12).*

In the end, we have chosen twelve natural cubic splines for the parameterisation with a more dense layering at shallow depth and less dense at larger depths (figure 2.9). We are able to resolve this parameterization for both Rayleigh and Love waves quite well. This is illustrated by the variance reductions for the spline parameterisation (figure 2.9). The variance reduction is smallest for the spline at 1500 km, this is due to the decrease in resolution at these depths ranges for the higher modes. The first spline also has a relatively small variance reduction for the spheroidal modes. This spline is situated in the crust where the toroidal modes are more sensitive hence the larger variance reduction for the toroidal modes. The best resolved splines are situated in the 60km to 800km depth range. In this range, we are able to resolve more than 85% for Love waves and 73% for Rayleigh waves. We obtain better variance reductions at larger depths (spline 8 to 12) for the toroidal modes than for the spheroidal modes. This can be explained by looking at the shapes of the sensitiviy curves for the toroidal and spheroidal higher modes that are sensitive to these depths. The spheroidal higher modes have their maximum sensitivity at shallow depths and a very broad peak sensitivity around 500 to 1500 km. The toroidal higher modes, on the other hand, have their maximum sensitivity at larger depths (depending on the specific mode from 500 to 1200 km) and the peak of this sensitivity is narrower. This indicates that it is easier to find coefficients for the toroidal modes to fit the target sensitivity kernel than for the spheroidal modes at larger depths, which explains the better variance reductions.

2.10 **Number of basis functions**

In the resolution test we have seen that we need a dense parameterisation at shallow depths and a less dense parameterisation at larger depths. The number of basis functions has an effect on the depth resolution of the shear wave velocity models. A fine parameterisation inherently allows a higher depth resolution than a coarse parameterisation, but also results in a larger number of basis functions (dimensions) in the model space search. As the number of dimensions increases, the computation time for the Neighbourhood Algorithm increases dramatically. Also, a fine parameterisation at every depth is not needed due to the decreasing resolution with depth of surface waves as illustrated in the previous section. From the Backus-Gilbert resolution test, we selected twelve splines that give resolution kernels close to the target depth kernels indicating we will be able to resolve shear wave velocity with that parameterisation. We still wanted to see the effect of this choice and tested three other parameterisations (figure 2.10) ; a 24 spline
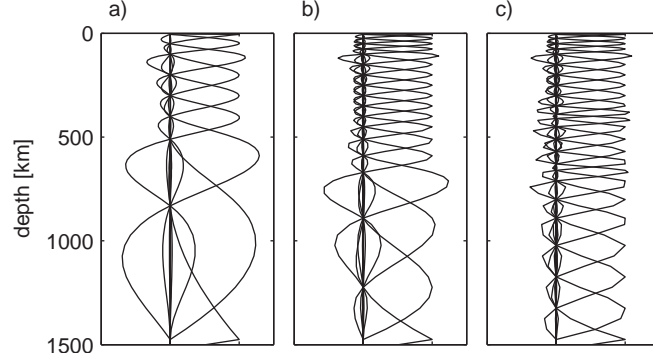
Figure 2.10: *Three spline parameterisations with depth, a 9 spline parameterisation (a), a 18 spline parameterisation (b) and a 24 spline parameterisation (c).*

parameterisation, a 18 spline parameterisation, and a 9 spline parameterisation. The parameterisations have thin layers at shallow depths and thicker layers as the depth increases to match the expected depth resolution of surface waves (see previous section).

First, we perturbed PREM using 9, 18 or 24 splines and calculated the exact synthetic seismogram of the perturbed model using mode summation. Second, we assumed PREM as a reference model and performed a model space search to find the best fitting shear wave velocity models to the exact seismograms. The model space search used either 9, 18 or 24 basis functions. We selected the best 100 shear wave velocity models that resulted from the model space search and calculated the mean misfit and standard deviation of the misfits. The position of the exact perturbed model in the model space is known and we can therefore also calculate the relative distance to the solution shear wave velocity model. We choose to express the relative distance in terms of perturbations from PREM, since the perturbations for the solution models are exactly known. The relative distance is defined as

$$d = \frac{1}{N} \sum_{i=1}^{N} \frac{(\gamma_i^0 - \gamma_i)^2}{(\gamma_i^0)^2}, \tag{2.15}$$

where d is the relative distance, N is the number of basis functions, γ^0 are the perturbations for the solution shear wave velocity model and γ are the perturbations for one of the 100 best shear wave velocity models we selected from the model space search. We calculate the relative distance for all 100 models and
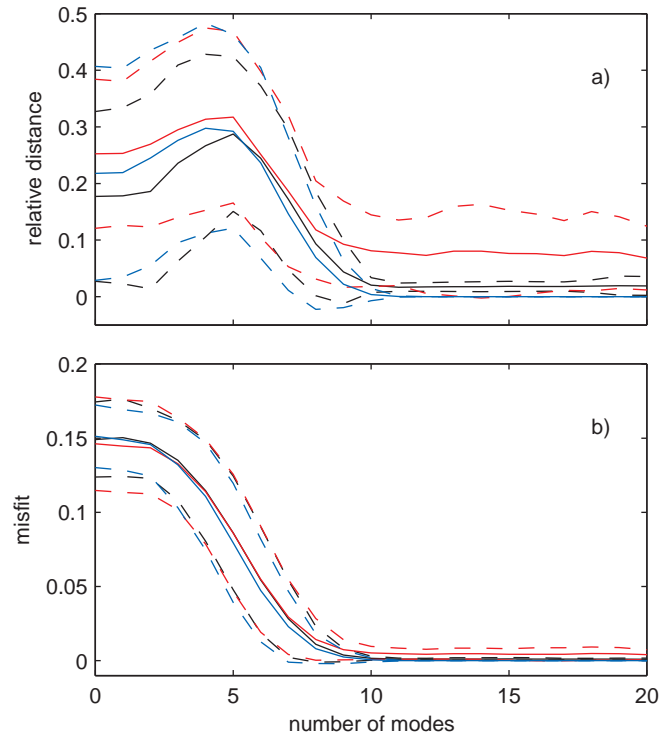
Figure 2.11: *(a) The relative distance to the solution and (b) the misfit as a function of the number of modes in the synthetic seismogram for a perturbed 24 spline model (black), a perturbed 18 spline model (red) and a perturbed 9 spline model (blue) fitted using 9 splines. Solid lines give the mean and the dashed lines the standard deviation.*

compute the mean and standard deviation. We only show three test results in this section since the results of the other tests are similar. Figure 2.11 shows the relative distance and misfit as a function of the number of higher modes used in the synthetic seismogram calculation for the perturbed 9, 18 or 24 spline model fitted using 9 splines. As the number of modes in the synthetic seismogram increases, the misfit and distance to the solution model decrease to zero within the standard deviations, which means that we are able to find the exact solution. The figure also shows that for all three tests, the behaviour is quite similar. It is initially surprising that we can fit a perturbed 24 spline model using only 9 parameters. This can be explained by examining the eigenfunction curves of modes 10 through 15 (figure 2.12). These modal eigenfunctions display two lobes in the 1000 to 1500

Figure 2.12: *Modal eigenfunctions for the indicated modes. The solid line represent the vertical component U and the dashed line the horizontal component V.*

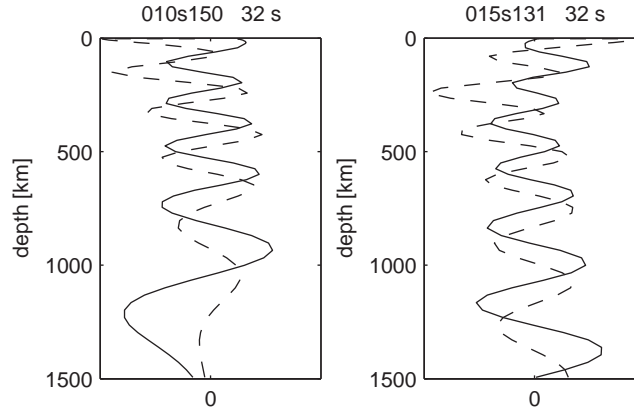km depth range. Even the 9 spline parameterisation has three lobes in the same depth range. Most seismograms contain a significant amount of information only below mode 10, which is why all the parameterisations we used are good enough. They are all capable of solving up to the fifteenth higher mode which we do not expect to be able to measure due to its low amplitude.

2.11   Shape of the basis functions

The parameterisation we chose according to the Backus-Gilbert resolution tests is given by twelve natural cubic splines. Other choices for the shape of the param- eterisation are also possible. Here we compared the difference in phase velocity for a twelve spline parameterisation and a twelve triangle parameterisation (fig- ure 2.13f,g) The triangles are defined such that their maximum value is positioned at the same depths as the splines maxima and their sensitivity is zero at the depths where the splines contain sidelobes. We chose this parameterisation to exclude changes in the position of the basis functions affecting the test. Figure 2.13a-e shows the difference in phase velocity for the fundamental and first four higher modes between a triangle and spline parameterisation. Also, we indicated the standard deviations obtained for these measurements. The phase velocity mea- surements are well within the standard deviations and are in fact very close to zero. Therefore we conclude that the shape of the basis functions does not signif-
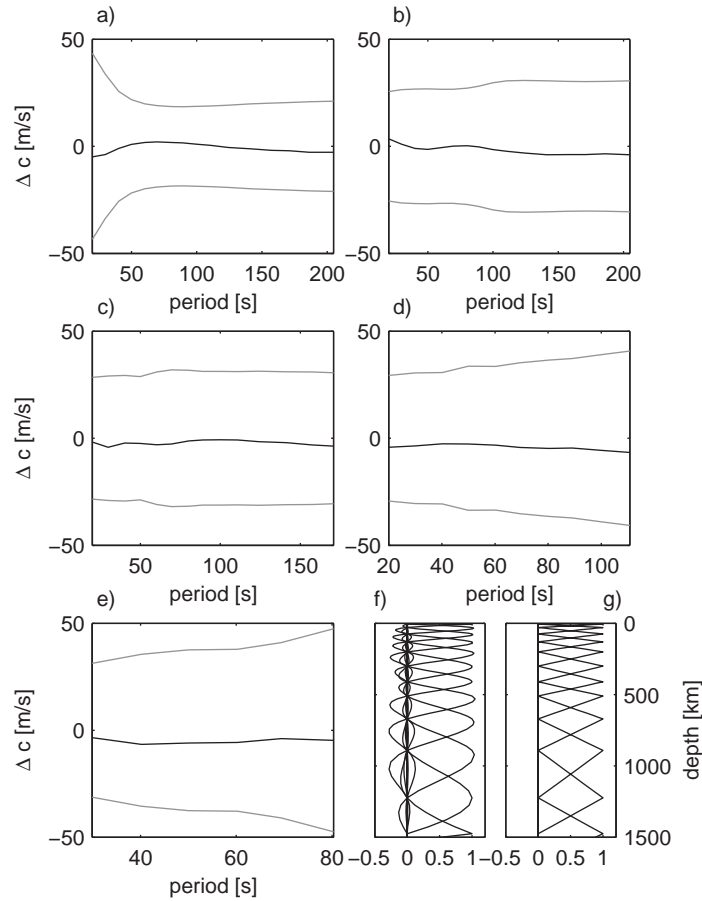
Figure 2.13: *Difference between phase velocity measurements using spline (f) or triangle (g) basis functions for the fundamental (a), first higher mode (b), second higher mode (c), third higher mode (d), fourth higher mode (e). The grey lines indicate the corresponding standard deviation.*

icantly influence the phase velocity measurements.

2.12 Neighbourhood Algorithm

Central in this research is the use of the Neighbourhood Algorithm (NA) by (Sambridge, 1999a,b). The Neighbourhood Algorithm consists of two parts. The first
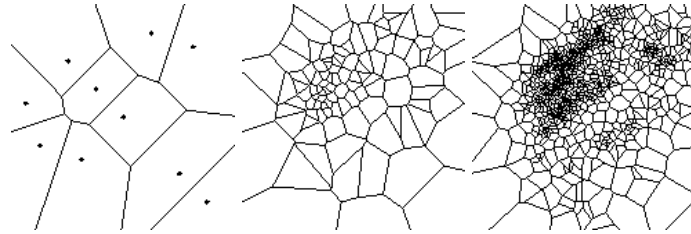
Figure 2.14: *quasi uniform random points and their Voronoi cells (right), the Voronoi cells about the first 100 (middle) and 1000 (left) samples generated by a Gibbs sampler using the neighbourhood approximation (from Sambridge, 1999a).*

part is a Monte Carlo sampling of the model space which makes use of a geometrical construct known as the Voronoi cell to drive the search in the model space. The Voronoi cells are defined as nearest neighbour regions using a suitable distance norm. The objective of the Neighbourhood Algorithm is to find an ensemble of models that preferentially samples the good data fitting regions in the model space, rather than a single optimal model. The NA first uniformly (or otherwise) generates an initial set of models (n_i) in the parameter space and calculates the misfit corresponding to these models and their Voronoi cells. The models are then ranked by lowest misfit. n_s new models are generated by performing a uniform random walk (using a Gibbs sampler, see Geman and Geman (1984); Rothmann (1986) for extensive explanations) in the Voronoi cells around the n_r lowest misfit models. Then the models are ranked again according to their lowest misfit and the Voronoi cells are recalculated. This is repeated for a selected number of iterations (n_{iter}). Figure 2.14 shows an example of the sampling of the model space using the NA. The higher density of the Voronoi cells indicates regions of better fit in the model space.

The second part of the NA consists of the appraisal problem or in other words, how to infer information from the ensemble of models. The NA first constructs an approximate posterior probability density (PPD) using the assumption of constant known PPD values in each Voronoi cell of the input ensemble. Using a Gibbs sampler, multiple random walks are performed in the model space. The random walk starts at some point (for example the best fit model) and takes a series of steps along each parameter axis in turn to obtain the next model. From this model, a new series of steps along each parameter axis in turn is performed to obtain the next model. This is repeated until the specified number of models is reached. After multiple random walks the distribution will asymptotically resemble the

Table 2.1: *frequency-time windows*

windows (t,ω)		
f (mHz)	Rayleigh (km/s)	Love (km/s)
5 - 10	3.7 ± 0.75	b - 3.8
10 - 20	3.75 ± 0.55	b - 3.8
	b - 4.3	
20 - 50	b - 4.3	b - 4.3

Definition of frequency-time windows for Rayleigh and Love seismograms. The time windows are defined using the group velocity (km/s).

approximate PPD (importance sampling). This resampled ensemble can be used in a Bayesian framework to construct statistical properties of the ensemble such as one or two dimensional marginals and the covariance matrix.

2.13   Model space search

Usually a model space search is used to solve highly non-linear problems. Our problem is linearized (equation 2.3) and we use the model space search to map the nullspace of the linear problem instead. The dimension of the model space is determined by the number of spline coefficients γ of equation 2.6 (in this case $N = 12$, since we use twelve natural cubic splines for the parameterisation). The Neighbourhood Algorithm performs a guided Monte Carlo type sampling of the model space using the values of the misfit between the real and the synthetic seismogram. The data and the synthetic are compared in different time and frequency windows (table 2.1). The windows are chosen such that the fundamental mode and the higher modes are included in the windows. The time b in table 2.1 depends on the epicentral distance. Below 35 $^\circ$ the time b corresponds to the time just before the arrival of the S wavetrain, between 35° and 70° the time b corresponds to the time just after the S and before the SS wavetrain etc. The misfit guides the model space search to area's of better fit. We prefer a relative least squares misfit since we expect it to be able to fit the higher modes better than an absolute misfit definition. The relative misfit is defined as

$$M = \sum_{\mathsf{w}=1}^{N} \sum_{i=1}^{L} \frac{(d_i - s_i)^2}{(d_{max}(\mathsf{w}))^2}, \tag{2.16}$$

where d is the seismogram, d_{max} the maximum amplitude of the seismogram for the window w and s the synthetic seismogram. The number of windows is N and the number of time samples in each window is L. The definition of the misfit influences the sampling of the Neighbourhood Algorithm. Therefore we considered another misfit function.

The absolute misfit is defined as

$$M_{abs} = \sum_{w=1}^{N} \sum_{i=1}^{L} (d_i - s_i)^2, \tag{2.17}$$

where d is the seismogram and s is the synthetic seismogram computed for all windows N and all time samples in the windows L. We use the same time and frequency windowing as in table 2.1. The relative misfit in equation 2.16 should weight the higher modes more since the misfit is defined relative to the maximum amplitude in the window which reduces effect of higher amplitudes for the fundamental mode. The absolute misfit should favor the fundamental mode since it gives more weight to higher amplitudes. We performed the waveform inversion for a test dataset using the relative and absolute misfit definitions and compared the best-fitting synthetic seismograms to the real seismograms. Figure 2.15 shows the effect of the absolute and relative misfit for one seismogram. Especially in the second frequency window (compare figure 2.15b with 2.15e) the higher mode fit is less when the absolute misfit is used. Although the difference is only small we prefer to use the relative misfit which improves the higher mode fit.

Figure 2.16 shows two examples of the best fitting waveform for a vertical (Rayleigh) and horizontal (Love) seismogram obtained using the relative misfit in the model space search. Also indicated are the different time and frequency windows for Rayleigh and Love waves. We have chosen two time windows in the second frequency window for Rayleigh, since for Rayleigh we are able to separate the fundamental mode from the higher modes. This will improve the relative misfit for these windows, since we use the maximum amplitude in the window to scale the misfit. For Love waves, we use only one time window since the fundamental mode and higher modes cannot be separated. The effect for the higher mode misfit is clearly shown in figure 2.16d. The Love wave higher mode fit is not as good as the Rayleigh higher mode fit. Also, Love wave seismograms are noisier, which affects the fit.
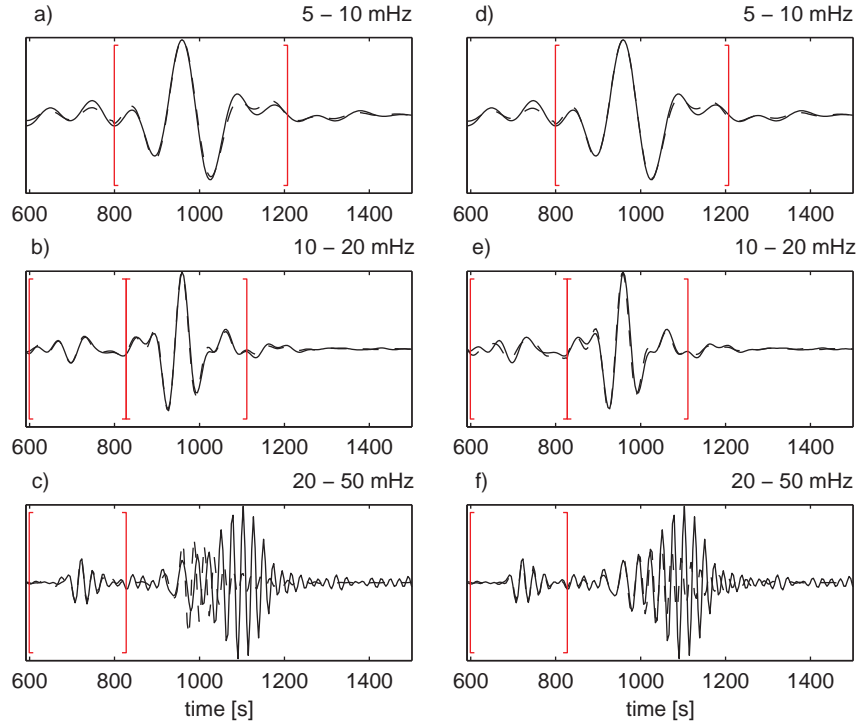
Figure 2.15: *Fit between the data (solid) and the synthetic seismogram (dashed) for the relative misfit (a)-(c) and absolute misfit (d)-(f) for the frequency and time windows indicated in red.*

2.14   Bayesian statistics

In the second part of the NA (Sambridge, 1999b), the ensemble of shear wave velocity models is resampled using importance sampling to construct a conditional posterior probability density function defined as

$$P(\mathbf{m}|\mathbf{d}) = \kappa \rho(\mathbf{m}) L(\mathbf{m}|\mathbf{d}), \tag{2.18}$$

where $\rho(\mathbf{m})$ is the prior probability distribution (depending on the parameterisation, equation 2.6; the model space size; and the forward theory, equations 2.3 and 2.5), κ is a normalisation constant and L is a likelihood function which represents the fit to the observations. For Gaussian error statistics the likelihood is
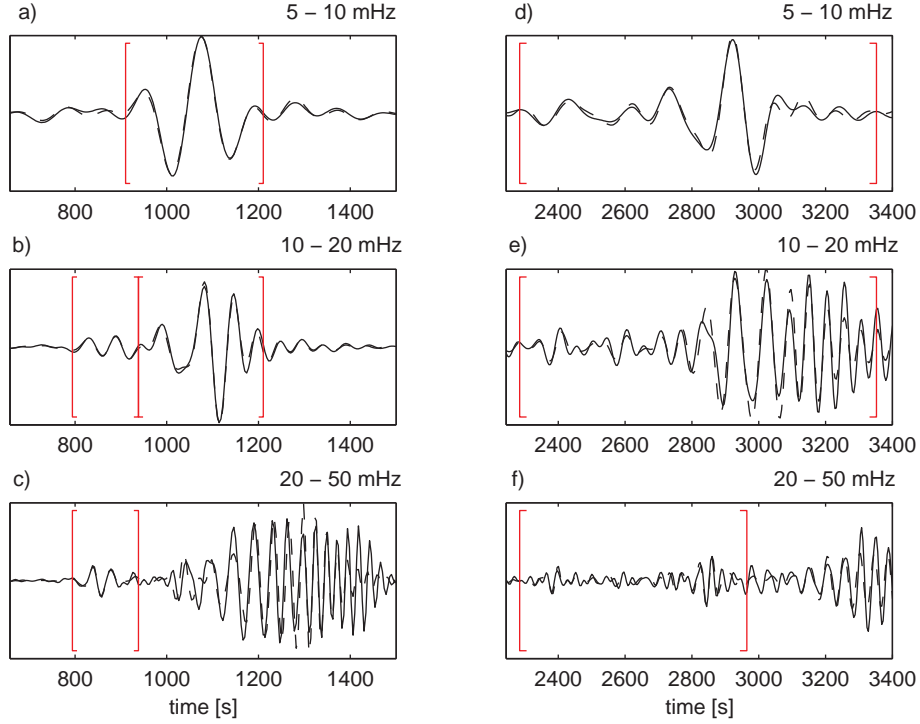
Figure 2.16: *Fit of the data with the synthetic for the (a)-(c) Rayleigh surface waves and (d)-(f) for Love surface waves in three frequency windows: (a),(d) 5-10mHz, (b),(e) 10-20mHz and (c),(f) 20-50mHz. The time windows are indicated in red.*

defined as

$$L(\mathbf{m}|\mathbf{d}) = \exp[-\frac{1}{2}(\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm})] \qquad (2.19)$$

where $\mathbf{C}_d$ is the data covariance matrix describing noise statistics. In this equation we recognise the χ^2 misfit ($\chi^2 = (\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm})$). The misfit M as defined in equation 2.16 is not a χ^2 misfit. For a χ^2 misfit we would need to know the data covariance matrix which is not obvious. The difficulty in obtaining the data covariance matrix $\mathbf{C}_d$ is that there are many sources of noise for real seismograms. Examples of noise are uncertainties in the source location, source mechanism, receiver, scattered waves or in general any waves that are not

accounted for in the theory. Therefore we define the likelihood as

$$L(\mathbf{m}|\mathbf{d}) = \exp(-\frac{1}{2}M/k). \tag{2.20}$$

Where M is the misfit as defined in equation 2.16 and k is a normalisation constant, which is chosen as the minimum misfit. In this way we assume the data covariance matrix is not known and our (scaled) minimum misfit is now one and all other misfits are larger than one. This resembles a χ^2 definition.

The ensemble of models is resampled using random walks through the model space. The number of walks and the number of steps in each walk are tuning parameters for the Bayesian stage of the Neighbourhood Algorithm. We tested different values for the tuning parameters and selected three walks of 500 models each which makes a total of 1500 resampled models, which is sufficient for the convergence to the posterior probability density function. The resampling algorithm can also be used to evaluate Bayesian indicators of any transformed parameters, that are a combination of the original parameters, in an identical manner to the original variables. We introduce phase velocity parameters defined for certain modes and periods, since phase velocities may be obtained by integrating a shear wave velocity model (defined by the coefficients γ_i) using the sensitivity kernels for the specific mode and period using equations 2.5 and 2.6. The sensitivity kernels are calculated for the reference model, which is the best shear wave velocity model that we obtained from AMI. We thus obtain a probability density surface for our original (shear wave velocity) parameters and transformed (phase velocity) parameters. From the probability density surface we can obtain one-dimensional marginals for one original/transformed parameter by integrating over all other original/transformed parameters. The Neighbourhood Algorithm also allows us to obtain two-dimensional marginals, correlation and covariances which show the relations between different parameters. The one-dimensional marginals for the phase velocities are Gaussian shaped, which is why we represent our phase velocity measurements as mean phase velocities and standard deviations.

2.15 Tuning the Neighbourhood Algorithm

The sampling stage of the Neighbourhood Algorithm only needs a couple of tuning parameters; the number of initial models (n_i), the number of iterations (n_{iter}), the number of new models sampled at each iteration (n_s) and the number of best misfit models (n_r) needed to guide the model space search. The initial models are generated at random in the model space. A large number of initial models means a more extensive preliminary search. The number of iterations should be large

enough for convergence of the model space search. The tuning parameters (n_r and n_s) determine how the model space is sampled. A large number for n_s and a small number of n_r means the Neighbourhood Algorithm will converge quickly onto the area's of better fit (but you might miss an area). A large number for n_r and a small number for n_s makes the model space search much broader.

We tested different values for the tuning parameters (n_r and n_s), from a very broad search to a very focused search and we selected $n_r = 5$ and $n_s = 10$ as the tuning parameters that provide us with efficient and reliable phase velocity measurements. We also decided to use 100 initial models (n_i) as a more preliminary search should give a good indication about the region of better fitting models. The number of iterations has a large effect on the computation time needed for the Neighbourhood Algorithm. Therefore, we would prefer a small number of iterations but still sample enough models to be sure of convergence. We performed the waveform inversion for our test dataset using model space searches with 100, 500, 1000, 2000 and 3000 iterations. Figure 2.17 and 2.18 show the difference in phase velocity and standard deviation for the waveform inversion runs. The results are shown for the fundamental and first four higher modes for one specific seismogram, but other seismograms of our test dataset showed similar results. The differences in phase velocity are small and well within the standard deviations. Also, the standard deviations do not change significantly. From this test, we concluded that we could use as few as 100 iterations and still obtain the same phase velocity measurements and standard deviations. However, we selected 500 iterations as a conservative choice, keeping in mind that the 100 iteration case displays the largest differences.

With these tuning parameters ($n_i = 100, n_s = 10, n_r = 5, n_{iter} = 500$) we sample 5100 models (100 initial models, 10 new models per iteration, 500 iterations) for each waveform inversion. In contrast with other direct search methods, the Neighbourhood Algorithm characterizes the whole range of models in the model space, not only the best fitting ones.

2.16   Do we need to resample the ensemble?

The resampling stage of the Neighbourhood Algorithm is a time consuming process, which leads to the question whether we need to resample the ensemble of shear wave velocity models to obtain standard deviations for the phase velocities, or take the approach of Yoshizawa and Kennett (2002) and obtain the standard deviations from the 1000 best models sampled in the first part of the NA. We compared the phase velocity for the best shear wave velocity model in the sampling
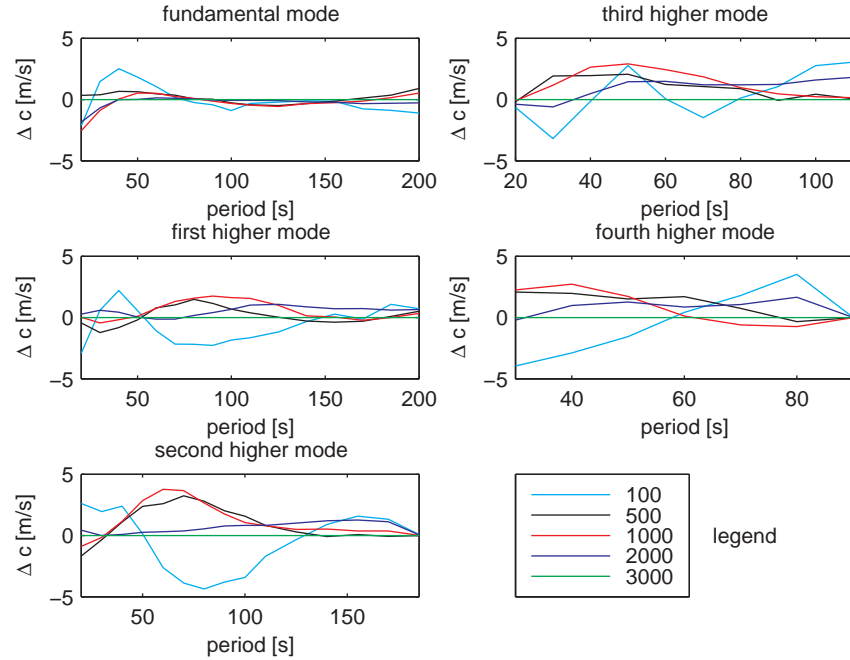
Figure 2.17: *The difference in phase velocity (Δc) for the fundamental (0), first (1), second (2), third (3) and fourth (4) higher mode for the different number of iterations indicated in color, the reference is the run with 3000 iterations*

stage, the average phase velocity of the 1000 best shear wave velocity models in the sampling stage and the most likely phase velocity obtained by the Bayesian stage of the Neighbourhood Algorithm (see figure 2.19). Figures 2.19 and 2.20 give the result for one seismogram. The differences between the phase velocity measurements are larger than one standard deviation from the second higher mode on. We also compared the standard deviations obtained by the Bayesian stage and the number of best shear wave velocity models ($\#M$) obtained from the sampling stage of the Neighbourhood Algorithm (figure 2.20). Even for a very thorough search of the model space (30,000 models) we obtain underestimated standard deviations using only the ensemble of shear wave velocity models from the sampling stage. We also compared the phase velocities and standard deviations for other seismograms and found similar results.

The one-dimensional marginals are projections of the probability density surface onto the corresponding axis (the phase velocity for a certain mode and fre-
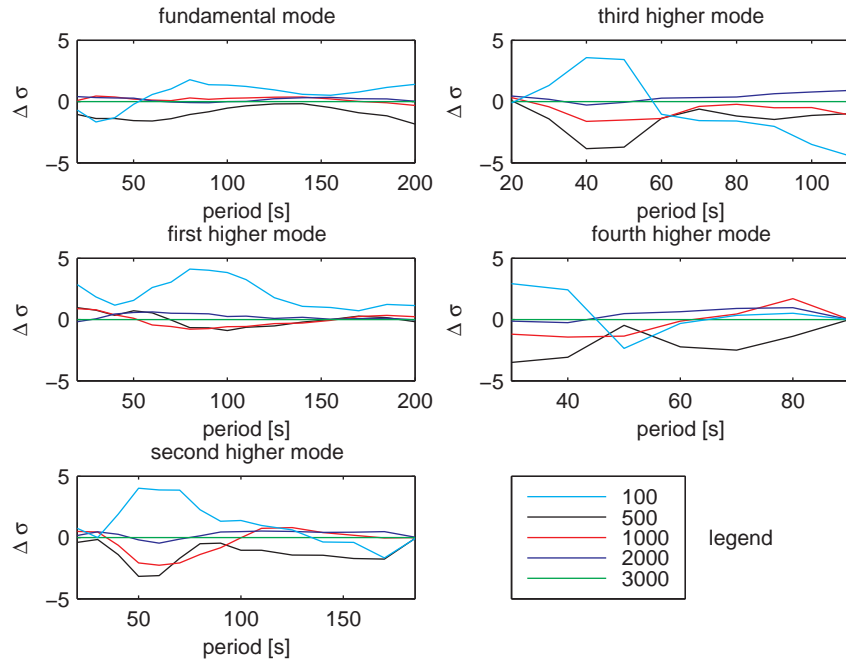
Figure 2.18: *The difference in standard deviation ($\Delta\sigma$) for the fundamental (0), first (1), second (2), third (3) and fourth (4) higher mode for the different iterations indicated in color, the reference is the run with 3000 iterations*

quency) and, thus, they display information on the entire probability density function. The ensemble of shear wave velocity models and their corresponding misfits obtained from the sampling stage of the NA, do not correspond directly to the probability density function and do not give a full indication of the statistical properties of the ensemble. The same was found by Sambridge (1999a, figure 8). The marginal distribution from the NA ensemble differs from the corresponding marginals determined by the resampling algorithm, indicating that the ensemble of models obtained in the sampling stage of the NA is not distributed according to the posterior probability distribution.

This test clearly indicates that we need to calculate the posterior probability density function to obtain meaningful phase velocity measurements and standard deviations.

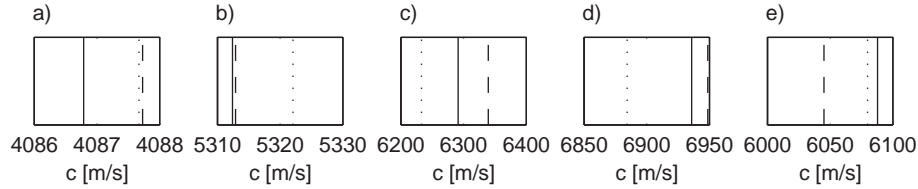Figure 2.19: *Phase velocities of the best model from NA (dashed line), the average model of NA of 1000 models (dotted) and the most likely model from the Bayesian part of the NA (line) for the (a) fundamental mode at 70 s, (b) first higher mode at 70s, (c) second higher mode at 70 s, (d) third higher mode at 60s and (e) fourth higher mode at 35s*
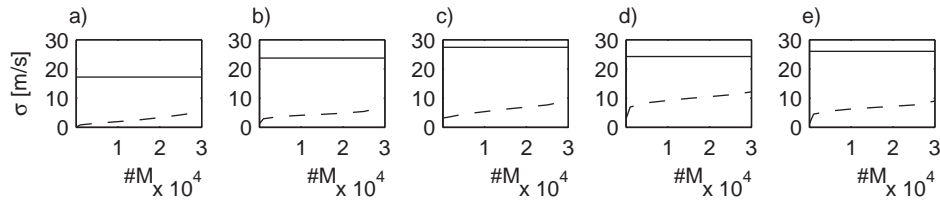
Figure 2.20: *Standard deviation of a number of models (#M) for the (a) fundamental mode at 70 s, (b) first higher mode at 70s, (c) second higher mode at 70 s, (d) third higher mode at 60s and (e) fourth higher mode at 35s. Solid lines indicate the standard deviation from the Bayesian stage of the NA, dashed lines the standard deviations from a number of best fitting models (#M) of the sampling stage of the NA.*

2.17   The prior

As seen in equation 2.18, the posterior probability density function depends on the prior and the likelihood, where the prior depends on the parameterisation, the search boundaries (model space size) and the forward theory. We tested the effect of the search boundaries on the phase velocity measurements and standard deviations. We selected search intervals of $\pm 2.5\%$, $\pm 4\%$, $\pm 7\%$ and $\pm 10\%$ and performed the waveform inversion. The results are shown in figure 2.21 and 2.22. Figure 2.21 shows that the differences in phase velocity are small in general but largest for the largest model space size, which is not surprising since the variation of $\pm 10\%$ around PREM is so large it is questionable whether we should use the Fréchet derivatives. Comparing the phase velocity measurements with the standard deviations (figure 2.22) we find that the differences in phase velocity are well within their standard deviations. The effect of the model space size on the stan-

Figure 2.21: *Differences in phase velocity (Δc) with the size of the model space. Black is a model space of $\pm$ 2.5%, red of $\pm$ 4%, blue of $\pm$ 7% and green of $\pm$10% for (a) the fundamental mode, (b) the first higher mode, (c) the second higher mode, (d) the third higher mode, (e) the fourth higher mode and (f) the fifth higher mode.*
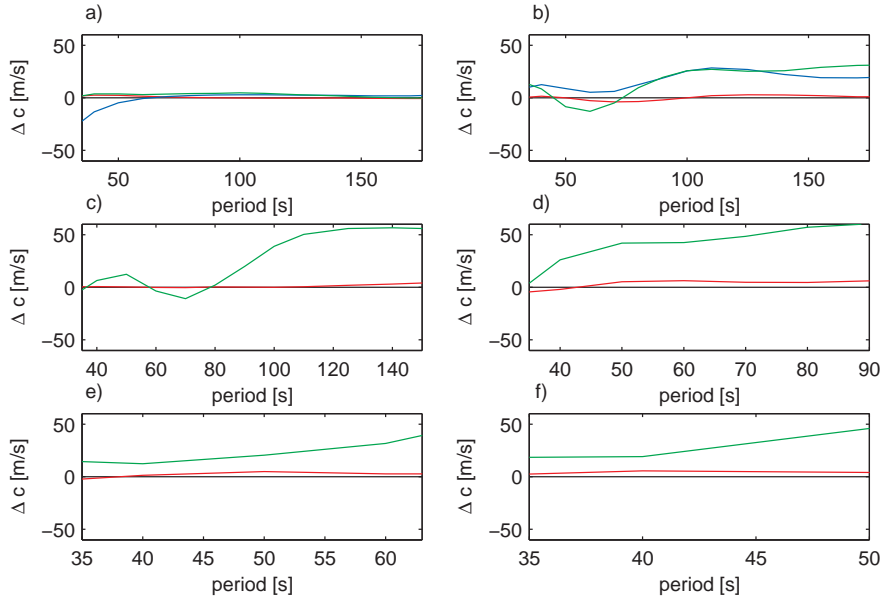
dard deviations however is large (figure 2.21). The standard deviations increase in size with increasing model space size. This is due to Central Limit Theorem, which states that the sum of n independent equally distributed random variables will approach a normally distributed random variable as n increases. The resulting standard deviation (but not the mean) is dependent on the standard deviations of the n random variables. The phase velocity marginals are a sum over all shear wave velocity marginals and thus tend to Gaussian distributions with a robust mean. The standard deviation of the phase velocity marginal depends on the standard deviations of the shear wave velocity marginals, and hence the a priori size of the model space. We are thus not able to obtain absolute standard deviations but the relative uncertainties are consistent between the modes and between seismograms. We compared the standard deviations for different search intervals for the fundamental mode measurements to the standard deviations of Trampert and Woodhouse (2001). They showed that uncertainties obtained by cluster analysis are in agreement with uncertainties obtained for comparing model predictions to
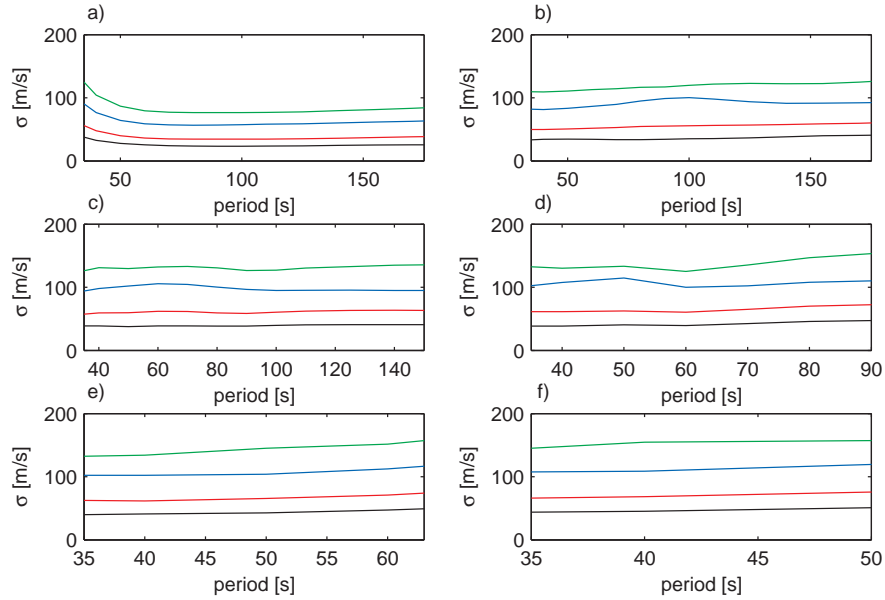
Figure 2.22: *Values of the standard deviation σ with size of the model space. Black is a model space of ± 2.5%, red of ± 4%, blue of ± 7% and green of ±10% for (a) the fundamental mode, (b) the first higher mode, (c) the second higher mode, (d) the third higher mode, (e) the fourth higher mode and (f) the fifth higher mode.*

real seismograms. We finally chose a model space size of $\pm 2\%$, which is large enough for the model space search, since the chosen reference model is the best model from AMI and hence (close to) our best model, and gives us standard deviations for the phase velocity measurements which are close to the standard deviations obtained by cluster analysis. This anchoring will turn the self consistent relative uncertainties between modes in realistic absolute uncertainties.

