GEOLOGICA ULTRAIECTINA

Mededeling van de
Faculteit Geowetenschappen
Universiteit Utrecht

No. 308

Assessing and improving seismic tomography models using 3-D numerical wave simulations

Ebru Bozdağ

The research described in this thesis was carried out at:
Faculty of Geosciences
Utrecht University
Budapestlaan 4
3584 CD Utrecht
The Netherlands

http://www.geo.uu.nl/Research/Seismology/

Printed by Wöhrmann Print Service, Zutphen

ISBN: 978-90-5744-169-1

Assessing and improving seismic tomography models using 3-D numerical wave simulations

Bepaling en verbetering van seismische tomografische modellen door 3-D numerieke golfsimulaties

(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 MAANDAG 18 MEI 2009 DES MIDDAGS TE 4:15 UUR

DOOR

Hatice Ebru Bozdağ

Geboren op 9 februari 1979, te Artvin, Turkije

Promotor: Prof. dr. J. Trampert

This thesis was partly accomplished with financial support from European Commission's Human Resources and Mobility Program Marie Curie Research and Training Network SPICE (Seismic wave Propagation and Imaging in Complex media: a European network) Contract No. MRTNCT-2003-504267 and from the Dutch National Science Foundation under grant number NWO:VICI-865.03.007.

The scientist does not study Nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful.
If Nature were not beautiful, it would not be worth knowing, and if Nature were not worth knowing, life would not be worth living.

Henri Poincaré

to my brother, Ünal Bozdağ

Members of the dissertation committee:

Prof. dr. H. Igel
Department of Earth Sciences
Ludwig-Maximilians-University, Germany

Prof. dr. J.-P. Montagner
Equipe de Sismologie
Institut de Physique du Globe de Paris, France

Prof. dr. W. Spakman
Faculty of Geosciences
Utrecht University, the Netherlands

Prof. dr. J. Tromp
Department of Geosciences
Princeton University, USA

Prof. dr. J. H. Woodhouse
Department of Earth Sciences
University of Oxford, UK

Contents

1 Introduction **1**

1.1 History of Seismic Tomography 1

1.1.1 The Earth's Mantle from Classical Seismic Tomography 2

1.1.2 Finite-Frequency Effects in Seismic Tomography . . 4

1.1.3 Numerical Methods in Seismic Tomography 5

1.2 Shortcomings in Global Mantle Tomography 6

1.3 Motivation and Outline 8

2 Spectral element method in 3-D wave simulation and seismic tomography **11**

2.1 Computation of synthetic seismograms 11

2.1.1 Seismological background 13

2.1.2 Spectral element method 15

2.2 Using 3-D simulations in seismic tomography 16

2.2.1 Adjoint methods 18

2.3 About SPECFEM3D 20

2.3.1 Meshing the globe 20

2.3.2 Velocity models 20

2.3.3 Model parameters 22

2.3.4 Source time function 23

2.3.5 Comparison of SEM and normal mode seismograms 23

2.3.6 Adjoint simulations 26

2.3.7 Technical Information 29

3 On crustal corrections in surface wave tomography **31**

3.1 Introduction 32

3.2 Method 34

3.2.1 Calculation of synthetic seismograms 34

3.2.2 Calculation of exact local phase velocity perturbations in Crust2.0 . . . 35
3.2.3 Calculation of phase shifts . . . 36
3.2.4 Measuring the time-shifts as a function of frequency 38
3.3 Results . . . 39
3.3.1 Some examples of time-shifts . . . 40
3.3.2 Time-shifts as a function of distance . . . 40
3.3.3 The effect of different crustal types . . . 50
3.3.4 Comparison of different methods . . . 53
3.4 Consequences for surface wave tomography . . . 55
3.5 Discussions . . . 59
3.6 Conclusions . . . 63

4 Assessment of tomographic mantle models using spectral element seismograms 65
4.1 Introduction . . . 66
4.2 Models entering the calculation of the synthetic seismograms 67
4.3 Measuring misfits . . . 69
4.3.1 Comparing real and SEM seismograms . . . 69
4.3.2 Measuring time-shifts . . . 71
4.3.3 Measuring amplitude differences . . . 74
4.4 Results . . . 75
4.4.1 Time-shift analysis . . . 75
4.4.2 Amplitude analysis . . . 79
4.5 Discussions . . . 88
4.6 Conclusions . . . 94

5 Assessing new and existing misfit functions for full waveform inversion 97
5.1 Introduction . . . 98
5.2 Misfit functions and associated adjoint sources . . . 102
5.2.1 Misfit kernels . . . 102
5.2.2 Instantaneous phase measurements . . . 105
5.2.3 Generalized cross-correlation measurements . . . 108
5.2.4 Waveform measurements . . . 112
5.2.5 Travel-time measurements . . . 112
5.2.6 Comparison of adjoint sources . . . 113
5.3 Results . . . 115
5.3.1 PREM vs. S20RTS+Crust2.0 . . . 115

5.3.2 Fixed upper mantle and crust 121
5.4 Discussions 130
5.4.1 Phase jumps 130
5.4.2 The effect of noise on simulations 131
5.4.3 Comparison of the misfit functions 135
5.5 Conclusions 135

6 Concluding remarks **137**

Bibliography **141**

Summary **153**

Samenvatting (Summary in Dutch) **155**

Acknowledgments **157**

Curriculum vitae **159**

Chapter 1

Introduction

Seismic tomography is the most powerful tool to image the Earth's interior. However, we have reached a stage where further refinements with classical techniques become very difficult. Advances in numerical methods and computational facilities are providing new opportunities in seismic tomography to enhance the resolution of tomographic mantle images. Meanwhile, an important step is to check the reliability of the current mantle images and the classical techniques using 3-D numerical simulation of seismic wavefield. In this thesis, using a spectral element method, we underline some shortcomings in global mantle tomography, particularly crustal effects on mantle images, and assess tomographic mantle models in order to understand to what extent the current models explain the real Earth's mantle. In the last part of the thesis, we propose new misfit functions such as instantaneous phase and cross-correlation differences between real and synthetic data with an aim to extract as much information as possible from a single seismogram.

1.1 History of Seismic Tomography

The illumination of the Earth's interior based on seismological observations goes back to the beginning of the 20th century following the developments in instrumentation which led to the installation of the first global seismic networks and the availability of collections of global seismic data. Based on the systematic behavior of body wave arrivals as a function of distance and time, the first order structures, such as the existence of the core (Oldham, 1906), the Moho discontinuity (Mohorovičić, 1909) and the existence of the

inner core (Lehmann, 1936) were rapidly identified. In the following years, the body wave travel-times were tabulated by Jeffreys & Bullen (1940) and are still used to locate earthquakes. This collection of travel times allowed to obtain the early velocity profiles of the Earth as a function of radius.

With the necessity of a common reference Earth model in seismology, Dziewonski & Anderson (1981) constructed an average model, known as PREM (Preliminary Reference Earth Model), as a function of the Earth's radius by modeling a large number of body wave travel-times, surface wave and free-oscillation data, which is still being used as a reference model. IASP91 of Kennett & Engdahl (1991) and an improved version ak135 of Kennett et al. (1995) were constructed using high frequency body wave travel-time measurements only.

It soon became clear that the radially symmetric models were not enough to explain the complexity of seismic signals. The first attempts to retrieve the lateral variations in the Earth's structure were done by Aki et al. (1977), Sengupta & Toksöz (1977) and Dziewonski et al. (1977). This marks the beginning of seismic tomography and it henceforth remained an indispensable tool in identifying the Earth's interior in 3-D dimension. Since the tomographic images are constructed based on perturbation theory, 1-D models are still commonly used as a background model in tomography.

1.1.1 The Earth's Mantle from Classical Seismic Tomography

Mapping the Earth's mantle in terms of its composition, thermal variations and structure is essential to understand the mantle dynamics and the surface processes such as the origin of hotspots and the forces behind plate motions. Furthermore, an accurate velocity structure is important for locating earthquakes and is required from an engineering point of view to assess seismic hazards and to detect nuclear explosions.

The Earth's mantle has been investigated by many researchers using different data sets such as body waves, surface waves, splitting functions of the Earth's free oscillations or a combination of these different data together at either a global or continental scale. For instance, P and S-wave arrival times were used to construct global and continental models by e.g., Dziewonski (1984); Zhou (1996); van der Hilst et al. (1997s); Kennett et al. (1998); Bijwaard & Spakman (2000); Boschi & Dziewonski (2000); Kárason & van der Hilst (2001); Kennett & Gorbatov (2004). Surface waves are

commonly used to map the crust and upper-mantle structure either using phase dispersion or direct waveform modeling (e.g., Trampert & Woodhouse, 1995; Ekström et al., 1997; Ritzwoller & Levshin, 1998; Shapiro & Ritzwoller, 2002). Waveforms of surface and S waves were also modelled using, for instance, cross-branch modal coupling (e.g., Li & Romanowicz, 1996) or partitioned waveform inversion (e.g., Lebedev et al., 2005). The measurements of splitting functions of the Earth's free oscillations (e.g., Woodhouse & Dahlen, 1978; He & Tromp, 1996; Resovsky & Ritzwoller, 1998) were used together with surface wave data to constrain the long wavelength structure of the deep interior (e.g., Masters et al., 1982; Woodhouse & Dziewonski, 1984; Beghein et al., 2002). Surface waves provide a better ray coverage particularly at the oceans to map the upper-mantle structure whereas lower mantle is better constrained by body waves. Therefore, it is common to integrate different data sets that are sensitive to different depths and properties of the Earth's interior in tomography to increase the resolution. A few examples that combine body wave, free oscillation and surface wave data are e.g., Su et al. (1994); Masters et al. (1996); Ritsema et al. (1999); Mégnin & Romanowicz (2000); Gu et al. (2001).

To have a better understanding of mantle processes, it is well known that anisotropy has to be taken into account. Seismic velocities, that are sensitive to the variations in material properties, can vary as a function of horizontal and vertical directions yielding *azimuthal* and *radial* anisotropy. Thus, detecting anisotropy is informative on the composition and dynamics of the mantle. Anisotropy was incorporated in seismic tomography, for instance, by Montagner & Tanimoto (1991); Ekström & Dziewonski (1998); Debayle & Kennett (2000); Trampert & van Heijst (2002); Trampert & Woodhouse (2003); Beghein et al. (2006); Visser et al. (2008).

In classical seismic tomography, it is common to separate phase and amplitude information and invert them individually. Body wave tomography is mostly based on travel-time measurements whereas surface wave experiments are based on phase measurements. Either travel-time or phase measurements allow to interpret the elastic properties of the Earth's structure. Due to practical difficulties, however, the number of studies on amplitude tomography is less compared to phase and travel-time tomography. The amplitude of seismic waves contains vast information on both elastic and anelastic parameters of the Earth's interior. The effect of lateral elastic heterogeneities on amplitudes was presented, for instance, by Woodhouse & Wong (1986); Selby & Woodhouse (2000). Thus, amplitude information

provides necessary constrains to identify small-scale elastic structures (e.g., Laske & Masters, 1996). Retrieving anelastic parameters is important to constrain the thermal and chemical structure of the Earth's interior. Attenuation also causes physical dispersion which effects the velocity of seismic waves (e.g., Kanamori & Anderson, 1977; Romanowicz, 1990). 3-D mantle attenuation models were obtained using surface wave (e.g., Romanowicz, 1995; Selby & Woodhouse, 2000; Gung & Romanowicz, 2004; Dalton & Ekström, 2006; Dalton et al., 2008) and body wave amplitudes (e.g., Bhattacharyya et al., 1996; Reid et al., 2001; Warren & Shearer, 2002). For recent reviews of seismic mantle tomography, the readers are invited to consult the review papers of Romanowicz (2003) and Trampert & van der Hilst (2005).

1.1.2 Finite-Frequency Effects in Seismic Tomography

Due to its simplicity and ease of implementation, classical seismic tomography is based on ray theory in which the energy of waves propagates along ray paths. Ray theory has provided robust images of long-wavelength structure of the Earth's interior. However, being a high frequency approximation, ray theory has some theoretical limitations and breaks down when the scale length of the heterogeneities are smaller than the width of the first Fresnel zone (Wang & Dahlen, 1995). The current resolution of the tomographic images has clearly reached the limits of ray theory (e.g., Passier & Snieder, 1995; Spetzler et al., 2001) therefore finite-frequency effects have started to become important.

Woodhouse & Girnuis (1982) and Snieder (1993) took into account the finite-frequency effects in surface waves by computing finite-frequency kernels, or so-called Fréchet kernels, using normal mode theory. Marquering et al. (1999) have shown that finite-frequency body waves are sensitive to 3-D structure off the ray path and paradoxically, their sensitivity is zero along the ray path although in the high frequency limit, finite-frequency theory converges to ray theory. Due to the banana shape of the kernels on a cross-section along the ray path and doughnut shape with a hole inside on the cross-section across the ray path, they named the finite-frequency kernels *banana-doughnut kernels*. This theory was afterwards extended by e.g., Tong et al. (1998); Dahlen et al. (2000); Hung et al. (2000); Zhao et al. (2000) and applied to tomography problems (e.g., Montelli et al., 2004). To improve the validity of smaller scale heterogeneities in surface wave studies 2-D (e.g., Spetzler et al., 2002; Yoshizawa & Kennett, 2002; Ritzwoller

et al., 2002) or 3-D (e.g., Zhou et al. 2004) finite-frequency kernels have recently been used. Although it has been proved that finite-frequency theory is a better forward theory than ray theory, currently there is no consensus whether this also improves the resolution of the tomographic models (e.g., Ritzwoller et al., 2002; Zhou et al., 2004; Sieminski et al., 2004; Trampert & Spetzler, 2006).

1.1.3 Numerical Methods in Seismic Tomography

The developments in computer facilities and the advances in numerical methods recently enabled us to simulate wave propagation in 3-D structures. Numerical simulation of wave propagation provides new opportunities to incorporate finite-frequency effects in seismic tomography. We are now able to compute finite-frequency sensitivity kernels in 3-D background models and combine different data observables having different frequency content. Since seismic tomography is a non-linear problem, the images will eventually improve in an iterative scheme based on numerical simulations.

Among all numerical techniques, the spectral element method stands out as the most successful to simulate the full wave propagation in complex 3-D Earth models (Capdeville et al., 2003; Chaljub et al., 2003; Chaljub & Valette, 2004; Komatitsch & Tromp, 2002a,b) in terms of its accuracy, ease of applicability to complex structures and computational time requirements. We are now at a stage to use the opportunity of incorporating the complete seismic wavefield in the forward theory and using 3-D reference models in seismic tomography. For instance, Bijwaard & Spakman (2000) tried to enhance global P-wave model based on a 3-D ray tracing method. Zhao et al. (2005) used a finite difference method and computed sensitivity kernels in 3-D reference models by storing the computed Green's functions. In this way, one can have access to the Hessian of the misfit function however, there are some practical difficulties such as tremendous increment in computational and storage tasks. Tromp et al. (2005) suggested to use adjoint techniques, as suggested by Tarantola (1984); Talangard & Courtier (1987) and extended, for instance, by Crase et al. (1990); Fink (1997); Pratt (1999); Akçelik et al. (2003), which provides an elegant way to compute the gradient in an inverse algorithm. They showed that the gradient of the misfit function may be obtained by the interaction of two wavefields, a forward and an adjoint, by performing only two numerical simulations for each source independent from number of receivers. The idea is similar to seismic reflector mapping (Claerbout, 1971) where the downgoing

wavefield can be obtained from the time-reversed upgoing wavefield. Some attempts to use adjoint tomography based upon a spectral element method have been done for a regional or continental scale using travel-times (e.g., Tape et al., 2007) or full waveform tomography (e.g., Fichtner et al., 2008).

The ultimate aim in global tomography is to perform full waveform inversion, that has become possible with the improvements of 3-D numerical wave simulations. It has been applied to some local problems (e.g., Pratt, 1999; Chen et al., 2007b,a). At global scale, however, it has only been used in a synthetic experiment based on a source stacking technique using a spectral element method by Capdeville et al. (2005). As Capdeville et al. (2005) pointed out, insufficient data coverage due to lack of stations particularly at the oceanic crust is one of the major obstructions in full waveform tomography at the global scale.

1.2 Shortcomings in Global Mantle Tomography

Many mantle models have been proposed by different groups using different data sets and different techniques. Tomographic models of the mantle at low degrees are, in general, in agreement with each other independent from the data type and inversion strategies (see Ritzwoller & Lavely, 1995; Trampert & Woodhouse, 2001). The discrepancy between the tomographic models, however, becomes visible as the resolution increases to map the smaller scale heterogeneities for several reasons. Therefore, there is not a general consensus on mapping the smaller scale heterogeneities, yet.

An important issue in global mantle tomography is that the thin but highly heterogeneous crust has a dramatic effect on seismic waves, particularly on surface waves, which is often as large as the mantle signal. Montagner & Jobert (1988) showed that shallow layer variations can be highly non-linear even at periods longer than 100 s. Therefore, having an accurate crustal model is a major requirement in mantle tomography. The crust is too thin to be resolved in classical seismic tomography thus in general crustal corrections are applied to surface waves (e.g., Ritsema et al., 1999; Boschi & Ekström, 2002; Trampert & Spetzler, 2006). In recent years, crustal corrections have been done using 3-D global crustal models such as Crust2.0 (Bassin et al., 2000) using first-order approximations. Crustal influence on body waves is less compared to surface waves however regional variations in the crustal and upper-mantle structures can mislead the hypocentral locations and smear into mantle models. In general, travel-

times are corrected based on the average of teleseismic rays arrive within a specific surface patch with respect to a reference model (e.g., Engdahl et al., 1998). Therefore, the crust has been standing out as one of the major problems to be addressed to improve the global mantle images.

At the current resolution of the tomographic images, finite-frequency effects are important. Although the finite-frequency theory is a better forward theory than ray theory, as mentioned in the previous section, there are controversial results on the advantage of the use of finite-frequency kernels in tomography. Trampert & Spetzler (2006) showed that same results can indeed be obtained from both ray theory and finite-frequency theory by only changing the damping of the tomographic images pointing out that the advantage of a better forward theory is actually lost due to problems of inverse techniques. Although damping is an important parameter in inverse theory to control the resolution of the tomographic images, the value is fixed in an arbitrary and a subjective manner.

Inverse theory, as Jackson (1972) stated, is the procedure to interpret *inaccurate, insufficient, and inconsistent data*. In seismic tomography, a set of model parameters, which are usually velocity of seismic waves, density, or anelasticity, are inferred through a set of observed data, that are recorded seismograms in our case. Seismic tomography, in this way, becomes an optimization problem where we try to minimize the difference between observed and predicted data through model parameters. There are several factors controlling the resolution of seismic tomography. The choice of parametrization is one of them. Horizontal and radial variations are parametrized by local adaptive grids or expanded into spherical harmonics, and polynomials or splines, respectively. Examples of adaptive local grids in travel-time tomography are found in e.g., Zhou (1996); van der Hilst et al. (1997s); Kennett et al. (1998); Boschi & Dziewonski (2000). In surface wave, free-oscillation or full waveform tomography, in general, spherical harmonics and splines are used to parametrize the horizontal and radial variations, respectively (e.g, Woodhouse & Dziewonski, 1984; Li & Romanowicz, 1996; Ekström & Dziewonski, 1998; Ritsema et al., 1999; Boschi & Ekström, 2002; Lebedev et al., 2005). It has been shown that the coarse or low-degree parameterizations may bias the tomographic models (e.g., Giardini et al., 1988; Snieder et al., 1991) however, Trampert & Snieder (1996) showed that this problem can be overcome by a weighting applied in the cost function. Different parametrization techniques may also

have different implications on the inverted model. Nevertheless, Boschi & Dziewonski (1999) reported that most discrepancies due to different parameterization techniques tend to disappear by a sufficiently strong regularization and discrepancies are more likely to be due to insufficient data coverage. Indeed, the uneven distribution of source and receivers on the Earth is one of the major restrictions in seismic tomography which is also one of the reasons that conceals the advantage of finite-frequency theory (Boschi et al., 2007). In general, due to the approximations used in the forward part of the seismic tomography, it is common to use secondary observables such as travel-time or phase of an isolated phase or only a small portion of a seismogram. Most of the time, the amplitude information is ignored which results to waste a large amount of information contained in a seismogram. Thus the ultimate aim in global tomography is to use full waveforms in the inversion. A way to remedy the insufficient data at the global scale is to try to extract as much information as possible from a single seismogram using the advantages of numerical simulations of wave propagation.

1.3 Motivation and Outline

Numerical methods have a crucial role in refining the current tomographic images as can be summarized by the following reasons: 1) They enable us to consider the full non-linearity of wave propagation, 2) we can now compute the data sensitivities in 3-D background models that allow us to handle the non-linearities better and, 3) as a result of (1), we can extract as much information as possible from a single seismogram to close the gap in data coverage without worrying to identify the seismic phases. Using the advantages of 3-D numerical simulations, the ultimate aim in seismic tomography is to use full waveforms to improve mantle models. 3-D numerical simulation of wave propagation is also a very powerful tool to investigate and understand the current status of seismic tomography. It is of great importance to underline the major shortcomings in seismic tomography as a guide towards using 3-D wave propagation in the inverse problem. In this thesis, we concentrate on global mantle tomography. Using 3-D numerical wave simulations, we tried to underline some of the characteristic problems of identification the mantle structures, as mentioned in the previous section, and suggest some ways of how to improve global tomographic images. All wave simulations were done by the spectral element method of Komatitsch

& Tromp (2002a,b) to compute synthetic seismograms and finite-frequency adjoint kernels. We particularly investigated the following problems:

Accuracy of crustal corrections: We first investigated the accuracy of crustal corrections and their effects on tomographic mantle images in **Chapter 3**. Using first-order approximations, e.g., ray theory and finite-frequency theory, we estimated crustal corrections for fundamental mode surface waves and compared them to those obtained from full 3-D wave simulations with the spectral element method. Based on the obtained results, we discussed the consequences of crustal corrections in global seismic tomography.

Assessment of mantle models: In order to understand to what extent the current tomographic mantle images represent the real Earth′s mantle, we compared real seismograms to those computed by the spectral element method in some 3-D mantle models in **Chapter 4**. In particular, we examined the effect of regularization on our interpretation of the mantle by choosing differently damped mantle models. We compared the results with those from 1-D crustal and mantle models to examine the source of the discrepancies between the real and synthetic seismograms. This study is particularly important for suggesting a way to assess how far the current tomographic images represent the Earth's structure.

Assessing new and existing misfit functions for full waveform inversions: Finite-frequency adjoint kernels based on numerical wave simulations allow us to visualize the 3-D sensitivity of the selected data which is controlled by the chosen misfit functions at the beginning of inverse problem. In **Chapter 5**, in order to extract as much information as possible from a single seismogram, we have developed new misfit functions based on the instantaneous phase difference between real and synthetic data, and the difference between the cross-correlation of observed data with synthetic data and the auto-correlation of data. Both demonstrated misfit functions allow to consider a wave train rather than an isolated single phase. We compared different misfit functions such as instantaneous phase, cross-correlation, waveform and travel-time differences using the adjoint kernel simulations based upon the spectral element code to have insight in the advantages/disadvantages of the chosen misfit functions.

For the completeness of the thesis, we give a brief background on the spectral element method and a description how we used the code in **Chapter 2**. We summarize our results in **Chapter 6**.

Chapter 2

Spectral element method in 3-D wave simulation and seismic tomography

The computation of synthetic seismograms is an essential part of seismic imaging. The developments in numerical methods for wave propagation in the last decade together with the advances in computational facilities will play a crucial role in seismic tomography. We are now able to simulate wave propagation in complex 3-D structures at either a regional or global scale using numerical methods with a high accuracy and the current challenge is to use this opportunity in seismic tomography to improve the current resolution of the tomographic images. This thesis is based on numerical simulations by a spectral element method to address some problems of classical seismic tomography and to suggest some possible solutions towards full waveform inversion. For the completeness of this thesis, in the following, we will give a brief overview of the spectral element and adjoint methods. We explain how we used the spectral element code of Komatitsch & Tromp (2002a,b). For more detailed information, the reader is referred to the original papers, cited as references, and the user manual of the spectral element code.

2.1 Computation of synthetic seismograms

The methods used to compute synthetic seismograms can roughly be classified into three groups: semi-analytical, asymptotic, and numerical methods.

Normal mode summation (e.g., Gilbert, 1970; Dahlen & Tromp, 1998) and the reflectivity method (e.g., Fuchs, 1968; Fuchs & Müller, 1971) are the most used semi-analytical methods where the synthetic seismograms are computed for a given 1-D Earth model (e.g., PREM by Dziewonski & Anderson (1981)) in which there are only radial variations of parameters. In normal mode summation, the wavefield is obtained by the summation of properly excited eigen functions which gives a good accuracy down to 8 s. Attenuation, gravity, transverse isotropy and discontinuities can easily be taken into account. In the reflectivity method, the reflection coefficients for a given layered medium are integrated in frequency-wavenumber domain. This is often used for higher frequency wavefields. However, the real Earth is more complex than a radially symmetric model. Since there are no analytical solutions for the elastodynamic wave equation in 2-D or 3-D models, we need to use either asymptotic or numerical methods to simulate the wave propagation in such complicated structures. Asymptotic methods, such as ray theory (e.g., Cerveny, 2001), or nonlinear asymptotic coupling theory (e.g., Li & Tanimoto, 1993), have been frequently used in seismology taking the advantage of reasonable computational requirements but limiting the resolution to the scale length of the heterogeneities larger than the width of the Fresnel zone (Wang & Dahlen, 1995).

Full wave propagation can be modeled by a numerical solution to the wave equation. A commonly used technique is the finite difference method (e.g., Olsen et al., 1995; Igel & Weber, 1996; Thomas et al., 2000) in which the numerical derivatives are estimated based on finite-difference equations. It is popular due to its simplicity and ease of implementation in uniform grids however the *strong* formulation of the wave equation can be problematic for applying boundary conditions accurately in complex structures. In addition, particularly surface waves may suffer from numerical dispersion. Pseudo-spectral methods (e.g., Tessmer et al., 1992; Carcione & Wang, 1993; Furumura et al., 1998; Igel, 1999), in which the wavefield is expanded into global polynomial basis functions, have similar problems due to the strong formulation of wave equation which prevents the application of these methods to the global simulation of 3-D wave propagation. To overcome these problems, normal mode coupling (e.g., Park, 1986; Lognoné & Romanowicz, 1990; Capdeville et al., 2000), in which the 3-D variations are considered in terms of the summation of normal modes computed for spherically symmetric models, and the direct-solution methods (e.g., Hara et al., 1991; Geller & Ohminato, 1994) have been used. Both mode coupling and

the direct-solution methods are computationally expensive to simulate wave propagation in 3-D structures and can be problematic to deal with the large variations in crustal thickness and to grid the entire globe. Finite element methods (e.g., Toshinawa & Ohmachi, 1992; Bao et al., 1998) are efficient to solve such problems. The disadvantage is that the use of low order polynomial degrees may be problematic for the spatial discretization. A way to remedy this problem is to increase the order of polynomials, however, it makes the method impractical due to an increase in computation time and inflexibility of using parallel algorithms.

Higher-order polynomials can more easily be taken into account using spectral methods. The spectral element method (SEM) is one of them which combines the flexibility of finite element methods with the accuracy of pseudo-spectral methods (Patera, 1984) and enables the accurate simulation of wave propagation in complex 3-D structures at either a regional or global scale. SEM enables the parallel simulation of wave propagation in complex 3-D structures. The early applications of the method were proposed in fluid mechanics (e.g. Patera, 1984) to solve incompressible Navier-Stokes equations. Fischer & Rønquist (1994) applied a parallel spectral element algorithm to calculate the incompressible flows in complex structures. The first use of the spectral element method in seismology started with 2-D simulations of seismic wave propagation (e.g. Priolo et al., 1994) and continued with the applications to 3-D wave propagation problems both at regional (e.g. Komatitsch & Vilotte, 1998; Seriani, 1998; Komatitsch & Tromp, 1999; Komatitsch et al., 2004) and global (e.g. Capdeville et al., 2003; Chaljub et al., 2003; Chaljub & Valette, 2004; Komatitsch & Tromp, 2002a,b; Komatitsch et al., 2002) scale. In the following, we give a brief overview of the spectral element method.

2.1.1 Seismological background

In seismology, it is convenient to use the Lagrange description to drive the equation of motion since the data, the seismograms, are the records of particle motions at a specific point on the Earth as a function of space and time. Using Cartesian coordinates and index notation, the relation between strain and displacement can be given as,

$$e_{ij} = \frac{1}{2}(s_{i,j} + s_{j,i}) \tag{2.1.1}$$

where e_{ij} is the second order symmetric strain tensor, $u_{i,j}$ and $u_{j,i}$ are the spatial derivatives of the displacement field $\mathbf{s}$. Using Hook's law, the linear relation between strain ($\mathbf{e}$) and stress ($\mathbf{T}$) tensors are given as,

$$T_{ij} = c_{ijpq} e_{pq} \tag{2.1.2}$$

where c_{ijpq} is a fourth order tensor built from the elastic constants. Because of internal symmetries, c_{ijpq} has 21 independent elastic parameters. For an isotropic medium, we only need two, the Lame parameters λ and μ,

$$c_{ijpq} = \lambda \delta_{ij} \delta_{pq} + \mu(\delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp}). \tag{2.1.3}$$

Inserting the constitutive equations into momentum equation,

$$\rho \partial_t^2 \mathbf{s} - \nabla \cdot \mathbf{T} = \mathbf{f} \tag{2.1.4}$$

we find the equation of motion,

$$\rho \ddot{\mathbf{s}} - (\lambda + \mu)\nabla^2 \mathbf{s} + \mu(\nabla^2 \mathbf{s})^T = \mathbf{f} \tag{2.1.5}$$

where $\mathbf{f}$ represents the seismic source which in the case of a point source at $\mathbf{x_s}$ can be written as,

$$\mathbf{f} = -\mathbf{M} \cdot \nabla \delta(\mathbf{x} - \mathbf{x_s}) S(t). \tag{2.1.6}$$

Here $\mathbf{M}$ represents the *moment tensor* and $S(t)$ is the *source-time function*.

To solve the equation of motion, we have to specify the boundary and initial conditions. The boundary conditions for the whole Earth are such defined that the traction must vanish at the free surface ($\hat{\mathbf{n}}.\mathbf{T} = \mathbf{0}$) and must be continuous at solid-solid, solid-fluid boundaries together with the displacement. The initial conditions are given as

$$\mathbf{s}(\mathbf{x}, t = 0) = 0, \qquad \partial_t \mathbf{s}(\mathbf{x}, t = 0) = 0. \tag{2.1.7}$$

2.1.2 Spectral element method

Unlike finite difference and pseudo-spectral methods, the spectral element method is, for computational convenience, based on an integral or so-called *weak* form of the equation of motion. This integral form is obtained by taking the dot product of the momentum equation eq. 2.1.4 with an arbitrary test vector $\mathbf{w}$ and integrating over a volume Ω which can be written as,

$$\int_{\Omega} \rho \mathbf{w} \cdot \partial_t^2 \mathbf{s} d^3\mathbf{r} = -\int_{\Omega} \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} + \mathbf{M} : \nabla \mathbf{w}(\mathbf{r_s}) S(t). \tag{2.1.8}$$

where the left-hand side of the equation gives the *mass matrix*. The right-hand side is obtained by integration by parts where the first term is called the *stiffness matrix* and the second term is associated with the source. The free surface condition is mutually taken into account since the integral is defined over the volume Ω. For solid and fluid regions, the wave equation is solved separately and than coupled together by the appropriate boundary conditions. This leads to formulate the problem individually for crust and mantle, outer core, and inner core (see Komatitsch & Tromp, 2002a). Considering the boundary conditions at each domain, eq. 2.1.8 can be written for the crust and mantle,

$$\begin{aligned} \int_M \rho \mathbf{w} \cdot \partial_t^2 \mathbf{s} d^3\mathbf{r} &= -\int_M \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} + M : \nabla \mathbf{w}(\mathbf{r_s}) S(t) \\ &\quad + \int_{CMB} p \hat{\mathbf{n}} \cdot \mathbf{w} d^2 \cdot \mathbf{r} \end{aligned} \tag{2.1.9}$$

where M and CMB stand for the mantle (and crust), and the core-mantle boundary, respectively. $\hat{\mathbf{n}}$ is the unit outward normal on the CMB and $p\hat{\mathbf{n}}$ is the traction in the fluid where p denotes the fluid pressure. For the fluid outer core,

$$\begin{aligned} \int_{OC} \kappa^{-1} w \partial_t^2 \chi d^3\mathbf{r} &= -\int_{OC} \rho^{-1} \nabla w \cdot \nabla \chi d^3\mathbf{r} + \int_{CMB} w \hat{\mathbf{n}} \cdot \partial_t \mathbf{s} d^2\mathbf{r} \\ &\quad - \int_{ICB} w \hat{\mathbf{n}} \cdot \partial_t \mathbf{s} d^2\mathbf{r} \end{aligned} \tag{2.1.10}$$

where OC and ICB stand for the outer core and the inner core boundary, respectively. κ is the adiabatic bulk modulus in fluid, $\mathbf{v}$ is the velocity in

fluid, χ is the scalar potential and w is the scalar test function. In this way, eqn. 2.9 and 2.10 correctly represents the fluid-solid interactions (Komatitsch & Tromp, 2002a). Finally, for the inner core the weak formulation is,

$$\int_{IC} \rho \mathbf{w} \cdot \partial_t^2 d^3\mathbf{r} = -\int_{IC} \nabla \mathbf{w} : \mathbf{T} d^3\mathbf{r} - \int_{ICB} p\hat{\mathbf{n}} \cdot \mathbf{w} d^2\mathbf{r} \tag{2.1.11}$$

where IC represents the inner core.

In numerical techniques, the accuracy of the simulations strongly depends on the quality of the mesh. Mesh design in SEM is the same as in finite element methods. The model volume is subdivided into a number of non-overlapping hexahedral finite elements. Each hexahedral element matches exactly with the sides of other neighbouring element which is called a *conforming mesh* in the literature. The size of the elements controls the resolution in the computed seismograms such that the smaller the size the shorter period you get in simulations.

Once the mesh is designed, the next step is to define the basis functions that represents the wavefield in each element. In SEM, the shape of each element is defined by low-degree polynomials whereas the basis functions that represent the unknown displacement wavefield and test vectors are defined by high order polynomials. Lagrange polynomials of degree 4 to 10 are commonly used as basis functions. The $n+1$ Lagrange polynomials are defined by $n+1$ control points which are chosen to be Gauss-Lobatto-Legendre points. This decision leads to a diagonal mass matrix which is one of the advantages of SEM over finite element methods since a diagonal mass matrix considerably reduces the cost in computation. The weak formulation of the equation of motion, that is solved at each element based on Gauss-Lobatto-Legendre polynomials, is then integrated over all elements.

2.2 Using 3-D simulations in seismic tomography

The accurate 3-D simulation of wave propagation by numerical methods is of great importance to seismic tomography because it can take the full nonlinearity of the problem into account. We have now the opportunity to use numerical wave propagation in 3-D Earth models to improve the seismic images of the Earth's interior. For instance, Bijwaard & Spakman (2000) tried to enhance global P-wave model based on a 3-D ray tracing method.

Zhao et al. (2005) used a finite difference method and computed sensitivity kernels in 3-D reference models by storing the computed Green's functions. In this way, one can have access to the Hessian of the misfit function but the difficulties are a tremendous increment in computational and storage tasks. Chen et al. (2007b,a) introduced 3-D numerical simulations into some local tomography problems using the scattering integral method based on a finite difference method. A way to reduce the computational costs is to use *adjoint methods*. The introduction of adjoint techniques in geophysics was by Tarantola (1984, 1988) where he gave the theoretical basis for the use of adjoint methods in exploration problems. He showed that Fréchet derivatives of a given misfit function, which gives the sensitivity of data with respect to model parameters, may be obtained by the "backpropagation" of data residuals. Talagrand & Courtier (1987) showed in an application to meteorological problems that one can obtain the gradient of the misfit function by using adjoint calculations. The theory was afterwards extended by e.g., Crase et al. (1990); Pratt (1999); Akçelik et al. (2003). Tromp et al. (2005) put the link between seismic tomography and time-reversal imaging (e.g., Fink, 1997) using adjoint simulations based on 3-D numerical wave propagation and demonstrated that the gradient of a given misfit function may be obtained by the interaction between the *forward* and the *adjoint* wavefields which requires only two numerical simulations per event. The sensitivity kernels computed by adjoint methods are finite-frequency kernels (e.g., Tromp et al., 2005; Liu & Tromp, 2006, 2008). At regional or continental scale, adjoint tomography has been put into practice by using traveltime measurements (Tape et al., 2007) and full waveform inversion (Fichtner et al., 2008) based upon a spectral element method.

In this thesis, we used adjoint simulations based upon the spectral element code of Komatitsch & Tromp (2002a,b) to compute data sensitivities in order to analyze the advantages/disadvantages of chosen misfit functions. For the completeness, in the following, we give some background information on adjoint simulations.

2.2.1 Adjoint methods

In inverse theory, the aim is to minimize the difference between a set of observed and predicted data defined through a misfit function by updating the model parameters. In classical tomography, we generally have access to the first and second derivative of the misfit function, called the *gradient* and the *Hessian*, respectively, thus the model parameters can be updated in one step. In adjoint tomography, the gradient is obtained through adjoint simulations, however, not the Hessian. Therefore, the inverse problem is solved iteratively based upon a *gradient* method such as conjugate gradient or steepest descent (see Tape et al., 2007).

As shown by Tromp et al. (2005), the gradient of a misfit function, which are the Fréchet derivatives, can be determined by the interaction of forward wavefield, from source to receiver, and adjoint wavefield, from receiver to source. This adjoint wavefield is constructed by an adjoint source which is derived from a particular misfit function between observed and synthetic seismograms and can be designed for a particular seismic observable. There are several measurement techniques in seismic tomography such as cross-correlation travel-time, amplitude ratio, or waveform difference, for which the construction of sensitivity kernels are presented in Tromp et al. (2005). For instance, the misfit function for the waveform difference is,

$$\chi(m) = \frac{1}{2} \sum_{r=1}^{N} \int_0^T \|\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) - \mathbf{d}(\mathbf{x_r}, t)\|^2 \, dt \tag{2.2.1}$$

where $\mathbf{s}$ and $\mathbf{d}$ denote the synthetic and the observed data, respectively, which are in general filtered and windowed within a certain frequency band and time window. Partial derivatives with respect to model parameters from this misfit would be,

$$\delta\chi = \sum_{r=1}^{N} \int_0^T [\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) - \mathbf{d}(\mathbf{x_r}, \mathbf{t})] \, \delta\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) dt \tag{2.2.2}$$

where $\delta\mathbf{s}$ is the perturbation in the displacement field $\mathbf{s}$ due to perturbations in model parameters $\delta\mathbf{m}$ which basically gives the Fréchet derivatives based on the Born approximation (Hudson, 1977; Wu & Aki, 1985). Using the reciprocity property of the Green's function, Tromp et al. (2005) derived the adjoint wavefield for the waveform misfit function by,

$$s_k^\dagger(\mathbf{x}^{'}, t^{'}) = \int_0^{t^{'}} \int_V G_{ki}(\mathbf{x}^{'}, \mathbf{x}; t^{'} - t) f_i^\dagger(\mathbf{x}, t) d^3\mathbf{x} dt \tag{2.2.3}$$

where G_{ki} is the Green's function reversed in time and $f_i^\dagger$ is the adjoint source obtained through the selected misfit function to compute the adjoint wavefield given as,

$$f_i^\dagger(\mathbf{x}, t) = \sum_{r=1}^{N} [s_i(\mathbf{x_r}, T - t) - d_i(\mathbf{x_r}, T - t)] \delta(\mathbf{x} - \mathbf{x_r}). \tag{2.2.4}$$

We can also write the gradient of the misfit function in terms of model perturbations,

$$\delta\chi = \int_V K_\rho(\mathbf{x}) \delta\rho(\mathbf{x}) + K_{c_{jklm}}(\mathbf{x}) \delta c_{jklm}(\mathbf{x}) d^3\mathbf{x} \tag{2.2.5}$$

where $\delta\rho$, and δc_{jklm} are the relative model perturbations in density and elastic parameters. The misfit kernels are then given by,

$$K_\rho(\mathbf{x}) = -\int_0^T \rho(\mathbf{x}) \mathbf{s}^\dagger(\mathbf{x}, T - t) \cdot \partial_t^2 \mathbf{s}(\mathbf{x}, t) dt \tag{2.2.6}$$

$$K_{c_{jklm}}(\mathbf{x}) = -\int_0^T \epsilon_{jk}^\dagger(\mathbf{x}, T - t) \epsilon_{lm}(\mathbf{x}, t) dt \tag{2.2.7}$$

where ϵ_{lm} and $\epsilon_{jk}^\dagger$ are the elements of the strain and the adjoint strain tensors, respectively. As can be noticed immediately, the kernels depend on forward and adjoint wavefields which can be obtained by two numerical simulations. The adjoint wavefield is controlled by an adjoint source which is dependent on a defined misfit from a particular observable.

The derivation of the adjoint equations based on a Lagrange multiplier method is presented by Liu & Tromp (2006) and in the review paper of Tromp et al. (2008). We will discuss the adjoint kernels in Chapter 5 introducing new misfit functions and deriving their adjoint sources.

2.3 About SPECFEM3D

A numerical method to compute synthetic seismograms needs to incorporate an anisotropic asthenosphere, a crustal model with rapid changes in Moho thickness, sharp fluid-solid discontinuities at the inner core and the mantle boundaries with the outer core, ellipticity, free surface topography/bathymetry, gravity, ocean effects and attenuation. This can be fully taken into account at either a global or regional scale by the spectral element method code of Komatitsch & Tromp (2002a,b). In the following, we give a brief description of the SEM code SPECFEM3D by Komatitsch & Tromp (2002a,b) which was used to compute all the synthetic seismograms and the adjoint simulations in this thesis.

