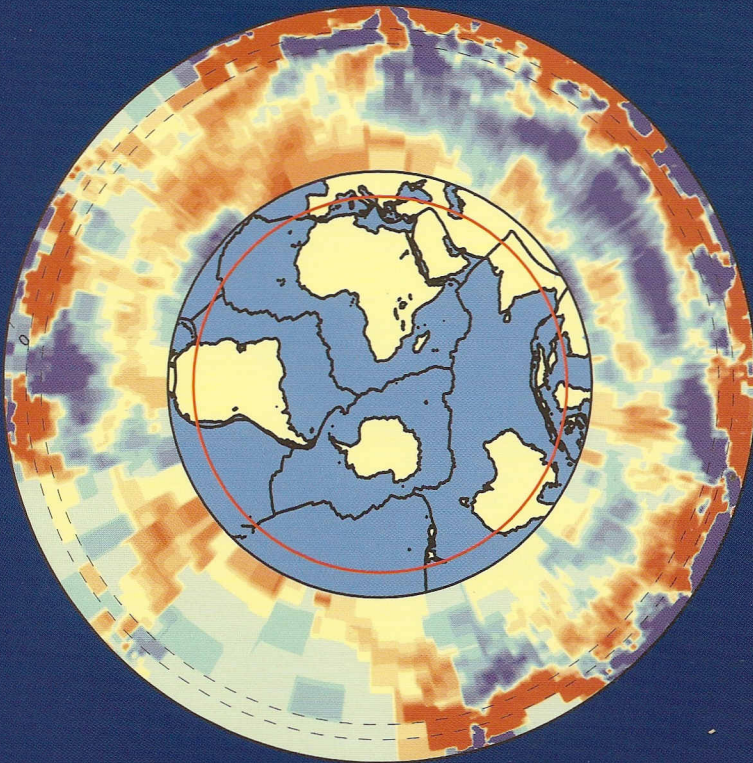


GEOLOGICA ULTRAIECTINA

Mededelingen van de  
Faculteit Aardwetenschappen  
Universiteit Utrecht

No. 181

# Seismic travel-time tomography for detailed global mantle structure



Harmen Bijwaard

GEOLOGICA ULTRAIECTINA

Mededelingen van de  
Faculteit Aardwetenschappen  
Universiteit Utrecht

No. 181

**Seismic travel-time tomography for  
detailed global mantle structure**

Harmen Bijwaard

Promotor: Prof. Dr. R.K. Snieder  
Institute of Earth Sciences  
Utrecht University

Co-promotor: Dr. W. Spakman  
Institute of Earth Sciences  
Utrecht University

Members of the dissertation committee:

Dr. E.R. Engdahl	National Earthquake Information Center United States Geological Survey, Denver
Prof. Dr. J.T. Fokkema	Subfaculty of Technical Earth Sciences Delft University of Technology
Prof. Dr. R. van der Voo	Department of Geological Sciences University of Michigan, Ann Arbor
Prof. Dr. N.J. Vlaar	Institute of Earth Sciences Utrecht University
Prof. Dr. M.J.R. Wortel	Institute of Earth Sciences Utrecht University

**Cover illustration:** Whole mantle-and-crust cross-section through the model of Chapter 4 along the great-circle indicated in red.



# **Seismic travel-time tomography for detailed global mantle structure**

Seismische reistijdentomografie voor  
gedetailleerde globale mantelstructuur

(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. H.O. VOORMA, INGEVOLGE HET BESLUIT VAN HET  
COLLEGE VOOR PROMOTIES IN HET OPENBAAR TE VERDEDIGEN  
OP MAANDAG 8 NOVEMBER 1999 DES MIDDAGS TE 2.30 UUR

DOOR

**Harmen Bijwaard**

Geboren op 13 September 1971, te Eindhoven



Promotor: Prof. Dr. R.K. Snieder  
Institute of Earth Sciences  
Utrecht University

Co-promotor: Dr. W. Spakman  
Institute of Earth Sciences  
Utrecht University

The research described in this thesis was carried out at:

Institute of Earth Sciences  
Utrecht University  
Budapestlaan 4  
3584 CD Utrecht  
The Netherlands

The investigations were supported by the Netherlands Geosciences Foundation (GOA) with financial aid from the Netherlands Organization for Scientific Research (NWO) through project 750.195.13. This work was conducted under the program of the Vening Meinesz School of Geodynamics.

ISBN: 90-5744-039-3

## List of publications (partially) incorporated in this thesis

- Bijwaard, H., and W. Spakman, Advancing global *S*-wave tomography toward the regional scale, *Geophys. J. Int.* (submitted), 1999.
- Bijwaard, H., and W. Spakman, Fast kinematic ray tracing of first and later arriving global seismic phases, *Geophys. J. Int.* (in press), 1999.
- Bijwaard, H., and W. Spakman, Nonlinear global *P*-wave tomography, *Geophys. J. Int.* (submitted), 1999.
- Bijwaard, H., and W. Spakman, Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sci. Lett.*, 166, 121-126, 1999.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Rangin, C., W. Spakman, M. Pubellier, and H. Bijwaard, Tomographic and geological constraints on subduction along the eastern Sundaland continental margin (south-east Asia), *Soc. Géol. Fr.* (submitted), 1999.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic inverse problems (manuscript in preparation), 1999.
- Spakman, W., and H. Bijwaard, Mantle structure and large scale dynamics of South-East Asia, in: *The GEODYnamics of S and SE Asia (GEODYSSSEA) Project*, Scientific Technical Report STR98/14, edited by P. Wilson and G.W. Michel, GeoForschungsZentrum Potsdam, pp. 313-339, 1998.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Mesozoic subducted slabs under Siberia, *Nature*, 397, 246-249, 1999.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.*, 171, 7-20, 1999.

For those who believe no  
explanation is necessary,  
while for those who do not believe  
no explanation is possible.

Franz Werfel



# Contents

<b>1</b>	<b>Introduction and summary</b>	<b>1</b>
	References	3
<b>2</b>	<b>A review of developments in seismic tomography</b>	<b>5</b>
2.1	Historic overview	5
2.2	Different tomography methods	6
2.3	Different tomography scales	7
2.4	Specific ingredients of travel-time tomography	8
2.4.1	Tomographic data	8
2.4.2	Model parameterizations	9
2.4.3	Tomographic inversions	10
2.4.4	Sensitivity analyses	11
2.5	Shortcomings and future directions	12
	References	13
<b>3</b>	<b>A brief summary of theory and irregular parameterization</b>	<b>23</b>
3.1	Travel-time tomography theory	23
3.2	Irregular cell parameterization of tomographic inverse problems	26
3.2.1	Introduction	26
3.2.2	Theory	27
3.2.3	Construction of an irregular cell model for global tomography	28
3.2.4	Conclusions	30
	References	30
<b>4</b>	<b>Closing the gap between regional and global travel-time tomography</b>	<b>33</b>
	Abstract	33
4.1	Introduction	34
4.2	Data	35
4.3	Parameterization and inversion	38
4.4	Data fit and model resolution	41
4.5	Results	45
4.6	Conclusions	58
	Acknowledgments	60
	References	61
<b>5</b>	<b>Regional interpretations of the global <i>P</i>-wave model</b>	<b>65</b>
5.1	Tomographic evidence for a narrow whole mantle plume below Iceland	65
	Abstract	65
5.1.1	Introduction	65
5.1.2	Description of the Iceland anomaly	66
5.1.3	Resolution of the Iceland anomaly	67
5.1.4	Interpretation of the Iceland anomaly	68

5.1.5	Conclusions	69
5.2	Mesozoic subducted slabs under Siberia	69
5.2.1	Paleogeographic background	70
5.2.2	Tomography results	70
5.2.3	Interpretation	70
5.2.4	Conclusions	71
5.3	Tethyan subducted slabs under India	72
5.3.1	Paleotectonic overview	72
5.3.2	Tomographic results	72
5.3.3	Interpretation	76
5.3.4	Conclusions	77
5.4	Tomographic constraints on the dynamics of Southeast Asia	77
5.4.1	Introduction	77
5.4.2	General description of the mantle below Southeast Asia	78
5.4.3	Detailed interpretations of anomalous structures	78
5.4.4	Conclusions	83
	References	83
<b>6</b>	<b>Advancing global S-wave tomography toward the regional scale</b>	<b>89</b>
	Summary	89
6.1	Introduction	89
6.2	Data, model parameterization and inversion	90
6.2.1	Data	90
6.2.2	Parameterization	90
6.2.3	Inversion	90
6.3	Results and resolution	91
6.3.1	Overview of the results	91
6.3.2	Resolution	92
6.3.3	Model description	93
6.3.4	Detailed comparisons of upper mantle regions	95
	6.3.4.1 Europe to Central Asia.	96
	6.3.4.2 Southeast Asia to Tonga-Kermadec	99
	6.3.4.3 The Caribbean	99
	6.3.4.4 South America	101
6.4	Discussion	102
6.5	Conclusions	104
	Acknowledgments	104
	References	104
<b>7</b>	<b>Fast kinematic ray tracing of first and later arriving global seismic phases</b>	<b>107</b>
	Summary	107
7.1	Introduction	107
7.2	General framework for ray tracing	108
7.3	The adjusted graph method	109
7.4	The extended ray perturbation method	111
7.5	Tests for precision and convergence speed	113

---

7.5.1	A comparison with analytical results	113
7.5.2	A synthetic experiment in a 2-D, spherical model	116
7.5.3	Tracing rays in a 3-D, spherical Earth model	118
7.6	Ray tracing for global nonlinear travel-time tomography	121
7.7	Conclusions	122
	Acknowledgments	123
	References	123
	Appendix A: Inclusion of velocity discontinuities in the perturbation method	124
<b>8</b>	<b>Nonlinear global <i>P</i>-wave tomography</b>	<b>129</b>
	Abstract	129
8.1	Introduction	129
8.2	Data and ray tracing	130
8.3	Parameterization and inversion	131
8.4	Results	134
	8.4.1 Data fit	134
	8.4.2 Resolution	135
	8.4.3 Cluster relocation vectors and station corrections	136
	8.4.4 The 3-D model obtained from nonlinear tomography	136
	8.4.5 Some comparisons between BSE and BSE_NL	139
8.5	Conclusions	146
	References	146
	<b>Appendix: Additional figures</b>	<b>149</b>
	<b>Samenvatting (Summary in Dutch)</b>	<b>175</b>
	<b>Acknowledgments</b>	<b>177</b>
	<b>Curriculum Vitae</b>	<b>179</b>



# Chapter 1

## Introduction and summary

Seismic waves provide the best way to investigate the Earth's deep interior, although the information they contain can be difficult to extract. Direct information about deeper Earth structure is only obtained from boreholes (up to some 12 km deep), but these do not penetrate through the Earth's crust, which constitutes only 0.5% of the Earth's radius (6371 km). The largest part of the Earth's interior is thus not sampled by direct methods. Other (indirect) methods, based on geological/geochemical data, gravity or (electro-)magnetic measurements, surface heat flow or rotation of the Earth, yield either detailed information about the shallow (say uppermost 50 km) structure or large-scale constraints on the deeper structure. High-frequency seismic data are, however, sensitive to heterogeneity as small as a few tens of kilometers in the upper mantle to a few hundred kilometers in the lowermost mantle (nearly 3000 km deep). Features of these scale-lengths have already been discovered in the upper mantle (e.g. Spakman et al., 1993) and may also exist in the deeper lower mantle. They arise from the convective state of the Earth's mantle in which cold plate material at the surface can subduct and sink, and hot material at the core-mantle boundary may rise.

Seismic travel-time tomography (the method employed in this thesis) uses the travel times of seismic waves to map the 3-D structure of the Earth's interior in terms of variations in seismic wavespeeds. These wavespeeds ( $P$ -velocity and  $S$ -velocity) depend on temperature, composition and pressure through the elastic moduli and density. Cold material sinking from the Earth's surface is expected to transmit seismic energy faster than normal, and vice versa hot material coming from the core-mantle boundary should transmit seismic energy relatively slowly. Although commonly interpreted as a temperature effect, the exact cause (temperature, pressure, composition or a combination thereof) of the variations in seismic velocity is still a matter of debate (e.g. Jackson, 1998). Hence, in this thesis no attempt will be made to further interpret the obtained seismic velocities in terms of these quantities.

The method of seismic tomography has been adapted from a 3-D imaging technique for the inside of the human body, the so-called CT-scan (Computerized Tomography) that is employed in medical imaging. Similar techniques are nowadays applied in a wide range of sciences such as meteorology (e.g. Fehmers, 1996), oceanography (e.g. Send et al., 1997), astronomy (e.g. Bennett et al., 1997), and helioseismology (e.g. Duvall et al., 1996) to image parts of the atmosphere, oceans, galaxies, and the interior of the sun, respectively. In the Earth Sciences several different forms of tomography (apart from travel-time tomography) have been developed, which are discussed in Chapter 2.

**The object of this thesis is to use travel-time tomography to focus and enhance the existing global image of the Earth's mantle and crust. This image is still rather blurred with respect to the considerably sharper pictures commonly obtained in regional studies. The improvement is basically obtained in two steps: In the first step a more or less conventional tomography method is applied to a reprocessed and sensibly averaged data set for a more sophisticated, irregular discretization of the study volume, and in the second step the (global) method itself is refined by introducing 3-D ray**

---

**tracing to calculate accurate seismic ray paths through heterogeneous velocity models (nonlinear tomography).**

The composition of this thesis is as follows. Chapter 2 gives a general overview of the developments and findings of seismic tomography and in particular of the different approaches to travel-time tomography.

Chapter 3 is basically divided in two parts: It first briefly summarizes the basic theory of travel-time tomography and the assumptions and approximations it is based on. This part is largely based on Spakman (1993), who gives an extensive derivation of all necessary equations. The second part deals with the efficient subdivision of the Earth's mantle and crust into irregular cells with sizes adapted to the amount of local seismic ray sampling. Due to the concentration of earthquakes near plate boundaries and the lack of seismic stations on the ocean floor, seismic ray density (or hitcount) is very inhomogeneously distributed throughout the Earth's interior. This has basically been ignored in previous global tomography studies, leading to large amounts of ill-determined model parameters. This part is a short summary of Spakman and Bijwaard (1999), who give a more elaborate treatise of the subject.

Chapter 4 gives a description of the global model and its resolution, derived by linearized *P*-wave tomography on the parameterization developed in Chapter 3. The main result is the retrieval of small-scale detail in well-sampled model volumes that is directly comparable to, and is sometimes improving on structure obtained in regional tomography studies. This chapter is identical to Bijwaard et al. (1998).

In Chapter 5, the model of Chapter 4 is exploited for a series of regional interpretations for well-resolved parts of the Eurasian mantle. Observations include the first rather detailed image of a whole mantle plume, rising from the core-mantle boundary (Bijwaard and Spakman, 1999a); the detection of the (so far) oldest subducted lithosphere, derived from the closure of the Mongol-Okhotsk ocean near the present-day location of Lake Baikal (Van der Voo et al., 1999a); the interior structure of a high-velocity anomaly commonly associated with the subduction of the Tethys ocean (Van der Voo et al., 1999b); and detailed patterns of subductions below Southeast Asia (Spakman and Bijwaard, 1998; Rangin et al., 1999).

Chapter 6 addresses the derivation of an *S*-wave tomography model in a similar fashion as for the *P*-wave model in Chapter 4. The main finding is a high correlation between the *P* and *S*-velocity models that increases for regions with higher ray sampling. This chapter is identical to Bijwaard and Spakman (1999b).

Chapter 7 describes the development of two fast and efficient 3-D ray tracing techniques that are prerequisite for the extension from linear to nonlinear tomography. A detailed comparison of both methods shows that only the ray perturbation method is computationally efficient enough to perform 3-D ray tracing on the entire data set used for linearized tomography (7.6 million rays). This chapter is practically identical to Bijwaard and Spakman (1999c).

Finally, Chapter 8 describes the *P*-wave tomography model derived from nonlinear tomography and some comparisons with the model of Chapter 4. The main conclusions are that nonlinear tomography does not lead to an overall dramatic change in heterogeneity patterns, but predominantly alters the solution in regions with considerable anomaly amplitudes (mainly in the uppermost mantle). This chapter is based on Bijwaard and Spakman (1999d).

## References

- Bijwaard, H., and W. Spakman, Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sci. Lett.*, 166, 121-126, 1999a.
- Bijwaard, H., and W. Spakman, Advancing global *S*-wave tomography toward the regional scale, *Geophys. J. Int.* (submitted), 1999b.
- Bijwaard, H., and W. Spakman, Fast kinematic ray tracing of first and later arriving global seismic phases, *Geophys. J. Int.* (in press), 1999c.
- Bijwaard, H., and W. Spakman, Nonlinear global *P*-wave tomography, *Geophys. J. Int.* (submitted), 1999d.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Bennett, C.L., M.S. Turner, and M. White, The cosmic Rosetta stone, *Phys. Today*, 11, 32-38, 1997.
- Duvall, T.L. Jr., S. D'Silva, S.M. Jefferies, J.W. Harvey, and J. Schou, Downflows under sunspots detected by helioseismic tomography, *Nature*, 379, 235-237, 1996.
- Fehmers, G., Tomography of the ionosphere, Ph.D. thesis, 132 pp., Eindhoven Univ., Eindhoven, Netherlands, 1996.
- Jackson, I., Elasticity, composition and temperature of the Earth's lower mantle: a reappraisal, *Geophys. J. Int.*, 134, 291-311, 1998.
- Rangin, C., W. Spakman, M. Pubellier, and H. Bijwaard, Tomographic and geological constraints on subduction along the eastern Sundaland continental margin (south-east Asia), *Soc. Géol. Fr.* (submitted), 1999.
- Send, U., G. Krahmhann, D. Mauuary, Y. Desaubies, F. Gaillard, T. Terre, J. Papadakis, M. Taroudakis, E. Skarsoulis, and C. Millot, Acoustic observations of heat content across the Mediterranean Sea, *Nature*, 385, 615-617, 1997.
- Spakman, W., Iterative strategies for non-linear travel time tomography using global earthquake data, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 190-226, Chapman and Hall, London, 1993.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic inverse problems (manuscript in preparation), 1999.
- Spakman, W., and H. Bijwaard, Mantle structure and large scale dynamics of South-East Asia, in: *The GEODYnamics of S and SE Asia (GEODYSSSEA) Project*, Scientific Technical Report STR98/14, edited by P. Wilson and G.W. Michel, GeoForschungsZentrum Potsdam, pp. 313-339, 1998.
- Spakman W., S. van der Lee, and R. van der Hilst, Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, 79, 3-74, 1993.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Mesozoic subducted slabs under Siberia, *Nature*, 397, 246-249, 1999a.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.*, 171, 7-20, 1999b.



# Chapter 2

## A review of developments in seismic tomography

This chapter provides a summary of the historic developments and the different forms of seismic tomography with a focus on the method of travel-time tomography, which is used in this study. This summary is not intended to be complete, but it is meant to draw the context for the research described in this thesis.

### 2.1 Historic overview

The purpose of seismic tomography is to map the 3-D variation in seismic velocities in the Earth's interior. Already early this century, the existence of differences in seismic velocities under oceans and continents was inferred (e.g. Gutenberg, 1924). However, attention long focussed on determining the 1-D, *radial* structure of the Earth (e.g. Jeffreys and Bullen, 1940), and only in the second half of this century *lateral* deviations from the radial structure were inferred from seismic travel times. In the late 1970s, the first formal travel-time inversions for 3-D velocity heterogeneity were performed. A robust method was introduced by Aki and Lee (1976) and Aki et al. (1977), in which seismic travel-time residuals were directly inverted for 3-D *P*-velocity structure in a regional block model. Their method (tomography *avant la lettre*) was also employed by Sengupta and Toksöz (1976), who computed whole mantle structure in  $10^\circ$  blocks. Simultaneously, Dziewonski et al. (1977) followed a different approach: Instead of dividing the Earth's mantle in cells, they employed spherical harmonic functions (up to degree 3) for the representation of its structure. Because this parameterization greatly reduced the number of unknowns in the inverse problem, they were able to calculate a direct (or generalized) inverse of the tomographic matrix even though a large number (700000) of *P*-residuals were used. In the early 1980s the introduction of efficient matrix solvers (Clayton and Comer, 1983; Nolet, 1985) proved to be a major step forward in refining the solutions for Earth structure. It allowed for the implementation of larger numbers of both seismic data and model parameters in tomographic inverse problems. As a result the first global inversions with a fine ( $5^\circ \times 5^\circ \times 100$  km) block parameterization appeared (Clayton and Comer, 1983: 75,000 blocks). Meanwhile the spherical harmonics representations were extended to higher degree (Nakanishi and Anderson, 1982: degree 7; Woodhouse and Dziewonski, 1984: degree 8) and the tomographic picture of the Earth's interior became more and more focussed. As ever more models were developed that showed similar but also contradictory features (see for example the review by Ritzwoller and Lavelly, 1995), questions were more strongly raised about the reliability of the images. Apart from ongoing improvements in data quality, the tomographic method itself has therefore been improved over the years in many different ways. For the travel-time tomography method the main advances consisted of the following. First of all, sensitivity analysis was introduced, because formal resolution could in many cases not be calculated due to the size of the inverse problem (e.g. Grand, 1987; Spakman and Nolet, 1988; Humphreys and Clayton, 1988). Secondly, the simultaneous solution for seismic velocities and earthquake hypocenters, which form an in principle coupled problem,

appeared (e.g. Spencer and Gubbins, 1980; Engdahl and Gubbins, 1987; Spakman and Nolet, 1988). Thirdly, nonlinear inversions were introduced for small-scale studies, in which models were calculated in an iterative fashion, alternating linear inversions with 3-D ray tracing, and thus accounting for seismic ray bending due to inferred velocity heterogeneity (e.g. Thomson and Gubbins, 1982; VanDecar, 1991; Papazachos and Nolet, 1997). And finally, irregular parameterizations were developed to account for the very inhomogeneous sampling of the Earth by seismic ray paths (e.g. Abers and Roecker, 1991; Fukao et al., 1992; Liu and Jin, 1993; Vesnaver, 1996; Curtis and Snieder, 1997). These improvements have led to better focussed and more reliable tomographic images, but the computational burden of the tomographic method has become enormous because of the extra hypocentral parameters, the sensitivity tests, irregular gridding, and the complicated, 3-D ray tracing. Hence, some of these improvements were up to now only applicable to small size problems, and had not been implemented in the construction of global Earth models.

In this thesis, a global irregular parameterization is developed (Chapter 3 and Spakman and Bijwaard, 1999), and linear and nonlinear inversions including 3-D ray tracing are performed for slownesses (the reciprocals of velocities), station statics and earthquake locations, simultaneously (Chapters 4, 6, and 8, and Bijwaard et al., 1998; Bijwaard and Spakman, 1999a; Bijwaard and Spakman, 1999b). This brings the current state of the art of global travel-time tomography more up to date with the developments in regional modeling.

## 2.2 Different tomography methods

Except for general methodological improvements, many different approaches to seismic tomography have been developed since the late 1970s. These different approaches are related to the different parts of seismograms that are exploited for tomographic purposes. At first mainly first arrivals from *P*-waves were used, but nowadays travel-time residuals from all kinds of *P*-phases (e.g. Humphreys et al., 1984; Engdahl and Gubbins, 1987; Morelli and Dziewonski, 1987; Van der Hilst and Engdahl, 1991) and *S*-phases (e.g. Grand, 1987; Kuo et al., 1987; Su and Dziewonski, 1991; Katzman et al., 1998) or combinations of both (e.g. Abers and Roecker, 1991; Robertson and Woodhouse, 1996; Su and Dziewonski, 1997; Kennett et al., 1998) are inverted. Apart from seismic body-wave travel times, surface wave dispersion (e.g. Romanowicz, 1982; Hadiouche and Jobert, 1988; Mocquet and Romanowicz, 1990; Bourjot and Romanowicz, 1992) also constitutes a powerful means to obtain insight in (upper) mantle structure. Furthermore, not only arrival times are useful but the addition of amplitudes (Neele et al., 1993) or the use of entire waveforms (e.g. Woodhouse and Dziewonski, 1984; Nolet, 1987; Snieder, 1988; Nolet, 1990; Cichowicz and Green, 1992; Trampert and Woodhouse, 1995; Li and Romanowicz, 1996) can provide valuable extra information. In general, this additional information relates to different parts of the Earth, and its inclusion therefore provides constraints on model volumes otherwise poorly covered by *P* data only. Moreover, the application of additional body wave phases results in a more accurate hypocenter determination (e.g. Engdahl et al., 1998). Apart from body and surface waves and their

waveforms, normal modes (e.g. Masters et al., 1982; Giardini et al., 1987; Snieder, 1993; Resovsky and Ritzwoller, 1999) and differential travel times (Darragh, 1985; Kuo et al., 1987; Woodward and Masters, 1991) have also been employed for seismic tomography. Normal mode inversions use the eigenfrequencies of the Earth, which are mainly sensitive to its large-scale structure; differential travel times are derived from combinations of different body wave arrivals and have the advantage of hardly depending on near source/receiver structure.

In this thesis, the  $P$ ,  $pP$ , and  $pwP$  data (Chapter 4 and Bijwaard et al., 1998) and the  $S$  data (Chapter 6 and Bijwaard and Spakman, 1999a) from the reprocessed data set of Engdahl et al. (1998) are used for travel-time tomography. The large number of first ( $P$ ) arrivals, together with the extra hypocenter constraints provided by the depth phases ( $pP$  and  $pwP$ ), allow for the imaging of small-scale  $P$ -velocity structure in regions with a dense ray coverage. The smaller number of  $S$  (later) arrivals do not allow for such a detailed global solution, but correlation between  $S$  and  $P$  anomalies increases with ray sampling.