2.18   How many overtones?

In principle, we can calculate the phase velocity for any mode from the one-dimensional shear wave velocity model. The important question is thus how many modes are constrained by each seismogram. To obtain a measure of the higher mode information in each seismogram, we investigate the unexplained variance

Figure 2.23: *Unexplained variance (V(m)) as a function of the number of modes in the synthetic seismogram (m). The red line gives the 25% range value.*

V as a function of the number of modes m in the synthetic seismogram defined as

$$V(m) = \frac{\sum_{i=1}^{T}[d_i - s_i(m)]^2}{\sum_{i=1}^{T} d_i^2}, m = 0, 1, ..., 30, \tag{2.21}$$

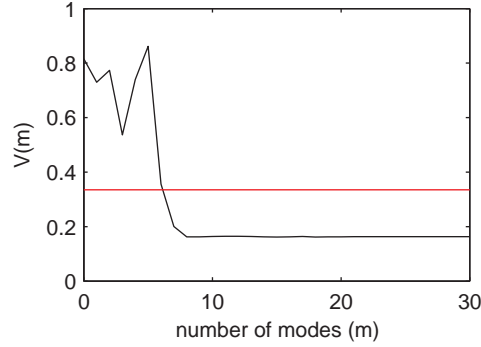where T is the number of time samples in the window (see equation 2.17). m is allowed to vary from the fundamental mode only ($m=0$) to up to 30 higher modes. The unexplained variance is only evaluated in the time-frequency window, which contains the most higher mode information and the least fundamental mode information (figure 2.16b (first time window only), c and f), because we want to obtain a measure of the higher mode information available in the seismogram. The unexplained variance with up to m modes $V(m)$ generally decreases with increasing m. The number of overtones constrained by a seismogram is defined as the smallest number of the modes which brings $V(m)$ below 25% of its range, where the range is defined as the difference between the maximum ($V(0)$) and the minimum ($V(30)$) value. This empirical threshold was chosen after visual inspection of the variance curves for numerous seismograms. Figure 2.23 shows an example. In this case we would select up to six higher modes.

In two cases we decide to measure the fundamental mode only: if we obtain a bad fit for the higher modes ($V(30) > 0.5$), or if there is no significant higher mode information in the seismogram ($V(0) < 0.2$). A final test calculates the unexplained variance ($V(30)$) for all three time-frequency windows. If more than 60% of the seismogram ($V(30) > 0.4$) remains unexplained the seismogram is discarded.

Chapter 3

Global Love wave overtone measurements

Love wave phase velocities for fundamental and higher modes are difficult to measure because the different modes cannot easily be separated. Following Yoshizawa and Kennett (2002), we generate suites of path specific one-dimensional shear wave velocity profiles using the Neighbourhood Algorithm (Sambridge, 1999a). From this family of $O(10^4)$ models both fundamental and higher mode phase velocities with mutually consistent uncertainties are calculated. We have fully automated the method and analysed over forty thousand Love wave seismograms from the GDSN and GEOSCOPE global networks from 1994-2004. Our phase velocity measurements agree remarkably well with previous studies, but we have been able to enlarge the available dataset dramatically. We present global Love wave phase velocity maps (up to the fifth overtone) with unprecedented resolution due to the improved path coverage. Comparing these maps to existing tomographic models, we discern evidence of significant anisotropy in the lower mantle around a depth of 1000 km in the Pacific.

3.1   Introduction

Phase velocity maps have mainly been constructed for fundamental mode surface waves. The sensitivity of fundamental modes, at commonly used periods up to 200 seconds, is limited to the upper 400 km. Surface wave tomography using

such modes, therefore, suffers from a limited depth penetration. The obvious way to increase depth penetration is to add higher mode information. The sensitivities of higher modes extend well below the transition zone and into the lower mantle. Techniques to measure higher mode surface wave phase velocities are mostly based on the separation of modes in the seismogram and can be roughly divided into two groups: methods that use clustering of stations (Nolet, 1975; Cara, 1979) or events (Stutzmann and Montagner, 1993; Beucler et al., 2003) and methods that use single seismograms (Van Heijst and Woodhouse, 1999; Yoshizawa and Kennett, 2002). The disadvantage of using clusters is that the distribution and number of phase velocity measurements is geographically very limited. Van Heijst and Woodhouse (1999) used a mode branch stripping technique to obtain phase velocity measurements of the fundamental and higher mode surface waves. This method is effective for seismograms with longer paths where the modes are reasonably well separated, and hence is difficult to apply to Love waves where fundamental and higher modes travel closely together. Yoshizawa and Kennett (2002) used a fully non-linear waveform inversion to obtain regional path specific multimode dispersion measurements. This method samples the model space for a depth dependent shear wave velocity model that fits the observations best. This best fitting shear wave model, not meant to be a direct representation of the Earth model, is interpreted as providing implicit information on the multimode dispersion for the given source-receiver path. In carefully chosen frequency windows, mode coupling for the first few modes can be restricted (Kennett, 1995). Furthermore, a direct interpretation of the path specific models is only valid for small lateral perturbation along the path, whereas less constrictive conditions apply to the corresponding multimode dispersion curves (Kennett and Yoshizawa, 2002). Yoshizawa and Kennett (2002) obtained approximate standard deviations by comparing the dipersion curves of the 1000 best shear velocity models. From the best fitting model, phase velocities are calculated without any mode separation, in a fully non-linear framework. We develop a fully automated implementation of this approach and improve the uncertainty analysis by calculating complete probability density functions for all phase velocities. This enables us to make multimode dispersion measurements with mutually consistent error estimates. The method is applied to measure fundamental and higher mode Love wave phase velocities on a global scale.

3.2   Method

In principle, the model space search to invert the seismogram for a 1D velocity model could include the full non-linearity of the forward problem. In practise, time constraints force us to limit the search around a reference model. The use of Fréchet derivatives introduces a dependence of the final result upon the reference model which we obtain from the Automated Multimode Inversion method (AMI, Lebedev et al. (2005)). This is a waveform inversion technique that uses multiple time and frequency windows to obtain a shear wave velocity model that explains both the fundamental mode as well as the higher modes in the seismogram. The safeguards build into AMI guarantee that the JWKB approximation is valid for all seismograms which provide measurements. The shear wave velocity model from AMI is used to calculate all eigenfunctions and Fréchet derivatives for the model space search. For the search itself we use the Neighbourhood Algorithm (NA, Sambridge (1999a,b)). NA involves two separate steps. The first step is a guided Monte Carlo sampling. A relative least squares misfit M guides the model space search to areas of better fit where the difference between the data (d) and the synthetic seismograms (s) is considered relative to the maximum amplitude in a specific window ($d_{max}(\mathsf{w})$):

$$M = \sum_{\mathsf{w}=1}^{N} \sum_{i=1}^{L} \frac{(d_i - s_i)^2}{(d_{max}(\mathsf{w}))^2} \tag{3.1}$$

where N is the number of time-frequency windows. and L is the number of time samples in each window. The time-frequency windows are chosen such that the fundamental mode is mainly excited in the lower frequency windows (5-20 mHz) and the higher modes are mainly excited in the higher frequency windows (20-50 mHz)(Figure 3.1a-c). The lower time limit in the first two windows corresponds to the group arrival with a velocity of 3.8 km/s, which should capture the complete Love wave fundamental mode as long as the perturbations with respect to PREM are not too large. The lower time limit of the third window corresponds to a group velocity of 4.3 km/s to include the higher modes and exclude the fundamental mode. The higher time limit in all windows depends on the epicentral distance, below 35° the time is set just before the arrival of the S wavetrain. Between 35° and 70° the time is set just after the S and before the SS wavetrain etc. For the determination of the left window boundaries, the S and multiple S arrival times are computed using the AK135 model (Kennett et al., 1995).
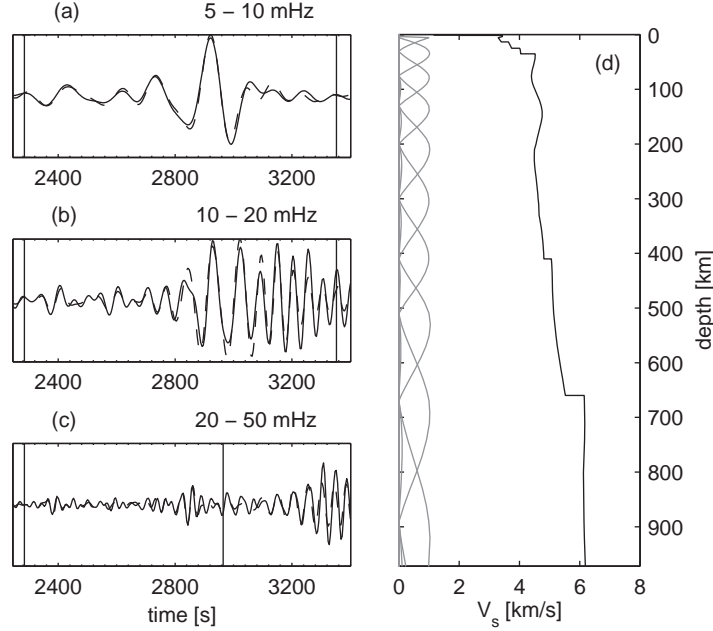
Figure 3.1: *Fit of the data with the synthetics (dashed lines) in three time-frequency windows (a:5-10 mHz, b:10-20 mHz, c:20-50 mHz) and the corresponding shear velocity model for the best fitting model with the spline parameterization (d). The time-windows are indicated by the vertical bars.*

The synthetic seismogram $s(\omega)$ in the frequency domain is calculated as a sum of modes m using the JWKB approximation:

$$s(\omega) = \sum_m A_m(\omega) exp[i\omega\Delta/(C_m^0(\omega) + \delta C_m(\omega))], \qquad (3.2)$$

where $A_m(\omega)$ is the complex amplitude of the modes, Δ is the source-receiver distance, $C_m^0(\omega)$ is the Love wave phase velocity in the reference model and $\delta C_m(\omega)$ is a Love wave phase velocity perturbation

$$\delta C_m(\omega) = \int_0^a [\frac{\partial C_m^0(\omega)}{\partial V_s(r)}\delta V_s(r) + \frac{\partial C_m^0(\omega)}{\partial \rho(r)}\delta\rho(r)]dr. \qquad (3.3)$$

The Fréchet derivatives $\partial C_m^0(\omega)/\partial V_s(r)$ and $\partial C_m^0(\omega)/\partial \rho(r)$ relate the change in shear wave velocity and density from the reference model to changes in the phase velocity. a is the radius of the Earth. The density perturbations are of secondary

importance and are scaled to the shear wave velocity perturbations. The scaling relation ($\xi = \partial ln\rho/\partial lnV_s$) is taken from Deschamps et al. (2001). We checked that different scaling relations did not alter our results. The attenuation is that of PREM. The change in shear wave velocity $\delta V_s(r)$ is parameterized through a set of 1-D basis functions $h_i(r)$:

$$\delta V_s(r) = \sum_{i=1}^{12} \gamma_i h_i(r), \tag{3.4}$$

where the γ_i are the coefficients to be found in the model space search. The boundaries of the model space are chosen such that $\pm 2\%$ changes (justification below) are allowed around the reference model.

The functions $h_i(r)$ are twelve natural cubic spline basis functions that span the shear velocity model in the crust, upper mantle and lower mantle up to a depth of 1500 km (figure 3.1d). The basis functions are spaced more densely in the crust and upper mantle to match the expected depth resolution of surface waves. We typically sample 5100 models per seismogram (n_s=10, n_r=5, 500 iterations, 100 initial models).

From this first sampling, no stable measurements can be estimated. The second part of the Neighbourhood Algorithm (Sambridge, 1999b) resamples the initial ensemble of models and constructs a conditional posterior probability density function given the seismograms **d**.

$$P(\mathbf{m}|\mathbf{d}) = \kappa\rho(\mathbf{m})L(\mathbf{m}|\mathbf{d}), \tag{3.5}$$

where $\rho(\mathbf{m})$ is the prior probability distribution (depending on the parameterization, eq. 3.4; search boundaries, $\pm 2\%$; and the forward theory, eq. 3.2 and 3.3) and $L(\mathbf{m}|\mathbf{d}) = exp(-\frac{1}{2}M/c)$ is a likelihood function which represents a fit to the observations. M is defined in eq. 3.1. κ and c are normalization constants.

The statistical properties of the ensemble are defined in a Bayesian framework and are evaluated using Monte Carlo integration (Sambridge, 1999b). We resample using 1500 models which is sufficient for convergence of the integrals. The results are presented as one-dimensional marginal probabilities for each model parameter by integrating over all other dimensions of the probability density function. The marginal probability densities for the coefficients γ_i (eq. 3.4) can easily be transformed into marginal probability densities for phase velocities of any mode at any period using equations (3.3) and (3.4). Because of the central limit theorem, we observe that the marginals for phase velocities are close to Gaussian, which can conveniently be described by its mean and standard deviation.

3.2.1   How Many Overtones?

In principle, we can calculate the phase velocity of any mode from the 1D Earth model, the important question is thus how many modes are constrained by each seismogram. To obtain a measure of the higher mode information in each seismogram, we investigate the unexplained variance V as a function of the number of modes K in the synthetic seismogram:

$$V(K) = \frac{\sum_{i=1}^{L}[d_i - s_i(K)]^2}{\sum_{i=1}^{T} d_i^2}, K = 0, 1, ..., 30, \qquad (3.6)$$

K is allowed to vary from the fundamental mode only (K=0) to up to 30 higher modes. The unexplained variance is only evaluated in the time-frequency window, which contains the most higher mode information and the least fundamental mode information (figure 3.1c), because we want to obtain a measure of the higher mode information available in the seismogram. The unexplained variance with up to K modes $V(K)$ generally decreases with increasing K. The number of overtones constrained by a seismogram is defined as the smallest number of the modes which brings $V(K)$ below 25% of its range, where the range is defined as the difference between the maximum ($V(0)$) and the minimum ($V(30)$) value. This empirical threshold was chosen after visual inspection of the variance curves for numerous seismograms.

In two cases we decide to measure the fundamental mode only: if we obtain a bad fit for the higher modes ($V(30) > 0.5$), or if there is no significant higher mode information in the seismogram ($V(0) < 0.2$). A final test calculates the unexplained variance ($V(30)$) for all three time-frequency windows. If more than 60% of the seismogram ($V(30) > 0.4$) remains unexplained, the seismogram is discarded. From a total of 310,000 seismograms, we measured 14.5% fundamental mode, 11.1% first, 10.1% second, 7.7% third, 4.8% fourth and 2.7% fifth higher mode dispersion curves.

3.3   Dispersion Measurements

We measured minor arc Love wave phase velocities using data from the GEO-SCOPE and GDSN global networks from 1994 to 2004. Figure 3.1a-c shows an example of the waveform fit obtained in three different time-frequency windows for the best fitting model (figure 3.1d) given by the model space search. The corresponding phase velocity measurements are shown in figure 3.2 for the fundamental mode and the first, second and third higher modes. The measurements agree
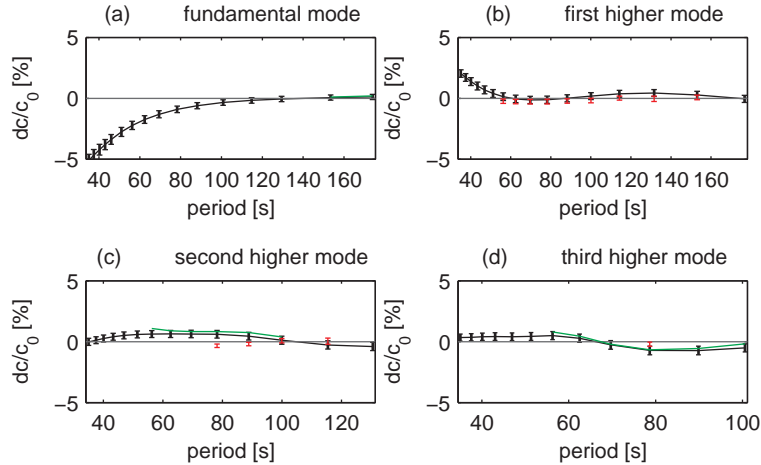
Figure 3.2: *Comparison of phase velocity measurements with respect to PREM for the fundamental mode (a), the first higher mode (b), the second higher mode (c) and the third higher mode(d). Indicated in red are measurements of Van Heijst and Woodhouse (1999) with standard deviations according to cluster analysis and in green measurements with AMI as in (Lebedev et al., 2006).*

well with existing measurements made by Van Heijst and Woodhouse (1999) (figure 3.2 and 3.3) and those made with AMI by Lebedev et al. (2006)(figure 3.2). We calculated uncertainties for Van Heijst's measurements by cluster analysis.

Our standard deviations do not depend on the quality of the original seismogram because of the definition of the misfit function (eq. 3.1); they depend mainly on the size of the modelspace. This is again a result of the central limit theorem, which states that the sum of n independent equally distributed random variables will approach a normally distributed random variable as n increases. The resulting standard deviation (but not the mean) is dependent on the standard deviations of the n random variables. The phase velocity marginals are a sum over all shear wave velocity marginals and thus tend towards Gaussian distributions, with a robust mean. The standard deviation of the phase velocity marginal depends on the standard deviations of the shear wave velocity marginals, and hence the *a priori* size of the modelspace. We can thus not obtain absolute uncertainties from NA, but the relative uncertainties betweeen modes are self consistent. We define the boundaries of the modelspace ($\pm 2\%$ around the AMI reference model) such that the resulting standard deviations for the fundamental mode match the standard
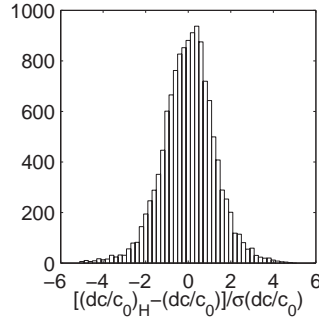
Figure 3.3: *Histogram comparing overtone phase velocity measurements (first, second and third overtone branch) of Van Heijst and Woodhouse (1999) ($(dc/c_0)_H$) with this study (dc/c_0) scaled by our standard deviations ($\sigma(dc/c_0)$). 16,756 ($\approx 65\%$) out of a total of 25,908 overlapping measurements fall within one standard deviation.*

mode	period [s] in PREM	number of measurements	$\chi^2_{initial}$	χ^2_{final}
3	46.95	24,102	4.21	3.28
4	62.77	15,065	2.00	1.75
5	56.29	8,515	2.16	1.96

Table 3.1: *Details for the phase velocity maps of figure 3.4*

deviations obtained by cluster analysis by Trampert and Woodhouse (2001). They showed that uncertainties for cluster analysis are in agreement with uncertainties obtained for comparing model predictions to real seismograms. The anchoring will thus turn the self consistent relative uncertainties between modes in realistic absolute uncertainties.

A compact representation of the measurements is to construct phase velocity maps. Figure 3.4b,d and f show global minor arc phase velocity maps for Love waves for the third, fourth and fifth overtone. The phase velocity maps were expanded on a spherical harmonic basis up to degree and order 20, following the same procedure as described in Trampert and Woodhouse (1995). The number of measurements, initial and final χ^2 of the phase velocity maps are given in table 3.1. The derivative damping was chosen to allow an easy comparison with predictions for the model S20RTS of Ritsema et al. (1999) where the crust was added using CRUST5.1 (Mooney et al., 1998) (Figure 3.4a,c and e). Even for
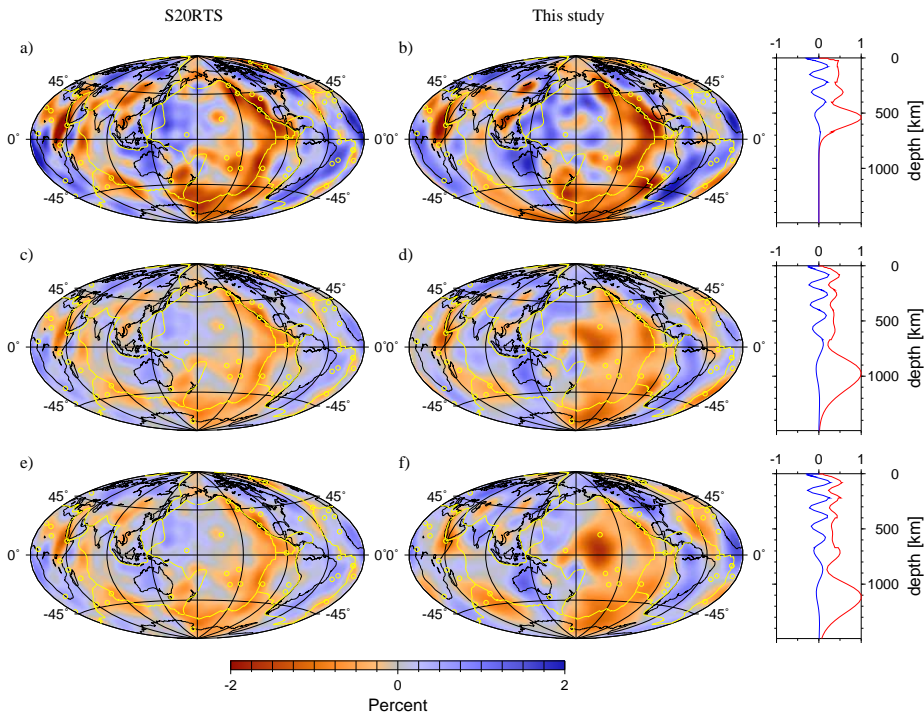
Figure 3.4: *Deviations in Love wave phase velocity dc/c_0 calculated for model S20RTS and the new Love wave measurements. (a),(b) third higher mode at 46.95s; (c),(d) fourth higher mode at 62.77s; (e),(f) fifth higher mode at 56.29s. To the right of the figure the sensitivity curves (red for V_s and blue for ρ) for the specific modes are displayed.*

the fifth overtone, the ray density is higher and more uniform then that given by Trampert and Woodhouse (1995). The resolution is then at least as good as that of figure 7a of Trampert and Woodhouse (1995). Some phase velocity maps (with bulk sensitivities shallower than 1000 km) agree well with the S20RTS prediction see (figure 3.4a and b), even though the S20RTS model does not contain Love wave information. For higher modes, with main sensitivities around 1000 to 1500 km, there is a discrepancy between the Love wave phase velocity maps and the S20RTS phase velocity maps (Figure 3.4c-d and 3.4e-f) in the Pacific. The S20RTS model is based on mainly Rayleigh equivalent surface waves. This discrepancy indicates a difference between SH and SV around 1000 to 1500 km, an indication of anisotropy.

3.4   Concluding Remarks

We measured higher mode Love wave phase velocities up to the fifth higher mode with corresponding uncertainties using a new, fully automatic procedure. The use of a model space sampling allows us to derive mutually consistent estimates of relative standard deviations between different overtone branches and from measurement to measurement. The phase velocities agree well with existing measurements (Van Heijst and Woodhouse, 1999; Lebedev et al., 2006), especially for the fundamental modes. The higher modes agree well within their standard deviation. The differences between the different techniques are caused by different theoretical formulations; branch stripping (van Heijst) or multiple frequency and time windows (Lebedev). The resulting phase velocity maps agree well with phase velocity maps predicted by the model S20RTS (Ritsema et al., 1999). Love wave phase velocity maps with high sensitivities between 1000 and 1500 km differ from the Rayleigh wave based S20RTS, giving an indication of anisotropy in the Pacific around the Pacfic superplume. The use of high quality overtone measurements should improve the resolution in the mid-mantle where the differences between existing models are largest (Romanowicz, 2003).

3.5   Acknowledgements

Chapter 4

Global anisotropic phase velocity maps for higher mode Love and Rayleigh waves

It is well established that the Earth's uppermost mantle is anisotropic, but there are no clear observations of anisotropy in the deeper parts of the mantle. Surface waves are well suited to observe anisotropy since they carry information about both radial and azimuthal anisotropy. Fundamental mode surface waves, for commonly used periods up to 200 s, are sensitive to structure in the first few hundred kilometers and therefore do not provide information on anisotropy below. Higher mode surface waves have sensitivities that extend to and beyond the transition zone, and should thus give insight about azimuthal anisotropy at greater depths. We have measured higher mode Love and Rayleigh phase velocities using a model space search approach, which provides us with consistent relative uncertainties from measurement to measurement and from mode to mode. From these phase velocity measurements, we constructed global anisotropic phase velocity maps. Prior to inversion, we determine the optimum relative weighting for anisotropy. We present global azimuthal phase velocity maps for higher mode Rayleigh waves (up to the sixth higher mode) and Love waves (up to the fifth higher mode) with corresponding average model uncertainties. The anisotropy we derive is robust within the uncertainties for all modes. Given the ray theoretical sensitivity kernels of Rayleigh and Love wave modes, the source of anisotropy is complex, but

mainly located in the asthenosphere and deeper. Our models show a good correspondence with other studies for the fundamental mode, but we have been able to achieve higher resolution.

4.1 Introduction

It is widely established that the Earth's upper mantle is anisotropic. The first observation of radial anisotropy was the discrepancy between Rayleigh and Love wave dispersion observed by Anderson (1961), Aki & Kaminuma (1963) and McEvilly (1964). Anisotropy was also observed in the azimuthal dependence of P_n velocities (Hess, 1964) and S-wave splitting in teleseismic SKS waves (Vinnik and Romanowicz, 1989). The first observation of azimuthal anisotropy, the azimuthal variation of phase velocities, was noted by Forsyth (1975) in the Pacific ocean. Radial and azimuthal anisotropy are both observed by surface waves, which is why these waves are well suited to study anisotropy. Radial and azimuthal anisotropy are the result of the same underlying anisotropy of the Earth's interior and were linked in a common mathematical framework by Montagner & Nataf (1986). The alignment (lattice preferred orientation or LPO) of intrinsically anisotropic minerals under strain in the mantle is assumed to be the major cause of upper mantle anisotropy (Karato, 1998a; Montagner, 1998). Anisotropy is thus an indicator of mantle deformation and flow. Therefore, it is critical to image anisotropy to understand the dynamics of the mantle.

Fundamental mode surface waves are well suited to provide information about anisotropy in the upper mantle (Tanimoto and Anderson, 1984; Nataf et al., 1984; Montagner and Tanimoto, 1991; Ekström and Dziewonski, 1998). The sensitivity of fundamental mode surface waves for commonly used periods up to 200 s is however limited to the upper 400 km of the Earth's mantle. The use of higher mode surface waves should increase our knowledge of anisotropy into the lower part of the upper mantle and the upper part of the lower mantle due to their greater sensitivity at depth compared to fundamental mode surface waves. In the last few years higher mode surface waves have been added to studies of anisotropy (Debayle and Kennett, 2000; Simons et al., 2002; Trampert and van Heijst, 2002; Maggi et al., 2006; Beucler and Montagner, 2006). The number of higher modes used in these studies varies due to the difficulty of measuring higher mode phase velocity, especially for Love waves since the higher modes arrive simultaneously with the fundamental mode. Trampert & van Heijst (2002) and Beucler & Montagner (2006) use phase velocities up to the second higher mode and Debayle & Kennett (2000) and Maggi et al. (2006) use phase velocities up to the fourth higher

mode. The number of measurements are often few (Debayle and Kennett, 2000) sometimes imposed by the clustering of events (Beucler and Montagner, 2006; Maggi et al., 2006; Sebai et al., 2006) and the geographical coverage is limited.

This study presents global azimuthal anisotropic phase velocity maps for fundamental and higher mode Love and Rayleigh waves up to the sixth higher mode consisting of a large number of measurements with consistent standard deviations. The phase velocities were measured using a model space search approach (Visser et al., 2007a; Yoshizawa and Kennett, 2002) which provides realistic consistent uncertainties on the phase velocity measurements. Following Trampert & Woodhouse (2003), we determine the optimum relative weighting of anisotropy prior to inversion and present global azimuthal anisotropic phase velocity maps up to the fifth higher mode for Love and up to the sixth higher mode for Rayleigh.

Finally, we analyse the resolution of the azimuthal anisotropic phase velocity maps and look at spectral leakage and trade-offs in particular. Spectral leakage is the effect of mapping small-scale structure not accounted for in the model expansion into the inverted low-degree structure. It arises as a result of uneven data coverage (Snieder et al., 1991). We suppress spectral leakage by using Laplacian damping, which increases the damping with increasing degree. This process effectively decreases the spectral leakage (Spetzler and Trampert, 2003) but also decreases resolution for higher and higher degrees. By looking at the off-diagonal terms of the resolution matrix, we find that the trade-off between parameters remains acceptably small.

4.2   Phase velocity measurements

We follow the approach of Yoshizawa & Kennett (2002) and measure phase velocities using a model space approach (Visser et al., 2007a). In principle, the model space search to invert for a 1-D velocity model could include the full non-linearity of the forward problem. This is very time consuming and therefore we chose to linearize the forward problem by centering the model space search around a reference model and using the Fréchet derivatives of this reference model to calculate the synthetic seismograms. This inherently introduces a dependence on the chosen reference model and the requirement that the chosen reference model should be close to our final model. We use the automated multimode inversion (AMI, Lebedev et al, 2005) to obtain a reference model for the model space search. AMI is a non-linear waveform inversion in multiple time and frequency windows which obtains the best shear wave velocity model that fits the seismogram. The time and frequency windows are chosen such that both the

Figure 4.1: *Minor arc Rayleigh wave azimuthal coverage for first higher mode (a), second higher mode (b), third higher mode (c), fourth higher mode (d), fifth higher mode (e) and sixth higher mode (f).*

fundamental mode and the higher modes are fitted. AMI also applies strict data quality criteria and ensures the validity of the JWKB approximation. The shear velocity model from AMI is close to the best shear wave velocity model which we find in the model space search. Differences between both models are small and largely due to the use of different parameterizations for the shear wave velocity models. AMI uses around 18 boxcar and triangle functions up to 1500 km and we use 12 natural cubic spline functions that span the crust, upper mantle and lower mantle up to a depth of 1500 km. We have experimented with the number and position of the spline functions and found that this parameterization is sufficient to resolve up to the fifteenth higher mode. We do not expect to resolve more than
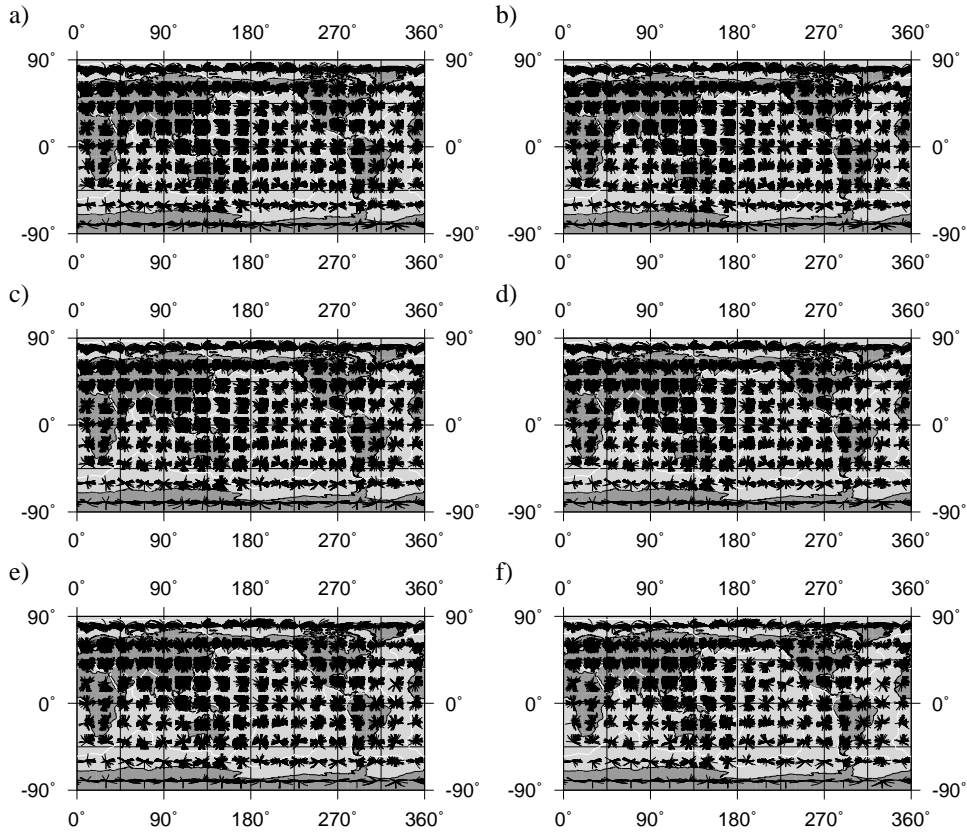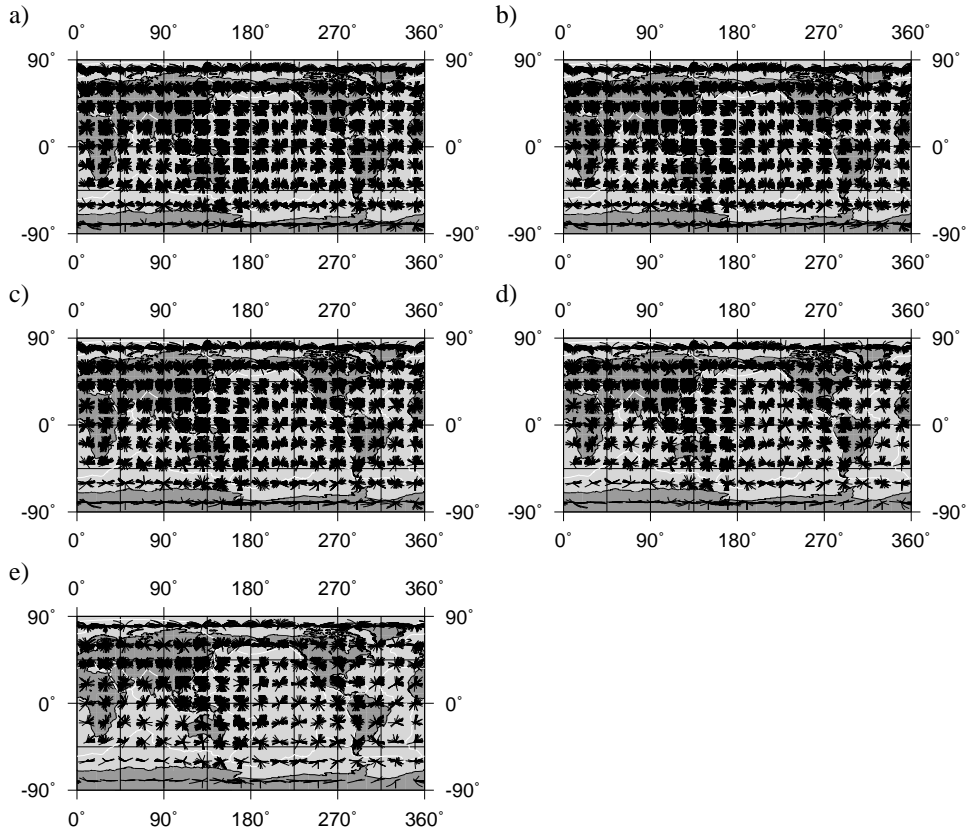
Figure 4.2: *Minor arc Love wave azimuthal coverage for first higher mode (a), second higher mode (b), third higher mode (c), fourth higher mode (d) and fifth higher mode (e).*

six higher modes which makes this parameterization more than sufficient for our purpose. The parameterization is more dense in the crust and upper mantle to match the expected depth resolution of surface waves. The difference between AMI and the model space search is that the first gives us one best fitting shear wave velocity model and the second gives us the whole ensemble of shear wave velocity models compatible with the seismograms which enables us to determine uncertainties. If we were only interested in the best fitting model, AMI would be sufficient. It is important to understand that AMI solves an ill-posed non-linear inverse problem. A solution is found by careful regularisation, but other solutions exist compatible within the data errors. A Monte Carlo search around the AMI solution finds all those other models and allows us to determine meaningful un-

windows (t,ω)		
f (mHz)	Rayleigh (km/s)	Love (km/s)
5 - 10	3.7 ± 0.75	b - 3.8
10 - 20	3.75 ± 0.55	b - 3.8
	b - 4.3	
20 - 50	b - 4.3	b - 4.3

Table 4.1: *Definition of frequency-time windows for Rayleigh and Love seismograms. The time windows are defined using the group velocity (km/s).*

certainties for the best fitting model. Usually, a model space search is used to solve highly non-linear problems. We use it to map the nullspace of a linearized problem instead.

For the model space search we use the Neighbourhood Algorithm (NA, Sambridge, 1999a). We search for the best fitting shear wave velocity model using a least squares misfit between the data and the synthetic seismogram defined in multiple time and frequency windows, see table 4.1. The frequency and time windows are chosen such that the fundamental and higher modes are included in the windows. For Love waves, it is not possible to separate the fundamental and higher modes since the group velocities are similar, thus we use a single time window for both waveforms. For Rayleigh waves, we separated the fundamental and higher mode waveforms in different time windows. The time b in table 4.1 depends on the epicentral distance, below 35° the time is set just before the arrival of the S wavetrain, between 35° and 70° the time is set just after the S and before the SS wavetrain etc. The synthetic seismograms are calculated using the JWKB approximation and the Fréchet derivatives which relate the change in compressional wave velocity, shear wave velocity and density from the reference model to a change in phase velocity. The compressional wave perturbations and the density perturbations are of secondary importance, thus we decided to scale these perturbations to the shear wave velocity perturbations thereby reducing the amount of parameters needed for the model space search. This choice significantly speeds up the model space search. For Rayleigh waves, we scale the compressional wave perturbations to the shear wave velocity perturbations with the scaling relation ($\mathcal{R} = \partial V_s/\partial V_p$) from Ritsema & van Heijst (2002). $\mathcal{R}$ increases linearly from 1.25 at the surface to 3.0 at the CMB. For Love and Rayleigh waves, the density perturbations are scaled to the shear wave perturbations using ($\xi = \partial ln\rho/\partial lnV_s$) from Deschamps et al. (2001). ξ varies between -0.1 to 0.2. The attenuation model employed is that of PREM (Dziewonski and Anderson, 1981).

The model space search provides us with an ensemble of shear wave velocity models and their fit to the data. This ensemble is resampled and transformed into a probability density surface in the second part of the NA (Sambridge, 1999b). The resampling algorithm can also be used to evaluate Bayesian indicators of any transformed parameters, that are a combination of the original parameters, in an identical manner to the original variables. We introduce phase velocity parameters defined for certain modes and periods since phase velocities may be obtained by integrating a shear wave velocity model using the sensitivity kernels for the specific mode and period. The sensitivity kernels are calculated for the reference model, which is the best shear wave velocity model that we obtained from AMI. We thus obtain a probability density surface for our original (shear wave velocity) parameters and transformed (phase velocity) parameters. From the probability density surface we can obtain one-dimensional marginals for each original/transformed parameter by integrating over all other original/transformed parameters. The advantage of our approach is that now we are able to solve for phase velocities for multiple modes and periods without having to separate the modes and without too much computation time. The shape of the phase velocity marginals is Gaussian, and therefore, we represent the one-dimensional marginals as a mean phase velocity and a standard deviation. The phase velocities are presented as changes with respect to PREM for convenience.

In the resampling process, we introduced specific phase velocity parameters which are obtained by integrating each resampled shear wave velocity model using the corresponding sensitivity kernel. In theory, we could obtain phase velocities for every higher mode and period. In practise, we know that not all higher modes are constrained by the seismogram. To evaluate the number of modes constrained in each seismogram we calculated the unexplained variance which is defined as the least squares misfit between the data and the synthetic normalized by the data. The unexplained variance is calculated as a function of the number of modes used in the synthetic seismogram calculation. For a perfect match between data and synthetic, the unexplained variance is zero. Since we calculate the unexplained variance only in the higher mode windows (table 4.1: two windows for Rayleigh, the middle and highest frequency band; and one window for Love, the highest frequency band), the unexplained variance is unity for a fundamental mode synthetic seismogram and should decrease to zero for a perfect full synthetic seismogram. The unexplained variance therefore is a function which, in practise, has its maximum for the fundamental mode only and decreases to some constant value for increasing higher modes. This determines the number of modes needed to explain the specific seismogram. The number of higher modes we measure

	Number of measurements	
	Rayleigh	Love
fundamental mode	63,628	45,179
first higher mode	54,035	34,859
second higher mode	52,457	31,704
third higher mode	48,762	24,102
fourth higher mode	40,606	15,065
fifth higher mode	31,637	8,514
sixth higher mode	21,626	

Table 4.2: *Number of minor arc phase velocity measurements for Rayleigh and Love wave fundamental and higher modes.*

is defined as the smallest number of higher modes which brings the unexplained variance below 25% of its range, where the range is defined as the difference between the largest and the smallest unexplained variance. Further details may be found in Visser et al. (2007).

The unexplained variance is also used to check when we obtain an overall bad fit for the higher modes and whether the seismogram contains no significant higher mode information. In such cases we decide to only measure the fundamental mode. Finally, if we obtain a bad fit for the whole seismogram (unexplained variance larger than 0.4), we discard the measurements.

We measured phase velocities for fundamental and higher mode Love and Rayleigh waves for seismograms that were recorded at the stations of the GDSN and GEOSCOPE networks from 1994 to 2004. The azimuthal coverage for the higher mode Rayleigh and Love wave measurements are shown in figures 4.1 and 4.2. We do not show the azimuthal coverage for the fundamental mode since it is similar to the minor arc coverage given by Trampert & Woodhouse (2002) for Rayleigh and Love waves. The number of measurements obtained for the fundamental and each of the higher modes is shown in table 4.2. We obtain the highest number of measurements for the fundamental mode followed by the first, second, third etc. higher modes. As explained above, the unexplained variance will decrease as the number of higher modes increases. For a seismogram with less higher mode information the unexplained variance decreases more rapidly and the number of higher modes we decide to measure is less. Also, noisier seismograms will lead to less higher modes that will be measured due to the faster decrease of the unexplained variance to an almost constant value. The almost constant value for higher modes indicates that we are trying to measure overtones not constrained

by the seismogram which is why we have to restrict the number of higher modes
we measure. Finally, the number of seismograms with significant second higher
mode information will be less than the number of seismograms with significant
first higher mode information and so on. This is mainly due to the smaller am-
plitudes of the higher modes which makes the contribution to the unexplained
variance smaller. We also obtain more measurements for Rayleigh than for Love
waves, because of the higher noise levels for Love wave seismograms.