2.3.1 Meshing the globe

The mesh, which is used for the 3-D global wave simulations, honors the first order (Moho at 24.4 km; upper-mantle discontinuities at 220 km, 400 km and 670 km; core-mantle and inner-core boundaries) and second order (600 km; 771 km and at top of the *D"*) discontinuities in PREM model (Dziewonski & Anderson, 1981). The mesh is doubled in size firstly below the Moho, secondly below 670 km, and thirdly above the inner-core boundary. The resolution in synthetics is controlled by the parameter that corresponds to the number of elements at the surface. The higher the number of surface elements the higher the time resolution but the longer the simulations take. For instance, when the number of surface elements equal to 128×128, a resolution down to approximately 34 s can be obtained whereas 96×96 surface elements leads to a resolution down to approximately 46 s. Using 96 processors, the first and latter take approximately 5 and 3.5 hours, respectively, to compute a 45 min seismogram.

2.3.2 Velocity models

In SPECFEM3D, several 1-D radially symmetric [isotropic and transversely isotropic PREM by Dziewonski & Anderson (1981); IASP91 by Kennett & Engdahl (1991); 1066A by Gilbert & Dziewonski (1975); AK135 by Kennett et al. (1995); REF by Kustowski et al. (2006)], 3-D crustal [Crust2.0 by Bassin et al. (2000)] and 3-D mantle models [S20RTS by Ritsema et al. (1999); S362ANI (radial anisotropy is confined in the uppermantle), S362W-

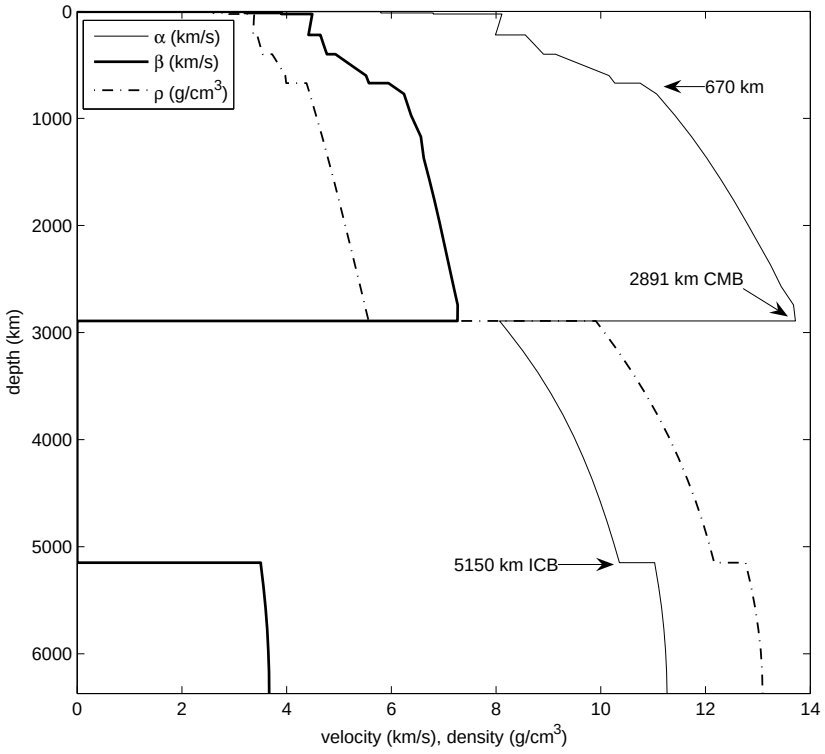

Figure 2.3.1: Preliminary Reference Earth Model PREM (Dziewonski & Anderson, 1981) describing the 1-D Earth's model in terms of P-wave velocity α, S-wave velocity β and density ρ.

MANI (anisotropy throughout the mantle) and S29EA (higher resolution in the uppermantle beneath Eurasia) by Kustowski et al. (2006)] are implemented.

In the SEM code, S20RTS is the default 3-D mantle model which uses transversely isotropic PREM (Figure 2.3.1) as a background model. The other 3-D mantle models are based on the 1-D model REF (there is also a version of S362ANI based on PREM). Since crustal corrections were applied before inversion, S20RTS is used together with Crust2.0. Crust2.0 is compiled from seismic refraction/reflection, receiver function and geological data where available and the rest of the model was interpolated. It

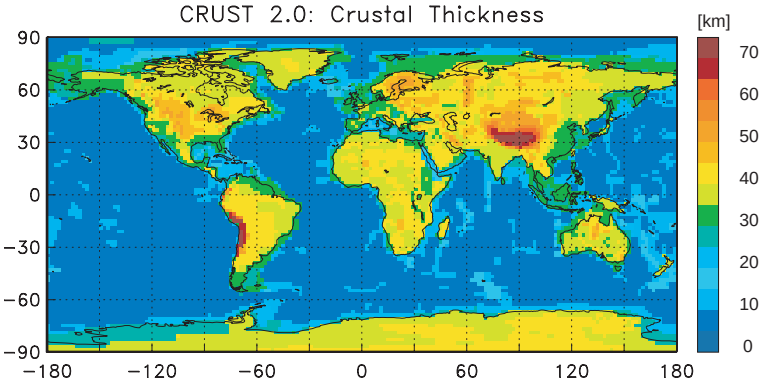

Figure 2.3.2: 3-D global crustal thickness variation from 3-D global crustal model Crust2.0 (Bassin et al., 2000). The model can be obtained from the website http://mahi.ucsd.edu/Gabi/rem.html.

has $2^\circ \times 2^\circ$ resolution which is an improved version of previously released $5^\circ \times 5^\circ$ resolution Crust5.1 (Mooney et al., 1998). Crustal thickness varies between 6.65 km (oceanic crust) and 75 km (underneath the Himalayas) (Figure 2.3.2). P-wave velocity, S-wave velocity and density varies between 5-6.2 km/s, 2.5-3.2 km/s and 2600-2800 kg/m^3, respectively, where the associated upper mantle values for PREM are 5.8 km/s, 3.2 km/s and 2600 kg/m^3. It should be noted that in SPECFEM3D, a smoothed version of Crust2.0 is used (see Komatitsch & Tromp (2002b) for a justification). In addition, the ice layer in Crust2.0 is excluded in the code. S20RTS (Figure 2.3.3) is a 3-D shear velocity model obtained through the inversion of Rayleigh waves, body wave travel-times and splitting functions of the Earth's normal modes. It is expanded into spherical harmonics to degree 20 and parametrized by 21 splines in the vertical direction. Shear velocity perturbations were inverted with respect to PREM. A different global model can easily be implemented in the code to simulate wave propagation. Any 3-D model which varies smoothly laterally can be used in the same way as S20RTS. This is done in our experiments in Chapter 4.

2.3.3 Model parameters

In simulations, it is possible to incorporate attenuation, gravity, ellipticity and rotation effects of the Earth which can be easily switched on in the parameter file. In the SEM code, ellipticity is taken into account as a function of depth based on Clairaut's equation (see Dahlen & Tromp, 1998),

the Cowling approximation is used for the gravity and Coriolis force is incorporated for the Earth's rotation. In addition, free-surface topography and bathymetry can also be incorporated with 5×5 min^2 ETOPO5 model (NOAA, 1988). The effect of the ocean load can be added on top of the model by turning the ocean option on without meshing the ocean column. By default, 1-D PREM attenuation, incorporating shear attenuation only, is used when attenuation is on (except that 1-D model 1066A uses an unpublished 1-D attenuation model from Scripps). 3-D attenuation models can also be implemented. In Figure 2.3.4, vertical, radial and transverse components of synthetic seismograms computed in 1-D and 3-D Earth models are shown. The 1-D model is transversely isotropic PREM with a 3-km thick ocean layer and 3-D model is S20RTS with Crust2.0 on top. The effect of the 3-D mantle and crust on body waves and surface waves are clearly seen both in phase and amplitude. During the simulations oceans, attenuation, gravity, rotation, ellipticity and topography of the Earth are considered except that in PREM seismograms topography was not used. Another note is that Earth's radius is taken 6371 km in all models except that for PREM with ocean layer on top the radius should be set to 6368 km.

2.3.4 Source time function

For global simulations, the source is represented by a point source. The Harvard Centroid-Moment Tensor (CMT) solution is used to describe the point source in the SEM code. The source time function can be incorporated either during or after the simulations based on the half duration defined in the CMT file. If the half duration is set to zero, a very short moment rate function is used to represent the source time function. The advantage of it is that, after the simulations are done, the synthetic seismograms can be convolved with a variety of source-time functions. If the half-duration is not zero, then, the seismogrmas are, by default, convolved by a Gaussian source-time function with half-width of the half duration. The Gaussian source used in the code is like a smoothed triangular source in shape (see Komatitsch & Tromp (2002a) and the manual for details).

2.3.5 Comparison of SEM and normal mode seismograms

The accuracy of SEM seismograms was tested by the comparisons of SEM seismograms computed in 1-D Earth models with normal mode seismo-

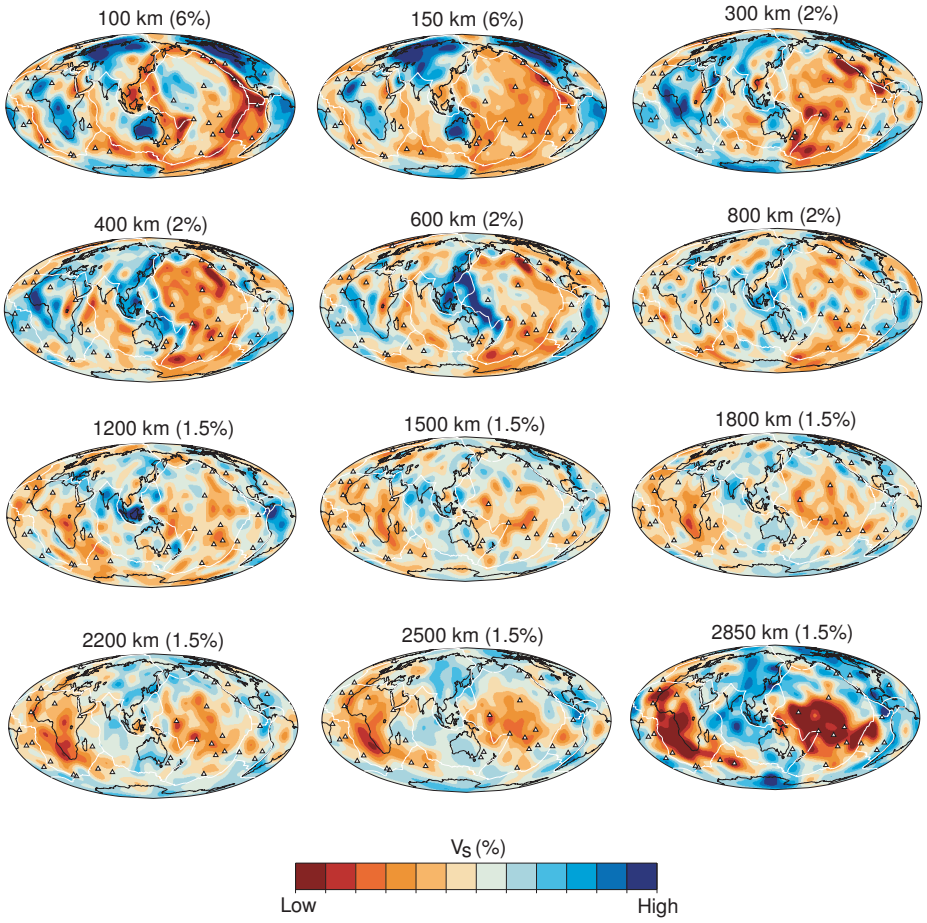

Figure 2.3.3: Shear velocity perturbation maps of the 3-D global mantle model S20RTS (Ritsema et al., 1999) plotted as depth slices.

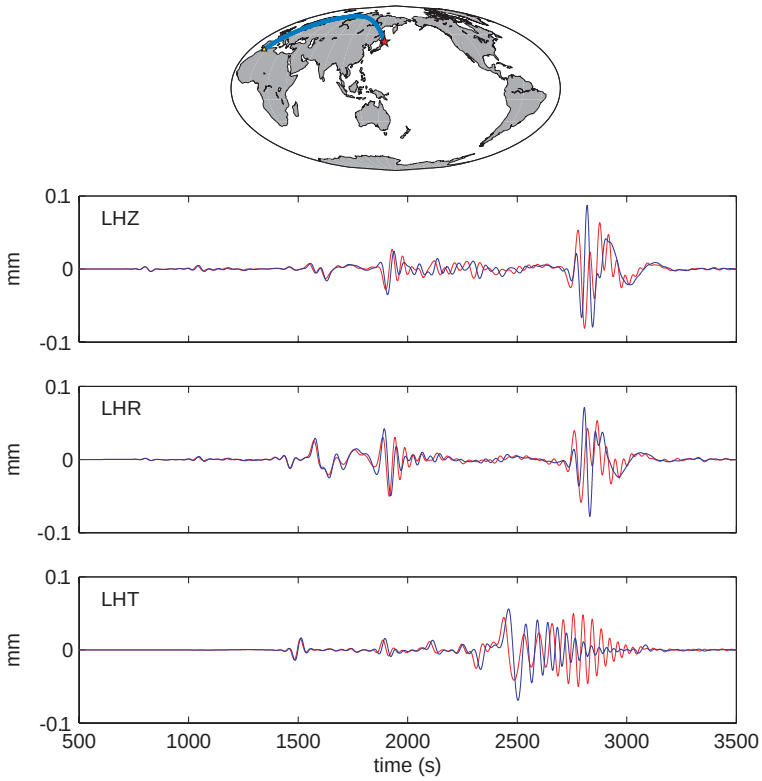

Figure 2.3.4: Synthetic seismograms computed in 1-D PREM (blue) and 3-D mantle S20RTS with 3-D crust Crust2.0 on top (red) for the Hokkaido earthquake (2004 November 28, $M_w = 7.0$) recorded at the station MELI ($\Delta = 97°$). The ray path is shown in the map where the earthquake and the station are denoted by a star and a triangle, respectively.

grams in Komatitsch & Tromp (2002a,b) (also see the manual of the code). It is however good to repeat the comparisons in order to check if the code is used in the right way. This benchmark is quite tricky since one must make sure that all parameters, such as attenuation, gravity, ellipticity, rotation etc., are the same in both methods. Furthermore, the PREM model used in the SEM code can be different from the original PREM model. Therefore, the best strategy is to take the velocity model from the SEM code and feed it into the normal mode code. We recommend that this is done in any benchmark or comparative studies (see Chapter 3). Such examples can be found in the package of the SPECFEM3D version 4.0. In Figure 2.3.5, we present a sample comparison between the SEM and normal-mode seismograms at vertical, radial and transverse components from a sample ray path where the both sets of synthetics perfectly match with each other.

2.3.6 Adjoint simulations

Adjoint simulations for one earthquake basically requires two numerical simulations, one for the forward wavefield and another for the adjoint wavefield. To initiate the adjoint simulation, one must first prepare an adjoint source based on a selected misfit function and give it as an input to the code at 3-components (vertical, north-south, east-west). This can typically be the cross-correlation of travel-time or waveform difference for a chosen observable. As default, in the SEM code, cross-correlation travel-time adjoint sources can easily be computed within a given window length. Other types of adjoint sources can also be used as demonstrated in Chapter 5. In the construction of the sensitivity kernels, there is a need to access both, forward and the time-reversed adjoint wavefields. However, it is problematic due to high storage and memory requirements. Therefore, in the SEM code, after the forward run, the last frame of the wave propagation is saved and then time-reversed backward simulation is started simultaneously with the adjoint simulation. This makes it possible to access the regular and the adjoint wavefields at a specific time t to construct the sensitivity kernels.

In Figure 2.3.6, sample finite-frequency adjoint kernels computed in isotropic PREM model for P-wave down to 20 s and 40 s is presented for a given source and receiver pair. The simulations were done in isotropic PREM model without an ocean layer on top. Gravity, ellipticity and rotation were not used. The banana-doughnut shape of the kernels can be seen very clearly. The kernels are frequency dependent, as the period increases the kernel gets broaden. In the high-frequency limit, the banana-doughnut

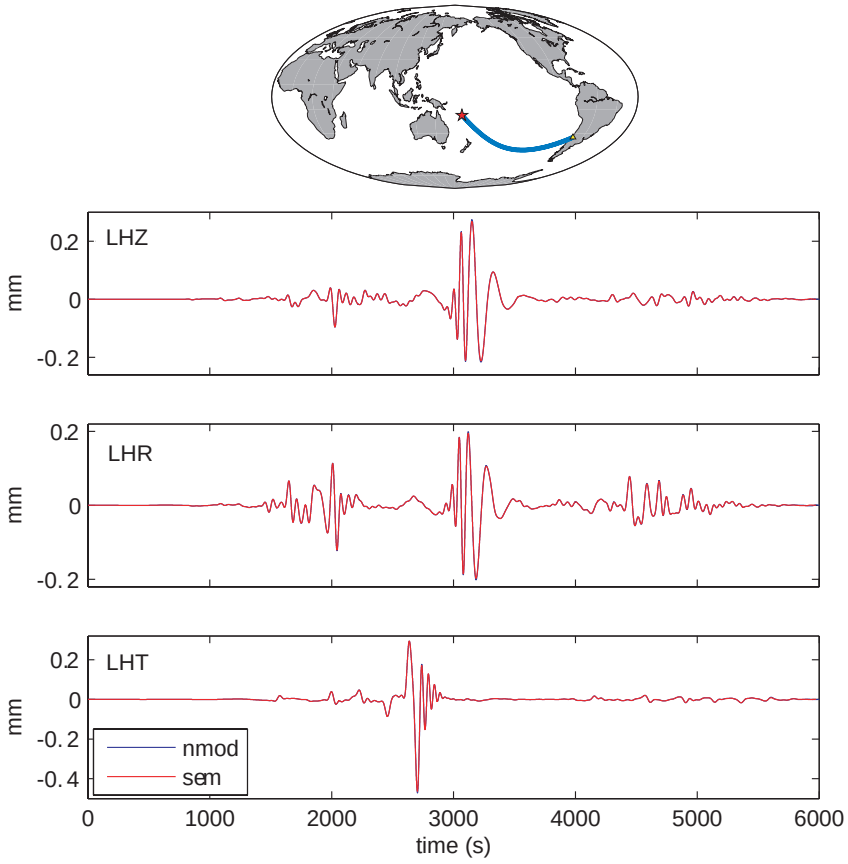

Figure 2.3.5: Comparisons of SEM seismograms with normal mode seismograms computed in transversely isotropic PREM model without ocean layer for the Vanuatu earthquake (1999 November 26, $M_w = 7$) recorded at the station PEL ($\Delta = 105°$). The ray path is shown in the map where the earthquake and the station are denoted by a star and a triangle, respectively. During simulations, gravity and attenuation were added. The seismograms can be found in the SPECFEM3D package.

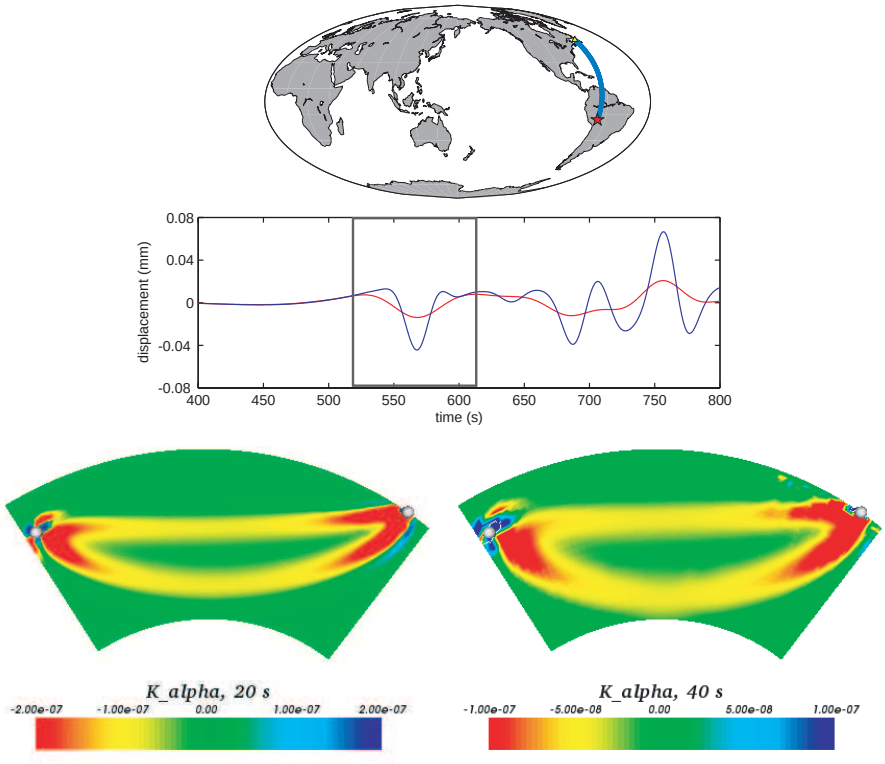

Figure 2.3.6: Cross-sections of the finite-frequency adjoint kernels between source and receiver for P-wave sensitivity for the P-waves filtered down to 20 s (blue seismogram) and 40 s (red seismogram). The source (Bolivia earthquake, 1994 June 9, $M_w = 8.0$) and the receiver (station DRLN, $\Delta = 63°$) are shown by the circles in the kernels. The ray path is shown in the map where the earthquake and the station are denoted by a star and a triangle, respectively.

kernel will collapse onto the ray path.

As shown by Tromp et al. (2005); Liu & Tromp (2008), attenuation can naturally be incorporated in adjoint simulations. However, due to technical difficulties it has not been implemented in adjoint simulations, yet.

2.3.7 Technical Information

SPECFEM3D strictly conforms to the Fortran95 standards. It uses parallel programing based on Message Passing Interface (MPI) to run on PC clusters. The simulations in Chapter 3-Chapter 4 and in Chapter 5 were performed using 54 and 96 processors of a 64 and 128-processor clusters, respectively, at Utrecht University Seismology Group.

Chapter 3

On crustal corrections in surface wave tomography

Summary

Mantle models from surface waves rely on good crustal corrections. We investigated how far ray theoretical and finite-frequency approximations can predict crustal corrections for fundamental mode surface waves. Using a spectral element method, we calculated synthetic seismograms in transversely isotropic PREM and in the 3-D crustal model Crust2.0 on top of PREM, and measured the corresponding time-shifts as a function of period. We then applied phase corrections to the PREM seismograms using ray theory and finite-frequency theory with exact local phase velocity perturbations from Crust2.0 and looked at the residual time-shifts. After crustal corrections, residuals fall within the uncertainty of measured phase velocities for periods longer than 60 s and 80 s for Rayleigh and Love waves, respectively. Rayleigh and Love waves are affected in a highly non-linear way by the crustal type. Oceanic crust affects Love waves stronger, while Rayleigh waves change most in continental crust. As a consequence, we find that the imperfect crustal corrections could have a large impact on our inferences of radial anisotropy. If we want to map anisotropy correctly, we should invert simultaneously for mantle and crust. The latter can only be achieved by using perturbation theory from a good 3-D starting model, or

The content of this chapter was published as: Bozdağ, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophysical Journal International*, **172**, 1066-1082, doi:10.1111/j.1365-246X.2007.03690.x.

implementing full non-linearity from a 1-D starting model.

3.1 Introduction

Much of our knowledge on upper mantle structure is based on surface wave experiments. However, the strong influence of the crust on the propagation of surface waves, even at long periods, makes the inversion for mantle structure very difficult. In most studies, crustal contributions, often as large as those from the mantle, are removed from surface wave data before constructing mantle models. Linear perturbation theory was used in the earlier examples of crustal corrections. Woodhouse & Dziewonski (1984) used two simple models for oceanic and continental crust and calculated local frequency perturbations for crustal corrections based on an ocean-continent function. Nataf et al. (1986) made crustal corrections considering crustal thicknesses, Pn and Sn velocities, ocean depth and topography defining two continental and five oceanic models. However, crustal thickness, which is the dominant factor in these corrections, varies too much for linear perturbation theory to remain valid. Montagner & Jobert (1988) showed that shallow layer variations are indeed strongly non-linear even at periods longer than 100 s. They proposed to model the non-linear effects by calculating exact phase velocities for three different crustal reference models, together with linear perturbations around these three models. This hybrid approach has recently been reviewed by Marone & Romanowicz (2007) and Kustowski et al. (1995). Li & Romanowicz (1996), on the other hand, pointed out that using a prior model for crustal corrections could bias the tomographic images. Instead of using a model for corrections, they allowed Moho depth to change during the tomographic inversion. This approach was also preferred by Shapiro & Ritzwoller (2002).

Today, detailed 3-D crustal models, compiled from a large set of refraction, reflection and geological data, are available (3SMAC, Nataf & Ricard 1996; Crust 5.1, Mooney et al. 1998; Crust2.0, Bassin et al. 2000). The development of computer facilities in recent years has made it possible to calculate the exact eigenfunctions at each point of such 3-D crustal models. Not only thickness variations, but also all changes of the structure of the crust can be taken into account. Once the exact local phase perturbations are calculated for each point of the 3-D crustal model, phase shifts along the (great circle) path can be determined using ray theoretical approximations (e.g., Ritsema et al., 1999; Boschi & Ekström, 2002; Trampert &

Spetzler, 2006). The basis of this path integral approximation originates from the analysis of Woodhouse (1974) who showed that, in the high frequency approximation, the local eigenfunctions at each point are identical to those determined from a spherically symmetric earth defined with the properties beneath that point. The accumulated phase of each individual mode along the path is the integral of the local phase slowness. However, in the presence of rapid structural changes compared to the wavelength of the surface wave, deviations from the great circle, and/or mode coupling and/or the complete breakdown of ray theory are to be expected. Commonly proposed extensions integrate over some influence zone rather than a ray path and use, to first order, 2-D horizontal kernels (e.g., Spetzler et al., 2002; Yoshizawa & Kennett, 2002; Ritzwoller et al., 2002) depending on the approximations made, or better, full 3-D kernels as recently advocated by Zhou et al. (2004).

Local phase perturbations due to the crustal structure are so large that both ray theory and Born theory break down (e.g., Wang & Dahlen, 1995; Zhou et al., 2005), at least for high frequency surface waves. Nevertheless, such corrections are constantly being made and much of our upper mantle knowledge depends on their accuracy. With the availability of the spectral element code (Komatitsch & Tromp, 2002a,b), we are for the first time in a position to quantify this accuracy. Crustal corrections cannot be handled by perturbation theory. This has clearly been demonstrated in the ray theoretical framework (Montagner & Jobert, 1988) and the Born theoretical framework (Zhou et al., 2005). The best remaining option is to integrate exact local phase shifts on the sphere along a ray path (Zhou et al., 2005). We will also investigate the use of 2-D Born kernels (Spetzler et al., 2002). Zhou et al. (2004) explained that the correct 3-D kernels reduce to 2-D kernels if we neglect mode coupling and assume forward scattering. Our reference model (as that of existing studies) is laterally homogeneous which means that forward scattering dominates (Snieder, 1988). Although not perfect, the 2-D kernels have the advantage that they allow to incorporate the vertical non-linearity of the local crustal structure similarly to the ray theoretical tests. Therefore, we decided to investigate their suitability for crustal corrections. We tested the accuracy of crustal corrections using the great circle approximation, exact ray theory and 2-D finite-frequency theory. The aim is not to investigate the validity of ray or finite-frequency theory, which has extensively been discussed in the literature, but rather to understand to what extent we can actually remove the

crustal signal from fundamental mode surface waves using these approximations. We computed synthetic seismograms in transversely isotropic PREM (Dziewonski & Anderson, 1981) and PREM with 3-D crustal model Crust2.0 (Bassin et al., 2000) on top, using the spectral element code (Komatitsch & Tromp, 2002a,b). We measured the time-shifts between these synthetic seismograms as a function of period and compared them with ray theoretical and finite-frequency predictions. It is worth noting that this is the best case scenario where we know the crust. In real seismograms, the actual crust is unknown and therefore our estimates will be on the optimistic side. In the following section, we give a brief outline of the methods and how we measured the time-shifts. In section 3, we analyse the results, and in section 4, we investigate the impact of imperfect corrections on surface wave tomography. Finally, a discussion and the conclusions of our findings are presented in section 5 and 6, respectively.

3.2 Method

3.2.1 Calculation of synthetic seismograms

It is now possible to simulate wave propagation in 3-D earth models using numerical techniques, at either a global or regional scale. The Spectral Element Method (SEM) is most successful for wave simulations in terms of its accuracy and applicability to complex structures. We used the SEM code by Komatitsch & Tromp (2002a,b) for the computation of synthetic seismograms. We selected 11 earthquakes from the global CMT catalogue (Table 3.2.1) and 253 stations distributed worldwide (Figure 3.2.1). We produced two sets of data corresponding to two different velocity models: (i) 1-D transversely isotropic PREM (Dziewonski & Anderson, 1981) with a 3 km-thick ocean layer on top, and (ii) the 3-D crustal model Crust2.0 (Bassin et al., 2000) with $2^\circ \times 2^\circ$ grid resolution on top of PREM including bathymetry, topography and the ocean (hereafter called PREM+Crust2.0). We also included gravity and attenuation in the calculations. The implemention of the models is described in Komatitsch & Tromp (2002a,b) and in the documentation of the code. The length of each seismogram is three hours, which is sufficient to observe the major arc surface waves. Based on the mesh we used in our simulations, the shortest period in our synthetics is approximately 30 s. Vertical and transverse components of sample seismograms for an oceanic and a continental path (A and B in Figure 3.2.1) are presented in Figure 3.2.2). The effect of the 3-D crust on fundamental

Table 3.2.1: List of earthquakes selected from the global CMT catalogue (www.globalcmt.org) to compute the synthetic seismograms.

Event name	Region	Date	Moment magnitude (M_w)	Depth (km)
011604D	Central Mid-Atlantic	16/01/2004	6.2	15
012904B	Southern East Pacific Rise	29/01/2004	6.1	15
020504B	Irian Jaya Region, Indonesia	05/02/2004	7.0	13
022304E	Samoa Islands Region	23/02/2004	6.1	12
031704C	Crete, Greece	17/03/2004	6.0	12
032704G	Xizang	27/03/2004	6.0	12
052804A	Northern and Central Iran	28/05/2004	6.3	22
060904C	Western Indian-Antarctic	09/06/2004	6.4	12
092804G	South of Africa	28/09/2004	6.3	12
100904E	Near Coast of Nicaragua	09/10/2004	6.9	39
110204F	Vancouver Island, Canada	02/11/2004	6.6	19

mode Rayleigh and Love waves can clearly be seen. Oceanic crust increases the phase speed whereas surface waves are delayed by travelling in the 3-D continental crust. Another important feature of the seismograms is that the 3-D crust changes the waveforms, and as expected mostly of Love waves.

3.2.2 Calculation of exact local phase velocity perturbations in Crust2.0

Following Woodhouse (1974), we calculated the local modes by solving the normal mode equations in a radially symmetric earth model corresponding to PREM+Crust2.0 at the desired point. From the local phase velocity in that model, we subtracted the PREM phase velocity to obtain the local phase velocity perturbations $\delta c/c_0$. To make the calculation with SEM and the normal mode code (eosani, courtesy of John Woodhouse) comparable, we needed to sample the Earth models at the same points. For the vertical sampling, we printed out the SEM mesh at the desired latitude and longitude and fed this into eosani. The remaining difference is that SEM interpolates using five Lagrange polynomials of degree 4 and eosani interpolates using cubic splines between two mesh points. We verified that this effect is negligible, and for a 1-D model we achieved quasi exact correspondence between SEM and normal mode seismograms at periods longer than 40 seconds (see similar comparisons in Komatitsch & Tromp (2002a)). Crust2.0 is sampled every $2^\circ \times 2^\circ$, but the SEM mesh corresponds to more

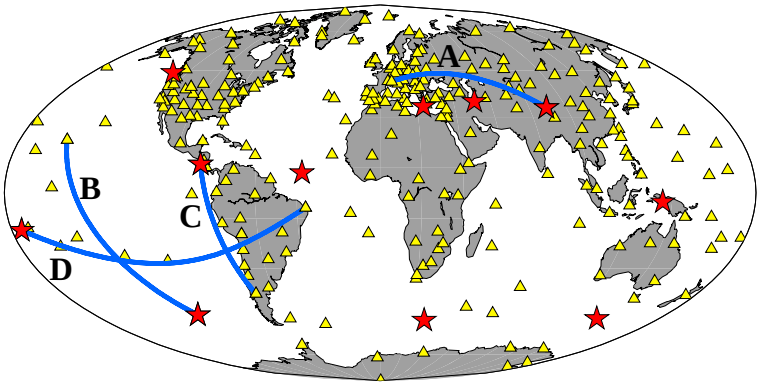

Figure 3.2.1: Distribution of 253 stations (triangles) and 11 earthquakes (stars) used to compute the synthetic seismograms. Blue lines are specific ray paths discussed in the text and used in Figures 3.2.2, 3.3.1 and 3.3.2.

than 1.8 million points at the surface. To avoid sharp edges, SEM uses local lateral smoothing of the crustal model (Komatitsch & Tromp, 2002b). In all correctness, we should calculate the local modes at all 1.8 million surface points, but this would put a tremendous time burden on the calculations (200 days rather than 2 days for all local modes). Instead, we only sampled at the original points given in Crust2.0 and smoothed the local phase velocity afterwards using various expansions. We used a local smoothing similar as in SEM, and a global spherical harmonic smoothing with different degrees of expansion as used in many existing studies.

3.2.3 Calculation of phase shifts

Once the local phase velocity perturbations are obtained, the total phase shift between source and receiver due to the crustal model is calculated by summing the perturbations either along the great circle path, the exact ray path or over the 2−D influence zone.

In the great circle approximation (GCA), phase velocity perturbations are integrated along the great circle path resulting in a phase shift

$$\delta\phi = -\frac{\omega}{c_0}\int_0^{\Delta}\frac{\delta c}{c_0}(\theta,\varphi)d\Delta. \tag{3.2.1}$$

c_0 is the phase velocity in PREM at the angular frequency ω, θ and φ are latitude and longitude, respectively, and Δ is the angular distance between

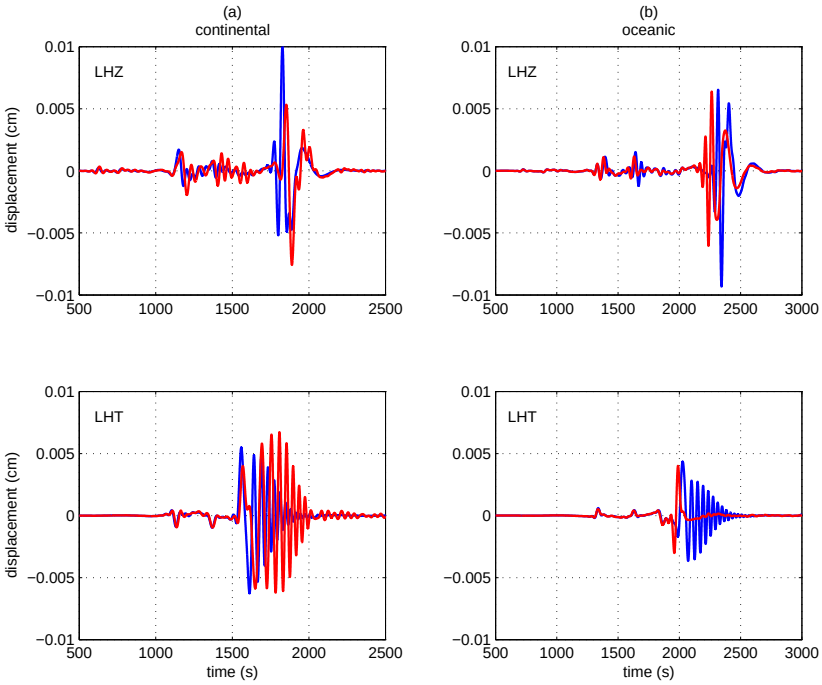

Figure 3.2.2: Vertical (LHZ) and transverse (LHT) components of sample SEM seismograms computed in PREM (blue lines) and PREM+Crust2.0 (red lines) for the continental (A) and oceanic (B) paths shown in Figure 3.2.1. Synthetics were computed for a) the Xizang earthquake (2004 March 27, $M_w = 6.0$) recorded at the station BNI in Europe ($\Delta = 62°$) and b) the Southern East Pacific Rise earthquake (2004 January 29, $M_w = 6.1$) recorded at the station MAUI in the Pacific ($\Delta = 80°$). Seismograms were bandpass filtered with corner frequencies 0.0025 and 0.025 Hz.

source and receiver. The sign is due to our Fourier transform convention.

The ray path, of course, depends on the local phase velocity. To take this non-linearity into account, the local perturbations can be summed along the exact ray path, which we refer to as exact ray theory (ERT). We performed ray tracing using the theory outlined in Woodhouse & Wong (1986). The phase shift can be obtained from

$$\delta\phi = -\frac{\omega}{c_0}\int_{ray}\frac{\delta c}{c_0}(\theta,\varphi)ds \tag{3.2.2}$$

GCA and ERT are formally the same; the only difference comes from the integration of perturbations along either the great circle or the exact ray path.

Both GCA and ERT are high frequency approximations, to take the finite-frequency of surface waves into account, the path integral can be replaced by an integral over a sensitivity kernel which we refer to as finite-frequency theory (FFT). The phase shift is then calculated as

$$\delta\phi = -\int\int_{sphere} K(\theta,\varphi)\frac{\delta c}{c_0}(\theta,\varphi)d\theta d\phi \tag{3.2.3}$$

where $K(\theta,\varphi)$ is the 2−D sensitivity kernel from Spetzler et al. (2002) calculated in PREM, using appropriate frequency averaging, but neglecting source and receiver contributions.

3.2.4 Measuring the time-shifts as a function of frequency

In this work, we are only interested in fundamental mode surface waves. To extract the fundamental mode R1, R2, G1, and G2, we applied a time-variable filter (Cara 1973) to the synthetic seismograms. The Fourier transform of the surface wave seismograms in terms of amplitude (A) and phase (ϕ) can be written as

$$S_P(\omega) = A_P(\omega)exp[\phi_P(\omega)] \tag{3.2.4}$$

and

$$S_{P+C}(\omega) = A_{P+C}(\omega)exp[\phi_{P+C}(\omega)] \tag{3.2.5}$$

where subscripts P and $P + C$ denote seismograms in PREM alone and PREM+Crust2.0, respectively. To determine the phase shift between these two seismograms, we cross-correlated $S_{P+C}(\omega)$ with $S_P(\omega)$ and measured the phase of the cross-correlogram as a function of frequency. After unwrapping the phase, the time-shifts between PREM+Crust2.0 and PREM seismograms are obtained by dividing the phase of the cross-correlogram by the angular frequency

$$\delta t(\omega) = \frac{\phi_{P+C}(\omega) - \phi_P(\omega)}{\omega} \tag{3.2.6}$$

In a similar way, the phase shift between seismograms PREM+Crust2.0 and PREM with crustal correction using ray theory or finite-frequency theory can be obtained. At all frequencies, the corrected PREM seismograms can be written as;

$$S_P^{cc}(\omega) = A_P(\omega) exp[\phi_P(\omega) + \delta\phi(\omega)] \tag{3.2.7}$$

where the calculation of $\delta\phi(\omega)$ is outlined in Section 3.2.3. The time-shift from the cross-correlation is now;

$$\delta t^{cc}(\omega) = \frac{\phi_{P+C}(\omega) - \phi_P(\omega) - \delta\phi(\omega)}{\omega}. \tag{3.2.8}$$

To avoid amplitude problems, we multiplied the amplitude of $S_P(\omega)$ and $S_P^{cc}(\omega)$ with $A_{P+C}(\omega)/A_P(\omega)$ before the cross-correlation to equalize all amplitudes.

3.3 Results

Comparing time-shifts before (δt) and after (δt^{cc}) correction, gives us a good idea on the effectiveness of the applied correction. Note that if the correction is perfect, δt^{cc} is close to zero. Time-shifts are a more convenient measure than relative phase shifts because they allow a direct interpretation in terms of cycles when compared to the period. Before analysing time-shifts, we excluded source-receiver distances $\Delta < 20°$, $160° < \Delta < 180°$, $180° < \Delta < 200°$ and $\Delta > 340°$ to avoid minor and major arc interferences which influence the phase. We also excluded the paths within 20° of the nodal plane of the radiation pattern where the phase can be severely distorted.

3.3.1 Some examples of time-shifts

Let us first concentrate on several individual paths shown in Figure 3.2.1. For the continental path (A), the crust delays the surface waves considerably (Figures 3.3.1 and 3.3.2), up to two cycles for short period Love waves. The bulk of the crustal signal is corrected, but it is not perfect. The different lateral smoothing strategies give consistent results with at most 6 s difference for short period Love waves. For the oceanic path (B), the observations are pretty similar. Notable is a large overprediction of the crustal effect for short period oceanic Love waves, much larger than for short period continental Rayleigh waves. This is most likely due to wavefront smoothing effects (Wang & Dahlen, 1995) where ray theoretical predictions tend to overestimate absolute phase anomalies. The difference between the different smoothing strategies is again of the same order and it is not clear which kernel performs best. Next, let us look at a path along an ocean-continent boundary (C). The most striking observation is that now FFT with local smoothing performs significantly better for both Rayleigh and Love waves. This is exactly what you would expect, the ray is either in the ocean or the continent, depending on the path, but the finite-frequency wave senses both which is of course much better modelled by the FFT kernel. The FFT kernel together with spherical harmonic smoothing performs well at long periods only. Short period waves sense the small scale variations in crustal gradients which are smoothed by the spherical harmonic expansion. Finally, for a path which is perpendicular to the ocean-continent boundary (D) the FFT advantage is not so clear any more. The nature of the kernel or the smoothing strategy matter little and all residuals are again within a few seconds of each other, all showing overcorrections.