### 2.3 Different tomography scales

Seismic tomography is nowadays applied at a variety of scales. It is used locally (Kissling, 1988) in exploration methods (e.g. Gouly, 1993), to map relatively small structures like magma chambers under ocean ridges (Dunn and Toomey, 1997) or volcanoes (e.g. Thurber, 1983; Achauer et al., 1988; Cardaci et al., 1993; Chiarabba et al., 1995), for detailed images of the crust and uppermost mantle (e.g. McNamara et al., 1997), back-arc regions (Zhao et al., 1997), mountain belts (Souriau and Granet, 1995) and fault zones (Chiarabba et al., 1997), and for earthquake source regions (e.g. Amato et al., 1992; Eberhart-Phillips and Michael, 1998). On the regional scale, mainly collision and subduction zones have been imaged. Regions for which (mainly body-wave) tomography models have been derived include (among others) Papua New Guinea (e.g. Abers and Roecker, 1991), the California Coast Ranges (Eberhart-Phillips, 1986), North America (Van der Lee and Nolet, 1997), the Caribbean (Van der Hilst and Spakman, 1989), Cascadia (VanDecar, 1991), the Aleutian Islands (Engdahl and Gubbins, 1987), Alaska (Zhao et al., 1995), Japan (e.g. Hirahara, 1981), the Northwest Pacific (e.g. Spakman et al., 1989), Peru (Cunningham et al., 1986), South America (Engdahl et al., 1995), Taiwan (Roecker et al., 1987), Tanzania (Ritsema et al., 1998), Europe and the Mediterranean (e.g. Spakman et al., 1993), Tonga-Kermadec (Van der Hilst, 1995), Australia (Zielhuis and Van der Hilst, 1996), Iceland (Wolfe et al., 1997), China (Liu and Jin, 1993), India (Srinagesh and Rai, 1996), and Indonesia (e.g. Widiyantoro and Van der Hilst, 1997). These include most of the well-sampled regions of the Earth's (upper) mantle. On the global scale, different depth intervals of the Earth's mantle and core have been imaged with seismic tomography: The upper mantle (e.g. Woodhouse and Dziewonski, 1984; Woodward and Masters, 1991; Zhang and Tanimoto, 1993), the entire mantle (e.g. Hager and Clayton, 1989; Tanimoto, 1990; Su et al., 1994; Vasco et al., 1995a; Trampert and Woodhouse, 1996; Zhou, 1996; Van der Hilst et al., 1997) and the entire Earth (mantle and core) (Morelli and Dziewonski, 1986; Morelli and Dziewonski, 1987; Vasco and Johnson, 1998). The amount of detail obtained generally trades off with study volume: Large-scale models



usually employ coarse parameterizations and hence cannot resolve small-scale features. Resolution also degrades with increasing depth as we are looking at the Earth's interior through an unpolished lens (Romanowicz, 1991) of crustal and upper mantle heterogeneities. Furthermore, seismological data are generally less sensitive to deeper structure due to the larger width of Fresnel zones (Nolet, 1992) and wavefront healing (Wielandt, 1987).

Up to now the main achievements obtained with seismic tomography have probably been the following. On the local scale, spreading centers and magma chambers (e.g. Spence and Gross, 1990; Dunn and Toomey, 1997; Zhao et al., 1997), small-scale tectonic features (e.g. Amato et al., 1992; Chiarabba et al., 1997), and ore bodies (e.g. Goult, 1993) have been imaged. On the regional scale, detailed pictures of subducting, partly aseismic slabs (e.g. Spakman et al., 1988; Van der Hilst et al., 1991; Hirahara and Hasemi, 1993; Spakman et al., 1993), rift zones (Davis et al., 1993; Slack et al., 1996), collision zones (Abers and Roecker, 1991), and mantle plumes (e.g. Wolfe et al., 1997; Kennett and Widiyantoro, 1999) have been constructed. Globally, the main findings have been: The correlation of shallow anomalies with tectonic features as ridges, subduction zones and cratonic shields (e.g. Su et al., 1992), the discovery of the so-called Pacific ring of high velocities (e.g. Su et al., 1994), the detection of large high velocity anomalies in the lower mantle associated with past subductions (indicating exchange of mass between upper and lower mantle) (e.g. Inoue et al., 1990; Fukao et al., 1992; Grand, 1994; Van der Hilst et al., 1997), the observation of strong heterogeneities near the top and base ( $D''$ ) of the mantle (e.g. Hager et al., 1985; Obayashi and Fukao, 1997), and the discovery of inner-core anisotropy (e.g. Morelli and Dziewonski, 1987; Giardini et al., 1987).

In this thesis, travel-time tomography is performed for the entire Earth's mantle. The obtained amount of detail is, however, directly comparable to regional investigations. Many of the above findings are confirmed in this study (such as the detailed images of subductions in the upper mantle and the extension of some of these anomalies into the lower mantle), and several significant new findings are added. These include the first detailed image of a whole mantle plume (section 5.1, and Bijwaard and Spakman, 1999c), the identification of the completely subducted Mongol-Okhotsk ocean (section 5.2, and Van der Voo et al., 1999a), and a substantiation of the complex subduction history of the Tethys ocean (section 5.3, and Van der Voo et al., 1999b).

## 2.4 Specific ingredients of travel-time tomography

Confirmation of some well-known features has come from numerous tomography studies, but the exact tomographic images depend on the precise recipe that is used to create them. In the following a brief overview is given of the specific choices that can influence the tomographic results in the different stages of a (body-wave) tomography study.

### 2.4.1 Tomographic data

As discussed above, the type of data (first or later arrivals,  $P$  or  $S$  waves) is very important but apart from that the source of the data (ISC, Geoscope, GDSN or other) may also differ

and influence the solution (Spakman, 1993; Gudmundsson et al., 1990). The number of data, their geographical distribution, and the applied selection criteria are also significant. The latest global inversions use typically  $O(10^7)$  residuals from as large a set of earthquakes as possible and selection criteria usually include: Number of recordings per event, reliability of onset, magnitude of earthquake, epicentral distance, depth of event, magnitude of the residual and the stability of the hypocentral location. The exact criteria may differ substantially from one study to another.

Generally, the travel-time data set that is to be used for tomography is calculated with respect to a specific (radial) reference Earth model. As the tomographic problem is in principle nonlinear, because travel times depend on seismic velocity and ray path geometry, and the latter is itself dependent on seismic velocity, the difference between this reference model and the real Earth should be small to allow for the linearization of the tomographic problem. Van der Hilst and Spakman (1989) have shown that the frequently used Jeffreys-Bullen reference model (Jeffreys and Bullen, 1940) introduces artifacts in the solution. For local/regional tomography a local/regional reference model is needed (e.g. Hadiouche and Jobert, 1988; Kissling et al., 1994), for global tomography a reference model should be used that includes the main upper mantle discontinuities such as the PREM (Dziewonski and Anderson, 1981), IASP91 (Kennett, 1991), SP6 (Morelli and Dziewonski, 1993), or ak135 (Kennett et al., 1995) models.

Finally, in studies that employ large amounts of seismic data, some form of data averaging is normally used whereby seismic rays are combined into ray bundles (e.g. Spakman et al., 1993). This has the advantage of reducing the size of the inverse problem, decreasing the heterogeneity in model sampling, and increasing the signal-to-noise ratio. However, care should be taken that data are not averaged over large model volumes since this will prevent the imaging of small-scale structure. For example, in Su and Dziewonski (1997) summary rays are constructed by averaging over  $5^\circ \times 5^\circ$  (equal surface area) caps, thereby retaining only data sensitivity for larger scale (500-1000 km) structure.

In this thesis, a selection of data from the Engdahl et al. (1998) data set is used (see Chapters 4 and 6, and Bijwaard et al., 1998; Bijwaard and Spakman, 1999a, for exact selection criteria). This data set consists mainly of reprocessed ISC data for which new hypocenters and travel times have been determined in the ak135 reference model. In the tomographic inversion, the data are combined in composite rays (see Spakman and Nolet, 1988) derived from event cluster volumes of  $30 \times 30 \times 30 \text{ km}^3$ .

#### *2.4.2 Model parameterizations*

After data selection, a model parameterization must be chosen that approximates the (unknown) reality fairly well and, for feasibility reasons, does not introduce too many unknowns at the same time. The two most used parameterizations are blocks/nodes (local basis functions) (e.g. Engdahl and Gubbins, 1987; Spakman, 1991; Piromallo and Morelli, 1997) and spherical harmonic functions (global basis functions) (e.g. Dziewonski and Woodhouse, 1987; Tanimoto, 1990; Masters et al., 1996). Other parameterizations include regionalized models (Hadiouche and Jobert, 1988; Nataf et al., 1986), which is a sort of block parameterization in itself or a "without blocks" approach following Tarantola and Nercessian (1984). An advantage of spherical harmonics is the straightforward comparison

with other spherical harmonic models of, for example, the geoid (e.g. Hager and Clayton, 1989), although comparison of individual coefficients of some spectral order may not be reliable (Snieder, 1993). The main, practical advantage of block (or cell) parameterizations is that their implementations have allowed for very detailed imaging: Some of the latest global block models have a discretization of  $1^\circ$ - $2^\circ$  (Zhou, 1996; Van der Hilst et al., 1997), whereas the latest spherical harmonic models use parameterizations up to degree 12 (Su et al., 1994) and 16 (Masters et al., 1996), which is comparable to  $15^\circ$ - $11^\circ$  cells. An exception in this respect is the model of Trampert and Woodhouse (1996), which is parameterized up to degree 40 ( $\sim 4.5^\circ$  cells). However, the lateral resolution in their model is only 1000-1700 km (or  $9^\circ$ - $15^\circ$ ). An additional advantage of cell models (and local basis functions in general) is the possibility to apply finer parameterization in areas with better sampling (e.g. Abers and Roecker, 1991; Liu and Jin, 1993; Sambridge and Gudmundsson, 1998).

Some authors have claimed that the resolving power obtained with the latest block models is not needed, because there are no small-scale anomalies in the deeper parts of the mantle (Gudmundsson et al., 1990; Su and Dziewonski, 1991). Spherical harmonic models indicate that indeed higher order harmonics have little power and that low degree harmonics are dominant (degree 2 according to e.g. Nataf et al., 1986; Su et al., 1994, or degree 5 and 7 according to Nakanishi and Anderson, 1982). In the case of Su et al. (1994), this may (partially) be caused by the implemented summary rays, which are constructed from  $5^\circ \times 5^\circ$  surface caps and hence average over large mantle volumes. From their block model, Hager and Clayton (1989) obtain a nearly flat power spectrum up to degree 12. Furthermore, deep subduction and mantle plumes are expected to produce small-scale structure ( $\sim 200$  km), and in regional models that go as deep as 1200-1700 km (e.g. Spakman, 1993; Liu and Jin, 1993; Van der Hilst et al., 1993; Grand, 1994) such small structure is actually retrieved. The fact that it is not seen in global models based on spherical harmonics and their power spectra might just be due to the low expansion cut-off and the spectral leakage this introduces, as is demonstrated by Snieder et al. (1991) and Trampert and Snieder (1996). Moreover, the Fresnelzone of some of the phases used in spherical harmonic studies might be too broad to image any small-scale structure. For an *ScS* wave with a travelling distance of 6000 km and a frequency below 1/45 Hz. (as in Su et al., 1994) the Fresnel zone becomes as large as 900 km (where the Fresnel radius was taken to be  $r_f \approx \sqrt{(\lambda D/8)}$ , with  $D$  for distance and  $\lambda$  for wavelength, which was calculated using a shear velocity of 6 km/s).

In this thesis, irregular cell parameterizations are employed with small cells (down to  $0.6^\circ \times 0.6^\circ \times 35$  km) in regions with a high ray density and large cells (up to  $9^\circ \times 9^\circ \times 400$  km) in regions with a low ray density. In both the *P* and *S* tomography results (Chapters 4 and 6, and Bijwaard et al., 1998; Bijwaard and Spakman, 1999a) detailed structure on scales as small as 100-200 km is resolved.

### 2.4.3 Tomographic inversions

The actual inversion of the data is another point that deserves proper attention. Whenever possible a direct inversion should be preferred because it allows for the computation of resolution and covariance matrices. However, for large inverse problems this is not feasible

and an approximate inversion needs to be performed. Approximate tomographic inversions have been realized with techniques such as ART (Algebraic Reconstruction Technique) (e.g. Humphreys et al., 1984), SIRT (Simultaneous Iterative Reconstruction Technique) (e.g. Nolet, 1985; Zhou and Clayton, 1990), CG (Conjugate Gradients) (e.g. Scales, 1987), BPT (Back Projection Technique) (Ho-Liu et al., 1989) and LSQR (Least Squares) (e.g. Nolet, 1985; Spakman, 1991). The different methods have been compared in several studies (Neumann-Denzau and Behrens, 1984; Nolet, 1985; Van der Sluis and Van der Vorst, 1987; Spakman and Nolet, 1988), which conclude that LSQR and CG seem to perform best. Vasco et al. (1994) point out that usually least squares solutions are calculated with these algorithms, but this may not be desired because the residual distribution is generally not Gaussian. They apply a modified LSQR algorithm that performs an  $l^p$  minimization and their optimum  $p$  appears to be  $p=1.2$ .

As mentioned before, a considerable improvement of the tomographic method that has up to now only been applied in local tomography studies, is the iterative inversion: Residuals are inverted for 3-D velocity structure (and hypocenters), and for the new model, new residuals are calculated and inverted again, and so on. In this way, ray bending due to inferred velocity heterogeneity is taken into account. Theoretically, the best approach is to invert simultaneously for velocities and hypocenters, but it is computationally efficient to use parameter separation (e.g. Inoue et al., 1990) and invert first for hypocenters and then for velocities. To calculate new residuals for the iterative procedure, a 3-D ray tracer is needed, but this slows down the computation and often some sort of approximate ray tracing is used (e.g. Thurber, 1983; Cardaci et al., 1993).

For the actual inversion of seismic data into velocities (and hypocenters), some specific choices need to be made. Because the tomographic matrix is usually inconsistent (due to data errors) and underdetermined (due to lack of data and inhomogeneous sampling), some form of regularization (or damping) needs to be applied to further constrain the solution. Apart from that, it could also be profitable to implement a scaling (or preconditioning) of the equations in order to improve the solutions for the most important parameters. In general, when inverting for different kinds of parameters (such as hypocenter locations and seismic velocities) attention should be paid to the scaling between them. There is no unique solution to the problem of finding the optimum damping and scaling parameters. The assessment of resolution together with a thorough evaluation of the results provide the only means in tuning them.

In this thesis, simultaneous tomographic inversions for cell slownesses, relocation vectors, and station statics are performed with the LSQR algorithm (basically following Spakman and Nolet, 1988). Every result that is presented has been obtained after extensive experimentation with different damping and scaling parameters. Furthermore, for the first time nonlinear (iterative) inversion is applied to the global problem (Chapter 8). For this purpose, the approximate 3-D ray tracer of Chapter 7 (Bijwaard and Spakman, 1999d) was used.

#### *2.4.4 Sensitivity analyses*

Generally, for global tomographic inverse problems with large numbers of data (several million) and large numbers of unknowns (several hundreds of thousands), it is

computationally unfeasible to calculate formal estimates of resolution (with perhaps the exception of Vasco et al., 1999, who calculate 5000 Lanczos vectors for a problem with 820000 data and 96000 unknowns). In the absence of resolution and covariance matrices some other measure has to be applied to assess the quality of the solution. Synthetic tests, in which some test structure is imaged with the applied ray set, approximate idealized estimates of resolution. Despite significant shortcomings of these tests (see Chapter 4 and Bijwaard et al., 1998), they still provide the only way to estimate resolution for some of the huge inverse problems encountered in global tomography. The impulse response test (e.g. Spakman and Nolet, 1988; Zhou, 1988) shows if and how input spikes are smeared in the solution; checkerboard tests (Inoue et al., 1990) and inversions for synthetic structures like slabs (Spakman et al., 1989) provide further insight in possible artifacts in the solution. The Permuted Data Test (Spakman, 1991), in which a randomly permuted version of the data vector is inverted, provides an upper estimate of the influence of incoherent data signal on the model. Inversions in which data are replaced by Gaussian noise, can also indicate how random data errors may contaminate the result, but Gaussian statistics underlie most of the inversion methods used, and such a test may thus provide a poor estimate of the influence of actual data errors that need not be Gaussian distributed. Apart from that, the actual noise level in the data is generally unknown and needs first to be estimated. This noise level may be obtained following the approach of Zhou (1988) and Gudmundsson (1990) who determine the variance of the data from events in decreasing volumes; The residual variance in an infinitesimal small volume should then (in the limit) represent the noise level. Other indicators of solution quality that have been employed, are: Hitcount (denser sampled regions are probably better resolved) or the ray density tensor (Kissling, 1988), variance reduction (how well are the data explained by the model), comparisons with models obtained using independent data (e.g. Grand et al., 1997) or the comparison of synthetic seismograms with seismograms that were not used for the inversion (e.g. Woodhouse and Dziewonski, 1984), inversions with shifted block interfaces (Cunningham et al., 1986; Kuge and Satake, 1987), and inversions with or without (jackknife) random subsets of the data (Inoue, 1993; Vasco et al., 1994). In this thesis, model resolution is estimated from impulse response and layer cake synthetic tests (see e.g. Chapter 4 and Bijwaard et al., 1998). Generally, the results from impulse response tests are used to infer some measure of lateral resolution, and the results from layer cake tests to infer vertical resolution. Apart from that, model variance is estimated with the above mentioned Permuted Data Test.

## 2.5 Shortcomings and future directions

The basic shortcoming of global tomographic inversion, is the absence of proper resolution and variance analyses due to the generally huge size of the problem. However, the idealized estimates derived from sensitivity tests, the increasing correspondence between different tomographic studies and their correlation with geology/tectonic reconstructions, together with some major theoretical improvements, have added to the credibility of tomographic images. One should, however, not forget that seismic tomography does in general not account for, e.g. source geometry/mechanism, anelasticity (with the exception of Romanowicz, 1995), scattering (with the exception of Snieder, 1988), the width of

Fresnelzones (with the exception of Vasco et al., 1995b), and anisotropy (with the exception of e.g. Nataf et al., 1986; Montagner and Nataf, 1988; Montagner and Jobert, 1988; Montagner and Tanimoto, 1991; Ekström and Dziewonski, 1998). Future improvements of tomographic results might thus come from the inclusion of these effects in either the forward or inverse problem, together with (among other things): the incorporation of more data with a better global coverage (especially ocean bottom array data), variable depths of upper mantle discontinuities in the inversion (as in Katzman et al., 1998; Vasco et al., 1998), the use of regionalized reference models that account for first order differences between for example sub-ocean and sub-continent velocities, and combined inversions with other geophysical data such as gravity measurements. These improvements would probably lead to a more accurate tomographic image that could provide new constraints on other geophysical studies involving, e.g. convection models, tectonic reconstructions and gravity interpretations.

## References

- Abers, G.A., and S.W. Roecker, Deep structure of an arc-continent collision: earthquake relocation and inversion for upper mantle  $P$  and  $S$  wave velocities beneath Papua New Guinea, *J. Geophys. Res.*, 96, 6379-6401, 1991.
- Achauer, U., J.R. Evans, and D.A. Stauber, High-resolution seismic tomography of compressional wave velocity structure at Newberry Volcano, Oregon Cascade Range, *J. Geophys. Res.*, 93, 10135-10147, 1988.
- Aki, K., and W.H.K. Lee, Determination of the three-dimensional velocity anomalies under a seismic array using first  $P$  arrival times from local earthquakes, 1, A homogeneous initial model, *J. Geophys. Res.*, 81, 4381-4399, 1976.
- Aki, K., A. Christoffersson, and E.S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, *Geophys. J. R. Astron. Soc.*, 82, 277-296, 1977.
- Amato, A., C. Chiarabba, L. Malagnini, and G. Selvaggi, Three-dimensional  $P$ -velocity structure in the region of the  $M_S=6.9$  Irpinia, Italy, normal faulting earthquake, *Phys. Earth Planet. Inter.*, 75, 111-119, 1992.
- Bijwaard, H., and W. Spakman, Advancing global  $S$ -wave tomography toward the regional scale, *Geophys. J. Int.* (submitted), 1999a.
- Bijwaard, H., and W. Spakman, Nonlinear global  $P$ -wave tomography, *Geophys. J. Int.* (submitted), 1999b.
- Bijwaard, H., and W. Spakman, Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sci. Lett.*, 166, 121-126, 1999c.
- Bijwaard, H., and W. Spakman, Fast kinematic ray tracing of first and later arriving seismic phases, *Geophys. J. Int.* (in press), 1999d.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Bourjot, L., and B. Romanowicz, Crust and upper mantle tomography in Tibet using surface waves, *Geophys. Res. Lett.*, 19, 881-884, 1992.
- Cardaci, C., M. Coviello, G. Lombardo, G. Patane, and R. Scarpa, Seismic tomography

- of Etna volcano, *J. Volc. Geotherm. Res.*, 56, 357-368, 1993.
- Chiarabba, C., A. Amato, and J.R. Evans, Variations on the NeHT high-resolution tomography method: A test of the technique and results for Medicine Lake Volcano, northern California, *J. Geophys. Res.*, 100, 4035-4052, 1995.
- Chiarabba, C., A. Amato, and M. Meghraoui, Tomographic images of the El Asnam fault zone and the evolution of a seismogenic thrust-related fold, *J. Geophys. Res.*, 102, 24485-24498, 1997.
- Cichowicz, A., and R.W.E. Green, Tomographic study of upper-mantle structure of the South African continent using wave-form inversion, *Phys. Earth Planet. Inter.*, 72, 276-285, 1992.
- Clayton, R.W., and R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel time data, *EOS Trans. AGU*, 64, 776, 1983.
- Cunningham, P.S., S.W. Roecker, and D. Hatzfeld, Three-dimensional *P* and *S* wave structures of southern Peru and their tectonic implications, *J. Geophys. Res.*, 91, 9517-9532, 1986.
- Curtis, A., and R.K. Snieder, Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, 62, 1524-1532, 1997.
- Darragh, R.B., Mapping of upper-mantle structure from differential (*PP-P*) travel time residuals, *Phys. Earth Planet. Inter.*, 41, 6-17, 1985.
- Davis, P., P. Slack, H.A. Dahlheim, W.V. Green, R.P. Meyer, U. Achauer, A. Glahn, and M. Granet, Teleseismic tomography of the continental rift zones, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 397-439, Chapman and Hall, London, 1993.
- Dunn, R.A., and D.R. Toomey, Seismological evidence for 3D melt migration beneath the East Pacific Rise, *Nature*, 388, 259-262, 1997.
- Dziewonski, A.M., and D.L. Anderson, Preliminary Reference Earth Model, *Phys. Earth Planet. Inter.*, 25, 297-356, 1981.
- Dziewonski, A.M., and J.H. Woodhouse, Global images of the Earth's interior, *Science*, 236, 37-48, 1987.
- Dziewonski, A.M., B.H. Hager, and R.J. O'Connell, Large-scale heterogeneity in the lower mantle, *J. Geophys. Res.*, 82, 239-255, 1977.
- Eberhart-Phillips, D., Three-dimensional velocity structure in northern California Coast Ranges from inversion of local earthquake arrival times, *Bull. Seismol. Soc. Amer.*, 76, 1025-1052, 1986.
- Eberhart-Phillips, D., and A.J. Michael, Seismotectonics of the Loma Prieta, California, region determined from three-dimensional  $V_p$ ,  $V_p/V_s$ , and seismicity, *J. Geophys. Res.*, 103, 21099-21120, 1998.
- Ekström, G., and A.M. Dziewonski, The unique anisotropy of the Pacific upper mantle, *Nature*, 394, 168-172, 1998.
- Engdahl, E.R., and D. Gubbins, Simultaneous travel time inversion for earthquake location and subduction zone structure in the central Aleutian Islands, *J. Geophys. Res.*, 92, 13855-13862, 1987.
- Engdahl, E.R., R.D. van der Hilst, and J. Berrocal, Imaging of subducted lithosphere beneath South America, *Geophys. Res. Lett.*, 22, 2317-2320, 1995.
- Engdahl, E.R., R.D. van der Hilst, and R.P. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination,

- Bull. Seismol. Soc. Amer., 88, 722-743, 1998.
- Fukao, Y., M. Obayashi, H. Inoue, and M. Nenbai, Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.*, 97, 4809-4822, 1992.
- Giardini, D., X.-D. Li, and J.H. Woodhouse, Three-dimensional structure of the Earth from splitting in free oscillation spectra, *Nature*, 325, 405-411, 1987.
- Gouly, N.R., Controlled-source tomography for mining and engineering applications, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 797-813, Chapman and Hall, London, 1993.
- Grand, S.P., Tomographic inversion for shear velocity beneath the North American plate. *J. Geophys. Res.*, 92, 14065-14090, 1987.
- Grand, S.P., Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.*, 99, 11591-11621, 1994.
- Grand, S.P., R.D. van der Hilst, and S. Widiyantoro, Global seismic tomography: A snapshot of convection in the Earth, *GSA Today*, 7, 1-7, 1997.
- Gudmundsson, O., J.H. Davies, and R.W. Clayton, Stochastic analysis of global travel time data: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, 102, 25-43, 1990.
- Gutenberg, B., *Der Aufbau der Erdkruste auf Grund geophysikalischer Betrachtungen*, *Z. Geophys.*, 1, 94-108, 1924.
- Hadiouche, O., and N. Jobert, Geographical distribution of surface wave velocities and 3-D upper mantle structure in Africa. *Geophys. J. Int.*, 95, 87-109, 1988.
- Hager, B.H., R.W. Clayton, M.A. Richards, R.P. Comer, and A.M. Dziewonski, Lower mantle heterogeneity, dynamic topography, and the geoid, *Nature*, 313, 541-545, 1985.
- Hager, B.H., and R.W. Clayton, Constraints on the structure of mantle convection using seismic observations, flow models, and the geoid, in: *Mantle Convection*, edited by W.R. Peltier, pp. 657-763, Gordon and Breach, New York, 1989.
- Hirahara, K., Three-dimensional seismic structure beneath southwest Japan: the subducting Philippine sea plate, *Tectonophysics*, 79, 1-44, 1981.
- Hirahara, K., and A. Hasemi, Tomography of subduction zones using local and regional earthquakes and teleseisms, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 519-562, Chapman and Hall, London, 1993.
- Ho-Liu, P., J.P. Montagner, and H. Kanamori, Comparison of iterative back-projection inversion and generalized inversion without blocks: case studies in attenuation tomography, *Geophys. J. Int.*, 97, 19-29, 1989.
- Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res.*, 93, 1073-1085, 1988.
- Humphreys, E., R.W. Clayton, and B.H. Hager, A tomographic image of mantle structure beneath southern California. *Geophys. Res. Lett.*, 11, 621-625, 1984.
- Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogata, Whole mantle *P*-wave travel time tomography, *Phys. Earth Planet. Inter.*, 59, 294-328, 1990.
- Jeffreys, H., and K.E. Bullen, *Seismological tables*, British Assoc. for the Advancement of Science, Gray-Milne Trust, 1940.
- Katzman, R., L. Zhao, and T.H. Jordan, High-resolution, two-dimensional vertical tomography of the central Pacific mantle using *ScS* reverberations and frequency-dependent travel times, *J. Geophys. Res.*, 103, 17933-17971, 1998.