4.3   Azimuthal anisotropy

In a slightly anisotropic medium the azimuthal dependence of the local phase
velocities of Rayleigh and Love surface waves is described as (Smith & Dahlen,
1973, 1975; Romanowicz & Snieder, 1998; Larson et. al., 1998).

$$\frac{dc}{c_0}(\omega, \psi) = \quad \alpha_0(\omega) + \alpha_1(\omega)\cos(2\psi) + \alpha_2(\omega)\sin(2\psi)$$
$$+\alpha_3(\omega)\cos(4\psi) + \alpha_4(\omega)\sin(4\psi), \tag{4.1}$$

where dc/c_0 is the relative phase velocity perturbation with respect to a spheri-
cally symmetric Earth model, ω the radial frequency and ψ is the azimuth along
the path. We follow the approach of Trampert & Woodhouse (2003) where the
local phase velocity perturbation is expanded in terms of generalized spherical
harmonics. This reduces equation 4.1 to

$$\mathbf{d} = \mathbf{Gm}. \tag{4.2}$$

Here $\mathbf{d}$ are the path-averaged phase velocity measurements, $\mathbf{m} = (\mathbf{m}_0, \mathbf{m}_2, \mathbf{m}_4)^T$
is the model vector corresponding to the spherical harmonic coefficients of the
0ψ, 2ψ and 4ψ terms. $\mathbf{G} = diag(\mathbf{G}_0, \mathbf{G}_2, \mathbf{G}_4)$ is the block diagonal matrix of
the path-averaged spherical harmonics, for the 0ψ, 2ψ and 4ψ terms. The number
of unknowns is $(L+1)^2$ for the 0ψ terms, $(2L+6)(L-1)$ for the 2ψ terms and
$(2L+10)(L-3)$ for the 4ψ terms. We choose L=40 for the isotropic term (0ψ)
and L=20 for the azimuthal terms (2ψ,4ψ), resulting in 3405 unknowns.

   The inverse problem is solved by minimizing the cost function

$$C = (\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm}) + \mathbf{m}^T \mathbf{C}_m^{-1}\mathbf{m}, \tag{4.3}$$

where $\mathbf{C}_d$ is the diagonal data covariance matrix which consists of the squared
standard deviations of the phase velocity measurements which are obtained from

χ^2	N-M							
	7500	10,000	15,000	20,000	30,000	40,000	50,000	65,000
1.0	0.029	0.025	0.021	0.018	0.015	0.013	0.011	0.010
1.5	0.044	0.038	0.032	0.027	0.022	0.019	0.017	0.015
2.0	0.059	0.051	0.042	0.036	0.029	0.025	0.023	0.020
2.5	0.073	0.063	0.053	0.045	0.037	0.032	0.029	0.025
3.0	0.088	0.076	0.063	0.054	0.044	0.038	0.034	0.030
3.5	0.103	0.089	0.074	0.063	0.051	0.044	0.040	0.035
4.0	0.117	0.101	0.084	0.072	0.059	0.051	0.048	0.040

Table 4.3: *The difference in χ^2 at the 99% significance level determined by the F-test as a function of the χ^2 and the number of independent parameters (N-M).*

the model space search. $\mathbf{C}_m$ is the diagonal model covariance, used to impose Laplacian smoothing. In its partitioned form the expressions are

$$(C_{m_o})_{jj} = \frac{1}{\lambda} \frac{1}{[l(l+1)]^2} \tag{4.4}$$

$$(C_{m_2})_{jj} = \frac{\theta_2}{\lambda} \frac{1}{[l(l+1)]^2} \tag{4.5}$$

$$(C_{m_4})_{jj} = \frac{\theta_4}{\lambda} \frac{1}{[l(l+1)]^2} \tag{4.6}$$

where λ is an overall damping parameter which controls the trade-off between the data misfit and smoothness. The parameters θ_2 and θ_4 control the relative strength of the anisotropy. For example, a value of 0.1 would give 10 times more weight to the isotropic terms relative to the anisotropic terms. Different values for θ_2 and θ_4 can be used to determine whether the data has a preference for anisotropy and if so, a preference for the 2ψ terms or for the 4ψ terms or for both. To compare inversions with different regularisations, we define, a reduced χ^2 as

$$\chi^2 = \frac{1}{N-M}(\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm}). \tag{4.7}$$

Where N is the number of data and M the trace of the resolution matrix. As the overall damping λ decreases, the trace of the resolution matrix (number of independent parameters) will increase and the reduced χ^2 will decrease, even if the misfit does not. A standard F-test (Bevington and Robinson, 1992) determines

if the difference between two χ^2 values is significant. Table 4.3 shows for a given χ^2 and number of free parameters $(N - M)$ the associated significant difference in χ^2 at the 99% confidence level. The meaning is that if for a given $N - M$, two χ^2 differ by more than this value, we are 99% sure that the misfit is better and that this inversion should be preferred.

4.4   Misfit curves for the higher modes

Following Trampert & Woodhouse (2003), we calculated misfit curves systematically changing λ for a fixed θ_2 and θ_4 for each of the higher modes to determine if the higher modes require anisotropy and if we can distinguish between the different anisotropic terms. The misfit curves (figures 4.3 and 4.4) show that for a small number of independent parameters, the isotropic parameterization $(\theta_2, \theta_4 = 10^{-5})$ explains the data best. As the number of independent parameters increases, the anisotropic parameterizations $(\theta_2, \theta_4 > 10^{-5})$ start to explain the data better than the isotropic parameterization. At around 500 independent parameters, the isotropic misfit curves flatten out, indicating that anisotropy is indeed required by the data, because they give a better misfit with a high confidence level. The F-test (Bevington and Robinson, 1992) gives the level of confidence with which the differences between the misfit curves is significant. For example, the first higher mode Rayleigh has a total of about 50,000 free parameters (the number of measurements - the trace of the resolution matrix) and a χ^2 of around 2.0 (figure 4.3a). According to table 4.3 the 99% significant difference is 0.023. At a trace of 1000, the difference between the isotropic and anisotropic misfit curves is 0.024, indicating indeed that we need anisotropy to explain our results. Beyond a trace of about 500, the differences between the isotropic and anisotropic misfit curves for all Rayleigh wave modes (figure 4.3) are significant with a high confidence level, indicating that anisotropy is needed to explain the phase velocity measurements. There are differences in misfit curves for different levels of anisotropic scaling, but these differences are not significant with a high confidence level.

As for Rayleigh waves, the difference between the isotropic and anisotropic misfit curves is 99% significant for all Love wave modes (figure 4.4). Again the data cannot distinguish between different levels of anisotropic scaling. For fundamental mode Love waves, we would expect a preference for the 4ψ term of anisotropy since the amplitude of the 4ψ sensitivity is much higher than the one for 2ψ. For higher mode Love waves, we do not expect a preference of one anisotropic term over the other since higher mode Love waves are sensitive to
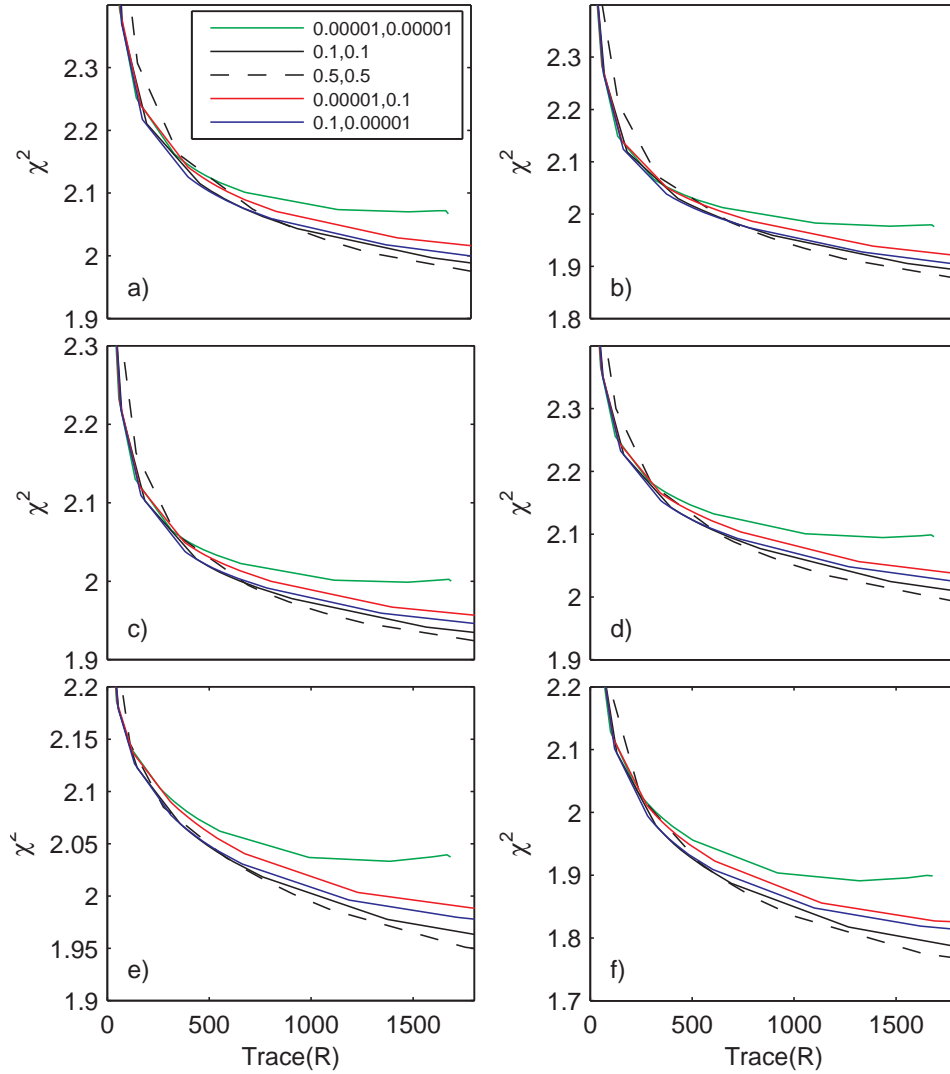
Figure 4.3: *Misfit curves for Rayleigh (a) first higher mode at 148.56 s, (b) second higher mode at 40.028 s, (c) third higher mode at 77.795 s, (d) fourth higher mode at 35.078 s, (e) fifth higher mode at 56.074 s and (f) sixth higher mode at 35.141 s. The legend shows different values for θ_2 and θ_4.*
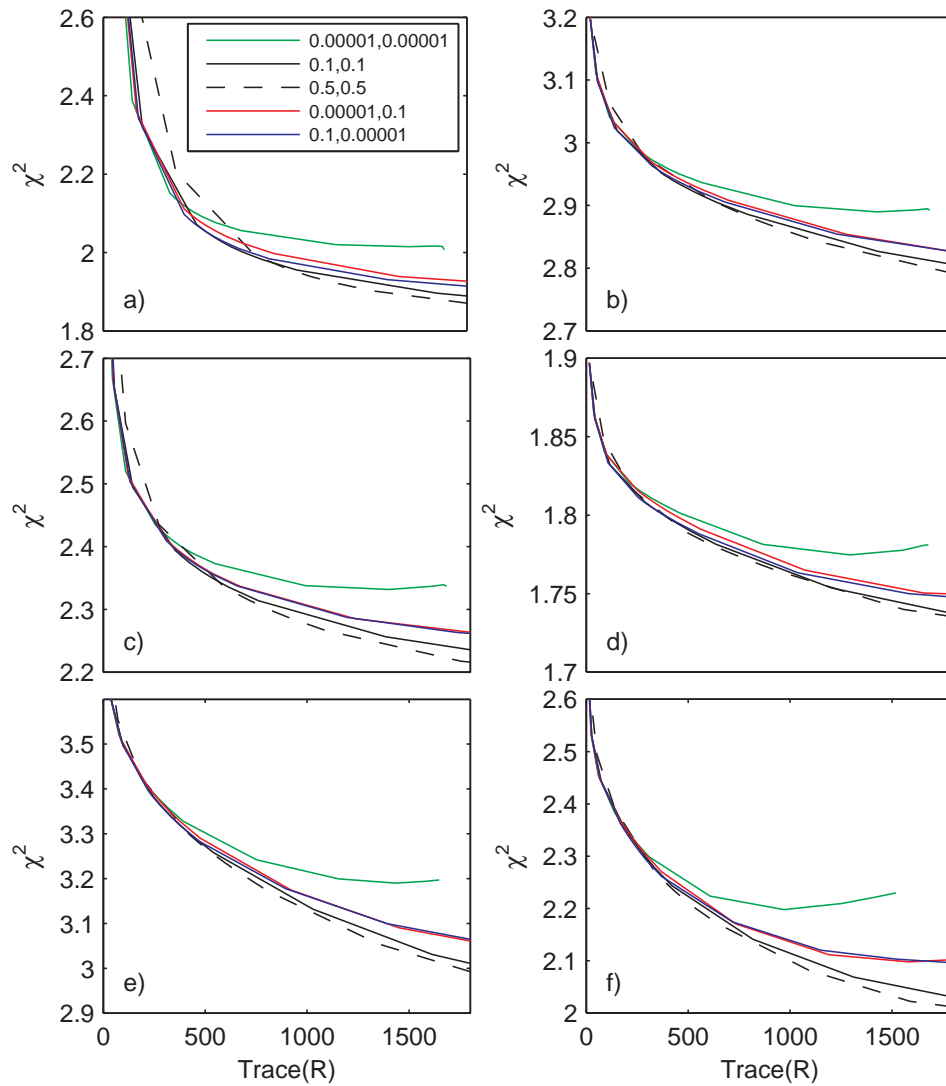
Figure 4.4: *Misfit curves for Love (a) fundamental mode at 153.46 s, (b) first higher mode at 153.07 s, (c) second higher mode at 40.02 s, (d) third higher mode at 78.66 s, (e) fourth higher mode at 35.06 s and (f) fifth higher mode at 35.12 s.The legend shows different values for θ_2 and θ_4.*
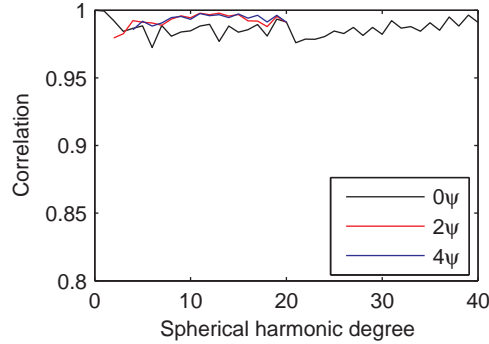
Figure 4.5: *Correlation as a function of spherical harmonic degree for minor and major arc vs. minor arc coverage.*

both the 2ψ as well as the 4ψ term of anisotropy. Earlier, a strong 2ψ term of anisotropy for fundamental Love waves has been found by Montagner & Tanimoto (1990). They ascribed this strong 2ψ term to Rayleigh-Love coupling, since fundamental mode sensitivity curves for Love waves only predict a strong 4ψ term of anisotropy. Trampert & Woodhouse (2003) found no statistical reason to include a 2ψ term and omitted it, based on asymptotic expectations. Since then, Sieminski et al. (2007) showed that Rayleigh-Love coupling is important (as speculated by Montagner & Tanimoto, 1990) and results in a high near source sensitivity for azimuthal parameters B-H. Furthermore, tilted uppermantle minerals with respect to the geographical reference system can result in high apparent values of B-H (Sieminski, personal communication, 2007). Although we find, similar to Trampert & Woodhouse (2003), no significant indication in favor of a 2ψ term, we choose to keep it based on a plausible reason for its existence. An important issue is to check whether the use of minor arc data alone (which results in poorer azimuthal coverage in the southern hemisphere) could bias our misfit curves. We computed synthetic data for a random anisotropic model (containing an isotropic, 2ψ and 4ψ term) and tested how well the random model could be retrieved by using a minor arc ray coverage alone and a minor and major arc ray coverage. We used the minor and major arc paths of Trampert & Woodhouse (2003). For the minor arc ray coverage we only took their minor arc paths. Figure 4.5 shows the correlation between the model retrieved by the minor and major arc coverage and the model retrieved by the minor arc coverage only. The correlations are very high (>0.95), indicating that essentially the same model is retrieved using both the mi-
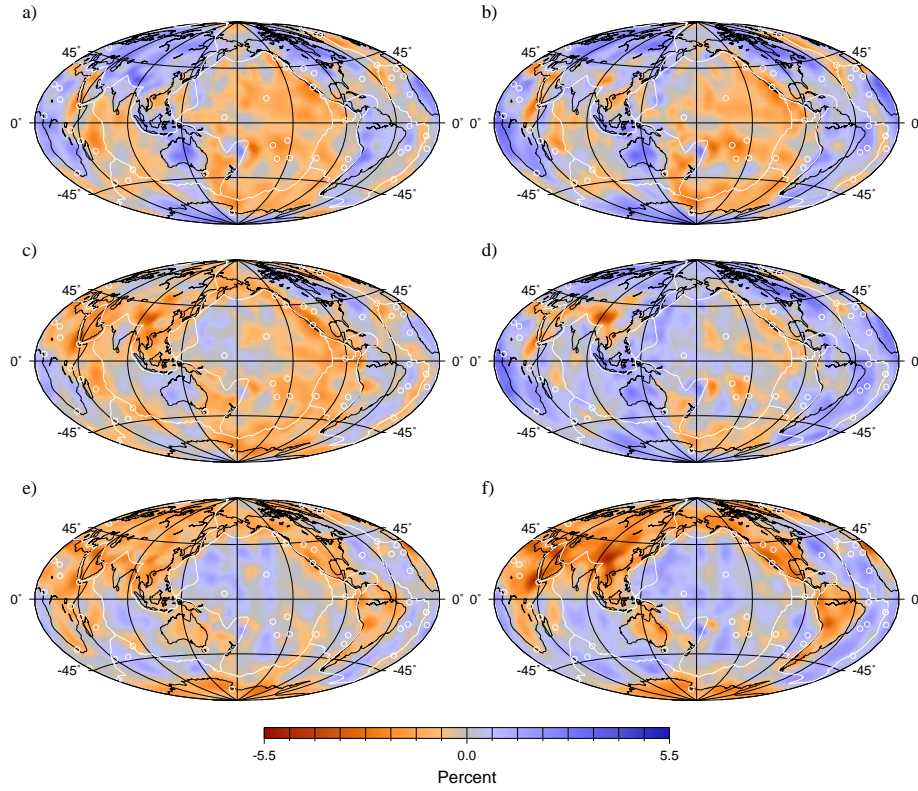
Figure 4.6: *Relative isotropic phase velocity maps with respect to PREM for Rayleigh (a) first higher mode at 148.56 s, (b) second higher mode at 40.028 s, (c) third higher mode at 77.795 s, (d) fourth higher mode at 35.078 s, (e) fifth higher mode at 56.074 s and (f) sixth higher mode at 35.141 s.*

nor and major arc coverage and using the minor arc coverage only. We further established the important point that including the 2ψ term does not change the 4ψ models. In summary, beyond 500 independent model parameters, azimuthal anisotropy is required by the data for all modes of Love and Rayleigh waves considered here. The prior strength of anisotropy cannot be determined from the data and has to be fixed by other arguments.

4.5 Azimuthally anisotropic phase velocity maps

We constructed azimuthally anisotropic phase velocity maps up to the first five higher mode Love and the first six higher mode Rayleigh wave phase velocities.
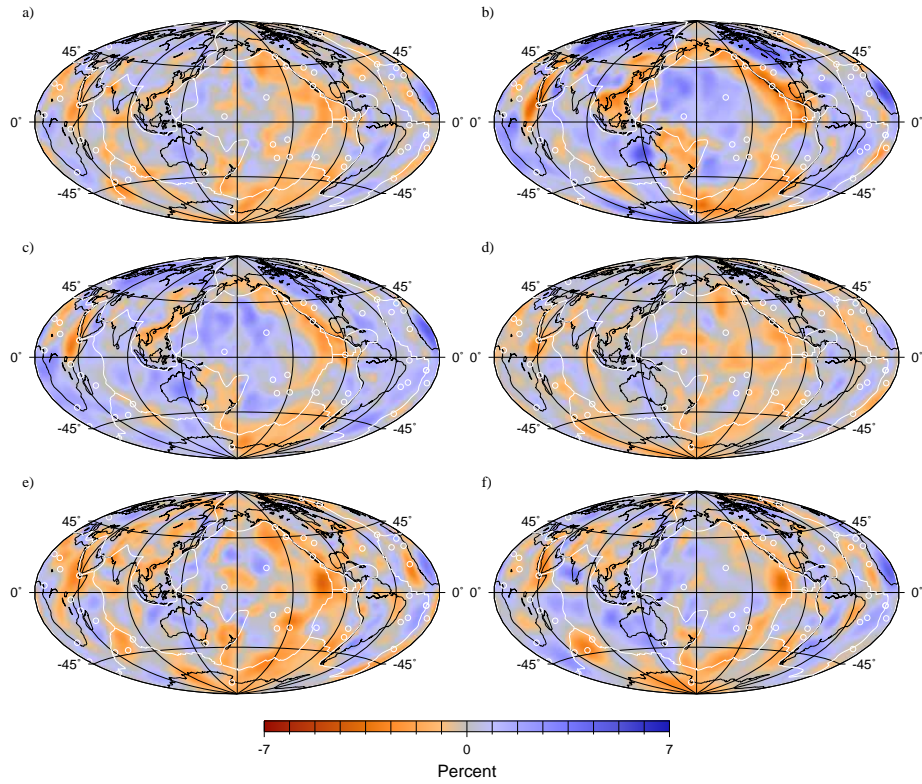
Figure 4.7: *Relative isotropic phase velocity maps with respect to PREM for Love (a) first higher mode at 153.07 s, (b) first higher mode at 40.16 s, (c) second higher mode at 40.02 s, (d) third higher mode at 78.66 s, (e) fourth higher mode at 35.06 s and (f) fifth higher mode at 35.12 s.*

The exact number of measurements used for the phase velocity maps is shown in table 4.2 and the rms uncertainty of a sample of the data is shown in table 4.4.

As seen in the previous paragraph, the data require azimuthal anisotropy but cannot decide upon its exact scaling. There is also no compelling reason to favor 2ψ or 4ψ terms only. An Occam-type argument guided us to choose a modest amount of anisotropy using $\theta_2 = \theta_4 = 0.1$. It should be noted that this is a prior constraint which will be overruled if the data require this locally. Because for fundamental modes our data quality seems superior compared to that used in Trampert & Woodhouse (2003), (smaller χ^2 for similar uncertainties in both data sets), we chose less overall damping to allow approximately 1000 independent parameters in the Rayleigh fundamental mode models. We have chosen an overall damping such that the relative model uncertainty remains constant for all modes.

Rayleigh							
mode	period	σ_{dc/c_0}(%)	σ_{dc}(m/s)	$trace(\mathbf{R})$	0ψ	2ψ	4ψ
000s097	100.393	0.48	19.56	1008	695	146	167
000s197	51.259	0.60	23.54	965	665	140	160
001s068	99.650	0.53	31.16	947	651	139	156
001s156	50.855	0.56	28.24	938	644	138	156
002s056	99.258	0.48	34.00	961	660	141	160
002s137	50.849	0.56	32.16	930	639	137	154
003s120	51.059	0.56	36.46	918	631	135	152
003s201	35.014	0.57	32.35	929	638	137	154
004s109	51.052	0.56	40.06	889	610	132	147
004s183	35.078	0.59	36.51	880	604	130	146
005s101	50.921	0.58	44.59	841	576	126	139
005s168	35.115	0.60	40.87	832	571	124	137
006s096	50.822	0.61	49.99	768	526	116	126
006s157	35.141	0.59	42.49	772	528	117	127
Love							
mode	period	σ_{dc/c_0}(%)	σ_{dc}(m/s)	$trace(\mathbf{R})$	0ψ	2ψ	4ψ
000t085	100.81	0.50	23.31	956	644	148	164
000t174	51.01	0.70	31.47	895	604	139	152
001t068	100.08	0.55	32.05	858	586	130	142
001t154	51.19	0.61	31.00	839	573	127	139
002t054	99.92	0.65	47.66	809	532	123	134
002t136	51.41	0.58	32.85	831	567	126	138
003t120	51.32	0.65	41.83	751	514	114	123
003t200	35.05	0.59	33.58	766	524	116	126
004t107	51.06	0.69	50.65	646	443	99	104
004t184	35.06	0.63	39.13	659	453	100	106
005t098	51.27	0.77	61.04	536	370	82	84
005t168	35.12	0.66	44.70	555	383	85	87

Table 4.4: *Relative and absolute rms data uncertainties (σ_{dc/c_0} and σ_{dc}), the total number of independent parameters ($trace(\mathbf{R})$) and the number of independent parameters for the isotropic (0ψ), 2ψ and 4ψ models.*

As a result, the phase velocity maps will have a decreasing resolution with increasing data uncertainty and/or decreasing number of data (see table 4.4). This choice is somewhat arbitrary. In view of a future depth inversion, ideally, we should have chosen for a constant resolution. The difference in the number of data between modes, however, is so large that the corresponding decrease in overall damping would have led to unrealistic amplitudes in some higher mode maps. The other extreme would have been to opt for an increasing uncertainty because the number of data constraints decreases. This would lead to seriously overdamped higher mode maps. A constant relative uncertainty in the phase velocity maps is an acceptable compromise between the two extremes. The isotropic phase velocity maps were expanded up to degree and order 40, while the azimuthal anisotropic phase velocity maps were expanded up to degree and order 20.

The isotropic models for Rayleigh and Love waves are shown in figures 4.6 and 4.7 for a number of different higher modes at the indicated periods. The fundamental mode maps are very similar to those of Trampert & Woodhouse (2003) with correlations of 0.91 (Rayleigh 40 seconds), 0.70 (Rayleigh 150 seconds), 0.87 (Love 40 seconds) and 0.79 (Love 150 seconds), and hence to other models by different research groups (see Becker et al. 2007, for a recent comparison). While the overtones generally show the strongest sensitivity to deeper mantle structure, it is interesting to note that high frequency Rayleigh waves of the fifth and sixth overtone are very sensitive to crustal structures. In general there is a good visual comparison between our maps and those of van Heijst (1997). Visser et. al. (2007) showed that higher modes can easily be measured with our technique. They showed isotropic degree 20 maps for illustration purposes. The correlation with our maps here up to degree 20 is around 0.80 for all Love wave overtones. The differences can be attributed to the neglect of anisotropy in the earlier paper.

The anisotropic contributions for the same higher modes and periods are shown in figures 4.8 and 4.9 for the 2ψ term and figures 4.10 and 4.11 for the 4ψ term. Figure 4.12 shows the rms amplitude averaged over the sphere of the 2ψ and 4ψ maps for Rayleigh and Love waves for the fundamental up to the sixth higher mode with corresponding uncertainties. The amplitudes of the 2ψ and 4ψ term are similar within their standard deviations. Most importantly, the amplitude remains positive within the uncertainties, indicating that the anisotropic models are robust and indeed required, for the chosen optimal scaling. The 2ψ amplitudes even remain robust within two standard deviations. Figures 4.13 and 4.14 show some chosen 2ψ correlations and the corresponding ray theoretical sensitivity kernels (Larson et al., 1998). For Rayleigh waves, the correlation of the fundamental
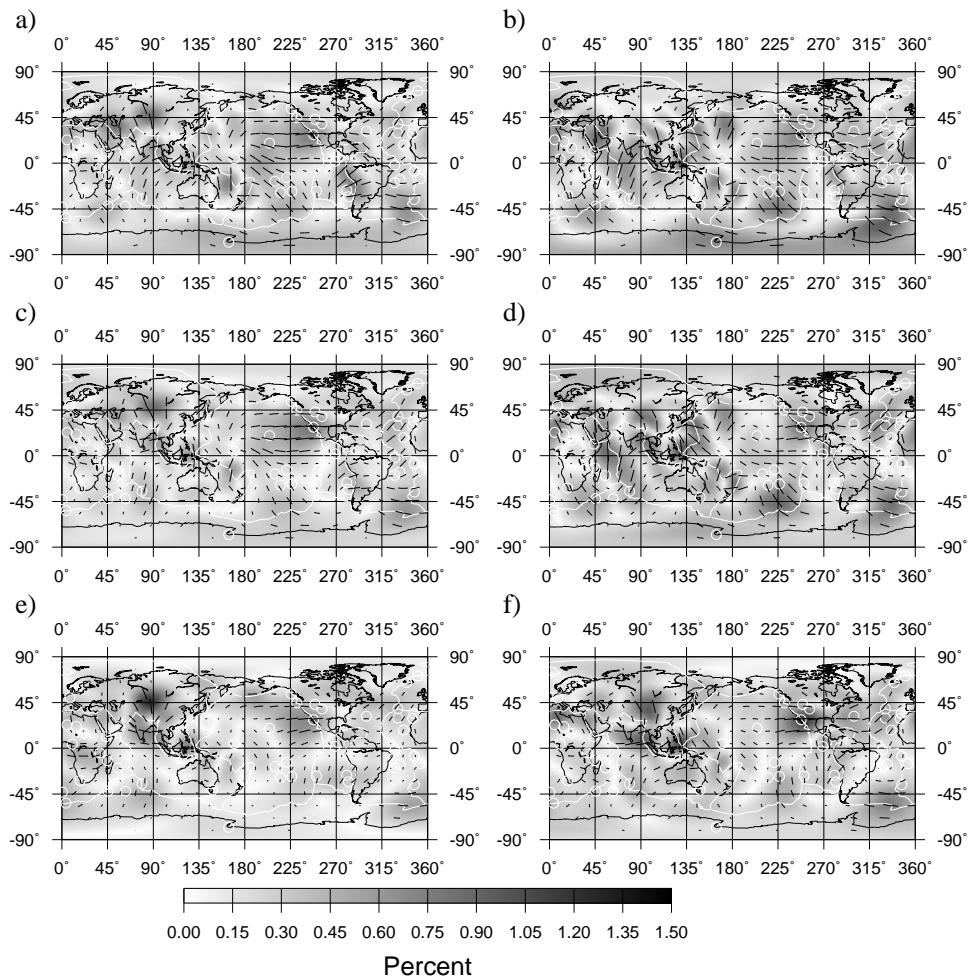
Figure 4.8: *Azimuthally anisotropic 2ψ phase velocity maps for Rayleigh. The grey scale in the background corresponds to the peak-to-peak amplitude of anisotropy expressed relative to the average phase velocity calculated from PREM. The black lines represent the fast directions which are also scaled to the amplitude shown in the background. The plate boundaries and hotspots are indicated in white. Panels (a) to (f) show the different modes and periods as indicated in figure 4.6.*

mode models with the first higher mode ones is high (figure 4.13a). In fact, the correlation of the fundamental mode with increasingly higher modes consistently shows high values. The corresponding sensitivities show that the B-H sensitivity
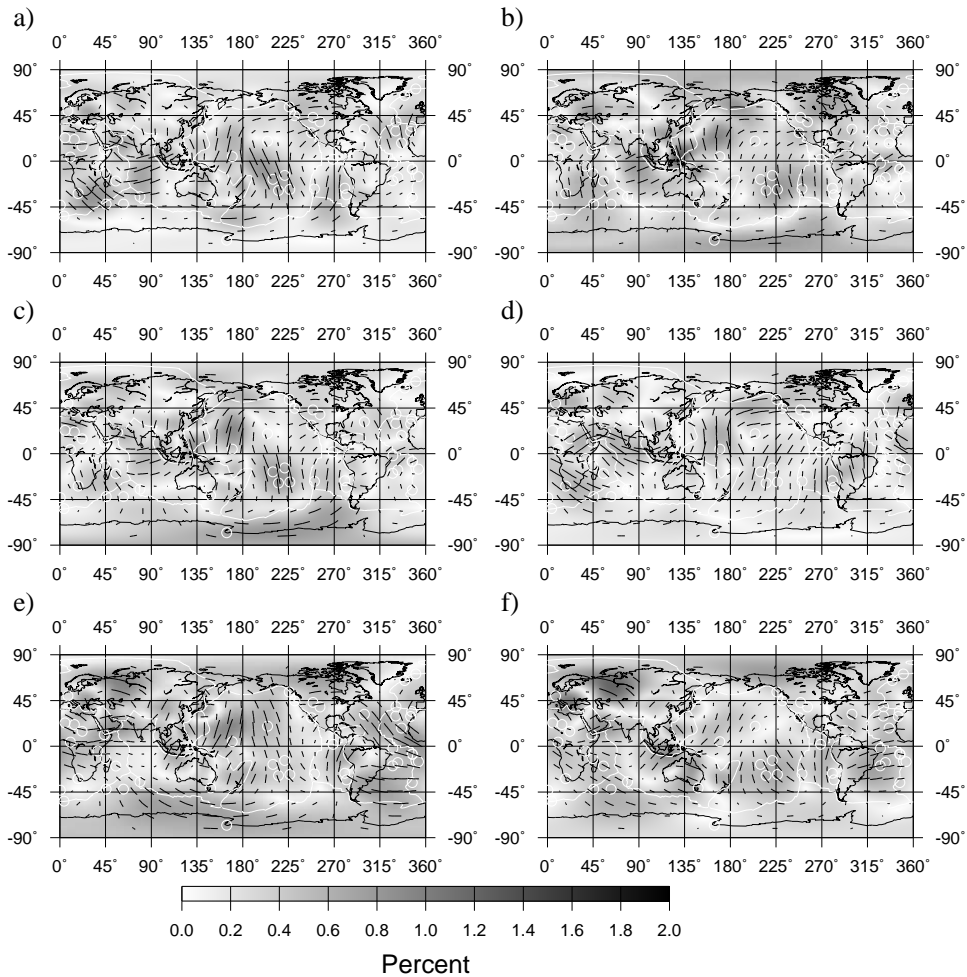
Figure 4.9: *Azimuthally anisotropic 2ψ phase velocity maps for Love higher modes. Panels (a) to (f) show the different modes and periods as indicated in figure 4.7.*

is mostly shallow for all modes while the G sensitivity changes with depth. This could indicate that B-H anisotropy is important for Rayleigh waves. We also find high correlations for modes where the most overlap is for deeper G, probably the transition zone anisotropy observed by Trampert & van Heijst (2002). For Love waves we obtain high correlations for G in the asthenosphere (figure 4.14). The correlation between the 2ψ models of the fundamental mode and first higher mode Love wave is quite low (< 0.5), which is not surprising since G sensitivity for the
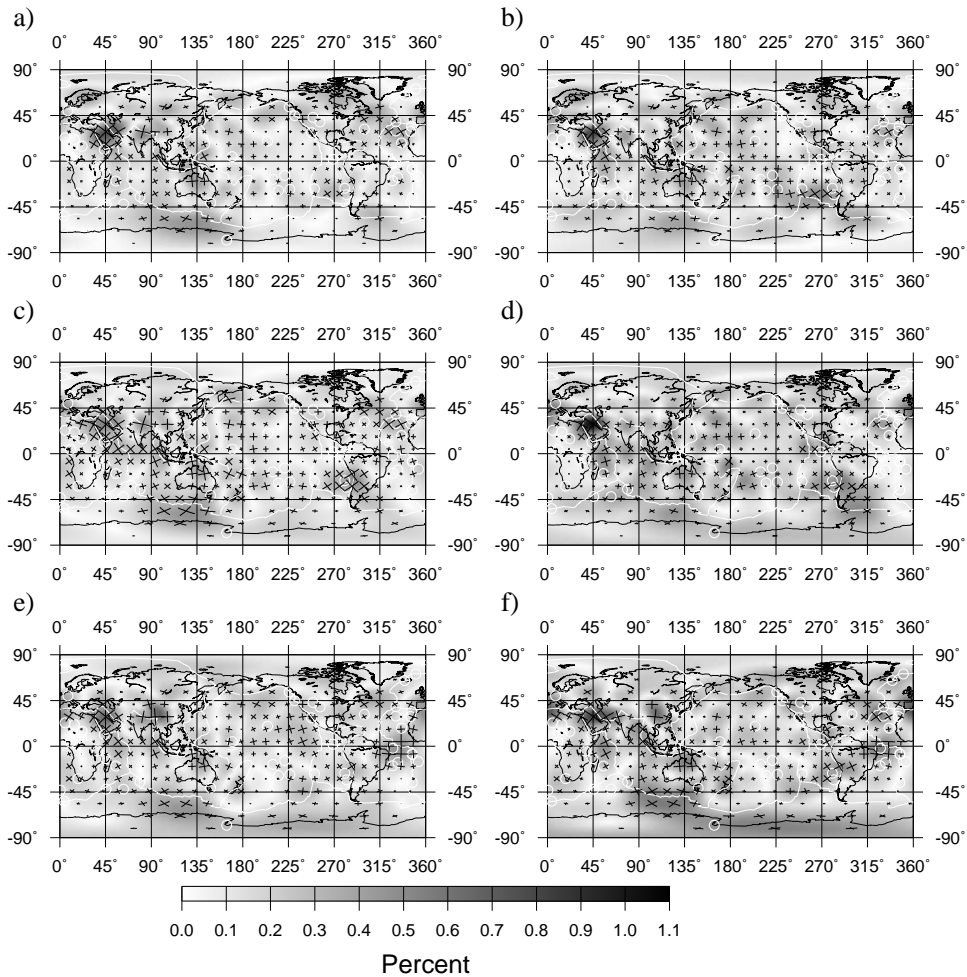
Figure 4.10: *Azimuthally anisotropic 4ψ phase velocity maps for Rayleigh higher modes. Panels (a) to (f) show the different modes and periods as indicated in figure 4.6.*

fundamental mode is almost zero while it is non-zero for the first higher mode. The most likely source of 2ψ anisotropy in fundamental mode Love waves is B-H (Sieminski et al., 2007), while for the overtones G dominates, hence a plausible low correlation. These few examples illustrate how complex the depth distribution of azimuthal anisotropy possibly is, and only a depth inversion will provide detailed information about the specific distribution of the anisotropy. This will require finite frequency kernels (Sieminski et al., 2007) that capture the strong
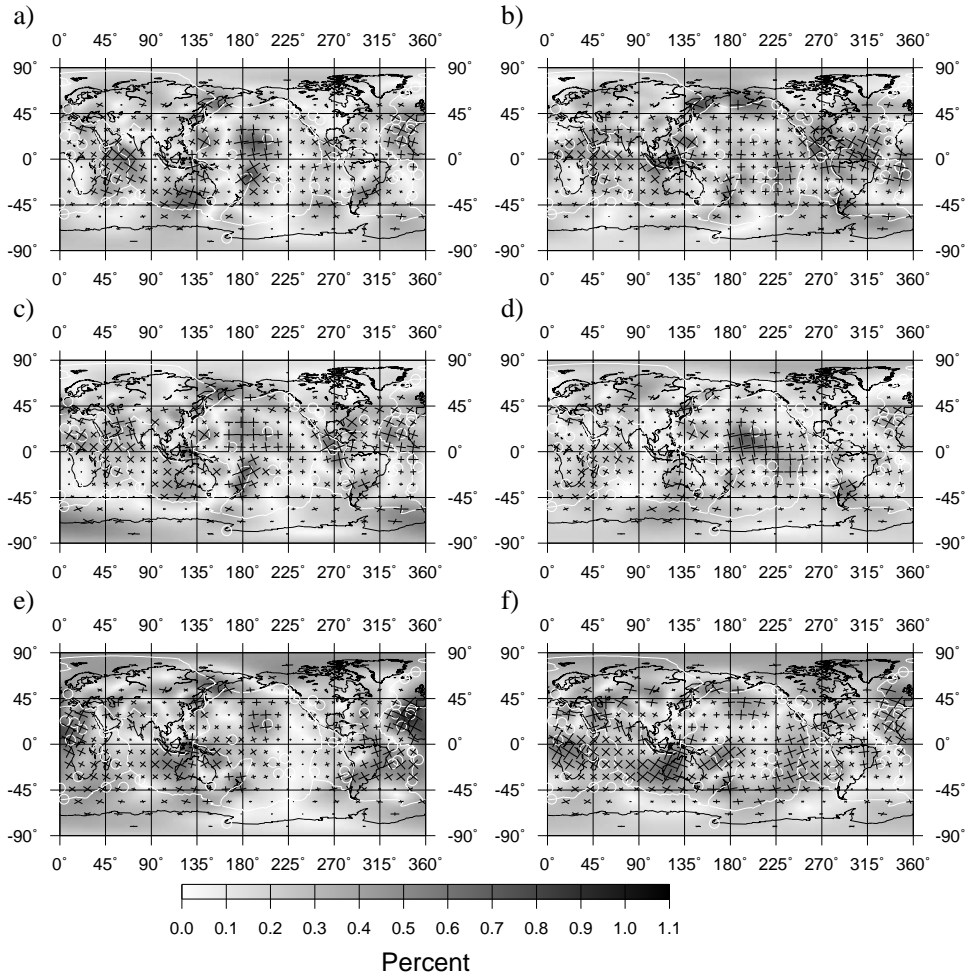
Figure 4.11: *Azimuthally anisotropic 4ψ phase velocity maps for Love higher modes. Panels (a) to (f) show the different modes and periods as indicated in figure 4.7.*

influence of path dependence and mode coupling for anisotropic parameters.