We cannot look at all the paths individually. In the following, we will examine the results statistically.

3.3.2 Time-shifts as a function of distance

To examine the influence of the source-receiver distance on the corrections, we grouped the time-shifts obtained from both minor and major arc measurements into six distance bins containing approximately the same amount of data. The histograms in Figure 3.3.3 show the time-shifts for 150 s Rayleigh waves. Grey histograms correspond to the time-shift (δt) (eq. 3.2.6) between PREM+Crust2.0 and PREM seismograms. Negative time-shifts correspond to cases where the crust advances the phase (i.e.

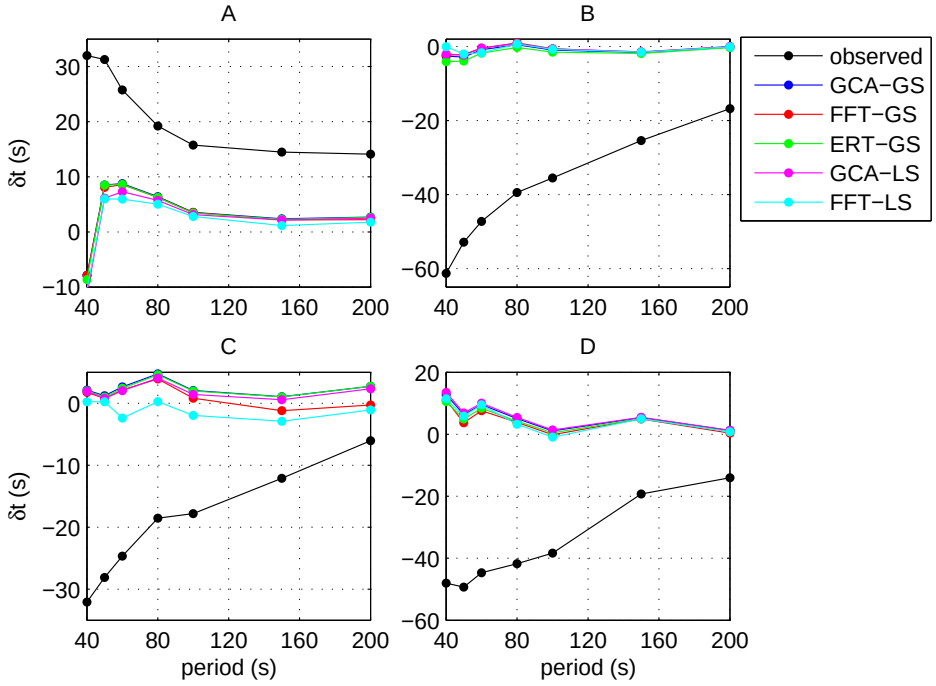

Figure 3.3.1: Observed (δt) and residual (δt^{cc}) time-shifts as a function of period for Rayleigh waves from; (A) the continental path, (B) the oceanic path, (C) the path along the ocean-continent boundary, (D) the path perpendicular to the ocean-continent boundary shown in Figure 3.2.1. LS and GS denote local smoothing and global smoothing (spherical harmonic degree 40), respectively.

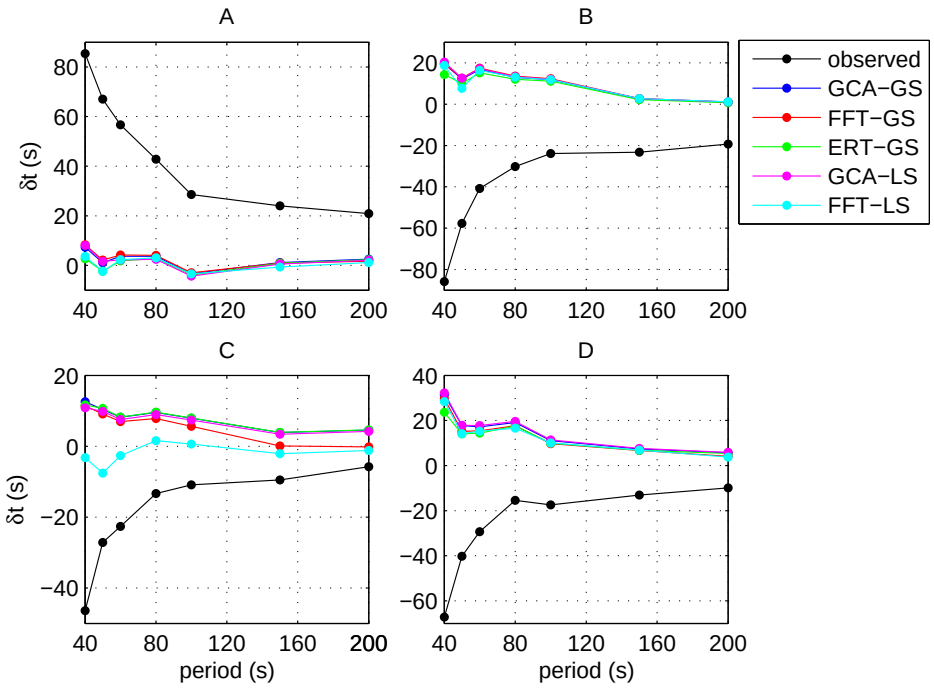

Figure 3.3.2: Same as Figure 3.3.1 but for Love waves.

thinner oceanic crust). Positive time-shifts correspond to paths with an average thicker crust than PREM. Red histograms correspond to residual time-shifts (δt^{cc}) (eq. 3.2.8) after correction. The corrections, calculated using GCA together with local smoothing, are not perfect, but remove a large part of crustal signal. We notice a slight shift of the residual histograms towards positive times as the distance increases. This is due to wavefront smoothing effects (Wang & Dahlen, 1995). Ray theoretical predictions tend to give extreme values for slow and fast paths. For longer distances, the oceanic paths dominate more and more and hence a shift of the residuals to the right. At short distances, there are about as many continental and oceanic paths and the residuals are nicely centred around zero. A similar picture is seen for Love waves at 150 s (Figure 3.3.4), although the corrections work slightly less well, because Love waves are more effected by the crust. At 80 s (Figures 3.3.5 and 3.3.6), the crustal effects are stronger, even bigger than a cycle for the longest paths. Still, the corrections manage

to bring the residuals down to less than half a cycle. The average positive time-shift for the residuals becomes stronger. At 40 seconds (Figures 3.3.7 and 3.3.8), the surface waves see the full effect of the crust and time-shifts up to five cycles are seen. Although the corrections account for much, for many paths, they do not manage to bring the residuals down to a fraction a cycle. Because the crustal effect is so strong on these short period surface waves, the wavefront smoothing effect is also most noticeable.

So far, the discussion remained rather qualitative. In order to quantify how good or bad the corrections are, we propose to compare the residual time-shifts to actual uncertainties in phase velocity measurements used in surface wave tomography. We chose uncertainties from the measurements published in the latest compilation of Trampert & Woodhouse (2001) who performed cluster analysis on similar paths. Table 3.3.1 shows the average standard deviations for $\delta c/c_0$ from this analysis. Figure 3.3.9 shows an example of residuals as a function of distance compared to the average measurement uncertainty $\sigma(\delta t) = \sigma(\delta c/c_0)\frac{x}{c_0}$, where x is the distance and c_0 is the reference phase velocity. The measurement uncertainties are assumed to be Gaussian distributed therefore the probability that the actual uncertainty is bigger than one standard deviation is 0.31731. From Figure 3.3.9, we can evaluate the probability p that our residuals are bigger than the same standard deviation. Bayesian statistics then tell us that the residuals are $p/0.31731$ times more likely than the actual measurement uncertainties to be bigger than one standard deviation. A ratio larger than 1 means that the residuals contain significant signal beyond the measurement uncertainties and are likely to bias models in tomographic inversions. Figures 3.3.10 and 3.3.11 show this ratio for different periods for Love and Rayleigh waves. One can repeat this analysis for two standard deviations. The results would look better because we would miss part of the offset of the residuals due to wavefront smoothing effects. However, this offset can potentially bias the tomographic models, and therefore we prefer to judge the residuals using one standard deviation of the measurement uncertainties.

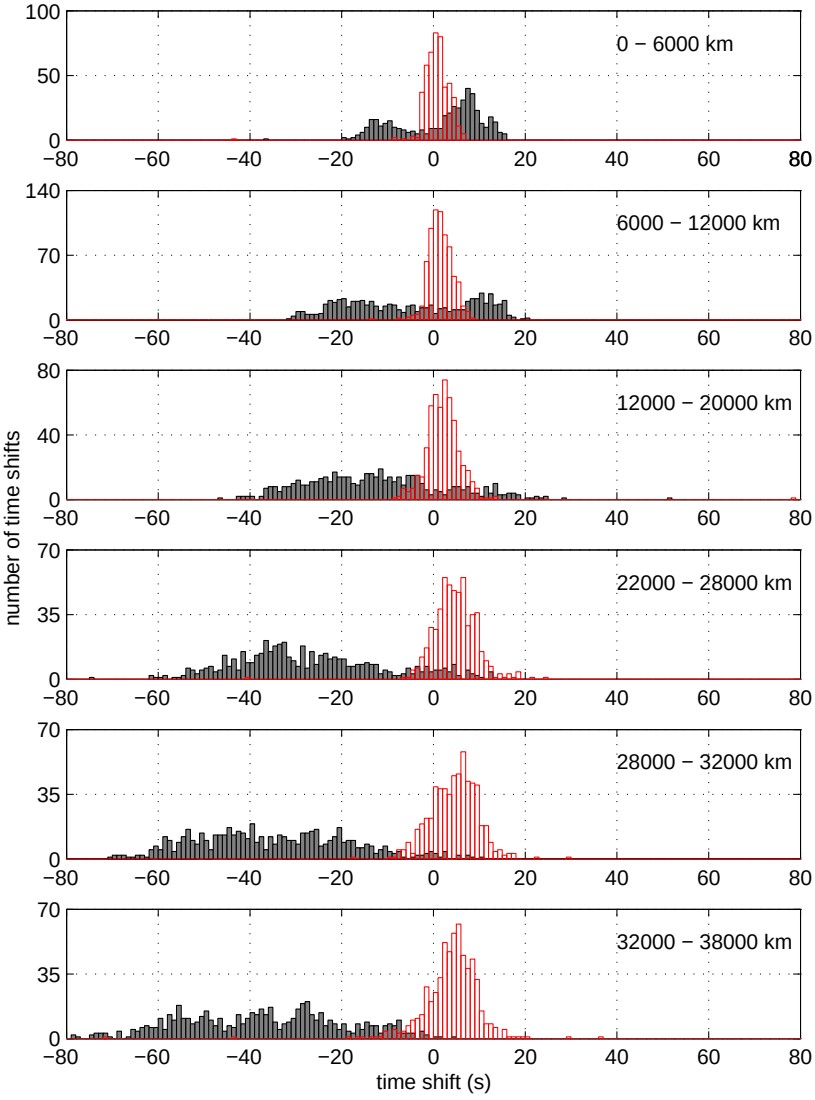

Figure 3.3.3: Histograms of time-shifts for 150 s Rayleigh waves from GCA and local smoothing as a function of distance. Grey and red bars are the observed (δt) and residual (δt^{cc}) time-shifts, respectively.

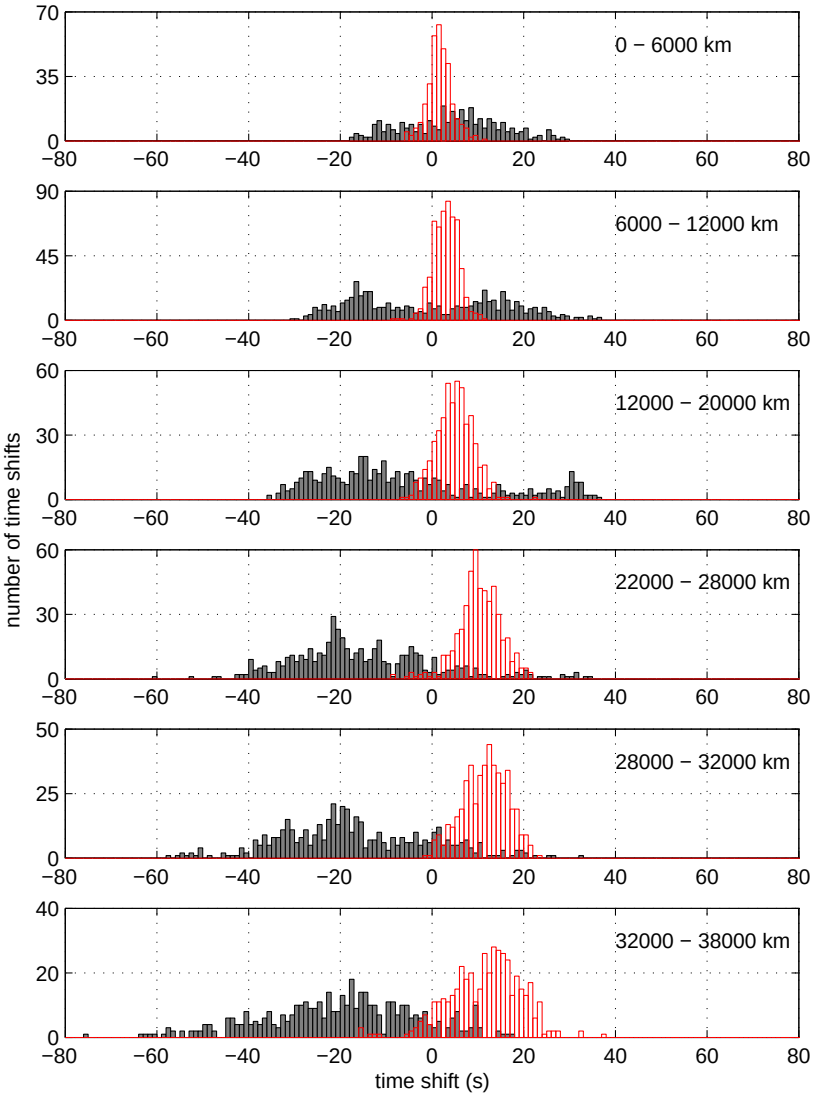

Figure 3.3.4: Same as Figure 3.3.3, but for 150 s Love waves.

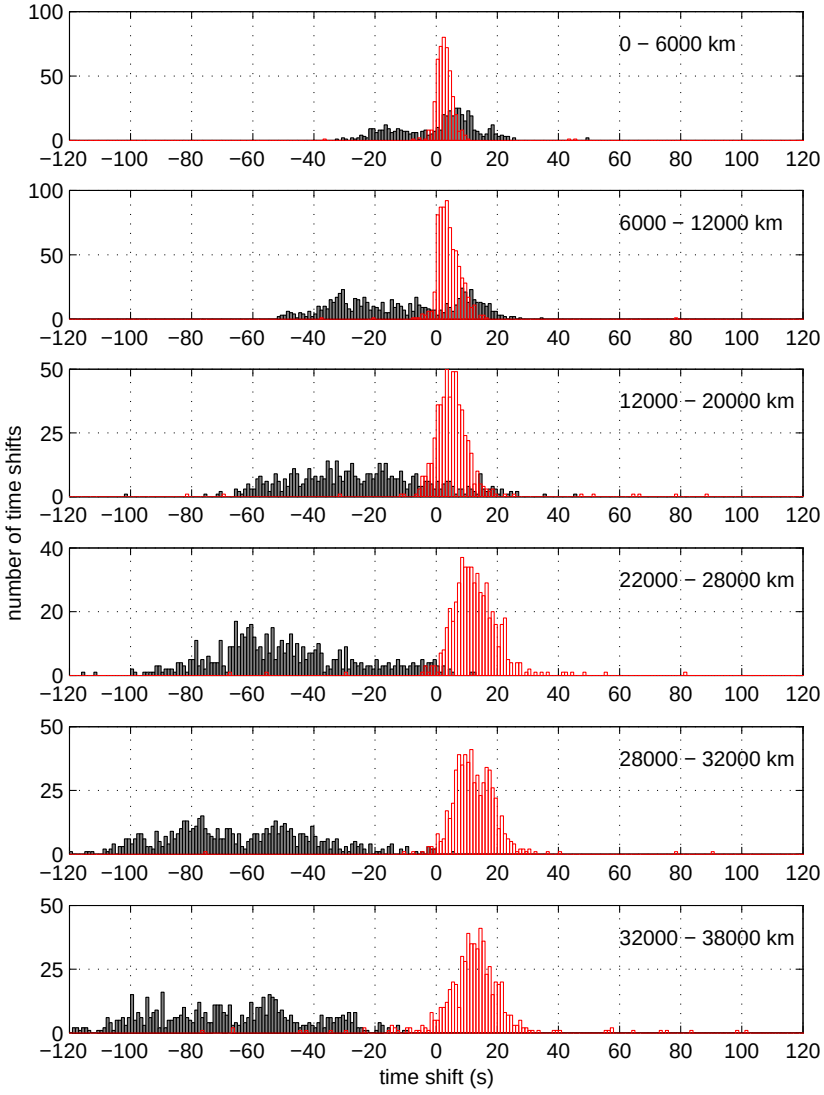

Figure 3.3.5: Same as Figure 3.3.3, but for 80 s Rayleigh waves.

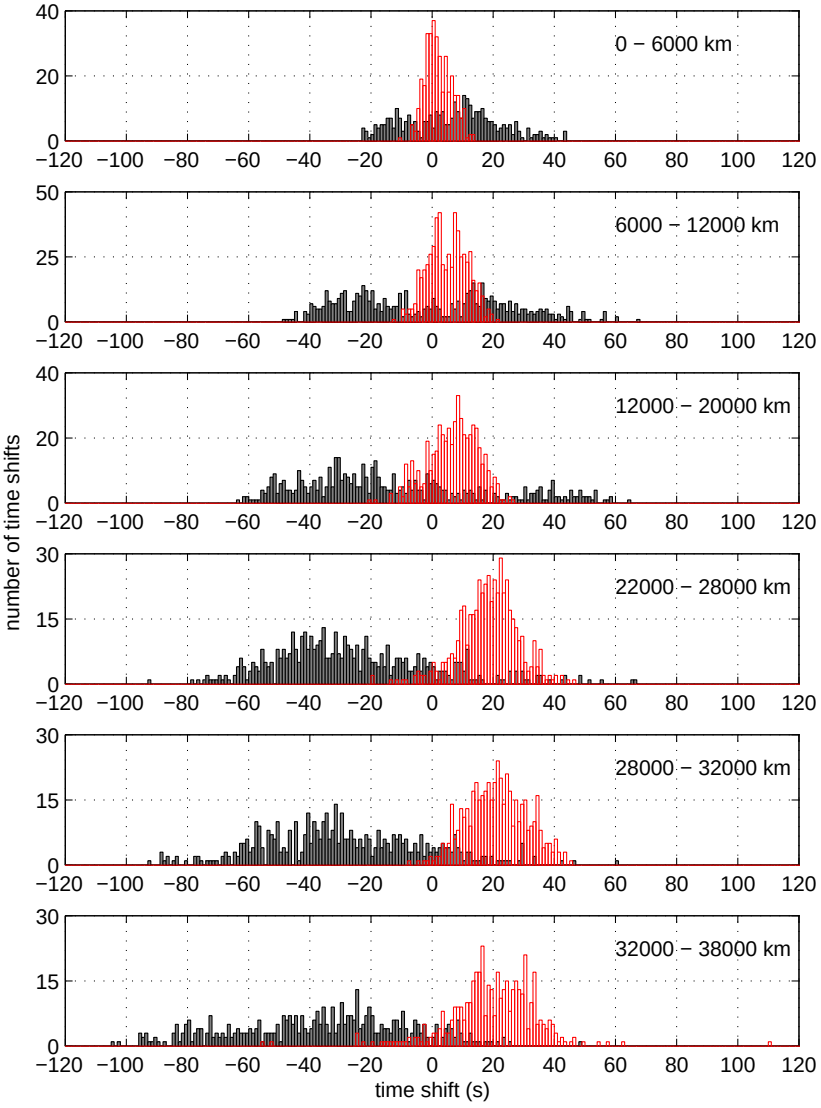

Figure 3.3.6: Same as Figure 3.3.3, but for 80 s Love waves.

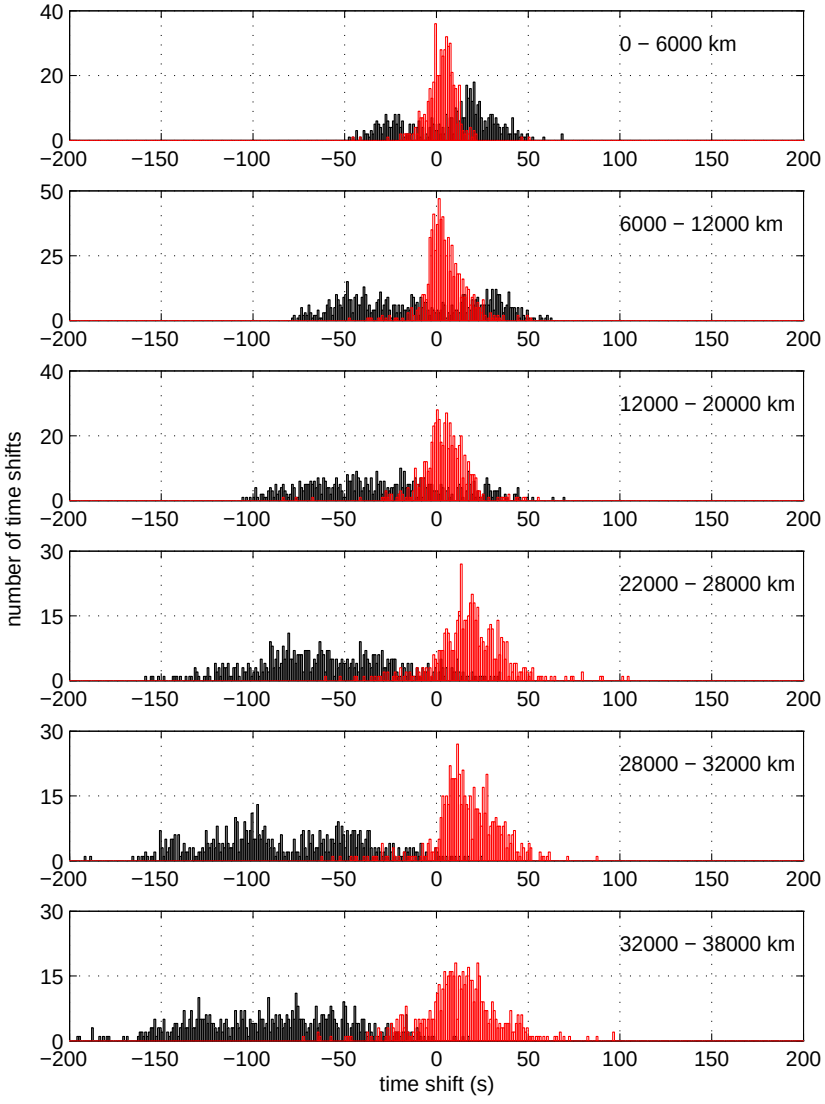

Figure 3.3.7: Same as Figure 3.3.3, but for 40 s Rayleigh waves.

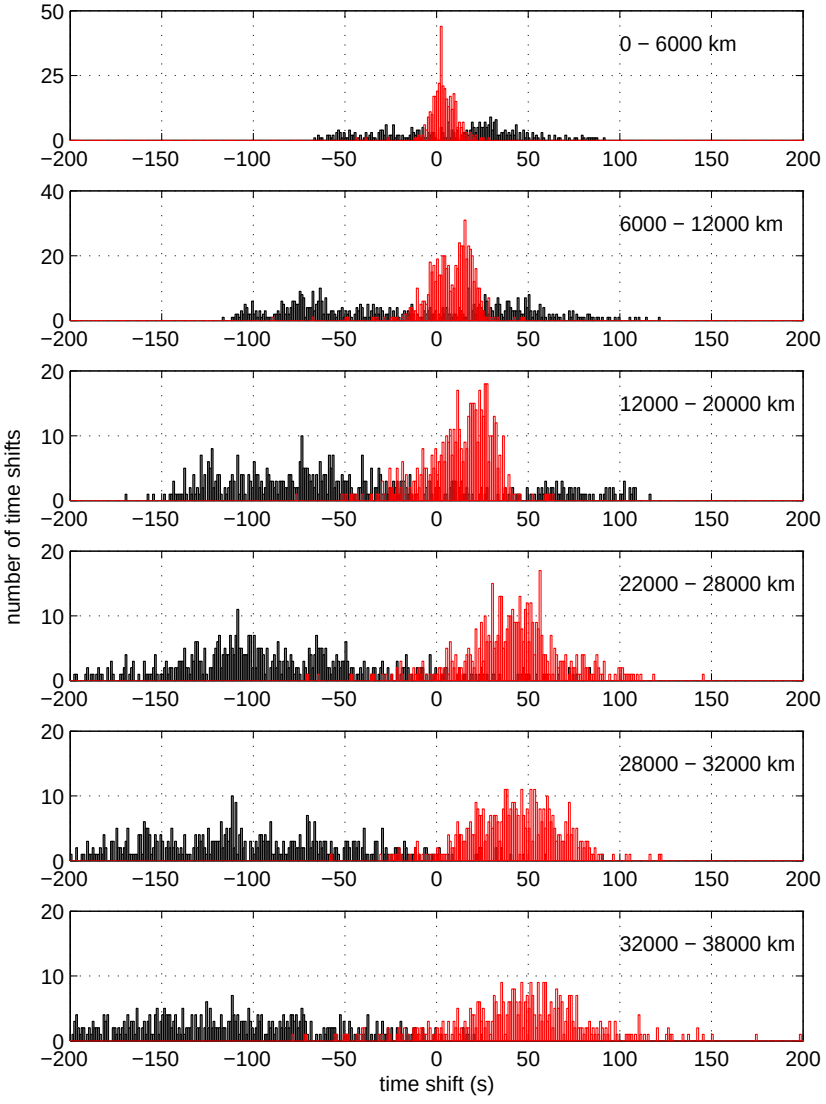

Figure 3.3.8: Same as Figure 3.3.3, but for 40 s Love waves.

Table 3.3.1: Average uncertainties for measurements of $\delta c/c_0$ (Trampert & Woodhouse, 2001) from cluster analysis.

Period	Average uncertainty for Rayleigh	Average uncertainty for Love
150 s	3.92×10^{-3}	5.64×10^{-3}
100 s	4.02×10^{-3}	4.50×10^{-3}
80 s	3.86×10^{-3}	4.44×10^{-3}
60 s	3.65×10^{-3}	4.29×10^{-3}
40 s	3.24×10^{-3}	4.22×10^{-3}

As one would expect, GCA crustal corrections work well for long period Love and Rayleigh waves which only experience limited effects from the crustal model. The residual time-shifts almost completely fall within one standard deviation of the measurement uncertainties and hence should be absorbed by the uncertainties without any trace in a mantle model if there are no systematics in the residuals. In general, the corrections are better for Rayleigh than Love waves which of course reflect the fact that Love waves sense the crust more than Rayleigh waves. Our statistics show that the corrections work reasonably well for Rayleigh waves from 60 s onwards, although the cut-off is somewhat distance dependent. For Love waves this situation is clearer and corrections work well from 80 s onwards based on our Bayesian criterion. Our analysis indicates that the corrections are good at longer periods, but there is the remaining issue that the residual histograms are not centred around zero. This bias could introduce artefacts into the tomographic models and only a depth inversion of the residuals (see Section 3.4) will reveal its importance.

3.3.3 The effect of different crustal types

We also investigated whether different crustal types might influence the effectiveness of the crustal corrections. We therefore regrouped the time-shifts according to the percentage of continental crust along the ray path. Corrections for Rayleigh and Love waves at periods longer than 80 s are similar although the results are slightly better for Rayleigh waves. The best results are obtained from purely oceanic paths where all the observed time-shifts are almost completely corrected to around zero. As the percentage of continental crust is increased, the residual time-shifts become larger.

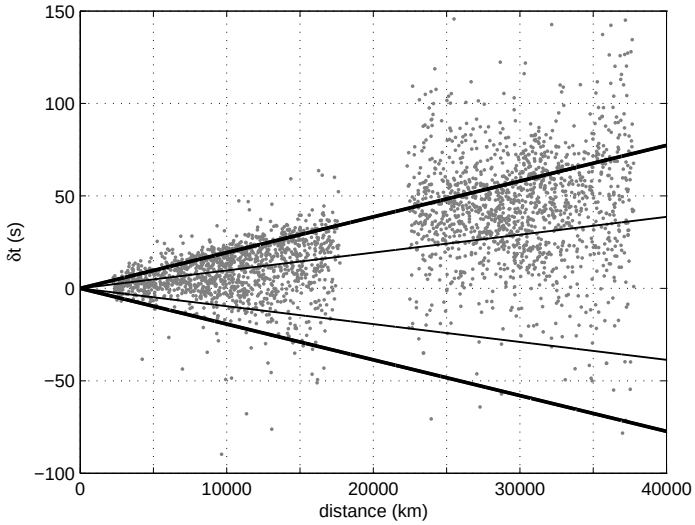

Figure 3.3.9: Scatter plot of residual time-shifts (δt^{cc}) as a function of distance for 40 s Love waves using GCA together with local smoothing. The average uncertainty determined from the actual phase velocity measurements of Trampert & Woodhouse (2001) is shown by thin (one standard deviation) and thick (two standard deviation) lines.

The largest time-shifts are obtained for the paths having ocean-continent transitions, but the main contribution to larger residual time-shifts comes from major arcs.

At short periods, we can see more clearly that Rayleigh and Love waves are affected in different ways by continental and oceanic crusts. In Figure 3.3.12, we present results for purely continental and purely oceanic paths for a common distance bin for 40 s Rayleigh and Love waves. Observed time-shifts up to 60 s in oceanic crust can effectively be corrected for 40 s Rayleigh waves whereas residual time-shifts are on average over-corrected by 15 s for Love waves. The latter is likely the wavefront smoothing effect identified by Wang & Dahlen (1995), where thin oceanic crust affects finite-frequency Love waves in a highly non-linear way. Although oceanic crust is mostly uniform, detailed comparisons of 1−D normal mode seismograms with an average oceanic crust and SEM seismograms in PREM+Crust2.0 have shown that mid-oceanic ridges and ocean islands, where the stations are located, have a considerable effect on the waveforms. Figure 3.3.10

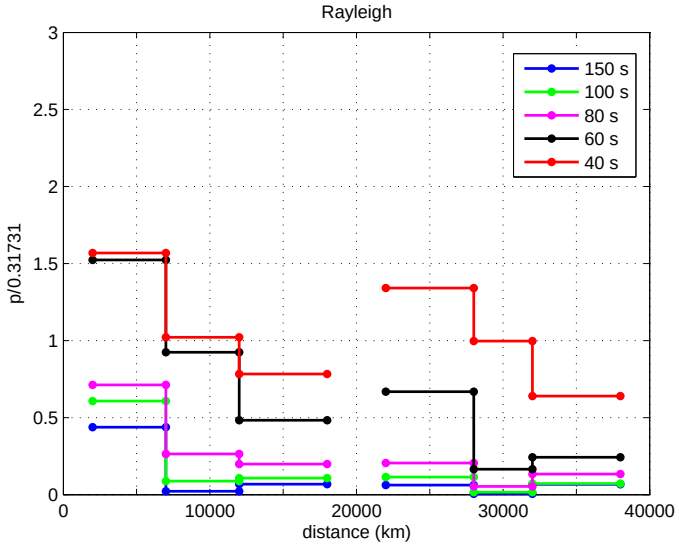

Figure 3.3.10: Likelihood factor as a function of distance for Rayleigh waves indicating how many times the residual time-shifts (δt^{cc}) are bigger than the actual phase measurement uncertainties.

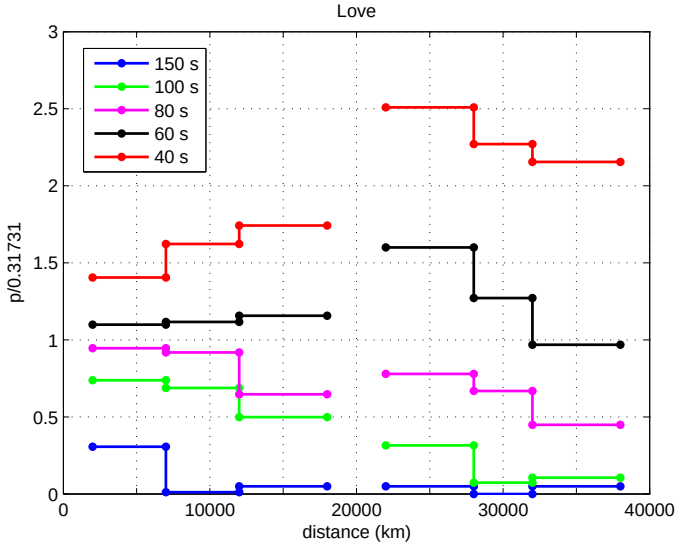

Figure 3.3.11: Same as Figure 3.3.10, but for Love waves.

shows this dramatic change of waveforms between 1−D and 3−D crusts. Corrections for purely continental paths are slightly better for 40 s Love waves than Rayleigh waves. The non-linear wavefront smoothing now affects Rayleigh waveforms more strongly. 40 s Rayleigh waves have higher sensitivity around the Moho depth of continental crust, whereas 40 s Love waves sense much shallower variations more strongly. This non-linearity of different crustal types elegantly explains our results. Non-linearity is stronger for Love waves than Rayleigh waves and linearized corrections work better on Rayleigh waves. For longer paths, oceanic crust is dominant and hence Love waves quickly deteriorate as a function of distance. For the short paths, continental crust is dominant, therefore $40-60$ s Rayleigh waves are worse affected than Love waves (Figures 3.3.10 and 3.3.11).

3.3.4 Comparison of different methods

So far we presented results for GCA using the same local smoothing of the crustal model as in the SEM calculations. It is interesting to know if crustal corrections based on global smoothing change the results. Most models expanded in spherical harmonics use crustal corrections expanded on spherical harmonics as well (e.g., Ritsema et al., 1999). Therefore, we checked the effect of different smoothing techniques (local smoothing and global smoothing by spherical harmonic expansion) on GCA. Up to degree 40, at 40 and 150 s, different smoothing strategies give statistically similar results (Figure 3.3.13). We checked that this is true for all other periods as well. We also compared local smoothing with higher degrees of spherical harmonic expansion ($l = 60$, l= 80) and the results remained the same.

It is further important to know if extensions to GCA are worth considering. We therefore implemented two commonly used methods using 2−D finite-frequency kernels and exact ray tracing. Comparisons between GCA and FFT, using local smoothing of the crustal model, show that residual time-shifts for 40 s and 150 s Rayleigh and Love waves are statistically similar (Figure 3.4.1). We observed that FFT shows a clear improvement, especially for Love waves that are most sensitive to the crustal heterogeneities, for the paths along the ocean-continent boundaries (see Figures 3.3.1 and 3.3.2). However, from a statistical point of view, given a realistic path coverage, the advantage of FFT is lost on average.

Exact ray tracing is difficult in complicated structures such as Crust2.0. Multipathing is so severe that it is difficult to find the correct ray path. To make ray tracing practical, we used a smooth version of the crustal model

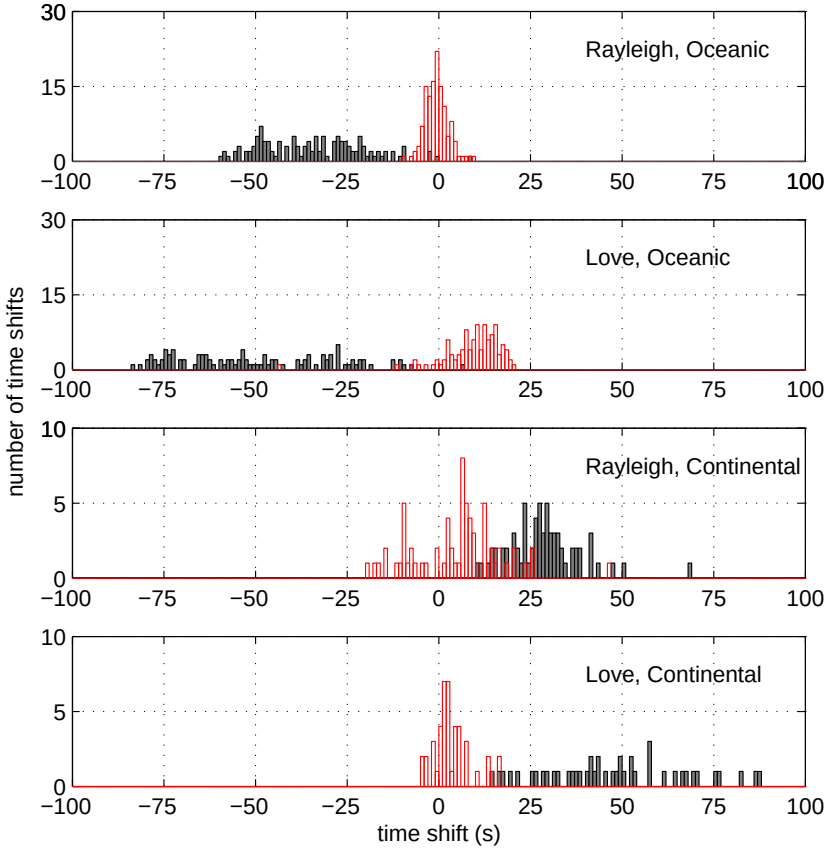

Figure 3.3.12: Histograms of time-shifts (GCA with local smoothing of the crust) for 40 s Rayleigh and Love waves for purely oceanic and purely continental paths. Grey and red bars are observed (δt) and residual (δt^{cc}) time-shifts, respectively.

and compared ERT to GCA in the crustal models expanded in spherical harmonics up to degree 40 and for minor arcs only. Scatter plots comparing GCA and ERT at 40 s and 150 s Rayleigh and Love waves (Figure 3.4.2) show the known Fermat bias (more points above the diagonal), which means that phase shifts along the exact ray path are always smaller than phase shifts along the great circle path (e.g., Dahlen & Tromp, 1998). Note that we plotted time-shifts between PREM+Crust2.0 and PREM seismograms rather than direct measurements of phase. ERT clearly improves crustal corrections for short period Love waves, which are more affected by wavefront smoothing effects. Because ERT predicts smaller phase shifts, the over-correction is less severe. However, this improvement (< 10 s) is modest compared to the residual time-shifts for 40 s Love waves (Figure 3.3.8) and does not really justify to try and handle the complexity of ray tracing in a rough crustal model.

3.4 Consequences for surface wave tomography

We found that residuals of crustal corrections fall within measurement uncertainties for longer period surface waves (longer than 60 s for Rayleigh and longer than 80 s for Love waves). This statistical analysis does not inform on systematics within these residuals apart from the shift of the histograms due to wavefront smoothing. To investigate how far the imperfect crustal corrections bias tomographic models, we simply inverted the residuals δt^{cc} (eq. 3.2.8) for a shear velocity model. We assumed that the GCA residuals are path averages and inverted 1731 minor and 1731 major arc data for Rayleigh and 1420 minor and 1420 major arc data for Love, sampled at $40, 60, 80, 100$ and 150 s, as described in Trampert & Spetzler (2006). The model is parameterized laterally into spherical harmonics up to degree 20 and vertically into 9 splines to a depth of 600 km and regularized using first derivative smoothing. We further assume that Love wave data produce an SH model and Rayleigh wave data an SV model (e.g., Ekström & Dziewonski, 1998).

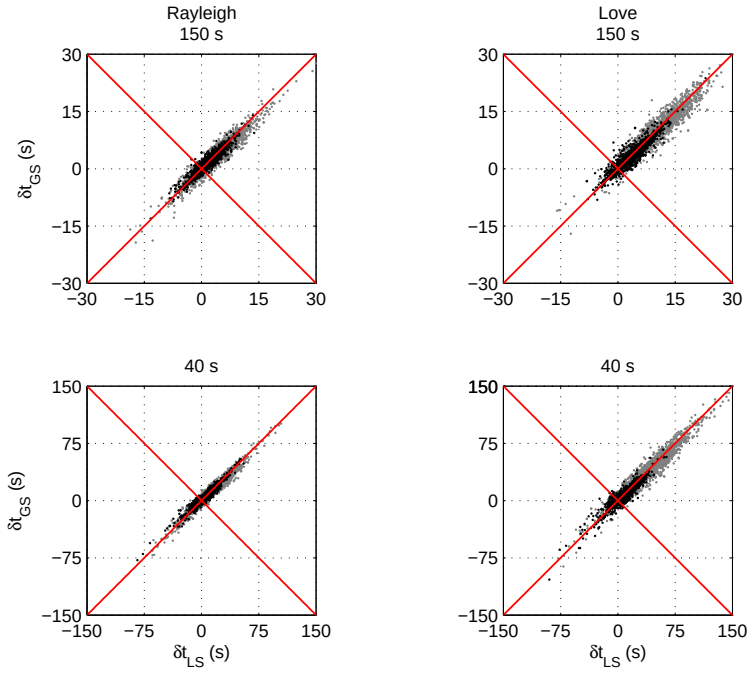

Figure 3.3.13: Scatter plots of residual time-shifts (δt^{cc}) using GCA with local smoothing (LS) and global spherical harmonic smoothing up to degree of 40 (GS) of the crustal correction at 150 s (upper plots) and 40 s (lower plots) Rayleigh and Love waves. Black and grey dots are the time-shifts from minor and major arc measurements.