- Kennett, B.L.N., IASPEI 1991 Seismological Tables, Research School of Earth Sciences, Australian National University, 1991.
- Kennett, B.L.N., S. Widiyantoro, A low seismic wavespeed anomaly beneath northwestern India: a seismic signature of the Deccan plume?, *Earth Planet. Sci. Lett.*, 165, 145-156, 1999.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, 122, 108-124, 1995.
- Kennett, B.L.N., S. Widiyantoro, and R.D. van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, 103, 12469-12493, 1998.
- Kissling, E., Geotomography with local earthquake data, *Rev. Geophys.*, 26, 659-698, 1988.
- Kissling, E., W.L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer, Initial reference models in local earthquake tomography, *J. Geophys. Res.*, 99, 19635-19646, 1994.
- Kuge, K., and K. Satake, Lateral segmentation within the subducting lithosphere: three-dimensional structure beneath the North Island, New Zealand, *Tectonophysics*, 139, 223-237, 1987.
- Kuo, B.-Y., D.W. Forsyth, and M. Wyssession, Lateral heterogeneity and azimuthal anisotropy in the North Atlantic determined from *SS-S* differential travel times. *J. Geophys. Res.*, 92, 6421-6436, 1987.
- Li, X.-D., and B. Romanowicz, Global mantle shear velocity model developed using nonlinear asymptotic coupling theory, *J. Geophys. Res.*, 101, 22245-22272, 1996.
- Liu, F., and A. Jin, Seismic tomography of China, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 299-318, Chapman and Hall, London, 1993.
- Masters, G., T.H. Jordan, P.G. Silver, and F. Gilbert, Aspherical earth structure from fundamental spheroidal mode data, *Nature*, 298, 609-613, 1982.
- Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London Ser. A*, 354, 1385-1411, 1996.
- McNamara, D.E., W.R. Walter, T.J. Owens, and C.J. Ammon, Upper mantle velocity structure beneath the Tibetan Plateau from *P<sub>n</sub>* travel time tomography, *J. Geophys. Res.*, 102, 493-505, 1997.
- Mocquet, A., and B. Romanowicz, Three dimensional structure of the upper mantle beneath the Atlantic ocean inferred from long-period Rayleigh waves, 1: Group and phase velocity distributions, *J. Geophys. Res.*, 94, 7449-7468, 1990.
- Montagner, J.P., and H.C. Nataf, Vectorial tomography, 1, Theory, *Geophys. J. R. Astron. Soc.*, 94, 295-307, 1988.
- Montagner, J.P., and N. Jobert, Vectorial tomography, 2, Application to the Indian ocean, *Geophys. J. Astron. Soc.*, 94, 309-344, 1988.
- Montagner, J.P., and T. Tanimoto, Global upper mantle tomography of seismic velocities and anisotropies, *J. Geophys. Res.*, 96, 20337-20351, 1991.
- Morelli, A., and A.M. Dziewonski, 3-D structure of the Earth's core inferred from travel time residuals. *EOS*, 67, 311, 1986.
- Morelli, A., and A.M. Dziewonski, The harmonic expansion approach to the retrieval of deep Earth structure, in: *Seismic Tomography*, edited by G. Nolet, pp. 251-274, Reidel, Dordrecht, 1987.

- Morelli, A., and A.M. Dziewonski, Body-wave traveltimes and a spherically symmetric *P*- and *S*-wave velocity model, *Geophys. J. Int.*, 112, 178-194, 1993.
- Nakanishi, I., and D.L. Anderson, Worldwide distribution of group velocity of mantle Rayleigh waves as determined by spherical harmonic inversion, *Bull. Seismol. Soc. Amer.*, 72, 1185-1194, 1982.
- Nataf, H.C., I. Nakanishi, and D.L. Anderson, Measurement of mantle wave velocities and inversion for lateral heterogeneities and anisotropy 3: inversion, *J. Geophys. Res.*, 91, 7261-7307, 1986.
- Neele, F., J. VanDecar, and R. Snieder, The use of *P* wave amplitude data in a joint inversion with travel times for upper mantle velocity structure, *J. Geophys. Res.*, 98, 12033-12054, 1993.
- Neumann-Denzau, G., and J. Behrens, Inversion of seismic data using tomographical reconstruction techniques for investigations of laterally inhomogeneous media, *Geophys. J. R. Astron. Soc.*, 79, 305-315, 1984.
- Nolet, G., Solving or resolving inadequate and noisy tomographic systems, *J. Comput. Phys.*, 61, 463-482, 1985.
- Nolet, G., Waveform tomography, in: *Seismic Tomography*, edited by G. Nolet, pp. 301-322, Reidel, Dordrecht, 1987.
- Nolet, G., Partitioned waveform inversion and two-dimensional structure under the Network of Autonomously Recording Seismographs, *J. Geophys. Res.*, 95, 8499-8512, 1990.
- Nolet, G., Imaging the deep earth: technical possibilities and theoretical limitations, in *Proceedings of the 25<sup>th</sup> ESC Symposium Barcelona 1990*, edited by A. Rocca and D. Mayer-Rosa, pp. 107-115, Servei Geològic de Catalunya, Barcelona, Spain, 1992.
- Obayashi, M., and Y. Fukao, *P* and *PcP* travel time tomography for the core-mantle boundary, *J. Geophys. Res.*, 102, 17825-17841, 1997.
- Papazachos, C., and G. Nolet, *P* and *S* deep velocity structure of the Hellenic area obtained by robust nonlinear inversion of traveltimes, *J. Geophys. Res.*, 102, 8349-8367, 1997.
- Piromallo, C., and A. Morelli, Imaging the Mediterranean upper mantle by *P*-wave travel time tomography, *Ann. Geofisica*, 40, 963-979, 1997.
- Resovsky, J.S., M.H. Ritzwoller, A degree 8 mantle shear velocity model from normal mode observations below 3 MHz, *J. Geophys. Res.*, 104, 993-1014, 1999.
- Ritsema, J., A.A. Nyblade, T.J. Owens, C.A. Langston, and J.C. VanDecar, Upper mantle seismic velocity structure beneath Tanzania, east Africa: Implications for the stability of cratonic lithosphere, *J. Geophys. Res.*, 103, 21201-21214, 1998.
- Ritzwoller, M.H., and E.M. Lavelly, Three-dimensional seismic models of the Earth's mantle, *Rev. Geophys.*, 33, 1-66, 1995.
- Robertson, G.S., and J.H. Woodhouse, Ratio of relative *S* to *P* velocity heterogeneity in the lower mantle, *J. Geophys. Res.*, 101, 20041-20052, 1996.
- Roecker, S.W., Y.H. Yeh, and Y.B. Tsai, Three-dimensional *P* and *S* wave velocity structures beneath Taiwan: deep structure beneath an arc-continent collision, *J. Geophys. Res.*, 92, 10547-10570, 1987.
- Romanowicz, B., Constraints on the structure of the Tibet Plateau from pure phase

- velocities of Love and Rayleigh waves, *J. Geophys. Res.*, 87, 6865-6883, 1982.
- Romanowicz, B., Seismic tomography of the Earth's mantle, *Annu. Rev. Earth Planet. Sci.*, 19, 77-99, 1991.
- Romanowicz, B., A global tomographic model of shear attenuation in the upper mantle, *J. Geophys. Res.*, 100, 12375-12394, 1995.
- Sambridge, M., and O. Gudmundsson, Tomographic systems of equations with irregular grids, *J. Geophys. Res.*, 103, 773-781, 1998.
- Scales, J.A., Tomographic inversion via the conjugate gradient method, *Geophysics*, 52, 179-185, 1987.
- Sengupta, M., and M.N. Toksöz, Three-dimensional model of seismic velocity variation in the Earth's mantle, *Geophys. Res. Lett.*, 3, 84-86, 1976.
- Slack, P.D., P.M. Davis, W.S. Baldrige, K.H. Olsen, A. Glahn, U. Achauer, and W. Spence, The upper mantle structure of the central Rio Grande Rift region from teleseismic *P* and *S* wave travel time delays and attenuation, *J. Geophys. Res.*, 101, 16003-16024, 1996.
- Snieder, R.K., Large-scale waveform inversions of surface waves for lateral heterogeneity, 2. Application to surface waves in Europe and the Mediterranean, *J. Geophys. Res.*, 93, 12067-12080, 1988.
- Snieder, R.K., Global inversions using normal modes and long-period surface waves, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 23-63, Chapman and Hall, London, 1993.
- Snieder, R.K., J. Beckers, and F. Neele, The effect of small-scale structure on normal mode frequencies and global inversions, *J. Geophys. Res.*, 96, 501-515, 1991.
- Souriau, A., and M. Granet, A tomographic study of the lithosphere beneath the Pyrenees from local and teleseismic data, *J. Geophys. Res.*, 100, 18117-18134, 1995.
- Spakman, W., Delay-time tomography of the upper mantle below Europe, the Mediterranean and Asia Minor, *Geophys. J. Int.*, 107, 309-332, 1991.
- Spakman, W., Iterative strategies for non-linear travel time tomography using global earthquake data, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 190-226, Chapman and Hall, London, 1993.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic inverse problems (manuscript in preparation), 1999.
- Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, pp. 155-187, Reidel, Dordrecht, 1988.
- Spakman, W., S. van der Lee, and R.D. van der Hilst, Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, 79, 3-74, 1993.
- Spakman, W., M.J.R. Wortel, and N.J. Vlaar, The Hellenic subduction zone: a tomographic image and its geodynamic implications, *Geophys. Res. Lett.*, 15, 60-63, 1988.
- Spakman, W., S. Stein, R.D. van der Hilst, and M.J.R. Wortel, Resolution experiments for NW Pacific subduction zone tomography, *Geophys. Res. Lett.*, 16, 1097-1100, 1989.
- Spence, W., and R.S. Gross, A tomographic glimpse of the upper mantle source of magmas of the Jemez lineament, New Mexico, *J. Geophys. Res.*, 95, 10829-

- 10849, 1990.
- Spencer, C., and D. Gubbins, Travel-time inversion for simultaneous earthquake location and velocity structure determination in laterally varying media, *Geophys. J. R. Astron. Soc.*, 63, 95-116, 1980.
- Srinagesh, D., and S.S. Rai, Teleseismic tomographic evidence for contrasting crust and upper mantles in south Indian Archean terrains, *Phys. Earth Planet. Inter.*, 97, 27-41, 1996.
- Su, W.-J., and A.M. Dziewonski, Predominance of long-wavelength heterogeneity in the mantle, *Nature*, 352, 121-126, 1991.
- Su, W.-J., and A.M. Dziewonski, Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth Planet. Inter.*, 100, 135-156, 1997.
- Su, W.-J., R.L. Woodward, and A.M. Dziewonski, Deep origin of mid-ocean-ridge seismic velocity anomalies, *Nature*, 360, 149-152, 1992.
- Su, W.-J., R.L. Woodward, and A.M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, 99, 6945-6980, 1994.
- Tanimoto, T., Long-wavelength *S*-wave velocity structure throughout the mantle, *Geophys. J. Int.*, 100, 327-336, 1990.
- Tarantola, A., and A. Nercessian, Three-dimensional inversion without blocks, *Geophys. J. R. Astron. Soc.*, 76, 299-306, 1984.
- Thomson, C.J., and D. Gubbins, Three-dimensional lithospheric modelling at NORSAR: linearity of the method and amplitude variations from the anomalies, *Geophys. J. R. Astron. Soc.*, 71, 1-36, 1982.
- Thurber, C.H., Earthquake locations and three-dimensional structure in the Coyote Lake area, central California, *J. Geophys. Res.*, 88, 8226-8236, 1983.
- Trampert, J., and R.K. Snieder, Model estimations biased by truncated expansions: possible artifacts in seismic tomography, *Science*, 271, 1257-1260, 1996.
- Trampert, J., and J.H. Woodhouse, Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.*, 122, 675-690, 1995.
- Trampert, J., and J.H. Woodhouse, High resolution global phase velocity distributions, *Geophys. Res. Lett.*, 23, 21-24, 1996.
- VanDecar, J.C., Upper-mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion, Ph.D. thesis, 165 pp., Univ. Of Wash., Seattle, 1991.
- Van der Hilst, R.D., Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, *Nature*, 374, 154-157, 1995.
- Van der Hilst, R.D., and E.R. Engdahl, On ISC *PP* and *pP* data and their use in delay-time tomography of the Caribbean region, *Geophys. J. Int.*, 106, 169-188, 1991.
- Van der Hilst, R.D., and W. Spakman, Importance of the reference model in linearized tomographic inversions: images of subduction below the Caribbean plate, *Geophys. Res. Lett.*, 16, 1093-1097, 1989.
- Van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, 386, 578-584, 1997.
- Van der Hilst, R.D., E.R. Engdahl, W. Spakman, and G. Nolet, Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature*, 353, 37-43, 1991.
- Van der Hilst, R.D., E.R. Engdahl, W. Spakman, Tomographic inversion of *P* and *pP* data

- for aspherical mantle structure below the northwest Pacific region, *Geophys. J. Int.*, 115, 264-302, 1993.
- Van der Lee, S., and G. Nolet, The upper-mantle *S* velocity structure of North America, *J. Geophys. Res.*, 102, 22815-22838, 1997.
- Van der Sluis, A., and H.A. van der Vorst, Numerical solution of large, sparse linear algebraic systems arising from tomographic problems, in: *Seismic Tomography*, edited by G. Nolet, pp. 49-83, Reidel, Dordrecht, 1987.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Mesozoic subducted slabs under Siberia, *Nature*, 397, 246-249, 1999a.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.*, 171, 7-20, 1999b.
- Vasco, D.W., and L.R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, 103, 2633-2671, 1998.
- Vasco, D.W., L.R. Johnson, and O. Marques, Global Earth structure: inference and assessment, *Geophys. J. Int.*, 381-407, 1999.
- Vasco, D.W., L.R. Johnson, and R.J. Pulliam, Lateral variations in mantle velocity structure and discontinuities determined from *P*, *PP*, *S*, *SS*, and *SS-SdS* travel time residuals, *J. Geophys. Res.*, 100, 24037-24059, 1995a.
- Vasco, D.W., J. John, E. Peterson, and E.L. Majer, Beyond ray tomography: Wavepath and Fresnel volumes, *Geophysics*, 60, 1790-1804, 1995b.
- Vasco, D.W., L.R. Johnson, R.J. Pulliam, and P.S. Earle, Robust inversion of IASP91 travel time residuals for mantle *P* and *S* velocity structure, earthquake mislocations, and station corrections, *J. Geophys. Res.*, 99, 13727-13755, 1994.
- Vesnaver, A.L., Irregular grids in seismic tomography and minimum-time ray tracing, *Geophys. J. Int.*, 126, 147-165, 1996.
- Widiyantoro, S., and R.D. van der Hilst, Mantle structure beneath Indonesia inferred from high-resolution tomographic imaging, *Geophys. J. Int.*, 130, 167-182, 1997.
- Wielandt, E., On the validity of the ray approximation for interpreting delay times, in: *Seismic Tomography*, edited by G. Nolet, pp. 85-98, Reidel, Dordrecht, 1987.
- Wolfe, C.J., I.Th. Bjarnason, J.C. VanDecar, and S.C. Solomon, Seismic structure of the Iceland mantle plume, *Nature*, 185, 245-247, 1997.
- Woodhouse, J.H., and A.M. Dziewonski, Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, 89, 5953-5986, 1984.
- Woodward, R.L., and G. Masters, Global upper mantle structure from long-period differential travel times, *J. Geophys. Res.*, 96, 6351-6377, 1991.
- Zhang, Y.-S., and T. Tanimoto, High-resolution global upper mantle structure and plate tectonics, *J. Geophys. Res.*, 98, 9793-9823, 1993.
- Zhao, D., D. Christensen, H. Pulpan, Tomographic imaging of the Alaska subduction zone, *J. Geophys. Res.*, 100, 6487-6504, 1995.
- Zhao, D., Y. Xu, D.A. Wiens, L. Dorman, J. Hildebrand, and S. Webb, Depth extent of the Lau back-arc spreading center and its relation to subduction processes, *Science*, 278, 254-257, 1997.
- Zhou, H.-W., How well can we resolve the deep seismic slab with seismic tomography?, *Geophys. Res. Lett.*, 15, 1425-1428, 1988.

- 
- Zhou, H.-W., A high-resolution *P* wave model for the top 1200 km of the mantle, *J. Geophys. Res.*, 101, 27791-27810, 1996.
- Zhou, H.-W., and R.W. Clayton, *P* and *S* travel time inversions for subducting slab under the island arcs of the northwest Pacific, *J. Geophys. Res.*, 95, 6829-6851, 1990.
- Zielhuis, A., and R.D. van der Hilst, Upper mantle shear velocity beneath eastern Australia from inversion of waveform SKIPPY portable arrays, *Geophys. J. Int.*, 127, 1-16, 1996.

# Chapter 3

## A brief summary of theory and irregular parameterization

This chapter is divided in two parts: A short summary of the essential travel-time tomography theory based on Spakman (1988), Spakman and Nolet (1988) and Spakman (1993), and the specific formulation of an irregular cell parameterization based on Spakman and Bijwaard (1999).

### 3.1 Travel-time tomography theory

Numerous derivations of the basic travel-time tomography equations exist in the literature (e.g. Van der Hilst and Engdahl, 1992; Zhang and Tanimoto, 1993; Vasco et al., 1994; Chiarabba et al., 1995). For completeness, this section is devoted to a short summary of the essential theory, based upon Spakman (1988), Spakman and Nolet (1988) and Spakman (1993).

In seismic travel-time tomography, as it is presented in this thesis, arrival times of body waves are used to determine wavespeeds, earthquake locations, and station statics, in a coupled inverse problem. Apart from a possible trade-off between the different parameters, an extra complication is introduced by the fact that the travel-time tomography problem is intrinsically nonlinear: Travel times depend on wavespeed and path geometry and the latter is in itself again dependent on wavespeed. Assuming that seismic body wave energy travels along rays and thereby neglecting Fresnelzones, this nonlinearity is expressed by the basic travel-time equation:

$$T = \int_L s \, dl \quad , \quad L = L(s) \quad (3.1)$$

where  $T$  is travel time,  $L$  represents the ray path,  $s$  stands for slowness (the reciprocal of velocity), and  $dl$  is the integration step length along the ray. In order to deduce slownesses (and thus velocities) by inversion of this equation, it needs to be linearized around an approximate ray path  $L_0$ . The accuracy of this approximation depends on how well  $L_0$  approximates  $L$ . Approximate ray paths can be defined by the introduction of a reference velocity model (which is usually 1-D) and corresponding earthquake locations. Together with inferred earthquake origin times, these generally explain the largest part of the observed arrival times. The remaining part, the residual or delay, can be used to infer 3-D variations of seismic velocities, improvements in hypocenter coordinates, and station statics. The delay is defined as the difference between observed and calculated arrival times or in terms of travel times as:

$$d = T(s) - T_0(s_0) + \Delta t^e + \Delta t^s + \epsilon = \int_L s \, dl - \int_{L_0} s_0 \, dl_0 + \Delta t^e + \Delta t^s + \epsilon \quad (3.2)$$

where  $d$  stands for delay, the subscripts  $0$  refer to reference model quantities,  $\Delta t^e$  and  $\Delta t^s$  are earthquake timing and station timing differences, and  $\epsilon$  represents an observational error (e.g. a phase misidentification or a picking error). Note that  $T_0$  is calculated from the reference source location through the reference velocity model to the reference station location. In reality, the earthquakes may have location errors in the reference medium of several kilometers. Station locations may also be in error (see e.g. Röhms et al., 1999), but this is ignored in the following.

Equation (3.2) can be linearized by invoking Fermat's principle (i.e. assuming that  $s_0$  is sufficiently close to  $s$  such that the difference between  $L_0$  and  $L$  is small and leads to travel-time differences of higher order), and using a first order Taylor expansion to account for the effect of source mislocations. These approximations lead to the following equation for the delay:

$$d = \int_{L_0} \Delta s \, dl_0 + \Delta \mathbf{x}^e \cdot \nabla_0 T_0 + \Delta t^s + \epsilon \quad (3.3)$$

where  $\Delta s = s - s_0$ ,  $\Delta \mathbf{x}^e$  is a vector of source mislocation coefficients together with origin time error, and  $\nabla_0$  is a gradient vector with respect to reference source position and origin time. The error  $\epsilon$  now also absorbs approximation errors (i.e. higher order terms).

The station term and source mislocation are discrete quantities, and hence for the transformation to a discrete problem only the slowness part needs to be parameterized. In this thesis this is performed with (irregular) orthonormal cell functions  $c_j$  that vanish everywhere except inside the cell they represent (cell  $j$ ). In this way the slowness anomaly  $\Delta s$  can be projected on the cell parameterization as follows:

$$\Delta s = \sum_{j=1}^N \Delta s_j c_j + \epsilon_p \quad (3.4)$$

where the sum extends over all ( $N$ ) cells, the  $\Delta s_j$  are the projection coefficients (cell amplitudes), and  $\epsilon_p$  represents a projection error. Inserting the parameterization in Equation (3.3), leads to:



$$d = \sum_{j=1}^N \Delta s_j \left( \int_{L_0} c_j dl \right) + \Delta x^e \cdot \nabla_0 T_0 + \Delta t^s + \epsilon \quad (3.5)$$

where the order of summation and integration has been changed and  $\epsilon$  now also includes the projection error. This equation can be rewritten for delay  $i$  (from a data set of  $M$  delays) as:

$$d_i \approx \sum_{j=1}^N \Delta s_j l_{ij} + \sum_{k=1}^4 \Delta x_k^e g_{ik} + \Delta s_l^s l_{il} \quad (3.6)$$

where the  $l_{ij}$  represent ray segments that follow from the integration of cell functions over the reference ray path, the  $g_{ik}$  stand for the elements of  $\nabla_0 T_0$ , and the station timing error has been replaced by an unknown cell slowness (for station  $l$ ) times a ray segment in that cell. The unknown error  $\epsilon$  has been dropped and implicitly adds to the error in  $d_i$ . Combining all unknowns ( $\Delta s_j$ ,  $\Delta x_k^e$ ,  $\Delta s_l^s$ ) into a single model vector  $\mathbf{m}$  and all coefficients ( $l_{ij}$ ,  $g_{ik}$ ,  $l_{il}$ ) into a matrix  $\mathbf{A}$ , leads to a tomographic matrix-vector equation of the form:

$$\mathbf{A}\mathbf{m} = \mathbf{d} \quad (3.7)$$

Solving this system of equations leads to a linearized solution to the tomography problem, consisting of cell slowness anomalies, earthquake relocation vectors, and station statics. This solution can be used in a ray tracing step to compute new (3-D) ray paths, travel times and travel-time residuals. The new residuals can again be inverted for slownesses, relocation vectors, and station statics, and so on. This iterative process is commonly denoted as nonlinear tomography (see Chapter 8), although it essentially constitutes a series of linear steps. The tomographic inversions in this process are identical to the inversion in linearized tomography except for the updated (3-D) ray paths and the correction of the travel-time residuals prior to each inversion step. In this thesis the residuals are corrected with Equation (8.4), but other corrections have also been suggested and lead to different iterative schemes (see Spakman, 1993). These schemes are mathematically equivalent, but they may lead to different solutions because of extra constraints that are needed to solve (3.7).

Due to modeling and observational errors the system of equations in (3.7) is generally inconsistent (e.g. two identical ray paths may have different travel times and residuals, because of picking errors). Furthermore, although the number of data usually exceeds the number of unknowns, the system is generally underdetermined as well (e.g. some cells in the model volume may not contain any ray paths). This requires a damped, least squares approach (e.g. Tarantola and Valette, 1982) to solve the system in an approximate way: Extra (damping) equations are added (e.g. Chapter 4 and Bijwaard et al., 1998, equation 3) to constrain ill-determined model parameters which leads to an augmented system of equations.

In practice, the augmented system is often too large to solve by direct (or generalized) inversion (e.g. using Singular Value Decomposition). Some size reduction can be obtained by combining ray paths in summary or composite rays and clustering the events, but for large tomographic problems this is not enough (e.g. the system of equations in Chapter 4 still contains 4.5 million equations and more than 400000 unknowns). Fortunately, several row-action methods are available to calculate an approximate solution. For the calculations in this thesis 200 iterations of the LSQR algorithm (Paige and Saunders, 1982) were used (see e.g. Chapter 4).