The fundamental mode 2ψ Rayleigh models visually agree at long wavelength with the models by Trampert & Woodhouse (2003), Ekström (2000) and in the Pacific with results obtained by Smith et al. (2004). Up to degree 8, we have a correlation of 0.49 with the 2ψ map of Trampert & Woodhouse (2003) for Rayleigh waves at 40 seconds. The first quantitative comparison between different azimuthally anisotropy models and geodynamic flow models was done by
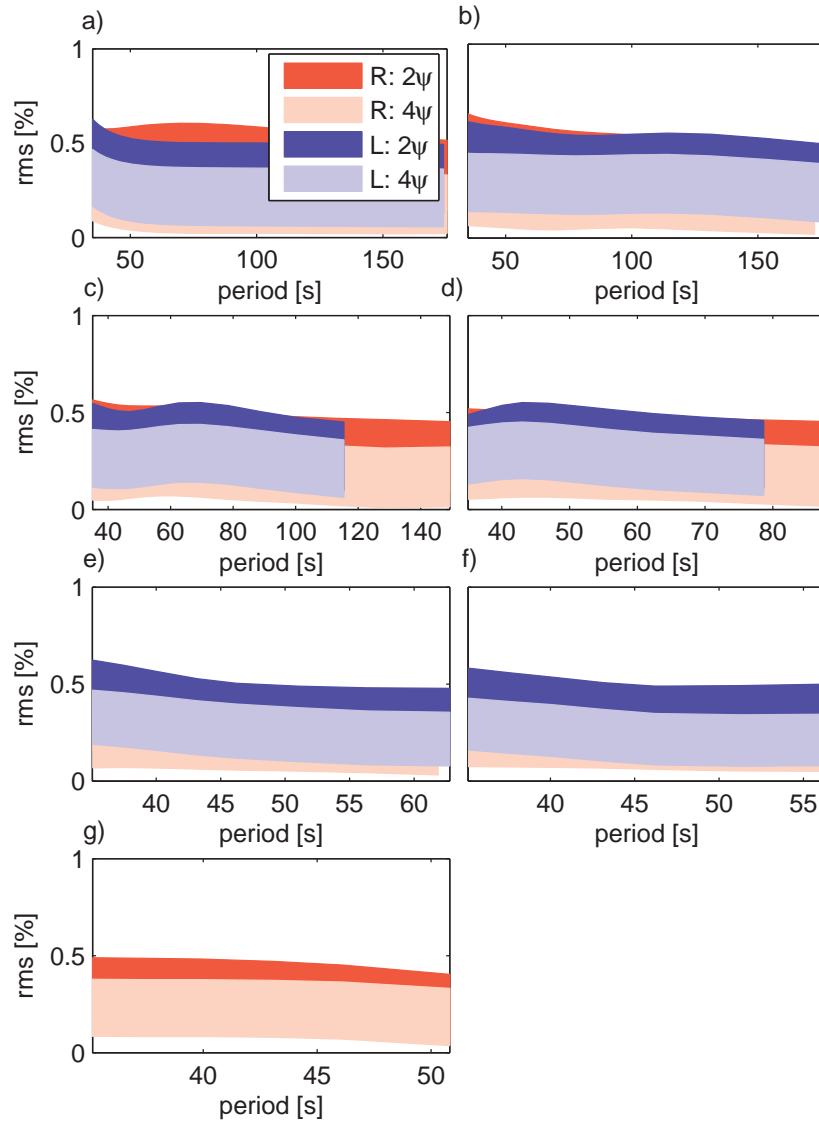
Figure 4.12: *The rms amplitude per unit sphere of the phase velocity maps of figures 4.8 to figure 4.11 for (a) the fundamental mode, (b) the first higher mode, (c) the second higher mode, (d) the third higher mode, (e) the fourth higher mode, (f) the fifth higher mode and (g) the sixth higher mode. The bands correspond to ± one standard deviation of the fixed average posterior uncertainty.*
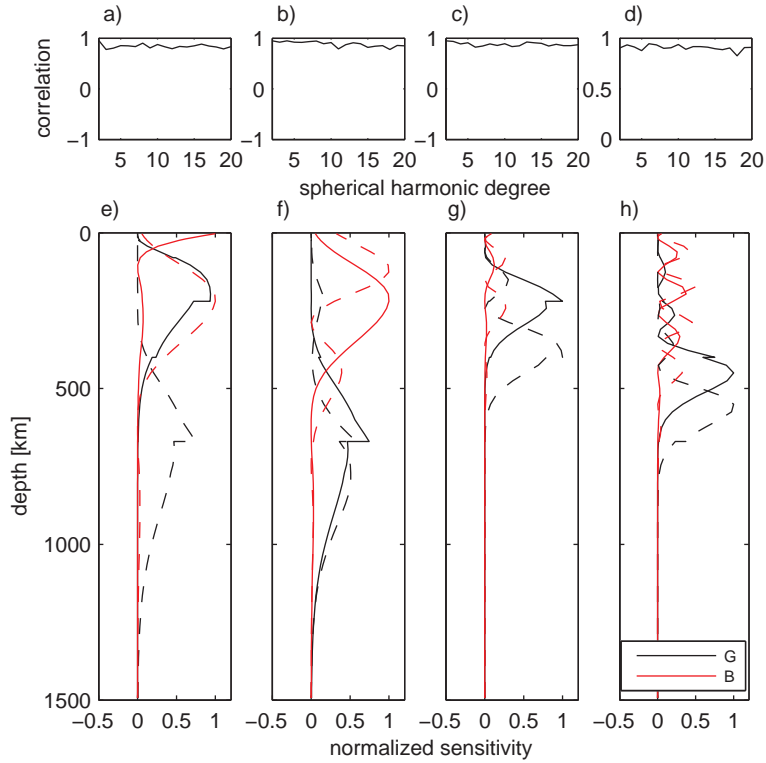
Figure 4.13: *Rayleigh 2ψ correlation (a-d) and 2ψ sensitivity kernels (e-h) of the fundamental mode at 151 s (solid) with the first higher mode at 149 s (dashed) (a,e), the first higher mode at 149 s (solid) with the second higher mode at 99 s (dashed) (b,f), the first higher mode at 40 s (solid) with the second higher mode at 40 s (dashed) (c,g) and the third higher mode at 35 s (solid) with the fourth higher mode at 35 s (dashed) (d,h). Sensitivity to H is not shown since it is similar to B sensitivity but opposite in sign.*

Becker et al. (2007). They found typical correlations between 0.18 and 0.47 indicating that our results are not at odds with any of these models. No comparison has been done for overtones. Only a detailed depth inversion can shed light on the geodymanic consequences of our 2ψ and 4ψ overtone maps.

Inverting for the azimuthal terms as well as the isotropic terms makes the isotropic maps become smoother for a given trace of the resolution. Decreasing overall damping will decrease the smoothness for both the isotropic, 2ψ and 4ψ phase velocity maps. We can split the total trace of the resolution matrix into the
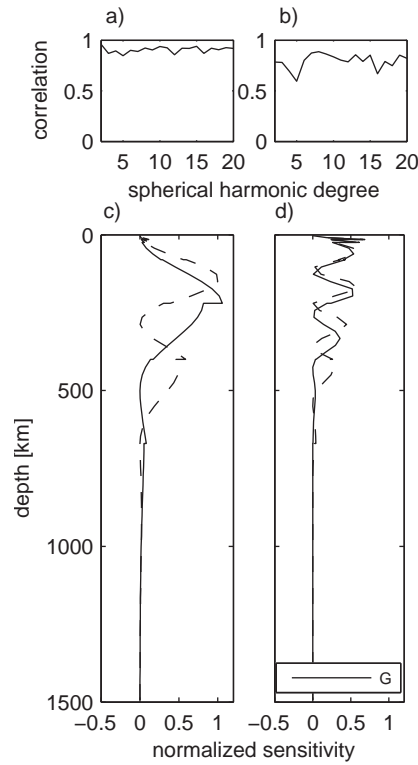
Figure 4.14: *Love 2ψ correlation (a,b) and 2ψ sensitivity kernels (c,d) of the first higher mode at 153 s (solid) with the second higher mode at 100 s (dashed) (a,c) and the third higher mode at 35 s (solid) with the fourth higher mode at 35 s (dashed) (b,d).*

trace for the isotropic, 2ψ and 4ψ terms separately (figure 4.15). These individual values are more meaningful for the phase velocity maps than the total trace. The number of independent parameters for the isotropic, 2ψ and 4ψ terms varies as a function of overall damping. For small numbers of independently inverted parameters, the isotropic parameters dominate. As the number of independently inverted parameters increases, the number of inverted 2ψ and 4ψ parameters increases. Table 4.4 shows the total number of independently inverted parameters and the number of isotropic, 2ψ and 4ψ parameters for some chosen Rayleigh and Love wave fundamental and higher modes. For the fundamental mode models, we can resolve on average up to 25 spherical harmonic degrees for the isotropic models, 8 spherical harmonic degrees for the 2ψ models and 9 spherical harmonic degrees
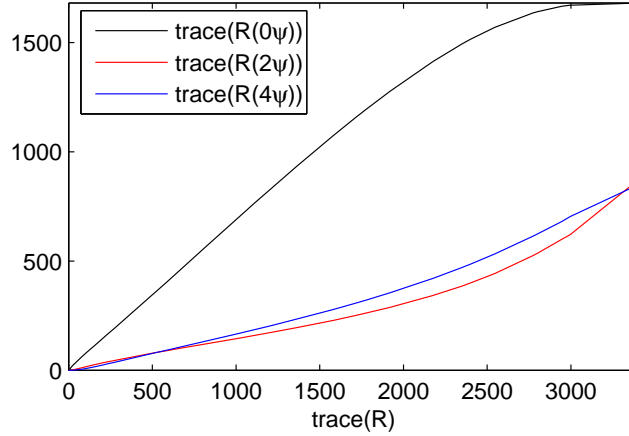
Figure 4.15: *The relation between the total trace of the resolution matrix and the trace of the resolution matrix separated for the isotropic, 2ψ and 4ψ terms for the fundamental mode Rayleigh wave at 151 seconds.*

for the 4ψ models. For the higher modes, the number of degrees we can resolve decreases to degree 18 for the isotropic models, degree 5 for the 2ψ models and degree 6 for the 4ψ models.

Shapiro & Ritzwoller (2002) use a rms data misfit as a measure of uncertainty for the phase velocity maps. They obtain values around 25 m/s for fundamental mode Rayleigh and between 25 and 40 m/s for the fundamental mode Love wave phase velocity maps. The rms data misfit values we obtain for the fundamental mode are between 24 m/s and 31 m/s for Rayleigh and 31 m/s to 45 m/s for Love waves. The rms data misfit values for the higher modes are somewhat larger. They vary for Rayleigh waves between 35 and 65 m/s and for Love waves between 45 and 75 m/s.

We are not so much interested in the posterior data uncertainties as in the posterior model uncertainties, which we need for a future depth inversion. The posterior model uncertainty is given by the posterior model covariance, defined as

$$\mathbf{C}_{\widetilde{m}} = (\mathbf{I} - \mathbf{R})\mathbf{C}_m(\mathbf{I} - \mathbf{R})^T + \mathbf{L}\mathbf{C}_d\mathbf{L}^T. \tag{4.8}$$

Where $\mathbf{C}_{\widetilde{m}}$ is the posterior model covariance, $\mathbf{C}_m$ is the prior model covariance (equation 4.4 to 4.6), $\mathbf{C}_d$ is the data covariance, $\mathbf{R}$ is the resolution matrix and $\mathbf{L}$ is the inverse operator which gives the estimated model $\widetilde{\mathbf{m}} = \mathbf{L}\mathbf{d}$. The square roots of the diagonal elements of the posterior model covariance can be interpreted as

error bars of the posterior values of the model parameters. An average posterior model uncertainty for a phase velocity map is obtained by taking the square root of the total power of the diagonal of the 0ψ, 2ψ and 4ψ terms averaged over the sphere. As explained above, we have chosen the overall damping (λ in $\mathbf{C}_m$) so that the average model uncertainty for $\widetilde{dc/c_0}$ is constant. The value has been chosen so as to invert for 1000 independent parameters for 100 s fundamental mode Rayleigh waves. This gives average relative model uncertainties $\sigma_{\widetilde{dc/c_0}}$ of 0.45%, 0.18% and 0.15% for the 0ψ, 2ψ and 4ψ maps (figure 4.16), respectively. The relative model uncertainty is much lower for the 2ψ and 4ψ maps due to the prior choice of a modest amount of anisotropy ($\theta_2 = \theta_4 = 0.1$). It is important to realise that a large part of $\mathbf{C}_{\widetilde{m}}$ comes from the prior information, therefore fixing $\mathbf{C}_{\widetilde{m}}$ will require different λ in $\mathbf{C}_m$ depending on $\mathbf{C}_d$ and the number of data, which will change $\mathbf{R}$ correspondingly.

The absolute uncertainties ($\sigma_{\widetilde{dc}}$, figure 4.16) in the fundamental mode Rayleigh isotropic models range from 15 m/s at short periods to 20 m/s at longer periods, the uncertainty of the 2ψ models range from 5 to 8 m/s and the uncertainty of the 4ψ models range from 5 to 7 m/s. For Love waves, the corresponding absolute uncertainties for the fundamental mode isotropic maps ranges from 14 to 20 m/s, for the 2ψ models from 5 to 8 m/s and the 4ψ models from 4 to 7 m/s. The rms data misfit values given earlier show uncertainties for both Love and Rayleigh wave azimuthal anisotropic models in the order of 25 m/s to 45 m/s for the fundamental mode. The absolute model uncertainties are a bit smaller but of the same order of magnitude as the rms data misfits, justifying the intuition of Shapiro & Ritzwoller(2002) to use the data misfits as average model uncertainties. The reason for this good correspondence is that the data misfit incorporates the prior information in equation 4.8 implicitly.

4.6   Resolution and Trade-off

The fifth higher mode Love wave data set has the lowest number of measurements. The number of measurements increases with lower overtone number (table 4.2). Nevertheless the pattern of ray density for the fifth higher mode Love wave is quite similar to the fundamental mode Rayleigh wave which contains the highest number of measurements. Trampert & Woodhouse (2003) converted the resolution matrix into averaging kernels. The relative phase velocity perturbation at a specific point on the Earth is an average of the true model over the whole Earth with weights (the averaging kernels). For a complete picture of resolution the averaging kernels have to be calculated at each point on the Earth's surface. Trampert
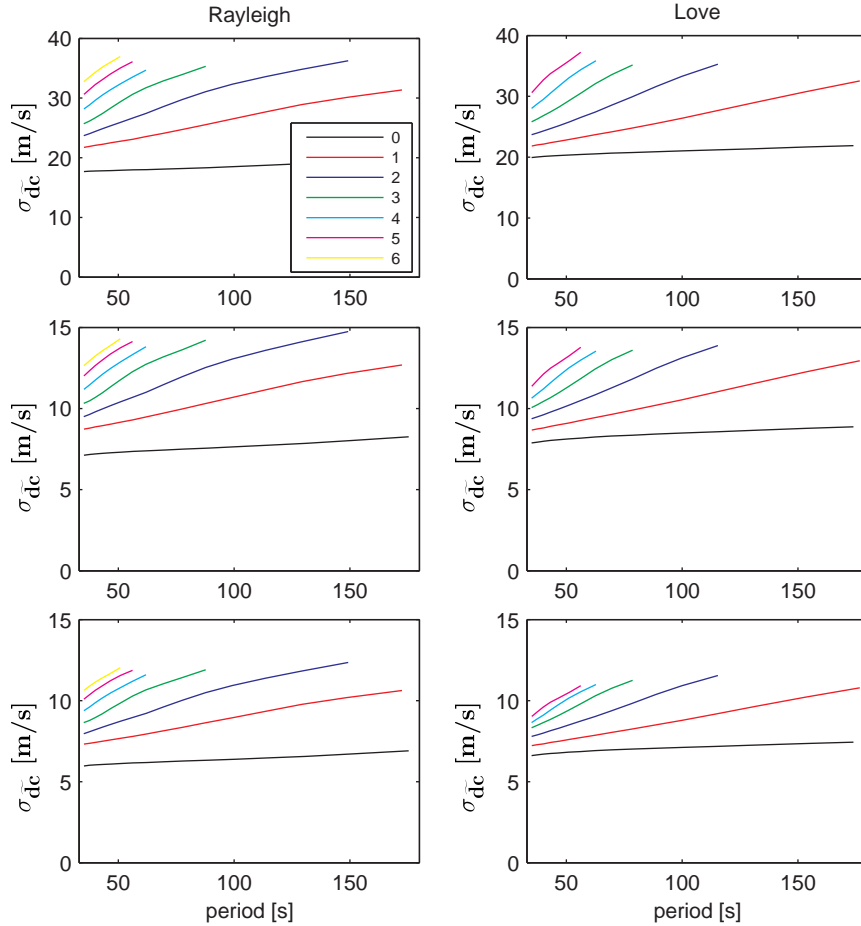
Figure 4.16: *Absolute standard deviations ($\sigma_{\widetilde{dc}}$) for the fundamental and first six higher mode Rayleigh wave isotropic (top left), 2ψ (middle left) and 4ψ (bottom left) phase velocity maps. The same for the fundamental and first five higher mode Love waves on the right.*

& Woodhouse (2003) chose to represent the averaging kernels by the radius of the central peak and called the maps resolving radii maps. The resolving radii are only dependent upon path coverage and the overall damping (λ) and give a good representation of the lateral resolution that can be achieved. The number of measurements of the fifth higher mode Love wave correspond to the number of
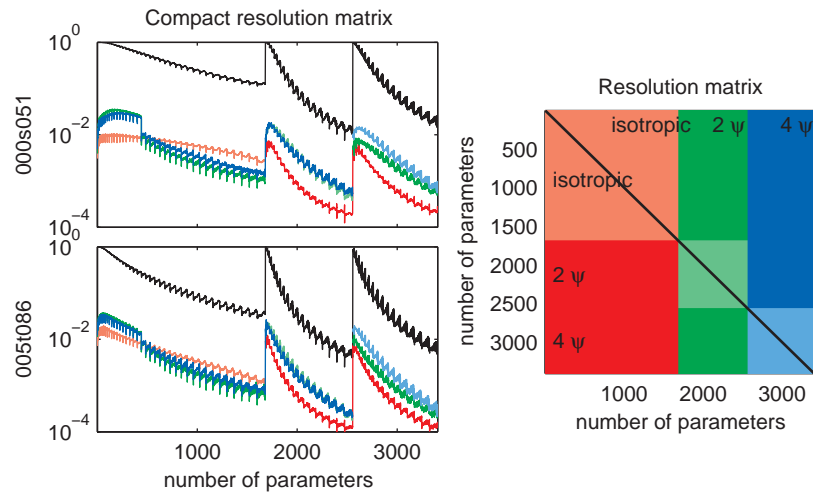
Figure 4.17: *The resolution matrix for the fundamental mode Rayleigh (top left figure) and the fifth higher mode Love (bottom left figure). The figure on the right shows how the left figures are related to the resolution matrix. Each row of the matrix is shown in the figures, where black is the trace of the resolution matrix, red is the rms of the isotropic parameters (not including the diagonal) for that row, green represents the rms of the 2ψ parameters for the specific row and blue the rms of the 4ψ parameters for the specific row. The off-diagonal terms within one parameter family are given by the lighter shaded colors, red for the isotropic parameters, green for the 2ψ parameters and blue for the 4ψ parameters. The darker shaded lines represent the trade-off between isotropic/2ψ/4ψ terms.*

measurements of the fundamental Love waves in Trampert & Woodhouse (1995) and the corresponding resolving radii map (figure7a in Trampert and Woodhouse, 1995) corresponds to the resolving radii map of the fifth higher mode Love wave (our worst data coverage). There is a high correspondence with the ray density map but the resolving radii test gives a better indication of the structures we are able to solve for.

There are three different issues that affect the resolution; spectral leakage, trade-off between the isotropic and anisotropic terms and damping. Spectral leakage is caused by the mapping of small-scale structure not accounted for in the model expansion into the inverted low-degree structure and is a result of uneven data coverage (Snieder et al., 1991). Spectral leakage can be suppressed by a time consuming operator (Trampert and Snieder, 1996) or approximately by Laplacian damping (Spetzler and Trampert, 2003) as in equations 4.4 to 4.6 ($\frac{1}{[l(l+1)]^2}$). The

price to pay for this Laplacian damping is that the higher the degree, the less it will be resolved and the diagonal peak of the resolution matrix will also broaden. Finally there will be trade-offs between the isotropic and anisotropic parameters. The resolution matrix provides information on the trade-offs and the broadening (figure 4.17). The diagonal of the resolution matrix shows the price we pay for the use of Laplacian damping. For higher degrees, the diagonal values of the resolution matrix decrease. The choice of the relative strength of anisotropy $(\theta_2, \theta_4 = 0.1, 0.1)$ causes the sharper decay for the 2ψ and 4ψ parameters. Table 4.4 shows the number of resolved parameters for certain modes given our choice of overall damping described above. The off-diagonal terms of the resolution matrix in figure 4.17 show the amount of broadening and trade-off between parameters. Fortunately, these values are small compared to the diagonal values. This holds for all higher modes.

4.7   Conclusions

We present global azimuthal anisotropic phase velocity maps for the fundamental modes and up to the sixth overtone for Rayleigh waves and up to the fifth overtone for Love waves. Phase velocities for fundamental and higher mode Love and Rayleigh waves were measured using a model space search approach (Visser et al., 2007a). The use of a model space search approach enables us to obtain realistic and consistent uncertainties on the phase velocities. The phase velocities are inverted to extract azimuthal anisotropic phase velocity maps. Following Trampert & Woodhouse (2003), we determine the optimum relative weighting prior to inversion. Both Love and Rayleigh fundamental and higher mode phase velocities require anisotropy according to the misfit curves. The relative weighting was chosen (in agreement with the significant difference of misfit curves) such that anisotropy is needed and equal for the 2ψ and the 4ψ terms of anisotropy. We have chosen the overall damping such that the relative uncertainty is constant in all maps. This causes the resolution to decrease with increasing data uncertainty and/or decreasing number of data.

The rms data misfits of the azimuthal anisotropic models for fundamental mode Rayleigh and Love waves are similar to values found by Shapiro & Ritzwoller (2002). The rms misfits for the higher modes are larger and vary between 35 and 65 m/s for Rayleigh and between 45 and 75 m/s for Love waves. The model uncertainties are smaller than the rms data misfits but of the same order. For the fundamental mode isotropic models, we obtain uncertainties up to 20 m/s and for the anisotropic models we obtain uncertainties up to 8 m/s.

The isotropic maps visually compare well with the isotropic higher mode maps of van Heijst (1997). We found a high correlation with the fundamental mode anisotropic maps of Trampert & Woodhouse (2003) and hence with equivalent work from other research groups (Becker et al., 2007). Indications are that the source of azimuthal anisotropy is complex and a detailed depth inversion, using finite frequency kernels, is needed to clarify this. Our efforts to provide maps for many overtones, should facilitate this final step in the search of deep anisotropy.

4.8   Acknowledgments

Chapter 5

Probability of radial anisotropy in the deep mantle

It is well established that the Earth's uppermost mantle is anisotropic, but observations of anisotropy in the deeper mantle have been more ambiguous. Radial anisotropy, the discrepancy between Love and Rayleigh waves, was included in the top 220 km of PREM, but there is no consensus whether anisotropy is present below that depth. Fundamental mode surface waves, for commonly used periods up to 200 s, are sensitive to structure in the first few hundred kilometers and therefore do not provide information on anisotropy below. Higher mode surface waves, however, have sensitivities that extend to and below the transition zone and should thus give insight into anisotropy at greater depths, but they are very difficult to measure. We previously developed a new technique to measure higher mode surface wave phase velocities with consistent uncertainties. These data are used here to construct probability density functions of a radially anisotropic Earth model. In the uppermost mantle, we obtain a high probability of faster horizontally polarized shear wave speed, likely to be related to plate motion. In the asthenosphere and transition zone, however, we find a high probability of faster vertically polarized shear wave speed, an indication of overall vertical flow. In the lower mantle, we see no significant shear wave anisotropy. This is consistent with results from laboratory measurements which show that lower mantle minerals are anisotropic but LPO is unlikely to develop in the pressure-temperature conditions present in the lower mantle.

5.1   Introduction

Radial and azimuthal anisotropy are different expressions of the underlying general anisotropy of the Earth's interior. The source of anisotropy in the mantle is usually assumed to be the alignment (lattice preferred orientation or LPO) of intrinsically anisotropic minerals under strain in the mantle (Karato, 1998a; Montagner, 1998). When detected, anisotropy can be an indicator of mantle strain and flow and improve our understanding of the dynamics of the mantle. Evidence for radial anisotropy was first inferred from the discrepancy between Rayleigh and Love waves by Anderson (1961), Aki and Kaminuma (1963) and McEvilly (1964). These observations prompted the inclusion of radial anisotropy in the upper 220 km, also referred to as the anisotropic zone, of the global reference Earth model PREM (Dziewonski and Anderson, 1981). It is now commonly accepted that the Earth is radially anisotropic at shallow depths (up to $\sim$200 km). There is, however, no consensus on whether radial anisotropy is present beyond the anisotropic zone. While earlier studies of radial anisotropy used fundamental mode surface waves (Tanimoto and Anderson, 1984; Nataf et al., 1984; Montagner and Tanimoto, 1991; Ekström and Dziewonski, 1998; Shapiro and Ritzwoller, 2002), in recent years higher mode surface waves have been added to studies of radial anisotropy (Debayle and Kennett, 2000; Gung et al., 2003; Beghein et al., 2006; Maggi et al., 2006; Panning and Romanowicz, 2006; Sebai et al., 2006; Marone et al., 2007) with a potential to yield constraints on deeper mantle dynamics, down to the transition zone and lower mantle. Radially anisotropic shear wave velocity models tend to agree at long wavelengths only (Panning and Romanowicz, 2006), suggesting large uncertainties in these models. These uncertainties depend on the regularisation, parameterisation, inverse method, data uncertainties etc. Model space search methods provide a way to obtain a full probability density function for the parameters through the mapping of the entire model space rather than just one preferred central value.

A previous (linearized) Monte Carlo model space search for radial anisotropy in seismic reference models of the mantle (Beghein et al., 2006) found no significant spherically averaged radial anisotropy beyond the anisotropic zone, while spherically averaged radial anisotropy was found up to 1000 km in an earlier study by Montagner and Kennett (1996). Panning and Romanowicz (2006) inverted for a three-dimensional radial anisotropic model and found faster vertically polarized shear wave speed associated with subducted slab material in the transition zone. There is a large consensus between the one-dimensional and three-dimensional radially anisotropic studies that the lower mantle is isotropic, except in D" (Kendall and Silver, 1996; Karato, 1998b; Panning and Romanowicz, 2004). An isotropic

lower mantle can be explained in terms of superplastic flow (Karato, 1998a), which does not result in any preferred orientation of minerals, even though, the minerals themselves are still highly anisotropic.

In this paper, we inverted the fundamental and higher mode phase, azimuthally averaged, velocity maps of Visser et al. (2007b) for a global radially anisotropic shear wave velocity model using a fully non-linear model space search approach. We use Rayleigh wave phase velocity maps for the fundamental and up to the sixth higher mode and Love wave phase velocity maps for the fundamental and up to the fifth higher mode. This provides us with a large dataset of higher modes, especially in comparison with previous radially anisotropic studies (Debayle and Kennett, 2000; Maggi et al., 2006; Panning and Romanowicz, 2006; Marone et al., 2007), where the number of higher mode measurements are often few and up to a relatively low higher mode (second to fourth higher mode). The use of a model space search approach in the inversion for shear wave velocities should provide us with realistic uncertainties. The phase velocity measurements were obtained using a model space search approach, yielding consistent uncertainties between all the measurements. These uncertainties have been propagated in the construction of the phase velocity maps (Visser et al., 2007b) and used as prior information here. The combination of a model space search and the large number of higher mode measurements should provide us with a global radially anisotropic model with an improved depth resolution and consistent uncertainties which in turn should give us insight into the mantle dynamics at larger depths in the mantle.

We invert Rayleigh and Love wave phase velocities separately to obtain global horizontally and vertically polarized shear wave velocity model. These models are then combined into a global radially anisotropic shear wave velocity model. While other studies (Ekström and Dziewonski, 1998; Debayle and Kennett, 2000; Maggi et al., 2006) have also used two separate inversions for the Rayleigh and Love wave data, they used a linearised approach. Ekström and Dziewonski (1998) showed that no significant bias was introduced by the use of isotropic sensitivity kernels and two separate inversions for V_{SH} and V_{SV}. Since we perform a fully non-linear model space search, we first need to validate the assumption of inverting the Rayleigh and Love wave phase velocities separately.

5.2   Depth inversion

This study presents the last stage of a three stage inversion as proposed by Kennett and Yoshizawa (2002). A traditional two-stage approach of multimode waveform tomography consists of obtaining one-dimensional velocity perturbations through

waveform fitting and inverting them, using the path average assumption, for a three-dimensional velocity model. The three stage approach consists of obtaining one-dimensional dispersion models through waveform fitting in the first stage, building multimode phase velocity models as a function of frequency using the path average assumption in the second stage and an inversion for local wave speed properties to obtain the three-dimensional velocity model in the third stage. The one-dimensional dispersion model in the first stage is regarded as a representation of the character of multimode dispersion along the source-receiver path. This is not as limiting an assumption as the path average assumption in the two stage approach. Yoshizawa and Kennett (2002) showed that multiple one-dimensional shear wave velocity models obtained through waveform fitting with a slight difference in misfit share the same dispersion characteristics indicating that the one-dimensional velocity model in the first stage may be regarded as a representation of the multimode dispersion characteristics along the source-receiver path.

In the first stage, we applied waveform fitting using a model space search approach (using the Neighbourhood Algorithm; Sambridge, 1999a,b) to obtain the fundamental and higher mode Love and Rayleigh wave phase velocity measurements (Visser et al., 2007a). In the second stage (Visser et al., 2007b), we inverted the fundamental and higher mode Love and Rayleigh wave phase velocity measurements for global isotropic and azimuthally anisotropic phase velocity models. The isotropic parts of the phase velocity models are now used in the third stage to obtain a radially anisotropic shear wave velocity model. Montagner and Nataf (1986) showed that radial anisotropy is dependent on the Love parameters (A, C, L, N, F) (Love, 1927) which describe a transversely isotropic medium, while azimuthal anisotropy is dependent on the other elastic parameters (B, H, E, G). By using the isotropic phase velocity models of the second stage (Visser et al., 2007b), we only have to worry about the five Love parameters. We perform a fully non-linear point-by-point depth inversion using a model space search approach. To keep the number of parameters low in this Monte Carlo search and inspired by previous work, we invert the Love and Rayleigh wave phase velocity models separately. Ekström and Dziewonski (1998) validated this approach for a linearised inversion. In the case of a fully non-linear approach, however, the validity of this approximation has not been shown. Therefore we first performed a test where we calculated Love and Rayleigh wave phase velocities for the anisotropic PREM model. We then separated the anisotropic PREM model in a horizontally polarized model (V_{PH}, V_{SH}) and a vertically polarized model (V_{PV}, V_{SV}) and calculated the Love and Rayleigh wave phase velocities separately assuming isotropy. We found that the resulting phase velocity differences are within the uncertainties of
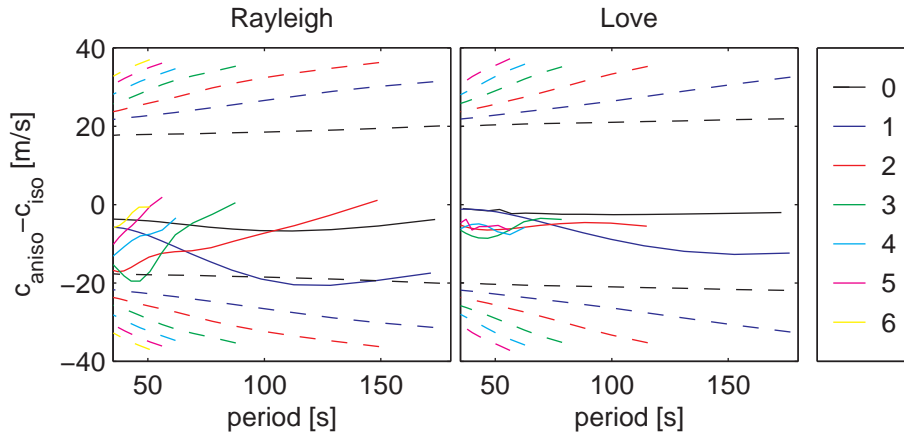
Figure 5.1: *The difference between the Love and Rayleigh phase velocities calculated assuming anisotropic and isotropic profiles for a location in the Baltic Shield ($58°N, 17°E$). The dashed lines indicate the uncertainties for the phase velocity models of Visser et al. (2007b). For Love, the isotropic model has V_S and V_P equal to V_{SH} and V_{PH} of the anisotropic model, for Rayleigh the isotropic model has V_S and V_P equal to V_{SV} and V_{PV} of the anisotropic model.*

the phase velocity models of Visser et al. (2007b). Since anisotropic PREM contains only shallow anisotropy, we performed the same test for the results of our depth inversion at a few locations on the Earth. The differences in the phase velocities calculated assuming isotropy or anisotropy are within the uncertainties of the phase velocity models (Fig. 5.1), indicating no significant difference between the two approaches. Therefore, we can invert Love and Rayleigh wave phase velocities separately resulting in a considerable gain in CPU time (days rather than weeks).

We selected 492 locations, covering the Earth's surface according to a 6-fold triangular tessellation (equal area representation, Wang and Dahlen (1995)). For each point, we calculated the local phase velocities for selected fundamental and higher mode isotropic Love and Rayleigh wave phase velocity maps. The sampling of the Earth's surface is comparable to that of a spherical harmonic expansion of degree and order 20 (Wang and Dahlen, 1995), which is similar to the resolution of the phase velocity maps of Visser et al. (2007b). The phase velocity measurements used for the building of the phase velocity maps were obtained using a model space search approach. This provided us with consistent uncertainties on the measurements as well as on the phase velocity maps as described in Visser
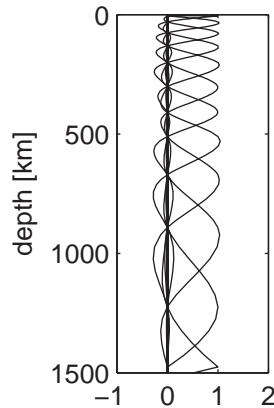
Figure 5.2: *Twelve natural cubic spline basis functions. The splines are numbered one to twelve from the top to the bottom (1500 km).*

et al. (2007b). At each location, we invert the local phase velocities of different modes, with the corresponding uncertainties. This will provide us with consistent posterior uncertainties given these prior uncertainties. The objective of the model space search is to find, for each location, the V_{SH} and V_{SV} model and the Moho depth that fits the observed phase velocities for Love and Rayleigh waves, respectively.

5.2.1   Parameterisation

We parameterize the shear wave velocity model using the same 12 natural cubic spline basis functions which have been used in the measurement stage (Fig. 5.2). The position and number of the spline basis functions were obtained after several tests with different parameterizations. A Backus-Gilbert resolution analysis showed that the twelve spline parameterization is optimal for the modes used here. The splines are more densely spaced in the upper mantle compared to the lower mantle to match the depth resolution of surface waves. As in Visser et al. (2007b), we scaled the compressional wave velocity and density to the shear wave velocity model. For the compressional wave velocity, we chose the scaling relation of Ritsema and Van Heijst (2002) and for the density the scaling relation of Deschamps et al. (2001). Scaling relations are often used in depth inversions (See for example, Ekström and Dziewonski, 1998; Shapiro and Ritzwoller, 2002; Gung et al., 2003; Panning and Romanowicz, 2006) to reduce the number of parameters in

the inverse problem to the best resolved parameters (V_{SH}, V_{SV}). Multiple studies (Ekström and Dziewonski, 1998; Gung et al., 2003) have shown that specific scaling relations did not affect the resulting velocity models much.

Crustal corrections are very important in surface wave tomography (Montagner and Jorbert, 1988; Mooney et al., 1998; Zhou et al., 2006; Marone et al., 2007; Bozdag and Trampert, 2007). Bozdag and Trampert (2007) showed that accurate crustal corrections are more difficult for Love waves, due to the higher sensitivity to crustal structure. Radially anisotropic shear wave velocity models (combinations of Rayleigh (V_{SV}) and Love (V_{SH}) data) are, therefore, most affected by improper crustal corrections. We therefore follow Li and Romanowicz (1996) and do not perform crustal corrections but add Moho depth as one additional parameter to the inversion. The initial crustal model is from Meier et al. (2007), who obtained a crustal model by inverting fundamental mode phase velocities using a neural network approach. The crustal model consists of an average shear wave velocity for the crust and a Moho depth. For the frequencies we use, Moho depth is the important parameter and crustal velocities matter little (Meier et al., 2007). We therefore keep the crustal velocities fixed and vary Moho depth only. The first spline coefficient is therefore fixed. The second spline is defined at the specific Moho depth for the tesselation location. So, our final velocity parameterisation consists of eleven splines from the Moho down to 1500 km and one extra parameter which is Moho depth. Both the V_{SH} as well as the V_{SV} inversion should provide similar Moho depths. Figure 5.3 shows that this is indeed the case for the V_{SV} and V_{SH} models. The Moho depths resulting from both inversions are consistent. The best fitting line through the points is for $Moho_{SV} = Moho_{SH}+0.5$ km and the standard deviation is 0.4 km. This range is well within the mean standard deviations of the Moho depths (3.0 km) from Meier et al. (2007), indicating that the differences between the Moho depths are not significant and the separate V_{SH} and V_{SV} inversions are consistent with each other.

For each tesselation point, we construct a shear wave velocity model, searching in a certain range around PREM (Dziewonski and Anderson, 1981), from the Moho down to 1500 km and adapt the Moho depth, searching around the model of Meier et al. (2007). The topography and the bathymetry, for the tesselation location, are fixed and taken from CRUST2.0 (Bassin et al., 2000). Below 1500 km, we assume PREM.

5.2.2 Model space search

For the model space search we use the Neighbourhood Algorithm (Sambridge, 1999a,b). The first part of the NA is a Monte Carlo search that uses the misfit
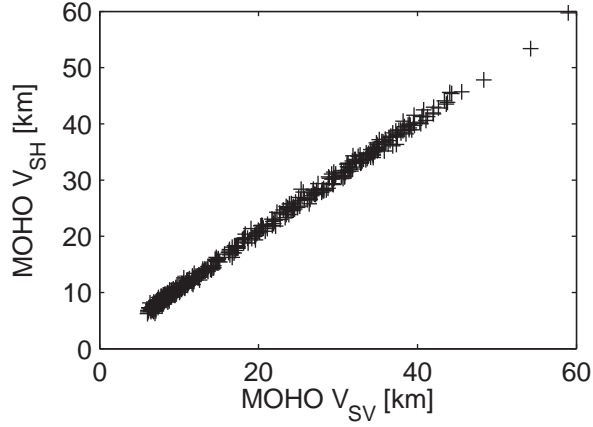
Figure 5.3: *Moho depths resulting from the inversion for V_{SV} (Rayleigh wave data) and the inversion for V_{SH} (Love wave data) for the 492 tesselation points.*

to guide the model space search to areas of better fit. The χ^2 misfit between the observed absolute phase velocities and the calculated phase velocities for each velocity model is defined as

$$\chi^2 = \frac{1}{N} \sum_{i=1}^{N} \frac{(c_{obs,i}^{L,R} - c_i^{L,R})^2}{(\sigma_{obs,i}^{L,R})^2}, \qquad (5.1)$$

where $c_{obs}^{L,R}$ are the observed phase velocities for Love (L) and Rayleigh (R) respectively and $\sigma_{obs}^{L,R}$ are the model uncertainties for the phase velocity maps (Visser et al., 2007b). $c_i^{L,R}$ are the calculated phase velocities.

The nature of the model space search is determined by a few tuning parameters: the number of initial models (n_i), the number of iterations (n_{iter}), the number of new models sampled at each iteration (n_s) and the number of best misfit models at each iteration (n_r). At each iteration, the existing models are ranked according to their fit. In the Voronoi cells (nearest neighbourhood cells) of the n_r best fit models, n_s new models are randomly chosen after which all the models are ranked again according to their fit. The tuning parameters (n_r and n_s) determine how the model space is sampled. A large number for n_s and a small number for n_r leads to a very focused search, where the disadvantage is that some areas of good fit may be missed by this search. A large number for n_r (for example, equal to n_s) leads to a much broader (but also slower) search. For each point in the model space,

a velocity model is constructed using the coefficients for the shear wave velocity splines, the change in Moho depth and the scaling relations between the shear wave velocity and the compressional wave velocity and density. For this velocity model, we compute the exact local eigenfunctions for the specific surface wave modes in our data and obtain the phase velocities for these modes. The problem is highly non-linear and, therefore, we need a very broad search (n_i=100, n_{iter}=500, n_s=100 and n_r=100) so as not to miss any well fitting areas. The total number of sampled models is 50100 per inversion.