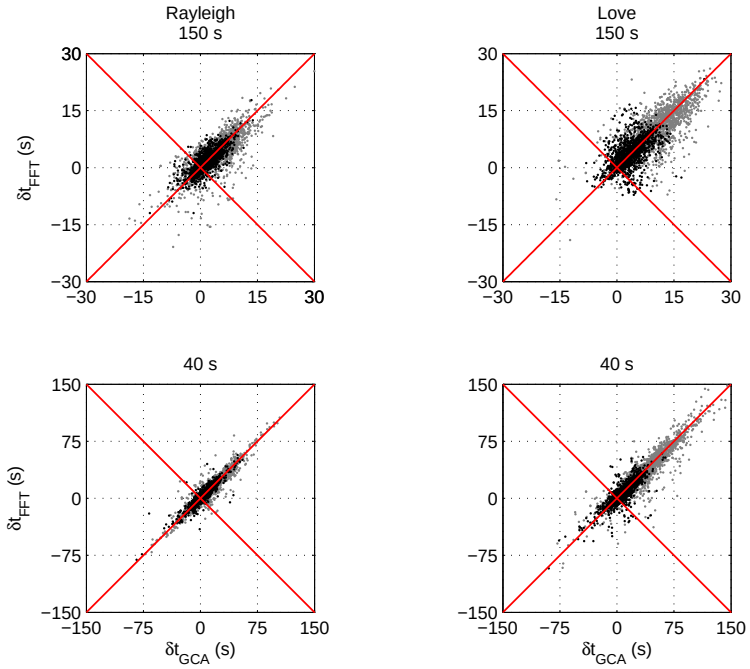

Figure 3.4.1: Scatter plots of residual time-shifts (δt^{cc}) for GCA and FFT using local smoothing at 150 s (upper plots) and 40 s (lower plots) Rayleigh and Love waves. Black and grey dots are the time-shifts from minor and major arc measurements.

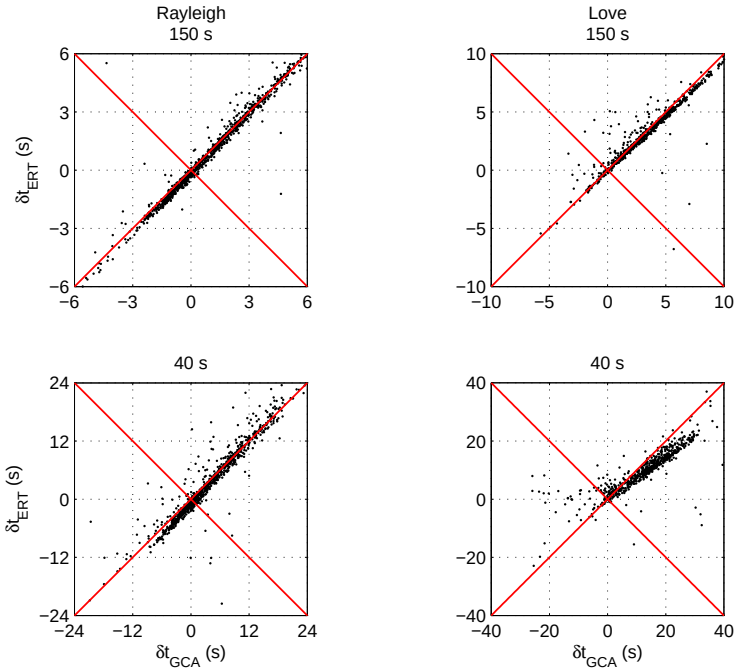

Figure 3.4.2: Scatter plots of residual time-shifts (δt^{cc}) for GCA and FFT using spherical harmonic smoothing up to degree 40 for 150 s (upper plots) and 40 s (lower plots) minor arc Rayleigh and Love waves.

Figure 3.5.1 shows the rms-amplitude of the models as a function of depth. As expected, the SH model has about $1.5 - 2$ times higher amplitude than the SV model in the first 400 km depth, reflecting the fact that Rayleigh wave crustal corrections are generally better than those for Love waves. For both, Love and Rayleigh waves, the amplitudes are strongest just below the crust. The amplitudes decay quickly and are reduced by about 70% at 100 km. They then continue to decay slowly, but remain significant to at least 400 km depth. The absolute amplitudes are difficult to determine because they strongly depend on regularization, but we can compare the amplitudes to a similarly damped more complete SV model using 1.5 million fundamental mode and overtone data (Figure 3, middle panel, in Trampert & Spetzler (2006)). We see that the SV model from the crustal residuals is about $2 - 4$ times smaller. This is reassuring, but equally important, the pattern of heterogeneity has no correlation to that

seen in published models. The SH model from crustal residuals, however, has stronger amplitudes and is only $1-2$ times smaller than the model of Trampert & Spetzler (2006). This is very important for inferences on seismic anisotropy. Following Ekström & Dziewonski (1998), we plot lateral variations of radial anisotropy ($\delta lnV_{SV} - \delta lnV_{SH}$, Figure 3.5.2). The most striking features are that the shallow anisotropy is strong (see also Figure 3.5.1) and overall changes sign between 50 and 150 km depth (we find a global correlation of -0.53 between these two maps). Such a sign change in anisotropy has been reported before (e.g., Ekström & Dziewonski, 1998) and hence could be due to imperfect crustal corrections. We further wondered what would happen if we only included frequencies for which the corrections are acceptable. Inspired by Figures 3.3.10 and 3.3.11 we only inverted frequencies from 60 and 80 s onwards for Rayleigh and Love, respectively. We saw that the bias in anisotropy is reduced by a factor of 1.5 in the first 100 km and no overall sign change occurs. Imperfect crustal corrections can thus bias surface wave tomography models (isotropic and anisotropic) throughout the upper mantle with an rms-amplitude shown in Figure 3.5.1. If high frequency (40 s) surface waves are used, the shallow apparent anisotropy can be strong, and have a sign change around 100 km depth, only because of imperfect crustal corrections.

3.5 Discussions

It is worrying, but not surprising that the biggest effect of imperfect crustal corrections is on inferences of radial anisotropy. Indeed, Levshin & Ratnikova (1984) showed that lateral variations in Moho thickness trade-off with radial anisotropy. This is easily understandable since the crustal corrections on average affect Love waves more strongly than Rayleigh waves and radial anisotropy is a measure of the difference. Since we can see the anisotropic bias even if we only use longer periods, the only option seems to invert simultaneously for mantle and crust. The experiments of Meier et al. (2007) show that the data we used are mainly sensitive to Moho thickness and to a much lesser extend to the crustal velocity structure. This suggests that it is sufficient to invert for mantle structure and Moho thickness, where care has to be taken to treat the strong non-linearity of the latter adequately.

Kennett (1995) looked at the validity of the path integral approximation in waveform fitting and concluded that a crustal thickness change up to

10 km can be accommodated using fundamental mode Love and Rayleigh waves of periods longer than 67 s. Crust2.0, of course, has bigger variations than those in Kennett (1995) and we therefore find that on average Love waves above 80 s and Rayleigh waves above 60 s are well corrected using a path integral approach. Kennett & Yoshizawa (2002) argue that the path integral approach is less restrictive for dispersion (phase velocity) models than for 3−D wave speed models. We found indeed that shear wave models are about 1.3 times smaller using the observed time-shifts (eq. 3.2.6) corrected from the crustal model at each frequency rather than the residual time-shifts which result from a waveform correction.

Our finite-frequency kernels do not guarantee on average a broader range of validity for waveform corrections. This is surprising, because we should be in a regime where structural variations are close to the wavelength of the surface waves. A possible explanation is that we used approximate 2−D kernels which neglect mode coupling and assume single forward scattering only. Experiments with 3−D kernels (Zhou et al., 2005) showed that they were also not effective to account for crustal structure using a 1−D reference model. This is due to the strong non-linearity identified by Montagner & Jobert (1988). 3−D kernels should work fine when calculated in a 3−D reference model close to the actual model, but requires measurements from 3−D models as well. Techniques for such calculations are starting to become available (Tromp et al., 2005), but still constitute a considerable computational challenge.

Exact ray theory gives slightly better results compared to those obtained from the great circle approximation. In general, ray tracing is a very challenging problem particularly for long ray paths in complex 3−D crustal models. Multipathing is the most important issue in ERT (see Ferreira, 2005). One of the ways to avoid multipathing is to perform full ray tracing, in which all possible ray paths are searched for, and then select the minimum-phase path. Other possibilities are to follow the evolution of the full wavefront (Vidale, 1988; Rawlinson & Sambridge, 2004). All these techniques are very time consuming and rather inefficient for crustal corrections on a global scale compared to the potential gain.

In our measurements, we only considered phase shifts due to propagation. Ferreira & Woodhouse (2006) pointed out that for Rayleigh waves the phase at the source strongly depends on azimuth, source depth and period. The source radiation pattern of Love waves does not show such a strong dependence. They found that phase shifts due to the source can

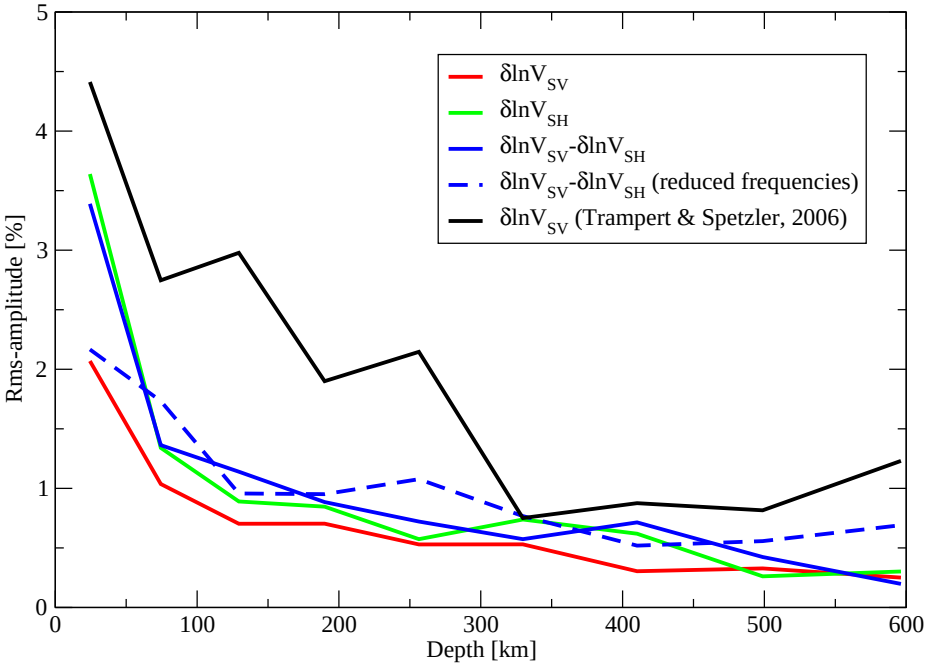

Figure 3.5.1: Rms-amplitude as a function of depth of the S-velocity models inverted from residual time-shifts of Rayleigh waves (red) and Love waves (green) in this study and a similarly damped Rayleigh wave model (Trampert & Spetzler, 2006) (black line). The rms-amplitude of radial anisotropy from residual time-shifts using all frequencies (blue line) and reduced frequencies (dashed blue line) are also shown.

be up to 10.2 s for Rayleigh waves for some specific paths which must be accounted for in the total phase shift of the seismograms. To investigate the source effect on our time-shift measurements, we made a simple test for the earthquakes used in this study. We computed synthetic seismograms imposing the 3−D crust in a area around the source, leaving the rest of the model untouched. We then measured the time-shifts with the PREM seismograms. We observed that time-shifts due to different source excitations are less than 2 s for both Rayleigh and Love waves which does not affect our results.

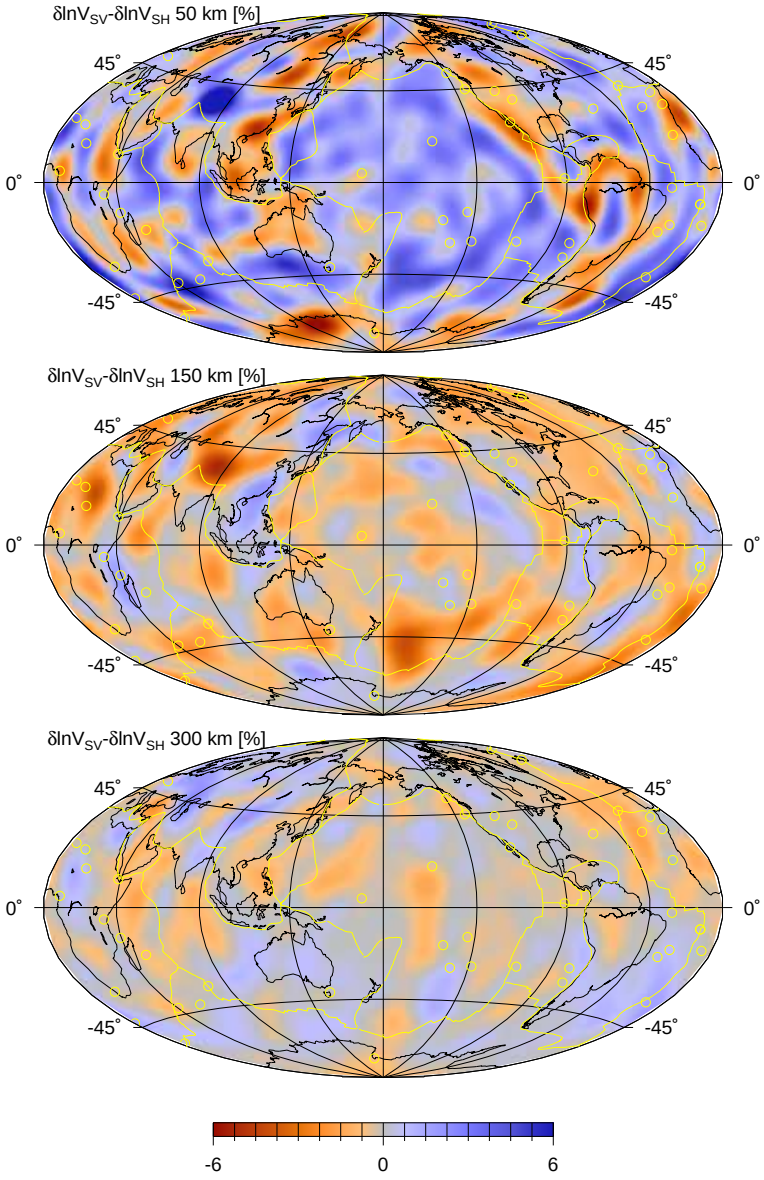

Figure 3.5.2: Lateral variations of radial anisotropy inverted from the residual time-shifts (δt^{cc}) of Love and Rayleigh waves using all frequencies at different depths.

3.6 Conclusions

We tested the accuracy of crustal corrections on fundamental mode surface waves using ray theory and 2−D finite-frequency theory. We compared synthetic seismograms computed in PREM and PREM+Crust2.0. Time-shift analysis shows that, on average, we can correct the phase of Rayleigh and Love wave seismograms for the periods longer than 60 s and 80 s, respectively. At longer periods, although the residual time-shifts are not perfect, they are within the uncertainties of the measured phase velocities. Nevertheless, the inversion of the residual time-shifts shows that they produce radial anisotropy comparable in strength to those of existing models at least in the uppermost mantle. Furthermore, observed sign changes in anisotropy could be artefacts of improper crustal corrections. This needs thorough investigation in future work.

Rayleigh and Love waves are differently affected by continental and oceanic crust. This is responsible for an apparent dependence of crustal corrections on distance and clearly highlights the non-linearity of the crustal influence. Therefore common extensions to GCA do not significantly improve the corrections. Higher order effects are important when starting from a 1−D reference model. First order crustal corrections on full seismograms should only be done from a 3−D reference model close to the crustal model. This implies being able to calculate kernels and make measurements in 3−D models. Such techniques are becoming available now, but still constitue a considerable computational challenge. Alternativley, inverting full seismograms from 1-D reference model should include crustal parameters using full non-linearity.

Acknowledgments

We are grateful to Jeroen Tromp for providing the SEM code, for his help with problems related to the code and for extended discussions. We would like to thank Ana Ferreira for her initial help with the local modes. Guust Nolet and an anonymous reviewer provided insightful comments which improved the paper considerably. Ying Zhou deserves special mention. Her questions during the review process allowed us to find a mistake in our local mode calculations for which we are very grateful. This project is funded by the European Commission's Human Resources and Mobility Program Marie Curie Research Training Network SPICE Contract No.

MRTNCT-2003-504267. Synthetic seismograms were calculated on a 64-processor cluster financed by the Dutch National Science Foundation under the grant number NWO:VICI-865.03.007.

Chapter 4

Assessment of tomographic mantle models using spectral element seismograms

Summary

We investigated the agreement between real seismograms and those predicted by long wavelength mantle models by looking at phase and amplitude differences. We computed the synthetic seismograms using a spectral element method with the 3-D mantle models and the appropriate crustal model on top. To check the phase agreement, we measured the time-shifts between the real and synthetic surface waves and body wave phases such as P, S, and SS using a cross-correlation technique. We also compared the amplitudes of real and synthetic seismograms in order to understand how well the models explain not only the phases but the whole waveforms. 3-D mantle models clearly improve the phase mismatch between real and synthetic seismograms. The remaining misfit, however, is so large that we cannot distinguish between the different tomographic models. We believe that this is mainly due to imperfect crustal corrections applied prior to the construction of the models. Amplitude mismatches are large and independent of 3-D mantle models. Our results suggest that, particularly for surface waves, the scattering and source effects together can at most ex-

The content of this chapter is under revision as: Bozdağ, E. & Trampert, J., 2009. Assessment of tomographic mantle models using spectral element seismograms, *Geophysical Journal International*.

plain half of the amplitude mismatches, pointing to a strong contribution of attenuation. More detailed modelling shows that scattering effects are more important for body waves.

4.1 Introduction

Mantle structure has extensively been studied by means of seismic tomography using body wave, surface wave, and free oscillation data. At the global scale, surface waves, both fundamental mode and overtones, provide a good sampling of the upper half of the mantle, particularly beneath oceans where the lack of stations can be problematic for body wave studies. With an increasing number of high-quality data in recent years, the resolution of tomographic images has considerably improved. Nevertheless in practice, the resolution of the models is still largely controlled by details of the inverse algorithm. Without exception, all current models are based upon some approximation to the wave equation. Due to its simplicity and ease of implementation, ray theory has been used in most surface wave studies either by averaging the phase shift along the great circle path (e.g., Ritsema et al., 1999; Boschi & Ekström, 2002) or the exact ray path (e.g., Ferreira & Woodhouse, 2006, 2007). However, ray theory is a high-frequency approximation and only valid when the scale length of the heterogeneities is larger than the width of the first Fresnel zone (Wang & Dahlen, 1995), which is often violated in current tomographic images (Spetzler et al., 2001). To improve the validity of the forward theory for smaller scale structures, the local phase velocity perturbations should be integrated over an influence zone using either 2-D (e.g., Spetzler et al., 2002; Yoshizawa & Kennett, 2002; Ritzwoller et al., 2002) or 3-D (e.g., Zhou et al., 2004) finite-frequency kernels. Currently there is no consensus whether this also improves the resolution of the tomographic models (e.g., Ritzwoller et al., 2002; Zhou et al., 2004; Sieminski et al., 2004; Trampert & Spetzler, 2006).

We are currently able to simulate wave propagation in 3-D earth models using numerical techniques without any approximation to the wave equation. Ideally, this new tool should be used for imaging. Although some efforts are being made (e.g., Tape et al., 2007), the computational cost is still prohibitive. A first step is to check how well current models are predicting full seismograms and not just the small part which has been used in the construction of the models. Until now, only a few qualitative comparisons have been published (e.g., Komatitsch & Tromp, 2002b; Komatitsch et al.,

2002, 2005; Tsuboi et al., 2003, 2004). Recently, Qin et al. (2009) have done a quantitative comparison of 3-D mantle models based on a spectral element method looking at long period minor and major arc surface waves (T $>$ 100 s). Our focus is rather to investigate the effect of regularization and try to understand how much we can interpret the Earth's interior based on our current knowledge. The ambiguity in choosing the appropriate regularization motivated us to select 3-D mantle models obtained with different levels of damping. We will present a full quantitative assessment of different models within the period range of 40-200 s. We compared real seismograms to synthetic ones by examining the phase and amplitude differences of both surface waves and body waves. The computation of synthetic seismograms was performed with the spectral element code (SEM) by Komatitsch & Tromp (2002a,b). We computed four sets of SEM seismograms using four different mantle models varying from very smooth to very rough (Trampert & Spetzler, 2006). We superimposed the 3-D mantle models on top of transversely isotropic PREM (Dziewonski & Anderson, 1981) and used Crust2.0 (Bassin et al., 2000) as a 3-D crustal model on top. In addition, we computed synthetic seismograms using the independent 3-D mantle model S20RTS (Ritsema et al., 1999) with Crust2.0 on top, PREM with Crust2.0 on top and PREM alone to separate effects of 1-D and 3-D crustal and mantle models on seismograms.

In the following section, we briefly describe the tomographic models used in this study. In Section 3, we illustrate how the phase and amplitude measurements were carried out and in Section 4, we present our results based on the measured time-shifts and the amplitude misfits. Finally, we discuss our results before presenting some general conclusions.

4.2 Models entering the calculation of the synthetic seismograms

We calculated SEM seismograms using four mantle models (Trampert & Spetzler, 2006) obtained for different levels of horizontal damping (see brown curve in Figure 1 in Trampert & Spetzler (2006)). The models were constructed from about 1.5 million Rayleigh wave phase velocity measurements of fundamental modes (Trampert & Woodhouse, 1995) and the first five overtones (van Heijst & Woodhouse, 1999). 2-D finite-frequency kernels (Spetzler et al., 2002) for phase velocity and asymptotic kernels for depth were used. We selected models obtained with horizontal smoothing

only. In Figure 4.2.1, sample cross-sections of the four mantle models are presented. It is clearly seen that highly damped models provide smooth and nice plots whereas smaller damping causes the amplitude of the models to increase and introduces more small scale heterogeneities. Model m_1 is built from 780 independent parameters and would correspond to one selected by a cautious seismologist who prefers smooth models. Model m_2 has a trace of the resolution of 1752 and would correspond to a model chosen by a criterion of maximum curvature of the L-curve. Model m_3 is built from 3108 independent parameters corresponds to a point where the reduced χ^2 does not significantly change any more. Finally model m_4 is constructed from 4489 independent parameters and is so rough that everybody would agree that instabilities are occurring. These models would obviously lead to different geological interpretations. The properties of the models can also be seen on the rms-amplitude plotted as a function of depth in Figure 4.2.2. From smooth to rough, the amplitude of the models gradually increases. Below 1000 km depth, the sensitivities of the data to mantle structure diminishes and this is where instabilities start to develop. This is particularly evident for model m_4. The synthetic seismograms computed for these four mantle models from a sample source-receiver path is shown in Figure 4.2.3 where the change both in phase and amplitude is observed for body and surface waves. Crustal corrections were applied to the phase velocity data prior to inversion by calculating the exact eigen frequencies at each grid point of Crust5.1 (Mooney et al., 1998) and integrating the crustal phase along the great circle path. In the computation of the full SEM seismograms, we added the 3-D crustal model Crust2.0 (Bassin et al., 2000) on top of the 3-D mantle models. Crust5.1 and Crust2.0 are very similar except for differences in grid resolution and sediments. This difference will not affect our results since the period range we considered is not sensitive enough to perceive the difference. We also computed synthetic seismograms with an independent mantle model S20RTS (Ritsema et al., 1999) which is parameterized in the same way of the mantle models of Trampert & Spetzler (2006) and has the same crustal correction. S20RTS, in addition, includes body wave and normal mode data and used a different smoothing strategy (norm damping). In order to appreciate the prediction power of 1-D and 3-D crustal models alone, we also computed some synthetic seismograms using PREM alone and PREM with Crust2.0 (hereafter called PREM+Crust2.0) on top. In all claculations, we included gravity, attenuation, oceans, topography and bathymetry, ellipticity, and rotation

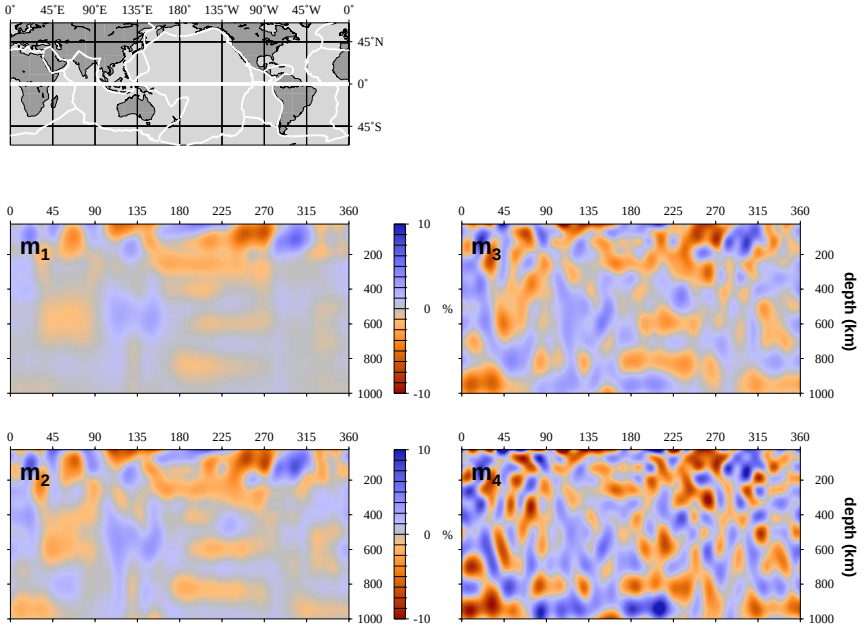

Figure 4.2.1: Sample cross-sections of the selected mantle models of Trampert & Spetzler (2006) along the equator. From m_1 to m_4, the damping decreases.

(except that in PREM alone, we did not use topography and bathymetry) to obtain synthetic seismograms as realistic as possible. The length of the synthetic seismograms is 60 minutes and the sampling rate is 0.26 s. Based on the mesh we used, the shortest period in our synthetics is approximately 30 s.

4.3 Measuring misfits

4.3.1 Comparing real and SEM seismograms

We checked the agreement between the real and synthetic seismograms by examining phase and amplitude differences. We compared 1374 pairs of seismograms using 22 shallow (CMT depth $\leq$ 50 km) and two deep (CMT depth $>$ 50 km) earthquakes having moment magnitudes between 6 and 8 (Table 4.3.1) which were used to compute synthetic seismograms for selected 3-D mantle models from Trampert & Spetzler (2006) with varying levels of damping. For S20RTS, PREM+Crust2.0 and PREM models,

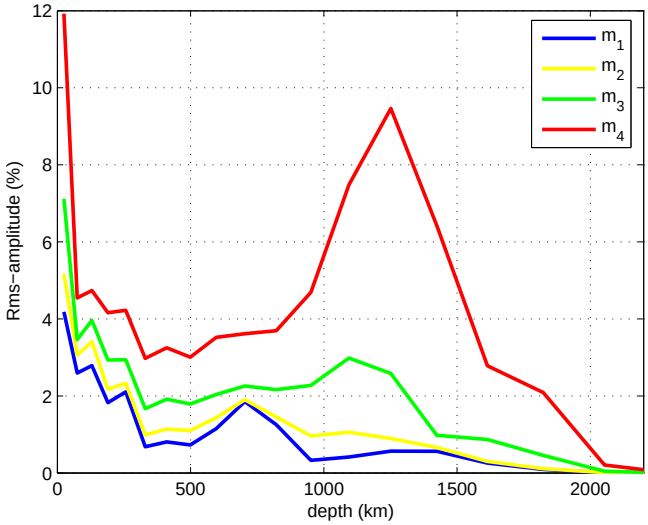

Figure 4.2.2: Rms-amplitudes of the mantle models of Trampert & Spetzler (2006) as a function of depth.

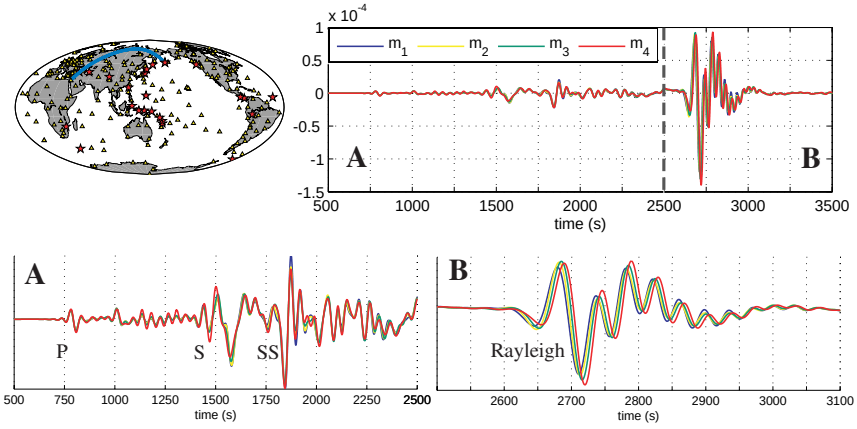

Figure 4.2.3: SEM seismograms computed in the selected mantle models of Trampert & Spetzler (2006) for the Rat Islands earthquake (2003 March 17, $M_w = 7.0$) recorded at the station EIL ($\Delta = 93°$). Body waves and Rayleigh waves are in A and B, respectively, to show the effect of damping on computed seismograms enlarged.

we only computed synthetic seismograms for a restricted number of earthquakes (see Table 4.3.1). To avoid interferences of the multiple-orbit waves, caustic source and receiver effects, we only considered source-receiver pairs with distances between $20^\circ < \Delta < 160^\circ$. For body waves, we used less source-receiver pairs because P, S and SS phases cannot be observed on all seismograms at all distances. We used broad-band vertical-component seismograms (sampling rate = 0.05 s) as real data. Data having high noise level or glitches were removed manually by comparing them to the synthetic seismograms. The Global CMT solution parameters were used to generate the SEM seismograms. To make the real and SEM seismograms comparable, we followed five steps: 1) We deconvolved the instrument response from the real seismograms. 2) We convolved the SEM seismograms with a Gaussian source time function with the half duration for each earthquake taken from the Global CMT solution, 3) We downsampled the real and the SEM seismograms, whose original sampling rates are 0.05 s and 0.26 s, respectively, to a common sampling rate of 1.3 s. 4) We applied a band-pass filter to both real and SEM seismograms between 35 s and 300 s to analyse the data in the period range 40 s - 200 s. 5) After synchronizing the real and SEM seismograms, we applied a common time window to make sure that the files contain the same number of points. In Figure 4.3.1, sample synthetic seismograms from 1-D and 3-D models are compared to the real ones in terms of waveform difference. As can clearly be seen, not only surface waves but also body waves are in good agreement both in phase and amplitude for the case of the 3-D mantle model m_1 although it is constructed by surface wave phase velocity measurements only. However, the waveform comparisons between the real and synthetic seismograms reveal clear mismatches. The mismatch from m_1 is, in general, smaller than those from PREM and PREM+Crust2.0 however there is still a large unexplained data in full seismograms. To quantify the agreement between the real and SEM synthetics, we examined the phase and amplitude differences for fundamental mode Rayleigh waves and the body wave phases P, S, and SS.

4.3.2 Measuring time-shifts

We examined the phase difference of surface and body waves in different ways. For surface waves, we first applied a time-variable filter (Cara, 1973) to both the real and SEM seismograms to extract the fundamental mode Rayleigh waves. We then took the Fourier transform of the surface waves

Table 4.3.1: List of earthquakes selected from the global CMT catalogue (www.globalcmt.org). Corresponding broad-band vertical component seismograms were obtained through IRIS (www.iris.edu) and synthetic seismograms were computed using the spectral element code by Komatitsch & Tromp (2002a,b). Synthetic seismograms for PREM, PREM+Crust2.0 and S20RTS models were only computed for earthquakes marked by a star.

Event name	Region	Date	Moment magnitude (M_w)	Depth (km)
111600B	New Ireland Region	16/11/2000	8.0	24
120600C	Turkmen SSR	06/12/2000	7.0	33
010901G	Vanuatu Islands	09/01/2001	7.0	114
011101A	Vancouver Island Region	11/01/2001	6.0	24
021801B	Prince Edward Islands Region	18/02/2001	6.0	15
022401A*	Molucca Passage	24/02/2001	7.0	43
101201E	South of Mariane Islands	12/10/2001	7.0	42
110901A	Panama - Costa Rica Border	09/11/2001	6.0	17
091602E	Papua New Guinea	16/09/2002	6.0	15
012703A	Turkey	27/01/2003	6.0	15
031703E*	Rat Islands, Aleutian Islands	17/03/2003	7.0	27
042403B	Kuril Islands	24/04/2003	6.0	44
062003D	Western Brazil	20/06/2003	7.0	556
103103A	Off East Coast of Honshu	31/10/2003	7.0	15
020504B*	Irian Jaya Region, Indonesia	05/02/2004	7.0	13
030804D	Northern Mid-Atlantic Ridge	08/03/2004	6.0	12
032704G	Xizang	27/03/2004	6.0	12
042704D	Vanuatu Islands	27/04/2004	6.0	17
100904E*	Near Coast of Nicaragua	09/10/2004	6.9	39
110904F	Solomon Islands	09/11/2004	6.9	12
112804I*	Hokkaido, Japan Region	28/11/2004	7.0	47
200602222219A*	Mozambique	22/02/2006	7.0	12
200608200341A	Scotia Sea	20/08/2006	7.0	17
200612261226A	Taiwan Region	26/12/2006	6.9	20

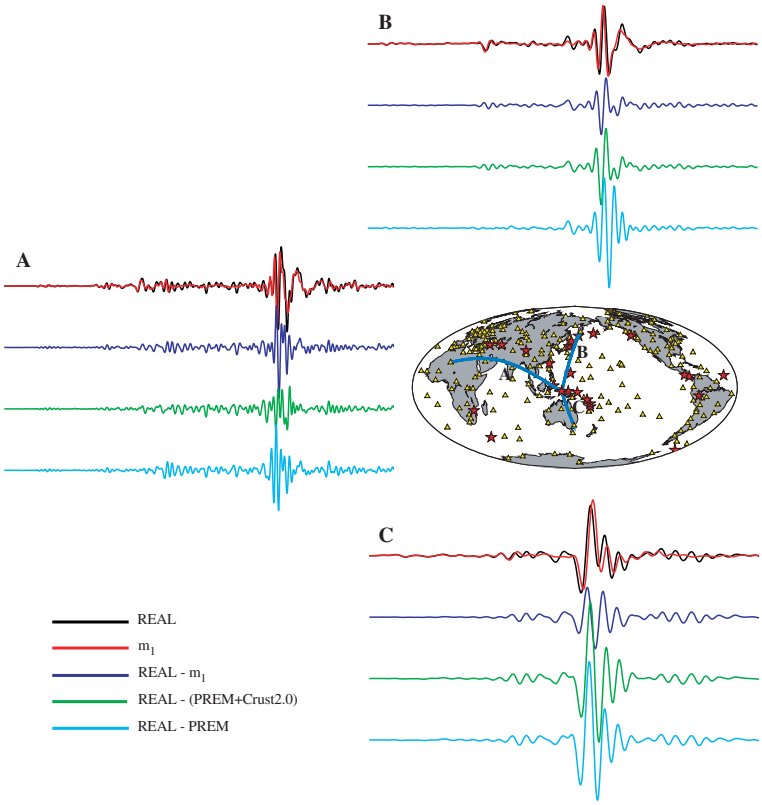

Figure 4.3.1: Waveform differences between real and SEM seismograms computed in the mantle model m_1 of Trampert & Spetzler (2006) with Crust2.0 (Bassin et al., 2000) on top (blue), PREM (Dziewonski & Anderson, 1981) with Crust2.0 on top (green), and PREM alone (cyan). On top of the waveform differences, real (black) and SEM seismograms computed in m_1 (red) are plotted together. The earthquake (Irian Jaya Region, 2004 February 5, $M_w = 7.0$) and the ray paths to the stations A) TAM ($\Delta = 128°$), B) PET ($\Delta = 60°$), and C) CAN ($\Delta = 34°$) are shown in the map where the distribution of earthquakes (red stars) and stations (yellow triangles) used in this study are marked as well.

which can be presented in terms of an amplitude (A) and a phase (ϕ)

$$S_{real}(\omega) = A_{real}(\omega)exp[\phi_{real}(\omega)] \tag{4.3.1}$$

and

$$S_{SEM}(\omega) = A_{SEM}(\omega)exp[\phi_{SEM}(\omega)] \tag{4.3.2}$$

where subscripts $real$ and SEM represent the real and synthetic seismograms, respectively and ω is the angular frequency. To determine the phase shift between these two seismograms, we cross-correlated $S_{real}(\omega)$ with $S_{SEM}(\omega)$ and measured the phase of the cross-correlogram as a function of frequency. After unwrapping the phase, the time-shifts between real and SEM seismograms are obtained by dividing the phase of the cross-correlogram by the angular frequency

$$\delta t_{Rayleigh}(\omega) = \frac{\phi_{real}(\omega) - \phi_{SEM}(\omega)}{\omega}. \tag{4.3.3}$$

Because body waves are not as dispersive as surface waves, we measured the time-shift between real and SEM body phases in time domain. We first applied time windows to extract the P, S, and SS phases. The time window was defined sufficiently large around the theoretical arrival times from IASPEI91 (Kennett & Engdahl, 1991) to capture the complete pulse. We Fourier transformed both and applied a Gaussian filter with a central period 60 s. The band-with of the Gaussian filter is 100% of the central frequency, which means that we consider the full waveform rather than a signal at a specific frequency. We then took the cross-correlation in the frequency domain and after taking the inverse Fourier transform, we measured the time-shift by picking the maximum of the cross-correlogram. We repeated the same procedure for each body wave phase P, S, and SS.

4.3.3 Measuring amplitude differences

The examination of the amplitude differences between the real and synthetic data help to understand how much the 3-D mantle models contribute to amplitudes. We extracted Rayleigh waves and P, S and SS as described for the phase shift measurements. After taking the Fourier transform, we calculated the amplitude spectra of the real and SEM seismograms without applying a Gaussian filter. It is more convenient to make the comparisons in frequency domain since, in a time series, it is difficult to measure the

amplitude independently from the phase which may distort the amplitude measurements. For the amplitude comparisons, we used the expression

$$\chi(\omega) = ln \left[\frac{A_{real}(\omega)}{A_{SEM}(\omega)} \right] \tag{4.3.4}$$

which shows the logarithmic ratio between the real and SEM amplitudes as a function of frequency. For body waves, we only show ratios at 40 and 60 s.

4.4 Results

4.4.1 Time-shift analysis

For Rayleigh waves, the measured time-shifts corresponding to the four selected models $m_1 - m_4$ are shown as histograms in Figure 4.4.1. We observe that the phase of the SEM seismograms, in general, match the real ones quite well at long periods. At 200 s, most of the time-shifts are within $\pm$10 s. The discrepancy between the real and SEM seismograms increases with decreasing period and the histograms broaden at short periods to around $\pm$ 25 s. Most interestingly, the histograms hardly allow to differentiate between the models. To further appreciate the effect of the 3-D mantle on synthetic seismograms, we made synthetics in PREM alone, in PREM+Crust2.0 and in S20RTS (obtained from an extended data set and a different regularisation strategy) and compared them to the real seismograms in a similar way (Figure 4.4.2). The time-shifts between the real and PREM seismograms can be up to 100 s at short period Rayleigh waves. Including the 3-D crustal model, on average, slightly reduces the time-shifts however it is very clear that a 3-D mantle model is needed to explain the phase of Rayleigh waves. The histograms also suggest that, on average, there is not much difference between using PREM alone or PREM+Crust2.0. This is of course not the case. The crust has a large effect on Rayleigh waves as demonstrated in Bozdağ & Trampert (2008) depending on distance and the portion of oceanic versus continental path. The histograms document that statistically one is not better off to explain the real data with just a crustal model. The results from S20RTS are, in general, in agreement with the results from the mantle models of Trampert & Spetzler (2006), particularly at long periods. Slight differences emerge for periods shorter than 100 s.

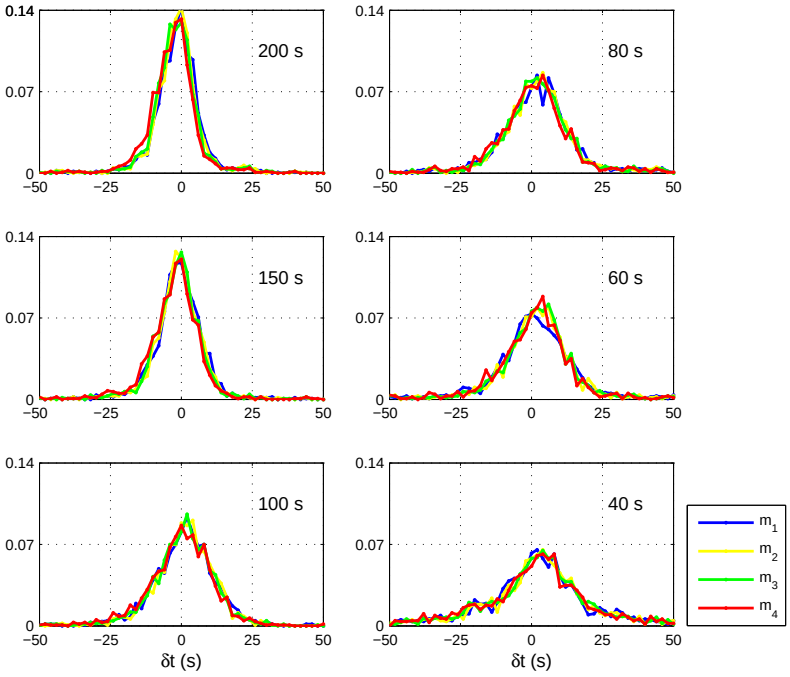

Figure 4.4.1: The histograms of Rayleigh wave time-shifts as a function of period measured between real and SEM seismograms computed in 3-D mantle models. From model m_1 to m_4, the damping decreases.