The use of a limited number of LSQR iterations (which is unavoidable due to the large computation time involved) is very similar to cutting off the eigenvalue spectrum with Singular Value Decomposition (Van der Sluis and Van der Vorst, 1987). It may be profitable to implicitly modify this spectrum and scale the model parameters to emphasize the more interesting parts of the solution. This finally leads to a tomographic matrix-vector equation of the form of Equation 4.2.

### 3.2 Irregular cell parameterization of tomographic inverse problems

In the first section of this chapter, the general travel-time tomography problem is parameterized with a set of unspecified, but orthonormal cell functions. This section, which is based on Spakman and Bijwaard (1999), deals with the specific formulation of an irregular cell parameterization.

#### 3.2.1 Introduction

In most global tomography studies, the Earth is parameterized with either local cell functions or spherical harmonics. The main advantage of cell parameterizations (or other local parameterizations) is their allowance for spatially detailed solutions in well-sampled model volumes (see also Chapter 2). So far, the global tomography applications of this parameterization have employed either equiangular or equal-surface-area (ESA) cells. For cell sizes that are commonly used in regional tomography studies (50-100 km), such a parameterization would lead to huge amounts of cells, many of which are poorly sampled by seismic rays (e.g. in the shallow mantle beneath the oceans). It therefore seems sensible to employ irregularly shaped cells, the sizes of which should depend on some measure of the resolving power of the applied data. This would lead to a large reduction in the number of unknowns and a better conditioning of the inverse problem.

Several authors have studied the implementation of irregular grids for tomographic purposes (e.g. Fukao et al., 1992; Michelini, 1995; Vesnaver, 1996; Curtis and Snieder, 1997; Sambridge and Gudmundsson, 1998). This resulted, however, in either computationally expensive parameterizations or rather limited irregular parameterizations consisting of stacks of regular grids. The method proposed in this chapter is based on an idea of Abers and Roecker (1991), in which the study volume is divided in small, regular basic cells which are subsequently used as the building blocks of larger, irregular cells. In the implementation of Abers and Roecker (1991) the construction of larger cells was done by hand, but this no longer feasible for global applications. The subject of this section is therefore mainly the

efficient automation of this process for applications to global travel-time tomography (Chapters 4 and 6, and Bijwaard et al., 1998; Bijwaard and Spakman, 1999).

### 3.2.2 Theory

As mentioned above, a drawback of many irregular parameterizations is the computational complexity they introduce. This is avoided here by the introduction of a basic grid of smallest cells to be used in the tomographic inversion. These non-overlapping and orthonormal cell functions  $b_j$  are zero everywhere except in basic cell  $j$  and serve as the building blocks for the larger, irregular cells. For computational efficiency they are equi-angular and positioned in layers of not necessarily equal thickness. Irregular, non-overlapping, and orthonormal cells  $c_k$  are defined as the sum of a particular set of basic cell functions:

$$c_k = \sum_{j=1}^{N_b} P_{kj} b_j \quad (3.8)$$

where  $N_b$  is the number of basic cells and  $P_{kj}$  is zero, except when  $b_j$  is a building block of  $c_k$ . In principle,  $P_{kj}$  is an  $N_i \times N_b$  sized projection matrix ( $N_i$  being the number of irregular cells), but since every basic cell can only be an element of a single irregular cell, every column of  $P_{kj}$  contains only one non-zero element.

After the implementation of these irregular cells in Equation (3.5), the tomographic matrix elements become ray arc lengths in the irregular cells. These may be computationally expensive to calculate which can be avoided because the arc lengths in the irregular cells are always a sum of arc lengths in the basic cells:

$$l_{ik}^c = \sum_{j=1}^{N_b} P'_{kj} l_{ij}^b \quad (3.9)$$

where the  $l_{ik}^c$  are matrix elements for the irregular grid, the  $l_{ij}^b$  are matrix elements for the basic grid, and the  $P'_{kj}$  are the projection coefficients, which become one when basic cell  $j$  is a building block of irregular cell  $k$ , and zero if this is not so. After the construction of the basic tomography matrix on the basic cells, for which the ray tracing requires the normal computational effort, this matrix can be projected with the projection matrix  $P'_{kj}$  on an irregular grid. Because every basic cell can only be an element of a single irregular cell, the matrix  $P'_{kj}$  contains only one non-zero element per column and can thus conveniently be stored in a pointer array  $k=p(j)$  of length  $N_b$ .

### 3.2.3 Construction of an irregular cell model for global tomography

This paragraph outlines the construction of the global, irregular parameterization that is used in Chapter 4. The basic idea behind this parameterization is that every irregular cell should be sampled by approximately the same number of rays, and that hence every cell slowness is determined by approximately the same number of data. This can only be achieved by varying cell volumes through the model. The solution to the inverse parameterization problem (determine all sizes such that hitcount has minimum variation) can be found by forward computation.

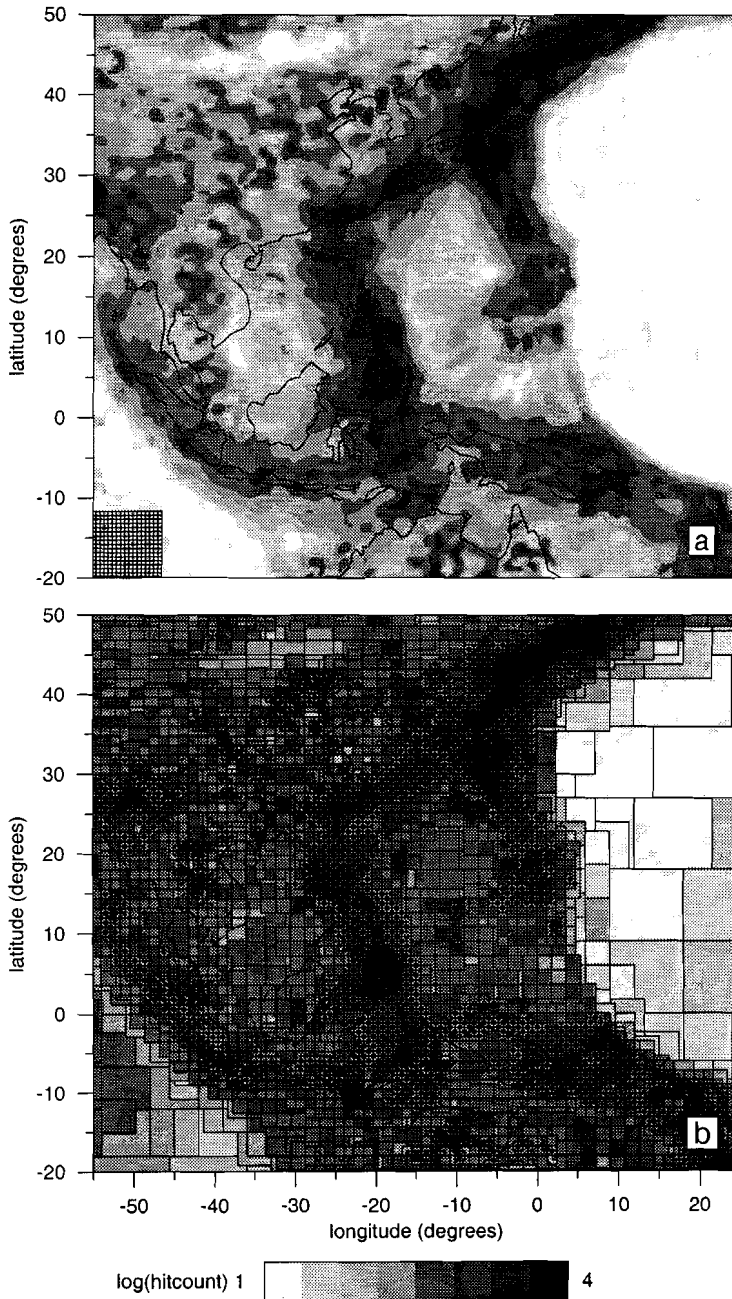
First, a basic grid of conical, equi-angular  $b_j$  cells of  $0.6^\circ$  longitude and latitude is defined throughout the Earth's mantle in 26 layers with boundaries at the following depths: 0, 35, 70, 120, 170, 230, 290, 350, 410, 470, 530, 595, 660, 760, 860, 980, 1100, 1250, 1400, 1600, 1800, 2000, 2200, 2400, 2560, 2720, and 2889 km. This yields a total number of  $N_b=600 \times 300 \times 26=4680000$  basic cells.

In order to build an irregular grid from the basic cells, a pointer  $k=p(j)$  needs to be defined. This is accomplished in the following steps. First, a series of ESA grids is constructed from the basic cells. For every ESA grid, a latitude interval  $\Delta\theta=\theta_2-\theta_1$ , centered around  $\theta_c=0.5(\theta_1+\theta_2)$ , determines a longitude interval  $\Delta\phi=\Delta\theta/\cos(\theta_c)$ . A chosen  $\Delta\theta=n \times 0.6^\circ$  hence gives a  $\Delta\phi$ , which is round off to the nearest  $m \times 0.6^\circ$ . Layer thicknesses are chosen such that the thickness of each ESA cell is approximately equal to its horizontal dimensions, but within the restrictions provided by the basic layering. To construct  $k=p(j)$ , the midpoints of all basic cells  $j$  are calculated, and from this the ESA cell number  $k$  is determined.

Following the procedure described above, ESA grids are determined for cell sizes of  $0.6^\circ$ ,  $1.2^\circ$ ,  $1.8^\circ$ ,  $2.4^\circ$ ,  $3.0^\circ$ ,  $4.2^\circ$ ,  $6.0^\circ$ , and  $9.0^\circ$ . The constructed pointer arrays then allow for the projection of the basic tomography matrix (for the  $0.6^\circ$  basic cells) on each ESA grid which yields ESA hitcount maps, which are expanded on the basic grid (i.e. every basic cell element obtains the hitcount value of the irregular cell it belongs to). In the next step these hitcount maps are used to break down large background cells into smaller cells wherever the hitcount remains above a target value. After some experimenting the target values were set at 500 rays per cell for the upper mantle, and 1000 rays per cell for the lower mantle. Smallest cells in the lower mantle are restricted to sizes larger than approximately 100 km. The initial  $k=p(j)$  is filled with ESA background cells of  $9.0^\circ$ .

The final irregular grid is then constructed by looping over the ESA grids from coarse to fine, and by looping over all basic cells for each ESA grid. In the inner loop, larger cells are replaced by smaller ones wherever the hitcount remains above the target value. This leads to updates of the pointer array  $k=p(j)$ , which is filled with new  $k$  values where smaller cells are being implemented. After finishing the loops, the final pointer array is sorted and the irregular cells are renumbered in closed sequence. Every basic cell then still uniquely relates to a single irregular cell. An example of the effect of the irregular parameterization on the hitcount pattern beneath southeast Asia is depicted in Figure 3.1. In general, the irregular cell parameterization derived here, is fully determined by the employed basic cells, the ESA hitcount maps, and the target hitcount values. Within these constraints the irregular cells minimize hitcount differences between adjacent cells.

The final parameterization contains 277565 irregular cells (out of 4680000 basic cells), which is approximately equal to the number of cells in an ESA grid with  $1.8^\circ$  cells in 20 layers. The irregular cells, however, allow for the determination of 60 km detail wherever



**Figure 3.1:** Example of hitcount patterns at a depth of 500 km below southeast Asia for the  $0.6^\circ$  ESA grid that is depicted in the lower left corner (a) and the irregular grid employed for global tomography in Chapter 4 (b). Notice that even the large background cells of the irregular grid do not contain many ray paths below the Pacific ocean.

allowed by the ray sampling. Compared to the construction of the tomography matrix on the basic cells (which costs approximately 10 hours computation time for the tomography problem in Chapter 4), the determination of the irregular grid is computationally cheap: each ESA hitcount map takes 30 min., and the final grid construction only 15 min. This makes experimenting with different irregular parameterizations rather cheap, once the basic tomography matrix has been computed.

### 3.2.4 Conclusions

An efficient method has been developed for the implementation of irregular cells in tomography problems. The method can easily be implemented in existing tomography software, because it only requires a projection of the conventional tomography matrix. The irregular grid is fully determined by a pointer array  $k=p(j)$  that describes which regular, basic cells  $j$  are the building blocks of irregular cells  $k$ . The main computational investment in the tomography problem remains the construction of the tomography matrix on the basic cells. The projection on irregular grids requires little CPU time, and hence allows for experimenting with different irregular grids. In general, these irregular grids may lead to a sensible reduction in the number of model parameters (by eliminating many small cells in poorly sampled model volumes), and a better conditioning of the inverse problem. For the particular case of global travel-time tomography as described in Chapter 4, an irregular grid was constructed with large cells in model parts that were poorly sampled by seismic rays and small cells in densely sampled regions, such that every cell slowness is determined by approximately the same number of seismic rays.

### References

- Abers, G.G., and S.W. Roecker, Deep structure of an arc-continent collision: Earthquake relocation and inversion for upper mantle  $P$  and  $S$  wave velocities beneath Papua New Guinea, *J. Geophys. Res.*, 96, 6379-6401, 1991.
- Bijwaard, H., and W. Spakman, Advancing global  $S$ -wave tomography towards the regional scale, *Geophys. J. Int.* (submitted), 1999.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Chiarabba, C., A. Amato, and J.R. Evans, Variations on the NeHT high-resolution tomography method: A test of the technique and results for Medicine Lake Volcano, northern California, *J. Geophys. Res.*, 100, 4035-4052, 1995.
- Curtis, A., and R.K. Snieder, Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, 62, 1524-1532, 1997.
- Fukao, Y., M. Obayashi, H. Inoue, and M. Nishii, Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.*, 97, 4809-4822, 1992.
- Michelini, A., An adaptive-grid formalism for travel time tomography, *Geophys. J. Int.*, 121, 489-510, 1995.
- Paige, C.C., M.A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.*, 8, 43-71, 1982.

- Röhm, A.H.E., J. Trampert, H. Paulssen, and R.K. Snieder, Bias in reported seismic arrival times deduced from the ISC bulletin, *Geophys. J. Int.*, 137, 163-174, 1999.
- Sambridge, M., and O. Gudmundsson, Tomographic systems of equations with irregular grids, *J. Geophys. Res.*, 103, 773-781, 1998.
- Spakman, W., Upper mantle delay time tomography, with an application to the collision zone of the Eurasian, African and Arabian plates, Ph.D. thesis, 200 pp., Utrecht university, 1988.
- Spakman, W., Iterative strategies for non-linear travel time tomography using global earthquake data, in: *Seismic Tomography; theory and practice*, edited by H.M. Iyer and K. Hirahara, pp. 190-226, Chapman and Hall, New York, 1993.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic inverse problems (manuscript in preparation), 1999.
- Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, pp.155-187, Reidel, Dordrecht, 1988.
- Tarantola, A, and B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.*, 20, 219-232, 1982.
- Van der Hilst, R.D., and E.R. Engdahl, Step-wise relocation of ISC earthquake hypocenters for linearized tomographic imaging of slab structure, *Phys. Earth Planet. Inter*, 75, 39-53, 1992.
- Vasco, D.W., L.R. Johnson, R.J. Pulliam, and P.S. Earle, Robust inversion of IASP91 travel time residuals for mantle *P* and *S* velocity structure, earthquake mislocations and station corrections, *J. Geophys. Res.*, 99, 13727-13755, 1994.
- Vesnaver, A.L., Irregular grids in seismic tomography and minimum-time ray tracing, *Geophys. J. Int.*, 126, 147-165, 1996.
- Van der Sluis, A., and H.A. van der Vorst, Numerical solution of large, sparse linear algebraic systems arising from tomographic problems, in: *Seismic Tomography*, edited by G. Nolet, pp. 49-83, Reidel, Dordrecht, 1987.
- Zhang, Y.-S., and T. Tanimoto, High-resolution global upper mantle structure and plate tectonics, *J. Geophys. Res.*, 98, 9793-9823, 1993.

# Chapter 4<sup>1</sup>

## Closing the gap between regional and global travel-time tomography

**Abstract.** Recent global travel-time tomography studies by Zhou (1996) and Van der Hilst et al. (1997) have been performed with cell parameterizations on the order of those frequently used in regional tomography studies (i.e., with cell sizes of  $1^{\circ}$ - $2^{\circ}$ ). These new global models constitute a considerable improvement over previous results that were obtained with rather coarse parameterizations ( $5^{\circ}$  cells). The inferred structures are, however, of larger scale than is usually obtained in regional models, and it is not clear where and if individual cells are actually resolved. This study aims at resolving lateral heterogeneity on scales as small as  $0.6^{\circ}$  in the upper mantle and  $1.2^{\circ}$ - $3^{\circ}$  in the lower mantle. This allows for the adequate mapping of expected small-scale structures induced by, for example, lithosphere subduction, deep mantle upwellings, and mid-ocean ridges. There are three major contributions that allow for this advancement. First, we employ an irregular grid of nonoverlapping cells adapted to the heterogeneous sampling of the Earth's mantle by seismic waves (Spakman and Bijwaard, 1998). Second, we exploit the global data set of Engdahl et al. (1998), which is a reprocessed version of the global data set of the International Seismological Centre. Their reprocessing included hypocenter redetermination and phase reidentification. Finally, we combine all data used ( $P$ ,  $pP$ , and  $pWP$  phases) into nearly 5 million ray bundles with a limited spatial extent such that averaging over large mantle volumes is prevented while the signal-to-noise ratio is improved. In the approximate solution of the huge inverse problem we obtain a variance reduction of 57.1%. Synthetic sensitivity tests indicate horizontal resolution on the scale of the smallest cells ( $0.6^{\circ}$  or  $1.2^{\circ}$ ) in the shallow parts of subduction zones decreasing to approximately  $2^{\circ}$ - $3^{\circ}$  resolution in well-sampled regions in the lower mantle. Vertical resolution can be worse (up to several hundreds of kilometers) in subduction zones with rays predominantly pointing along dip. Important features of the solution are as follows: 100-200 km thick high-velocity slabs beneath all major subduction zones, sometimes flattening in the transition zone and sometimes directly penetrating into the lower mantle; large high-velocity anomalies in the lower mantle that have been attributed to subduction of the Tethys ocean and the Farallon plate; and low-velocity anomalies continuing across the 660 km discontinuity to hotspots at the surface under Iceland, east Africa, the Canary Islands, Yellowstone, and the Society Islands. Our findings corroborate that the 660 km boundary may resist but not prevent (present day) large-scale mass transfer from upper to lower mantle or vice versa. This observation confirms the results of previous, global mantle studies that employed coarser parameterizations.

---

<sup>1</sup> This chapter has been published as: Bijwaard, H., W. Spakman, and E.R. Engdahl, *Closing the gap between regional and global travel time tomography*, *J. Geophys. Res.*, 103, 30055-30078, 1998.



## 4.1 Introduction

Imaging the three-dimensional velocity structure of the Earth with seismic tomography was pioneered in the late 1970s. An early local tomography study by Aki et al. (1977) used a lithosphere parameterization of 20 km blocks, whereas in early global studies by Sengupta and Toksöz (1976) and by Dziewonski et al. (1977) the Earth's mantle was parameterized with  $10^\circ$  cells and spherical harmonics up to degree 3 (which is in resolution comparable with  $60^\circ$  cells), respectively. In the early 1980s the global parameterization was refined considerably by Clayton and Comer (1983), who implemented 100 km thick,  $5^\circ$  wide cells. Meanwhile, spherical harmonic representations were extended to include degree 7 (Nakanishi and Anderson, 1982) and degree 8 (Woodhouse and Dziewonski, 1984). Although the general tomography method has improved over the years in many ways, the global cell parameterization scales have, until recently, not decreased any further. For example, Inoue et al. (1990) inverted  $P$  data for velocities in  $5.6^\circ$  cells, Zhang and Tanimoto (1993) used surface wave data for a  $5^\circ$  cell model, and Vasco et al. (1995) employed  $P$ ,  $PP$ ,  $S$ ,  $SS$ , and  $SS-SdS$  phases for a model with  $6^\circ$  cells. Parameterization in terms of spherical harmonics has been extended to degree 12 (Su et al., 1994) and 16 (Masters et al., 1996), which is in resolution comparable to  $15^\circ$ - $11^\circ$  cells. However, lateral heterogeneity caused by, for example, lithosphere subduction, deep mantle upwellings, and mid-ocean ridges, is expected to occur on a scale of  $0.5^\circ$ - $2^\circ$  in the Earth's mantle. Several regional studies (e.g. Zhou and Clayton, 1990; Engdahl et al., 1995; Spakman et al., 1993; Van der Hilst, 1995; Widiyantoro and Van der Hilst, 1996; Zhao et al., 1995) have, indeed, imaged subducting slabs at this scale, which was not resolved in the aforementioned global studies.

Only recently, model parameterizations in global  $P$ -wave travel-time tomography studies have been refined to a scale that approaches the small cells on the order of  $1^\circ$  that have frequently been used in regional studies. Van der Hilst et al. (1997) use  $2^\circ$  cells in a whole mantle model, and Zhou (1996) parameterizes the upper 1200 km of the Earth's mantle with a series of grids, the smallest cells of which are only  $1^\circ$  in longitude and latitude. However, neither study discusses the resolution and interpretation of structures at the smallest cell scales. Van der Hilst et al. (1997) focus on lower mantle structure and mainly discuss large-scale (more than  $5^\circ$ ) high-velocity lower mantle anomalies, which they interpret as subducted Mesozoic ocean lithosphere that in some places seems to be connected with well-known regional slab structures. Zhou (1996) estimates resolution only on scales of  $5^\circ$  and  $15^\circ$ , and his inversion results, do in our opinion, hardly show structures on a scale of  $1^\circ$ . He does show high-velocity anomalies below some subduction zones, but not all known slabs are imaged, and those that are, do not always resemble the plate-like slabs in regional studies.

The aim of this study is to take recent improvements in global tomography a step further, attempting to resolve upper mantle heterogeneity on scales as small as  $0.6^\circ$  in densely sampled regions and thereby to close the resolution gap between global and regional mantle tomography studies. Because the resolving power decreases in the lower mantle because of nongeometrical ray effects (Wielandt, 1987; Nolet, 1992), we aim to resolve lower mantle structure of  $1.2^\circ$ - $3^\circ$  and larger. We thereby expect to fully exploit the resolving power of the available seismic data in the entire mantle volume. However, we do not take into account the bending of seismic rays by velocity anomalies (which may be significant for  $0.6^\circ$  cells). This is the subject of further study (Bijwaard et al., 1998).

The main contributions that lead to our advancement are the following: First, the implementation of a model parameterization with cells of variable sizes adapted to the very inhomogeneous sampling of the Earth's mantle by the seismic rays used (Spakman and Bijwaard, 1998). Second, the implementation of the reprocessed global data set of Engdahl et al. (1998) (which was also used by Van der Hilst et al. (1997)) that is computed against the reference Earth model (ak135) of Kennett et al. (1995). Finally, we combine all  $P$ ,  $pP$ , and  $pwP$  data into nearly 5 million narrow ray bundles such that averaging over large mantle volumes is avoided while the signal-to-noise ratio is improved. The combination of these contributions leads to a better determined inverse problem.

Apart from improving on the global modeling done so far, there are several advantages of our global model over the existing set of regional models. The global mantle model corroborates findings in regional mantle studies and tectonic implications inferred from these (such as the fate and behavior of subducted lithosphere and the deep origins of hotspots) in a global framework, and it enables us to compare regions in which similar tectonic processes take place. Furthermore, it also shows detailed upper mantle features of well-sampled regions that have never been studied in detail such as the Scotia Sea, northwest Africa, Iran, and the Solomon Islands. In addition, the proposed model can serve as input for several other types of study, such as tectonic reconstructions and global 3-D earthquake relocation. Finally, regional solutions can be influenced by the signal and noise of teleseismic data acquired outside the model region (Masson and Trampert, 1997). It turns out that the differences between our global model and existing regional solutions (e.g. Spakman et al., 1993; Van der Hilst et al., 1993) are largest in the lower mantle part of the regional models, which may indicate the mapping of structure from outside these regions into the regional models. This does, however, not affect the major findings of these studies.

This paper should be considered an introduction to our global mantle model of  $P$ -wave velocity heterogeneity and the applied method. Detailed interpretations and further comparisons with results from sensitivity analyses will be given elsewhere. Our primary aim is to present the model and to show the improvement obtained in imaging global mantle structure. In our opinion, the results close the resolution gap with regional mantle studies.

## 4.2 Data

We exploit a reprocessed version of the International Seismological Centre (ISC) data set supplemented with recent data from the U.S. Geological Survey's National Earthquake Information Center and from several temporary deployments of seismic stations (Engdahl et al., 1998). The data processing included phase reidentification, travel-time recalculation through the ak135 reference model (Kennett et al., 1995), and hypocenter redetermination (including  $P$  and  $S$ ,  $PKiKP$ ,  $PKPdf$ , and the teleseismic depth phases  $pP$ ,  $pwP$ , and  $sP$  in the location procedure). The data set comprises over 82000 well-constrained earthquakes and a total of 12 million first- and later-arriving seismic phases observed in the period 1964-1995. Engdahl et al. (1998) conclude from regionally systematic location shifts and a reduction of scatter in Wadati-Benioff zone seismicity that hypocenter locations have been significantly improved. Furthermore, density plots of travel-time residuals against epicentral distance no longer display the well-known dependence of ISC delay times on epicentral distance that indicates deviations of the reference model velocities from the layer-averaged

real Earth. From this database, we select 7.6 million teleseismic (i.e., from epicentral distances larger than  $25^\circ$ )  $P$ ,  $pP$ , and  $pwP$  data with travel-time residuals between  $-3.5$  and  $+3.5$  s and regional ( $< 25^\circ$ )  $P$  phases with absolute residuals smaller than  $7.5$  s. For  $pP$  and  $pwP$  phases, only events with hypocenters deeper than  $35$  km are included. All data are corrected for the Earth's ellipticity and station elevations, and the  $pP$  and  $pwP$  data are also corrected for bounce point topography and water depth, respectively.