The model space is searched around a reference model. The reference model is PREM with the crust of the specific latitude-longitude location taken from Meier et al. (2007) and the topography and bathymetry information taken from CRUST-2.0 (Bassin et al., 2000). In the upper mantle we allow a change of $\pm$10%, in the transition zone a change of $\pm$5% and in the lower mantle a change of $\pm$2.5% with respect to the reference model. We, further, allow the Moho depth to vary by $\pm$5.0 km. The decrease in the model space size with depth is motivated by results from previous shear wave velocity modelling (Su and Dziewonski, 1997; Ritsema et al., 1999; Panning and Romanowicz, 2006). This first part of the Neighbourhood Algorithm produces an ensemble of velocity models with their corresponding fit (equation 5.1) to the observed phase velocities.

5.2.3   Bayesian information

The second part of the NA (Sambridge, 1999b) extracts information from the whole ensemble of models. It computes the conditional posterior probability density function ($P(\mathbf{m}|\mathbf{d})$) of the model ($\mathbf{m}$) given the data ($\mathbf{d}$) as

$$P(\mathbf{m}|\mathbf{d}) = \kappa\rho(\mathbf{m})L(\mathbf{m}|\mathbf{d}), \tag{5.2}$$

where $\rho(\mathbf{m})$ is the prior probability distribution which depends on the parameterisation, the search boundaries and the forward theory, κ is a normalisation constant and $L(\mathbf{m}|\mathbf{d})$ is a likelihood function representing the fit to the observations defined as $L(\mathbf{m}|\mathbf{d}) = \exp(-1/2\chi^2)$. The NA first constructs an approximate posterior probability density (PPD) for the ensemble of models by assuming constant known PPD values in the Voronoi cells and then performs a couple of random walks using a Gibbs sampler (Geman and Geman, 1984; Rothmann, 1986). After multiple random walks, the distribution will asymptotically resemble the approximate posterior probability density function. This resampled ensemble can be used in a Bayesian framework to infer information from the ensemble such as one or two-dimensional marginals and the covariance matrix. The one-dimensional
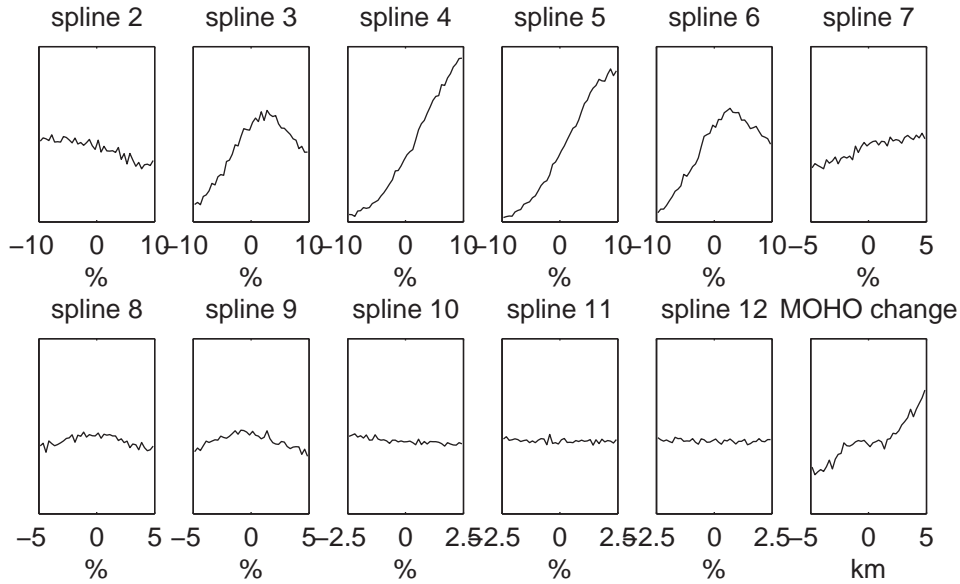
Figure 5.4: *One-dimensional marginals indicating the change in V_{SH} inversion parameters (Fig. 5.2) from the reference model at the Baltic Shield $(58°N, 17°E)$ location. The limits on the x-axis give the limit of the prior marginal.*

marginals of the separate V_{SH} and V_{SV} inversions can be jointly resampled to obtain one-dimensional marginals of anisotropic and isotropic anomalies. We define the Voight average isotropic shear wave velocity (Babuska and Cara, 1991) as

$$V_S^2 = \frac{2V_{SV}^2 + V_{SH}^2}{3},$$

(5.3)

and the shear wave anisotropy as

$$\xi = \frac{V_{SH}^2}{V_{SV}^2}.$$

(5.4)

5.3   A detailed example

We illustrate our approach with an example for a location on the Baltic Shield $(58°N, 17°E)$. We perform the Rayleigh and Love wave inversions and obtain one and two dimensional marginals that provide the full information on the entire ensemble of shear wave velocity models. Figures 5.4 and 5.5 show the one-
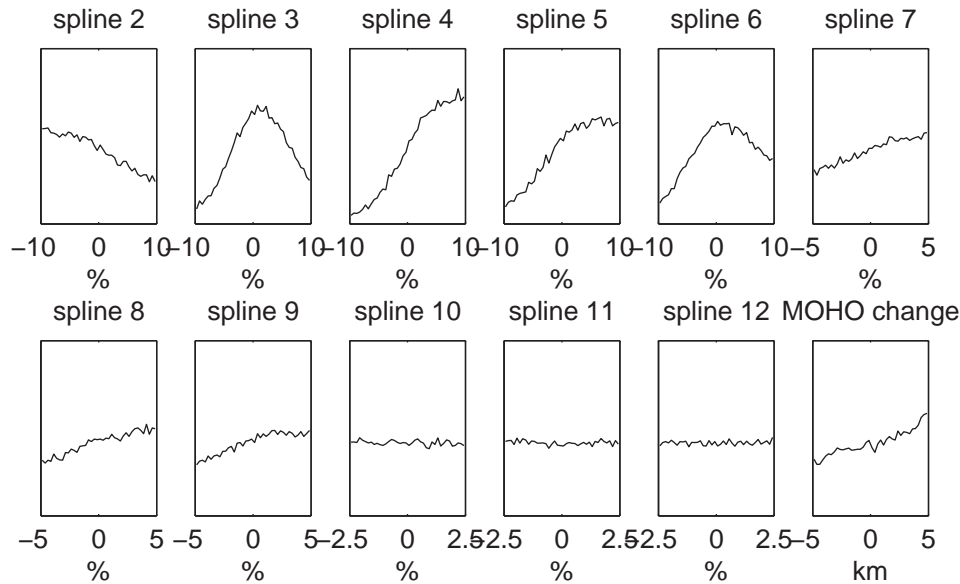
Figure 5.5: *One-dimensional marginals indicating the change in V_{SV} inversion parameters (Fig. 5.2) from the reference model at the Baltic Shield ($58°N, 17°E$) location. The limits on the x-axis give the limit of the prior marginal.*

dimensional marginals for the V_{SH} and V_{SV} inversions, respectively. The one-dimensional marginals show how well we are able to resolve the individual spline coefficients (Fig. 5.2). Spline coefficients three to six are relatively well resolved, there are clearly defined areas of higher probability, while spline coefficients ten to twelve are completely unresolved (flat). From this we can infer that at this location, we are able to resolve V_{SH} and V_{SV} best from 75 km to 400 km, but we are unable to resolve shear wave velocity from 800 km to 1500 km. Comparing the one-dimensional marginals for the V_{SH} and V_{SV} inversions, we notice that the areas of highest probability are quite similar for both inversion indicating modest anisotropy. The two-dimensional marginals (Fig. 5.6) are important to identify trade-offs which show as diagonal alignments. Trade-offs exist, but they are weak compared to our inability to resolve shear wave speed at certain depths. In general, the one-dimensional marginals (Fig. 5.4, 5.5) are not Gaussian, and not even symmetric which reflects the non-linearity of the problem. This means that we cannot represent them by a simple mean and a standard deviation.

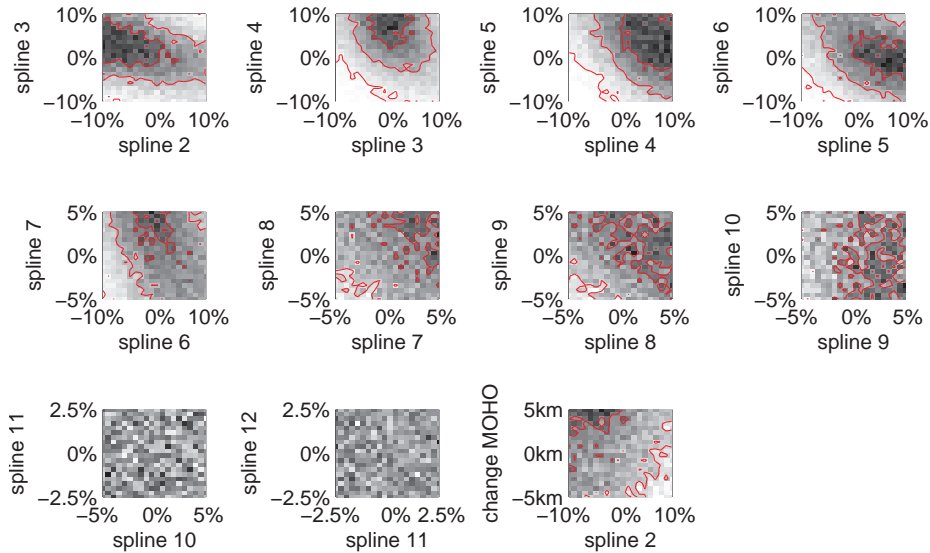We jointly resample the one-dimensional marginals for V_{SH} and V_{SV} to ob-

Figure 5.6: *Two-dimensional marginals for the inversion for V_{SV} at the Baltic Shield ($58°N, 17°E$) location. The contour lines indicate $1\sigma, 2\sigma$ and 3σ. Darker shading means higher probability. The limits on the axis give the limit of the prior marginal. The reference is the reference model for the specific point.*

tain the one-dimensional marginals for the isotropic shear wave velocity and radial anisotropy (equations 5.3 and 5.4). Figure 5.7 shows the one-dimensional marginals for anisotropy at the example location. Comparing the prior marginals (limits of the x-axis) with the posterior marginals, we have now obtained information on all spline coefficients (all marginals of the spline coefficients show a clearly defined maximum). This may seem surprising since the separate marginals for V_{SH} and V_{SV} show no information gain (flat marginals) for spline coefficients ten to twelve. The theorem for the association of probability density functions (see statistical textbooks) explain this. A particular case of association is the sum which is governed by the Central Limit Theorem. Here, the formation of ξ is more complicated and non-linear. Nevertheless, general properties remain the same: the result is a peaked probability density and its moments depend strongly on the individual spreads. It is difficult to have an intuition for the results, but for an unresolved spline for V_{SH} and V_{SV} (e.q. #9 in Fig. 5.7) the standard devivation for ξ in percent is more than twice the initial sampling interval. For a resolved spline (e.g. #3) the standard deviation for ξ in percent is the same as the initial sampling
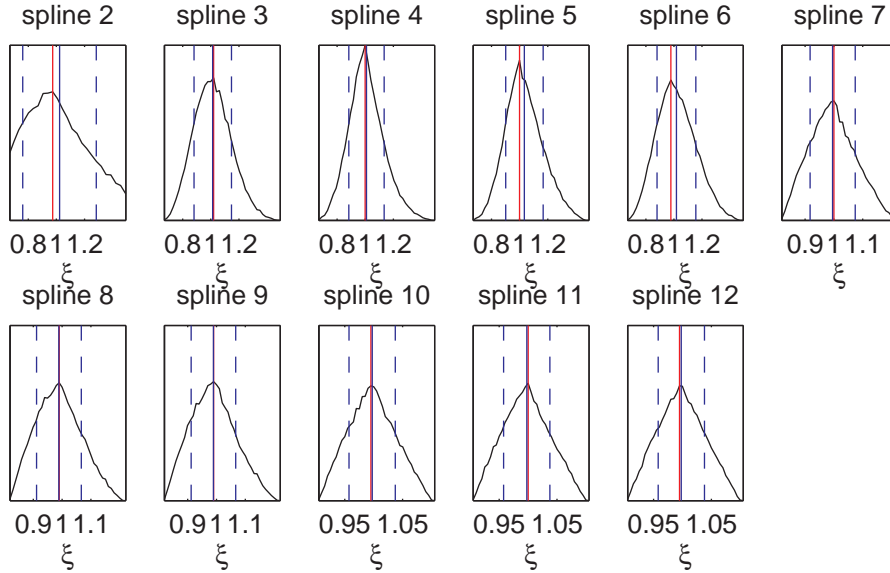
Figure 5.7: *One-dimensional posterior marginals for the anisotropic parameters at the Baltic Shield ($58°N$, $17°E$) location. The red line indicates the maximum value, the blue line indicates the mean and the dashed blue lines the mean $\pm$ one standard deviation. The limits on the x-axis give the limit of the prior marginal.*

interval. The standard deviations are not too meaningful, since most marginals are skewed, except for the ones corresponding to unresolved parameters. Still this gives a feeling of what to expect.

5.4 Spherically averaged anisotropy

We performed the depth inversion and obtained the one-dimensional marginals for anisotropy for all tesselation locations. The one-dimensional marginals at each location are now averaged to compute the spherically averaged anisotropy at each depth. We performed the sum by resampling the individual marginals. The result is governed by the Central Limit Theorem and therefore the average probability density at each depth is nearly Gaussian. It is thus meaningful to represent its mean and standard deviation. Figure 5.8 shows the spherically averaged anisotropy. In the anisotropic zone, the positive ($V_{SH} > V_{SV}$) spherically averaged anisotropy corresponds quite well to anisotropic PREM as well as the results
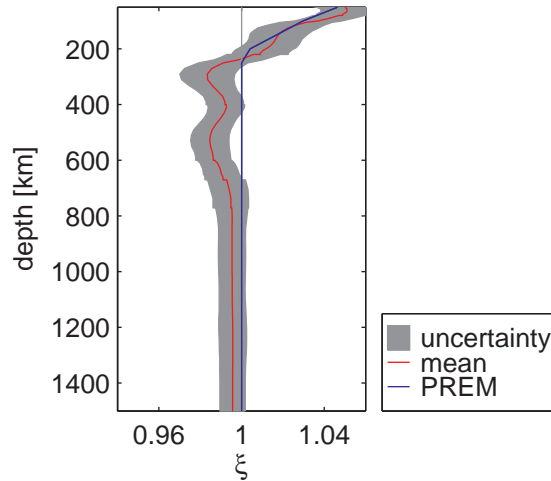
Figure 5.8: *Spherically averaged anisotropy. Also indicated are the 95% confidence levels (two standard deviations) and the anisotropic PREM model (Dziewonski and Anderson, 1981).*

obtained by previous studies (Montagner and Kennett, 1996; Beghein et al., 2006; Zhou et al., 2006). At 220 km, we observe a sign change in the average anisotropy from positive ($V_{SH} > V_{SV}$) to negative ($V_{SV} > V_{SH}$) anisotropy, which was also observed by Montagner and Kennett (1996); Beghein et al. (2006); Zhou et al. (2006), although Beghein et al. (2006) concluded that it is not significant due to the large uncertainties in their linearised inversion. We find significant (95% confidence or larger than two standard deviations) negative average anisotropy from 220 km down to the transition zone. The change in the sign of anisotropy could indicate a change from predominantly horizontal flow in the lithosphere and asthenosphere to predominantly vertical flow in the deeper mantle assuming that anisotropy is caused by the lattice preferred orientation of intrinsically anisotropic mantle minerals by finite strain due to mantle flow. The peak in negative anisotropy around 300 km was also observed by Zhou et al. (2006). The significant negative anisotropy continues through the transition zone which disagrees with Montagner and Kennett (1996) who found positive anisotropy in the transition zone. In the lower mantle, we find no significant average anisotropy in agreement with previous studies.
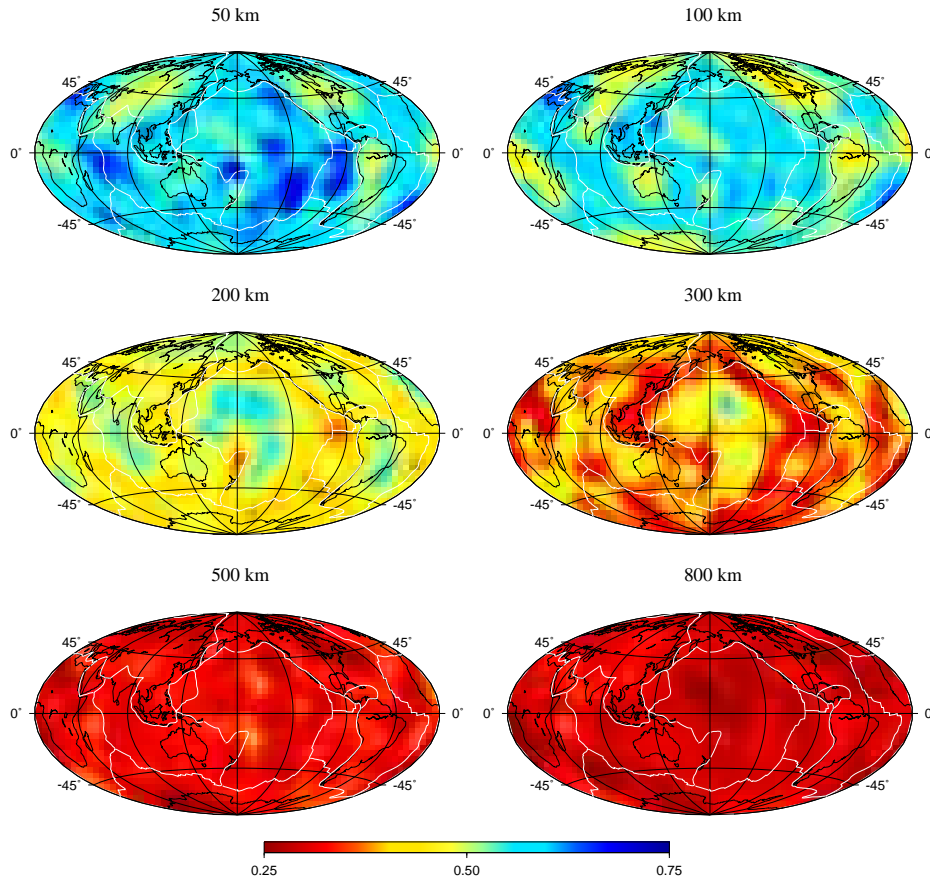
Figure 5.9: *Maps of probability of anisotropy ($V_{SH} > V_{SV}$).*

5.5 How probable is laterally varying anisotropy?

Our individual posterior probability density functions for ξ are clearly skewed (Fig. 5.7), which makes it difficult to represent them by a mean and a standard deviation. But our posterior probability density functions allow us to calculate the probability that V_{SH} is larger than V_{SV}, for instance, which is the area under the curve of the one-dimensional marginal for which ξ is larger than one. Figure 5.9 shows the distribution of the total probability of positive ($V_{SH} > V_{SV}$) anisotropy for various depths. Since the total area under a probability density function is one ($P(\xi > 1) + P(\xi < 1) = 1$), the low probabilities of positive anisotropy (Fig. 5.9) show the high probabilities of negative ($V_{SH} < V_{SV}$) anisotropy. In the anisotropic zone, we find a high probability of positive ($V_{SH} > V_{SV}$) anisotropy,
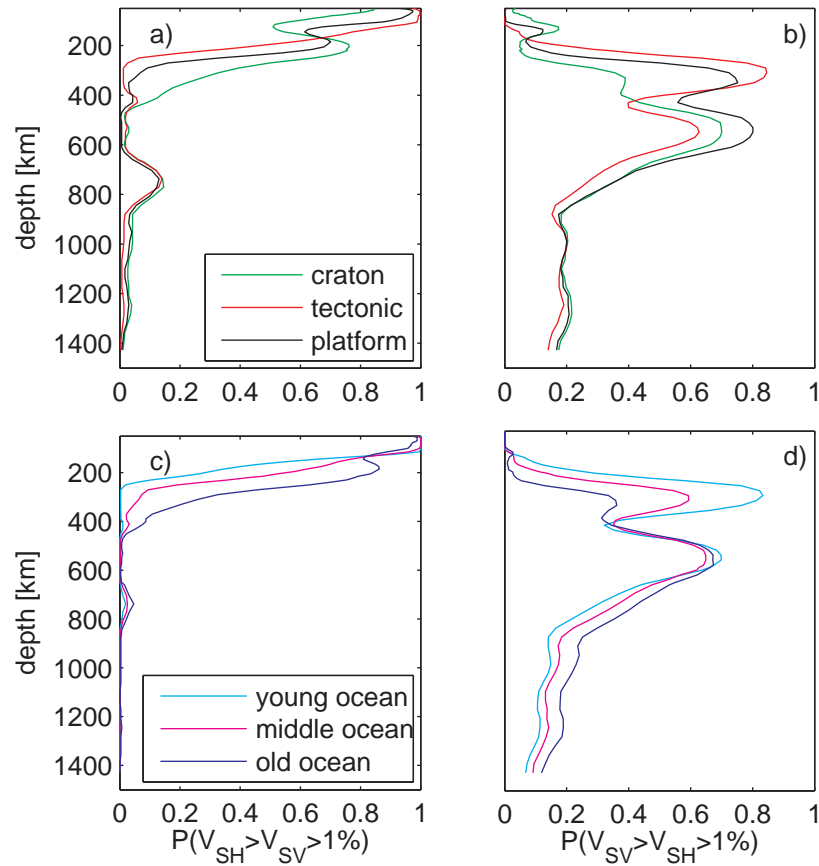
Figure 5.10: *The probability that the amplitude of anisotropy is larger than 1% for different tectonic areas. The definition of the tectonic areas is taken from 3SMAC (Nataf and Richard, 1996). Young oceans correspond to oceanic crust younger than 50Ma, middle oceans correspond to oceanic crust between 50Ma and 100Ma and old oceans correspond to oceanic crust older than 100Ma. The probability of a higher than 1% positive anisotropy for cratonic, tectonic and platform areas (a), young, middle and old oceans (c) and the probability of a higher than 1% negative anisotropy for cratonic, tectonic and platform areas (d), young, middle and old oceans (d).*

except for the cratonic areas. At 300 km, we see a change to a high probability for negative anisotropy associated mainly with subduction zones and mid-ocean ridges. Below the transition zone, we find a high probability of negative anisotropy, but this does not give any details about the amplitude of anisotropy.
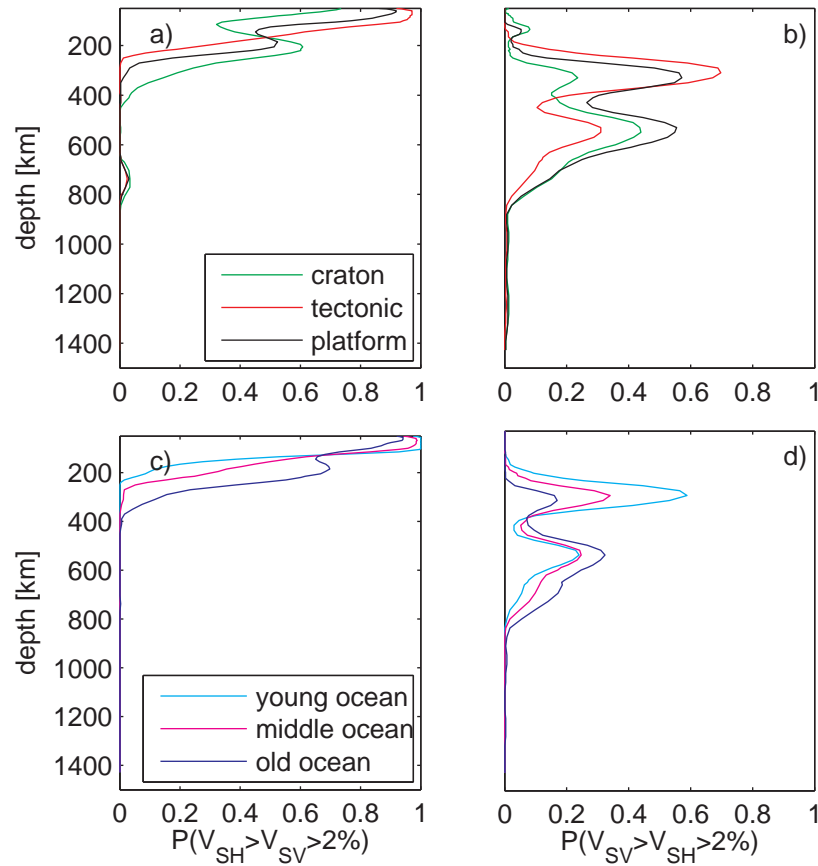
Figure 5.11: *The probability that the amplitude of anisotropy is larger than 2% for different tectonic areas. The definition of the tectonic areas is taken from 3SMAC (Nataf and Richard, 1996). Young oceans correspond to oceanic crust younger than 50Ma, middle oceans correspond to oceanic crust between 50Ma and 100Ma and old oceans correspond to oceanic crust older than 100Ma. The probability of a higher than 2% positive anisotropy for cratonic, tectonic and platform areas (a), young, middle and old oceans (c) and the probability of a higher than 2% negative anisotropy for cratonic, tectonic and platform areas (b), young, middle and old oceans (d).*

Just as easily, our marginals allow us to compute the probability that anisotropy is larger than 1% ($P(|\xi| > 1.01)$) or larger than 2% ($P(|\xi| > 1.02)$). Figures 5.10 and 5.11 show the probability of anisotropy with an amplitude larger than 1% and 2%, respectively for different tectonic regions (defined from 3SMAC, Nataf

and Richard, 1996). We computed the average probabilities over these regions by resampling the one-dimensional marginals. From this averaged one-dimensional marginal for the region, we computed the probability of anisotropy with an amplitude larger than 1% and 2%. Overall, we find the same pattern as for the average anisotropy; a high probability of positive anisotropy in the anisotropic zone, and a high probability of negative anisotropy down through the transition zone with two peaks around 300 km and 550 km. The probability of a high amplitude (>2%) positive anisotropy in the anisotropic zone is high (>0.95) (Fig. 5.11). The anisotropy in the transition zone is likely smaller in amplitude since only the probability of negative anisotropy with an amplitude larger than 1% is as high as 0.6-0.8. In the lower mantle, the probability that the amplitude of anisotropy is larger than 1% is exceedingly low. If present, the amplitude of anisotropy in the lower mantle is too small to be mapped with any confidence.

5.6   Discussion

In the uppermost mantle we find a high probability of anisotropy with fast horizontally polarized shear waves in the oceans and continents (Fig. 5.9). The amplitude of the anisotropy is likely to be large (>2%, figure 5.11). The probability of a large amplitude of anisotropy shows a difference between different regions. The oceanic areas show the highest probabilities, while the cratonic areas show the lowest probabilities down to 200 km (Fig. 5.10a, 5.11a). From 200 to 400 km, the cratonic areas and old oceans show a higher probability of positive ($V_{SH} > V_{SV}$) anisotropy. This corresponds roughly to an earlier observation by Gung et al. (2003), who found fast horizontally polarized shear wave anisotropy underneath oceans from 80 to 250 km and underneath cratons from 250 to 400 km. They explained this by a low-viscosity asthenospheric channel at different depths underneath oceans and continents.

From 200 km to 400 km we find prominent features of fast vertically polarized shear wave anisotropy at mid-ocean ridges and subduction zones (Fig. 5.9). The tectonic regions (Fig. 5.10, 5.11) and young oceanic regions show indeed much higher probability of negative ($V_{SV} > V_{SH}$) anisotropy from 200 to 400 km. The probability of a significant amplitude (>1%/2%) of negative ($V_{SV} > V_{SH}$) anisotropy (Fig. 5.10, 5.11) shows a peak at 300 km, for all tectonic areas. The probability that the amplitude of negative anisotropy is more than 1% is more than 0.8 for the young oceans and tectonic areas. The probability that the amplitude is larger than 2% is 0.6 for the mid-ocean ridges and 0.7 for the tectonic areas, indicating a possible amplitude difference between the mid-ocean ridges

and subduction zones. This agrees with the models of Gung et al. (2003); Panning and Romanowicz (2006); Zhou et al. (2006) who also found negative anisotropy associated with mid-ocean ridges and subduction zones at these depths. Zhou et al. (2006) found negative anisotropy at mid-ocean ridges visible from 120 km down to the transition zone. Figure 5.10d shows that the probability of negative anisotropy, with an amplitude larger than 1%, is different for young oceans and middle aged oceans from about 120 km down to the transition zone. This corresponds to the finding of Zhou et al. (2006).

If we assume that anisotropy is caused by the lattice preferred orientation of intrinsically anisotropic minerals under strain in the mantle, we observe evidence of predominantly horizontal flow in the anisotropic zone and predominantly vertical flow below. The horizontal flow in the lithosphere has probably been frozen in at the time of the formation of the lithosphere or at the last major episode of its deformation while the horizontal flow in the asthenosphere is probably due to plate motion. Down from about 120 km we observe evidence of vertical flow at mid-ocean ridges, and down from about 200 km we also observe evidence of vertical flow at subduction zones. The vertical flow associated with the mid-ocean ridges and subduction zones extends at least down to the transition zone.

In the transition zone we find in general a high probability of radial anisotropy with fast vertically polarized shear waves (P<0.40, Fig. 5.9). Panning and Romanowicz (2006) found anisotropy with fast vertically polarized shear waves associated with subduction zones in the transition zone. The total probability of large (>2%) anisotropy (Fig. 5.11b,d) shows a peak at 550 km, but is at most 0.5. The amplitude of anisotropy in the transition zone is likely between 1% and 2% (compare Fig. 5.10b,d and 5.11b,d). Also, the amplitude of negative anisotropy seems to be lower for the oceanic areas. The observed anisotropy in the transition zone could be explained by quasi vertical flow in the subduction zones. The mechanism of the anisotropy could be the alignment of spinel crystals or the alignment of pockets of strongly contrasting garnetite from oceanic crust (Karato, 1998a).

Although figure 5.9 shows a large probability of fast vertically polarized shear wave anisotropy, the probability of a significant amplitude (>1%) is low (Fig. 5.10 b,d). The lower mantle is most likely isotropic, which corresponds to earlier findings of Panning and Romanowicz (2006) and Meade et al. (1995). An isotropic lower mantle could be explained by superplastic flow (Karato, 1998a), because in superplastic flow the minerals do not align in preferred orientations. Even though the minerals themselves are highly anisotropic, seismic waves would see an isotropic lower mantle.

5.7   Conclusions

We performed the last step in a three-stage inversion (Yoshizawa and Kennett, 2002) for radially anisotropic structure of the mantle. In the first stage, we applied waveform fitting using a model space search approach to obtain fundamental and higher mode Love and Rayleigh wave phase velocity measurements (Visser et al., 2007a). The second stage (Visser et al., 2007b) consisted of inverting the fundamental and higher mode Love and Rayleigh phase velocity measurements to obtain global isotropic and azimuthally anisotropic phase velocity maps. In the third stage, presented here, we invert the isotropic phase velocity maps, including their uncertainties, for Love and Rayleigh waves separately to obtain a global V_{SH} and V_{SV} model. We invert the phase velocity maps using a fully non-linear model space search approach. We tested that we could invert Love and Rayleigh wave phase velocities separately. The model space search provides us with the whole ensemble of V_{SV} and V_{SH} models and we resample these ensembles to obtain an ensemble of isotropic and anisotropic models. Since we know not only the best anisotropic model but the whole ensemble of models we can compute the total probability of positive ($V_{SH} > V_{SV}$) or negative ($V_{SV} > V_{SH}$) anisotropy as well as compute the probability that the amplitude of anisotropy is above a certain amplitude (1%,2%).

We find a high probability of anisotropy with fast horizontally propagating shear waves (horizontal flow), in the upper mantle down to 200 km. For cratons, this fast horizontally propagating shear wave anisotropy (horizontal flow) is found down to 400 km. The amplitude of positive anisotropy in the uppermost mantle is likely to be large (>2%) in the lithosphere and decreases down to 200 km. In the lithosphere, the observed anisotropy could be related to anisotropy frozen in at the time of formation or last significant deformation. From about 120 km, we find a high likelihood of fast vertically polarized shear wave anisotropy (vertical flow) associated with mid-ocean ridges and from about 200 km the fast vertically polarized shear wave anisotropy is also associated with subduction zones. This extends down to the transition zone. The amplitude of this anisotropy just above the transition zone (300 km) is probably large (>2%). The transition zone is dominated by fast vertically polarized shear wave anisotropy (vertical flow), although the amplitude is likely lower (between 1% and 2%). The lower mantle appears isotropic.

5.8 Acknowledgments

Chapter 6

Summary and Conclusions

In this thesis we present all three stages of the inversion approach proposed by Kennett and Yoshizawa (2002). The three stage inversion approach consists of obtaining fundamental and higher mode Love and Rayleigh wave phase velocity measurements through waveform fitting in the first stage, combining them into multimode phase velocity models using the path average assumption in the second stage and an inversion for local shear wave speed properties to obtain a three-dimensional shear wave velocity model in the third stage. Yoshizawa and Kennett (2002) showed that multiple one-dimensional shear wave velocity models obtained through waveform fitting with a slight difference in misfit share the same dispersion characteristics, indicating that the phase velocity measurements may be regarded as a representation of the multimode dispersion characteristics along the source-receiver path. This is not as restricting as the assumption of the path average approximation in a regular two stage inversion approach where the shear wave velocity model is regarded as an average over the source receiver path. At each stage, particular care has been taken to assess the uncertainties.

In chapter 2, we present the first stage of the three stage inversion approach, which consists of measuring fundamental and higher mode Love and Rayleigh wave phase velocity measurements through waveform fitting. We present the fully automated procedure to measure the phase velocities and all the tests we performed to validate the procedure. For the waveform inversion we use a model space search approach. The advantage of a model space search approach is that it enables us to obtain consistent uncertainties for the phase velocity measurements. The model space is given by twelve natural cubic spline functions that together represent a one-dimensional shear wave velocity model. The synthetic seismogram, calculated from this shear wave velocity model, is compared to the

observed seismogram using a certain misfit criteria. For the model space search we use the Neighbourhood Algorithm (Sambridge, 1999a,b). The Neighbourhood Algorithm samples preferentially in areas of better fit. In the end, the model space search provides us not just with one one-dimensional shear wave velocity model but with an ensemble of one-dimensional shear wave velocity models and their corresponding fit to the seismogram. This ensemble is used to construct the posterior probability density function for phase velocities of specific modes. The one-dimensional marginals are Gaussian shaped, and we can therefore represent the phase velocity measurements by a mean and a standard deviation. We tested each step of the method extensively as described in chapter two. For example, we studied the parameterization (theoretical resolution, number and shape of the basis functions), misfit criteria, convergence of the model space search, prior information, use of Bayesian statistics and so on. An important issue is the dependence of the standard deviations on the range of the model space. This is solved by choosing the range of the model space such that our standard deviations agree with those from cluster analysis. Trampert and Woodhouse (2001) showed that uncertainties obtained by cluster analysis are in agreement with uncertainties obtained for comparing model predictions to real seismograms. This anchoring turns the consistent relative uncertainties to consistent absolute uncertainties.

Chapter three presents the Love wave phase velocity measurements. Love wave higher mode phase velocity measurements are more difficult to measure since the fundamental mode and higher modes travel closely together which causes an overlap of the fundamental mode and higher mode waveforms. The higher mode phase velocity measurements compare well to other studies (Van Heijst and Woodhouse, 1999; Lebedev et al., 2006), with 65% of our higher mode phase velocity measurements falling within one standard deviation compared to the measurements of Van Heijst and Woodhouse (1999). The resulting phase velocity maps agree well with the S20RTS model by Ritsema et al. (1999), except in the Pacific ocean between 1000 and 1500 km. Since S20RTS is based on Rayleigh equivalent waves this could indicate radial anisotropy ($V_{SH} < V_{SV}$) around the Pacific superplume.

In chapter four, we present the second stage of the three stage approach, where we build multimode phase velocity maps as a function of frequency using the path average assumption. We invert the fundamental and higher mode Love and Rayleigh wave phase velocities for global isotropic and azimuthally anisotropic phase velocity maps. Prior to inversion, we determine the optimum relative weighting for the isotropic and azimuthally anisotropic terms. We found that all fundamental and higher mode measurements require anisotropy. Spe-

cial care was taken to obtain the posterior model uncertainties, needed for the depth inversion in the next chapter. The isotropic higher mode models compare well to the models of Van Heijst and Woodhouse (1999) and the fundamental mode azimuthally anisotropic models correlate well with the models of Trampert and Woodhouse (2003) and hence with equivalent work from other research groups (Becker et al., 2007). We further examined the effect of spectral leakage, trade-offs between the isotropic and anisotropic terms and damping on the resolution matrix and found that the trade-offs are small. The average resolution for the isotropic models is of degree 25, the 2ψ models of degree 8 and the 4ψ models of degree 9. Indications are that the source of azimuthal anisotropy is complex and a detailed depth inversion is needed to clarify this.

Finally in chapter five, we present the last stage and invert the phase velocity maps to obtain a radially anisotropic shear wave velocity model using a model space search approach. For 492 locations on the Earth (equivalent to spherical harmonic degree 20), we invert the local azimuthally averaged phase velocity maps for radial anisotropy taking the full non-linearity into account. We separately inverted the Rayleigh wave phase velocities for a vertically polarized shear wave velocity model and the Love wave phase velocities for a horizontally polarized shear wave velocity model and combine the shear wave velocity models to obtain an isotropic and radially anisotropic shear wave velocity model. We checked that this separation was permissible. Since we use a model space search approach, we not only find the best model but the whole ensemble of models, giving the posterior probability density functions for the vertically and horizontally polarized shear wave velocity model and the isotropic and radially anisotropic shear wave velocity model. The one-dimensional marginals are not Gaussian and, therefore, cannot be represented by an average and a standard deviation. For the anisotropic model, we decided to compute the total probability of radial anisotropy and likewise that the amplitude of anisotropy is above 1% or 2%. We find a lithosphere dominated by fast horizontally polarized shear wave anisotropy (horizontal flow), with a significant amplitude (>2%) except underneath cratons. This anisotropy is probably frozen in at the time of the formation of the lithosphere. The asthenosphere is dominated by fast vertically polarized shear wave anisotropy (vertical flow), with a probability of more than 70% that the amplitude is large (>2%) at 300 km. The fast vertically polarized shear wave anisotropy is associated with mid-ocean ridges (from 120 km) and subduction zones (from 200 km). The transition zone is also dominated by fast vertically polarized shear wave anisotropy (vertical flow), although the amplitudes are likely smaller (1%-2%). Finally, the lower mantle appears to be mostly isotropic (or anisotropic with a very low am-

plitude, <1%), which corresponds to earlier findings.

This seems contrary to the observation in chapter three of radial anisotropy ($V_{SV} > V_{SH}$) in the Pacific. If we, however, compute the probability that $V_{SV} > V_{SH}$ for the Pacific area shown in chapter 3, we find a high probability (75 - 80 %) that this area is radially anisotropic. The probability that the amplitude of this radial anisotropy is larger than 1% is 30% at 1000 km and decreases to 20% at 1500 km. These probabilities are higher than the probabilities shown in figure 5.10 for other tectonic areas. Therefore the results in chapter five do not exclude radial anisotropy in the Pacific as observed in chapter three, a large amplitude (>1%) is however not very likely (30%). Furthermore, the isotropic model in chapter 3 is an average velocity model and can be seen as one realization, while the radially anisotropic model in chapter 5 represents a posterior probability density function which contains all possible anisotorpic velocity models.

So, in conclusion we have shown that measuring higher modes for single seismograms is possible up to the sixth higher mode using a waveform inversion. We obtained a large dataset of higher mode measurements (>350,000), using a fully automated approach and inverted them to obtain isotropic and azimuthally anisotropic phase velocity maps. We have also shown that we can invert the azimuthally averaged phase velocity maps to obtain a radially anisotropic shear wave velocity model, using a fully non-linear approach which was possible with up-to-date computing power. Throughout the whole study we have paid careful attention to the uncertainties, which proved vital in the search for significant radial anisotropy. For future work, we suggest to apply a detailed depth inversion using finite frequency kernels to the azimuthal anisotropic phase velocity maps to provide more information on the depth extent of significant azimuthal anisotropy.

Bibliography

Aki, K., Christofferson, A., and Husebye, E. (1977). Determination of the three-dimensional structure of the lithosphere. *J. Geophys. Res.*, 82:277–296.

Aki, K. and Kaminuma, K. (1963). Love waves from the Aleutian shock of March 9, 1957. *Bulletin of the Earthquake Research Institute*, 41:243–259.

Amaru, M. (2007). *Global travel time tomgraphy with 3-D reference models*. D. Phil. Thesis, Utrecht University.

Anderson, D. L. (1961). Elastic wave propagation in layered anisotropic media. *Journal of Geophys. Res.*, 66:2953–2963.

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

Bassin, C., Laske, G., and Masters, G. (2000). The current limits of resolution for surface wave tomography in North America. *EOS Trans AGU*, 81:F897.

Becker, T. W., Ekström, G., Boschi, L., and Woodhouse, J. H. (2007). Length scales, patterns, and origin of azimuthal seismic anisotropy in the upper mantle as mapped by Rayleigh waves. *Geophys. J. Int.*, 171:451–462.

Beghein, C., Trampert, J., and van Heijst, H. J. (2006). Radial anisotropy in seismic reference model of the mantle. *Journal of Geophys. Res.*, 111:B02303, doi:10.1029/2005JB003728.