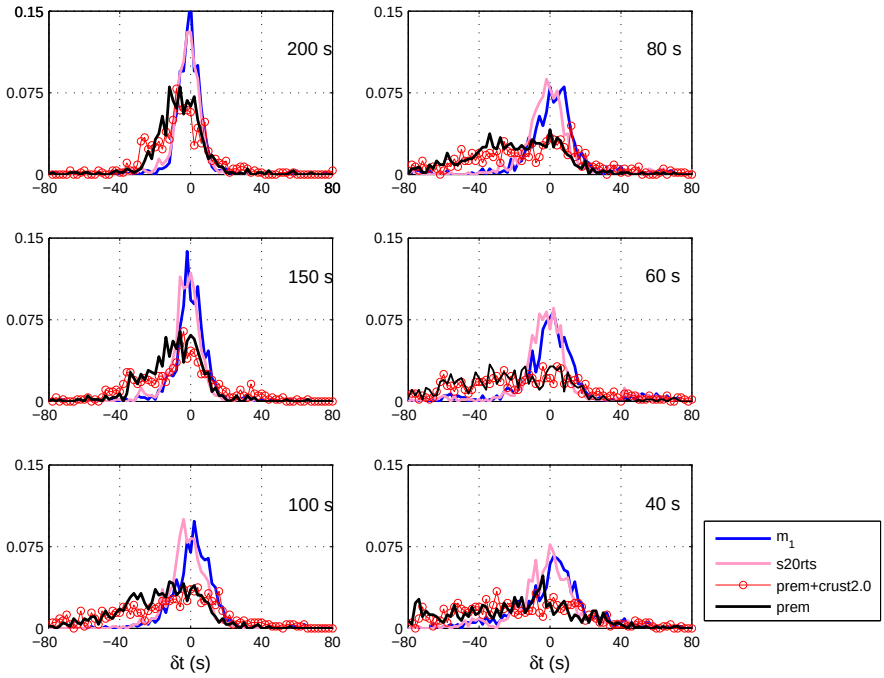

Figure 4.4.2: Same as Figure 4.4.1, but for different 1-D and 3-D models. The restricted ray paths correspond to the marked earthquakes only in Table 4.3.1.

To quantify the difference between the model predictions, we computed the probability that the measured time-shifts between the real and SEM seismograms are smaller than those for the smoothest model m_1. The numbers demonstrate that it is hard to distinguish which mantle model explains the real data best (Figure 4.4.3). Amongst the models of Trampert & Spetzler (2006) a slightly better fit to the real data is provided by the mantle model m_2. S20RTS shows a slight improvement, particularly at 40 seconds, but does worse at the longest periods. We will address this in the discussion section. While the predictions of all 3-D mantle models are very close to each other, synthetics for PREM and PREM+Crust2.0 are much worse, only having a 25% chance of making a better Rayleigh wave prediction than any of the mantle models. It is important to stress that although the model predictions are statistically similar when compared to real data, the model predictions themselves are different. This is most easily seen by making histograms of time-shifts between synthetics from different models. Some examples are shown in Figure 4.4.4. For the mantle models, these differences are fairly uniform as a function of period and reach 10 s. This is only a small fraction of the period for 200 s Rayleigh waves but represents a quarter of a cycle at 40 s. We clearly see that the mantle models provide a large contribution compared to PREM+Crust2.0 and that the crustal signal is large compared to PREM alone.

Body waves carry information from deeper parts whereas fundamental mode surface waves are mainly influenced by the first a few hundred km of the mantle. The measured time-shifts between the real and synthetic P, S, and SS-phases are shown as histograms in Figure 4.4.5. P-waves are, in general, less affected by lateral heterogeneities than S-waves because the corresponding velocity anomalies are much smaller. The P-wave histograms are narrow and vary between $\pm$ 5 s while those for SS-waves reach more than $\pm$ 10 s. In addition, we see a slight shift in the histograms towards positive times similar to surface wave results. S20RTS results are in agreement with those from m_1 with a slight improvement. Similar to Rayleigh wave observations, 3-D mantle has the most effect compared to 3-D crust alone. For P-waves there is an 80% chance that the results from m_1 and m_2 are equal. Even for PREM alone there is a 40% chance that the predictions are equal to m_1. For S-waves, m_4 is clearly the worst and S20RTS is the best, which is reassuring because the latter included S-wave travel-time measurements. For SS-waves, a 3-D mantle model clearly helps to improve the misfits and m_4 gives the worst match to the real seismograms. Results

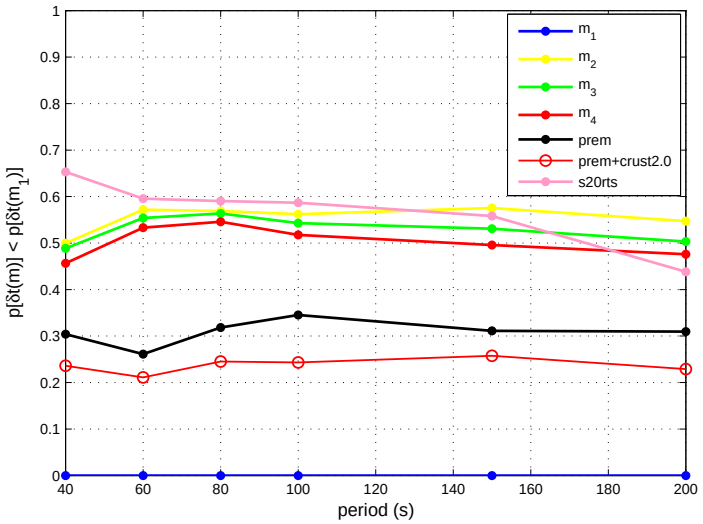

Figure 4.4.3: The probability that the Rayleigh wave time-shifts as a function of period between real and SEM seismograms for any model are smaller than those between real and SEM seismograms for the smoothest mantle model m_1.

for m_1, m_2 and S20RTS are very similar. The full probabilities may be found in Table 4.4.1. As for surface waves, it is hard to distinguish which mantle model predicts real seismograms best, while the models clearly make different predictions (Figure 4.4.6). The difference between m_4 and m_1 is bigger than that between S20RTS and m_1, showing that S20RTS is much similar to smooth models. For P-waves, the 3-D mantle has the biggest effect whereas for S-waves, the 3-D crust has a comparable effect to those from mantle models. The crustal effect is observed even more for SS-phases. It is interesting to note that crustal corrections are quite important and that their values are different from ray theoretical estimations (J. Ritsema, personal communication). This suggests that crustal corrections for body waves need to be evaluated for the correct frequency content.

4.4.2 Amplitude analysis

We also looked at amplitude ratios between the real and SEM seismograms for various models. As for the phase measurements, we considered Rayleigh waves, P, S, and SS-phases. In all synthetics, the attenuation model is that

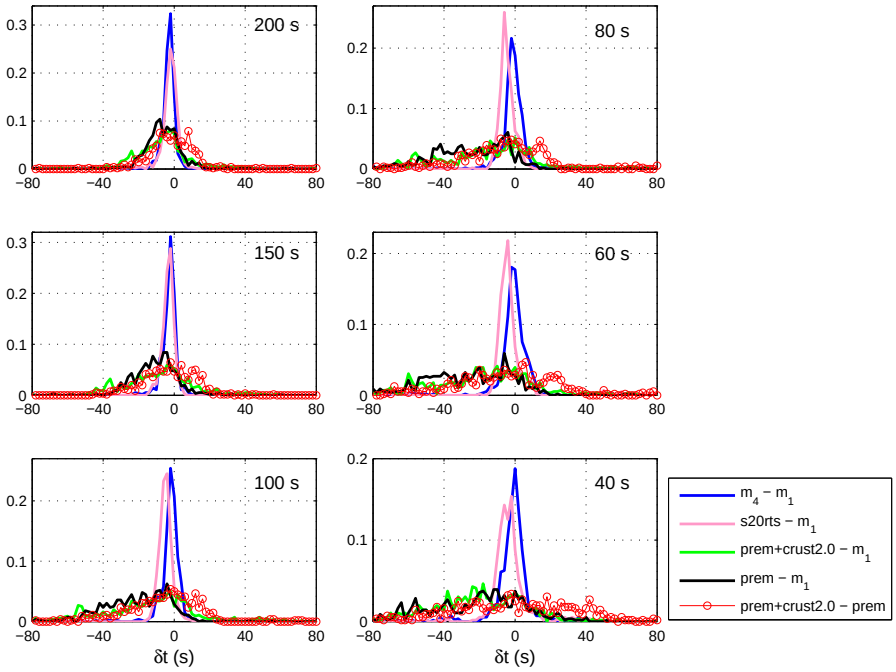

Figure 4.4.4: The histograms of Rayleigh wave time-shifts as a function of period between synthetic seismograms corresponding to various 1-D and 3-D models.

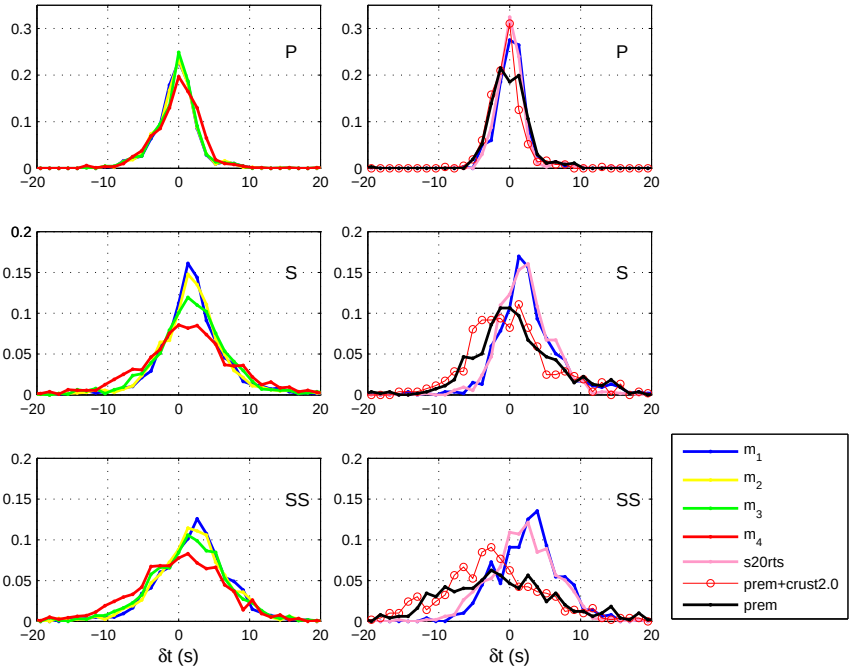

Figure 4.4.5: The histograms of body wave time-shifts at 60 s measured between real and SEM seismograms computed in 3-D mantle models (m_1 to m_4) with different dampings (left), and in different 1-D and 3-D models (right). In figures in the right column, the restricted ray paths correspond to the marked earthquakes only in Table 4.3.1.

Table 4.4.1: The probabilities that the body wave time-shifts at 60 s between real and SEM seismograms for any model are smaller/larger than, or equal to those between real and SEM seismograms for the smoothest mantle model m_1.

P-wave

	p[δt(m)<δt(m_1)]	p[δt(m)=δt(m_1)]	p[δt(m)>δt(m_1)]
$m = m_2$	9%	80%	11%
$m = m_3$	15%	66%	19%
$m = m_4$	20%	43%	37%
m =S20RTS	22%	57%	21%
m =PREM+Crust2.0	27%	40%	33%
m =PREM	15%	46%	39%

S-wave

	p[δt(m)<δt(m_1)]	p[δt(m)=δt(m_1)]	p[δt(m)>δt(m_1)]
$m = m_2$	23%	45%	32%
$m = m_3$	28%	26%	46%
$m = m_4$	28%	19%	53%
m =S20RTS	40%	33%	27%
m =PREM+Crust2.0	38%	18%	44%
m =PREM	46%	12%	42%

SS-wave

	p[δt(m)<δt(m_1)]	p[δt(m)=δt(m_1)]	p[δt(m)>δt(m_1)]
$m = m_2$	30%	39%	31%
$m = m_3$	36%	23%	41%
$m = m_4$	34%	18%	48%
m =S20RTS	38%	33%	29%
m =PREM+Crust2.0	45%	1%	54%
m =PREM	31%	1%	68%

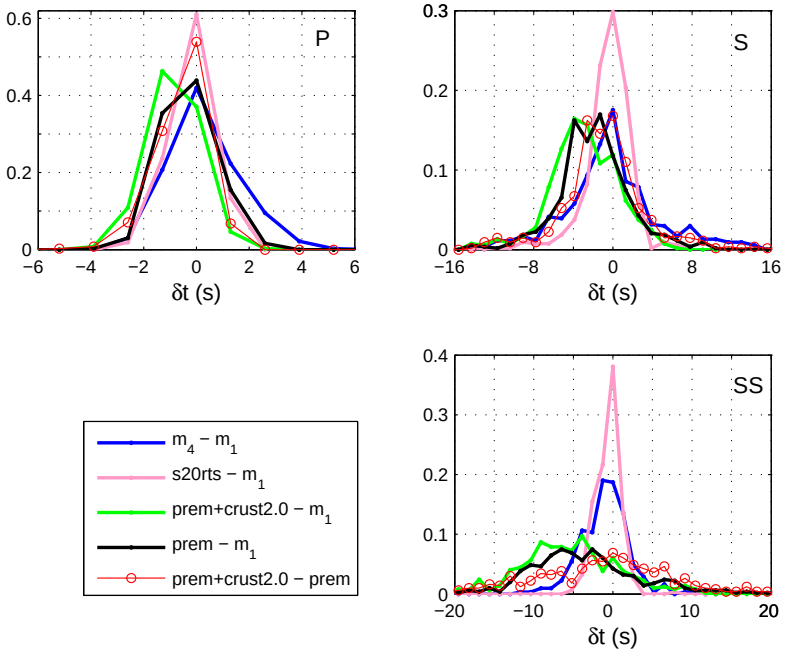

Figure 4.4.6: The histograms of the body wave time-shifts at 60 s between synthetic seismograms corresponding to various 1-D and 3-D models.

of PREM. For Rayleigh waves, all mantle models from Trampert & Spetzler (2006) give similar amplitude misfits (Figure 4.4.7). For long periods, the agreement with the real seismograms is much better (mostly within a factor of 1.5) and as the period decreases the misfits can be up to a factor of 7 at 40 s. In general, the histograms are skewed towards negative values indicating that the amplitudes of Rayleigh waves are overestimated in the SEM seismograms. S20RTS also gives similar results whereas we start observing some slight differences for PREM and PREM+Crust2.0 at short periods (Figure 4.4.8). A detailed statistical analysis does not show much improvement of amplitude misfits for 3-D models compared to PREM. On average, all models, 1-D or 3-D, give the same probabilities of fit. To understand the amplitude predictions in more detail, we compared the models directly by taking the ratios of the respective amplitude anomalies (Figure 4.4.9). We compared the 3-D mantle models with respect to m_1. The difference naturally increases with decreasing damping thus we only present results for m_4/m_1 where we observed the most discrepancy in amplitudes. For predicting amplitudes, S20RTS is more similar to smooth models, as for phase. We also compared PREM+Crust2.0 with respect to m_1 and PREM to isolate the scattering contributions due to the 3-D mantle and 3-D crust only. The comparison of PREM with m_1 shows the total scattering contributions due to the 3-D crust and mantle. Although there are some differences between the 3-D mantle models, it is noteworthy that the 3-D model hardly matters for amplitudes, indicating that scattering effects in the mantle are minor for these periods. The scattering due to 3-D crust and mantle together is bigger but remains small compared to the observed amplitude anomalies between real and the synthetic seismograms.

For body waves, we compared the amplitude spectra of P, S and SS-waves at 40 and at 60 s. On average, mantle models from different dampings differ more for S and SS-waves compared to P-waves (Figure 4.5.1). There is again a slight over-estimation of the synthetic amplitudes, but much less than for surface waves. The results from S20RTS are, in general, in agreement with m_1 and the results from PREM and PREM+Crust2.0 only do slightly worse (Figure 4.5.2). As the period decreases, the histograms start broadening showing that the scattering effects increase as one would expect. Comparing of the amplitude misfits from different models, we find that the width of the histograms due to a 3-D mantle and/or 3-D crust is comparable to that of the amplitude ratio between m_4 and m_1. Scattering due to 3-D crust and mantle together can be up to a factor of 1.6 for P

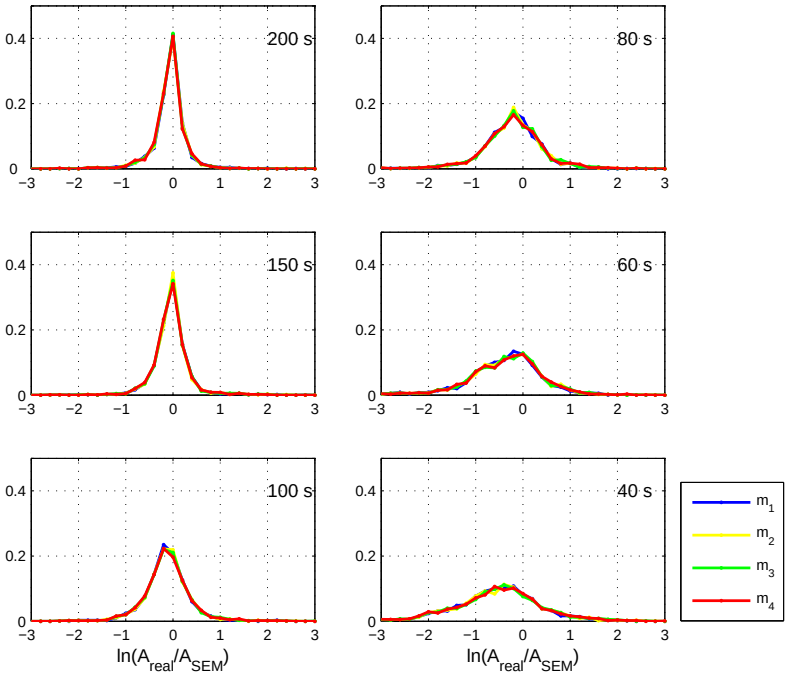

Figure 4.4.7: The histograms of Rayleigh wave amplitude ratios as a function of period between real and SEM seismograms computed in 3-D mantle models. From model m_1 to m_4, the damping decreases.

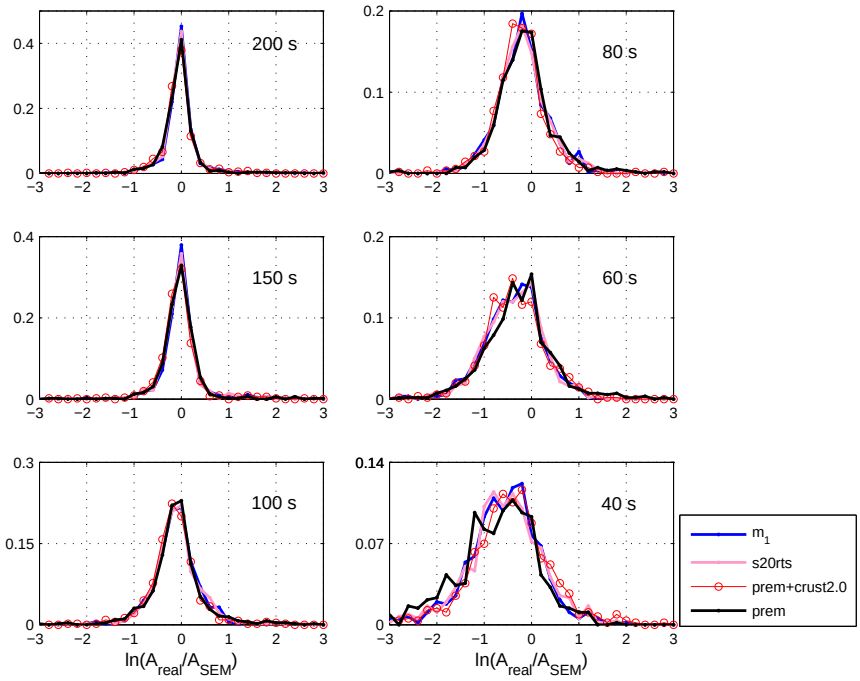

Figure 4.4.8: Same as Figure 4.4.7 but for different 1-D and 3-D models. The restricted ray paths correspond to the marked earthquakes only in Table 4.3.1.

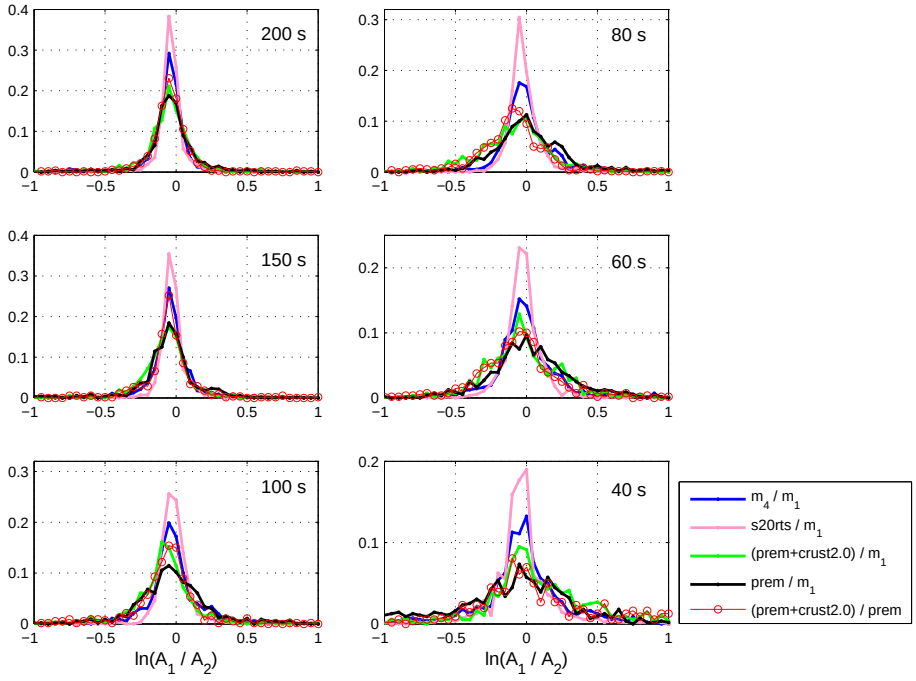

Figure 4.4.9: The histograms of amplitude ratios for Rayleigh waves as a function of period between synthetic seismograms corresponding to various 1-D and 3-D models.

waves and more than a factor of 2.7 for SS-waves (Figure 4.5.3).

4.5 Discussions

It is encouraging to see that 3-D mantle models bring synthetic seismograms closer to observed data in phase. It is surprising, however, that quite diverse mantle models cannot be distinguished on average, although their individual predictions are all different. For amplitudes, all models considered make hardly any difference to amplitudes. The question is then what can be done to improve models so that they predict better phases and amplitudes. Let us first discuss the individual contributors to the phase before turning to the amplitude.

At least for surface waves, a strong contribution to the overall phase is from the crust. The 3-D crust, compared to PREM, advances the phase of Rayleigh waves in oceans and slows it down in continents producing positive and negative phase shifts, respectively, with respect to real seismograms. Thus, there is, on average, not much difference between the results from PREM and PREM+Crust2.0 for seismograms which randomly sample oceans and continents. Classical surface wave tomography applies crustal corrections to dispersion measurements before inverting for mantle structure. Bozdağ & Trampert (2008) showed that these corrections are not perfect and these significant errors can map into 3-D mantle structure. While it is clear that the 3-D mantle models improve the phase a lot compared to PREM+Crust2.0 (Figure 4.4.2), the histograms of unexplained time residuals are very similar to those incurred by imperfect ray theoretical crustal corrections (Bozdağ & Trampert, 2008), both as a function of period and epicentral distance. These imperfect crustal corrections have a bigger effect on the phase than different mantle models (compare Figure 4.4.2 and 4.4.4), and compared to real seismograms cannot identify potentially better mantle models. The crustal contribution to body waves is less than to Rayleigh waves but the discrepancy between the real and SEM seismograms is also smaller for body waves. Although we have not done a detailed study of classical body wave crustal corrections, they seem to be different when the full frequency content of the body wave is considered. This assumes that Crust2.0 represents the real Earth which might not be the case of course. Another issue is how the crust is implemented in the SEM code. Capdeville & Marigo (2008) reported that in spectral element methods, different smoothing strategies applied to crustal models

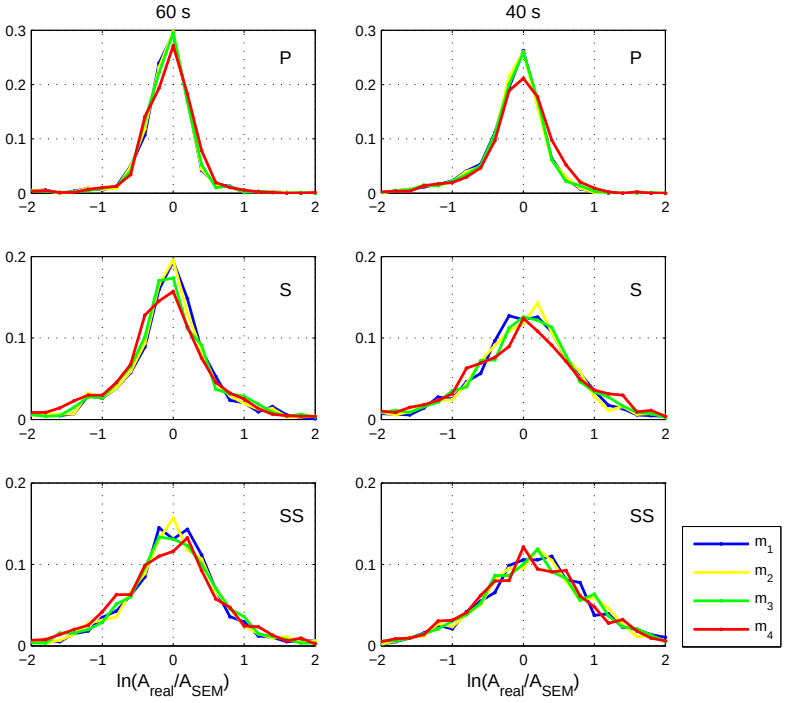

Figure 4.5.1: The histograms of body wave amplitude ratios at 60 and 40 s between real and SEM seismograms computed in 3-D mantle models. From model m_1 to m_4, the damping decreases.

may change the response. Because in the SEM code a smoothed version of Crust2.0 is used, errors could be introduced in our time-shift measurements. All in all, it is very likely that the crust is responsible for not being able to distinguish the models, whether it is the crustal correction, the crustal model itself or its implementation in the code. If this is indeed the case, progress to better mantle models can only be made if the crust and mantle are jointly imaged without a correction step.

While the source and receiver phase shifts are correctly modelled in SEM given the source parameters and the crustal model, the phase measurements which are used to construct the models usually neglect them (e.g., Trampert & Woodhouse, 1995; Ekström et al., 1997; van Heijst & Woodhouse, 1999; Visser et al., 2008). A recent study by Ferreira & Woodhouse (2007) has shown that the receiver phase-shift is negligible for both Rayleigh and Love

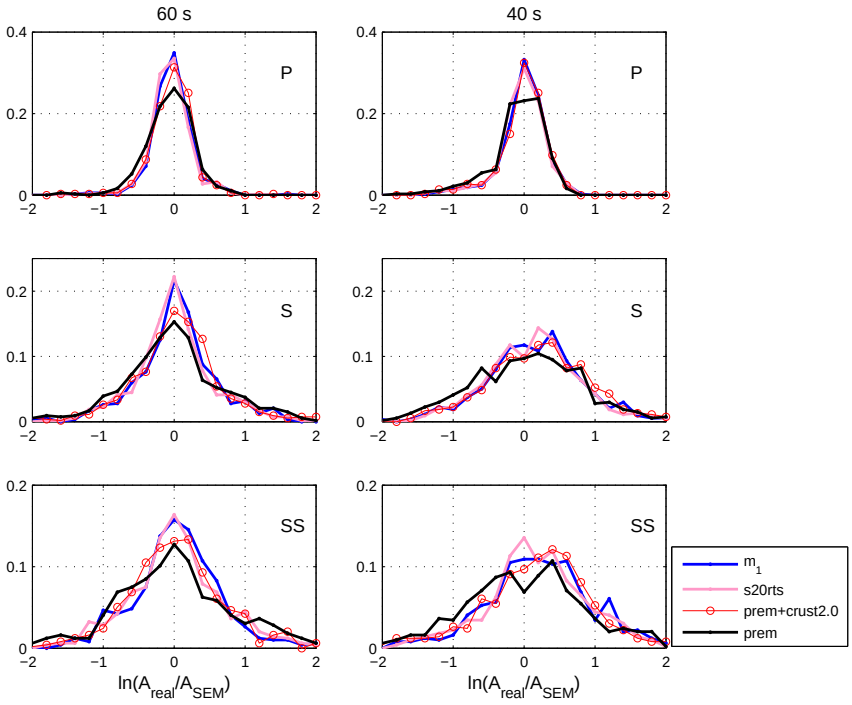

Figure 4.5.2: Same as Figure 4.5.1 but for different 1-D and 3-D models. The restricted ray paths correspond to the marked earthquakes only in Table 4.3.1.

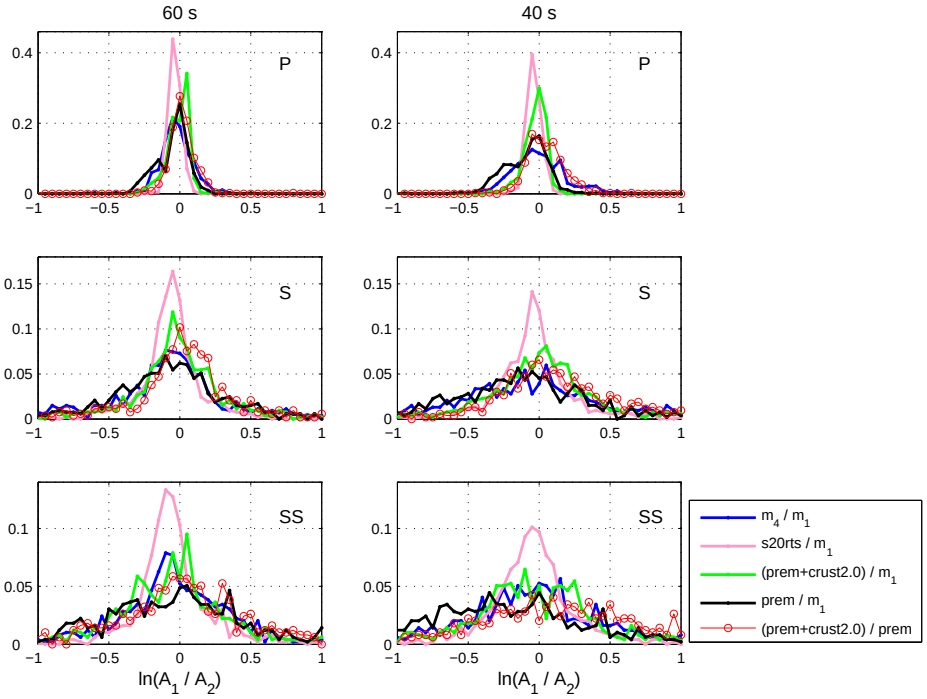

Figure 4.5.3: The histograms of amplitude ratios for body wave phases at 60 and 40 s between synthetic seismograms corresponding to 1-D and 3-D models. The restricted ray paths correspond to the marked earthquakes only in Table 4.3.1.

waves. The source phase shift due to lateral heterogeneities around source can be up to 10 s for Rayleigh waves for some specific ray paths as shown by Ferreira & Woodhouse (2006) however, on average, it is small compared to our observed discrepancies. The important question is if we used the right source parameters in SEM. The results of Ferreira & Woodhouse (2007) show that source parameters obtained for more detailed source inversions are compatible with CMT parameters, and therefore do not change the data misfit dramatically. Although we cannot exclude source effects to be responsible for part of our observed mismatches, we think that crustal effects are more likely.

The amplitudes of the SEM seismograms are clearly larger than those of the real ones. The amplitude of the seismograms depends on elastic and anelastic terms together with source and receiver factors (e.g., Dalton & Ekström, 2006). The elastic term, or focussing/defocussing of seismic waves, is due to lateral heterogeneities in velocity models (e.g., Woodhouse & Wong, 1986; Wang & Dahlen, 1995). 3-D attenuation of the Earth's mantle has been investigated using either surface waves (e.g., Romanowicz, 1995; Selby & Woodhouse, 2000; Gung & Romanowicz, 2004; Dalton & Ekström, 2006; Dalton et al., 2008) or body waves (e.g., Bhattacharyya et al., 1996; Reid et al., 2001; Warren & Shearer, 2002) quantifying the anelastic contribution to amplitudes. Smooth models improve the misfit slightly more than the rough ones but not enough. This could be an indication of a problem in the gradient of the heterogeneities in the models. For Rayleigh waves, we find amplitude anomalies are mostly within a factor of 1.3 due to focussing/defocussing of the 3-D crust and/or mantle (Figure 4.4.9). However, the amplitude factors between the real and all SEM seismograms are much larger (Figures 4.4.7 and 4.4.8) and thus the remaining difference should be due to attenuation, source, and receiver effects. Ferreira & Woodhouse (2007) showed that source and elastic contributions are equally important for surface wave amplitudes, but that receiver effects can be neglected for vertical component seismograms. They also showed that inverting for source parameters does not dramatically change the CMT source parameters. Dalton & Ekström (2006) found that source factors are mostly within 0.8 to 1.2 for 75 s Rayleigh waves and are getting smaller at longer periods. According to these observations, the source and elastic contributions could roughly amount to a factor of 1.7 together, still only half of what we observe in Figures 4.4.7 and 4.4.8. This could point to a large contribution from attenuation to the amplitudes, almost as much as

the total contribution from elastic and source effects. Zhou (2009) computed 3-D sensitivity kernels of phase delays and amplitudes for perturbations in anelasticity for fundamental mode surface waves and reported that scattering effects due to 3-D velocity models are larger than those from 3-D anelastic models. If this is the case, then much of the amplitude misfit should be due to the uncertainty in source parameters and unexplained elastic structures in the velocity models. However, although the amplitudes might be less sensitive to the 3-D variations in anelastic structure, 1-D attenuation model can still be an important factor. As shown in Resovsky et al. (2005), anealsticity in the crust and uppermantle is poorly constrained in 1-D shear Q models and depending on the data set, Q values can vary between 250 to more than 1000. The uncertainty in the 1-D Q models may have a critical importance on amplitudes since 3-D models are constructed as a perturbation from them. All in all, we have to consider both source and anelastic parameters to explain the amplitudes in inversions. It is also interesting to see that, on average, 3-D models do not improve the amplitudes at all compared to PREM. Such an observation was already reported by Ferreira & Woodhouse (2007), where they compared some real seismograms with synthetics computed by great circle approximation and full ray theory and by Qin et al. (2009) based on a spectral element method.

The different models change body wave amplitudes far more. Especially S and SS-phases are very much affected by focussing/defocussing. Amplitude anomalies from 3-D mantle and/or 3-D crust are almost as large as the observed amplitude anomalies between the real and SEM seismograms and increase with decreasing period. This is in agreement with Tibuleac et al. (2003) who concluded that P-wave amplitudes were dominated by elastic effects at higher periods. While focussing/defocussing is more important to explain the amplitudes of body waves, in the period range of this study, attenuation and/or source effects still play a big role as demonstrated by the similarity of all anomaly histograms in Figures 4.5.2 and 4.5.3.

The selected mantle models of Trampert & Spetzler (2006) were surprisingly indistinguishable when the corresponding synthetic seismograms are compared to real ones. S20RTS shows small but noticeable differences in the predictions. We used S20RTS with its P-wave model to compute the synthetic seismograms. Scaling P-wave as in the models of Trampert & Spetzler (2006) brings the results closer to each other. The selected models of Trampert & Spetzler (2006) all use the same regularization strategy (horizontal damping) with a varying damping factor. S20RTS uses norm

damping. If you compare S20RTS with any of the others, the correlation is above 0.9 in the uppermost mantle, except at the shallowest depth, just below the crust where the correlation is only 0.4. Changing the regularization in the models of Trampert & Spetzler (2006) from horizontal to vertical smoothing or norm damping, brings the correlation in this layer also to 0.9. This suggests that the model just below the crust is to a great extend undetermined by data. A bad crustal correction will mostly effect the models just below the crust and in the absence of data constraint effect the model more or less depending on the choices of regularization.

4.6 Conclusions

We made quantitative comparisons between real seismograms and those predicted by various 3-D mantle models using the SEM code. We found that the models clearly improved the phase mismatch for surface as well as body waves, but surprisingly we could hardly distinguish between the mantle models. Amplitude discrepancies were little affected by any of the models and even PREM alone made just as good predictions.

Although there is a slight tendency that smooth models are just a little better, our findings are a good reminder that current tomographic images contain (or lack) a lot of structure because of a certain choice of regularization. While the latter has a dramatic effect on the interpretation little of the details are currently constraint by the seismograms.

Based on our previous work, we suggest that the imperfect crustal corrections are the most likely reason why we cannot distinguish the models in terms of phase mismatches. We observed a large amplitude misfit which can not fully be explained by scattering effects. For surface waves, the scattering effects are relatively independent of period down to 40 s at least and attenuation has a comparable effect to those from elastic and source effects together. For body waves, the focussing effect is much larger and increases with decreasing period. The 1-D attenuation model we used in this study might bias the amplitude mismatches particularly at short periods sensitive to the upper-mantle where a high attenuation is observed. This subject needs further investigation.

Much signal remains unexplained in phase and amplitude. Progress in explaining the phase can only be made if we treat crust and mantle simultaneously. Attenuation is notoriously difficult as demonstrated by the lack of agreement between various studies, but essential to explain

the amplitudes. Full waveform inversion is probably the most elegant way forward, but needs to address the full elastic, anelastic and source problem.

Acknowledgment

We gratefully acknowledge the availability of broad-band seismograms from the IRIS/USGS (IU), the USAF/USGS (GT) and the New China Digital Seismograph Network (IC) from USGS/Albuquerque Seismological Laboratory; the IRIS/IDA Network (II) from University of California, Scripps Institute of Oceanography; the Canadian National Seismograph Network (CN) from Geological Survey of Canada; the GEOSCOPE Network (G) from Institut de Physique du Globe de Paris; the GEOFON Network (GE) from GEOFON Program and the Pacific21 Network (PS) from University of Tokyo through the IRIS Data Management Center. Guust Nolet's comments and suggestions about body-wave amplitudes were very helpful. This project was funded by the European Commission's Human Resources and Mobility Program Marie Curie Research and Training Network SPICE Contract No. MRTNCT-2003-504267. Synthetic seismograms were calculated on a 128-processor cluster financed by the Dutch National Science Foundation under the grant number NWO:VICI-865.03.007.

Chapter 5

Assessing new and existing misfit functions for full waveform inversion

Summary

The resolution of tomographic images is naturally dependent on the data coverage which controls the sensitivity kernels. It is well known that different parts of seismograms are sensitive to different parts of the Earth structure. In classical tomography, in general, either travel-time/phase information of isolated phases or only a small portion of the seismic data are used. We propose new misfit functions based on instantaneous phase and cross-correlation measurements to extract as much information as possible from a single seismogram which, in the ideal case, is the full seismogram. Both misfit functions are advantageous in different situations. The instantaneous phase misfit is based on the instantaneous phase difference of the observed and synthetic data and allows to separate the phase and amplitude information completely. Different observables in a seismogram can be equally weighted which allows the use of a complete wavetrain. Cross-correlation misfit is based on the difference between the cross-correlation of observed data with the predicted ones and the auto-correlation of the data. We made a qualitative comparison of the new misfit functions with existing ones based on waveform and travel-time measurements in order to have a better insight on the advantages/disadvantages of the chosen misfit functions by computing finite-frequency adjoint kernels based upon a spectral

element method.

Due to an inherent amplitude dependence, the CC measurements favor the higher amplitude parts of wavetrains. The amplitude dependence, however, provides a robust way of suppressing the noise in the signals. Instantaneous phase measurements give the same emphasis to every phase in a seismogram. This allows the use of full seismograms as demonstrated by synthetic experiments. Possible challenges in instantaneous phase measurements are the noise in the signals and the phase jumps. These problems can easily be detected by examining the envelope and the instantaneous phase difference. Possible solutions are to cut the problematic parts of the seismograms, to use high quality data only or to start inversions with long period seismograms which are less affected by these problems.

Our comparisons showed that cross-correlation measurements are similar to travel-time measurements in case of a single phase but due to their amplitude dependent nature they have properties similar to waveform misfit functions. Instantaneous phase measurements are complementary to classical misfit functions and successful in equally weighting the different phases in a seismogram. The instanteneous phase measurements are promising and show a way to exploit as much information as possible from a single seismogram.

5.1 Introduction

The ultimate aim of seismic tomography is to take advantage of the advances in the theory of seismic wave propagation, numerical techniques, and the increase in computational facilities, in order to use full waveforms in the inversion. So far, there have been some full waveform inversions in regional studies (e.g., Pratt, 1999; Chen et al., 2007b) however it remains a challenge for global tomography. Capdeville et al. (2005) proposed a full waveform inversion on a global scale in a synthetic experiment based on a source-stacking technique with a spectral element method as a forward theory. They focused on long period data to reduce the effect of non-linearity and pointed out the problems of full waveform inversion such as data coverage and in case of using full seismograms the dominance of surface waves in the residuals. Insufficient data coverage resulting from the uneven distribution of earthquakes and stations on the globe is one of the major restrictions at the global scale. A way to remedy this problem would be to use as much information as possible from a single seismogram

by defining appropriate misfit functions.