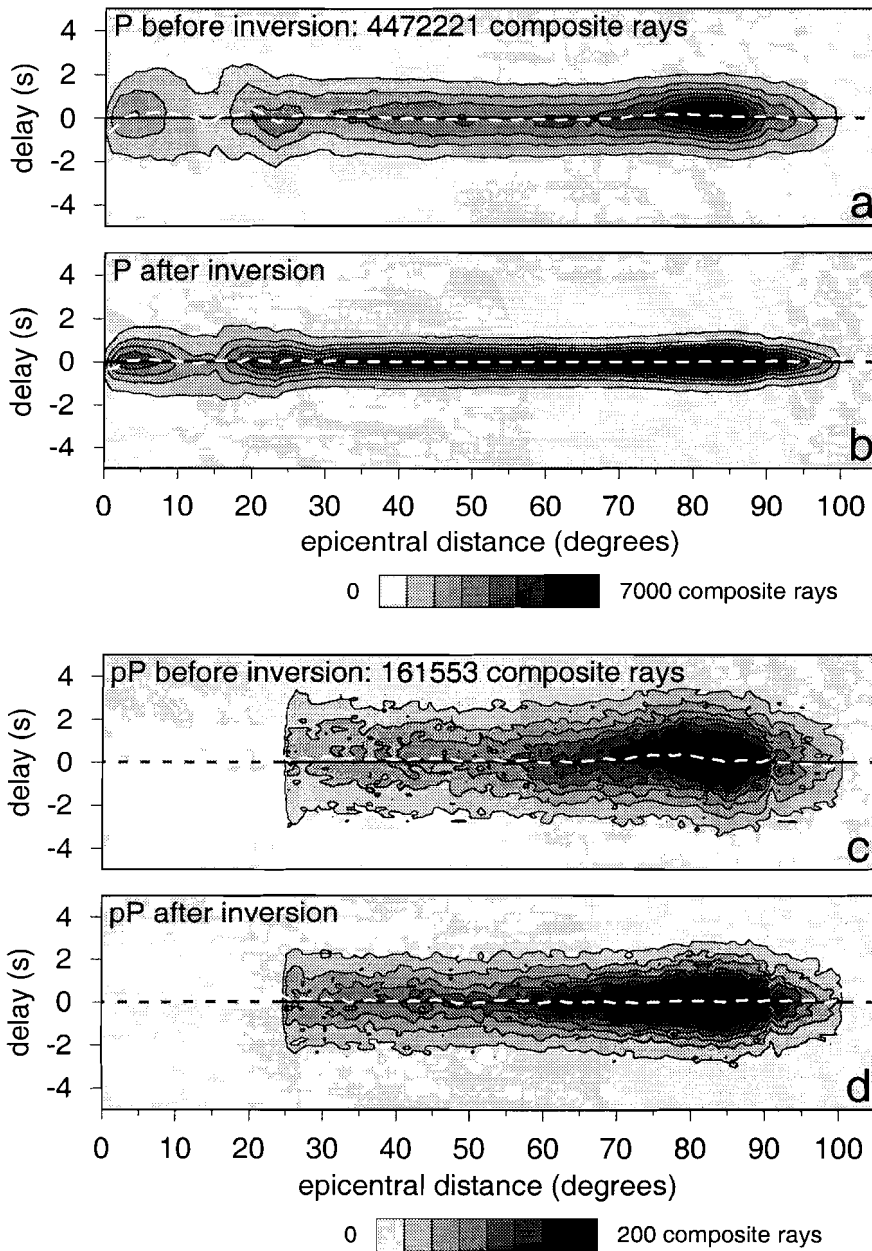
The data set of Engdahl et al. (1998) has also been used for tomographic purposes by Van der Hilst et al. (1997), who grouped the  $P$  and  $pP$  data into summary rays. In general, the use of ray bundles reduces the number of data and thus leads to a smaller inverse problem. Furthermore, the combined data are distributed more equally through the model, and the signal-to-noise ratio is increased. We therefore bundle the rays as well, but we combine the data into composite rays (Spakman and Nolet, 1988) constructed from 1 to 220 individual rays (the maximum number of single rays in the data set that fall within a ray bundle). A summary ray is a single ray that represents some average of a ray bundle; a composite ray is a ray bundle forged from all rays from an event cluster volume of, in our case,  $30 \times 30 \times 30$  km to a single station. A summary ray intersects only the cells traversed by the single ray it is represented by, whereas a composite ray intersects all cells traversed by rays in the ray bundle. We calculate composite ray residuals as the average of the original delays, allowing rays with sharp initial onsets to weigh twice as heavily as those with emergent onsets. Furthermore,  $P$  (and  $Pn$ ),  $pP$ , and  $pwP$  phases are weighted with their specific standard deviations as determined from the raw data ( $1.3$ ,  $1.4$ , and  $1.4$  s, respectively). The small event cluster volume of our composite rays leads to many more, narrow ray bundles ( $4.7$  million) than the number of summary rays that have been used by Van der Hilst et al. (1997) (namely,  $500000$ ). This prevents the averaging of data over large ray bundle volumes, which is prerequisite for achieving small-scale resolution.

The clustering of the  $7.6$  million single rays into  $4.7$  million composite rays divided over  $34000$  event clusters reduces the data variance by  $16.7\%$ . In order to adapt to the different ray bundle sizes, we weighted the data used prior to inversion:

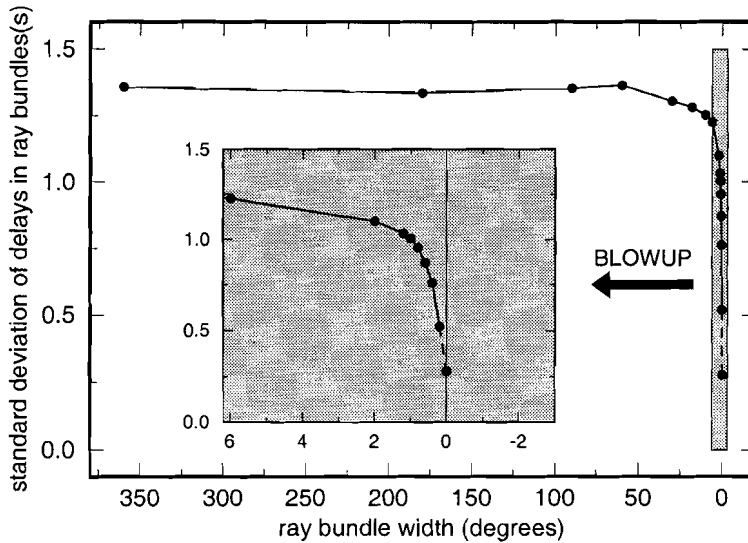
$$W_{rb}^{-1} = \sqrt{\frac{\sum_i w_i^2 (\langle dt \rangle - dt_i)^2}{\sum_i w_i^2}} \quad (4.1)$$

where  $W_{rb}$  represents the ray bundle weight,  $w_i$  is the weight of ray  $i$  in the ray bundle,  $dt_i$  is the delay of ray  $i$ , and  $\langle dt \rangle$  is the average delay of the ray bundle. The total weights of the different composite rays were, however, restricted to vary over one order of magnitude only. Residual density plots for the data used are shown in Figures 4.1a and 4.1c. Although the data distribution is flat, we still observe small fluctuations in the averages as a function of epicentral distance (denoted by the dashed white lines), indicating small deviations of the reference model from the best possible average for the data used.

To estimate the noise level in the data, we examined the delay time variance in ray bundles which are decreasing in width (following Gudmundsson et al. (1990)). For perfect measurements one expects the residual variance to drop to zero in infinitely narrow bundles,



**Figure 4.1:** Density plots of (a)  $P$  delays before inversion, (b)  $P$  delays after inversion, (c)  $pP$  delays before inversion, and (d)  $pP$  delays after inversion. Dashed white lines denote average residual per epicentral distance.

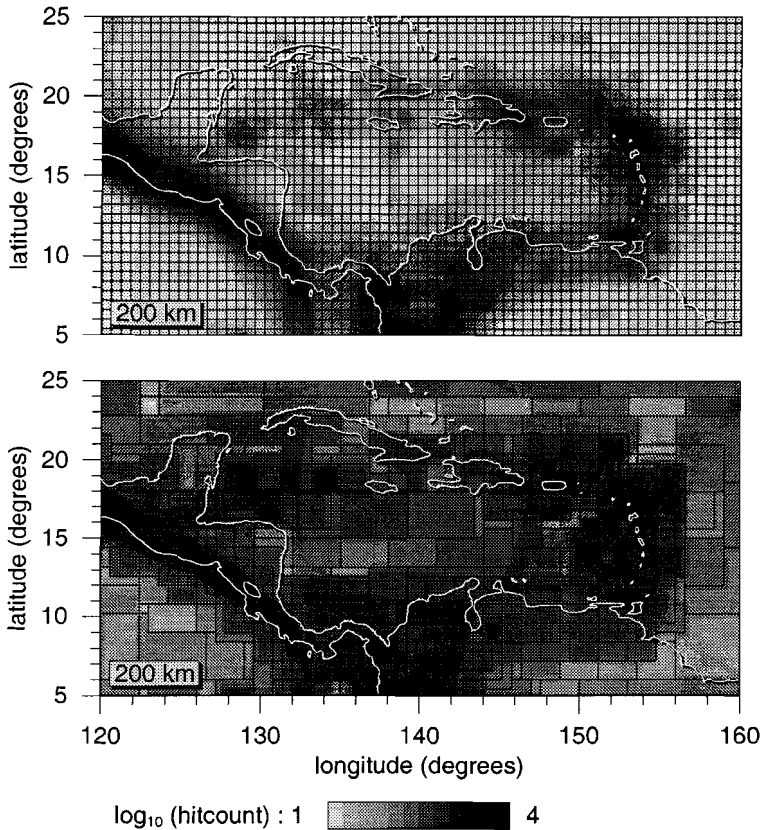


**Figure 4.2:** Standard deviation of residuals in decreasing ray bundles, following the approach of Gudmundsson et al. (1990). The dashed line denotes extrapolation to zero width.

but by extrapolation of the values for the narrowest ray bundles we arrive at a limit value for the standard deviation of approximately 0.3 s (Figure 4.2). According to Gudmundsson et al. (1990) this represents an upper limit for the standard error in the applied data leading to a signal-to-noise ratio of almost 4. This ratio is approximately 2 for the original ISC data, which indicates that the reprocessed data are more accurate.

### 4.3 Parameterization and inversion

The two most popular classes of basis functions for model parameterization in global seismic tomography are spherical harmonic functions (e.g. Su et al., 1994; Masters et al., 1996; Tanimoto, 1990; Woodhouse and Dziewonski, 1984) and local discrete cell functions (e.g. Inoue et al., 1990; Vasco et al., 1995; Zhang and Tanimoto, 1993; Zhou, 1996; Van der Hilst et al., 1997; Grand et al., 1997). A disadvantage of a parameterization based on spherical harmonics is that the relatively low degree truncation of the spherical harmonic expansion of the slowness field may lead to spectral leakage as a result of the uneven sampling of the mantle by seismic rays (Snieder et al., 1991; Trampert and Snieder, 1996). Cell parameterizations for tomography studies aiming at a global solution have up to now constituted either regular grids, based on an equiangular division in longitude and latitude (e.g. Inoue et al., 1990; Van der Hilst et al., 1997) or equal surface area (ESA) grids (e.g. Hager et al., 1985; Vasco et al., 1995). The disadvantage of using regular or ESA cell grids is that when aiming at resolving small-scale structure (i.e., using small cells) large mantle volumes are severely overparameterized as compared to their poor ray sampling. This also holds for the complex cell model of Zhou (1996), which constitutes a stack of regular grids



**Figure 4.3:** Hitcount plots of the Caribbean for a  $0.6^\circ$  regular grid (top plot) and the irregular grid used (bottom plot) at 200 km depth. Notice how the variable cell sizes in the bottom plot lead to much smaller differences in hitcount between cells than observed in the top plot. Cell sizes range in the bottom plot from  $0.6^\circ$  to  $6.0^\circ$ .

with different cell sizes. Furthermore, in the type of parameterization applied by Zhou (1996), the solution in small and large cells is not independent, which may lead to the projection of small-scale structure on an overlapping large cell. Overparameterization will, in general, necessitate the use of rather severe regularization to constrain the solution in sparsely sampled regions.

In order to reduce overparameterization, but at the same time retain the possibility to resolve structure at small scales where allowed by the data, a method has been developed for the parameterization of large linear inverse problems using irregular cells (see section 3.2). The irregular cell model used in the present study has been discussed in Chapter 3. Underlying the grid construction algorithm is the idea to build nonoverlapping larger cells from a regular network of small cells (Abers and Roecker, 1991). The irregular grid is constructed in a completely automated way using hitcount as a constraint on cell volume. In short, we employ a mantle grid of nonoverlapping cells in which cell dimensions vary from

$0.6^\circ \times 0.6^\circ \times 35$  km up to  $9.0^\circ \times 9.0^\circ \times 400$  km such that the cell volumes minimize the variation of cell hitcount (i.e., the number of rays traversing a cell) between adjacent cells. The cells can have irregular shapes; the only restriction is that each cell is constructed from an integer number of basic ( $0.6^\circ$ ) cells. Because of that, hitcount cannot be equalized completely, but differences between cells are reduced considerably. This leads to a model with small cells in densely sampled regions and large cells in regions of low ray density (illustrated in Figure 4.3). Of course, hitcount does not imply resolution, but it can be used to yield a parameterization in which each model parameter is constrained by about the same number of data. In the upper mantle each irregular cell is traversed by at least 500 composite rays (with the exception of extremely poorly sampled regions such as the Pacific); in the lower mantle this limit is raised to 1000 composite rays per cell. This effectively limits the minimum cell size in the lower mantle to 100 km. Regarding the expected width of Fresnel zones for the data used in the lower mantle (Nolet, 1992), it will be impossible to resolve smaller cells. The construction and implementation of the irregular grid goes at the expense of a negligible amount of computation time compared to the usual investment when using regular grids, and it does not complicate the inversion (see section 3.2 and Spakman and Bijwaard, 1998). The main advantage of using sampling dependent cell grids is that they reduce the number of parameters thereby improving the conditioning of the inverse problem while still allowing for the resolution of small detail wherever warranted by the data. A regular cell grid based on the smallest cell sizes we use (laterally  $0.6^\circ$  in the upper mantle,  $1.2^\circ$  up to 1100 km depth, and  $1.8^\circ$  down to the core-mantle boundary (CMB) and varying in thickness from 35 km in the upper mantle to 200 km in the lower mantle) would contain 2.5 million cells. Instead, our irregular cell model consists of only 277,000 cells which is, for instance, comparable to the 275,000 cells in the regular cell model of Van der Hilst et al. (1997). However, in the upper mantle about 25  $0.6^\circ \times 0.6^\circ \times 35$  km cells can be fitted in their  $2^\circ \times 2^\circ \times 100$  km cells, allowing for a much more detailed solution where possible. The actual inversion of the travel-time residuals is based on the approach of Spakman and Nolet (1988), who aim for a joint inversion of cell slowness anomalies, relocation vectors (including origin time shifts), and station statics. We determine an approximate least squares solution of the following matrix equation:

$$\begin{pmatrix} \mathbf{A} \\ \mu \mathbf{D} \end{pmatrix} \mathbf{S}^{-\frac{1}{2}} \mathbf{m}' = \begin{pmatrix} \mathbf{d} \\ \mathbf{0} \end{pmatrix} \quad (4.2)$$

where  $\mathbf{A}$  contains arc lengths of rays in cells and relocation and station coefficients,  $\mu$  is a damping factor that increases with depth from 2000 to 15000,  $\mathbf{D}$  is a matrix of damping coefficients,  $\mathbf{S}$  is a diagonal scaling matrix,  $\mathbf{m}'$  is the scaled model vector:  $\mathbf{m}' = \mathbf{S}^{1/2} \mathbf{m}$ ,  $\mathbf{m}$  is the model vector (consisting of 277000 cell slownesses,  $4 \times 34000$  event cluster relocation parameters, and 5000 station parameters), and  $\mathbf{d}$  is the data vector (consisting of 4.7 million weighted composite delays).

The scaling matrix  $\mathbf{S}$  has been implemented to experiment with the weighting of model parameters in the inversion (Spakman and Nolet, 1988). This is useful, because only an approximate solution can be computed in a reasonable amount of CPU time. We obtain this solution after a limited number of iterations (200) of the least squares solver LSQR (Paige and Saunders, 1982). Within this number of iterations the scaling can weigh some model

parameters down and accentuate others in the solution. With cell volume scaling one can, for example, weigh large cells down, which prevents the long ray arc lengths in these cells from dominating the inversion and allows for a better solution in small cells. It turned out that an inversion with hitcount scaling (to weigh the generally better sampled cells of the lower mantle down) in combination with cell volume scaling gives a good result (for cell  $j$ ,  $S_j = h_j v_j$ , where  $h$  and  $v$  are hitcount and cell volume, respectively), that is, an acceptable resolution for both large and small cells in the upper or lower mantle, with a small emphasis on the convergence of the solution in the smaller cells as these potentially contain the most interesting part of the solution.

Although the use of a sampling dependent cell parameterization may result in a better conditioned inverse problem, hitcount is not a sufficient criterion for optimal design of model space structure (Curtis and Snieder, 1997). Explicit regularization is still needed to constrain underdetermined model parameters and to suppress the influence of data errors (both explicit reading errors and implicit modeling/linearization errors). After extensive testing we have adopted a lateral second-derivative regularization that generally increases with depth (governed by  $\mu$  in (4.2)) and cell volume. The following damping equation has been implemented for irregular cell  $k$ :

$$v_k \left( m_k - \frac{1}{N} \sum_{l=1}^N m_l \right) = 0 \quad (4.3)$$

where  $v_k$  is a factor depending on cell volume between 0.7 and 1.0,  $m_k$  is the model parameter for cell  $k$  and the sum extending over all ( $N$ ) laterally adjacent cells of cell  $k$  in the irregular grid. Equation (4.3) is restricted to lateral regularization only in order to obtain maximum depth resolution.

Implicitly, cell volume relates to the degree of resolution expected. Our synthetic experiments indicate that a regularization independent of cell volume degrades spatial resolution of the smaller detail in the model. The weights  $v_k$  are such that they impose stronger regularization on the solution in larger cells. Together with damping coefficients for the relocation and station parameters (which are chosen such that these obtain reasonable values), the  $v_k$  make up the rows of matrix  $\mathbf{D}$  in (4.2).

This type of parameterization-dependent regularization does not impair the main advantages of using an irregular cell grid. Furthermore, features of the solution that are described below could be identified in test inversion results with various types of regularization and are thus taken to be robust with respect to the influence of damping on the solution.

#### 4.4 Data fit and model resolution

On top of the 16.7% variance reduction obtained from the grouping of data into composite rays, we gain another 57.1% variance reduction in the inversion. Although the total variance reduction amounts to approximately 70% with respect to the original (ISC) data, it still implies that a large part of the residuals remains unexplained, which is very common in seismic tomography. For example, Inoue et al. (1990) obtained 34% variance reduction; Vasco et al. (1995) obtained 35% (for the  $P$  data); Zhou (1996) obtained only 20%; and Van



der Hilst et al. (1997) obtained almost 50% variance reduction. The standard deviation after inversion is 0.9 s, which is considerably larger than the estimated average data error (0.3 s). This is probably due to underparameterization of small-scale structure (model errors), approximations in the inversion procedure, and nonlinear wave front healing effects that are not accounted for. On the basis of results of Gudmundsson et al. (1990), Spakman (1993) estimates the amount of positive delay that is consumed by wave front healing at approximately 0.5 s for a ray at 70° epicentral distance, which is in good agreement with values obtained by Snieder and Sambridge (1992).

Figures 4.1b and 4.1d display the  $P$  and  $pP$  density plots of all residuals after inversion. By comparison with the density plots before inversion (Figures 4.1a and 4.1c) the 57.1% variance reduction is clearly visible as the decreased spread of the data for all epicentral distances. The dashed white lines (which are very similar for  $P$  and  $pP$ ) represent the averages of all residuals per epicentral distance. These (nonzero) averages have decreased in the inversion, implying that small but systematic deviations from the reference model may have been mapped into the inversion result (Van der Hilst and Spakman, 1989).

Owing to the size of the inverse problem it is impossible to calculate a resolution matrix to formally assess the reliability of the results. In large-scale tomography studies, model resolution is therefore usually estimated using sensitivity analysis (Spakman and Nolet, 1988; Humphreys and Clayton, 1988). In such an analysis, travel-time residuals caused by synthetic heterogeneity of an artificial input model are calculated for the employed seismic rays. After the addition of normally distributed noise the inversion algorithm is applied to the synthetic residuals in an attempt to recover the synthetic structures. This method, however, overestimates the resolving power of the data. It neglects nonlinear errors due to the bending of seismic rays by velocity anomalies (which probably is significant for 0.6° cells), reference model biases, and systematic earthquake mislocations (Spakman, 1991; Van der Hilst et al., 1993). Furthermore, it has been shown by Lévêque et al. (1993) that small-scale resolution, implied by a synthetic test with small-scale anomalies, does not guarantee resolution on larger scales. Finally, such tests do not measure model uncertainty: the normally distributed noise is a poor substitute for (systematic) data inconsistencies because of model parameterization and linearization or station errors that may account for biases in the solution (Spakman, 1991; Van der Hilst et al., 1993). Model uncertainty can be addressed with so-called permuted data tests (Spakman and Nolet, 1988; Spakman, 1991), which test the hypothesis that the actual data are pure noise. In a permuted data test the elements of the delay time vector are randomly permuted, while keeping the order of the matrix equation fixed. With this action any correlation between ray paths and delay times is destroyed, and the permuted data can be considered pure noise with exactly the same bulk statistics (average, variance, and frequency distribution) as the actual data. We performed five of these tests to construct variance maps much like those previously obtained from the inversion of normally distributed noise by Inoue et al. (1990). These maps (not shown) indicate that the estimated errors, which should be regarded as upper limits for model uncertainty, are throughout the model much smaller than the anomaly amplitudes. The average error is 0.21% (of the ref. velocity) with a standard deviation of 0.16%.

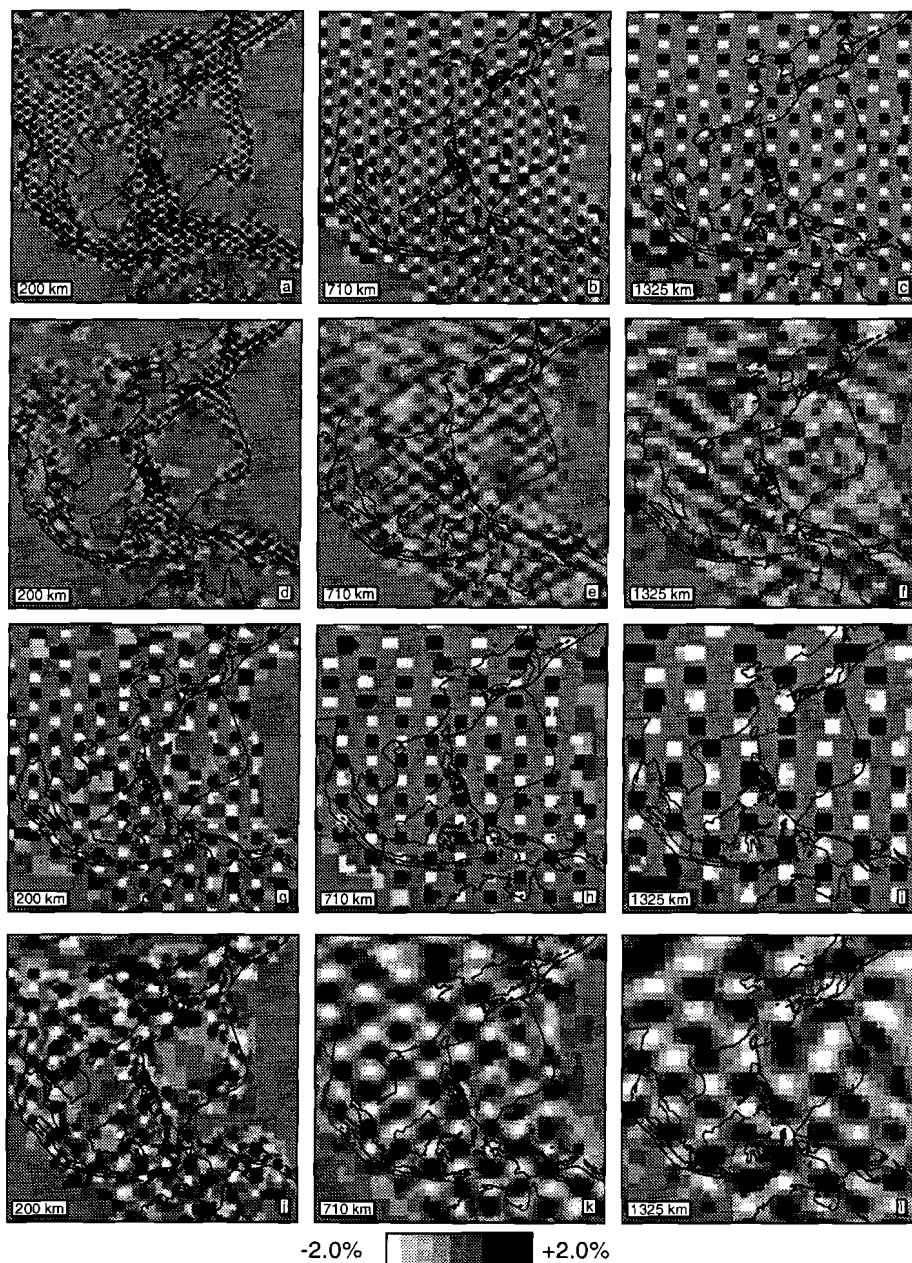
A very popular sensitivity experiment to estimate model resolution is the so-called checkerboard test (e.g. Inoue et al., 1990; Fukao et al., 1992; Vasco et al., 1995; Zhou, 1996), in which a (3-D) checkerboard pattern of high- and low-velocity anomalies is used as a synthetic input model. However, in regions where the pattern of the checkerboard is

reproduced but not its amplitude, this test does not permit one to infer that the along-ray-path smearing of anomalies causes the amplitudes to be underestimated. We therefore prefer impulse response or 'spike' tests (Spakman and Nolet, 1988), in which anomalous cells of alternately positive and negative sign are spaced a few cells apart in every direction, thus enabling us to investigate the smearing of the spiked anomalies by the inversion. The spike models we apply are generated on an ESA grid and then projected onto the irregular cells. Wherever spikes are smaller than the irregular cells, the irregular cells reflect the cell-volume-averaged value of all spikes they contain. Wherever spikes are larger than the irregular cells, each irregular cell simply obtains the spike value. Examples of such irregular spike models are shown in Figure 4.4.