Beucler, E. and Montagner, J. (2006). Computation of large anisotropic seismic heterogeneities (CLASH). *Geophys. J. Int.*, 165:447–468.

Beucler, E., Stutzmann, E., and Montagner, J. P. (2003). Surface wave higher mode phase velocity measurements using a roller-coaster type algorithm. *Geophys. J. Int.*, 155:289–307.

Bevington, P. R. and Robinson, D. K. (1992). *Data reduction and error analysis for the physical sciences*. McGraw-Hill, New York.

Bijwaard, H. and Spakman, W. (1999). Tomographic evidence for a narrow whole mantle plume below iceland. *Earth Planet. Sci. Lett.*, 166:121–126.

Bijwaard, H. and Spakman, W. (2000). Non-linear global p-wave tomography by iterated linearized inversion. *Geophys. J. Int.*, 141:71–82.

Bijwaard, H., Spakman, W., and Engdahl, E. (1998). Closing the gap between regional and global travel time tomography. *J. Geophys. Res.*, 103:30,055–30,078.

Boschi, L. (2003). Measures of resolution in global body wave tomography. *Geophys. Res. Lett.*, 30:doi: 10.1029/2003GL018222.

Boschi, L. and Dziewonski, A. (1999). High- and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling. *J. Geophys. Res.*, 104:25,567–25,594.

Bozdag, E. and Trampert, J. (2007). On crustal corrections in surface wave tomography. in press.

Capdeville, Y., Stutzmann, E., and Montagner, J. P. (2000). Effect of a plume on lon-period surface waves computed with normal mode coupling. *Physics of the*

Earth and Plantary interiors, 119:57–74.

Cara, M. (1978). Regional variations of higher Rayleigh-mode phase velocities:a spatial-filtering method. *Geophys. J. 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. and Leveque, J. (1987). Waveform inversion using secondary observables. *Geophys. Res. Lett.*, 10:1046–1049.

Dahlen, F. A. and Tromp, J. (1998). *Theoretical global seismology*. Princeton University Press, Princeton.

Deal, M., Nolet, G., and van der Hilst, R. (1999). Slab temperature and thickness from seismic tomography 1. method and application to Tonga. *J. Geophys. Res.*, 104:28,789–28,802.

Debayle, E. and Kennett, B. L. N. (2000). Anisotropy in the Australian upper mantle from Love and Rayleigh waveform inversion. *Earth and Planetary Science Letters*, 184:339–351.

Deschamps, F., Snieder, R., and Trampert, J. (2001). The relative density to shear velocity scaling in the uppermost mantle. *Physics of the Earth and Plantary Interiors*, 124:193–211.

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

Durek, J. J., Ritzwoller, M. H., and Woodhouse, J. H. (1993). Constraining upper mantle anelasticity using surface wave amplitudes. *Geophys. J. Int.*, 114:249–272.

Dziewonski, A. M. (1984). Mapping the lower mantle: deformation of lateral heterogeneity in P velocity up to degree and order 6. *J. Geophys. Res.*, 89:5929–5952.

Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary Reference Earth model. *Physics of the Earth and Plantary Interiors*, 25:297–356.

Dziewonski, A. M., Hager, B. H., and O'Connell, R. J. (1977). Large-scale heterogeneities in the lower mantle. *J. Geophys. Res.*, 82:239–255.

Ekström, G. (2000). Mapping the lithosphere and asthenosphere with surface waves: Lateral structure and anisotropy. *The History and Dynamics of Global Plate Motions, Geophys. Monogr. Ser.*, 121:211–238.

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

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

Ferreira, A. M. G., Woodhouse, J. H., Visser, K., and Trampert, J. (2007). Inverstigating the anisotropic shear wave velocity structure of the Earth's mantle. talk given at the American Geophysical Union meeting.

Forsyth, D. W. (1975). The early structural evolution and anisotropy of the oceanic upper mantle. *Geophys. J. R. astr. Soc.*, 43:103–162.

Fukao, Y. (1992). Seismic tomography of the Earths mantle - geodynamic implications. *Science*, 258:625–630.

Geman, A. E. and Geman, D. (1984). Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Trans. Patt. Analysis Mach. Int.*,

6:721–741.

Godey, S., Deschamps, F., Trampert, J., and Snieder, R. (2004). Thermal and compositional anomalies beneath the north-american continent. *J. Geophys. Res.*, 109:B01308, doi: 10.1029/2002JB002263.

Gung, Y., Panning, M., and Romanowicz, B. (2003). Global anisotropy and the thickness of continents. *Nature*, 422:707–711.

Hess, H. H. (1964). Seismic anisotropy of the uppermost mantle under oceans. *Nature*, 203:629–631.

Ishii, M. and Tromp, J. (1999). Normal-mode and free-air gravity constraints on lateral variations in velocity and density of earth's mantle. *Science*, 285:1231–1236.

Ishii, M. and Tromp, J. (2001). Even-degree lateral variations in the Earth's mantle constrained by free oscillations and the free-air gravity anomaly. *Geophys. J. Int.*, 145:77–96.

Kaban, M. and Schwintzer, P. (2001). Oceanic upper mantle structure from experimental scaling of vs and density at different depths. *Geophys. J. Int.*, 147:199–214.

Kárason, H. and van der Hilst, R. D. (2001). Tomograhic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (PKP, Pdiff). *J. Geophys. Res.*, 106:6569–6587.

Karato, S. (1998a). Seismic anisotropy in the deep mantle, boundary layers and the geometry of mantle convection. *Pure appl. Geophys.*, 151:565–587.

Karato, S.-I. (1998b). Some remarks on the origin of seismic anisotropy in the D" layer. *Earth Planets Space*, 50:1019–1028.

Kendall, J.-M. and Silver, P. G. (1996). Constraints from seismic anisotropy on the nature of the lowermost mantle. *Nature*, 381:409–412.

Kennett, B., Widiyantoro, S., and van der Hilst, R. (1998). Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle. *J. Geophys. Res.*, 103:12,469–12,493.

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

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

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

Lévêque, J., Rivera, L., and Wittlinger, G. (1993). On the use of checkerboard test to asses the resolution of tomogrphic inversions. *Geophys. J. Int.*, 115:313–318.

Larson, E. W. F., Tromp, J., and Ekström, G. (1998). Effects of slight anisotropy on surface waves. *Geophys. J. Int.*, 132:654–666.

Lebedev, S., Meier, T., and van der Hilst, R. D. (2006). Asthenospheric flow and origin of volcanism in the Baikal Rift area. *Earth and Planetary Science Letters*. in press.

Lebedev, S., Nolet, G., Meier, T., and van der Hilst, R. D. (2005). Automated multimode inversion of surface and S waveforms. *Geophys. J. Int.*, 162:951–964.

Li, X. D. and Romanowicz, B. (1996). Global mantle shear wave velocity model developed using nonlinear asymptotic mode coupling theory. *J. Geophys.*

Res., 101:22,245–22,272.

Love, A. E. H. (1927). *A treatise on the Theory of Elasticity*. Cambridge University Press.

Maggi, A., Debayle, E., Priestley, K., and Barruol, G. (2006). Multimode surface waveform tomography of the Pacific Ocean: a closer look at the lithospheric cooling signature. *Geophys. J. Int.*, 166:1384–1397.

Marone, F., Gung, Y., and Romanowicz, B. (2007). Three-dimensional radial anisotropic structure of the North-American upper mantle from inversion of surface waveform data. *Geophys. J. Int.*, 171:206–222.

Marquering, H. and Snieder, R. (1996). Shear-wave velocity structure beneath europe, the northeastern atlantic and western asia from waveform inversions including surface wave mode coupling. *Geophys. J. Int.*, 124:258–278.

Masters, G., Laske, G., Bolton, H., and Dziewonski, A. (2000). The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure. *Geophysical Monograph*, 117:63–87.

McEvilly, T. V. (1964). Central US crust-upper mantle structure from Love and Rayleigh wave phase velocity inversion. *Bulletin of the Seismological Society of America*, 54:1997–2015.

Meade, C., Silver, P. G., and Kaneshima, S. (1995). Laboratory and seismological observations of lower mantle isotropy. *Geophys. Res. Lett.*, 22:1293–1296.

Meier, U., Curtis, A., and Trampert, J. (2007). Fully nonlinear inversion of fundamental mode surface waves for a global crustal model. *Geophys. Res. Lett.*, 34:L16304,

doi:10.1029/2007/GL030989.

Mitchel, R. (1980). Array measurements of higher mode Rayleigh wave dispersion: an approach utilizing source parameters. *Geophys. J. R. atr. Soc.*, 63:311–331.

Montagner, J. and Tanimoto, T. (1990). Global anisotropy in the upper mantle inferred from the regionalization of phase velocities. *J. Geophys. Res.*, 95:4794–4819.

Montagner, J. P. (1986). 3-Dimensional structure of the Indian Ocean inferred from long period surface waves. *Geophys. Res. Lett.*, 13:315–318.

Montagner, J. P. (1998). Where can seismic anisotropy be detected in the Earth's mantle? *Pure appl. geophysics*, 151:223–256.

Montagner, J. P. and Jorbert, N. (1988). Vectorial tomography - ii. Application to the Indian Ocean. *Geophys. J. Int.*, 94:309–344.

Montagner, J. P. and Kennett, B. L. N. (1996). How to reconcile body-wave and normal-mode reference earth models. *Geophys. J. Int.*, 125:229–248.

Montagner, J. P. and Nataf, H. C. (1986). A simple method for inverting the azimuthal anisotropy of surface waves. *Journal of Geophys. Res.*, 91:511–520.

Montagner, J. P. and Tanimoto, T. (1991). Global upper mantle tomography of seismic velocities and anisotropies. *J. Geophys. res.*, 96:20337–20351.

Montelli, R., Nolet, G., and Dahlen, F. A. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303:338 – 343.

Mooney, W. D., Krasnova, M., Chulick, G. S., and Detweiler, S. (2004). Crust2.1 : A revised global crustal model at 2 degrees.

Mooney, W. D., Laske, G., and Masters, G.

(1998). CRUST 5.1: a global crustal model at 5 degrees by 5 degrees. *J. Geopys. Res.*, 136:727–748.

Nataf, H. C., Nakanishi, I., and Anderson, D. L. (1984). Anisotropy and shear velocity heterogeneity in the upper mantle. *Geophys. Res. Lett.*, 11:109–112.

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

Nataf, H.-C. and Richard, Y. (1996). 3SMAC: an a priori tomographic model of the upper mantle based on geophysical modeling. *Phys. Earth. Planet. Int.*, 95:101–122.

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

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 waveform inversion and two-dimensional structure under the network of autonomously recording seismographs. *J. Geophys. Res.*, 95:8499–8512.

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

Okal, E. and Jo, B. (1987). Stacking investigations of the dispersion of higher order mantle rayleigh waves and normal modes. *Physics of the Earth and Planetary Interiors*, 47:188–204.

Panning, M. and Romanowicz, B. (2006). A three-dimensional radially anisotropic model of shear velocity in the whole mantle. *Geophys. J. Int.*, 167:361–379.

Panning, M. P. and Romanowicz, B. A. (2004). Inferences on flow at the base of the Earth's mantle based on seismic anisotropy. *Science*, 303:351–353.

Resovsky, J. and Trampert, J. (2003). Using probabilistic tomography to test mantle velocity-density relationships. *Earth Planet. Sci. Lett.*, 215:121–134.

Ritsema, J. and Van Heijst, H. (2002). Constraints on the correlation of P- and S-wave velocity heterogeneity in the mantle from P,PP,PPP and PKPab traveltimes. *Geophys. J. Int.*, 149:482–489.

Ritsema, J., van Heijst, H. J., and Woodhouse, J. H. (1999). Complex shear wave velocity structure imaged beneath Africa and Iceland. *Science*, 286:1925–1928.

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

Romanowicz, B. (1995). A global tomographic model of shear attenuation in the upper-mantle. *J. Geophys. Res.*, 100:12,375–12,394.

Romanowicz, B. (2003). Global mantle tomography: Progress status in the past 10 years. *Annu. Rev. Earth Planet. Sci.*, 31:303–328.

Romanowicz, B. and Snieder, R. (1988). A new formalism for the effect of lateral heterogeneity on normal modes and surface waves, ii: General anisotropic perturbations. *Geophys. J. R. atr. Soc.*, 93:91–99.

Rothmann, D. H. (1986). Automatic estimation of large residual statics corrections. *Geophysics*, 51:332–346.

Saltzer, R., van der Hilst, R., and Karason, H. (2001). Comparing P and S wave het-

erogeneity in the mantle. *Geophys. Res. Lett.*, 28:1335–1338.

Sambridge, M. (1999a). Geophysical inversion with a neighbourhood algorithm-I. Searching a parameter space. *Geophys. J. Int.*, 138:479–494.

Sambridge, M. (1999b). Geophysical inversion with a neighbourhood algorithm-II. Appraising the ensemble. *Geophys. J. Int.*, 138:727–746.

Sebai, A., Stutzmann, E., Montagner, J. P., Sicilia, D., and Beucler, E. (2006). Anisotropic structure of the African upper mantle from Rayleigh and Love wave tomography. *Physics of the Eath and Planetary Interiors*, 155:48–62.

Selby, N. D. and Woodhouse, J. H. (2002). The Q structure of the upper mantle: constraints from Rayleigh wave amplitudes. *J. Geophys. Res.*, 107:933–940.

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

Sieminski, A., Liu, Q., Trampert, J., and Tromp, J. (2007). Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods. *Geophys. J. Int.*, 168:doi:10.1111/j.1365–246X.2006.03261.x.

Simons, F., van der Hilst, R., Montagner, J., and Zielhuis, A. (2002). Multimode Rayleigh wave inversion for heterogeneity and azimuthal anisotropy of the Australian upper mantle. *Geophys. J. Int.*, 151:738–754.

Smith, D. B., Ritzwoller, M. H., and Shapiro, N. M. (2004). Stratification of anisotropy in the Pacific upper mantle. *Journal of Geophys. Res.*, 109:B11309, doi:10.1029/2004JB003200.

Smith, M. L. and Dahlen, F. A. (1973).

The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium. *J. Geophys. Res.*, 78:3321–3333.

Smith, M. L. and Dahlen, F. A. (1975). Correction to: 'The azimuthal dependence of Love and Rayleigh wave propagation in a slightly anisotropic medium'. *J. Geophys. Res.*, 80:1923.

Snieder, R., Beckers, J., and Neele, F. (1991). The effect of small-scale structure on normal mode frequencies and global inversions. *J. Geophys. Res.*, 59:501–515.

Snieder, R. and Trampert, J. (2000). *Linear and nonlinear inverse problems, in Geomatic methods for the analysis of data in the earth sciences*. Springer, Berlin.

Soldati, G. and Boschi, L. (2005). The resolution of whole Earth seismic tomographic models. *Geophys. J. Int.*, 161:143–153.

Spetzler, J. and Trampert, J. (2003). Implementing spectral leakage corrections in global surface wave tomography. *Geophys. J. Int.*, 155:532–538.

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

Su, W. and Dziewonski, A. (1997). Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle. *Physics of the Earth and Plantary Interiors*, 100:135–156.

Tanimoto, T. and Anderson, D. L. (1984). Mapping convection in the mantle. *Geophys. Res. Lett.*, 11:287–290.

Tanimoto, T. and Anderson, D. L. (1985). Lateral heterogeneity and azimuthal anisotropy of the upper mantle: Love and Rayleigh waves 100-250s. *J. Geophys. Res.*, 90:1842–1858.

Trampert, J. and Snieder, R. (1996). Model estimates biased by truncated expansions: Possible artifacts in seismic tomography. *Science*, 271:1257–1260.

Trampert, J. and van Heijst, H. J. (2002). Global azimuthal anisotropy in the transition zone. *Science*, 296:1297–1299.

Trampert, J. and 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. and Woodhouse, J. H. (2001). Assesment of global phase velocity models. *Geophy. J. Int.*, 144:165–174.

Trampert, J. and Woodhouse, J. H. (2003). Global anisotropic phase velocity maps for fundamental mode surface waves between 40 and 150 s. *Geophys. J. Int.*, 154:154–165.

van der Hilst, R. and Kárason, H. (1999). Compositional heterogeneity in the bottom 1000 kilometers of earth's mantle: toward a hybrid convection model. *Science*, 283:1885–1887.

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

Van Heijst, H. J. (1997). *New constraints on the seismic structure of the Earth from surface wave overtone phase velocity measurements*. D. Phil. Thesis, University of Oxford.

Van Heijst, H. J. and Woodhouse, J. H. (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. and Johnson, L. (1998). Whole earth structure estimated from seismic arrival times. *J. Geophys. Res.*, 103:2633–2671.

Vinnik, L. and P., M. J. (1996). Shear wave splitting in the mantle Ps phases. *Geophys. Res. Lett.*, 23:2449–2452.

Vinnik, L. P. Farra, V. and Romanowicz, B. (1989). Azimuthal anisotropy in the Earth from observations of SKS at GEOSCOPE and NARS broad-band stations. *Geophys. J. Int.*, 79:1542–1558.

Visser, K., Lebedev, S., Trampert, J., and Kennett, B. L. N. (2007a). Global love wave overtone measurements. *Geophysical Research Letters*, 34:L03302, doi:10.1029/2006GL028671.

Visser, K., Trampert, J., and Kennett, B. L. N. (2007b). Global anisotropic phase velocity maps for higher mode Love and Rayleigh waves. accepted.

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

Wessel, P. and Smith, W. H. F. (1995). New version of the generic mapping tools released. *EOS Trans. AGU*, 76:329.

Widiyantoro, S., Kenett, B., and van der Hilst, R. (1999). Seismic tomography with P and S data reveals lateral variations in the rigidiry of deep slabs. *Earth and Planetary Science Letters*, 173:91–100.

Woodhouse, J. H. and Dziewonski, A. M. (1984). Mapping the upper mantle: three-dimensional modeling of the Earth structure by inversion of seismic waveforms. *J. Geophys. Res.*, 89:5953–5986.

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

Zhao, D. (2001). Seismic structure and origin of hotspots and mantle plumes.

Earth Planet. Sci. Lett., 192:251–256.

Zhou, Y., Nolet, G., Dahlen, F. A., and Laske, G. (2006). Global upper-mantle structure from finite-frequency surface wave tomography. *J. Geophys. Res.*, 111:B04304, doi:10.1029/2005JB003677.

Appendix A

Isotropic phase velocity maps

This appendix shows some of the isotropic phase velocity models which we obtained in chapter 4. We show the phase velocity models at selected periods for the fundamental mode up to the sixth higher mode Rayleigh and the fundamental mode up to the fifth higher mode Love surface waves.
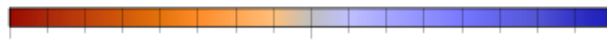
Figure A.1: *Color scale of figures in appendix A, the maximum amplitude is given in percent above each figure.*

35.099s (6%)   40.197s (6%)

51.259s (6%)   62.305s (6%)

78.129s (6%)   100.393s (5%)

114.372s (4.5%)   131.097s (4.5%)

151.193s (4.0%)   175.393s (4.0%)
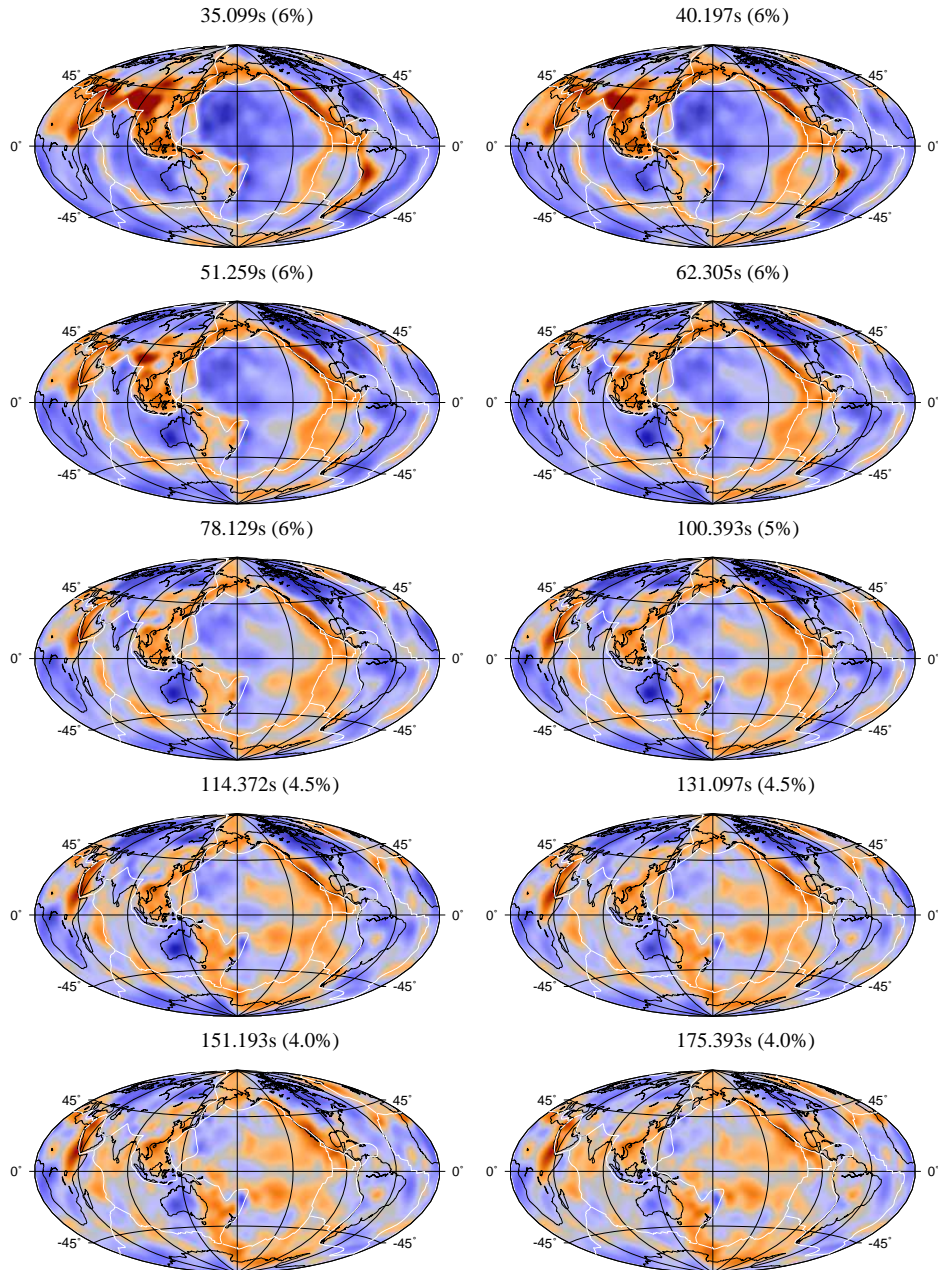
Figure A.2: *Relative isotropic phase velocity maps with respect to PREM for the fundamental mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.095s (5.0%)   40.142s (5.0%)

50.855s (5.0%)   61.868s (5.0%)

77.575s (5.0%)   99.650s (4.5%)

113.719s (4.5%)   129.311s (4.5%)

148.560s (4.0%)   172.519s (3.5%)
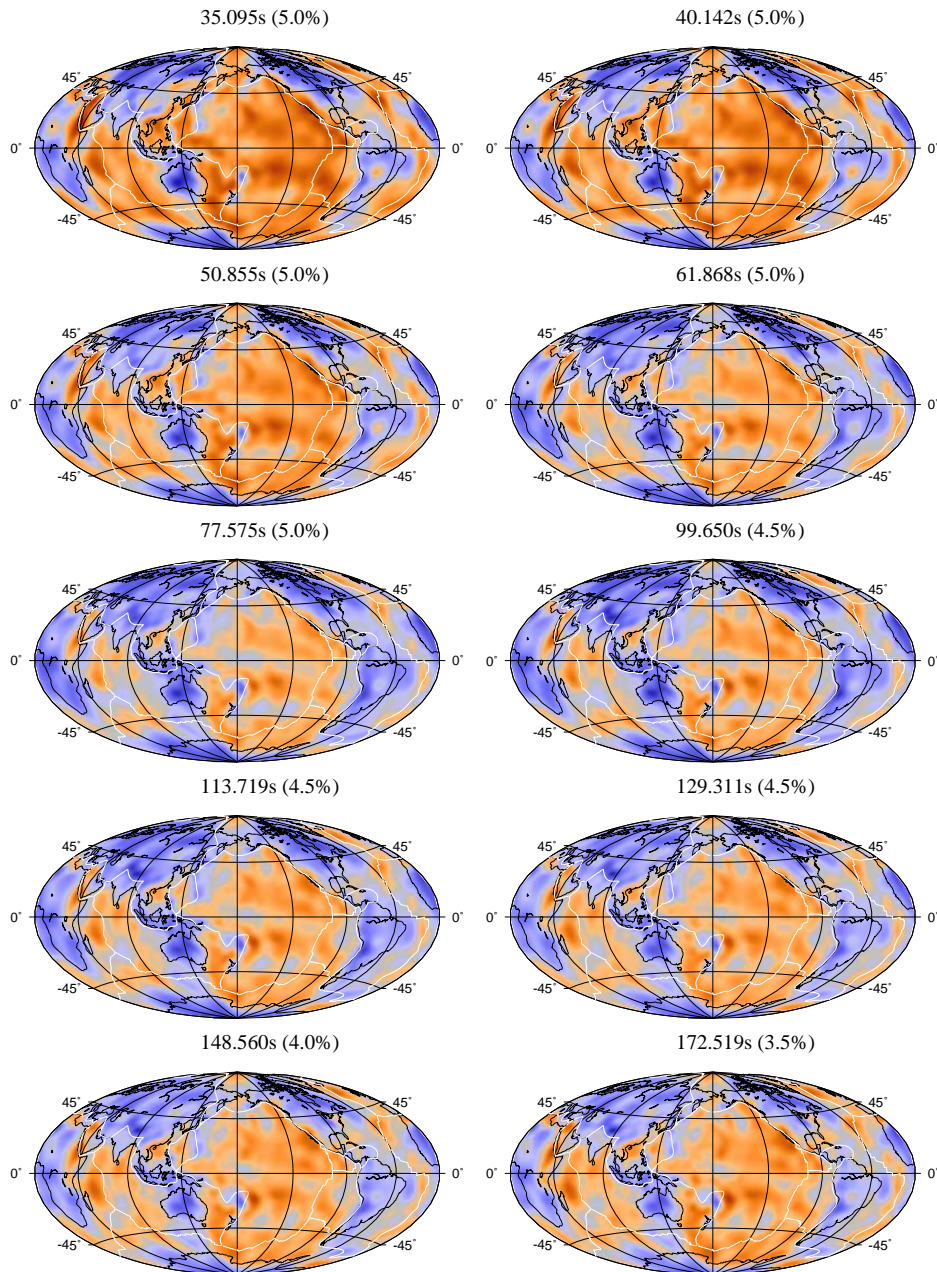
Figure A.3: *Relative isotropic phase velocity maps with respect to PREM for the first higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.018s (4.0%)   40.028s (4.0%)

50.849s (4.0%)   61.990s (4.0%)

77.618s (4.0%)   99.258s (3.5%)

113.551s (3.5%)   128.635s (3.5%)

149.275s (3.5%)
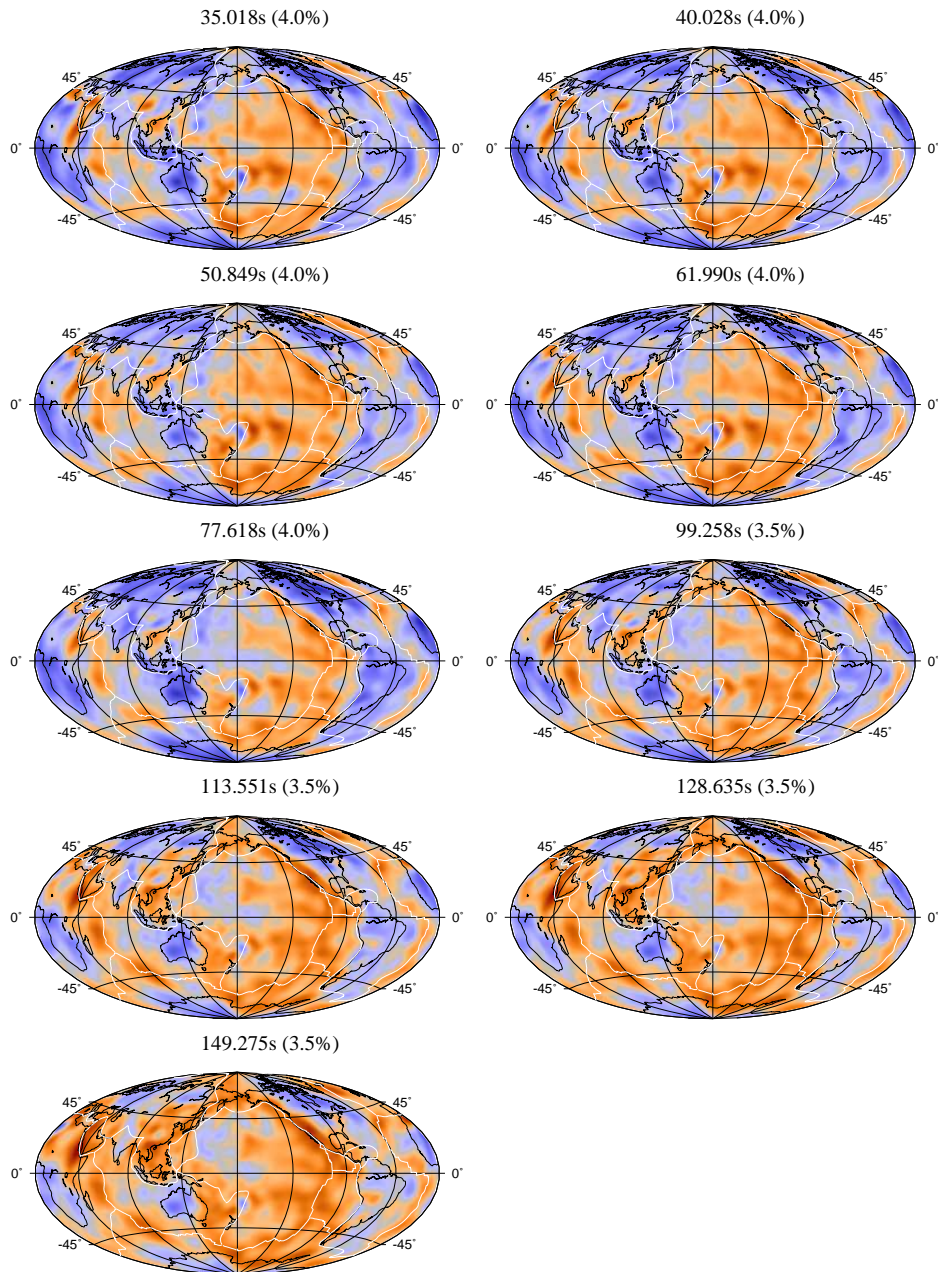
Figure A.4: *Relative isotropic phase velocity maps with respect to PREM for the second higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*
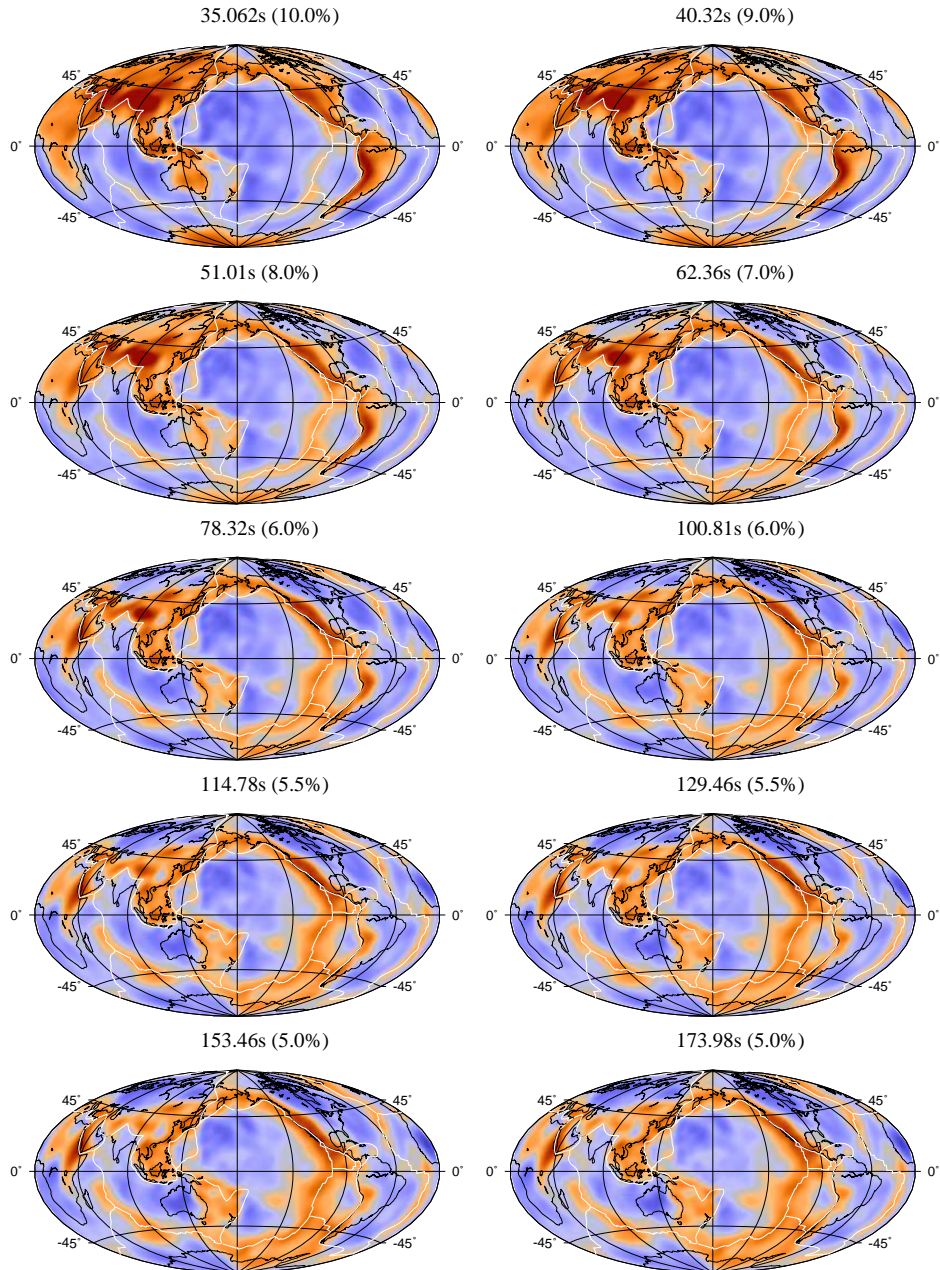
35.014s (3.5%)   40.026s (3.5%)

51.059s (4.5%)   61.799s (4.5%)

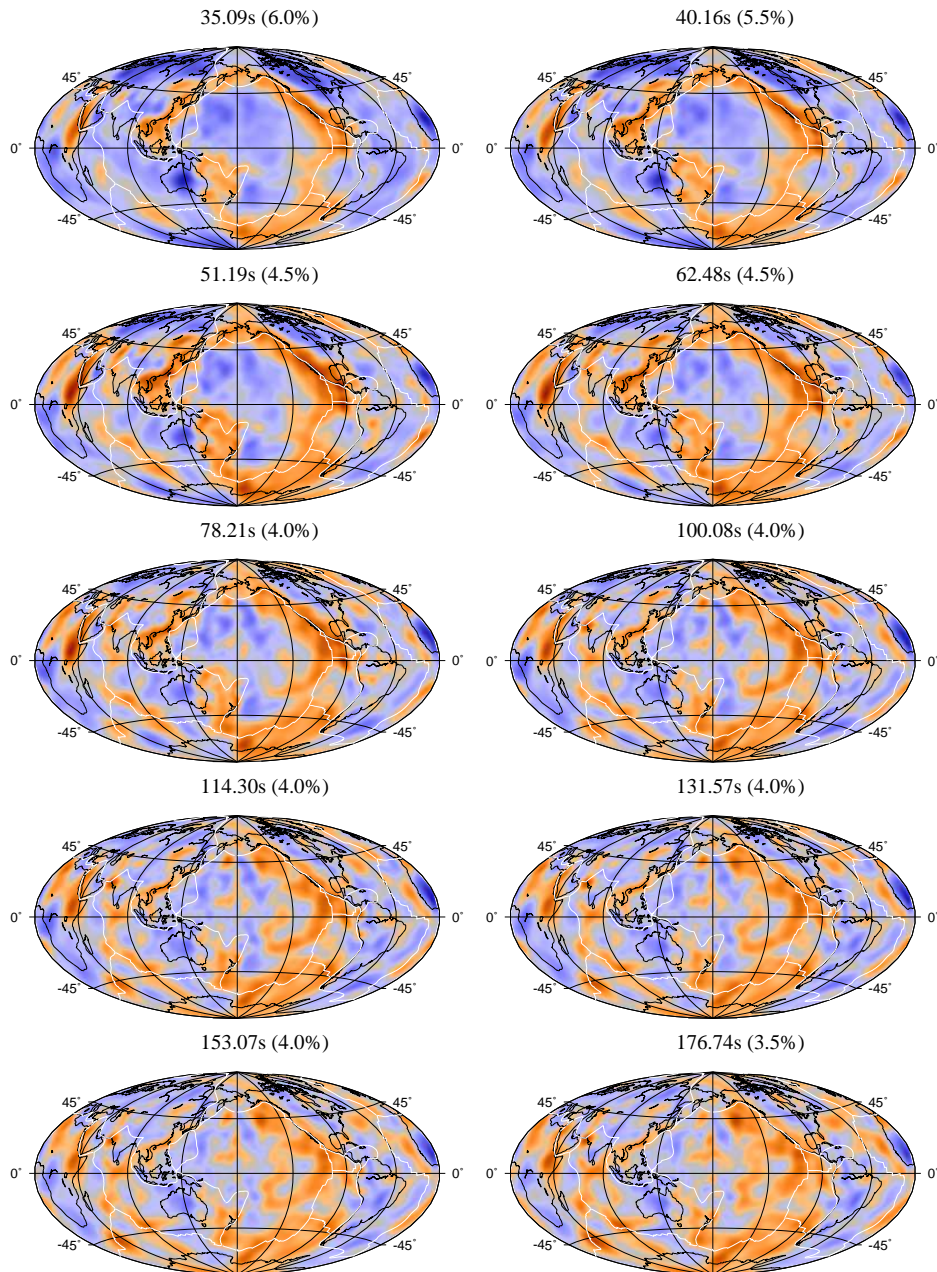77.795s (4.0%)   87.727s (4.0%)
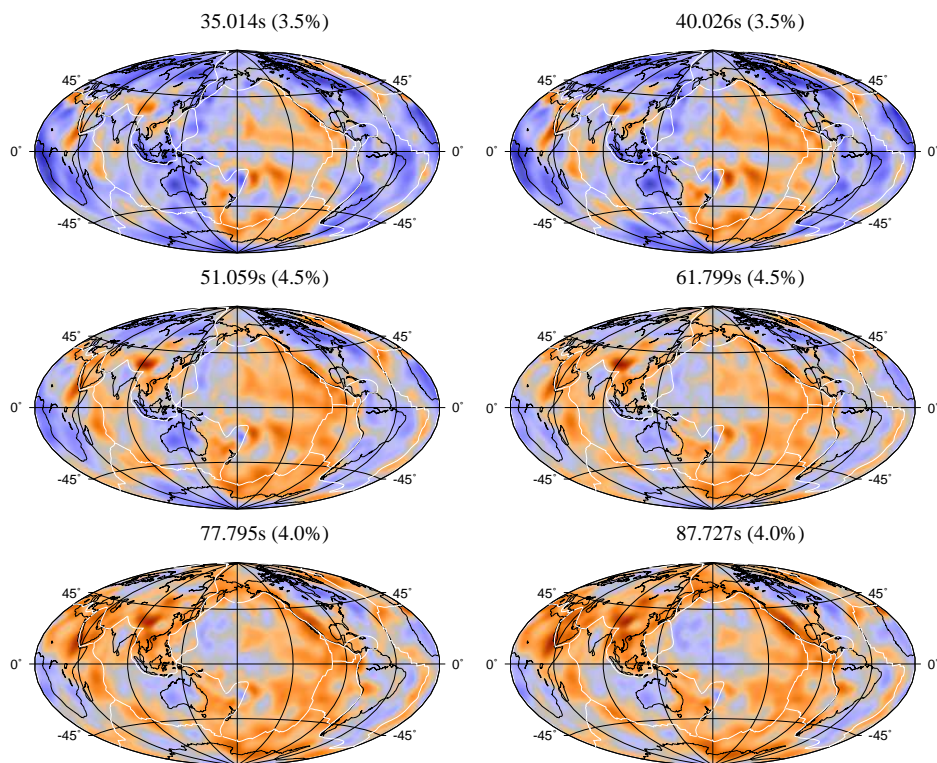
Figure A.5: *Relative isotropic phase velocity maps with respect to PREM for the third higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.0784s (4.5%)   40.100s (5.0%)