In seismic tomography, in general, the data are secondary observables such as travel-times, phase information or the waveforms of a small portion of the full seismograms. An important factor controlling the usable data is the forward theory. Ray based tomography is limited to use well-isolated primary body-wave phases (e.g., Zhou, 1996; Boschi & Dziewonski, 2000) or surface waves where the fundamental and higher modes are well separated from each other (e.g., Trampert & Woodhouse, 1995; Ekström et al., 1997). Its advantage is the ease of implementation and reasonable computational requirements. It is also common to integrate different data sets to increase the resolution (e.g., Su et al., 1994; Masters et al., 1996; Ritsema et al., 1999; Mégnin & Romanowicz, 2000; Gu et al., 2001). Using higher order approximations, Li & Romanowicz (1996); Mégnin & Romanowicz (2000); Gung & Romanowicz (2004) constructed global models based on waveforms obtained by isolating all identifiable observables in the seismograms. Consideration of finite-frequency effects of wave propagation in seismic tomography is one way to improve the resolution of tomographic images (e.g., Montelli et al., 2004). However, as reported by Boschi et al. (2007), the application of the finite-frequency theory based on 1-D reference models still requires the use of well-isolated phases and hence insufficient data coverage hampers the advantage of the finite-frequency theory.

The progress in numerical techniques and computational facilities made it possible to solve the wave equation numerically in 3-D Earth models (e.g., Komatitsch & Vilotte, 1998; Komatitsch & Tromp, 1999; Capdeville et al., 2003). This has two major consequences for seismic tomography: 1) in the forward problem, the full non-linearity of wave propagation can be taken into account, 2) the finite-frequency kernels can be computed numerically in 3-D background models. This has provided an invaluable opportunity to increase the usable data in seismic tomography. The challenge is that the computation of Fréchet derivatives can be computationally expensive. Using 3-D numerical simulations, Zhao et al. (2005) proposed to compute the Green's functions in 3-D models and store them as a function of space and time which requires high memory facilities. Tromp et al. (2005) addressed this problem by combining 3-D numerical simulations with adjoint techniques (e.g., Tarantola, 1984, 1988; Fink, 1997; Talagrand & Courtier, 1987; Crase et al., 1990; Pratt, 1999; Akçelik et al., 2003). The idea is that for a chosen set of observables and a defined misfit function, the Fréchet derivatives, that define the sensitivity of the data with respect to the model

parameters, can be computed by only two numerical simulations, one for the forward and one for the adjoint wavefield. An alternative way of introducing 3-D numerical simulations into tomography is using the scattering integral method as applied by Chen et al. (2007b) to some local problems.

The Fréchet kernels strongly depend on the chosen misfit function. In seismic tomography, the misfit function is often based on cross-correlation travel-time measurements (e.g., Luo & Schuster, 1991; Zhao et al., 2000; Marquering et al., 1999; Dahlen et al., 2000), relative amplitude variations (e.g., Dahlen & Baig, 2002; Ritsema et al., 2002) or waveform differences (e.g., Tarantola, 1984, 1988; Nolet, 1987). Luo & Schuster (1991) combined travel-time measurements with full wave simulations by defining the misfit function based on the cross-correlation between observed and synthetic phases. Liu & Tromp (2006) presented finite-frequency adjoint sensitivities based on travel-time measurements computed in 3-D background models for some regional phases. Adjoint kernels for different global observables computed in a 1-D background model were illustrated in Liu & Tromp (2008). This approach provides a robust estimation of the phase or time difference between two isolated phases. The drawback is that phases have to be picked from seismograms. Some automated phase picking algorithms have recently been released (e.g., Maggi et al., 2009) particularly tailored to adjoint tomography. Tape et al. (2007) showed a way to implement adjoint techniques in seismic tomography based on these cross-correlation travel-time measurements using a conjugate gradient method at a regional scale. This technique was extended to generalized seismological data functionals (GSDF) (Gee & Jorden, 1992) in which the misfit function is represented by the frequency dependent phase-delay and amplitude differences between the observed and synthetic data (e.g., Chen et al., 2007b). This allows to characterize the change in waveforms. However the separation of phases might not always be very practical. Although a full seismogram can be considered rather than a phase, time domain waveform misfit functions (e.g., Tarantola, 1984; Nolet, 1987) defined by the difference between the observed and synthetic data can behave in a highly non-linear way with respect to the velocity model (e.g., Gauthier et al., 1986; Luo & Schuster, 1991). Furthermore, such misfit functions emphasize high amplitude phases. Fichtner et al. (2008) proposed a full waveform inversion scheme based on time-frequency analysis which separates the phase and amplitude information. It provides an elegant way to deal with non-linearities under the regime of the Born approximation. This technique can suffer from

cycle slips which is treated by isolating the non-problematic parts with a designed Gaussian filter.

Using 3-D numerical simulations, adjoint techniques offer the possibility to compute the sensitivity of an arbitrary part of the signal in 3-D structures. In this study, we propose new misfit functions based on the instantaneous phase and the complete cross-correlation measurements with the aim to extract as much information as possible from a single seismogram which, in ideal case, is not cut into pieces. We computed data sensitivities for several synthetic experiments using adjoint techniques based upon the spectral element method of Komatitsch & Tromp (2002a,b) which allows us to analyze the properties of the different misfit functions.

The instantaneous phase misfit is based on the difference between the instantaneous phases of the observed and synthetic data. Instantaneous phase measurements are widely used in exploration seismics (e.g., Taner et al., 1979; Perz et al., 2004; Barnes, 2007) to increase the resolution in the obtained images. This approach allows to separate phase and amplitude information similar to the work of Fichtner et al. (2008) in the time-frequency domain. The instantaneous phase however is a pure time-domain approach which requires less data processing. The complete cross-correlation misfit (to avoid confusion, from now on, we refer this new misfit function as cross-correlation misfit and classical cross-correlation travel-time misfit as travel-time misfit) is based on the difference between the cross-correlations of observed and synthetic data, and the auto-correlation of the observed data. It can be considered as an extension of cross-correlation travel-time measurements to a wavetrain rather than a single phase. In order to have a better insight into the advantages/disadvantages of the chosen misfit functions, we compared the corresponding kernels with those from waveform and travel-time differences that are widely used in seismic tomography.

In the following, we derive the adjoint sources for instantaneous phase and cross-correlation misfits. In Section 5.3, we present the sensitivity kernels computed by adjoint techniques for some synthetic experiments. Finally, we discuss our findings in the context of global mantle tomography and give general conclusions in Sections 5.4 and 5.5, respectively. Although our focus is on global problems, the results are applicable to other scales of seismic tomography.

5.2 Misfit functions and associated adjoint sources

In this part, we present the derivations of the adjoint sources for the newly introduced misfit functions based on instantaneous phase and cross-correlation differences. In addition, we give brief information on the adjoint sources of commonly used misfit functions, waveform and travel-time measurements. In the following sections, we are going to compare the different misfit functions by computing their adjoint sensitivity kernels using the spectral element method by Komatitsch & Tromp (2002a,b).

5.2.1 Misfit kernels

In seismic tomography, we try to extract optimum information on the model parameters describing the Earth's interior from a set of observed seismograms. The model parameters are updated by minimizing a defined misfit between observed and synthetic data. In adjoint tomography, the gradient of a misfit function can be computed through the interaction of the forward wavefield with the adjoint wavefield which is generated by the back-propagation of data residuals. The inverse problem is solved iteratively based on a gradient method such as conjugate gradient or steepest descent.

Let us define a misfit function in a generalized form,

$$\chi(m) = \sum_{r=1}^{N} \int_0^T g(\mathbf{x_r}, t, \mathbf{m}) dt \tag{5.2.1}$$

where r denotes the number of receivers and $g(\mathbf{x_r}, t, \mathbf{m})$ can be any kind of misfit between observed and synthetic data observed at $\mathbf{x_r}$. t is time and $\mathbf{m}$ are the model parameters. The gradient of the misfit function, which gives the Fréchet derivatives is,

$$\delta\chi = \sum_{r=1}^{N} \int_0^T \partial_{\mathbf{s}} g(\mathbf{x_r}, t, \mathbf{m}) \cdot \delta\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) dt \tag{5.2.2}$$

where $\delta\mathbf{s}(\mathbf{x_r}, t, \mathbf{m})$ is the perturbations in the displacement wavefield with respect to the model parameters. In general, $\delta\mathbf{s}$ is expressed using the Born approximation (Hudson, 1977; Wu & Aki, 1985),

$$\delta s_i(\mathbf{x},t,\mathbf{m}) = \\ -\int_0^t \int_V [\delta\rho(\mathbf{x}',\mathbf{m})G_{ij}(\mathbf{x},\mathbf{x}';t-t')\partial_{t'}^2 s_j(\mathbf{x}',t',\mathbf{m}) \\ +\delta c_{jklm}(\mathbf{x}',\mathbf{m})\partial_k' G_{ij}(\mathbf{x},\mathbf{x}';t-t')\partial_l' s_m(\mathbf{x}',t',\mathbf{m})]d^3\mathbf{x}'dt' \quad (5.2.3)$$

where ρ is the density, c_{jklm} is the forth-order elastic tensor, and $\delta\rho$ and δc_{jklm} are the associated perturbations. If we insert eq. 5.2.3 into 5.2.2, the gradient becomes,

$$\delta\chi = \\ -\sum_{r=1}^{N}\int_0^T \partial_{\mathbf{s}} g(\mathbf{x_r},t,\mathbf{m}) \int_0^t \int_V [\delta\rho(\mathbf{x}',\mathbf{m})G_{ij}(\mathbf{x},\mathbf{x}';t-t')\partial_{t'}^2 s_j(\mathbf{x}',t',\mathbf{m}) \\ +\delta c_{jklm}(\mathbf{x}',\mathbf{m})\partial_k' G_{ij}(\mathbf{x},\mathbf{x}';t-t')\partial_l' s_m(\mathbf{x}',t',\mathbf{m})]d^3\mathbf{x}'dt'. \quad (5.2.4)$$

Using the reciprocity of Green's function (Aki & Richards, 1980; Dahlen & Tromp, 1998) and reversing it in time, it is convenient to define the *adjoint wavefield*,

$$s_k^\dagger(\mathbf{x}',t',\mathbf{m}) = \int_0^{t'} \int_V G_{ki}(\mathbf{x}',\mathbf{x_r};t-t')f_i^\dagger(\mathbf{x_r},t,\mathbf{m})d^3\mathbf{x}dt \quad (5.2.5)$$

where $f_i^\dagger$ is the *adjoint source* given by,

$$f_i^\dagger(\mathbf{x_r},t,\mathbf{m}) = \sum_{r=1}^{N} \partial_{s_i} g(\mathbf{x_r},t,\mathbf{m})\delta(\mathbf{x}-\mathbf{x_r}). \quad (5.2.6)$$

For a more complete derivation of the adjoint wavefield, please see Tromp et al. (2005); Liu & Tromp (2006); Tromp et al. (2008).

For an isotropic material, the gradient of a given misfit function can alternatively be written in the form of,

$$\delta\chi = \int_V [K_\rho(\mathbf{x})\delta\ln\rho(\mathbf{x}) + K_\mu(\mathbf{x})\delta\ln\mu(\mathbf{x}) + K_\kappa(\mathbf{x})\delta\ln\kappa(\mathbf{x})]d^3\mathbf{x} \quad (5.2.7)$$

where K_ρ, K_μ, K_κ are the Fréchet derivatives with respect to the model parameters density (ρ), shear modulus (μ) and bulk modulus (κ), respectively. Comparing to eq. 5.2.4, the Fréchet kernels can be written,

$$K_\rho(\mathbf{x}) = -\int_0^T \rho(\mathbf{x})\mathbf{s}^\dagger(\mathbf{x}, T-t)\cdot\partial_t^2\mathbf{s}(\mathbf{x},t)dt \tag{5.2.8}$$

$$K_\mu(\mathbf{x}) = -\int_0^T 2\mu(\mathbf{x})\mathbf{D}^\dagger(\mathbf{x}, T-t) : \mathbf{D}(\mathbf{x},t)dt \tag{5.2.9}$$

$$K_\kappa(\mathbf{x}) = -\int_0^T \kappa(\mathbf{x})[\nabla\cdot\mathbf{s}^\dagger(\mathbf{x}, T-t)][\nabla\cdot\mathbf{s}(\mathbf{x},t)]dt. \tag{5.2.10}$$

where $\mathbf{D}$ and $\mathbf{D}^\dagger$ are the forward and adjoint traceless strain deviators, respectively. This shows that the kernels can be obtained by two numerical simulations, one for the forward and one for the adjoint wavefields. Using the relations between the elastic moduli and the parameters density, P-wave (α) and S-wave (β) velocities, the gradient can also be written in the form of,

$$\delta\chi = \int_V [K_\rho^{'}(\mathbf{x})\delta\ln\rho(\mathbf{x}) + K_\beta(\mathbf{x})\delta\ln\beta(\mathbf{x}) + K_\alpha(\mathbf{x})\delta\ln\alpha(\mathbf{x})]d^3\mathbf{x} \tag{5.2.11}$$

where $K_\rho^{'}$, K_β and K_α are given as,

$$K_\rho^{'} = K_\rho + K_\kappa + K_\mu \tag{5.2.12}$$

$$K_\beta = 2\left(K_\mu - \frac{4}{3}\frac{\mu}{\kappa}K_\kappa\right) \tag{5.2.13}$$

$$K_\alpha = 2\left(\frac{\kappa + \frac{4}{3}\mu}{\kappa}K_\kappa\right). \tag{5.2.14}$$

As can be noticed immediately, for a given seismic wavefield, the misfit kernels depend on the adjoint wavefield. The adjoint wavefield is controlled by an adjoint source which is dependent on a defined misfit from a particular observable. Thus the key point to initiate an adjoint simulation is to define an adjoint source for a defined misfit function. In our case, the misfit functions are based on instantaneous phase, cross-correlation, waveform and travel-time measurements. The unit of the kernels will be dependent on the chosen misfit function as well.

5.2.2 Instantaneous phase measurements

A signal having no negative-frequency component is called an *analytic signal* ($\widetilde{f}(t)$) which can be constructed from a real signal ($f(t)$) and its Hilbert transform ($\mathcal{H}\{f(t)\}$),

$$\widetilde{f}(t) = f(t) - i\mathcal{H}\{f(t)\} \tag{5.2.15}$$

where the Hilbert transform of a signal defined as,

$$\tilde{f}(t) = -\frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{f(\tau)}{t-\tau} d\tau. \tag{5.2.16}$$

Due to singularities at $t = \tau$, the integral is defined using the Cauchy principle value. The analytic signal can be written in terms of instantaneous amplitude $A(t)$ and phase $\phi(t)$,

$$\widetilde{f}(t) = A(t)e^{i\phi(t)}. \tag{5.2.17}$$

This form allows to separate the phase and amplitude in the time domain rather than in the frequency domain. The instantaneous phase difference between the observed and synthetic seismograms in the time domain is completely independent from amplitude information. We could also define a misfit function based on the instantaneous amplitude difference however we restrict our discussion here to the phase information only.

We define the square instantaneous phase misfit as,

$$\chi(m) = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \left\|\phi_r(t,m) - \phi_r^{obs}(t)\right\|^2 dt. \tag{5.2.18}$$

where $\phi_r(t,m)$ and $\phi_r^{obs}(t)$ denote the instantaneous phase of synthetic and observed seismograms, respectively, as a function of time t and at receiver r for a given Earth model m. The observed and synthetic seismograms are, in general, filtered and windowed within a certain frequency band and time window. The gradient of the misfit function is then,

$$\delta\chi = \sum_{r=1}^{N}\int_0^T [\phi_r(t,m) - \phi_r^{obs}(t)]\delta\phi_r(t,m)dt \tag{5.2.19}$$

where $\delta\phi_r$ is the perturbation in instantaneous phase of the synthetic seismograms due to a perturbation in model parameters δm. From eq. 5.2.15, $\phi_r(t,m)$ is defined as,

$$\phi_r(t,m) = \arctan \frac{\Im\left\{\tilde{s}(t,m)\right\}}{\Re\left\{\tilde{s}(t,m)\right\}} \tag{5.2.20}$$

where $\tilde{s}(t,m)$ is the analytic signal of synthetic data $s(t,m)$ given by,

$$\tilde{s}(t,m) = s(t,m) - i\mathcal{H}\left\{s(t,m)\right\}. \tag{5.2.21}$$

The perturbations in instantaneous phase then becomes,

$$\delta\phi_r(t,m) = \delta\left(\frac{\Im\left\{\tilde{s}(t)\right\}}{\Re\left\{\tilde{s}(t)\right\}}\right) / \left[1 + \left(\frac{\Im\left\{\tilde{s}(t)\right\}}{\Re\left\{\tilde{s}(t)\right\}}\right)^2\right]. \tag{5.2.22}$$

After some algebra, we obtain,

$$\delta\phi_r(t,m) = \frac{\mathcal{H}\left\{s(t,m)\right\}\delta s(t,m) - s(t,m)\delta\mathcal{H}\left\{s(t,m)\right\}}{s(t,m)^2 + \mathcal{H}\left\{s(t,m)\right\}^2}. \tag{5.2.23}$$

We can write $\delta H\left\{s(t,m)\right\}$ as,

$$\delta\mathcal{H}\left\{s(t,m)\right\} = \frac{\partial\mathcal{H}\left\{s(t,m)\right\}}{\partial s(t,m)}\delta s(t,m) \tag{5.2.24}$$

and using the definition of the partial derivative,

$$\delta\mathcal{H}\left\{s(t,m)\right\} = \frac{\mathcal{H}\left\{s(t,m) + \delta s(t,m)\right\} - \mathcal{H}\left\{s(t,m)\right\}}{\delta s(t,m)}\delta s(t,m) \tag{5.2.25}$$

where $\delta s(t,m)$ is the perturbation in synthetic displacement wavefield with respect to the model parameters m. Using the linearity property of the Hilbert transform, it reduces to,

$$\delta\mathcal{H}\left\{s(t,m)\right\} = \frac{\mathcal{H}\left\{\delta s(t,m)\right\}}{\delta s(t,m)}\delta s(t,m). \tag{5.2.26}$$

Rearranging eq. 5.2.26, we obtain,

$$\delta\mathcal{H}\left\{s(t,m)\right\} = -\tan[\delta\phi_r(t,m)]\delta s(t,m). \tag{5.2.27}$$

We can expand the tangent function into power series and neglecting the higher order terms, we are left with,

$$\delta\mathcal{H}\left\{s(t,m)\right\} = -\delta\phi_r(t,m)\delta s(t,m) \tag{5.2.28}$$

If we substitute eq. 5.2.28 into 5.2.23, $\delta\phi_r$ becomes,

$$\delta\phi_r = \frac{\left[\mathcal{H}\left\{s(t,m)\right\} + s(t,m)\delta\phi(t,m)\right]\delta s(t,m)}{s(t,m)^2 + \mathcal{H}\left\{s(t,m)\right\}^2}. \tag{5.2.29}$$

We can neglect $s(t,m)\delta\phi_r(t,m)\delta s(t,m)$ since it is a second order term, eq. 5.2.29 then becomes, to first order,

$$\delta\phi_r = \frac{\mathcal{H}\left\{s(t,m)\right\}\delta s(t,m)}{s(t,m)^2 + \mathcal{H}\left\{s(t,m)\right\}^2}. \tag{5.2.30}$$

Now we can rewrite the gradient of the misfit function by inserting eq. 5.2.30 into 5.2.19,

$$\delta\chi = \sum_{r=1}^{N}\int_0^T [\phi_r(t,m) - \phi_r^{obs}(t)]\frac{\mathcal{H}\left\{s(t,m)\right\}\delta s(t,m)}{s(t,m)^2 + \mathcal{H}\left\{s(t,m)\right\}^2}dt \tag{5.2.31}$$

where we can immediately define the *adjoint source* as,

$$\begin{aligned} f_i^\dagger(\mathbf{x_r},t,\mathbf{m}) &= \sum_{r=1}^{N}\left[\phi_{r_i}(\mathbf{x_r},t,\mathbf{m}) - \phi_{r_i}^{obs}(\mathbf{x_r},t)\right] \\ &\quad \frac{\mathcal{H}\left\{s_i(\mathbf{x_r},t,\mathbf{m})\right\}}{s_i(\mathbf{x_r},t,\mathbf{m})^2 + \mathcal{H}\left\{s_i(\mathbf{x_r},t,\mathbf{m})\right\}^2}\delta(\mathbf{x}-\mathbf{x_r}). \end{aligned} \tag{5.2.32}$$

The adjoint source is the instantaneous phase difference multiplied by a weighting function (the second term in the equation). If there is no phase difference, the adjoint source will automatically be zero. The weighting

function equalizes the amplitude of the signal. The denominator, which is the square of the envelope of the synthetic data, weights the adjoint source in a way that the maximum amplitude parts of the signal will interfere with a low amplitude adjoint source and vice versa. This makes the adjoint kernel independent of the signal amplitude.

To have a better insight about the concept of instantaneous phase, let us consider two hypothetic Gaussian signals (Figure 5.2.1). The two signals are the same except that there is 4 s time-shift between them. The instantaneous phase of each signal monotonically varies from $-\pi/2$ to $\pi/2$. Their difference is symmetric and is maximum where the two signals cross. The weighting function has its minimum value at the maximum value of the synthetic signal due to the normalization with the square of the envelope.

5.2.3 Generalized cross-correlation measurements

We define the cross-correlation misfit function as,

$$\chi(m) = \frac{1}{2}\sum_{r=1}^{N}\int_{T_1}^{T_2}\left\|\Gamma_r^s(\mathbf{x_r},t,\mathbf{m}) - \Gamma_r^d(\mathbf{x},t)\right\|^2 dt \tag{5.2.33}$$

where,

$$\Gamma_r^s(t) = \frac{1}{N_d}\int_0^T s_i(\tau)d_i(t+\tau)d\tau \tag{5.2.34}$$

$$\Gamma_r^d(t) = \frac{1}{N_d}\int_0^T d_i(\tau)d_i(t+\tau)d\tau \tag{5.2.35}$$

are the cross-correlation between the synthetic and observed data, and the auto-correlation of the data, respectively. N_d is the normalization factor given by,

$$N_d = \int_0^T d(t)^2 dt. \tag{5.2.36}$$

The gradient of the misfit function then becomes,

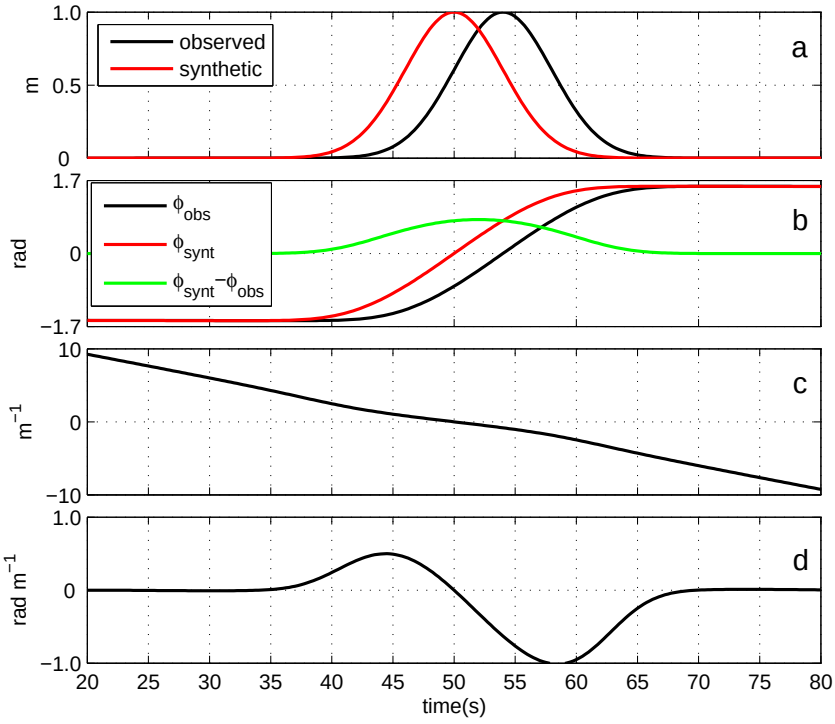

Figure 5.2.1: An example for instantaneous phase measurements for the two Gaussian signals shown in (a). In (b), instantaneous phases of each signal (red and black) and their difference (green); in (c), the weighting function; and in (d), the adjoint source which is obtained by the multiplication of the phase difference in (b) and the weighting function in (c) are presented.

$$\delta\chi = \sum_{r=1}^{N} \int_{T_1}^{T_2} \left[\Gamma_r^s(\mathbf{x_r}, \tau, \mathbf{m}) - \Gamma_r^d(\mathbf{x}, \tau)\right] \delta\Gamma_r^s(\mathbf{x}, \tau, \mathbf{m}) d\tau \tag{5.2.37}$$

and the perturbations in cross-correlation function is,

$$\delta\Gamma_r^s(\tau) = \frac{1}{N_d} \int_0^T d_i(t+\tau)\delta s(t) dt \tag{5.2.38}$$

where $\delta s(t)$ denotes the perturbations in synthetic seismograms due to perturbations in model parameters δm. If we substitute eq. 5.2.38 into 5.2.37,

$$\delta\chi = \frac{1}{N_d} \sum_{r=1}^{N} \int_0^T \int_{T_1}^{T_2} \left[\Gamma_r^s(\mathbf{x_r}, \tau, \mathbf{m}) - \Gamma_r^d(\mathbf{x_r}, \tau)\right] d_i(t+\tau)\delta s_i(t, \mathbf{m}) d\tau dt. \tag{5.2.39}$$

The *cross-correlation adjoint source* then is,

$$\begin{aligned} f_i^\dagger(\mathbf{x_r}, t, \mathbf{m}) &= \frac{1}{N_d} \sum_{r=1}^{N} \int_{T_1}^{T_2} \left[\Gamma_{r_i}^s(\mathbf{x_r}, \tau, \mathbf{m}) - \Gamma_{r_i}^d(\mathbf{x_r}, \tau)\right] \\ &\quad d_i(T - t + \tau)\delta(\mathbf{x} - \mathbf{x_r}) d\tau. \end{aligned} \tag{5.2.40}$$

The cross-correlation measurements are simply based on the difference between the cross-correlation of observed and synthetic data, and the auto-correlation of observed data, convolved with the observed data themselves. Let us again consider the same Gaussian signals as discussed in the previous section. Since our observed and synthetic data have the same waveform, Γ_d and Γ_s also have similar waveforms but Γ_s is shifted 4 s towards positive side (Figure 5.2.2). Thus the time difference between their pick values gives the time-shift between the signals. However, the adjoint source does not only consider the time-shift at the peak values as in travel-time measurements and is therefore a generalization of the latter. Due to the convolution of Γ_s-Γ_d with the observed data, the adjoint source will be dependent on amplitude.

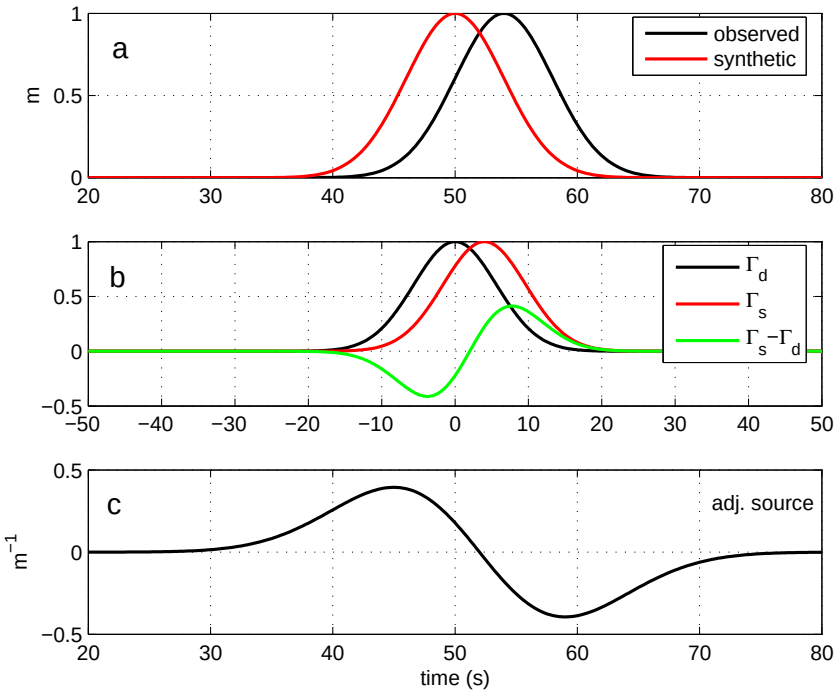

Figure 5.2.2: An example for cross-correlation measurements for the two Gaussian signals shown in (a). In (b), cross-correlation of the Gaussian signals (red), auto-correlation of the observed data (black) and the difference between them (green); and in (c), the adjoint source which is obtained by the convolution of the difference between cross-correlation and auto-correlation with observed data are presented.

5.2.4 Waveform measurements

A common way of incorporating amplitude together with phase information in seismic tomography is done by defining the *waveform misfit function* (e.g., Tarantola, 1984, 1987, 1988; Nolet, 1987) as,

$$\chi(m) = \frac{1}{2}\sum_{r=1}^{N}\int_0^T \|\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) - \mathbf{d}(\mathbf{x_r}, t)\|^2 \, dt \tag{5.2.41}$$

where $\mathbf{s}$ and $\mathbf{d}$ denote the synthetic and the observed data, respectively. We closely follow the notation of Tromp et al. (2005). The gradient of the misfit function is then,

$$\delta\chi = \sum_{r=1}^{N}\int_0^T [\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) - \mathbf{d}(\mathbf{x_r}, \mathbf{t})]\, \delta\mathbf{s}(\mathbf{x_r}, t, \mathbf{m}) dt \tag{5.2.42}$$

where $\delta\mathbf{s}$ is the perturbation in the displacement field $\mathbf{s}$ due to perturbations in model parameters δm. Thus the *waveform adjoint source* is defined as,

$$f_i^\dagger(\mathbf{x}, t, \mathbf{m}) = \sum_{r=1}^{N}[s_i(\mathbf{x_r}, T-t, \mathbf{m}) - d_i(\mathbf{x_r}, T-t)]\delta(\mathbf{x} - \mathbf{x_r}). \tag{5.2.43}$$

5.2.5 Travel-time measurements

Travel-time misfit is a widely-used measurement technique in the literature and is defined as the travel-time difference between synthetic and observed data given as,

$$\chi(m) = \frac{1}{2}\sum_{r=1}^{N}[T_r(m) - T_r^{obs}]^2 \tag{5.2.44}$$

where $T_r(m)$ is the predicted travel-time of a selected phase at the receiver r and T_r^{obs} is the corresponding observed travel-time. We again follow the notation of Tromp et al. (2005). The gradient of the travel-time misfit function is,

$$\delta\chi = \sum_{r=1}^{N} [T_r(m) - T_r^{obs}]\delta T_r \tag{5.2.45}$$

where δT_r is the perturbations due to model perturbations δm and may be written in the form of cross-correlation between observed and synthetic waveform data (e.g., Luo & Schuster, 1991; Marquering et al., 1999; Dahlen et al., 2000),

$$\delta T_r = \frac{1}{N_r} \int_0^T w_r(t)\partial_t s_i(\mathbf{x_r}, t, \mathbf{x_r})\delta s_i(\mathbf{x_r}, t, \mathbf{x_r})dt \tag{5.2.46}$$

where N_r is the normalization factor given by,

$$N_r = \int_0^N w_r(t)s_i(\mathbf{x_r}, t, \mathbf{m})\partial_t^2 s_i(\mathbf{x_r}, t, \mathbf{m})dt \tag{5.2.47}$$

and $w_r(t)$ is an appropriate window which isolates a specific phase. Inserting eq. 5.2.46 into eq. 5.2.45, the *travel-time adjoint source* becomes,

$$f_i^\dagger(\mathbf{x}, t, \mathbf{m}) = \sum_{r=1}^{N} [T_r(m) - T_r^{obs}] \frac{1}{N_r} w_r(T-t)\partial_t s_i(\mathbf{x_r}, T-t, \mathbf{m})\delta(\mathbf{x_r} - \mathbf{x}). \tag{5.2.48}$$

5.2.6 Comparison of adjoint sources

The Fréchet kernels depend on the misfit function via the adjoint source. It is therefore instructive to look at the adjoint sources. In Figure 5.2.3, the adjoint sources are shown for the misfit functions defined as the instantaneous phase difference (IP), cross-correlation difference (CC), waveform difference (WF) and travel-time difference (TT) using the two Gaussian signals considered in the previous sections (see Figures 5.2.1, 5.2.2). For single phase waveforms all misfit functions give similar adjoint sources although not identical. The width of the adjoint sources change, WF is the narrowest and CC is the widest. The CC measurements can be thought as an extension to the TT measurements. Among all the misfit functions we consider here, IP and TT do not depend on amplitude but the latter is only defined for a well isolated phase and not applicable to a complicated waveform.

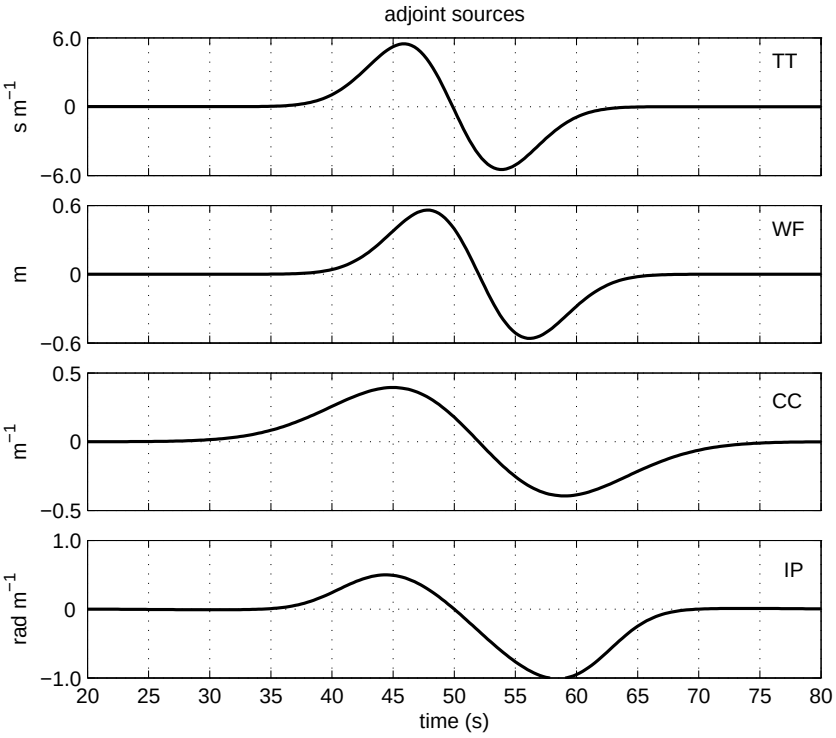

Figure 5.2.3: Adjoint sources computed for the Gaussian signals in Figures 5.2.1 and 5.2.2 using travel-time (TT), waveform (WF), cross-correlation (CC) and instantaneous phase (IP) measurements

.

5.3 Results

We considered the misfit functions based on TT, WF, CC and IP measurements for different waveforms and computed the corresponding sensitivities using adjoint simulations based upon the 3-D numerical wave simulation by the spectral element method (SEM) of Komatitsch & Tromp (2002a,b) for selected ray paths. To have a better understanding on how selected misfit functions behave, we did purely synthetic experiments. We first used 1-D isotropic PREM and 3-D mantle model S20RTS with 3-D crustal model Crust2.0 (S20RTS+Crust2.0) to be used as *synthetic* and *observed* data, respectively. We used an explosive source to make the kernels simple and present the results for the vertical component. We then fixed the crust and upper mantle structures in both sets of SEM seismograms. We used S20RTS+Crust2.0 to compute the *observed* seismograms. For the *synthetic* seismograms, the crust and uppermantle is S20RTS+Crust2.0 and the lower mantle is PREM (S20RTS$_{um}$+Crust2.0). In this case, we used the original CMT solution of a real earthquake and present results for the transverse component.

5.3.1 PREM vs. S20RTS+Crust2.0

We computed seismograms in 1-D isotropic PREM and in the 3-D mantle S20RTS with 3-D crust Crust2.0 on top (S20RTS+Crust2.0) which play the role of *synthetic* and *observed* data, respectively. We modified the CMT solution of the Rat Islands earthquake (March 17, 2003, $M_w = 7.0$) to an explosive source in order to eliminate the directionality of the source and to simplify the content of the kernels. We present the results for the wavefield recorded at the station DBO ($\Delta = 40^\circ$) (path A in Figure 5.3.1).

Before computing the adjoint sources, we applied a band-pass filter to the SEM seismograms between 42 s and 500 s. The seismograms with marked phases on are shown in Figure 5.3.2. The instantaneous phase shift difference clearly shows the difference in phase at each time step between these two seismograms. The maximum phase shift occurs for the surface waves and multiple surface reflections.

We applied cosine windows to extract individual parts of the seismograms to understand the properties of the different misfit functions. We then computed the adjoint kernels considering the full seismograms using body and surface waves together. We qualitatively compared the computed finite-frequency adjoint kernels from different misfit functions. Note that,

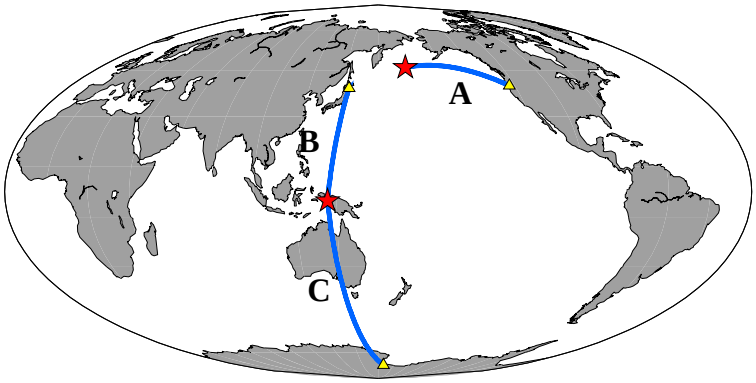

Figure 5.3.1: The ray paths of the seismograms used in this study. The path A corresponds to the Rat Island earthquake (March 17 2003, $M_w = 7.0$, depth=27 km) recorded at the station DBO. Note that for the synthetic experiment in Section 5.3.1, the CMT solution of the earthquake was modified to an explosive source. The paths B and C correspond to the Irian Jaya earthquake (February 2 2004, $M_w = 7.0$, depth=13 km) recorded at the stations ERM and CASY, respectively.

for the P-wavetrain, surface waves and full seismograms, we did not use TT measurements since they are not designed to be used for multiple phases.

P-waves: There is approximately a 2.5 s difference between the observed and the synthetic P-waves as shown in Figure 5.3.3. The filtered adjoint sources between 42s-500 s, which is the frequency band of our seismograms, for TT, WF, CC, and IP measurements are also shown in Figure 5.3.3. The adjoint sources are not identical, however, they are similar in shape with the biggest difference for IP measurements. These similarities can also be seen on the adjoint kernels (Figure 5.3.4). The adjoint kernel from TT measurements shows a typical banana-doughnut shape representing the P-wave sensitivity in the mid-mantle. WF and CC kernels are similar whereas the IP kernel has a less pronounced hole. Due to our frequency content, the kernels are rather large.

P-wavetrains: Here we set the time window between 350 s and 650 s where we observe multiple P-wave arrivals, P, pP, PP, PcP (Figure 5.3.5). All misfit functions have now different adjoint sources. WF and CC pick out P and PP-waves in the adjoint sources and WF shows pronounced oscillations for the whole signal due to amplitude changes. IP emphasizes the end of the P-wave. The adjoint kernels for WF and CC show a mixture of P and PP sensitivities where the WF kernel is more oscillatory as can be

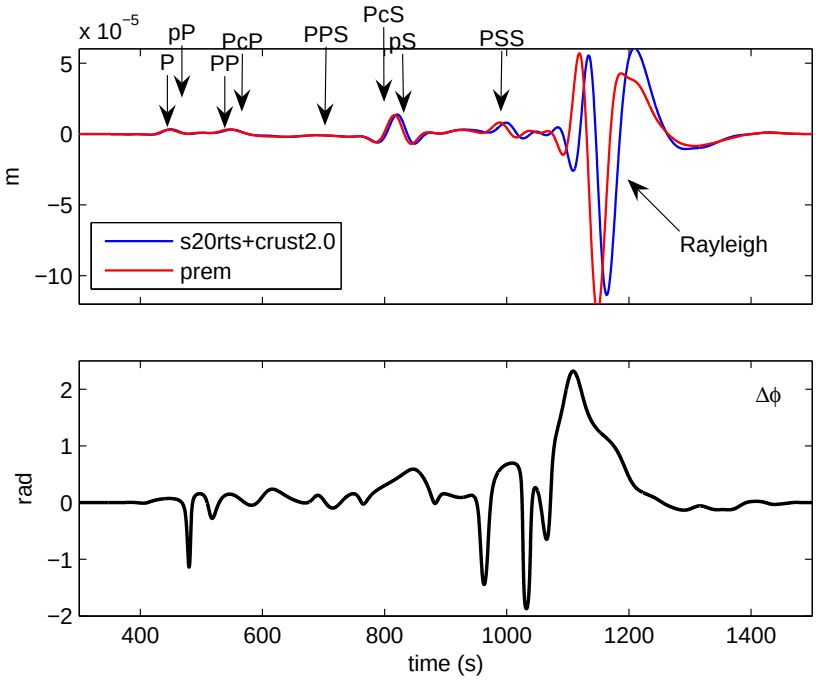

Figure 5.3.2: Vertical component SEM seismograms computed for PREM and S20RTS+Crust2.0 models which play the role of *synthetic* and *real* seismograms, respectively, with the marked seismic phases on top (upper plot), and the instantaneous phase difference between them as a function of time (lower plot). The seismograms are from the path A presented in Figure 5.3.1.