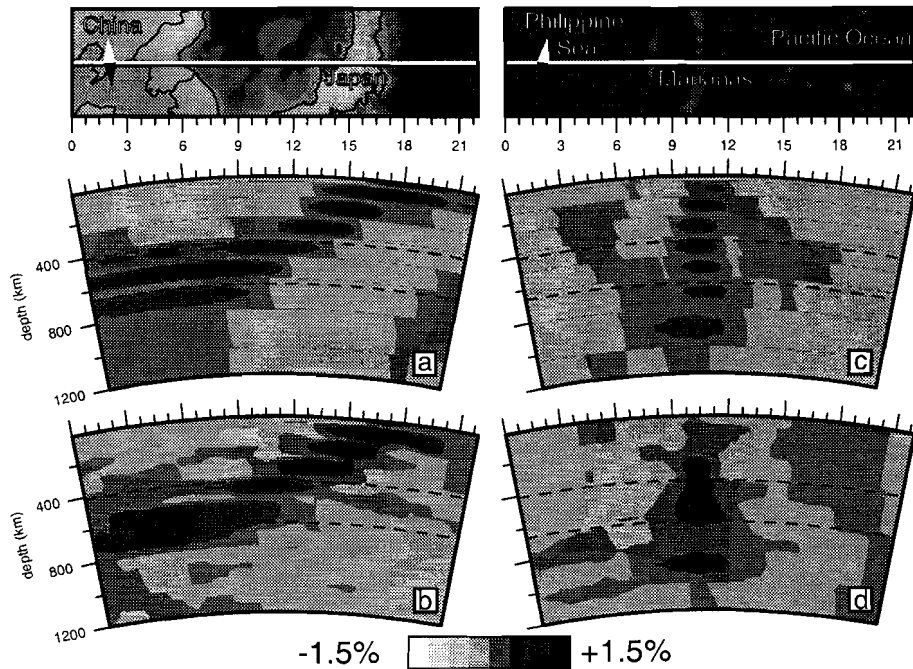
For the following implemented cell sizes:  $0.6^\circ$ ,  $1.2^\circ$ ,  $1.8^\circ$ ,  $2.4^\circ$ ,  $3.0^\circ$ ,  $4.2^\circ$ , and  $6.0^\circ$ , we perform these spike tests using normally distributed noise with a standard deviation of 0.5 s added to the synthetic data, aiming to resolve spikes with amplitudes of 5%. In general, these tests show reasonably good recovery (i.e., the pattern is clearly visible, but smearing may reduce the amplitudes) on all scales within small cells in the top 300 km, that is, in all subduction zones plus most of Eurasia (with the exception of northern Siberia) and the United States (an example for southeast Asia is shown in Figure 4.4). Only in a few areas, we obtain smallest scale (65 km) recovery which decreases slowly below 300 km to about 100 km at the 660 km discontinuity, but the total recovered area (on a scale of 250 km) increases and includes parts of Africa, Australia, and the northern Atlantic. In the upper part of the lower mantle, the area recovered on a scale of 150-300 km becomes larger and includes the boundaries of the oceans, entire Australia and northern Siberia. In the middle of the lower mantle, nearly the entire northern hemisphere and Australia, and southeast Asia are resolved on a scale of 300 km. In the lowermost mantle, spike resolution decreases to a scale of about 350 km for the best sampled regions in Eurasia and parts of the Indian Ocean. In the Pacific, we observe at this depth severe streaking effects due to an anisotropic distribution of ray directions.

Apart from spike tests, we perform a series of tests more dedicated to the shapes and amplitudes imaged in the inversion of the actual data. These tests are much like the frequently used slab tests (e.g. Spakman et al., 1989; Fukao et al., 1992; Van der Hilst, 1995). However, slab tests are not aimed at investigating whether the internal continuity of the inferred slab is real or whether it may be due to smearing along rays. We therefore follow a slightly different approach. For our tests, synthetic models are constructed from the real solution. This is done in the following way: we remove all positive (or negative) anomalies, smooth the model, and add a background anomaly value to obtain a zero mean. We then remove all anomalies at specific depth intervals which leads to a layer-cake type of synthetic model (see, e.g. Figures 4.5a and 4.5c). We compute synthetic data for this model and finally add normally distributed noise.

We can make two inferences on the basis of the inversion results obtained with this kind of synthetic input models. First, the observed decrease in amplitude with depth in the real solution (see section 4.5) is not amplified in the synthetics and may thus be genuine, which indicates that lower mantle wavespeed heterogeneity is of smaller amplitude than upper mantle heterogeneity. This is consistent with the findings of Vasco et al. (1995), Gudmundsson et al. (1990), and Inoue et al. (1990), who find upper mantle heterogeneity on the order of 2-3% and lower mantle heterogeneity on the order of 0.5-1%. These values are probably both underestimates of the real amplitudes, because of the applied damping and



**Figure 4.4:** Synthetic spike models and inversion results for southeast Asia: (a)  $0.6^\circ$  spike model at 200 km depth, (b)  $1.2^\circ$  spike model at 710 km depth, and (c)  $1.8^\circ$  spike model at 1325 km depth. (d), (e), and (f) plots show the corresponding inversion results. (g), (h), and (i) plots are for spike models of  $1.8^\circ$ ,  $2.4^\circ$ ,  $3.0^\circ$  and (j), (k), and (l) plots for their inversion results, respectively. In general, the first six plots show input and output at the lower resolution boundary, the last six depict spikes of sizes that are well-resolved.



**Figure 4.5:** (a) and (c) input and (b) and (d) output of a synthetic test to investigate vertical resolution in subduction zones. The left plots ((40°N, 120°E) - (34.8°N, 147°E)) depict the Japan subduction zone; the right plots ((16°N, 135°E) - (18.5°N, 157.9°E)) the Marianas. The smearing in the upper mantle part of the Marianas subduction zone is caused by rays preferentially traveling obliquely downward (in the Japan subduction zone this is compensated by many upward rays).

possibly wave front healing. This also holds for our solution, but the synthetic tests indicate that amplitudes are estimated fairly well.

Second, the layer-cake model allows us to investigate the vertical smearing of heterogeneity in realistic structures such as slabs and plumes. The results indicate vertical smearing (over 35-65 km thick layers) in regions of the upper mantle where the predominant ray direction is near vertical (see, e.g. Figures 4.5c and 4.5d) but indicate better recovery in regions with more diverse ray directions (see, e.g. Figures 4.5a and 4.5b). In the lower mantle the layer-cake models (with 100-200 km thick layers) seem to be well-resolved in regions with adequate ray sampling. In general, we shall only show and discuss model features for which the synthetic spike tests and tests dedicated to specific structures indicate that they are resolved on the scale of the discussed structure unless indicated otherwise.

## 4.5 Results

From the matrix inversion we obtain the model vector, which consists of station corrections, cluster relocation vectors, origin time errors, and cell slowness anomalies. The station

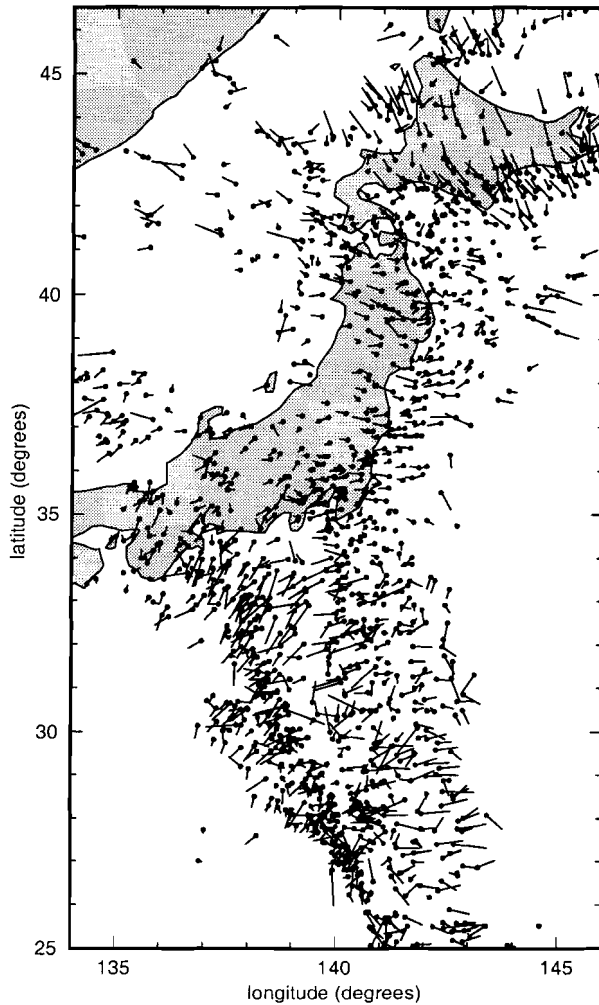
corrections correlate well with average station residuals before inversion. This implies that they mainly serve to remove average station residuals (i.e., the average of all composite ray residuals from phases recorded at the station). The average station residuals are regionally systematic (not shown) and are thus not due to random station errors. For stations with substantial azimuthal coverage they probably represent real, shallow structure directly beneath the station.

The event cluster relocation vectors are of the same order of magnitude (10 km) as the formal hypocentral errors determined by Engdahl et al. (1998). In general, they neither increase nor decrease the amount of clustering of hypocenters, though the original locations were determined with the inclusion of *S*, *sP*, *PKiKP*, and *PKP<sub>df</sub>*, which are not included in the inversion. We compared the original and the new locations of a subset of 111 events for which the real locations are known from independent data (e.g. explosions). The average length of the 3-D mislocation vector (9.9 km) for these events is, compared to their original mislocation in ak135, reduced in the inversion by 18%. Furthermore, for the other event clusters the direction of relocation is regionally systematic. The systematic trends are the combined result of mislocation due to 3-D heterogeneity and station network geometry for a particular cluster. An example of the pattern of horizontal relocations is shown in Figure 4.6: in the Japan and Izu-Bonin subduction zones, we observe a trend in the relocations from mostly relocation to the northwest beneath Hokkaido, to the west beneath Honshu, and to the southwest in the Izu-Bonin subduction zone. In general, the origin time shifts show a strong correlation with the depth relocation (not shown), which illustrates the well-known trade-off between hypocenter depth and origin time.

The imaged velocity heterogeneity is depicted in the form of layer solutions (smoothed over distances of  $0.6^\circ$ ) for a few selected depths in Figure A1 (see appendix). Above the transition zone (0-410 km depth) the anomalies are mainly confined to the continents, where most seismic stations reside (with the exception of Antarctica), and the subduction zones, where most events originate. The oceanic parts and the mid-ocean ridges (with the exception of the North Atlantic Ridge) are scarcely illuminated by seismic rays, and even large cells (where implemented) have a relatively low hitcount and are therefore not resolved. The applied regularization pushes the solution toward the reference model in these areas. Below the transition zone (660 km - CMB), the suboceanic regions are better sampled, and the tomographic image becomes more complete, although large parts of the southern hemisphere remain poorly sampled.

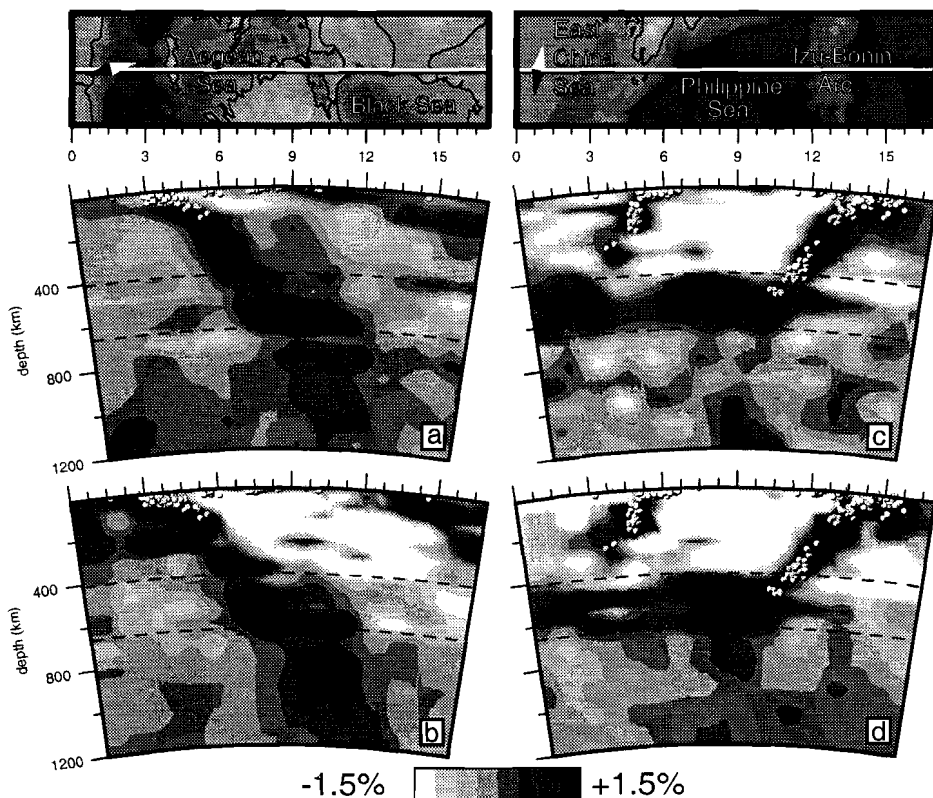
Above the transition zone (Figure A1a) we observe underneath large parts of the continents high-velocity anomalies with amplitudes of 3-4% and roots extending to 300-400 km (not shown), although depth resolution is generally insufficient to determine the exact depth extent. These zones correspond with continental shields in Fennoscandia, Baltica, Arabia, southern Africa, Australia, Canada, and Greenland, which have previously been identified in several other global travel-time tomography models (e.g. Van der Hilst et al., 1997; Zhou, 1996; Vasco et al., 1995), but which are usually better resolved in lateral extent by surface wave studies (e.g. Zhang and Tanimoto, 1993). A prominent shield that is not imaged in our model lies beneath the Amazon where ray sampling is very poor.

The best sampled mid-ocean ridge, in the northernmost Atlantic, is characterized by very low velocities (consistent with, for example, Vasco et al. (1995); Zhou (1996); Inoue et al. (1990)) throughout the upper mantle. We also find negative anomalies beneath the tectonically active regions of the Dead Sea transform, the Red Sea spreading zone, and the



**Figure 4.6:** Horizontal relocations for subcrustal events beneath Japan and in the Izu-Bonin subduction zone. Circles denote the locations before inversion; lines point in the direction of relocation and have exaggerated lengths ( $3\times$ ). Notice the systematic patterns with relocations mainly to the northwest beneath Hokkaido, to the west beneath Honshu, and to the southwest in the Izu-Bonin subduction zone.

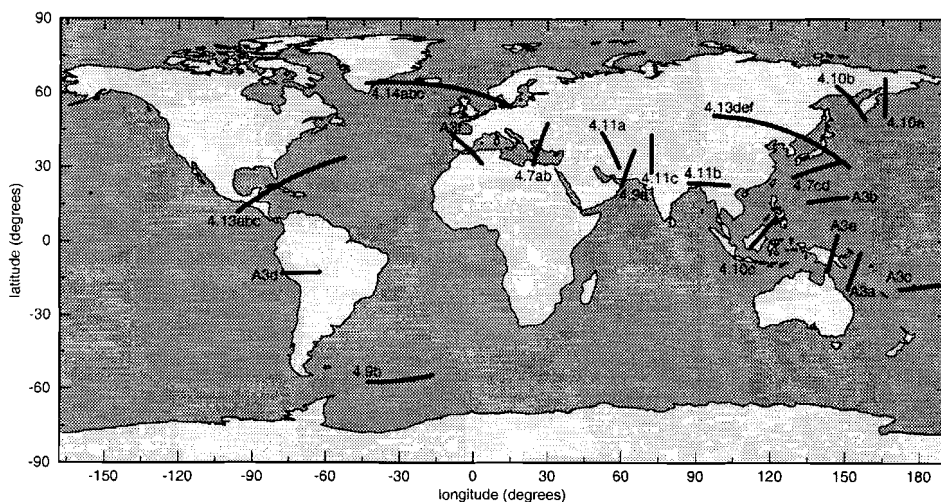
east Africa rift. Other prominent, shallow low velocities can be associated with orogenic belts (e.g. in Europe, in Asia Minor, and in North America), and back arc basins associated with major subduction zones (e.g. the Aegean, the Java and Banda Seas, parts of the South China and East China Seas, the Japanese and Philippine Seas, the Aleutian Basin, the South Fiji Basin, and parts of the Caribbean). Many of these features have been observed in several regional studies (Spakman et al., 1993; Mohan and Rai, 1995; Widiyantoro and Van der Hilst, 1996; Fukao et al., 1992; Engdahl and Gubbins, 1987; Van der Hilst, 1995; Van der



**Figure 4.7:** Comparison between the presented global solution and regional models in two cross-sections: (a) section across the Aegean ( $31.2^{\circ}\text{N}$ ,  $24^{\circ}\text{E}$ ) - ( $47.6^{\circ}\text{N}$ ,  $30^{\circ}\text{E}$ ) from model EUR89B (Spakman et al., 1993), (b) same section as Figure 4.7a from the global model, (c) section across the Izu-Bonin arc ( $26^{\circ}\text{N}$ ,  $129^{\circ}\text{E}$ ) - ( $32^{\circ}\text{N}$ ,  $147.2^{\circ}\text{E}$ ) from model NWP90 (Van der Hilst et al., 1991), and (d) same section as Figure 4.7c from the global model. Notice that the regional models contain more small-scale, low-amplitude anomalies in the lower mantle than the global solution. Note that the regional results were obtained against somewhat different reference models than the global result.

Hilst and Spakman, 1989) and in previous global studies, though usually with less detail (e.g. Van der Hilst et al., 1997; Zhou, 1996; Vasco et al., 1995; Su et al., 1994).

The most striking features above the transition zone are the narrow, elongated high-velocity anomalies corresponding with the major subduction zones. In previous global studies many of these slabs could not be identified in a coarse cell model, because they were ‘sandwiched’ between low-velocity anomalies associated with back arc basins and possible contributions from frictional heating. As an example, we compare our solution in southeast Asia in Figure A2 with the model of Van der Hilst et al. (1997). Anomalies associated with subduction beneath the Sunda and Banda arcs, the Philippines, the Marianas, and New Guinea are either faintly visible or totally absent in Figure A2a (Van der Hilst et al., 1997) but are clearly



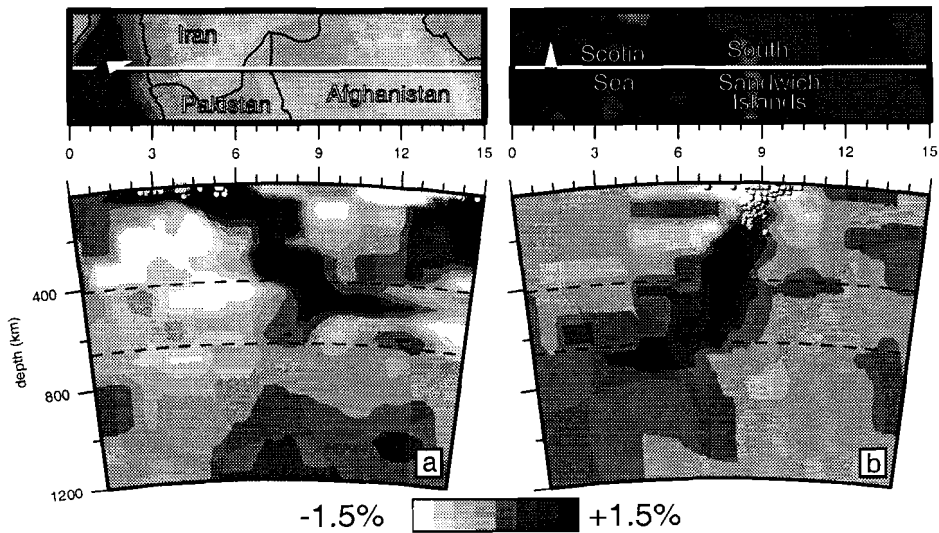
**Figure 4.8:** Overview of the locations of cross-sections shown in Figures 4.7, 4.9-4.11, 4.13, 4.14, and A3. All cross-sections have a small map with a compass needle (white pointing north) for orientation.

present in Figure A2b (our solution).

In general, we find clear, approximately 150 km thick plate-like anomalies that confirm and that can be compared with regional studies of the zones of large-scale subduction of Tonga-Kermadec (Van der Hilst, 1995), Indonesia (Fukao et al., 1992; Puspito et al., 1993; Widiyantoro and Van der Hilst, 1996), Japan (Zhou, 1988; Spakman et al., 1989; Zhou and Clayton, 1990; Van der Hilst et al., 1991; Fukao et al., 1992; Van der Hilst et al., 1993), the Aleutians (Engdahl and Gubbins, 1987), Cascadia (VanDecar, 1991), the Caribbean (Van der Hilst and Spakman, 1989; Van der Hilst, 1990), South America (Engdahl et al., 1995), and the Mediterranean (Spakman, 1991; Blanco and Spakman, 1993; Spakman et al., 1993). Two examples of a comparison between a regional study and our model are shown in Figure 4.7 (locations of all cross-sections discussed here can be found in Figure 4.8). Notice that the lower parts of the sections from the regional models show small-scale anomalies that are not present in the global solution. This may indicate the mapping of structure from outside the study region into the model (Masson and Trampert, 1997). We can add to the total sum of these regional (subduction zone) models some less well-studied subduction zones like those beneath the Solomon Islands (connecting Vanuatu subduction to New Guinea), the Scotia Sea, Pakistan (Makran subduction zone), and Algeria where we observe slab signatures as well (see Figure 4.9). We note, however, that although the geometry of these anomalies seems to be resolved, their internal structure is not. Vertical resolution is limited in these areas where most rays travel obliquely downward.