51.052s (4.5%)   61.871s (4.5%)
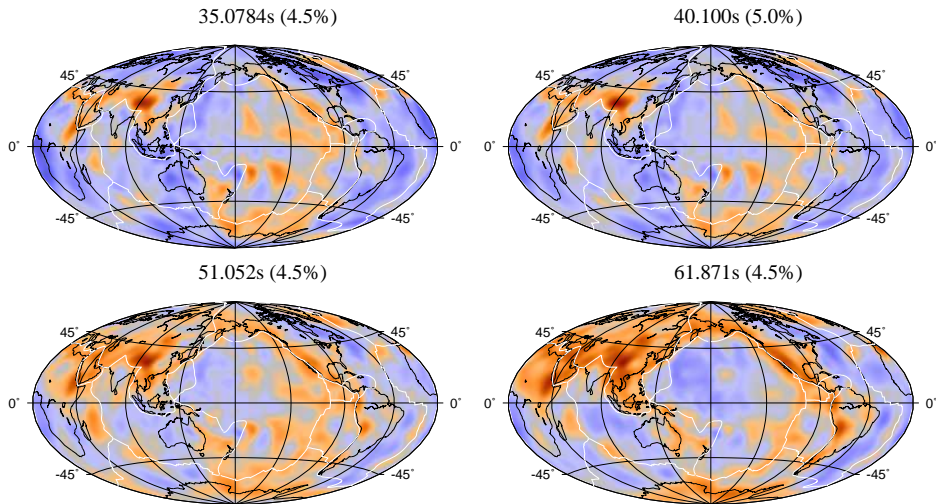
Figure A.6: *Relative isotropic phase velocity maps with respect to PREM for the fourth higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*
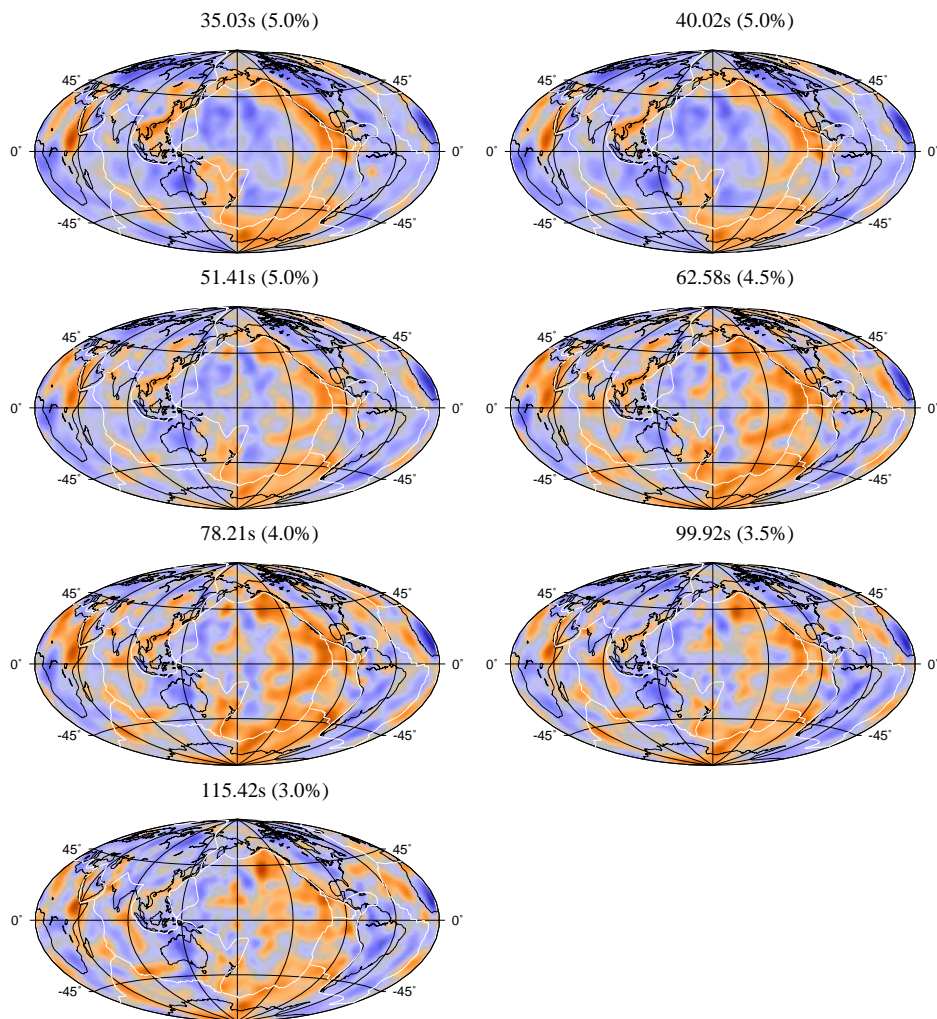
35.115s (5.5%)   40.019s (5.5%)

50.921s (4.0%)   56.074s (4.0%)
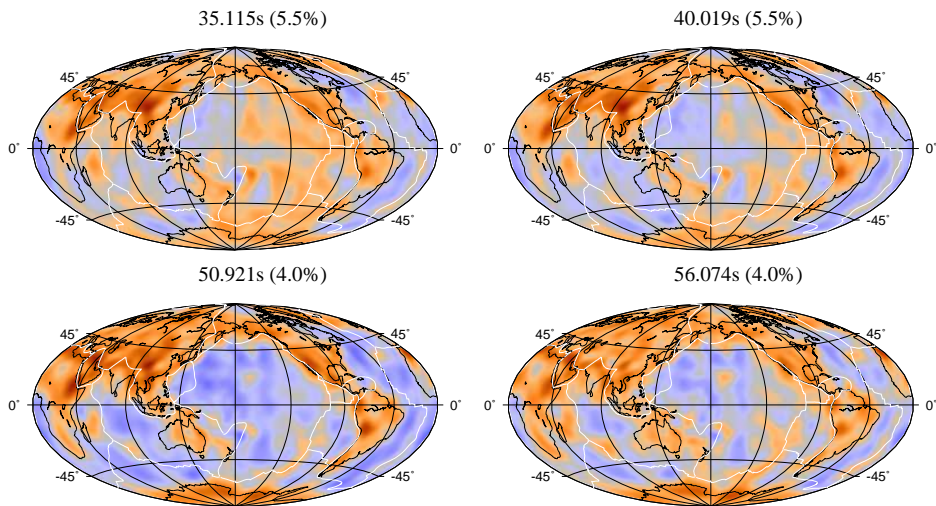
Figure A.7: *Relative isotropic phase velocity maps with respect to PREM for the fifth higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*
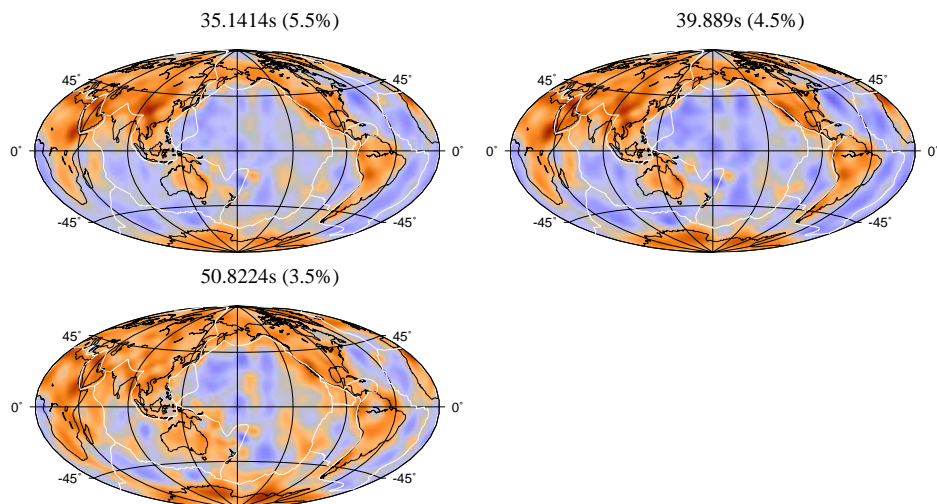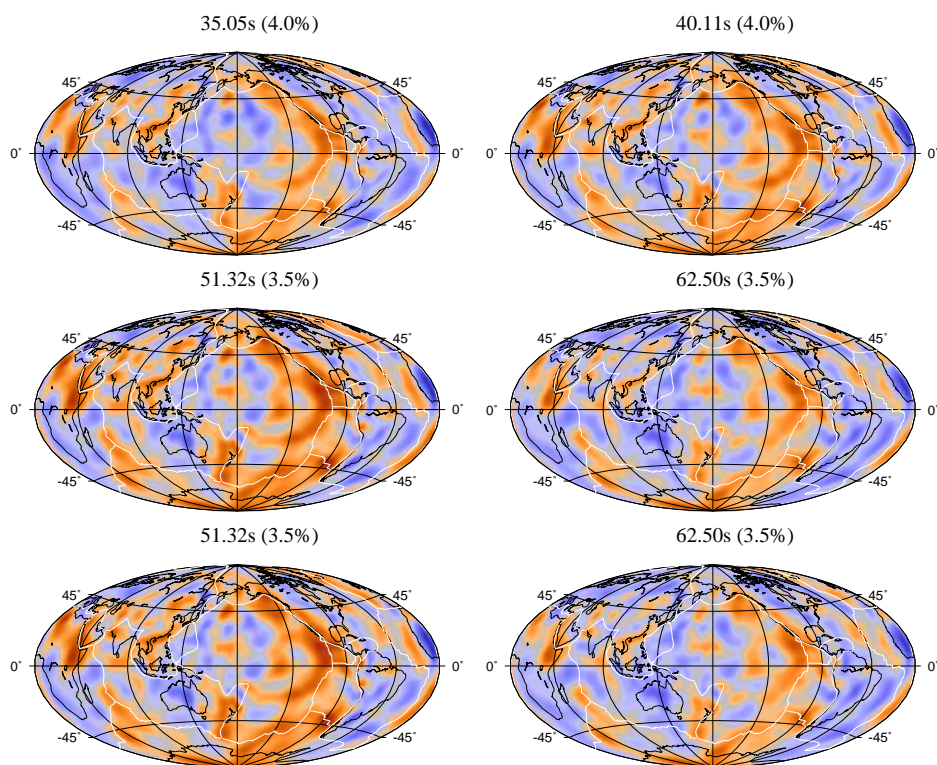
Figure A.8: *Relative isotropic phase velocity maps with respect to PREM for the sixth higher mode Rayleigh at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.062s (10.0%)   40.32s (9.0%)

51.01s (8.0%)   62.36s (7.0%)

78.32s (6.0%)   100.81s (6.0%)

114.78s (5.5%)   129.46s (5.5%)

153.46s (5.0%)   173.98s (5.0%)

Figure A.9: *Relative isotropic phase velocity maps with respect to PREM for the fundamental mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.09s (6.0%) 40.16s (5.5%)

51.19s (4.5%) 62.48s (4.5%)

78.21s (4.0%) 100.08s (4.0%)

114.30s (4.0%) 131.57s (4.0%)

153.07s (4.0%) 176.74s (3.5%)

Figure A.10: *Relative isotropic phase velocity maps with respect to PREM for the first higher mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.03s (5.0%)   40.02s (5.0%)

51.41s (5.0%)   62.58s (4.5%)

78.21s (4.0%)   99.92s (3.5%)

115.42s (3.0%)

Figure A.11: *Relative isotropic phase velocity maps with respect to PREM for the second higher mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.05s (4.0%)　　　　　　　　40.11s (4.0%)

51.32s (3.5%)　　　　　　　　62.50s (3.5%)

51.32s (3.5%)　　　　　　　　62.50s (3.5%)

Figure A.12: *Relative isotropic phase velocity maps with respect to PREM for the third higher mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.06s (4.0%)   40.01s (4.0%)

51.06s (3.0%)   62.77s (3.5%)
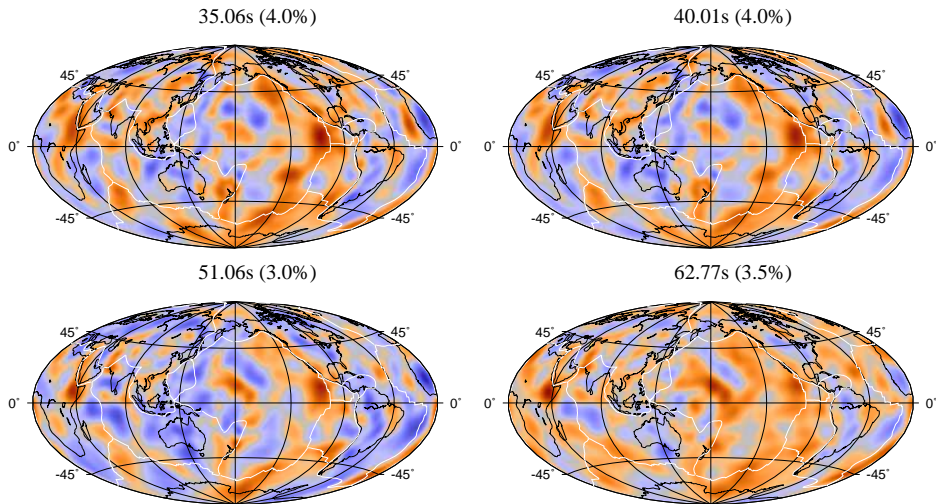
Figure A.13: *Relative isotropic phase velocity maps with respect to PREM for the fourth higher mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

35.12s (4.5%)   40.06s (4.0%)

51.27s (3.0%)   56.29s (3.0%)
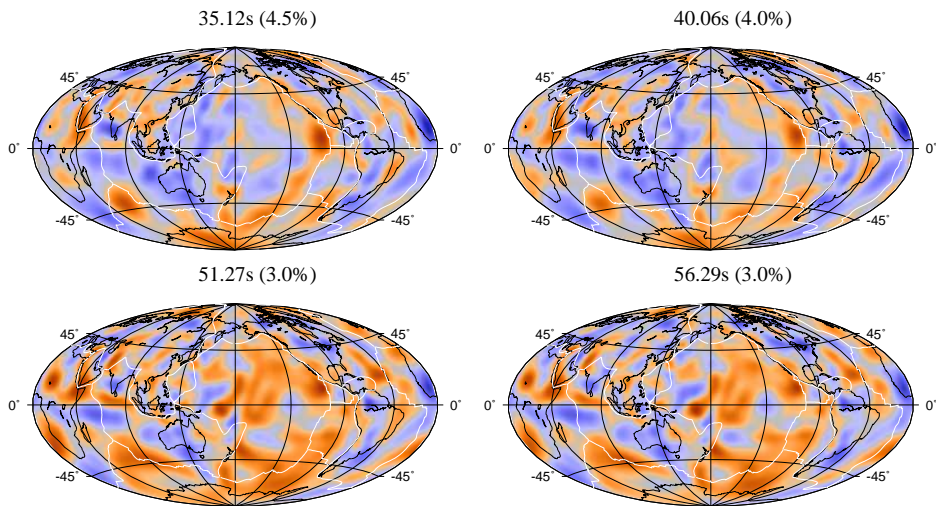
Figure A.14: *Relative isotropic phase velocity maps with respect to PREM for the fifth higher mode Love at the indicated periods. The maximum amplitude of the color scale is indicates in percent.*

Appendix B

Anisotropic phase velocity maps

This appendix shows the azimuthally anisotropic 2ψ and 4ψ phase velocity maps obtained in chapter 4. We show the azimuthally anisotorpic phase velocity models at selected periods for the fundamental up to the sixth higher mode Rayleigh and the fundamental up to the fifth higher mode Love waves. In each figure, the grey scale in the background corresponds to the peak-to-peak amplitude of anisotorpy expressed relative to the average phase velocity calculated from PREM. The black lines correspond to the fast directions which are also scaled to the amplitude shown in the background. The plate boundaries and hotspots are indicated in white.
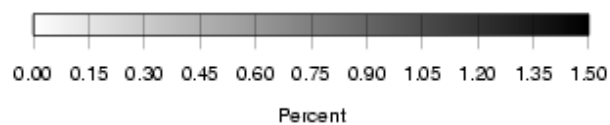
Figure B.1: *Color scale of figures in appendix B*

35.099s   62.305s

100.393s   175.393s
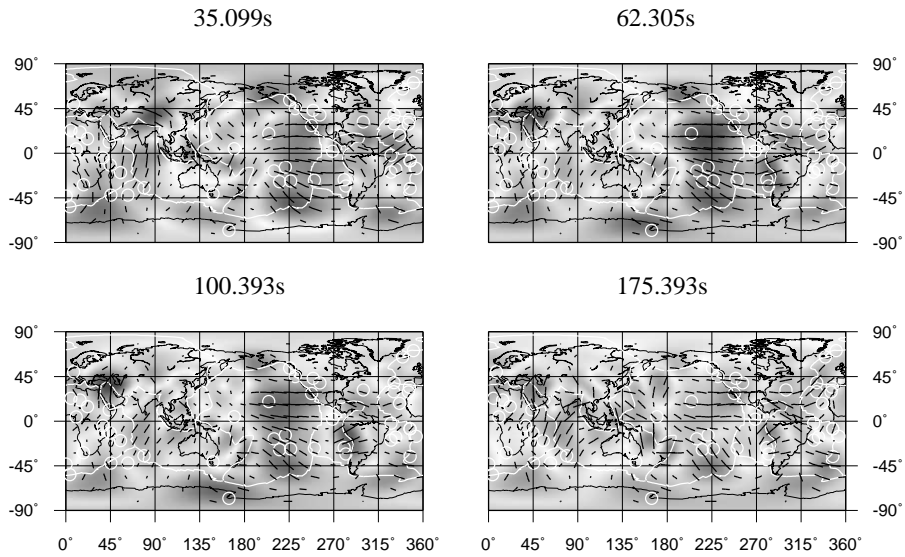
Figure B.2: *Azimuthal anisotropic 2ψ phase velocity maps for the fundamental mode Rayleigh at the indicated periods.*

35.095s   61.866s

99.650s   172.519s
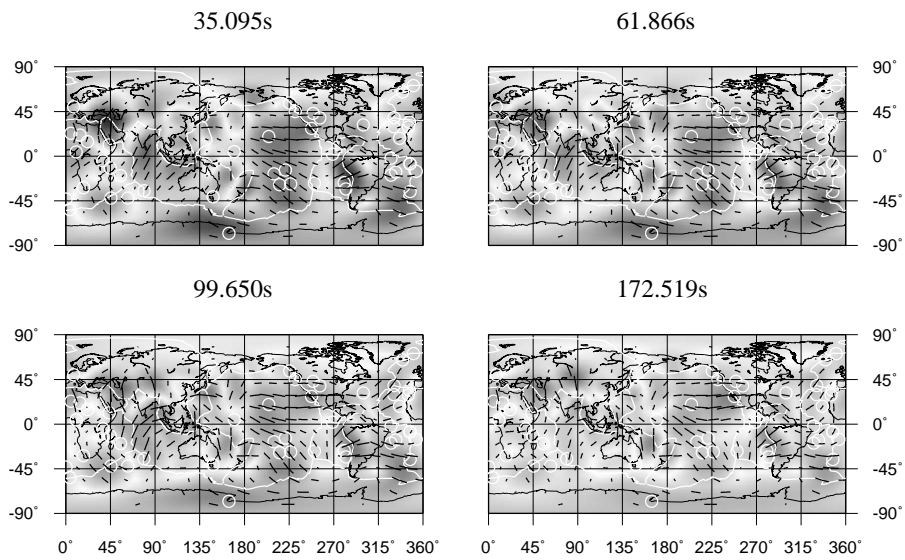
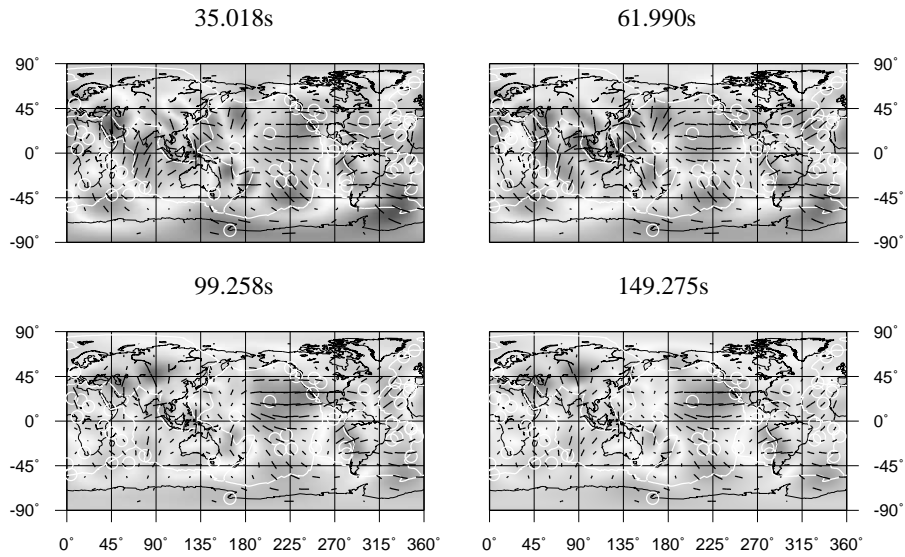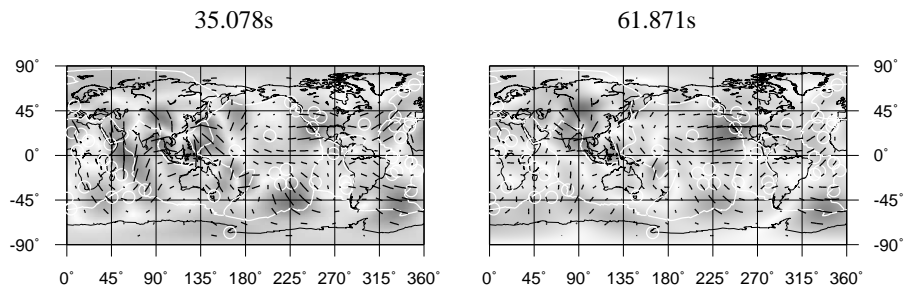Figure B.3: *Azimuthal anisotropic 2ψ phase velocity maps for the first higher mode Rayleigh at the indicated periods.*

35.018s   61.990s

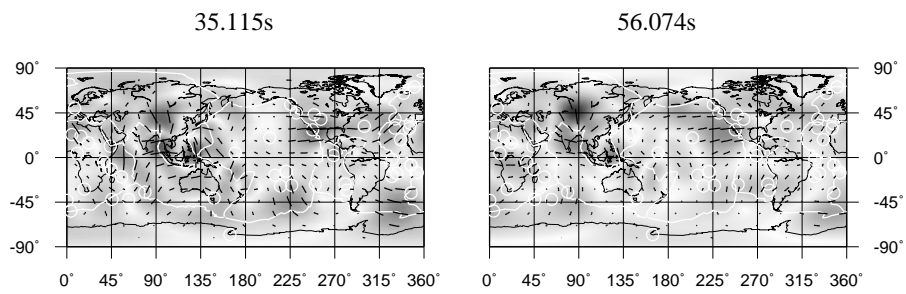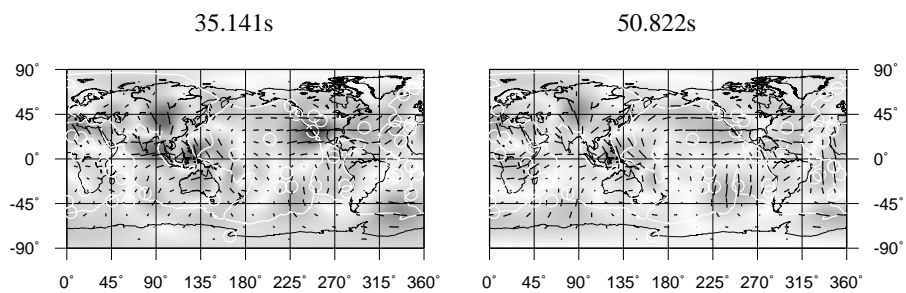99.258s   149.275s

Figure B.4: *Azimuthal anisotropic 2ψ phase velocity maps for the second higher mode Rayleigh at the indicated periods.*

35.014s   87.727s
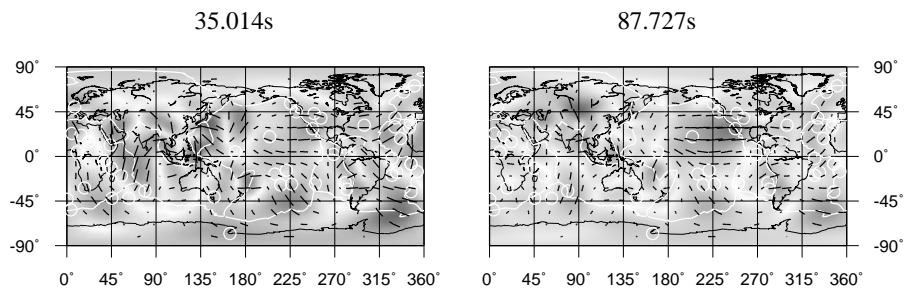
Figure B.5: *Azimuthal anisotropic 2ψ phase velocity maps for the third higher mode Rayleigh at the indicated periods.*

35.078s   61.871s

Figure B.6: *Azimuthal anisotropic 2ψ phase velocity maps for the fourth higher mode Rayleigh at the indicated periods.*
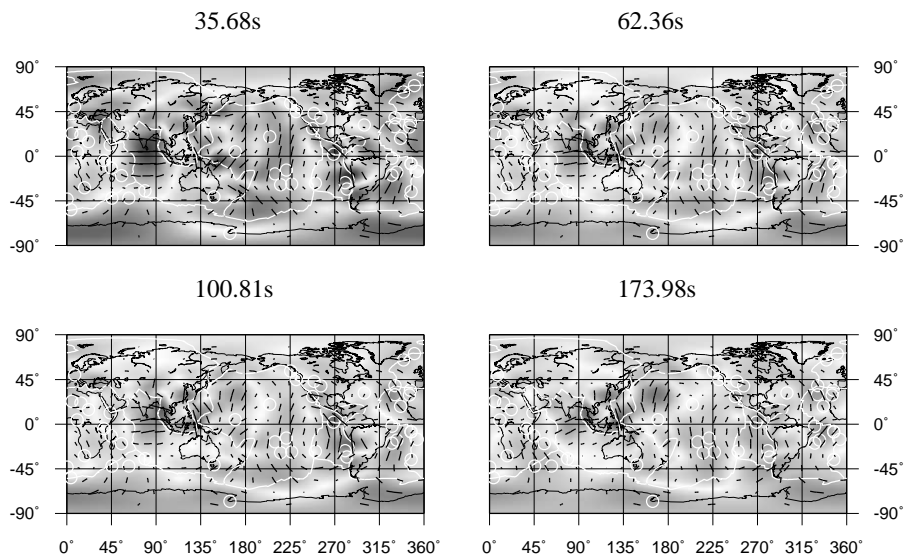
35.115s   56.074s

Figure B.7: *Azimuthal anisotropic 2ψ phase velocity maps for the fifth higher mode Rayleigh at the indicated periods.*

35.141s   50.822s

Figure B.8: *Azimuthal anisotropic 2ψ phase velocity maps for the sixth higher mode Rayleigh at the indicated periods.*

35.68s

62.36s

100.81s

173.98s

Figure B.9: *Azimuthal anisotropic 2ψ phase velocity maps for the fundamental mode Love at the indicated periods.*

35.09s

62.48s

100.08s

176.74s
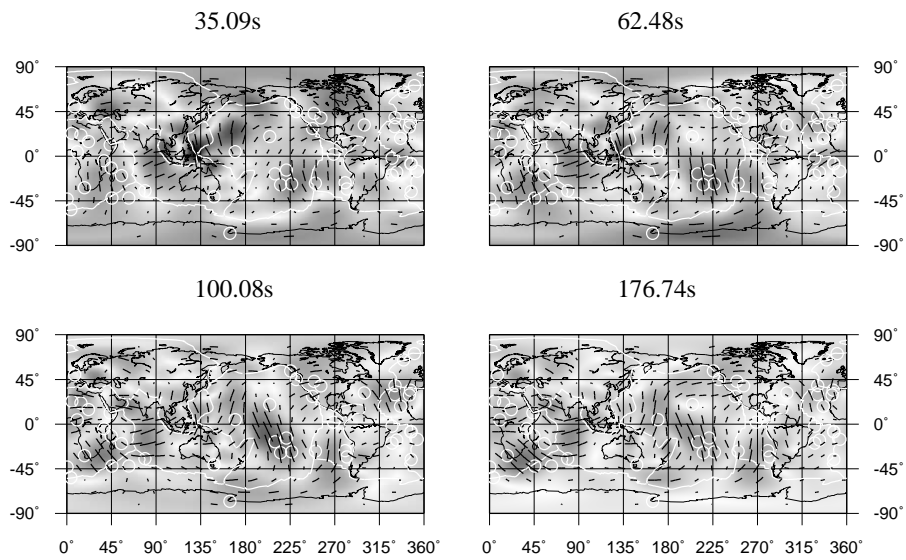
Figure B.10: *Azimuthal anisotropic 2ψ phase velocity maps for the first higher mode Love at the indicated periods.*

35.03s   51.41s
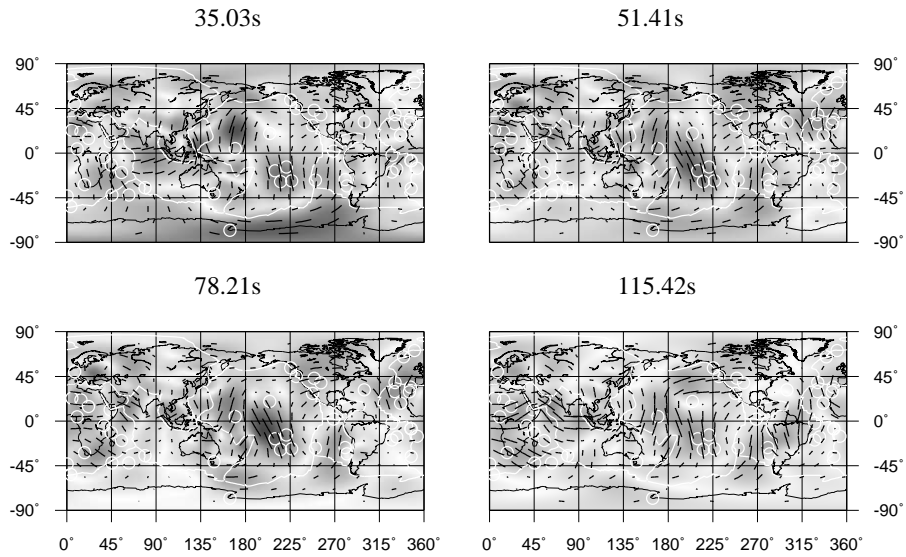

78.21s   115.42s

Figure B.11: *Azimuthal anisotropic 2ψ phase velocity maps for the second higher mode Love at the indicated periods.*
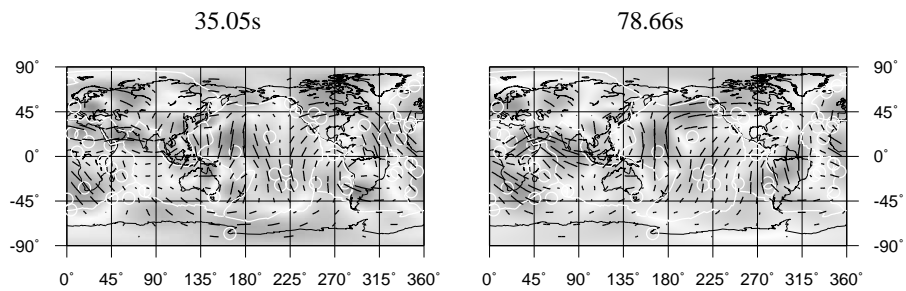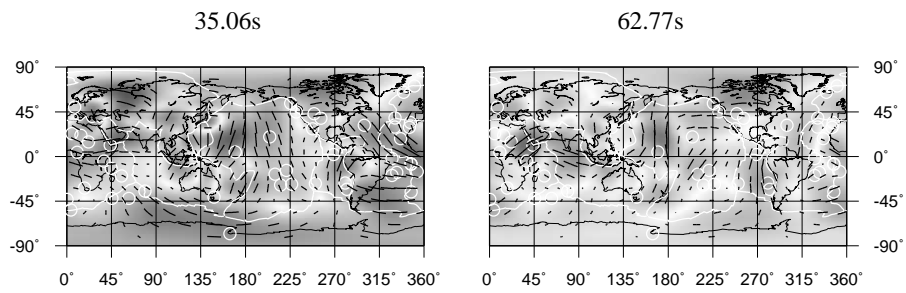
35.05s   78.66s

Figure B.12: *Azimuthal anisotropic 2ψ phase velocity maps for the third higher mode Love at the indicated periods.*

35.06s 62.77s

Figure B.13: *Azimuthal anisotropic 2ψ phase velocity maps for the fourth higher mode Love at the indicated periods.*
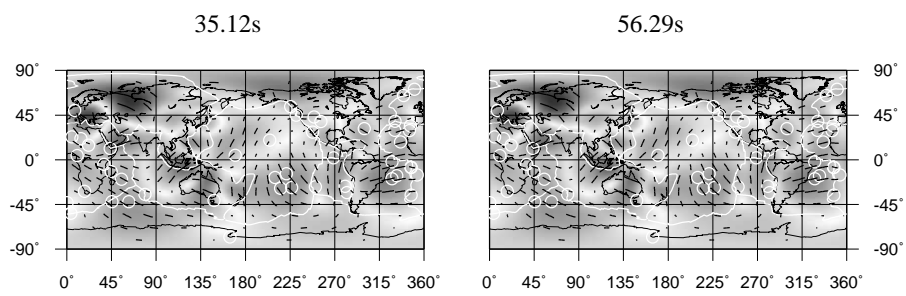
35.12s 56.29s

Figure B.14: *Azimuthal anisotropic 2ψ phase velocity maps for the fifth higher mode Love at the indicated periods.*

35.099s   62.305s

100.393s   175.393s

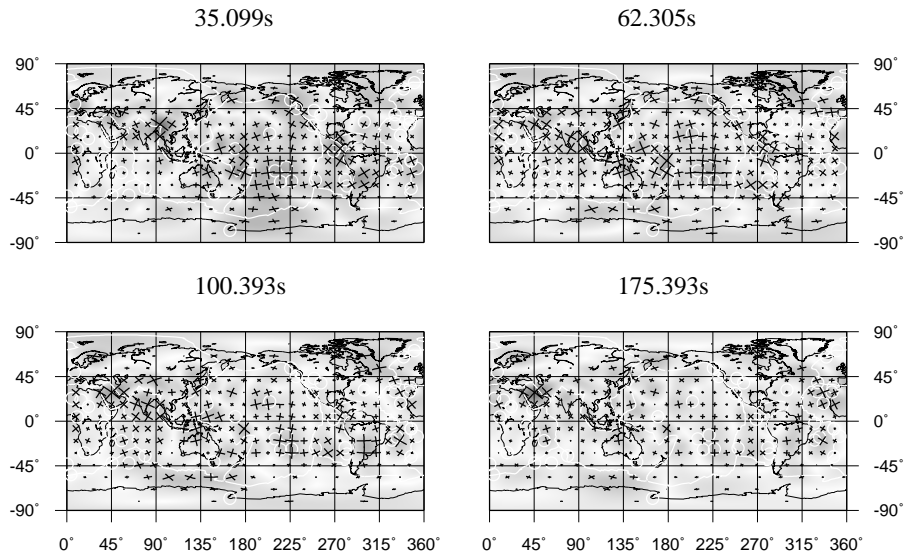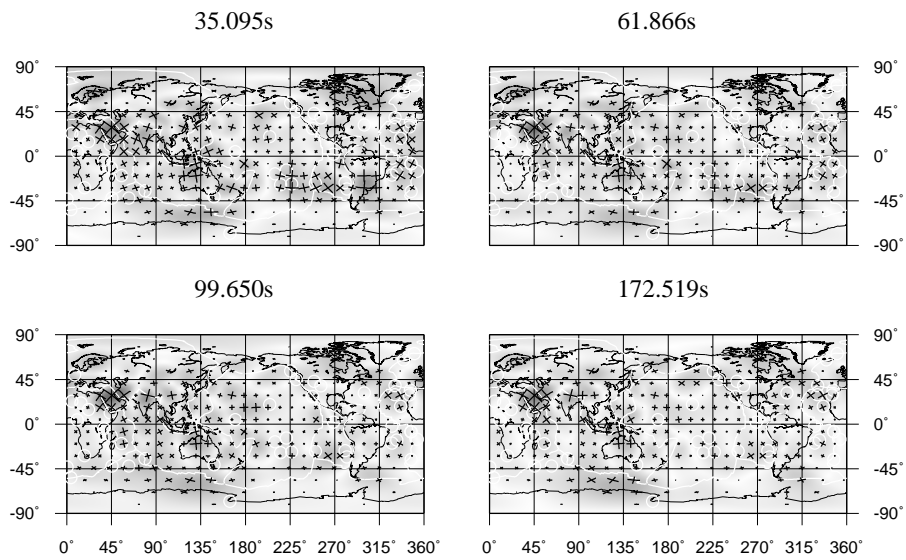Figure B.15: *Azimuthal anisotropic 4ψ phase velocity maps for the fundamental mode Rayleigh at the indicated periods.*

35.095s   61.866s

99.650s   172.519s

Figure B.16: *Azimuthal anisotropic 4ψ phase velocity maps for the first higher mode Rayleigh at the indicated periods.*

35.018s

61.990s

99.258s

149.275s

Figure B.17: *Azimuthal anisotropic 4ψ phase velocity maps for the second higher mode Rayleigh at the indicated periods.*

35.014s

87.727s

Figure B.18: *Azimuthal anisotropic 4ψ phase velocity maps for the third higher mode Rayleigh at the indicated periods.*

35.078s   61.871s

Figure B.19: *Azimuthal anisotropic 4ψ phase velocity maps for the fourth higher mode Rayleigh at the indicated periods.*
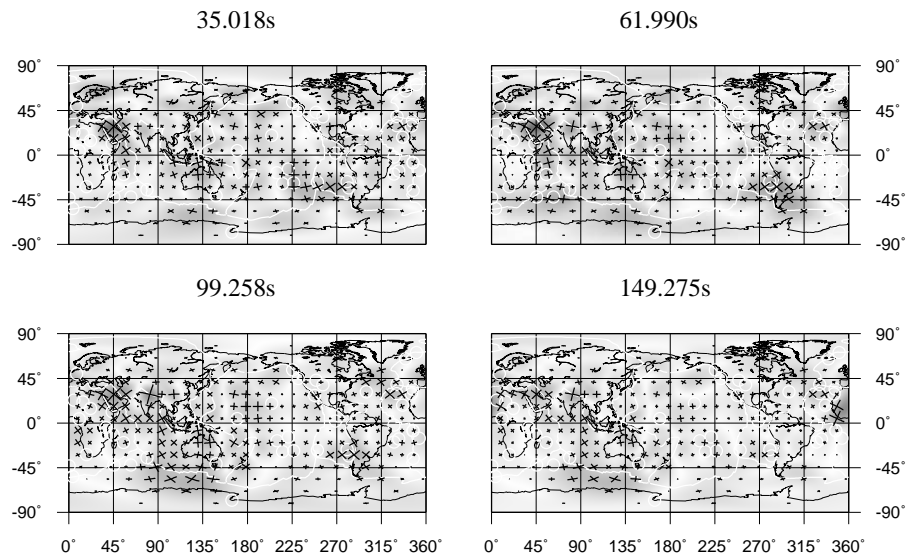
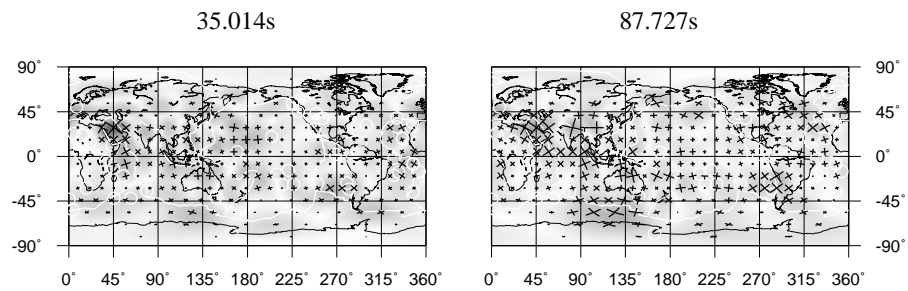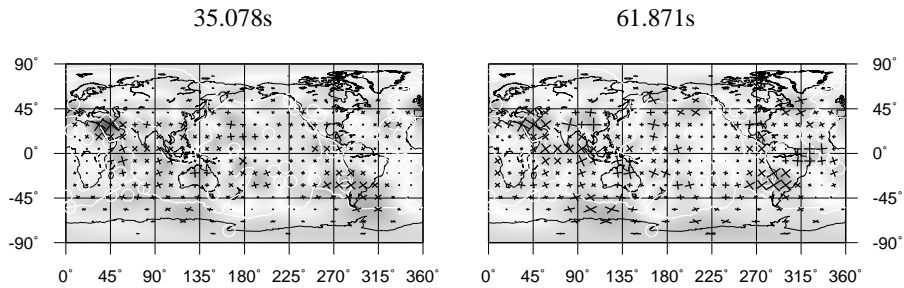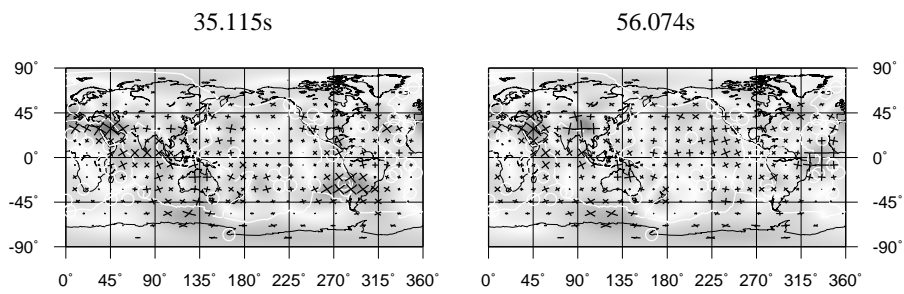35.115s   56.074s