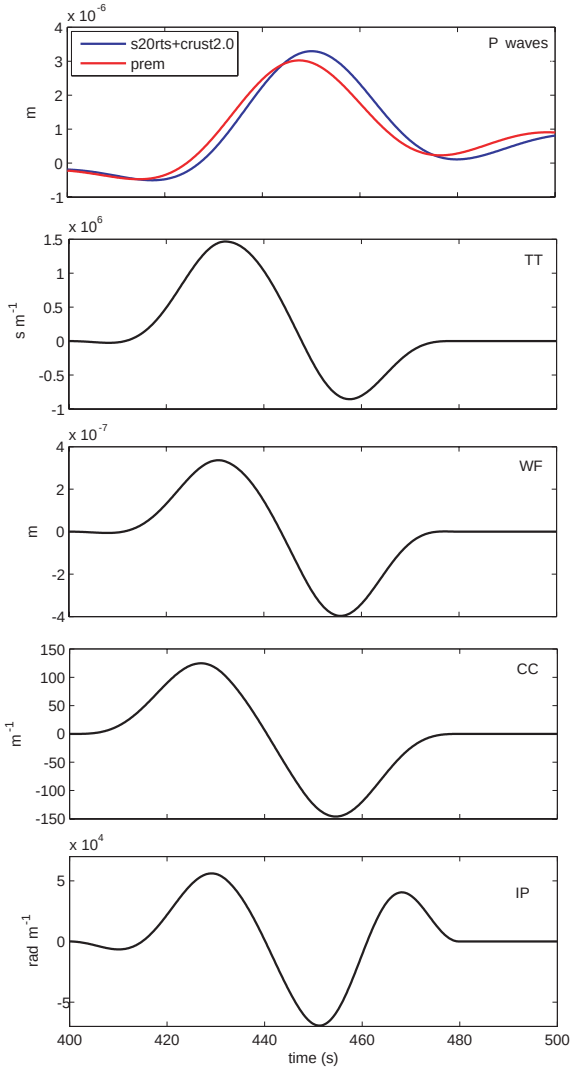

Figure 5.3.3: Adjoint sources computed for the P-waves (top plot) from the seismograms presented in Figure 5.3.2 using travel-time (TT), waveform (WF), cross-correlation (CC) and instantaneous phase (IP) measurements.

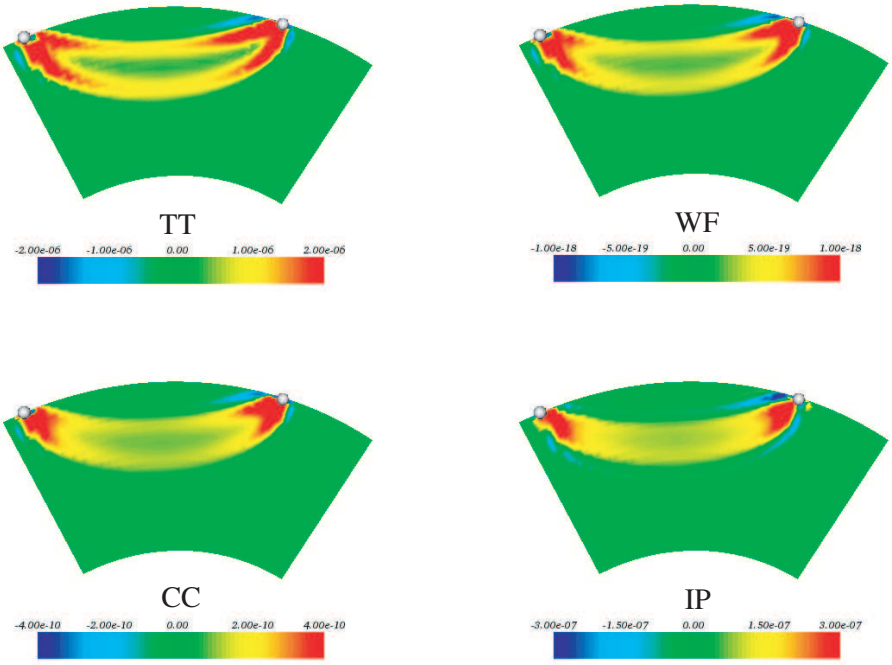

Figure 5.3.4: Source-receiver cross-sections of the K_α finite-frequency adjoint kernels for P-waves computed by the adjoint sources presented in Figure 5.3.3. The units of the kernels are; s^2/m^3 for travel-time misfit (TT), s/m for waveform misfit (WF), s/m^3 for cross-correlation misfit (CC), and $rad.s/m^3$ for instantaneous phase misfit (IP).

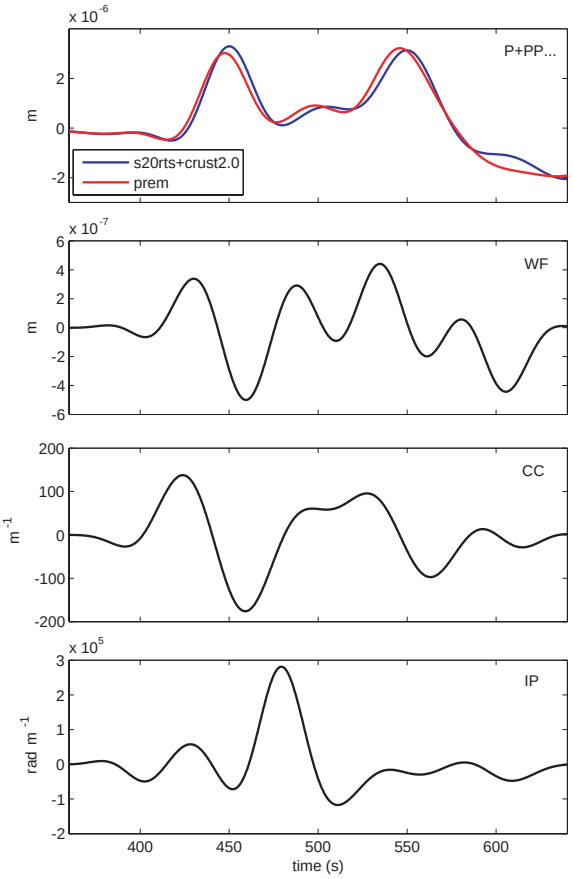

Figure 5.3.5: The same as Figure 5.3.3 but for P-wavetrain (P+PP..) and without TT measurements.

expected from the adjoint source (Figure 5.3.6). In the IP kernel, P-wave sensitivity is dominant and has nicely developed higher Fresnel zones. In all kernels, PcP sensitivity is very small compared to the other phases.

Rayleigh waves and the overtones: We kept the time window large enough to capture not only the fundamental mode but also overtones and multiple surface reflections such as the PSS phase. Since the adjoint source from IP measurements equally weights all the seismic phases, low amplitude parts of the signal has the highest amplitude in the adjoint source (overtones at 1050 s and long period surface waves at 1400 s, Figure 5.3.7). WF and CC measurements favor the high amplitude parts as expected. The misfit kernels are of course all sensitive to the uppermost part of the Earth's structure but in different ways. When we look at the sensitivities at the surface, we observe that WF and CC kernels are similar to amplitude kernels and more sensitive to the fundamental mode (Figure 5.3.8). The IP kernel is narrower due to the interaction of the fundamental mode with surface reflections.

Full seismograms: For a single phase or a small portion of the seismogram, the kernels for different misfit functions are quite similar except for small scale structures and the amplitude of the kernels. The interesting question is how these different measurement techniques behave when we consider all phases together, i.e., the entire seismogram. In Figure 5.3.9, we present the adjoint sources for WF, CC and IP measurements. Since WF and CC misfit functions are amplitude dependent, they immediately pick out the high amplitude parts of the signal which are the surface waves. Thus, they focus on surface structures (Figure 5.3.10). However, IP measurements provide additional information. The highest sensitivity is for surface waves since they present a large phase shift due to the velocity models we used to compute the synthetics. However, IP measurements also show P-wave sensitivity, which is less than surface waves, not because of the amplitude but because the phase shift is smaller for P-waves (see Figure 5.3.2). WF and CC kernels have no P-wave sensitivity although there are differences in the P-waves.

5.3.2 Fixed upper mantle and crust

The maximum velocity heterogeneity occurs in the uppermost part of the Earth and therefore we get the maximum phase shift for surface waves in shallow earthquake seismograms. In order to investigate the effect of amplitude variations in seismograms further, we fixed the uppermantle and

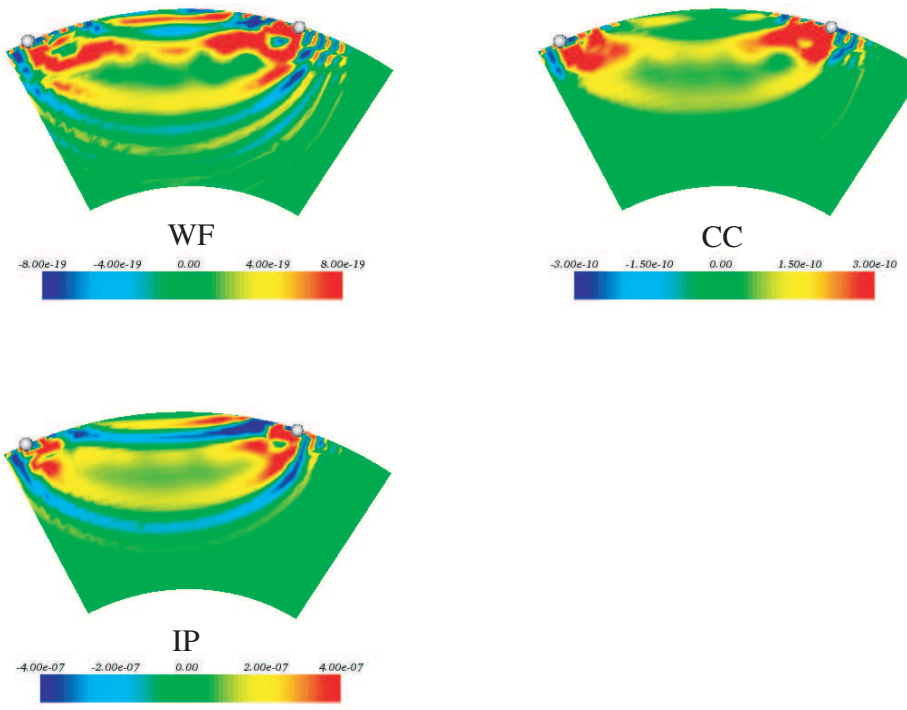

Figure 5.3.6: Source-receiver cross-sections of the K_α finite-frequency adjoint kernels for P-wavetrain computed by the adjoint sources presented in Figure 5.3.5. The units of the kernels are; s/m for waveform misfit (WF), s/m^3 for cross-correlation misfit (CC), and $rad.s/m^3$ for instantaneous phase misfit (IP).

crustal part of the velocity models in both sets of seismograms. The SEM seismograms computed in S20RTS+Crust2.0 and S20RTS$_{um}$+Crust2.0 are *observed* and *synthetic* data, respectively. We computed the synthetics for the Irian Jaya earthquake (5 February, 2004, $M_w = 7.0$) and present the results for the wavefield recorded at the station ERM ($\Delta = 46°$) (path B in Figure 5.3.1).

The *observed* and the *synthetic* seismograms and the instantaneous phase difference between the seismograms are shown in Figure 5.3.11. The surface waves in both SEM seismograms are almost the same with some minor differences and S, SS and ScS phases differ due to their sensitivities to deeper structures. In Figure 5.3.12, the adjoint sources obtained from WF, CC and IP measurements are shown filtered between 42-500 s

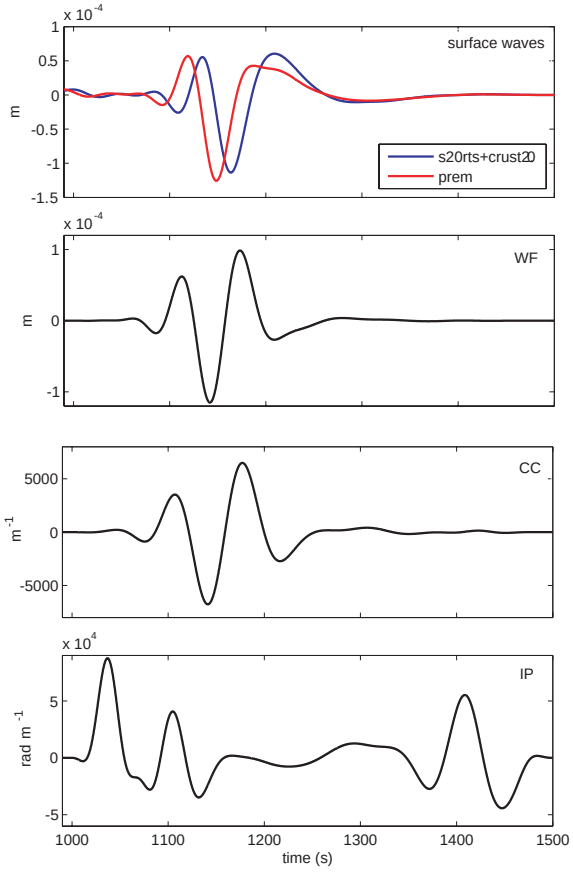

Figure 5.3.7: The same as Figure 5.3.3 but for surface waves and without TT measurements.

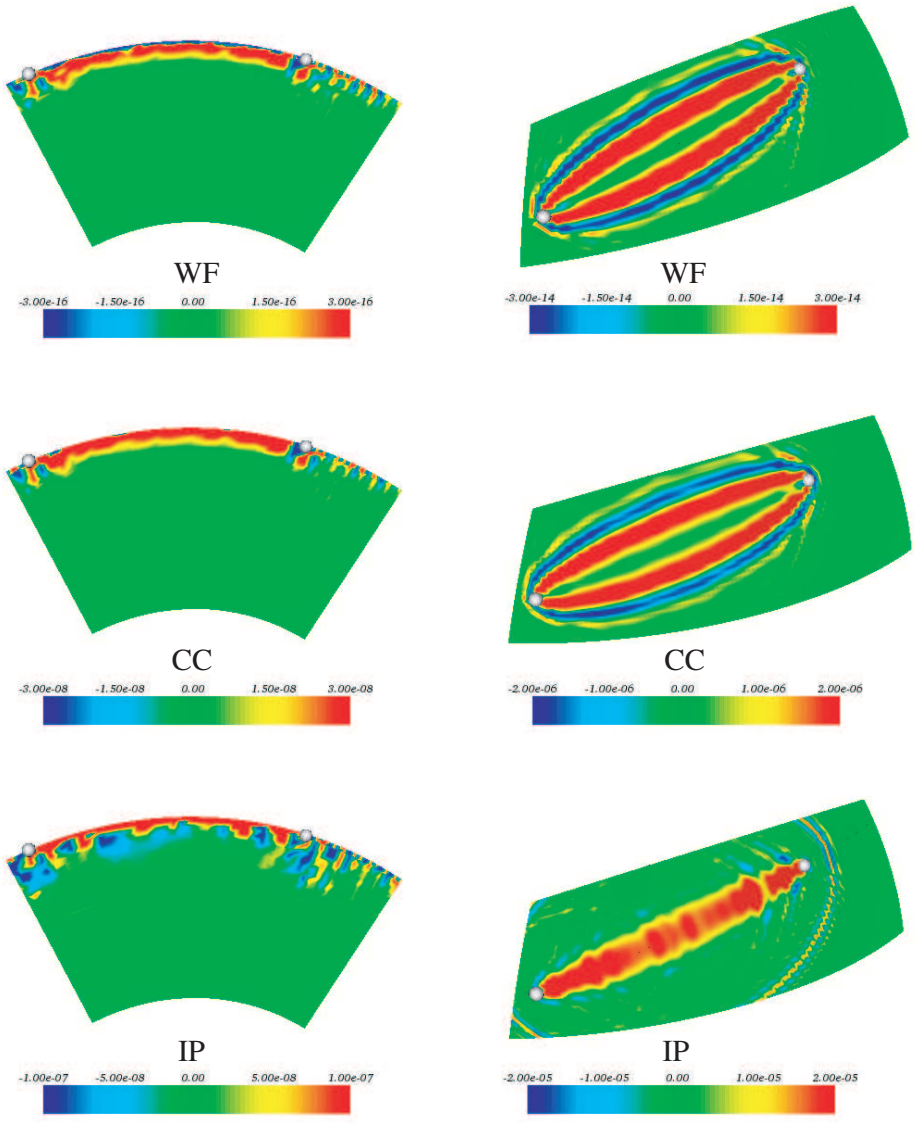

Figure 5.3.8: Source-receiver cross-sections (left column) and horizontal slices at the surface (right column) of the K_α finite-frequency adjoint kernels for surface waves computed by the adjoint sources presented in Figure 5.3.7. The units of the kernels are; s/m for waveform misfit (WF), s/m^3 for cross-correlation misfit (CC), and $rad.s/m^3$ for instantaneous phase misfit (IP).

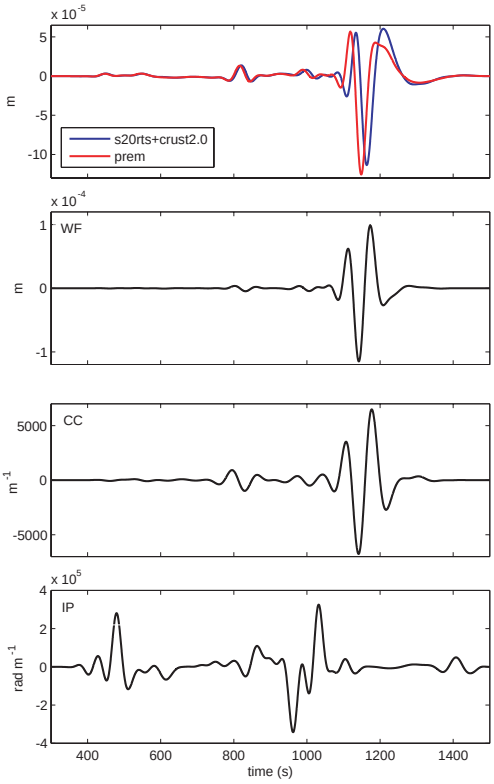

Figure 5.3.9: The same as Figure 5.3.3 but for the full seismograms and without TT measurements.

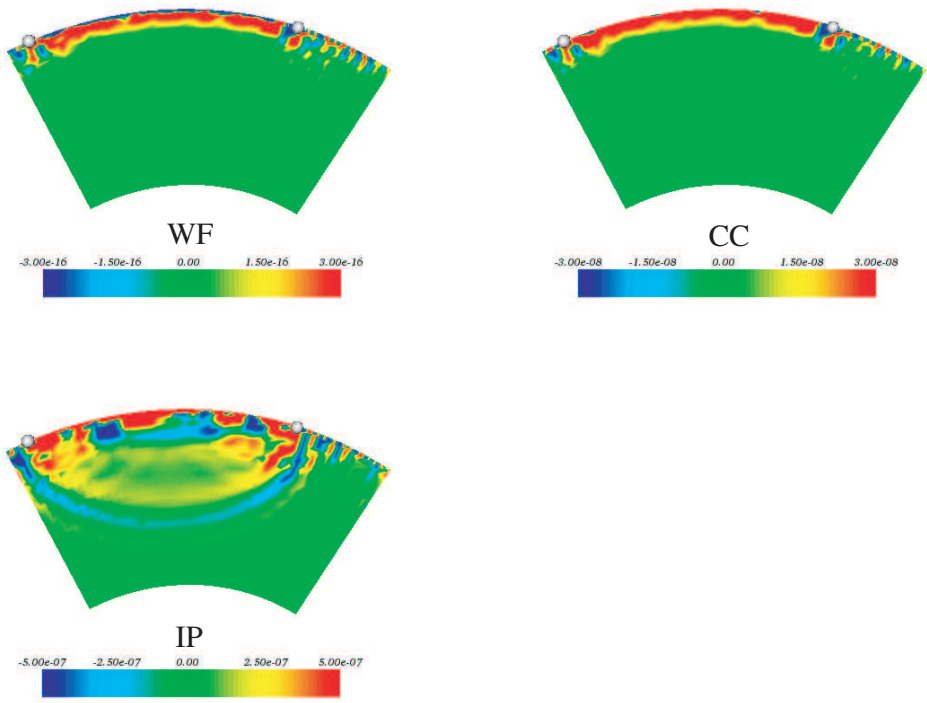

Figure 5.3.10: Source-receiver cross-sections of the K_α finite-frequency adjoint kernels for full seismograms computed by the adjoint sources presented in Figure 5.3.9. The units of the kernels are; s/m for waveform misfit (WF), s/m^3 for cross-correlation misfit (CC), and $rad.s/m^3$ for instantaneous phase misfit (IP).

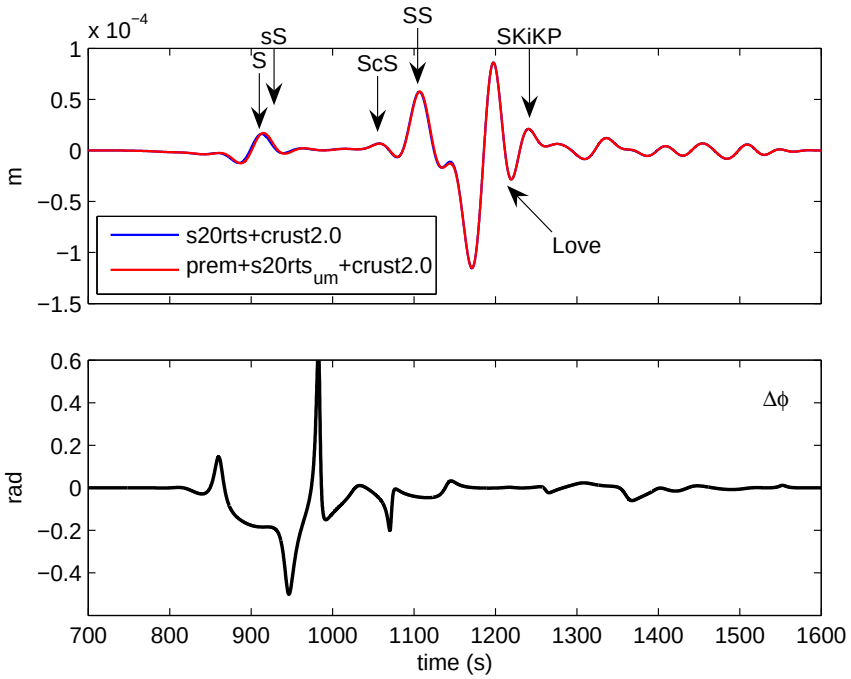

Figure 5.3.11: Transverse component SEM seismograms computed for S20RTS$_{um}$+Crust2.0 and S20RTS+Crust2.0 models which play the role of *synthetic* and *real* seismograms, respectively, with the marked seismic phases on top (upper plot), and the instantaneous phase difference between them as a function of time (lower plot). The seismograms are from the path B presented in Figure 5.3.1.

in the same way as the seismograms. WF and CC measurements naturally favor the high amplitude parts of the data. Due to the convolution with the *observed* data, the CC adjoint source shows more oscillations. IP emphasis the parts where it detects phase differences at the end of S and at the beginning od SS, normalizing the amplitudes of different phases in a single seismogram. The difference between amplitude dependent WF and CC with IP is very clear on their sensitivity kernels (Figure 5.3.13). WF and CC measurements are most sensitive to SS phase where IP is almost equally sensitive to S and SS.

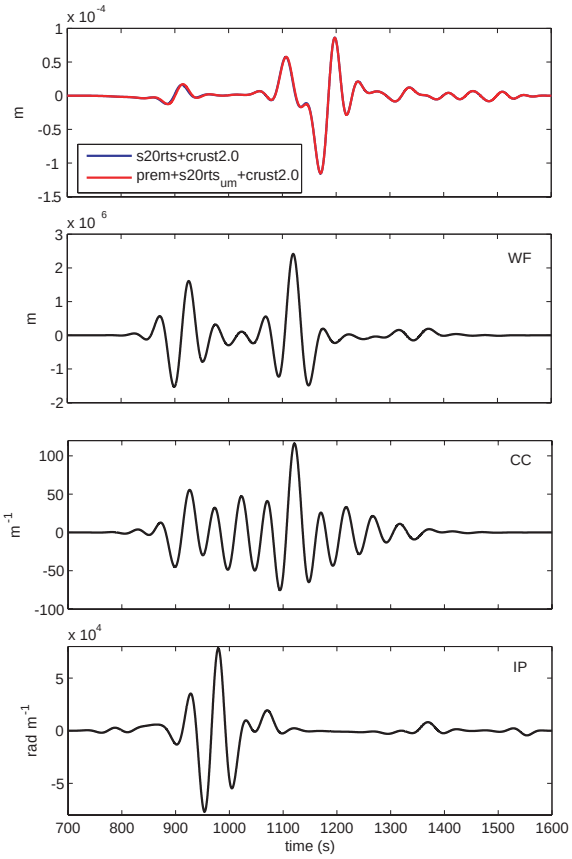

Figure 5.3.12: Adjoint sources computed for the full seismograms (top plot) also presented in Figure 5.3.11, using waveform (WF), cross-correlation (CC) and instantaneous phase (IP) measurements.

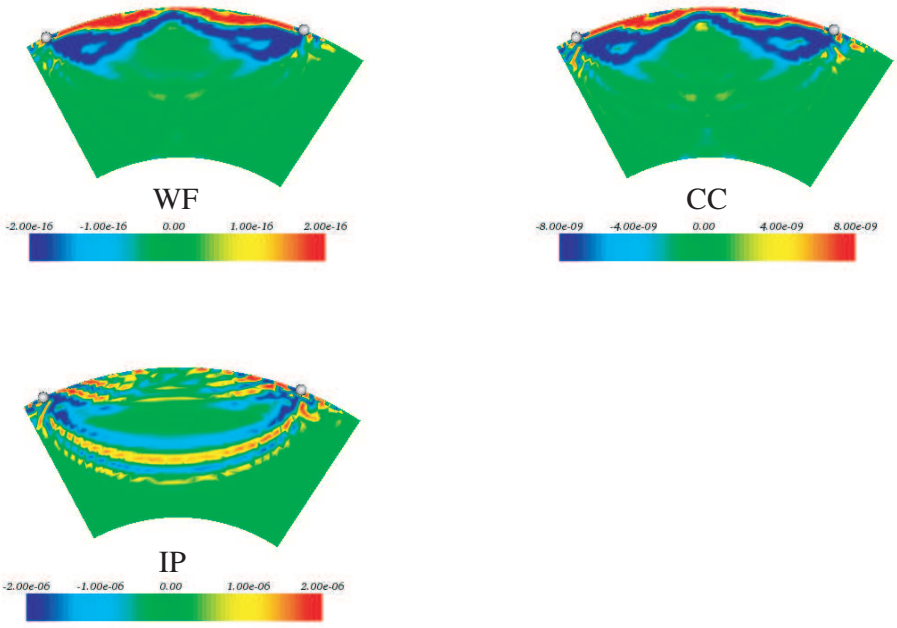

Figure 5.3.13: Source-receiver cross-sections of the K_β finite-frequency adjoint kernels for full seismograms computed by the adjoint sources presented in Figure 5.3.12. The units of the kernels are; s/m for waveform misfit (WF), s/m^3 for cross-correlation misfit (CC), and $rad.s/m^3$ for instantaneous phase misfit (IP).

5.4 Discussions

The chosen misfit function determines the information we can get out of the data. Considering single phases, although not the same, all corresponding kernels have similar characteristics. For longer waveforms, WF and CC measurements have similar kernels, but IP measurements differ because they are completely independent from amplitude information.

In case of purely synthetic seismograms, IP measurements, for the whole seismograms, have well behaved properties. However, in practice, their use might not be straightforward. IP measurements are amplitude independent, a reliable estimation of the phase requires a minimum level of amplitude present in the signal. We are in the regime of the Born approximation thus we assume that the phase differences are within one cycle at most. Noise content of signals could effect the low amplitude part of the signal and randomize the phase. We are going to discuss these issues in detail below.

5.4.1 Phase jumps

In IP measurements, the phase difference is measured between $\pm\pi$ rad. Thus, cycle slips occur whenever there is more than one cycle difference between observed and synthetic data. We present such an example in Figure 5.4.1 (the seismograms are from the path C in Figure 5.3.1). We observe phase jumps between the two signals around 1800 s and right before surface waves around 2050 s. The cycle slips are likely to occur when one seismogram has completed one cycle and the other one has not, which is controlled by the zero crossings. Cycle slips can also occur whenever the two phases are very much out of phase. Since we are working within the limits of Born theory, our observed and synthetic data should differ by $\pm\pi$ at most. It is easy to detect such cycle slips in the instantaneous phase difference and to eliminate these parts of the seismograms by windowing similar to Fichtner et al. (2008). In order to avoid cycle slip problem in phase measurements, Ekström et al. (1997) used long period waveforms first and then gradually increased the frequency content of the data in the next iterations of the inversion. Since the phase jumps should be less at long periods (see Figure 5.4.2), a similar approach can be used in our case as well.

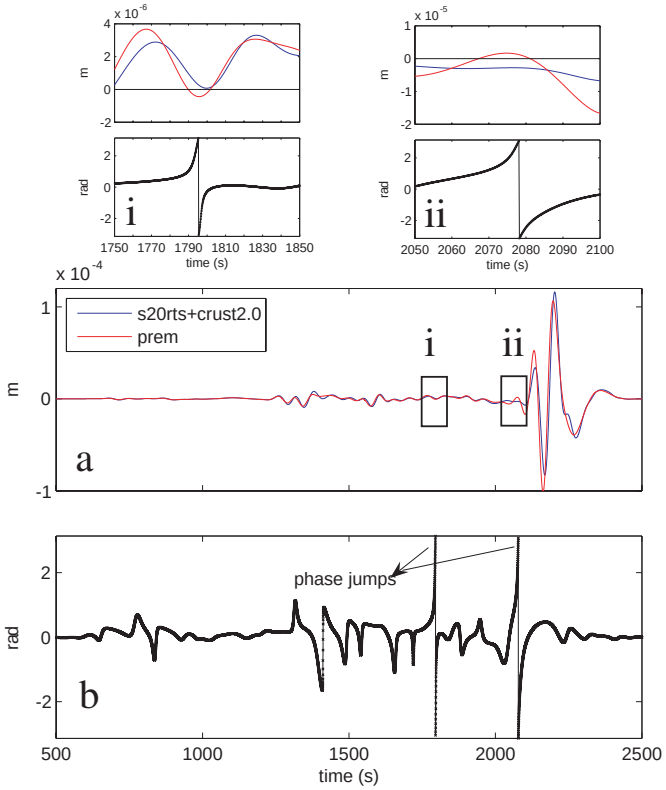

Figure 5.4.1: An example for phase jumps occur in instantaneous phase measurements. a) SEM seismograms from the path C presented in Figure 5.3.1 were bandpass filtered between 40 s and 500 s, b) instantaneous phase difference between the seismograms. In (i) and (ii), the parts of the seismograms where the phase jumps are observed are enlarged.

5.4.2 The effect of noise on simulations

All experiments so far have been based on noise free synthetic experiments. An important issue, however in real cases is the noise content of the data. The effect of noise should be small in the case of WF and CC measurements since they favor the high amplitude parts of the data. IP measurements are amplitude independent and every wiggle has equal weight, thus this could potentially increase the noise contribution in the kernels compared to the amplitude dependent misfit functions. The best solution is of course to use high quality seismograms. One obvious strategy is to isolate parts of the

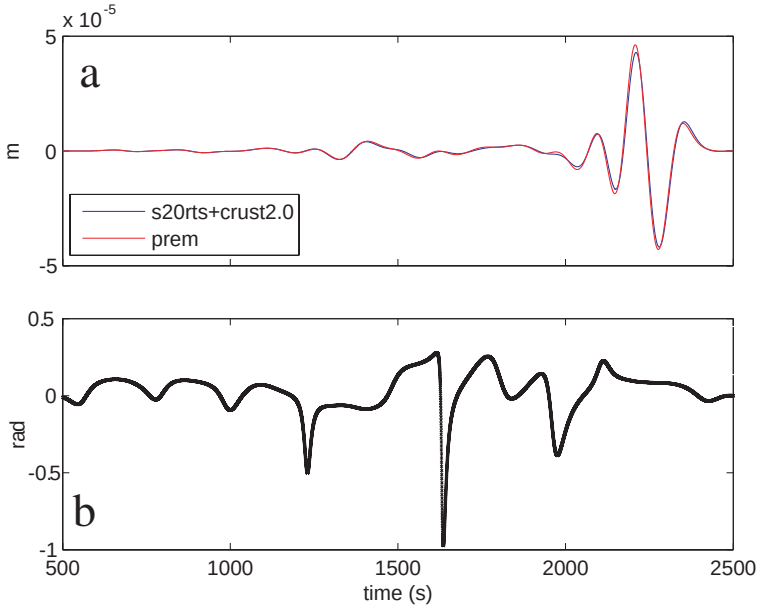

Figure 5.4.2: The same as Figure 5.4.1 but the seismograms were bandpass filtered between 100 s and 500 s.

signal with a high signal-to-noise ratio (e.g., Maggi et al., 2009).

In Figures 5.4.3 and 5.4.4, we did some tests with noisy data, to see how the IP measurements would be sensitive to the noise content. We used similar seismograms as presented in Section 5.3.1 and performed our experiments by adding different levels of noise to the S20RTS+Crust2.0 seismograms. In the first example (Figure 5.4.3), we produced uniformly distributed random numbers and scaled them with 50% of the maximum of the P-wave amplitude. In this case, the P-wave can still be identified and the noise level does not have a dramatic effect on the kernels. The sensitivity kernel is very similar to the one from noise free data. When we increase the noise level to be comparable with P-wave amplitude, we loose the P-wave sensitivity. However, it does not introduce any structure in the kernel which does not present in the data. This is probably because, random noise does not constructively interfere with the forward wavefield during the adjoint simulations.

In the second example we used similar seismograms as presented in Section 5.3.2, where we had fixed the uppermantle and crustal models in real and synthetic data. This time, random noise was scaled to 10% and

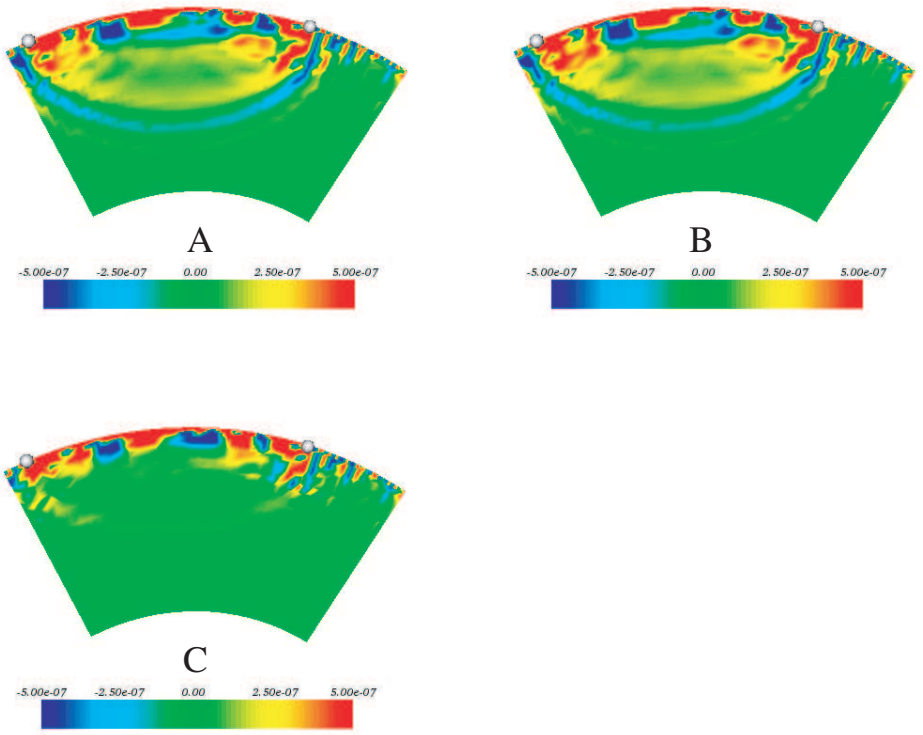

Figure 5.4.3: Noise analysis on sensitivity kernels computed for the full seismograms presented in Figure 5.3.2 using IP measurements. Uniformly distributed random noise was generated and scaled to 50% (B) and 100% (C) of the maximum of the P-wave and added to the observed data. In (A), the kernel from noise free data is presented.

50% of the maximum amplitude of S-wave. The multiply reflected waves and higher mode surface waves are partly to completely under the noise level. We observe that the higher the noise level, the stronger the amplitude of the S-kernels. The sensitivity kernel having 10% noise is similar to the one from noise free data. In the case of 50% noise, the clear separation of S and SS phases is less apperent. However, the major structure of the kernel remains the same as in the kernel from noise free data.

We observe that in the case of high signal to noise ratio, the sensitivity kernels are very similar to noise free kernels. When the noise level increases, interestingly the sensitivities do not change dramatically. The noisy part of the seismogram relative to the less noisy part looses sensitivity.

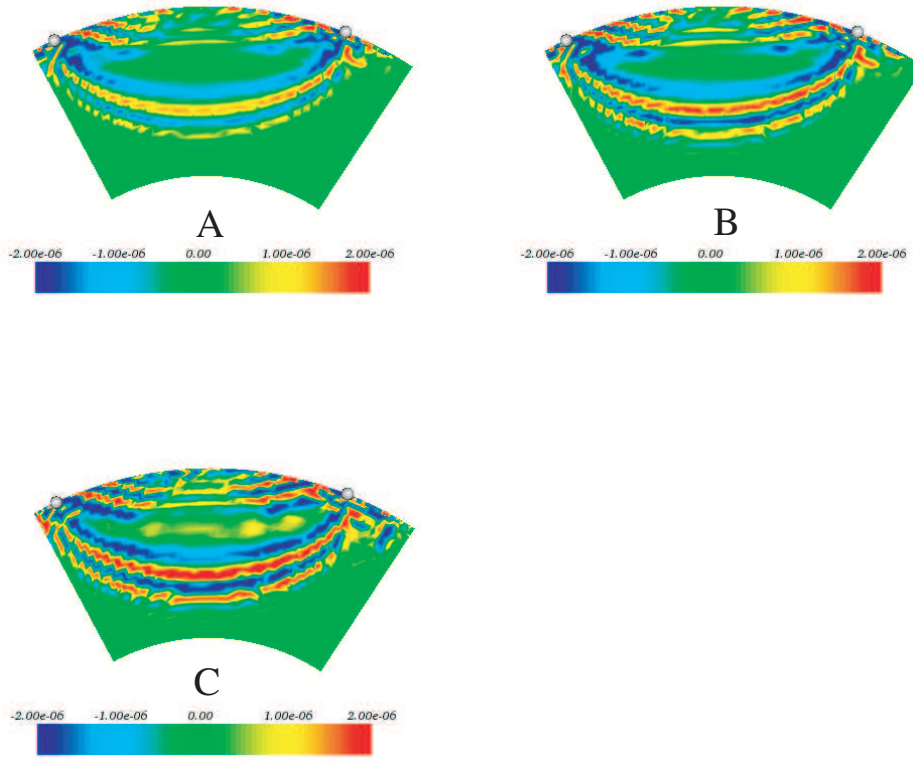

Figure 5.4.4: Noise analysis on sensitivity kernels computed for the full seismograms presented in Figure 5.3.11 using IP measurements. Uniformly distributed random noise was generated and scaled by 10% (B) and 50% (C) of the maximum of the S-wave and added to the observed data. In (A), the kernel from noise free data is presented.

5.4.3 Comparison of the misfit functions

For a single phase, all misfit functions give similar sensitivities. For longer waveforms, misfit functions have their own way of sensing the Earth's interior.

TT measurements are designed for single phases or very similar waveforms which have to be identified first. This gives a robust estimation of the phase shift and kernel (e.g, Luo & Schuster, 1991; Marquering et al., 1999; Dahlen et al., 2000; Tape et al., 2007). However, the drawback is that it can be time consuming to pick every phase in a seismogram. Recently, automatic phase picking algorithms have been released (e.g. Maggi et al., 2009) particularly to be used in adjoint tomography. General CC measurements suggest a way to look at the phase shifts of a group of phases together. However, since the adjoint source of the CC measurements is convolved with observed data, the CC measurements are amplitude dependent and behave like the WF measurements. This amplitude dependence of the misfit functions favors the higher amplitude parts in a wavetrain. This can be a disadvantage that we loose the contribution of the low amplitude phases in the sensitivity kernels. Amplitude dependence of the misfit functions, on the other hand, provide a good way of suppressing the noise contribution.

IP measurements suggest a good way to cope with strong amplitude differences in a seismogram and allow to treat even surface waves and body waves together by equally weighting every wiggle. It is similar to the approach of Fichtner et al. (2008) based on the time-frequency analysis of Kristekova et al. (2006) which also separates the phase and amplitude information. The advantages of IP are less data processing and easier implementation. It can also be easier to handle phase jumps in time domain.

5.5 Conclusions

We proposed new misfit functions based on the cross-correlation and instantaneous phase differences for seismic tomography. We determined the associated adjoint sources and examined their properties by computing finite-frequency adjoint kernels based upon the spectral element method. We also compared the newly introduced misfit functions with more commonly used waveform and cross-correlation travel-time misfit functions.