Apart from these major subduction zones, we detect in a few areas small isolated anomalies that might be associated with localized subduction. In the western part of the Aleutians, we observe in Figure 4.10a north of the slab below the trench a low-amplitude, high-velocity anomaly below the Koryakskoye mountains of Russia, coinciding with a seismically active region. Similar features are observed in Figures 4.10b and 4.10c northwest of Kamchatka

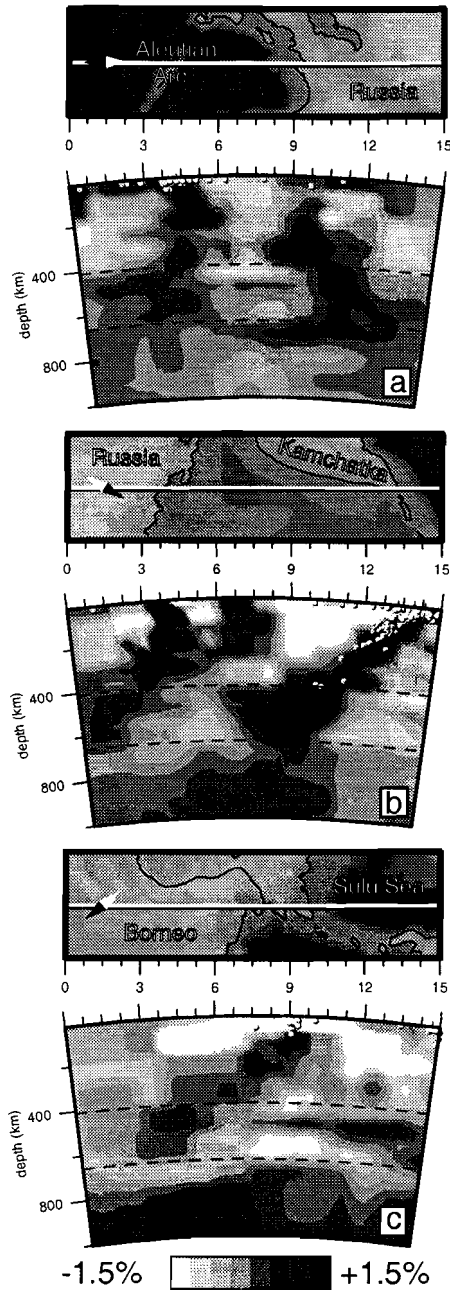




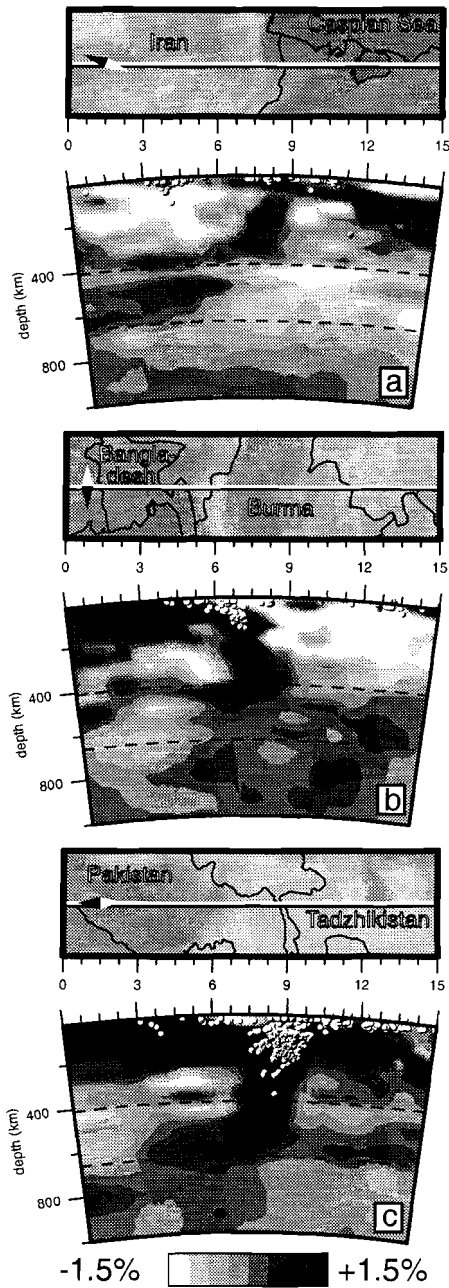
**Figure 4.9:** Cross-sections through (a) the Makran subduction zone (Pakistan) ((22.5°N, 60°E) - (36.8°N, 65.1°E)) and (b) the Scotia Sea subduction zone (South Sandwich Islands) ((57.5°S, 43°W) - (54.5°S, 16.5°W)). The geometry of these anomalies is resolved, but their internal structure is not.

and in northeast Borneo. Sensitivity tests indicate that these anomalies are real, but their exact geometry is not resolved. In the continental collision zones in India and Iran we observe high velocities indicating a thickened lithosphere (up to 300 km) and again very localized subduction in the Pamir region in northern India (e.g. Roecker, 1993; Mohan and Rai, 1995), in northern Burma, and possibly in the Elburz in north Iran (displayed in Figure 4.11). The anomalies associated with localized subduction are reasonably well-resolved in lateral extent but suffer from vertical smearing between 350 and 660 km depth.

As noted by several authors (e.g. Su et al., 1994; Masters et al., 1996), the overall picture changes considerably in the transition zone (410-660 km depth). The signature of the continental shields has disappeared, and we observe a very different pattern of high and low velocities (see Figure A1b). The main positive anomalies now underlie Europe and the large subduction zones. Most of the negative back-arc anomalies have been replaced by broad positive anomalies caused by the apparent flattening of subduction in the transition zone (Van der Hilst et al., 1991; Fukao et al., 1992). Sensitivity tests indicate that the broad, positive anomalies are reasonably well-resolved in lateral extent in the major subduction zones of the Mediterranean, of the Pamir, of the Sunda and Banda arcs, below the Philippines, below the Marianas, and below Taiwan, in the Japanese and Ochotsk Seas, below the eastern part of the Aleutians, in the Caribbean, below the Andes, in the Scotia Sea, below the Solomon Islands, and to the west of Tonga. This indicates that the observed blurring is not due to lack of resolution. However, at this depth the model may be biased by a few tenths of % due to the aforementioned nonzero data averages. Furthermore, misidentified phases resulting from the overlapping of travel-time curve triplications related to the seismic discontinuities at 410 and 660 km may also affect the solution.



**Figure 4.10:** Cross-sections through (a) the Aleutian arc ((51°N, 166°E) - (66°N, 166°E)), (b) Kamchatka ((63°N, 146.5°E) - (49.4°N, 158°E)), and (c) Borneo ((2°S, 112.5°E) - (9.5°N, 122.2°E)). In all three sections we observe peculiar high-velocity anomalies (apart from the obvious slab related anomalies) that might be related to subduction. The presence of these anomalies is resolved, but their exact extent is not.

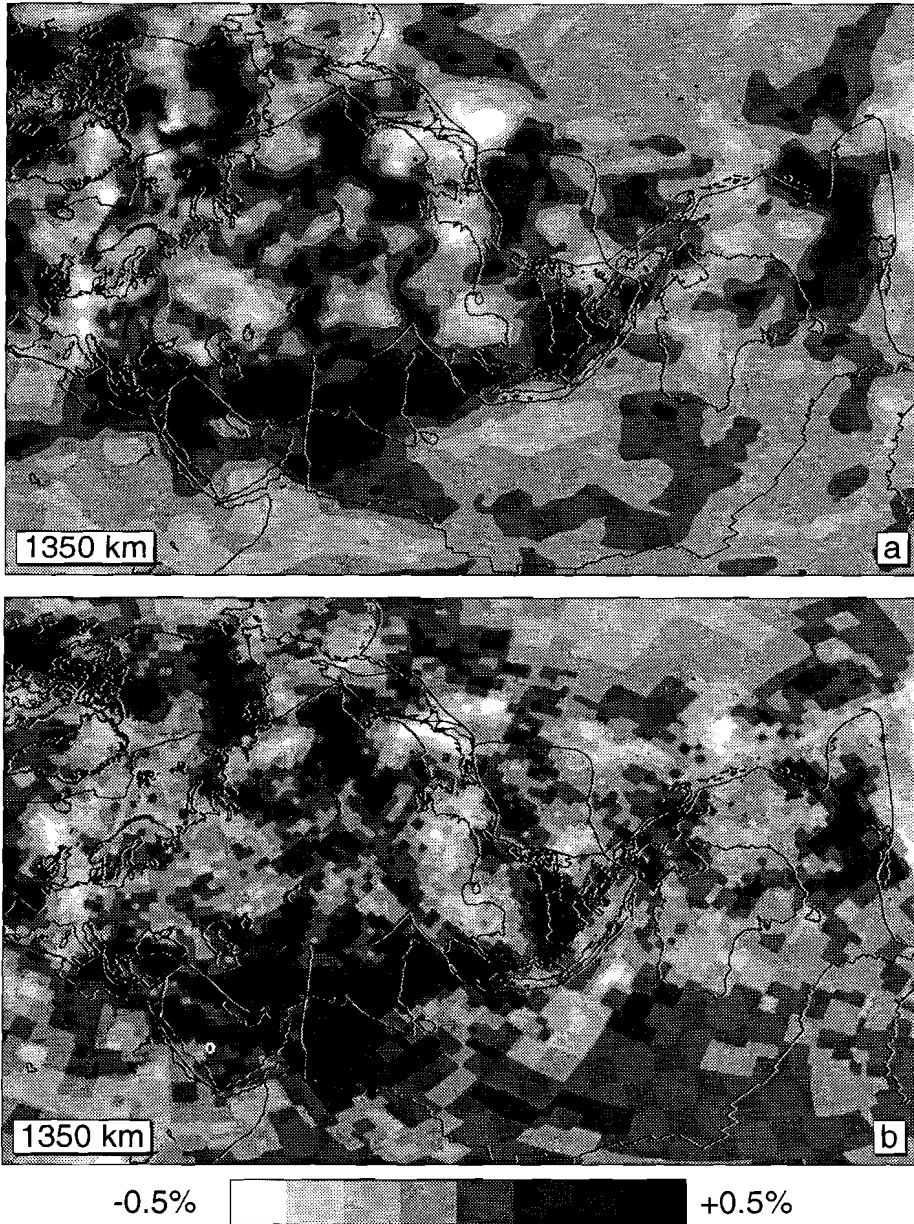


**Figure 4.11:** Cross-sections through (a) the Elburz Mountains (northern Iran) ((30°N, 59°E) - (43.9°N, 51.9°E)), (b) Burma and Bangladesh ((23.5°N, 87°E) - (22.7°N, 103.3°E)), and (c) the Pamir (Himalayas) ((28°N, 72°E) - (43°N, 72°E)) displaying continental subduction. Lateral resolution is reasonable; smearing in the vertical direction occurs between 350 and 660 km.

The large low-velocity anomalies can partly be spatially correlated to rifts (east Africa) and hotspots (Iceland, Hawaii, and the Deccan traps), but it may also be possible that the reference velocities are biased to higher values by the preferential (subduction zone) sampling of the seismic rays that were used in constructing the reference model (although care was taken by Kennett et al. (1995) to avoid this). This could be the cause of some of the low velocities we observe around subduction zones, although frictional heating could also be important. Synthetic tests indicate that the low velocities surrounding the slabs are not artefacts caused by, for instance, a possible instability in the inversion.

In general, the image is much more diffuse than in the upper 400 km. This does not seem to be caused by lack of lateral resolution. Although resolution for the  $0.6^\circ$  cells is limited, we can still by ray geometry resolve  $1.2^\circ$  structures in large areas, but this is at the limit of what physically can be resolved regarding the Fresnel zone width for teleseismic rays at this depth.

We observe a general continuity of structures across the 660 km discontinuity. Especially, the prominent subduction-related anomalies can easily be traced into the lower mantle, where they continue to widen. This widening may have multiple causes: the gradual decrease in resolving power with depth, the broadening of Fresnel zones with ray length (Nolet, 1992), and/or a possible slowing down of the slab at the phase boundary with its associated viscosity jump. Around 800 km depth (Figure A1c) many positive anomalies can still be clearly associated with subducted slabs in the upper mantle, although most have lost their plate-like geometry. The subducted lithosphere around Indonesia converges into one huge high-velocity anomaly (lateral resolution is on the order of  $2.4^\circ$ ), the anomalies associated with Vanuatu and Tonga-Kermadec subduction seem to merge at this depth (in accord with Van der Hilst (1995)), and the high-velocity anomalies beneath South America now continue across the Caribbean well into North America as a remnant of the subducted Farallon plate (in accord with Grand (1994)). A similar feature can be observed below Eurasia, where the Aegean anomaly continues eastward across Turkey and high velocities below India connect with those beneath Indonesia. Around 1300 km depth (Figure A1d) the high velocities stretch from the eastern Mediterranean to Borneo to form the Tethys anomaly (in accord with Van der Hilst et al. (1997) and Grand et al. (1997)). A comparison between the result obtained by Van der Hilst et al. (1997) and our solution for this region is shown in Figure 4.12. The results are very similar, although our solution has somewhat higher amplitudes and shows finer and resolved detail of this large structure. At this depth, the model amplitudes are substantially lower ( $\sim 0.5\%$ ) than in the upper mantle ( $\sim 2\%$ ), but the anomalies are resolved on a scale of 200-300 km. The Tethys and Farallon anomalies dominate the image, but smaller positive anomalies west of Japan and New Zealand also connect to upper mantle subduction of the Pacific plate. Many of the low velocities can, at this depth, still be spatially correlated with hotspots at the surface: Iceland, the Canary Islands and Cape Verde, east Africa, Society Islands, and Yellowstone. In general, our lower mantle results confirm observations by Van der Hilst et al. (1997) and Grand et al. (1997) but also add some more detail. The large-scale anomalies sometimes have an internal structure that has not been imaged before. The Tethys anomaly, for instance, consists of at least two segmented parallel and linear high-velocity bands: one stretching from Turkey to India and on to Borneo and the other stretching from Greece across Arabia into the Indian Ocean. These may reflect the complex history of subduction of the Tethys ocean (see also section 5.3). An explanation could be that these parallel zones result from the subduction of two major parts of the



**Figure 4.12:** Comparison of the solution for the Tethys region at 1350 km depth in an oblique geographical projection (centered along a great circle starting at  $(45^{\circ}\text{N}, 15^{\circ}\text{W})$  with azimuth  $50^{\circ}$  and length  $190^{\circ}$ ) between (a) the model of Van der Hilst et al. (1997) and (b) our model (unsmoothed).

Tethys ocean, separated by a ridge system of unknown geometry. The more southerly located anomaly would then be the image of the former northern segment of the Tethys, which was subducted first. The northern anomaly would result from the subduction of the southern part of the Tethys, which connected to northern India and which has overridden the already subducted northern segment. The distinction between northern and southern neo-Tethys basins follows from the fact that the Indian continent was attached to the southern Tethys and that the last active subduction of Tethyan age lithosphere presently occurs below the Aegean and below the central Sunda arc. Other intriguing features are the low-velocity anomalies connecting the Cape Verde and Canary Islands hotspots across northwest Africa with the Hoggar hotspot (Dautria, 1988) and the low velocities beneath western Europe, where remnants of recent volcanic activity are present in the Massif Central (Granet et al., 1995) and the Eifel, connecting with low velocities below the Iceland hotspot. Up to 1800 km depth the picture hardly changes, and the main features can easily be traced, but the amplitudes decrease and around 1900 km the image consists of many small and low-amplitude anomalies (Figure A1e). Van der Hilst et al. (1997) raise the question of whether this part of the lower mantle might be a transitional interval in which downwellings transform from a planar geometry to a more cylindrical one. There are, however, only two seemingly planar anomalies identified in the lower mantle immediately above this interval: The Tethys and Farallon anomalies. The absence of the large sheet-like Tethys and Farallon anomalies in the lower part of the mantle may just reflect that the bulk of the lithosphere material has not reached this depth yet. If one would remove these prominent anomalies from the model, the anomaly pattern from 1300 to 1800 km would show similar scale and amplitude anomalies as deeper in the lower mantle. Moreover, we do not observe a change towards a more cylindrical shape for all high-velocity anomalies that continue across the 1800-2300 km interval. For instance, the high-velocity anomaly beneath Siberia (see Figure A1e) that is imaged from about 1500 km depth to the CMB has a distinct, elongated Z shape that persists all the way to the D'' layer, where it joins a large high-velocity anomaly. This Z-shaped anomaly may have resulted from older subduction than that of the Tethys and Farallon plates and hence has had more time to reach the lower part of the mantle (see section 5.2). Hence in our model we find no clear evidence for a transitional interval between 1800 and 2300 km.

From about 2200 km to the CMB the pattern grows toward the often observed degree 2 pattern (e.g. Su et al., 1994) with low velocities beneath Africa and most of the Pacific and high velocities below Asia and the Americas (Figure A1f). These huge anomalies with somewhat higher amplitudes than we observe in the middle of the lower mantle do exhibit some internal structure, but resolution is rather limited at this depth, and severe streaking effects are visible in the northwest Pacific because of preferential ray directions. The elongated linear high-velocity anomaly below the Americas resembles a subduction-related anomaly, but the anomaly is not connected with the Farallon anomaly. The observed high velocities beneath Asia and the northwest Pacific may locally be connected with the subducting slab beneath Japan (in accord with Van der Hilst et al. (1997)), which corroborates the slab graveyard origin of these anomalies (see Wyssession (1996) for a review of this hypothesis).

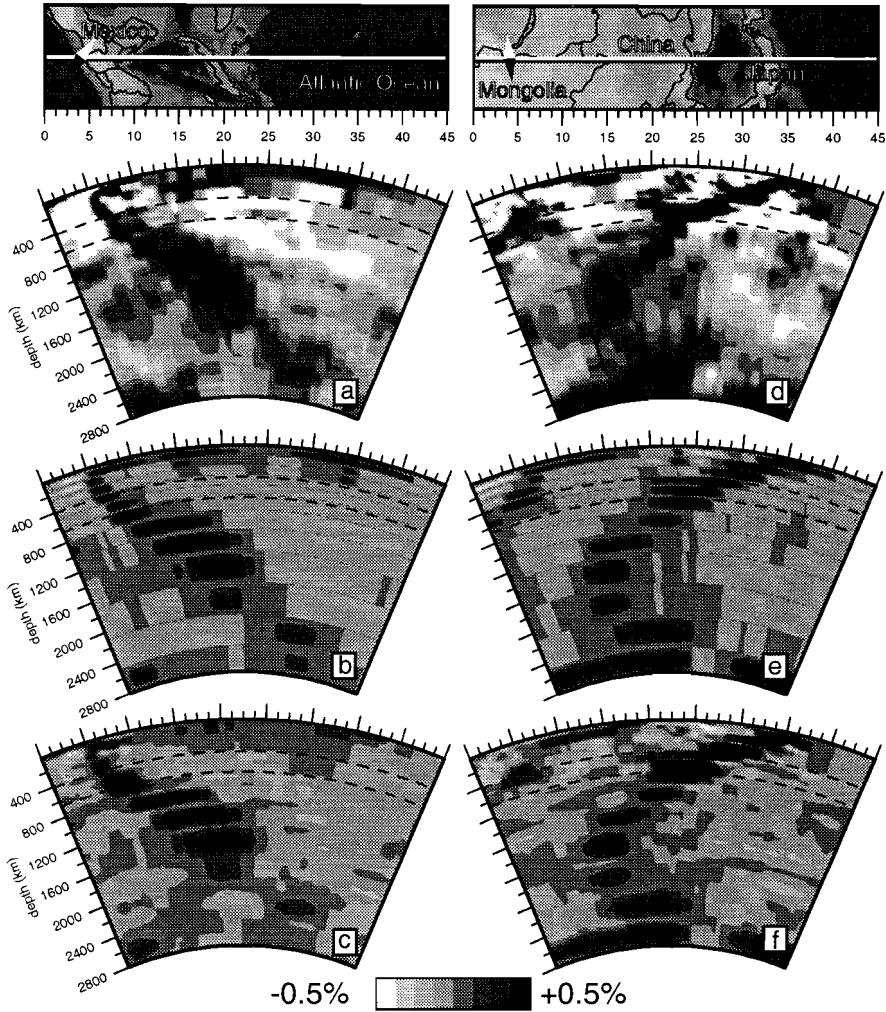
We refrain from interpreting the imaged lower mantle heterogeneity in terms of temperature, pressure, or compositional anomalies. Although it is evident that *P*-wave velocity depends on these three parameters, we feel there is at present insufficient information to sensibly

quantify the relative contributions of temperature, pressure, and compositional anomalies to seismic wave speed.

As mentioned before, a global high-resolution model enables us to compare small-scale tectonic features and to put them in a global perspective. The most detailed parts of our model are the densely sampled subduction zones and different types of slabs are imaged in these regions (see Figure A3), but the continuity and extent of these slabs are not always sufficiently resolved. In the slab cross-sections shown in Figures 4.9-4.11 and in Figure A3, we do not, in general, show well-known major subduction zones, but instead we focus on smaller, less well-known slabs that often have not been imaged on a regional scale before. Where subduction has continued over longer time spans, we observe deep slabs, flattening in the transition zone underneath the western Mediterranean (Figure A1b), the Banda arc, northern Philippines, the Solomon Islands (Figure A3a), and the Izu-Bonin trench. Resolution may be questionable beneath the Solomon Islands, but these flattenings can be resolved beneath the western Mediterranean and also below the Banda arc and in the Izu-Bonin subduction zone where they were previously identified by Puspito et al. (1993), Widiyantoro and Van der Hilst (1996), and Van der Hilst et al. (1991). Subduction-related anomalies penetrate the 660 km discontinuity underneath Java (Fukao et al., 1992; Puspito et al., 1993; Widiyantoro and Van der Hilst, 1996) and the Marianas (Figure A3b) (Van der Hilst et al., 1991), and they penetrate the 660 km discontinuity, after initial flattening in the transition zone, underneath the Kuriles (Fukao et al., 1992) and Tonga (Figure A3c) (Van der Hilst, 1995). Vertical resolution is mediocre in the upper mantle beneath Java, the Marianas, and Tonga, but we seem to be able to identify their lower mantle continuations. In several subduction zones other forms of slab deformation occur. Below Venezuela and below the Peru-Bolivia region (Figure A3d) we observe a kink in the high-velocity anomaly (Engdahl et al., 1995) that seems to be resolved. Below New Guinea (Figure A3e) we may be imaging the splitting of a slab at the 660 km discontinuity, but the northward extending, upper mantle high-velocity part may also be the result of smearing along rays. We find anomalies suggesting possible detachment of slabs below Ecuador, southern Mexico, Italy, and Albania (consistent with Spakman et al. (1993)) and south Spain (Figure A3f) (consistent with Blanco and Spakman (1993)). Resolution may not be sufficient beneath Mexico and Ecuador, but it probably is sufficient beneath parts of southern Europe. We do not find any indication of detachment underneath Sumatra (Widiyantoro and Van der Hilst, 1996, 1997), where we do seem to have enough vertical resolution to detect such a feature (see also section 5.4). We do also not detect detachment beneath Greece, but estimates of the vertical resolution in this region are not as good as obtained by Spakman et al. (1993), who also used a large amount of regional data.

The dynamics behind some of the different types of slab behavior have been investigated through numerical modeling and laboratory experiments. In these studies the observed flattening and/or penetration at the 660 km discontinuity has been suggested to be caused by different rates of trench migration (Griffiths et al., 1995; Christensen, 1996; Olbertz et al., 1997). Numerical modeling has indicated that detachment of slabs may occur shortly after the cessation of subduction (Olbertz, 1997; Wong A Ton and Wortel, 1997).

As previously stated, some of the slab-related anomalies connect in the global perspective to anomalies at 1300 km depth (under southeastern Europe, Indonesia, Tonga, Central America, and Japan). Figure 4.13 demonstrates the continuation of the high-velocity anomalies beneath Central America and Japan all the way down to the CMB. This is



**Figure 4.13:** Cross-sections across (a)-(c) Central America ((12°N, 96°W) - (33.6°N, 52.6°W)) and (d)-(f) Japan ((51°N, 97°E) - (30.2°N, 151.5°E)). Figures 4.13a and 4.13d show the solutions, Figures 4.13b and 4.13e show synthetic input models, and Figures 4.13c and 4.13f show the synthetic solutions. Notice that vertical resolution ranges from 50 to 200 km in these sections, indicating that the continuity of structures in our real solution is not caused by smearing.

consistent with the results of Van der Hilst et al. (1997) and Grand et al. (1997). Results of tests dedicated to specific structures (Figures 4.13b, 4.13c, 4.13e, and 4.13f) show that these continuations of subduction to the CMB are not caused by a lack of vertical resolution. Although low-velocity regions are, in general, not as well-resolved as high-velocity subduction zones, it is still possible to observe the surface expression and the deep roots of some of these anomalies. The low velocities underneath the Afar Triangle and the African rift valley seem to extend several hundreds of kilometers downward and appear to have a



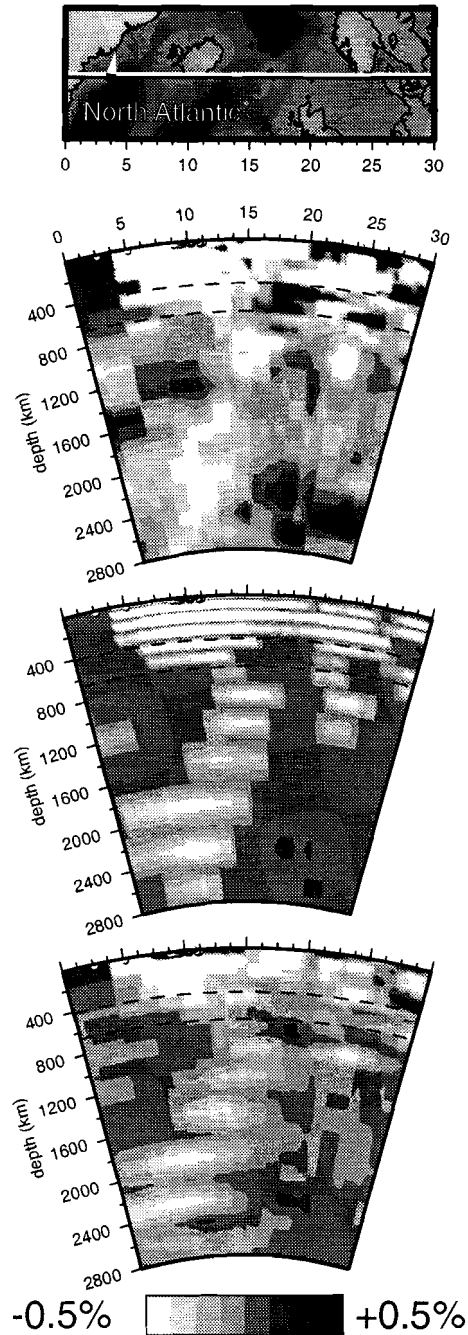
common root under central Africa (where horizontal resolution is limited to  $3^{\circ}$ - $6^{\circ}$ ) which extends all the way to the CMB. This root has been observed by many other global tomography studies (e.g. Inoue et al., 1990; Su et al., 1994; Vasco et al., 1995; Van der Hilst et al., 1997) but without a connection to the upper mantle. We find similar though smaller-scale anomalies underneath the Canary Islands, Yellowstone, and Iceland, extending to at least 1200 km, 1500 km, and the CMB, respectively, but only Iceland is reasonably well-resolved in our model. Figure 4.14 displays a section across Iceland together with a sensitivity result which clearly demonstrates the presence of a continuous low-velocity anomaly from the upper mantle to the CMB beneath Iceland (see also section 5.1). Smearing could be significant in the upper mantle, but a local tomography study (Wolfe et al., 1997) has already demonstrated the presence of low velocities in the upper mantle below Iceland to at least 400 km depth.