Figure B.20: *Azimuthal anisotropic 4ψ phase velocity maps for the fifth higher mode Rayleigh at the indicated periods.*

35.141s   50.822s

Figure B.21: *Azimuthal anisotropic 4ψ phase velocity maps for the sixth higher mode Rayleigh at the indicated periods.*

35.06s 62.36s

100.81s 173.98s
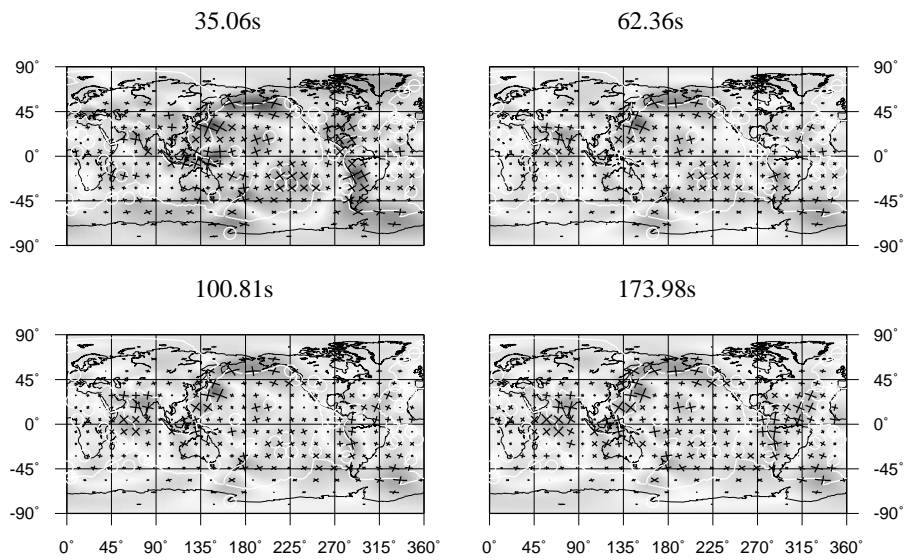
Figure B.22: *Azimuthal anisotropic 4ψ phase velocity maps for the fundamental mode Love at the indicated periods.*

35.09s 62.48s

100.08s 176.74s
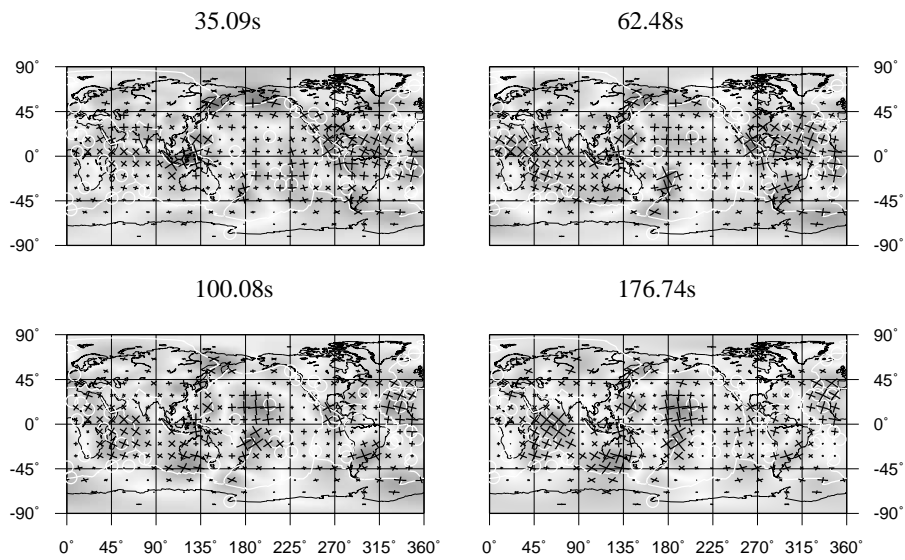
Figure B.23: *Azimuthal anisotropic 4ψ phase velocity maps for the first higher mode Love at the indicated periods.*

35.09s   51.41s

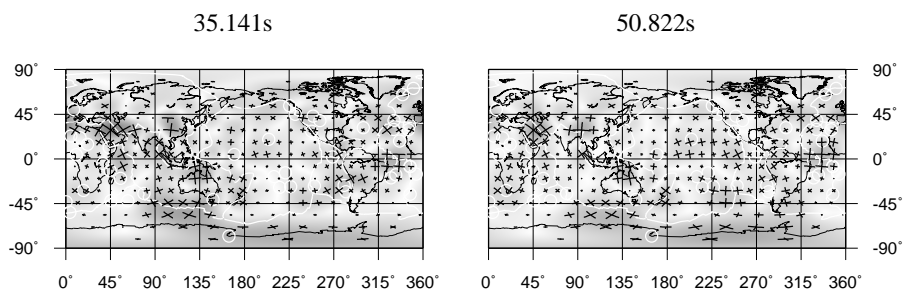78.21s   115.42s

Figure B.24: *Azimuthal anisotropic 4ψ phase velocity maps for the second higher mode Love at the indicated periods.*

35.05s   78.66s

Figure B.25: *Azimuthal anisotropic 4ψ phase velocity maps for the third higher mode Love at the indicated periods.*

35.06s   62.77s

Figure B.26: *Azimuthal anisotropic 4ψ phase velocity maps for the fourth higher mode Love at the indicated periods.*

35.12s   56.29s

Figure B.27: *Azimuthal anisotropic 4ψ phase velocity maps for the fifth higher mode Love at the indicated periods.*
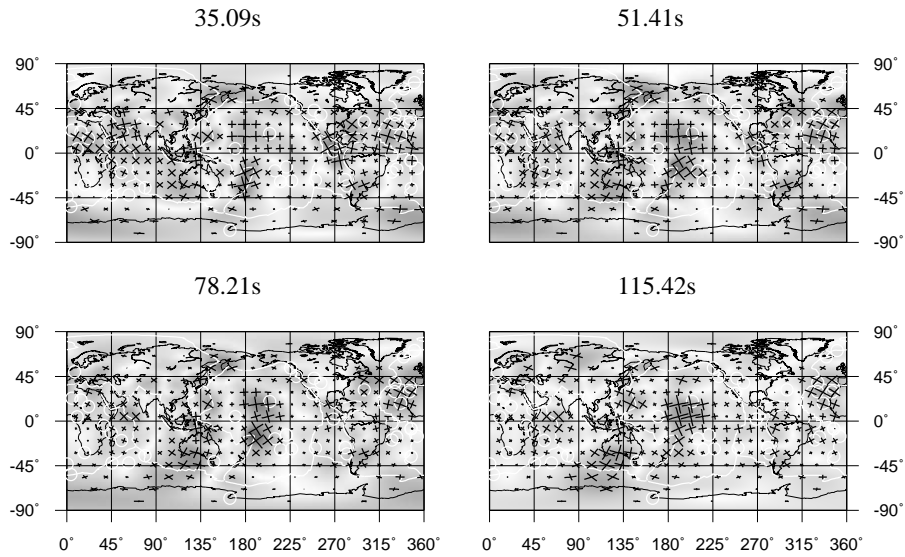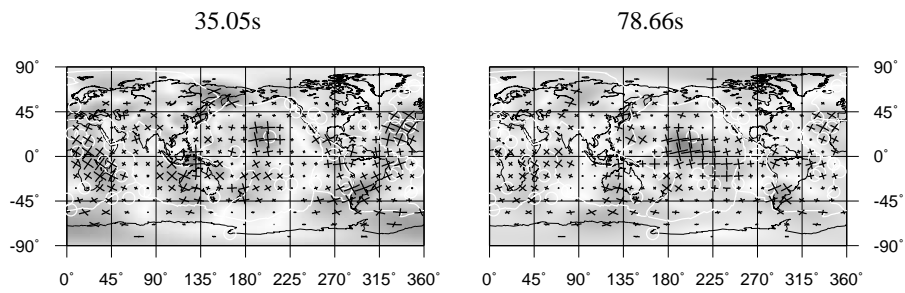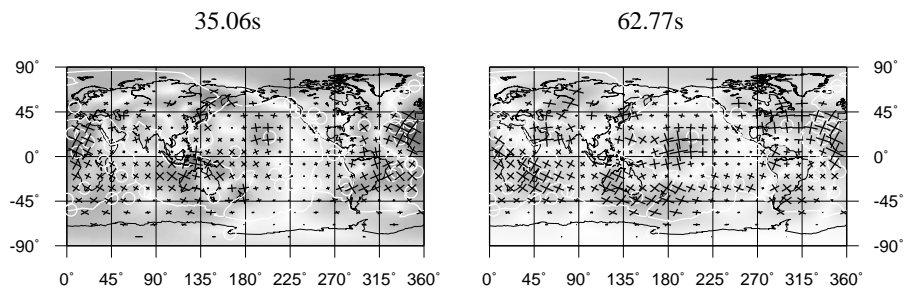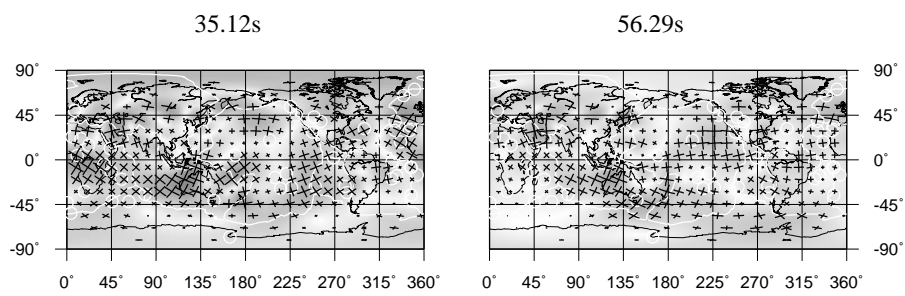
Appendix C

Shear wave velocity maps

This appendix shows a V_{SH}, V_{SV} and an isotropic velocity model obtained in chapter 5. As the velocity models are taken from the mean of the one-dimensional marginals, they represent one realization of the velocity models out of many possibilities. The low amplitudes are caused by the choice of plotting the mean of the one-dimensional marginal. They would be higher when plotting the maximum probability models (drawn from the maximum of the one-dimensional marginals). This relation between the mean and maximum probability models can be seen in figure 5.7. The patterns of higher and lower velocity are very similar to results obtained by Ferreira et al. (2007), the main difference is a lower resolution (spherical harmonic degree 20) and lower amplitudes (mean of the one-dimensional marginals). The velocity models are plotted with respect to the global mean at the indicated depths. On top of each panel we indicate the depth (km), the reference mean velocity (m/s) and the variation with respect to this mean (%).

Figure C.1: *Color scale of figures in appendix C, the maximum amplitude is given in percent above each figure.*

50km, 4560.81m/s (5.0%)   100km, 4465.3m/s (5.0%)

150km, 4417.08m/s (5.0%)   200km, 4391.4m/s (5.0%)

300km, 4626.82m/s (4.0%)   400km, 4895.03m/s (2.0%)

500km, 5168.97m/s (2.0%)   600km, 5457.29m/s (2.0%)

800km, 6243.87m/s (1.0%)   1000km, 6368.21m/s (0.5%)

Figure C.2: *Relative horizontally polarized shear wave velocity maps with respect to the mean as indicated. The maximum amplitude of the color scale is indicates in percent.*

50km, 4441.06m/s (5.0%)   100km, 4366.41m/s (5.0%)

150km, 4366.83m/s (5.0%)   200km, 4356.99m/s (5.0%)

300km, 4656.44m/s (4.0%)   400km, 4905.17m/s (2.0%)

500km, 5191.3m/s (1.5%)   600km, 5481.64m/s (1.5%)

800km, 6246.76m/s (1.0%)   1000km, 6368.02m/s (0.5%)

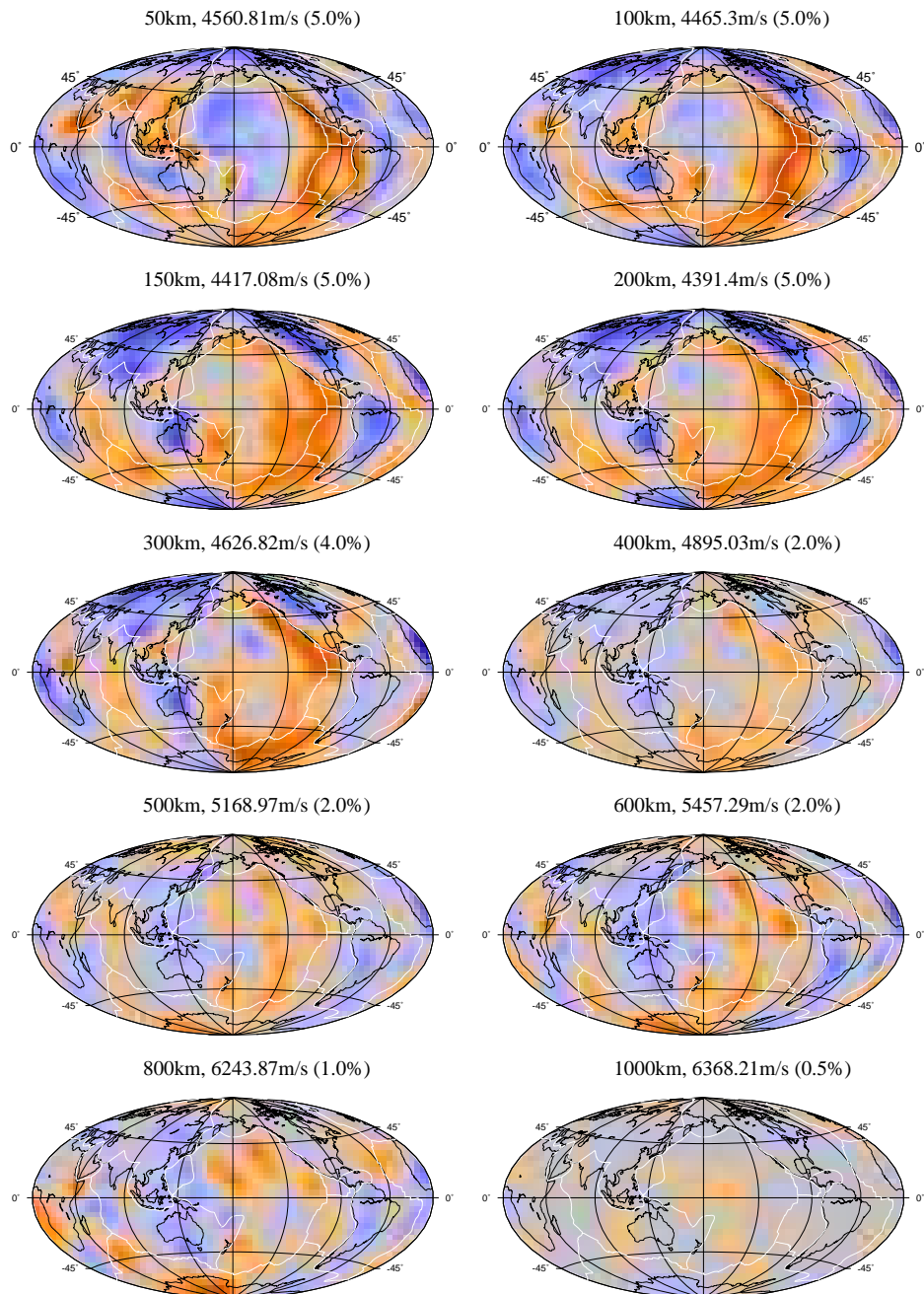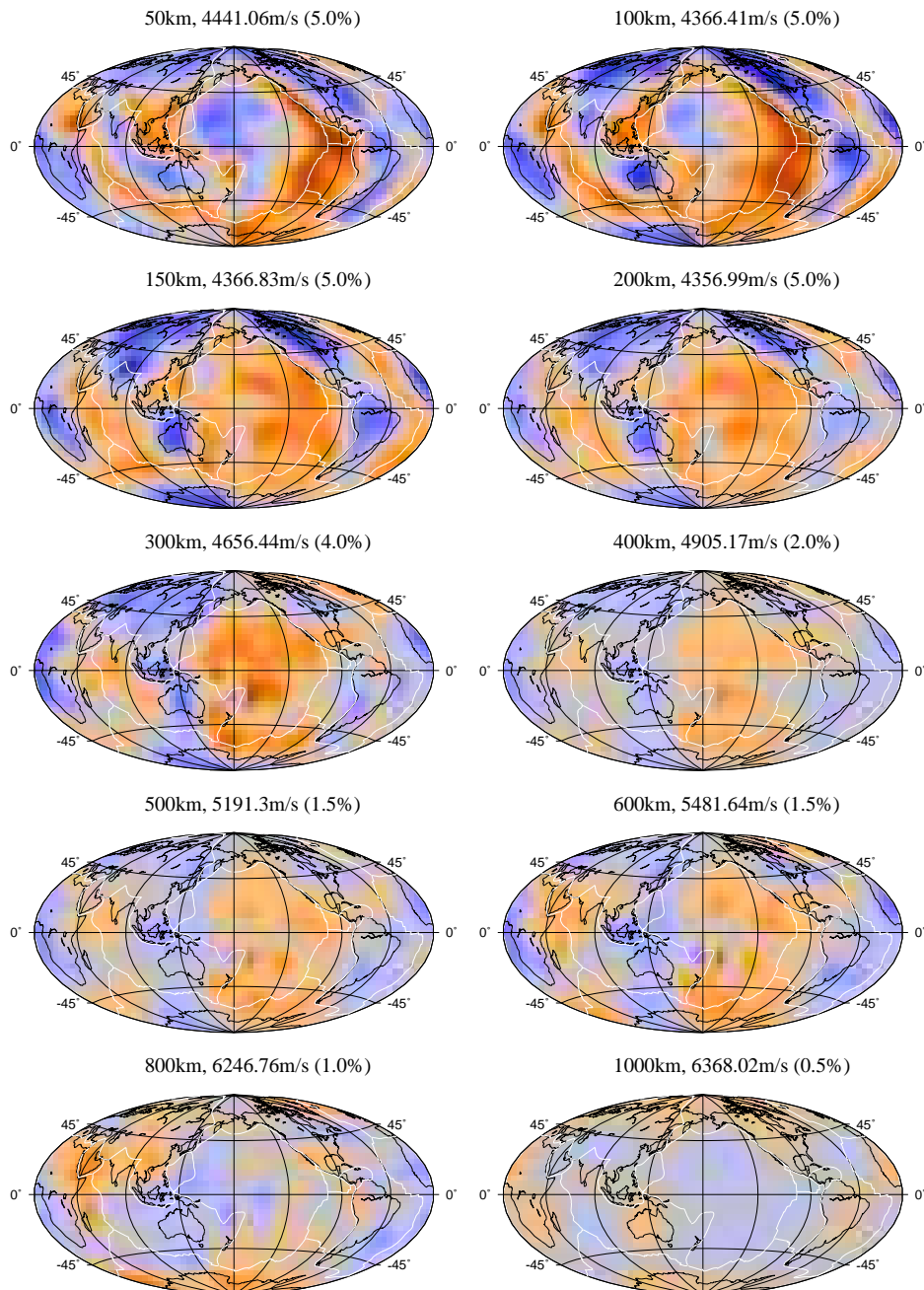Figure C.3: *Relative vertically polarized shear wave velocity maps with respect to the mean as indicated. The maximum amplitude of the color scale is indicates in percent.*
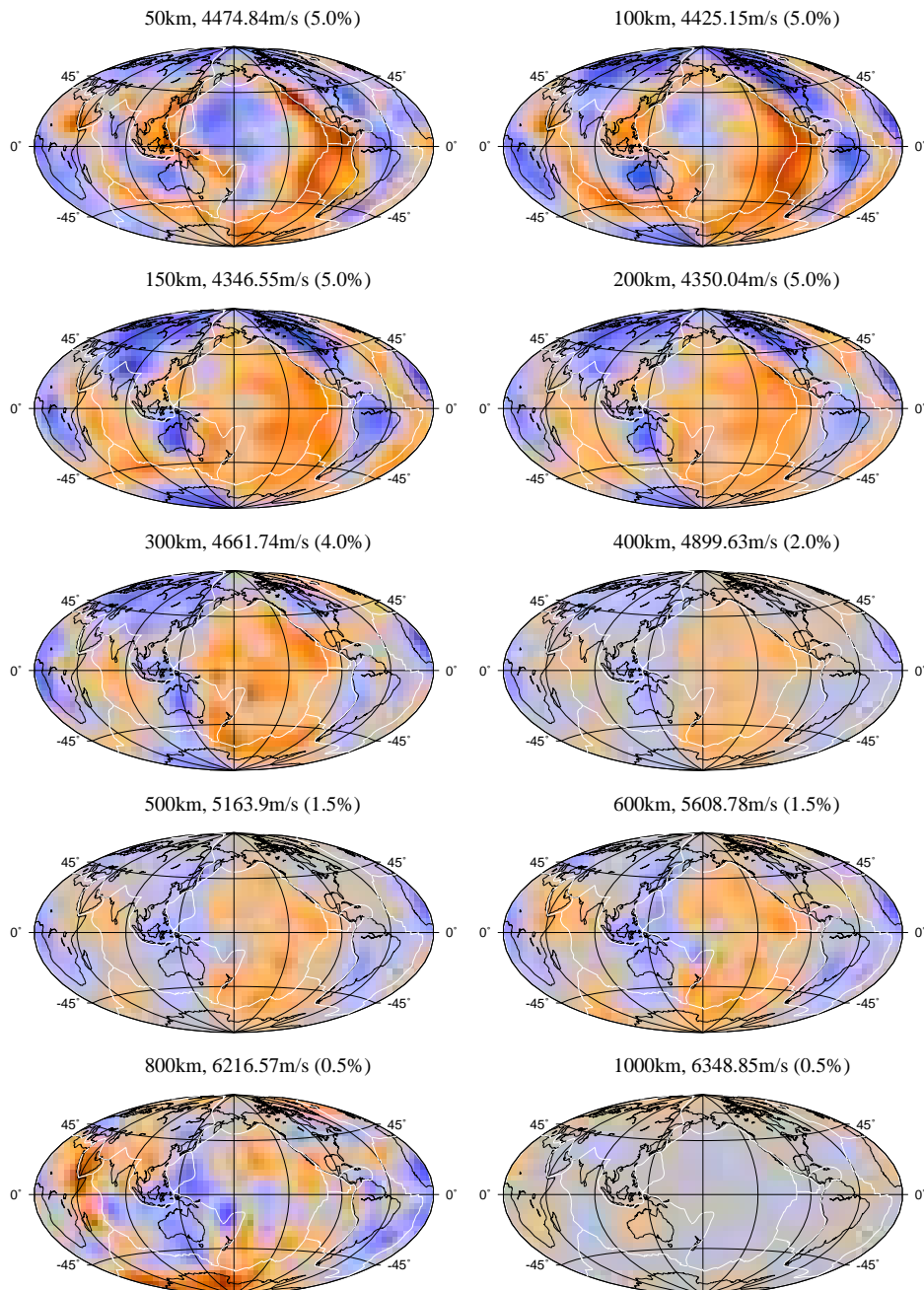
Figure C.4: *Relative isotropic shear wave velocity maps with respect to the mean as indicated. The maximum amplitude of the color scale is indicates in percent.*

Samenvatting en conclusies (Summary and conclusions in Dutch)

Overal op de wereld worden aardbevingen geregistreerd met seismometers. De golven die door de aarde reizen na een aardbeving geven informatie over het gesteente waar ze doorheen hebben gereisd. Dus als wij hier in Utrecht een aardbeving registeren die plaatsvond in Japan, dan geeft het seismogram informatie over het hele pad dat de golven hebben afgelegd van Japan tot in Utrecht. Er zijn verschillende aardbevingsgolven die op verschillende tijden aankomen in het seismogram. Eerst komen de drukgolven aan die door de aarde heen reizen; we noemen deze de P golven. Deze worden gevolgd door transversale golven die ook door de aarde heen reizen, de S golven. Deze golven worden op enige afstand gevolg door de oppervlaktegolven. De oppervlaktegolven reizen langs het aardoppervlak en hebben de grootste amplitude, waardoor ze ook de meeste schade aanrichten na een aardbeving. De oppervlaktegolven die in het horizontale vlak geregistreerd worden op de seismometer worden Love-golven genoemd en de oppervlakte golven die op het vertikale vlak van de seismometer geregistreerd worden, worden Rayleigh-golven genoemd. Oppervlaktegolven zijn dispersief, wat wil zeggen dat elke frequentie op een andere tijd aankomt. De lage frequenties komen eerst aan, gevolgd door de hoge frequenties. De snelheid waarmee elke frequentie reist wordt de fasesnelheid genoemd. De oppervlaktegolven noemen we ook wel de grondtoon. De boventonen van oppervlaktegolven komen in het seismogram aan net voor de grondtoon en reizen dus met een hogere snelheid. De meeste energie van de grondtoon reist door de bovenste 400 km van de aarde, terwijl de energie van de boventonen door diepere structuren heenreist, afhankelijk van de frequentie van de boventoon. Boventonen geven dus meer informatie over diepere structuren.

In dit proefschrift presenteren we de drie stadia inversiebenadering zoals voorgesteld door Kennett en Yoshizawa (200). Het doel van deze inversiebenadering is het verkrijgen van een S-snelheidsmodel met de diepte. Een simpel voorbeeld van een inversie is de som x=2y. Als je y weet, kan je meteen x berekenen, maar als je x weet en je wilt eigenlijk y weten moet je de inverse berekenen, namenlijk y=1/2x. In het eerste stadium verkrijgen we de fasesnelheidsmetingen van

161

de grondtoon en boventonen van Love- en Rayleigh-golven door middel van
een golfvorminversie. Deze metingen worden gebruikt in het tweede stadium om
fasesnelheidskaarten te maken. Daarbij nemen we aan dat de gemeten fasesnel-
heden per seismogram de gemiddelde fasesnelheden aangeven van het pad tussen
de oorsprong (aardbeving) en de ontvanger (seismometer). Het derde stadium is
een inversie van de fasesnelheidskaarten voor een S-snelheidsmodel met de diepte,
ook wel een diepte-inversie genoemd. In een normale twee stadia inversiebenader-
ing geeft het eerste stadium de meting van S-snelheidsprofielen met de diepte per
seismogram en het tweede stadium de inversie voor een S-snelheidsmodel met
de diepte. In deze benadering word aangenomen dat de S-snelheid gemeten per
seismogram de gemiddelde S-snelheid is tussen de oorsprong en de ontvanger.
Deze aanname is echter niet altijd waar. Yoshizawa en Kennett (2002) hebben
laten zien dat verschillende S-snelheidsprofielen in het eerste stadium dezelfde
fasesnelheden hebben. De gemiddelde pad aanname in de drie stadia inversie is
daarom veel beter dan in een normale twee stadia inversiebenadering.

In hoofdtuk twee presenteren we het eerste stadium van de inversie: het meten
van fasesnelheden voor de grondtoon en boventonen van Love en Rayleigh op-
pervlaktegolven. Het is heel moeilijk om de fasesnelheid van de boventonen
te meten, omdat ze een veel kleinere amplitude hebben in vergelijking met de
grondtoon en vaak (bijna) tegelijkertijd met de grondtoon aankomen. De sig-
nalen van de grondtoon en boventonen kun je dus vaak niet van elkaar schei-
den. Om de fasesnelheden te meten gebruiken we een golfvorminversie. In dit
geval weten we de golfvorm maar we willen de fasesnelheid weten. De manier
waarop wij dit oplossen is door middel van een Monte Carlo zoektechniek. In
een Monte Carlo zoektechniek doorzoek je de modelruimte van je probleem. Elk
punt in onze model ruimte correspondeerd met een golfvorm die we vergelijken
met onze geobserveerde golfvorm, door middel van een misfit criterium. Als
we de golfvorm vinden die (bijna) gelijk is aan onze geobserveerde golfvorm
(minimale misfit tussen de geobserveerde en berekende golfvorm) dan hebben we
ook de fasesnelheden ontdekt van de geobserveerde golfvorm. Het mooie van
het gebruik van een modelruimte-onderzoek is dat je niet alleen weer waar je
beste model in de model ruimte zit, maar je weet ook alle andere modellen in de
model ruimte met hun misfit. Nu kunnen we (Bayesiaanse) statistiek toepassen
op de modelruimte en hieruit kunnen we de waarschijnlijkheid van waardes van
elke dimensie van de modelruimte berekenen. Dus voor elke van onze fasesnel-
heden krijgen we een waarschijnlijkheidscurve van de waardes van de fasesnel-
heid. Deze waarschijnlijksheidscurves zijn vrij Gaussisch van vorm en daarom
kunnen we de fasesnelheden weergeven met een gemiddelde waarde (die gelijk is

aan de maximale waarschijnlijkheid) en een standaarddeviatie (een variatie). Dus een typische gemeten fasesnelheid is bijvoorbeeld 4 km/s±40 m/s. Het berekenen van de waarschijnlijkheidscurve is nog nooit gedaan voor fasesnelheden en zo krijgen we dus niet alleen de fasesnelheden maar ook de onzekerheden die erbij horen. Omdat we de hele golfvorm gebruiken krijgen we niet alleen de fasesnelheden voor de grondtoon maar ook die van de boventonen. Hoofstuk twee legt uit hoe deze volledig geautomatiseerde methode werkt en laat alle testen zien (de validatie van de methode).

Omdat de fasesnelheden voor de hogere ordes oppervlaktegolven voor Love-golven nog veel moeilijker te meten zijn dan die voor Rayleigh, gaat hoofdstuk drie speciaal over de gemeten fasesnelheden van de Love-boventonen. De fase-snelheden worden vergeleken met vorige studies en we concluderen dat 65% van onze fasesnelheden binnen één standaarddeviatie vallen in vergelijking met de metingen van Van Heijst en Woodhouse (1999). De gemeten fasesnelheden per seismogram bevatten informatie over het pad dat de golven hebben afgelegd van de bron (de aardbeving) naar de ontvanger (seismometer). Als je de fasesnelheden meet van een grote hoeveelheid seismogrammen, kun je deze combineren tot fase-snelheidskaarten van de aarde. Deze fasesnelheidskaarten hebben we vergeleken met een eerder gepubliceerd S-snelheidsmodel (S20RTS) en we vinden een goede overeenkomst.

In een anisotroop medium hangt de snelheid van seismische golven die reizen door dit medium af van de richting waarin de seismische golven reizen. Een sim-pel geval is een horizontaal gelaagd medium. De snelheid van seismische golven die hier verticaal doorheenreizen (bijvoorbeeld Rayleigh-golven) zal kleiner zijn dan de snelheid van seismische golven die hier horizontaal (bijvoorbeeld Love-golven) doorheen reizen. Dit wordt ook wel radiale anisotropie genoemd. De anisotropie is onafhankelijk van de bewegingsrichting in het horizontale vlak. Als we dit gelaagd medium een beetje schuin zetten, zullen de seismische golven die horizontaal door het medium heenreizen nog steeds sneller zijn, maar de snelheid zal nu ook afhangen van de richting van de horizontale seismische golven door het medium. Dit wordt ook wel azimutale (hoek) anisotropie genoemd. De snelheid van de seismische golven hangt nu ook af van het azimut (de hoek) van voortplan-ting door het medium. De oorzaak van anisotropie in de aarde is waarschijnlijk de oriëntatie van mineralen in de aarde. Als mineralen onder spanning staan, gaan ze zich zo richten dat de snelle richtingen oplijnen in de richting van de span-ning. Een snelle vertikale snelheid wordt zo waarschijnlijk veroorzaakt door een vertikale stroming, en een snelle horizontale snelheid wordt waarschijnlijk veroor-zaakt door een horizontale stroming. In het geval van azimutale anisotropie hangt

de (hoge) horizontale snelheid af van de richting van de horizontale stroming. In de plaattektoniek worden aardplaten gecreëerd bij mid-oceanische ruggen en zinken ze in de aardmantel bij de subductiezones. De stroming in de ondiepe aardmantel is dus van de mid-oceanische rug naar de subductiezone. Een horizontale snelheid met deze stroming mee zal hoger zijn dan een loodrecht erop vanwege de oriëntatie van de mineralen in de stroming.

In hoofdstuk vier presenteren we het tweede stadium van de inversie: het maken van de isotrope en azimutale anisotrope fasesnelheidskaarten van de aarde door middel van een inversie van de gemeten fasesnelheden van hoofdstuk twee. Voor de inversie bepalen we de onderlinge verhoudingen voor de isotrope en azimutale anisotrope termen. We vinden dat de fasesnelheden van de grondtoon en boventonen alleen verklaard kunnen worden met azimutale anisotropie. Verder hebben we veel aandacht geschonken aan het bepalen van de modelonzekerheden, omdat we deze nodig hebben in het volgende hoofdstuk voor de diepte-inversie. De isotrope boventoon fasesnelheidskaarten en de azimuthale anisotrope fasesnelheidskaarten van de grondtoon komen overeen met die van andere studies. Als laatste kijken we naar de eventuele lekkage van de isotrope informatie in de azimutale anisotrope kaarten en andersom en we vinden dat dit effect verwaarloosbaar is.

In hoofdstuk vijf presenteren we het derde stadium van de inversie: het inverteren van de fasesnelheidskaarten van hoofdstuk vier om een S-snelheidsmodel te verkrijgen met de diepte. Voor 492 locaties op de aarde inverteren we de fasesnelheidskaarten waar de azimuthale anisotropie eruit gehaald is voor een radiaal anisotroop diepteprofiel met een volledig niet-lineaire inversie. De Rayleigh-fasesnelheidskaarten worden geïnverteerd voor een SV profiel (een vertikaal gepolarizeerd S-snelheidsprofiel) en de Love-fasesnelheidskaarten worden geïnverteerd voor een SH profiel (horizontaal gepolarizeerd S-snelheidsprofiel). De SV en SH profielen worden gecombineerd in een isotroop en anisotroop profiel. We gebruiken weer een modelruimte-onderzoek, wat ons niet alleen het beste model oplevert, maar ook andere modellen in de modelruimte met hun misfit. De waarschijnlijkheidscurves zijn dit keer echter niet Gaussisch, waardoor we de profielen niet kunnen presenteren als een gemiddelde en een standaarddeviatie. Daarom hebben we besloten om de waarschijnlijkheid van radiale anisotropie te berekenen voor verschillende tektonische gebieden op aarde (oceaan, continent, tektonische gebieden, cratons, platforms), evenals de waarschijnlijkheid dat de amplitude van de radiale anisotropie boven de 1% of 2% uitkomt. De lithosfeer (0-100 km) wordt gedomineerd door snelle horizontaal gepolarizeerde S-anisotropie (horizontale stroming), met een significante amplitude (>2%), behalve onder cratons. Deze

anisotropie werd waarschijnlijk veroorzaakt door de vorming van de lithosfeer en is nu 'bevroren'. De asthenosfeer (100-220 km) wordt gedomineerd door snelle vertikaal gepolarizeerde S-anisotropie (vertikale stroming). Ook vinden we een waarschijnlijkheid van meer dan 70% dat de amplitude van de anisotropie boven de 2% ligt op 300 km. De snelle vertikaal gepolarizeerde S-anisotropie wordt geassocieerd met de mid-oceanische ruggen, waar materiaal uit de aardmantel naar boven komt, en subductie zones, waar materiaal in de aardmantel zinkt. De transitiezone (400-670 km) wordt ook gedomineerd door snelle vertikaal gepolarizeerde S-anisotropie, maar de amplitudes zijn waarschijnlijk kleiner (1%-2%). De ondermantel (670 - 1500 km) lijkt vooral isotroop (of anisotroop met een zeer kleine amplitude). Dit komt overeen met eerdere studies.

We hebben laten zien dat we boventonen kunnen meten voor enkele seismogrammen tot de zesde boventoon met een golfvorminversie. De grote hoeveelheid metingen (>350,000), verkregen met een volledig geautomatiseerde benadering, worden geïnverteerd voor isotrope en azimutale anisotrope fasesnelheidskaarten. Verder hebben we laten zien dat we deze fasesnelheidskaarten kunnen inverteren voor een radiaal anisotroop model gebruik makende van een volledig niet-lineaire benadering, wat alleen mogelijk was met moderne computerkracht. Verder hebben we veel aandacht geschonken aan de onzekerheden, die heel belangrijk bleken in de zoektocht naar significante radiale anisotropie. In de toekomst zou het interessant zijn om een diepte-inversie toe te passen op de azimuthale anisotropie fasesnelheidskaarten om meer informatie te verkrijgen over het dieptebereik van de azimuthale anisotorpie.

Dankwoord (Acknowledgments)

First of all, I would like to thank my promotor and daily supervisor Jeannot Trampert. After my previous supervisor Rob van der Hilst left for M.I.T., United States you kindly agreed to be my PhD supervisor. Even though the number of PhD students almost tripled since I started, you always find time to help all of your PhD students. Thank you for your help and guidance. I would also like to thank my second promotor Brian Kennett. Even though Australia is quite far away, we have always managed to talk in person at least one time per year. Thank you as well for the carefull reading and comments on my sometimes overly long reports and monthly e-mails, even though they were sometimes a bit jumbled. I would like to thank my 'paranimfen' Jojanneke van den Berg and Christine Franke, for being there for me. Christine, thank you for all the dinners (especially the crème brûlée) and the many laughs resulting in me almost choking. Jojanneke, thank you for all the games we played and your friendship. My office mates, first Robbert van Vossen and later Wouter Kimman were quite patient and I enjoyed our discussions and chats. The seismology group in Utrecht is a very diverse and social group and I am very glad I could be part of it. So thank you, Alison Malcolm, Ueli Meier, Ilaria Mosca, Ebru Bozdag, Xiaomei Zhang, Anne Sieminski, Arie van Wettum, Hanneke Paulssen and Sergei Lebedev, I enjoyed our many teatimes and discussions. All my other colleagues are thanked as well Céline Tirel, Thomas Geenen, Joost van Summeren , Roberta de Franco, Bahjat Alhammoud, Pasha Karami Arokhloo, Marzieh Baes, Gabriela Tanasescu, Karin Ruckstuhl, Steven van Benthem, Kabir Roy-Chowdhury, Wim Spakman, Arie van den Berg, Paul Meijer, Rob Govers and Rinus Wortel. Thank you as well to all my 'old' colleagues Stefan Carpentier, Tim van Zon, Caroline Beghein, Frederic Deschamps, Joe Resovsky, Maisha Amaru, Annemarie Muntendam-Bos. Henk van der Meer, bedankt voor het koffiezetten door de jaren heen. Ik was altijd blij als het koffietijd was vooral in de periode dat ik vroeg begon met werken. Joop Hoofd en Theo van Zessen, bedankt voor jullie steun met alle computergerelateerde issues, Linux in het algemeen en het bijhouden van de clusters, zonder jullie was mijn onderzoek ook niet mogelijk geweest. Jacqueline Landsheer, dank je wel voor de administratieve hulp en de vele gesprekken tijdens de koffiepauzes. Verder natuurlijk al mijn vrienden heel erg bedankt voor de mooie tijden, lekkere dineetjes en spelletjes spelen. Dus bedankt, Robbert en Jojanneke, Erik (veel succes met je promotie) en Miriam, Patrick (ook veel succes met je promotie) en Meike, Edgar en Jacqueline en Annemarie en Jan-Willem. Mijn zus Inge die altijd klaar staat

167

voor anderen, bedankt voor het luisteren en voor je steun en vriendschap. Kees, nuchtere broer, bedankt voor het bijwonen van onze LAN-dagen, het uitlenen van computerspelletjes en de gezelligheid tijdens onze etentjes. Maarten, mijn schoonbroer, ook bedankt voor de LAN dagen en het aansteken van onze interesse naar aquariums. Je vrouw Linda en jullie kleintje Max ook bedankt voor de gezellige momenten. Mijn schoonouders Joke en Hans, bedankt voor de steun en de gastvrijheid en de onvergetelijke vakantie samen in Bali. Natuurlijk wil ik ook mijn ouders, Gerda en Kees, bedanken voor hun steun, nuchtere kijk op dingen, shoppen in Utrecht, gezellige dineetjes, sinterklaas surprises, en de vele gezellige dagen. Als laatste wil ik graag mijn man, Peter van Thienen, bedanken voor zijn steun en liefde door de jaren heen, eerst als mijn vriend en collega, later als mijn man en het laatste jaar ook weer als mijn collega. Ik vind het heerlijk bij jou!

Curriculum vitae

May 8, 1979   Born in Biddinghuizen, the Netherlands

1991 - 1997   Secondary school at CCNV in Harderwijk, The Netherlands

1997 - 2002   Study of Geophysics at Utrecht University in Utrecht, The Netherlands

2001 - 2002   Masters research project at Leeds University, U. K.,

2002 - 2003   PhD student at Utrecht University in Utrecht, The Netherlands,
              supervisor: Prof. dr. R. D. van der Hilst

2003 - 2008   PhD student at Utrecht University in Utrecht, The Netherlands,
              supervisor: Prof. dr. J. Trampert