When we consider isolated phases, all misfit functions more or less give similar results. CC measurements are similar to WF measurements. They favor the high amplitude part of the seismograms, but are relatively in-

sensitive to noise. Instantaneous phase measurements provide an easy and efficient way to deal with different observables in a single seismogram. Unlike amplitude dependent misfit kernels, the sensitivity kernels have equally weighted contributions from all phases. Care has to be taken in the presence of phase jumps and noise. It is encouraging that a reasonable noise content does not have a dramatic effect on the kernels. On the other hand, phase jumps and noise level, if problematic, can be identified very easily by the instantaneous phase difference between the signals and envelope. The problematic parts in the seismograms can then easily be removed. IP provides complementary information to other amplitude dependent misfit functions in the finite-frequency adjoint kernels.

Acknowledgment

We would like to thank Anne Sieminski for her help with the adjoint simulations and the discussions on cross-correlation difference measurements. This project is funded by the European Commissions Human Resources and Mobility Program Marie Curie Research Training Network SPICE Contract No. MRTNCT-2003-504267. Numerical simulations were performed on a 128-processor cluster financed by the Dutch National Science Foundation under the grant number NWO:VICI865.03.007.

Chapter 6

Concluding remarks

We have reached the limits of ray theory in the current resolution of the seismic tomography images. The advances in the numerical simulation of wave propagation have created the opportunity to refine the tomographic models by taking into account the full non-linearity of the problem. Numerical simulation of the wave propagation in 3-D structures also provides an efficient way to assess the current tomographic images. This thesis is dedicated to underline some shortcomings of classical seismic tomography at the global scale using 3-D numerical wave propagation by a spectral element method (Komatitsch & Tromp, 2002a,b). Based on our observations, we suggest some possible solutions to improve the images. In addition, we propose new ways to exploit as much information as possible from a single seismogram by defining new misfit functions and identifying new usable data in the seismograms in combination with 3-D numerical simulations.

In **Chapter 3**, we analyzed the accuracy of crustal corrections on fundamental mode surface waves and their effect in global mantle tomography. We estimated crustal corrections using ray theoretical and 2-D finite frequency approximations and compared them to those obtained from 3-D wave simulation by a spectral element method. We observed that crustal effects can not be completely removed from seismograms and produce errors larger than those in phase velocity measurements at periods shorter than 60 and 80 s for Rayleigh and Love waves, respectively. As a result of different impact of crustal corrections on Rayleigh and Love waves, the inversion of residual time-shifts produce radial anisotropy in mantle comparable in strength to those of existing models at least in the uppermost mantle. Furthermore, we observed a sign change in anisotropy caused by

imperfect crustal corrections which is similar to what has been interpreted so far. As a consequence, our observations indicate that part of mantle anisotropy could be uncorrected crustal signal. This subject needs a thorough investigation in future work. Rayleigh and Love waves are affected by the crustal structure in a different way. The nonlinear influence of the crust on surface waves becomes more visible at longer source-receiver distances. Since the crust has a highly nonlinear behavior, extensions to great circle approximation such as exact ray theory or finite frequency theory do not improve the results suggesting that the higher order effects are important to describe the crust. The possible solutions are the use of either 3-D reference models, which is becoming available with numerical methods, or inverting mantle together with crust to take the nonlinearity of crust into account.

In **Chapter 4**, we quantitatively compared real seismograms with synthetic ones computed in various 3-D mantle models constructed by phase measurements by looking at phase and amplitude differences of surface waves and some body wave phases such as P, S, and SS waves, in order to understand how far current tomographic images represent the real Earth's mantle. In particular, we examined the effect of damping on our interpretation of the mantle by choosing differently damped mantle models. In addition, by comparing syntehtic seismograms computed in various 1-D and 3-D crustal and mantle models, we tried to understand the elastic and anelastic contribution of the models in tomography. It is surprising that although the differently damped mantle models are remarkably different from each other, from a statistical point of view, we cannot distinguish which model explains the mantle better than the others. Based on the results in the previous chapter, we think that this is likely due to large crustal signal remaining after imperfect crustal corrections. As expected, the 3-D tomographic models mainly improve the phases of the seismograms both in surface and body waves compared to 1-D models. However, there is still vast information in seismograms both in phase and amplitude to be explained. In our comparisons, SEM amplitudes are on average larger than the amplitudes of the real seismograms. Amplitudes can partly be explained by the scattering and source effects, but our results are likely pointing to an important contribution from anelastic properties of wave propagation. Attenuation seems more important for surface waves whereas body waves are more affected by focusing/defocusing effects.

Another important issue in global seismic tomography is the lack of data coverage due to uneven distribution of earthquakes and stations on the globe. Unlike the approximations to the wave equation that restrict the usable data, numerical methods offer an invaluable opportunity to model full seismograms. In **Chapter 5**, we defined new misfit functions for the full waveform inversion in order to extract as much information from a single seismogram as possible. We proposed two new misfit functions based on instantaneous phase and cross-correlation measurements to be used for full waveform inversions. In order to have a better insight into the properties of the different misfit functions, we compared them to existing ones such as waveform and travel-time measurements that are widely used in tomography. We qualitatively compared the new and existing misfit functions by computing finite frequency sensitivity kernels for a source-receiver pair using adjoint methods based upon a spectral element method. Cross-correlation mesurements are a kind of extension to classical travel-time measurements in terms of allowing to measure the phase or time shifts in a wavetrain. It provides a robust way of suppressing the noise due to the amplitude dependence of the misfit function. However, this gives the emphasis to high amplitude phases while restricting the contribution of small amplitude phases in the sensitivity kernels similar to waveform measurements. Thus, CC measurements should be used for wavetrains where the phases have comparable amplitudes in order not to loose information. Instantaneous phase measurements provide an effective way to measure the phase difference independently from amplitude information. Due to the natural normalization involved in the adjoint source, it is successful in weighting every observable in a seismogram which theoretically allows us to use a full seismogram. This may enhance the noise, however, our simulations show that noise does not have a dramatic effect on the sensitivity kernels. Another possible difficulty might occur due to phase jumps. These problems can be easily detected by examining the instantaneous phase difference. Possible solutions are to cut out the problematic parts of the seismograms or to start inversion using long period seismograms.

3-D numerical simulations are clearly the way forward to enhance the resolution of the current tomographic images. Using a large part of the seismograms is possible in one simulation using certain misfit functions. However, we also need to consider the anelastic properties of the Earth even for phase measurements alone. In addition, the crustal problem has to be addressed during inversion by inverting the mantle together with the

crust omitting the correction step. Due to highly non-linear behavior of the crust, crustal corrections to the waveforms can only be applied using 3-D background models. Considering the full non-linearity of wave propagation undoubtedly provides the opportunity to increase the amount of usable data without worrying about defining phases in a seismogram. The chosen misfit function however has also a crucial control over the data sensitivity. In this respect, instantaneous phase measurements provide complementary information to the classical misfit functions. The results from instantaneous phase measurements are promising to exploit more information from a single seismogram.

Bibliography

Akçelik, V., Bielak, J., Biros, G., Epanomeritakis, I., Fernandez, A., Ghattas, O., Kim, E., O'Hallaron, D., & Tu, T., 2003. High resolution forward and inverse earthquake modeling on terascale computers, *Proc. ACM/IEEE Supercomputing SC'2003 Conference*, published on CD-ROM and at www.scconference.org/sc2003.

Aki, K. & Richards, P., 1980. *Quantitative Seismology, Theory and Methods*, W.H. Freeman, San Francisco, CA.

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

Bao, H., Bielak, J., Ghattas, O., Kallivokas, L., O'Hallaron, D., & adn J. Xu, J. S., 1998. Large-scale simulation of elastic wave propagation in heterogeneous media on parallel computers, *Comput. Methods Appl. Mech. Eng.*, **152**, 85–102.

Barnes, A. E., 2007. A tutorial on complex seismic trace analysis, *Geophysics*, **72(6)**.

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

Beghein, C., Resovsky, J., & Trampert, J., 2002. P and s tomography using normal mode and surface wave data with a neighbourhood algorithm, *Geophys. J. Int.*, **149**, 646–658.

Beghein, C., Trampert, J., & van Heijst, H. J., 2006. Radial anisotropy in seismic reference model of the mantle, *J. Geophys. Res.*, **101**, B02303, doi:10.1029/2005JB003728.

Bhattacharyya, J., Masters, G., & Shearer, P., 1996. Global lateral variations of shear wave attenuation in the upper mantle, *J. Geophys. Res.*, **101**, 22273–22289.

Bijwaard, H. & Spakman, W., 2000. Nonlinear global p-wave tomography by iterated linearized inversion, *Geophys. J. Int.*, **141**, 71–82.

Boschi, L. & Dziewonski, A., 1999. High and low-resolution images of the Earth's mantle: Implications of different approaches to tomographic modeling, *J. Geophys. Res.*, **104**, 25567–25594.

Boschi, L. & Dziewonski, A., 2000. Whole Earth tomography from delay times of P, PcP and PKP phases: Lateral heterogeneities in the outer core and radial anisotropy in the mantle?, *J. Geophys. Res.*, **105(B6)**, 13675–13696.

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

Boschi, L., Ampuero, J., Peter, D., Mai, P., Soldati, G., & Giardini, D., 2007. Petascale computing and resolution

in global seismic tomography, *Phys. Earth Planet. Inter.*, **163**, 245–250.

Bozdağ, E. & Trampert, J., 2008. On crustal corrections in surface wave tomography, *Geophys. J. Int.*, **172**, 1066–1082.

Capdeville, Y. & Marigo, J.-J., 2008. Shallow layer correction for Spectral Element like methods, *Geophys. J. Int.*, **172**, 1135–1150, doi:10.1111/j.1365–246X.2007.03703.x.

Capdeville, Y., Stutzmann, E., & Montagner, J., 2000. Effect of a plume on long-period surface waves computed with normal-mode coupling, *Phys. Earth Planet. Inter.*, **119**, 54–71.

Capdeville, Y., Chaljub, E., Vilotte, J., & Montagner, J., 2003. Coupling the spectral element method with a modal solution for elastic wave propagation in global Earth models, *Geophys. J. Int.*, **152**, 34–67.

Capdeville, Y., Gung, Y., & Romanowicz, B., 2005. Towards global earth tomography using the spectral element method: a technique based on source stacking, *Geophys. J. Int.*, **162**, 541–554.

Cara, M., 1973. Filtering of dispersed wavetrains, *Geophys. J. R. astr. Soc.*, **33**, 65–80.

Carcione, J. & Wang, P., 1993. A chebyshev collocation method for the wave equation in generalized coordinates, *Comp. Fluid Dyn. J.*, **2**, 269–290.

Cerveny, V., 2001. *Seismic Ray Theory*, Cambridge University Press, Cambridge.

Chaljub, E. & Valette, B., 2004. Spectral-element modeling of three-dimensional wave propagation in a self-gravitating Earth with an arbitrarily stratified outer core, *Geophys. J. Int.*, **158**, 131–141.

Chaljub, E., Capdeville, Y., & Vilotte, J., 2003. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids, *J. Comput. Phys.*, **187(2)**, 457–491.

Chen, P., Jordan, T., & Zhao, L., 2007a. Full three-dimensional tomography:a comparison between the scattering-integral and the adjoint wavefield methods, *Geophys. J. Int.*, **170**, 175–181.

Chen, P., Zhao, L., & Jordan, T., 2007b. Full 3D tomography for the crustal structure of the Los Angeles region, *Bull. Seismol. Soc. Am.*, **97(4)**, 1094–1120.

Claerbout, J. F., 1971. Toward a unified theory of reflector mapping, *Geophysics*, **36**, 467–481.

Crase, E., Pica, A., Noble, M., McDonald, J., & Tarantola, A., 1990. Robust elastic non-linear waveform inversion: application to real data, *Geophysics*, **55**, 527–538.

Dahlen, F. & Tromp, J., 1998. *Theoretical Global Seismology*, Princeton University Press, Princeton, New Jersey.

Dahlen, F. A. & Baig, A., 2002. Fréchet kernels for body wave amplitudes, *Geophys. J. Int.*, **150**, 440–466.

Dahlen, F. A., Hung, S. H., & Nolet, G., 2000. Fréchet kernels for finite-frequency traveltimes - I. Theory, *Geophys. J. Int.*, **141**, 157–174.

Dalton, C. A. & Ekström, G., 2006. Global models of surface wave attenuation, *J. Geophys. Res.*, **111**, B05317, doi:10.1029/2005JB003997.

Dalton, C. A., Ekström, G., & Dziewonski, A. M., 2008. The global attenu-

ation structure of the upper mantle, *Geophys. J. Int.*, **in press**.

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

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

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

Dziewonski, A. M., Hager, B., & O'Connell, R., 1977. Large scale heterogeneities in the lower mantle, *J. Geophys. Res.*, **82**, 239–255.

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

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

Engdahl, E., van der Hilst, R., & Buland, R., 1998. Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Am.*, **88(3)**, 722–743.

Ferreira, A., 2005. *Seismic surface waves in the laterally heterogeneous earth*, DPhil thesis, University of Oxford, Oxford.

Ferreira, A. & Woodhouse, J., 2006. Long-period seismic source inversions using global tomographic models, *Geophys. J. Int.*, **166**, 1178–1192.

Ferreira, A. & Woodhouse, J., 2007. Source, path, and receiver effects on seismic surface waves, *Geophys. J. Int.*, **168**, 109–132.

Fichtner, A., Kennett, B. L. N., Igel, H., & Bunge, H.-P., 2008. Theoretical background for continental and global scale full-waveform inversion in the time-frequency domain, *Geophys. J. Int.*, **175**, 665–685.

Fink, M., 1997. Time reversed acoustics, *Phys. Today*, **50**, 34–40.

Fischer, P. & Rønquist, E., 1994. Spectral-element methods for large scale parallel navier-stokes calculations, *Comput. Methods Appl. Mech. Eng.*, **116**, 69–76.

Fuchs, K., 1968. The reflection of spherical waves from transition zones with arbitrary depth-dependent elastic moduli and density, *J. Phys. Earth*, **16**, Special issue, 27.

Fuchs, K. & Müller, G., 1971. Computation of synthetic seismograms with the seflectivity method and comparison with observations, *Geophys. J. R. Astron. Soc*, **23**, 417–433.

Furumura, T., Kennett, B., & Furumura, M., 1998. Seismic wavefield calculation for laterally heterogeneous whole earth models using the pseudospectral method, *Geophys. J. Int.*, **135**, 845–860.

Gauthier, O., Virieux, J., & Tarantola, A., 1986. Two-dimensional nonlinear inversion of seismic waveforms: Numerical results, *Geophysics*, **51(7)**, 1387–1403.

Gee, L. & Jorden, T., 1992. Generalized seismological data functionals, *Geophys. J. Int.*, **111**, 363–390.

Geller, R. & Ohminato, T., 1994. Computation of synthetic seismograms and their partial derivatives for heterogeneous media with arbitrary nat-

ural boundary conditions using the Direct Solution Method, *Geophys. J. Int.*, **116**, 421–446.

Giardini, D., Li, X.-D., & Woodhouse, J. H., 1988. Splitting functions of long-period normal modes of the Earth, *J. Geophys. Res.*, **93**, 13716–13742.

Gilbert, F., 1970. Excitation of the normal modes of the Earth by earthquakes sources, *Geophys. J. R. Astron. Soc*, **22**, 223.

Gilbert, F. & Dziewonski, A., 1975. An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Philos. Trans. R. Soc. London A*, **278**, 187–269.

Gu, Y. J., Dziewonski, A. M., Su, W. J., & Ekstr´om, G., 2001. Models of mantle shear velocity and discontinuities in the pattern of lateral heterogeneities, *J. Geophys. Res.*, **106**, 11169–11199.

Gung, Y. & Romanowicz, B., 2004. Q tomography of the upper mantle using three-component long-period waveforms, *Geophys. J. Int.*, **157**, 813–830.

Hara, T., Tsuboi, S., & Geller, R., 1991. Inversion for laterally heterogeneous structure using a laterally heterogeneous starting model: preliminary results, *Geophys. J. Int.*, **104**, 523–540.

He, X. & Tromp, J., 1996. Normal mode constraints on the structure of the earth, *J. Geophys. Res.*, **101**, 20053–20082.

Hudson, J., 1977. Scattered waves in the coda of P, *J. Geophys.*, **43**, 359–374.

Hung, S. H., Dahlen, F. A., & Nolet, G., 2000. Fréchet kernels for finite frequency traveltimes - ii. examples, *Geophys. J. Int.*, **141**, 175–203.

Igel, H., 1999. Wave propagation in three-dimensional spherical sections by the Chebtshev spectral method, *Geophys. Res. Lett.*, **136**, 559–566.

Igel, H. & Weber, M., 1996. P-SV wave propagation in the whole mantle using high-order finite differences:application to lowermost mantle structure, *Geophys. Res. Lett.*, **23**, 415–418.

Jackson, D., 1972. Interpretation of inaccurate, insufficient, and inconsistent data, *Geophys. J. R. astr. Soc.*, **28**, 97–109.

Jeffreys, H. & Bullen, K., 1940. *Seismological Tables*, British Association Seismological Committee, London, UK.

Kanamori, H. & Anderson, D. L., 1977. Importance of physical dispersion in surface wave and free oscillation problems: Review, *Rev. Geophy.*, **15**, 105–112.

Kárason, H. & van der Hilst, R. D., 2001. Tomographic imaging of the lowermost mantle with differential times of refracted and diffracted core phases (PKP, Pdiff), *J. Geophys. Res.*, **106(B4)**, 6569–6587.

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

Kennett, B. & Engdahl, E. R., 1991. Traveltimes for global earthquake location and phase identification, *Geophys. J. Int.*, **105**, 429–465.

Kennett, B. & Gorbatov, A., 2004. Seismic heterogeneity in the mantle - strong shear wave signature of slabs from joint tomography, *Phys. Earth Planet. Inter.*, **146**, 87–100.

Kennett, B. & Yoshizawa, K., 2002. A reappraisal of regional surface wave tomography, *Geophys. J. Int.*, **150**, 37–44.

Kennett, B., Engdahl, E., & Buland, R., 1995. Constraints on seismic velocities in the earth from traveltimes, *Geophys. J. Int.*, **122**, 108–124.

Kennett, B., Widiyantoro, S., & van der Hilst, R. D., 1998. Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, **103**, 12469–12493.

Komatitsch, D. & Tromp, J., 1999. Introduction to the spectral-element method for 3-D seismic wave propagation, *Geophys. J. Int.*, **139**, 806–822.

Komatitsch, D. & Tromp, J., 2002a. Spectral-element simulations of global seismic wave propagation–I. Validation, *Geophys. J. Int.*, **149**, 390–412.

Komatitsch, D. & Tromp, J., 2002b. Spectral-element simulations of global seismic wave propagation–II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.*, **150**, 303–318.

Komatitsch, D. & Vilotte, J., 1998. The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, *Bull. Seismol. Soc. Am.*, **88(2)**, 368–392.

Komatitsch, D., Ritsema, J., & Tromp, J., 2002. The spectral-element method, Beowulf computing, and global seismology, *Science*, **298**, 1737–1742.

Komatitsch, D., Liu, Q., Tromp, J., Süss, P., Stidham, C., & Shaw, J. H., 2004. Simulations of ground motion in the Los Angeles Basin based upon the spectral-element method, *Bull. Seismol. Soc. Am.*, **94**, 187–206.

Komatitsch, D., Tsuboi, S., & Tromp, J., 2005. The Spectral-Element Method in Seismology, *editors A. Levander and G. Nolet, AGU Monograph Series 157*, pp. 205–228.

Kristekova, M., Kristek, J., Moczo, P., & Day, S., 2006. Misfit criteria for quantitative comparison of seismograms, *Bull. Seismol. Soc. Am.*, **96(5)**, 1836–1850.

Kustowski, B., Dziewonski, A., & Ekström, G., 1995. Nonlinear crustal corrections for normal-mode seismograms, *Bull. Seismol. Soc. Am.*, **97**, 1756–1762, doi:10.1785/0120070041.

Kustowski, B., Dziewonski, A., & Ekström, G., 2006. Modeling the anisotropic shear-wave velocity structure in the earth's mantle on global and regional scales, *EOS*, **87(52)**, Abstract S1E–02, Fall Meet. Suppl.

Laske, G. & Masters, G., 1996. Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, **101**, 16059–16075.

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

Lehmann, I., 1936. P', *Trav. Sci. Sect. Seis. U.G.G.I Toulouse*, **14**, 3–31.

Levshin, A. & Ratnikova, L., 1984. Apparent anisotropy in inhomogeneous media, *Geophys. J. R. astr. Soc*, **74**, 65–69.

Li, X. & Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.*, **101**,

22245–22272.

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

Liu, Q. & Tromp, J., 2006. Finite-frequency kernels based on adjoint methods, *Bull. Seismol. Soc. Am.*, **96**, 2383–2397.

Liu, Q. & Tromp, J., 2008. Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods, *Geophys. J. Int.*, **174**, 265–286.

Lognoné, P. & Romanowicz, B., 1990. Modeling of coupled normal modes of the Earth: the spectral method, *Geophys. J. Int.*, **102**, 365–395.

Luo, Y. & Schuster, G. T., 1991. Wave-equation traveltime tomography, *Geophysics*, **56**, 645–653.

Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J., 2009. An automated time-window selection algorithm for seismic tomography, *Geophys. J. Int.*, p. in press.

Marone, F. & Romanowicz, B., 2007. Non-linear crustal corrections in high resolution waveform seismic tomography, *Geophys. J. Int.*, **170**, 460–467.

Marquering, H., Dahlen, F. A., & Nolet, G., 1999. Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox, *Geophys. J. Int.*, **137**, 805–815.

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

Masters, G., Johnson, S., Laske, G., & Bolton, H., 1996. A shear velocity model of the mantle, *Philos. Trans. R. Soc. Lond. A*, **354**, 1385–1411.

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

Meier, U., Curtis, A., & Trampert, J., 2007. Fully nonlinear inversion of fundamental mode surface waves for a global crustal model, *Geophys. Res. Lett.*, **34**, L16304, doi:10.1029/2007GL030989.

Mohorovičić, A., 1909. Das Beben vom 8. X, *Jahrbuch met. obs. Zagreb*, **9**, 1–63.

Montagner, J. & Jobert, N., 1988. Vectorial tomography – II. Application to the Indian Ocean, *Geophys. J. Int.*, **94**, 309–344.

Montagner, J. & Tanimoto, T., 1991. Global upper mantle of seismic velocities and anisotropies, *J. Geophys. Res.*, **96(B12)**, 20337–20351.

Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E., & Hung, S., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343.

Mooney, W., Laske, G., & Masters, T., 1998. Crust5.1: a global crustal model at $5^\circ \times 5^\circ$, *J. Geophys. Res.*, **103**, 727–747.

Nataf, H., Nakanishi, I., & Anderson, D., 1986. Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy. Part III: inversion, *J. Geophys. Res.*, **91**, 7261–7307.

NOAA, 1988. National Oceanic and Atmospheric Administration (NOAA) product information catalog - ETOPO5 Earth Topography 5-minute digital model, *Technical*

Report, pp. 171, US Department of Commerce, Washington, DC.

Nolet, G., 1987. Waveform tomography, in Seismic Tomography, with Applications in Global Seismology and Exploration Geophysics, pp. 301–322, ed. Nolet, G., Reidel, Dordrecht.

Oldham, R. D., 1906. The Constitution of the Interior of the Earth, as Revealed by Earthquakes, *Q. J. Geol. Soc.*, **62**, 456–475.

Olsen, K., Archuleta, R., & Matarese, J., 1995. Magnitude 7.75 earthquake on the San Andreas fault: three-dimensional ground motion in Los Angles, *Science*, **270**, 1628–1632.

Park, J., 1986. Synthetic seismograms from coupled free oscillations: the effects of lateral structure and rotation, *J. Geophys. Res.*, **91**, 6441–6464.

Passier, M. & Snieder, R., 1995. Using differential waveform data to retrieve local S-velocity structure or path-averaged S-velocity gradients, *J. Geophys. Res.*, **100**, 24061–24078.

Patera, A., 1984. A spectral-element method for fluid dynamics:laminar flow in a channel expansion, *J. Comput. Phys.*, **54**, 468–488.

Perz, M., Sacchi, M., & O'Byrne, A., 2004. Instantaneous phase and the detection of lateral wavelet stability, *The Leading Edge*, **23**, 639–643.

Pratt, R., 1999. Seismic waveform inversion in the frequency domain, Part I: Theory and verification in a physical scale model, *Geophysics*, **64**, 888–901.

Priolo, E., Carcione, J., & Seriani, G., 1994. Numerical simulation of interface waves by high-order spectral modeling techniques, *J. Acoust. Soc. Am.*, **95(2)**, 681–693.

Qin, Y., Capdeville, Y., Montagner, J.-P., Boschi, L., & Becker, T. W., 2009. Reliability of mantle tomography models assessed by spectral element simulation, *Geophys. J. Int.*, doi:10.1111/j.1365-246X.2008.04032.x.

Rawlinson, N. & Sambridge, M., 2004. Wavefront evolution in strongly heterogeneous layered media using the fast marching method, *Geophys. J. Int.*, **156**, 631–647.

Reid, F., Woodhouse, J., & van Heijst, H., 2001. Upper mantle attenuation and velocity structure from measurements of differential s phases, *Geophys. J. Int.*, **145**, 615–630.

Resovsky, J. S. & Ritzwoller, M. H., 1998. New and refined constraints on three-dimensional Earth structure from normal modes below 3 mHz, *J. Geophys. Res.*, **103**, 783–810.

Resovsky, J. S., Trampert, J., & der Hilst, R. D. V., 2005. Error bars for the global seismic Q profile, *Earth planet. Sci. Lett*, **230**, 413–423.

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

Ritsema, J., Rivera, L., Komatitsch, D., Tromp, J., & van Heijst, H., 2002. The effects of crust and mantle heterogeneity on PP/P and SS/S amplitude ratios, *Geophys. Res. Lett.*, **29**, doi:10.1029/2001GL013831.

Ritzwoller, M. & Lavely, E. M., 1995. Three-dimensional seismic models of the earth's mantle, *Rev. Geophys.*, **33**, 1–66.

Ritzwoller, M. & Levshin, A., 1998. Eurasian surface wave tomography:

group velocities, *J. Geophys. Res.*, **103**, 4839–4878.

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

Romanowicz, B., 1990. The upper mantle degree 2: Constraints and inferences from global mantle wave attenuation measurements, *J. Geophys. Res.*, **95**, 11051–11071.

Romanowicz, B., 1995. A global tomographic model of shear attenuation in the upper mantle, *J. Geophys. Res.*, **100**, 12375–12394.

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

Selby, N. D. & Woodhouse, J. H., 2000. Controls on rayleigh wave amplitudes: attenuation and focusing, *Geophys. Res. Lett.*, **142**, 933–940.

Sengupta, M. & Toksöz, N., 1977. Three-dimensional model of seismic velocity variation in the earth's mantle, *Geophys. Res. Lett.*, **3**, 84–86.

Seriani, G., 1998. 3-D large-scale wave propagation modeling by a spectral element method on a Cray T3E multiprocessor, *Comput. Method. Appl. Mech. Engrg.*, **164**, 235–247.

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

Sieminski, A., Lévêque, J.-J., & Debayle, E., 2004. Can finite-frequency effects be accounted for in ray theory surface wave tomography?, *Geophys. Res. Lett.*, **31**, L24614, doi:10.1029/2004GL021402.

Snieder, R., 1988. Large-scale waveform inversions of surface-waves for lateral heterogeneity, *J. Geophys. Res.*, **93**, 12055–12065.

Snieder, R., 1993. Global inversions using normal mode and long-period surface waves, *Seismic Tomography: Theory and Practice*, pp. 23–63, Chapman and Hall, London, UK.

Snieder, R., Beckers, J., & Neele, F., 1991. The effect of small-scale structure on normal mode frequencies and global inversions, *J. Geophys. Res.*, **96**, 501–515.

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

Spetzler, J., Trampert, J., & Snieder, R., 2002. The effect of scattering in surface wave tomography, *Geophys. J. Int.*, **149**, 755–767.

Su, W. J., Woodward, R. L., & Dziewonski, A. M., 1994. Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, **99**, 4945–4980.

Talagrand, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.

Talangard, O. & Courtier, P., 1987. Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, **113**, 1311–1328.

Taner, M. T., Koehler, F., & Sheriff, R. E., 1979. Complex seismic trace analysis, *Geophysics*, **44(6)**, 1041–1063.

Tape, C., Liu, Q., & Tromp, J., 2007. Finite-frequency tomography using adjoint methods - Methodology and examples using membrane surface waves, *Geophys. J. Int.*, **168**, 1105–1129.

Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266.

Tarantola, A., 1987. *Inverse Problem Theory*, Elsevier, Amsterdam, the Netherlands.

Tarantola, A., 1988. Theoretical background for the inversion of seismic waveforms, including elasticity and attenuation, *Pure appl. Geophys.*, **128**, 365–399.

Tessmer, C., Kessler, D., Kosloff, D., & Behle, A., 1992. Multi-domain Chebyshev-Fourier method for the solution of the equations of motion of dynamic elasticity, *J. Comput. Phys.*, **100**, 355–363.

Thomas, C., Igel, H., Weber, M., & Scherbaum, F., 2000. Acoustic simulation of P-wave propagation in a heterogeneous spherical earth: numerical method and application to precursor waves to PKPdf, *Geophys. J. Int.*, **141**, 6441–6464.

Tibuleac, I. M., Nolet, G., Michaelson, C., & Koulakov, I., 2003. P wave amplitudes in a 3-D earth, *Geophys. J. Int.*, **155**, 1–10.

Tong, J., Dahlen, F. A., Nolet, G., & Marquering, H., 1998. Diffraction effects upon finite-frequency travel times: a simple 2-D example, *Geophys. Res. Lett.*, **25**, 1983–1986.

Toshinawa, T. & Ohmachi, T., 1992. Love wave propagation in three-dimensional sedimentary basin, *Bull. Seismol. Soc. Am.*, **82**, 1661–1667.

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

Trampert, J. & Spetzler, J., 2006. Surface wave tomography: finite-frequency effects lost in the null space, *Geophys. J. Int.*, **164**, 394–400.

Trampert, J. & van der Hilst, R. D., 2005. Towards a quantitative interpretation of global seismic tomography, *Geophysical Monograph*, **160**, 47–62.

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

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

Trampert, J. & Woodhouse, J., 2001. Assessment of global phase velocity models, *Geophys. J. Int.*, **144**, 165–174.

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

Tromp, J., Tape, C., & Liu, Q., 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels, *Geophys. J. Int.*, **160**, 195–216.

Tromp, J., Komatitsch, D., & Liu, Q., 2008. Spectral-element and adjoint methods in seismology, *Commun. Comput. Phys.*, **3**, 1–32.

Tsuboi, S., Komatitsch, D., Ji, C., &

Tromp, J., 2003. Broadband modeling of the 2002 Denali fault earthquake on the Earth Simulator, *Phys. Earth Planet. Inter.*, **139**, 305–312.

Tsuboi, S., Komatitsch, D., Ji, C., & Tromp, J., 2004. Modeling of global seismic wave propagation on the Earth Simulator, *Journal of Earth Simulator*, **1**, 57–66.

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

van Heijst, H. J. & Woodhouse, J., 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.

Vidale, J., 1988. Finite-difference calculations of traveltimes, *Bull. Seismol. Soc. Am.*, **78**, 2062–2076.

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

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

Warren, L. M. & Shearer, P. M., 2002. Mapping lateral variations in upper mantle attenuation by stacking p and pp spectra, *J. Geophys. Res.*, **107(B12)**, 2342, doi:10.1029/2001JB001195.

Woodhouse, J., 1974. Surface waves in a laterally varying layered structure, *Geophys. J. R. astr. Soc.*, **37**, 461–490.

Woodhouse, J. & Dahlen, F. A., 1978. The effect of a general aspherical perturbation on the free oscillations of the Earth, *Geophys. J. R. astr. Soc.*, **61**, 335–354.

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

Woodhouse, J. & Girnuis, T. P., 1982. Surface wavess and free oscillations in a regional earth model, *Geophys. J. R. astr. Soc.*, **68**, 653–673.

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

Wu, R. & Aki, K., 1985. Scattering characteristics of elastic waves by an elastic heterogeneity, *Geophysics*, **50**, 582–595.

Yoshizawa, K. & Kennett, B., 2002. Determination of the influence zone for surface wave paths, *Geophys. J. Int.*, **149**, 440–453.

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

Zhao, L., Jordan, T. H., Olsen, K., & Chen, P., 2005. Fréchet kernels for imaging regional earth structure based on three-dimensional reference models, *Bull. Seismol. Soc. Am.*, **95**, 2066–2080.

Zhou, H. W., 1996. A high resolution p wave model of the top 1200 km of the mantle, *J. Geophys. Res.*, **101**, 27791–27810.

Zhou, Y., 2009. Surface-wave sensitivity to 3-D anelasticity, *Geophys. J. Int.*, **in press**.

Zhou, Y., Dahlen, F., & Nolet, G., 2004. Three-dimensional sensitivity kernels for surface wave observables, *Geophys. J. Int.*, **158**, 142–168.

Zhou, Y., Dahlen, F., Nolet, G., & Laske, G., 2005. Finite-frequency effects in global surface-wave tomography, *Geophys. J. Int.*, **163**, 1087–1111.

Summary

We have reached a stage in seismic tomography where further refinements with classical techniques become very difficult. Advances in numerical methods and computational facilities are providing new opportunities in seismic tomography to enhance the resolution of tomographic mantle images. 3-D numerical simulations of seismic wavefields also allow us to check the reliability of the current mantle images and the classical techniques. In this thesis, using 3-D wave simulations by a spectral element method, we present three specific studies which underline some shortcomings in global mantle tomography.

We first investigated crustal corrections for fundamental mode surface waves. We compared the crustal corrections estimated by first order approximations such as great circle approximation, exact ray theory and finite-frequency theory, to those obtained from 3-D wave simulations. We observed that crustal corrections can produce errors larger than those in phase measurements at periods longer than 60 and 80 s for Rayleigh and Love waves, respectively. Rayleigh and Love waves are differently influenced by the crust. Therefore imperfect crustal corrections have potential to produce radial anisotropy which may bias our interpretation of anisotropy in the upper-mantle. Extensions to the great circle approximation do not improve the results because of the highly non-linear behaviour of the crust. Thus either the mantle has to be inverted together with the crust or 3-D background models have to be used for the measurements.

We then compared real data with synthetic seismograms computed in various 3-D mantle models. We particularly investigated the effect of damping on global mantle models. In addition, we tried to examine the elastic and anelastic contribution of the models by comparing seismograms computed in 1-D and 3-D crustal and mantle models. Our results show that different 3-D mantle models give statistically similar results although they are different from each other. We observed that a large part of the seismograms remained unexplained, particularly the amplitudes. 3-D velocity models are not enough to explain amplitude information alone. Attenuation may have an important contribution to surface-wave amplitudes whereas body waves are more affected by scattering effects.

Lastly, we proposed new misfit functions based on instantaneous phase and cross-correlation measurements for full waveform inversion. In addi-

tion, we qualitatively compared them to classical ones such as waveform and travel-time misfits by computing the finite-frequency sensitivity kernels using adjoint methods based upon a spectral element method. We observed that instantaneous phase measurements provide complementary information to the other misfit functions and are promising to exploit more information from a single seismogram.

Samenvatting (Summary in Dutch)

Aan de hand van 3-D numerieke golfvoortplantingssimulaties met de spectrale elementen methode, presenteren we drie studies die tekortkomingen in globale manteltomografie blootleggen. Ten eerste hebben we korstcorrecties voor de grondtoon van oppervlaktegolven onderzocht. We constateerden dat deze korstcorrecties fouten kunnen veroorzaken die groter zijn dan de standaarddeviatie van fasemetingen voor korte perioden en daarom onze interpretatie van anisotropie in de bovenmantel kunnen beïnvloeden. Dit betekent dat we bij inversies, ofwel een oplossing moeten zoeken voor zowel de mantel en de korst, of dat we een 3-D achtergrondmodel moeten gebruiken. Ten tweede vergeleken we synthetische seismogrammen, berekend met 3-D modellen van de mantel, met echte seismogrammen. Verschillende aardmodellen gaven, statistisch gezien, vergelijkbare resultaten. Grote gedeeltes van de seismogrammen bleef echter onverklaard, met name de amplitude. Snelheidsmodellen alleen zijn niet afdoende om de amplitudes te verklaren. Attenuatie kan een belangrijke rol spelen bij de amplitude van oppervlaktegolven. Ten slotte poneren we nieuwe misfit-functies gebaseerd op verschillen in instantane fase en in kruiscorrelaties. We vergelijken de functies met klassieke misfit-functies door geadjungeerde kernels te berekenen. Metingen van de instantane fase zijn veelbelovend om meer informatie uit individuele seismogrammen te verkrijgen.

Acknowledgments

I had a great experience, lots of fun and challenges during the last four years I spent in Utrecht. Here I would like to thank everybody who contributed to this thesis in one way or another during all these years.

This thesis would have never been accomplished without the help and guidance of my supervisor Jeannot Trampert. I am greatly indebted to him for introducing me to the world of "global seismology" and for the chance he gave me to work on such an interesting and challenging project. I appreciate very much the time he spent with me, his constant support and his encouragement. He was unbelievably patient with my endless questions and over-stressed moments. I learnt a lot from him; not only seismology but also the importance of being positive and patient in science. Many thanks for everything, Jeannot!

The backbone of this research is the spectral element method, as cited many times in the thesis, thus I would like to acknowledge Komatitsch & Tromp once again for making the code available. Thanks everybody who contributed to the development of the code as well. I am very grateful to Jeroen Tromp for his help with the problems related to the code, all discussions, and his comments about my work.

I appreciated very much the time that the dissertation committee spent to read my thesis and all their comments: Heiner Igel, Jean-Paul Montagner, Wim Spakman, Jeroen Tromp and John Woodhouse. I also would like to thank Christine Thomas, Sergei Lebedev, Laura Cobden and Hanneke Paulssen for accepting being in the committee in my defense.

It was a great opportunity to be a part of SPICE community. Thanks to all SPICErs for the great scientific atmosphere and pleasurable social events in the workshops and meetings. I gratefully acknowledge the financial support from the SPICE project to my research.

I am grateful to Hanneke Paulssen for her detailed answers to my frequent questions and many fruitful discussions. Thanks to Anne Sieminski for her help with adjoint simulations, all discussions and advice which also continued after she moved to Grenoble. I sincerely thank Ana Ferreira due to her interest in my work, all the scientific/non-scientific conversations we had in the SPICE meetings and in Utrecht.

I express my gratitude to Argun Kocaoğlu, my former supervisor, and Abdullah Karaman who encouraged me to start this PhD and remained

always interested in my work. I wish to thank all my friends and colleagues in Istanbul Technical University.

I had great friends and colleagues in Utrecht. Special thanks go to Roberta & Ueli for their friendship and extra attention to help me get used to living in the Netherlands. Thanks Roberta also for her effort to teach me cycling. I have always enjoyed the delightful discussions with Ueli about nearly everything. A big thanks go to Sergei for all scientific discussions, encouragement, and many social activities. I enjoyed a lot many spontaneous dinners and drinks with Céline. Thanks to all the past and present people in the Earth Sciences department for many scientific and social occasions: Mei -my officemate-, Karin, Peter, Robbert, Frederic, Florian, Tedi, Alison, Arie, Maisha, Benoît, Marzieh, Jelle, Jeroen, Sandrine, Karin R., Pasha, Bahjat, Stefan, Tim, Gabriela, Tonie, Steven, Rinus, Rob, Kabir, Paul, Henk, Sanne, Ildiko, Pınar & Nuri. Wouter and Thomas were very helpful with the Dutch translation of the summary, and thanks Joost (have fun in Hawaii!) for proofreading it. Jacqueline was always very kind and extremely helpful with all administrative issues. Without the help of Theo, I would have never succeeded in finishing my endless simulations on the cluster. Joop is thanked for all his help related to various computer problems. Special thanks go to my paranimfen Ilaria and Ahmet. Thanks Ilaria for extra-strong espresso supply, delicious Italian dinners, being a perfect housemate and all the enjoyable times we spent together. Ahmet, the flying cheeseman(!), thanks a lot for making me laugh by showing up in my office, mostly for coffee or tea, during my most stressed moments.

A big "thank you" goes to my second housemate Elif for all hilarious moments, cooking together, parties, concerts and her great DVD collection. Thanks to friends Nazem & Ali, Özkan, Gülseli, İrem, and my very special former flatmates in Cambridgelaan; Vera, Manuela and Rodrigo for all the fun we shared together. Very special thanks go to my dearest friends Didem and Doğa. No need for big words, just thanks for proving that the distance does not matter for real friendships.

Many thanks to my parents Zehra and Hilmi Bozdağ for their love and continuous support throughout my life. Thanks to my sister and her husband İnci and Can Küçükler, my dear niece Ceren, my sister-in-law Buket Pekgöz-Bozdağ and my sweet nephew Efe. I appreciate the effort they put in to be in my defense ceremony. Finally, the biggest thanks go to my brother, Ünal Bozdağ, without whom I could have never succeeded it. I have always felt so lucky to have him always right by my side.

Curriculum vitae

February 9, 1979	Born in Artvin, Turkey
1992-1996	**High School** Kadir Has Lisesi Istanbul, Turkey
1996-2000	**BSc in Geophysics** Department of Geophysical Engineering Istanbul Technical University Istanbul, Turkey
2000-2002	**MSc in Geophysics** Institute of Science and Technology Istanbul Technical University Istanbul, Turkey
2002-2005	**Research Assistant** Department of Geophysical Engineering Istanbul Technical University Istanbul, Turkey
2005-2009	**PhD in Seismology** Department of Earth Sciences Utrecht University Utrecht, the Netherlands