We suspect that the approximately 400-700 km wide low-velocity anomaly observed below Iceland is the blurred image of an even narrower rising plume. The conduits of plumes are usually estimated to have a diameter of some 100-200 km (e.g. Ribe and Christensen, 1994), which would be difficult to image. However, Steinberger (1998) argues that especially Iceland may not be a typical plume, but rather a more broad upwelling, because it does not seem to have a hotspot track. In any case, our result in Figure 4.14 gives seismological evidence for the presence of a relatively narrow low-velocity channel all across the (lower) mantle. The upwelling is tilted toward the south-southeast and has a broad root zone that extends to below the western Atlantic and Greenland.

Figures 4.13 and 4.14 constitute important evidence for the continuation of both high and low velocities across the 660 km discontinuity, a depth that is quite well-resolved in the model. These observations again (e.g. Van der Hilst et al., 1997; Grand et al., 1997) indicate that the 660 km discontinuity may resist but not prevent (present day) large-scale mass transfer from the upper to the lower mantle and vice versa.

## 4.6 Conclusions

The lack of sampling in large parts of the Earth's mantle due to the uneven distribution of earthquakes and seismic stations prevents a detailed solution for the entire model space. However, the irregular cell parameterization applied with cell sizes adapted to the degree of local sampling together with the application of a vast amount of reprocessed ISC data allows for the first time for a simultaneous solution for long-wavelength structures and small-scale details wherever warranted by the resolving power of the data. The irregular parameterization not only reduces the number of unknowns drastically, it probably better conditions the inverse problem of determining seismic wave velocities, event cluster relocation parameters, and station parameters from 4.7 million narrow composite rays. The limited sampling of the lower mantle together with the broad Fresnel zones of the seismic waves (Nolet, 1992) and probably wave front healing effects (Wielandt, 1987) prevent a very detailed solution in the lower mantle part of the model. This part of our model confirms and adds more detail to recent results obtained by Van der Hilst et al. (1997) and Grand et al. (1997). Apart from the huge Tethys and Farallon plate anomalies



**Figure 4.14:** Cross-sections through Iceland ((64°N, 44°W) - (54.7°N, 14.5°E)): (a) real solution, (b) synthetic input, and (c) synthetic output. Notice that nearly all layers in the lower mantle can be resolved separately, indicating that the continuity of the low-velocity plume in Figure 4.14a is not caused by vertical smearing.

several low-velocity anomalies are imaged, which continue from the lower mantle up to the Earth's surface. Although the resolution of these features is usually less than for subduction zones, their continuation seems to be resolved. Below Iceland, we have imaged a relatively narrow low-velocity anomaly, the image of possibly a mantle plume, stretching all the way from the CMB to the surface. This may be the first (blurred) image of a narrow plume (see section 5.1). This continuity of both positive and negative anomalies clearly points to the occurrence of mass transport across the 660 km discontinuity.

In the upper mantle we are able to resolve structures locally on a scale of 65 km, which is comparable to and in some cases better than results from regional mantle studies. For the first time, detailed upper mantle structures are imaged adequately on the global scale. The global model confirms many regionally obtained results (especially in the upper mantle), puts them in a global framework, and adds several reasonably resolved regions such as the Scotia Sea, parts of northwest Africa, Iran, and the Solomon Islands for which detailed tomography results have never been published before. Furthermore, it allows for the comparison of the structure of different regions (such as subduction zones) without being hampered by different inversion methods or limited model extent. In comparison with our global model, regional mantle studies often exhibit small-scale anomalies in their lower mantle parts, which may be caused by signal and noise acquired by teleseismic data outside the model space.

In general, all subduction zones are well resolved in lateral extent and some also in depth extent, which allows for the detailed observation and comparison of slab behavior. Above the transition zone, slabs may show kinks (Venezuela and Peru-Bolivia), perhaps slabs can split (not well resolved below New Guinea), and they may detach (Spain, Italy, and Albania), which confirms several regional results. At the 660 km discontinuity, we find clear continuations into the lower mantle (e.g. below Java, southeastern Europe, the Marianas, Japan, Central America, and Tonga) as well as flattening on top of it (e.g. below the western Mediterranean, the Banda arc, and the Izu-Bonin trench) and penetration after flattening (e.g. below Tonga and the Kuriles). These findings are in general agreement with results from regional studies and indicate that the 660 km discontinuity may resist but not prevent penetration of slabs into the lower mantle.

Here, we have tried to give an overview of the most striking observations in our global mantle model of *P*-wave velocity heterogeneity. This should be considered an introduction to the model and the applied method. Detailed interpretations and further comparisons with results from sensitivity tests will be given elsewhere as these would make this paper unduly long. Our primary aim was to present the model and to show the improvement obtained in the global imaging of mantle structure. The results close the resolution gap with regional mantle studies, which is an important step toward a better understanding of the dynamics and interaction of upper and lower mantle.

**Acknowledgments.** We are grateful for support from the Netherlands Geosciences Foundation (GOA) with financial aid from the Netherlands Organization for Scientific Research (NWO) through Pioneer project PGS 76-144, through hardware grant NWO-750.396.02, and for support of H.B. through grant NWO-750.195.13. We thank R.D. Van der Hilst and an anonymous referee for helpful comments, and R.D. Van der Hilst for supplying us with an early version of the data set. H.B. thanks dr. Kok, a naïve reader, for

critical comments on an early version of the manuscript. This work was conducted under the program of the Vening Meinesz School of Geodynamics.

## References

- Abers, G.G., and S.W. Roecker, Deep structure of an arc-continent collision: Earthquake relocation and inversion for upper mantle *P* and *S* wave velocities beneath Papua New Guinea, *J. Geophys. Res.*, 96, 6379-6401, 1991.
- Aki, K., A. Christoffersson, and E.S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, *Geophys. J. R. Astron. Soc.*, 82, 277-296, 1977.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Preliminary results from nonlinear global travel-time tomography (abstract), *Ann. Geophys.* 16, 33, 1998.
- Blanco, M.J., and W. Spakman, The *P*-wave velocity structure of the mantle below the Iberian Peninsula: Evidence for subducted lithosphere below southern Spain, *Tectonophysics*, 221, 13-34, 1993.
- Christensen, U.R., The influence of trench migration on slab penetration into the lower mantle, *Earth Planet. Sci. Lett.*, 140, 27-39, 1996.
- Clayton, R.W., and R.P. Comer, A tomographic analysis of mantle heterogeneities from body wave travel times (abstract), *EOS Trans. AGU*, 64 (45), 776, 1983.
- Curtis, A., and R.K. Snieder, Reconditioning inverse problems using the genetic algorithm and revised parameterization, *Geophysics*, 62, 1524-1532, 1997.
- Dautria, J.-M., Relations entre les hétérogénéités du manteau supérieur et le magmatisme en domaine continental distensif: Exemple des basaltes alcalins du Hoggar (Sahara Central, Algérie) et de leurs enclaves, Ph.D. thesis, Cent. Géol. et Géophys. de Montpellier, Montpellier, France, 1988.
- Dziewonski, A.M., B.H. Hager, and R.J. O'Connell, Large-scale heterogeneities in the lower mantle, *J. Geophys. Res.*, 82, 239-255, 1977.
- Engdahl, E.R., and D. Gubbins, Simultaneous travel time inversion for earthquake location and subduction zone structure in the central Aleutian Islands, *J. Geophys. Res.*, 92, 13855-13862, 1987.
- Engdahl, E.R., R.D. Van der Hilst, and J. Berrocal, Imaging of subducted lithosphere beneath South America, *Geophys. Res. Lett.*, 22, 2317-2320, 1995.
- Engdahl, E.R., R.D. Van der Hilst, and R.P. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seismol. Soc. Amer.*, 88, 722-743, 1998.
- Fukao, Y., M. Obayashi, H. Inoue, and M. Nenberg, Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.*, 97, 4809-4822, 1992.
- Grand, S.P., Mantle shear structure beneath the Americas and surrounding oceans, *J. Geophys. Res.*, 99, 11591-11621, 1994.
- Grand, S.P., R.D. Van der Hilst, and S. Widiyantoro, Global seismic tomography: A snapshot of convection in the Earth, *GSA Today*, 7, 1-7, 1997.
- Granet, M., G. Stoll, J. Dorel, U. Achauer, G. Poupinet, and K. Fuchs, Massif Central (France): New constraints on the geodynamical evolution from teleseismic tomography, *Geophys. J. Int.*, 121, 33-48, 1995.
- Griffiths, R.W., R.I. Hackney, and R.D. Van der Hilst, A laboratory investigation of effects

- of trench migration on the descent of subducted slabs, *Earth Planet. Sci. Lett.*, 133, 1-17, 1995.
- Gudmundsson, O., J.H. Davies, and R.W. Clayton, Stochastic analysis of global traveltime data: Mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, 102, 25-43, 1990.
- Hager, B.H., R.W. Clayton, M.A. Richards, R.P. Comer, and A.M. Dziewonski, Lower mantle heterogeneity, dynamic topography and the geoid, *Nature*, 313, 541-545, 1985.
- Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic travel time problems, *J. Geophys. Res.*, 93, 1073-1085, 1988.
- Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogota, Whole mantle *P*-wave travel time tomography, *Phys. Earth Planet. Inter.*, 59, 294-328, 1990.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, 122, 108-124, 1995.
- Lévêque, J.-J., L. Rivera, and G. Wittlinger, On the use of the checker-board test to assess the resolution of tomographic inversions, *Geophys. J. Int.*, 115, 313-318, 1993.
- Masson, F., and J. Trampert, On ACH, or how reliable is regional teleseismic delay time tomography?, *Phys. Earth Planet. Inter.*, 102, 21-32, 1997.
- Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London Ser. A*, 354, 1385-1411, 1996.
- Mohan, G., and S.S. Rai, Large-scale three-dimensional seismic tomography of the Zagros and Pamir-Hindukush regions, *Tectonophysics*, 242, 255-265, 1995.
- Nakanishi, I., and D.L. Anderson, Worldwide distribution of group velocity of mantle Rayleigh waves as determined by spherical harmonic inversion, *Bull. Seismol. Soc. Amer.*, 72, 1185-1194, 1982.
- Nolet, G., Imaging the deep earth: technical possibilities and theoretical limitations, in: *Proceedings of the 25th ESC Symposium Barcelona 1990*, edited by A. Rocca and D. Mayer-Rosa, pp. 107-115, Servei Geològic de Catalunya, Barcelona, Spain, 1992.
- Olbertz, D.C., The long-term evolution of subduction zones: A modelling study, Ph.D. thesis, Utrecht Univ., Utrecht, Netherlands, 1997.
- Olbertz, D.C., M.J.R. Wortel, and U. Hansen, Trench migration and subduction zone geometry, *Geophys. Res. Lett.*, 24, 221-224, 1997.
- Paige, C.C., and M.A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.*, 8, 43-71, 1982.
- Puspito, N.T., Y. Yamanaka, T. Miyatake, K. Shimazaki, and K. Hirahara, Three-dimensional *P*-wave velocity structure beneath the Indonesian region, *Tectonophysics*, 220, 175-192, 1993.
- Ribe, N.M., and U.R. Christensen, Three-dimensional modeling of plume-lithosphere interaction, *J. Geophys. Res.*, 99, 669-682, 1994.
- Roecker, S.W., Tomography in zones of collision: Practical considerations and examples, in: *Seismic Tomography; theory and Practice*, edited by H. M. Iyer and K. Hirahara, pp. 584-612, Chapman and Hall, New York, 1993.
- Sengupta, M., and M.N. Toksöz, Three-dimensional model of seismic velocity variation in the Earth's mantle, *Geophys. Res. Lett.*, 3, 84-86, 1976.
- Snieder, R., and M. Sambridge, Ray perturbation theory for traveltimes and ray paths in 3-D

- heterogeneous media, *Geophys. J. Int.*, 109, 294-322, 1992.
- Snieder, R., J. Beckers, and F. Neele, The effect of small-scale structure on normal mode frequencies and global inversions, *J. Geophys. Res.*, 96, 501-515, 1991.
- Spakman, W., Delay-time tomography of the upper mantle below Europe, the Mediterranean, and Asia Minor, *Geophys. J. Int.*, 107, 309-332, 1991.
- Spakman, W., Iterative strategies for non-linear travel time tomography using global earthquake data, in: *Seismic Tomography; theory and Practice*, edited by H. M. Iyer and K. Hirahara, pp. 190-226, Chapman and Hall, New York, 1993.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic problems (abstract), *Ann. Geophys.* 16, 28, 1998.
- Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N. J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, pp. 155-187, Reidel, Dordrecht, 1988.
- Spakman, W., S. Stein, R.D. Van der Hilst, and R. Wortel, Resolution experiments for NW Pacific subduction zone tomography, *Geophys. Res. Lett.*, 16, 1097-1100, 1989.
- Spakman, W., S. Van der Lee, and R.D. Van der Hilst, Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, 79, 3-74, 1993.
- Su, W., R.L. Woodward, and A.M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, 99, 6945-6980, 1994.
- Tanimoto, T., Long-wavelength *S*-wave velocity structure throughout the mantle, *Geophys. J. Int.*, 100, 327-336, 1990.
- Trampert, J., and R. Snieder, Model estimations biased by truncated expansions: possible artifacts in seismic tomography, *Science*, 271, 1257-1260, 1996.
- VanDecar, J.C., Upper-mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion, Ph.D. thesis, 165 pp., Univ. of Wash., Seattle, 1991.
- Van der Hilst, R.D., Tomography with *P*, *PP* and *pP* delay-time data and the three-dimensional mantle structure below the Caribbean region, Ph.D. thesis, Utrecht Univ., Utrecht, Netherlands, 1990.
- Van der Hilst, R.D., Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, *Nature*, 374, 154-157, 1995.
- Van der Hilst, R.D., and W. Spakman, Importance of the reference model in linearized tomography and images of subduction below the Caribbean plate, *Geophys. Res. Lett.*, 16, 1093-1096, 1989.
- Van der Hilst, R.D., E.R. Engdahl, W. Spakman, and G. Nolet, Tomographic imaging of subducted lithosphere below northwest Pacific island arcs, *Nature*, 353, 37-43, 1991.
- Van der Hilst, R.D., E.R. Engdahl, and W. Spakman, Tomographic inversion of *P* and *pP* data for aspherical mantle structure below the northwest Pacific region, *Geophys. J. Int.*, 115, 264-302, 1993.
- Van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, 386, 578-584, 1997.
- Vasco, D.W., L.R. Johnson, R.J. Pulliam, Lateral variations in mantle velocity structure and discontinuities determined from *P*, *PP*, *S*, *SS*, and *SS-SdS* travel time residuals, *J. Geophys. Res.*, 100, 24037-24059, 1995.

- Widiyantoro, S., and R.D. Van der Hilst, Structure and evolution of lithospheric slab beneath the Sunda arc, Indonesia, *Science*, 271, 1566-1570, 1996.
- Widiyantoro, S., and R.D. Van der Hilst, Mantle structure beneath Indonesia inferred from high-resolution tomographic imaging, *Geophys. J. Int.*, 130, 167-182, 1997.
- Wielandt, E., On the validity of the ray approximation for interpreting delay times, in: *Seismic Tomography*, edited by G. Nolet, pp. 85-98, D. Reidel, Norwell, Mass., 1987.
- Wolfe, C.J., I.Th. Bjarnason, J.C. VanDecar, and S.C. Solomon, Seismic structure of the Iceland mantle plume, *Nature*, 385, 245-247, 1997.
- Wong A Ton, S.Y.M., and M.J.R. Wortel, Slab detachment in continental collision zones: An analysis of controlling parameters, *Geophys. Res. Lett.*, 24, 2095-2098, 1997.
- Woodhouse, J.H., and A.M. Dziewonski, Mapping the upper mantle: Three-dimensional modelling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, 89, 5953-5986, 1984.
- Wysession, M E., Imaging cold rock at the base of the mantle: The sometimes fate of slabs?, in: *Subduction: Top to Bottom*, *Geophys. Monogr. Ser. Vol. 96*, edited by G. E. Bebout et al., pp. 369-384, AGU, Washington, D.C., 1996.
- Zhang, Y., and T. Tanimoto, High-resolution global upper mantle structure and plate tectonics, *J. Geophys. Res.*, 98, 9793-9823, 1993.
- Zhao, D., D. Christensen, and H. Pulpan, Tomographic imaging of the Alaska subduction zone, *J. Geophys. Res.*, 100, 6487-6504, 1995.
- Zhou, H., How well can we resolve the deep seismic slab with seismic tomography?, *Geophys. Res. Lett.*, 15, 1425-1428, 1988.
- Zhou, H., A high-resolution *P* wave model for the top 1200 km of the mantle, *J. Geophys. Res.*, 101, 27791-27810, 1996.
- Zhou, H., and R. W. Clayton, *P* and *S* wave travel time inversions for subducting slab under the island arcs of the northwest Pacific, *J. Geophys. Res.*, 95, 6829-6851, 1990.

# Chapter 5

## Regional interpretations of the global *P*-wave model

This chapter provides four detailed interpretations of well-resolved regions in the global model derived in Chapter 4. Section 5.1 describes the mantle volume below Iceland and was published as Bijwaard and Spakman (1999a); section 5.2 deals with the deep structure below Siberia and represents a summary of Van der Voo et al. (1999a); section 5.3 gives a detailed interpretation of the Tethyan anomalies and summarizes Van der Voo et al. (1999b); finally, in section 5.4 small-scale structure below southeast Asia is interpreted following Spakman and Bijwaard (1998a), and Rangin et al. (1999).

### 5.1 Tomographic evidence for a narrow whole mantle plume below Iceland

**Abstract.** Iceland is known to be a center of both hotspot and Mid-Atlantic Ridge activity. A plume origin for the hotspot has been suggested by observational, laboratory and numerical investigations as well as by several seismological studies. The latter have, however, not yielded a detailed image of the entire mantle volume below Iceland. We present tomographic images of a narrow low velocity anomaly below Iceland, extending from the core-mantle boundary (CMB) to the surface. From these images we infer a bent plume with a diameter  $\leq 500$  km, rising from a broad root zone ( $>1000$  km) and culminating in a plume top with a diameter of  $\sim 1200$  km. The anomaly amplitude can, if only controlled by temperature, be converted into excess temperatures in the upper and lower mantle on the order of 200-300 K and  $\sim 200$  K, respectively. The wavy shape of the anomaly together with its lateral branches, possibly indicates that the position of the Iceland hotspot is not stationary.

#### 5.1.1 Introduction

Iceland has long been known to be the locus of both a segment of the Mid-Atlantic Ridge (MAR) and a hotspot (Wilson, 1963; Schilling, 1991). On the one hand, the neovolcanic zones on Iceland are similar to oceanic spreading center segments and are commonly interpreted as on-land extensions of the MAR; on the other hand, the excessive volcanism with a geochemically distinct signature (Schilling, 1973; Klein and Langmuir, 1987), the extremely thick (16 km) oceanic crust (White and McKenzie, 1989), and the large (1000 km across or more) topographic swell (Crough, 1983) all indicate the simultaneous presence of a hotspot.

Since Morgan (1971) hotspots have been associated with mantle 'plumes': Relatively narrow channels of hot mantle upwellings. Laboratory experiments (Whitehead and Luther, 1975; Griffiths and Campbell, 1990) and numerical models (e.g. Ribe et al., 1995; Ito et al., 1996) have indicated that plumes emerge from boundary layers in which a temperature difference of 500 K is enough for plumes to rise to the bottom of the lithosphere (Albers and Christensen, 1996).



Although numerical modeling and laboratory experiments have helped to gain insight into the dynamics of plumes, several important questions remain unanswered. First of all, it is unclear whether plumes may originate at the upper-to-lower mantle discontinuity (White and McKenzie, 1989; Fitton et al., 1997) and/or at the CMB (Griffiths and Campbell, 1990; Davies, 1998), although recent evidence points toward the CMB as the source for at least the Hawaiian hotspot (Brandon et al., 1998). Second, from their small relative motion, the positions of hotspots have been inferred to be more or less stationary (e.g. Duncan and Richards, 1990). However, plumes may tilt due to mantle currents (Skilbeck and Whitehead, 1978), and kinematic models of hotspot mobility due to mantle flow indicate hotspot motion of several mm/yr (Steinberger and O'Connell, 1998). Finally, theoretical predictions, laboratory experiments and numerical models yield different estimates of the size of the plume conduit (50-500 km radius) and its temperature difference with respect to the surrounding mantle (100-800 K) (e.g. Loper and Stacey, 1983; Ito et al., 1996; Thompson and Tackley, 1998).

Several local seismological studies have tried to image plumes and to answer some of these questions (Nataf and VanDecar, 1993; Pritchard et al., 1996; Wolfe et al., 1997; Shen et al., 1998). From these studies the signature of a plume has been inferred below Iceland up to 1500 km depth (Pritchard et al., 1996) with between 100 and 400 km depth a diameter of ~300 km and an excess temperature of 200-300 K (Wolfe et al., 1997). On the global scale, the signature of the Iceland plume has been picked up as well (e.g. Zhang and Tanimoto, 1993; Su et al., 1994; Zhou, 1996; Grand et al., 1997). Most global studies image low velocities in parts of the upper or lower mantle below Iceland. Only Su et al. (1994) find a continuous low velocity column, but their degree 12 parameterization cannot adequately depict structures smaller than approximately 1500 km in lateral dimension.

Recently, Bijwaard et al. (1998a) (Chapter 4) constructed a relatively high-resolution global mantle tomography model from the data set of Engdahl et al. (1998). The implementation of a large amount of narrow ray bundles and the application of an irregular cell parameterization (Spakman and Bijwaard, 1998b) in which cell sizes depend on hitcount allowed for a detailed solution in well-sampled mantle volumes. The lateral cell dimensions in their model vary below Iceland from 65-200 km in the upper mantle to 150-400 km in the lower mantle. The vertical cell dimensions increase from 35 km in the lithosphere to 200 km in the lower mantle.

This section focuses on the description, resolution and interpretation of the low-velocity anomaly we observe below Iceland and the surrounding Atlantic. This anomaly, which extends from the CMB to the surface, is inferred to be the first rather detailed image of a whole mantle plume.

### *5.1.2 Description of the Iceland anomaly*

Figure A4 shows the low velocities (with respect to reference model ak135 (Kennett et al., 1995)) below Iceland, stretching from the CMB to the surface. The low velocity anomaly is not shaped as a cylindrical column, but instead it is a tilted feature slightly broader in N-S direction than in E-W direction. The width of the central part of the anomaly is rather constant (400-700 km in E-W direction) from D'' to the transition zone, but above the transition zone it increases to about 1500 km, nearly the entire width of the North Atlantic

ocean. The center of the anomaly lies beneath Iceland in the upper mantle, shifts towards the Faeroe Islands at 1000 km depth, then returns to below southern Iceland at 1900 km depth and from there on migrates in the direction of Greenland.

High amplitudes are observed in the lithosphere (>5%), decreasing to ~1% at the upper-to-lower mantle transition. There is a dip in amplitude around 800 km depth (0.3%), but from there on the amplitude remains rather constant at 0.5% throughout the lower mantle until D'', where it decreases again to below 0.3%.

In cross-section the anomaly may appear to be a rather simple structure, but the layer sections in Figure A5 show that it is not. We observe at several depths lateral branches or connections of the low velocity heterogeneity: From 200 to 700 km a strong anomaly branches off towards eastern Greenland. From 800 to 1200 km a parallel zone of low velocities develops from Norway to Ireland. The Norwegian part of this anomaly disappears around 1300 km, but the remainder starts to connect with central Europe. At depths from 900 to 1400 km a semicircular belt of low velocities exists from Iceland to the Eifel, the Massif Central, the Hoggar hotspot (northern Africa), and the Canary Islands and Cape Verde (not shown). From 1700 km to the D'', low velocities cover a vast area, initially from northwestern Europe across the Atlantic to eastern Greenland and then from Iceland to western Greenland and the western Atlantic. Hence, the relatively narrow mid-mantle anomalies are rooted in a broad lowermost mantle low-velocity zone.

As mentioned by Bijwaard et al. (1998a) (Chapter 4) the Iceland anomaly is one of several low wavespeed structures that can spatially be associated with hotspots. Low velocities are, for example, also present below Yellowstone, the Society Islands, the Canary Islands, and central Africa. The mantle below Iceland proves to be one of the places for which we can demonstrate good resolution for the entire mantle plume.

### 5.1.3 Resolution of the Iceland anomaly

The spatial resolution in the global mantle model of Bijwaard et al. (1998a) (Chapter 4) was estimated with synthetic sensitivity experiments. These included, first of all, so-called 'spike-tests' (Spakman and Nolet, 1988) in which artificial travel times (supplemented with random noise with a standard deviation of 0.5 s), calculated for a synthetic spike structure, were inverted. These tests indicate that the lateral resolution below Iceland and the eastern Atlantic ocean is on the order of 3° above the transition zone. In the transition zone, the area recovered on a scale of 3° increases to include parts of the western Atlantic and Greenland, but the space between spikes starts to fill up due to smearing effects. Down to 1000 km, 3° spikes can still be resolved (now below the entire Atlantic ocean, Iceland and Greenland), but from there on resolution gradually decreases to 6° at the top of D''. In D'' even for 6° spikes some smearing occurs, but at this depth these have dimensions of only ~300 km. In summary: Lateral resolution for the Iceland anomaly is estimated at 300-500 km throughout the entire mantle.

To assess the vertical resolution of the Iceland anomaly we constructed a synthetic layer-cake model similar to Bijwaard et al. (1998a) (Chapter 4). Figure A4b indicates that in the lower mantle, where separate layers of the layer-cake are easily retrieved, the Iceland anomaly is well-resolved in the vertical direction on a scale of 120-200 km, but that vertical resolution in the upper mantle for 200 km thick layers is mediocre. Similar to the real

solution, the layer-cake results have low amplitudes around 800 km depth.

The linear tomography approach of Bijwaard et al. (1998a) (Chapter 4) does not take into account the bending of seismic rays due to the inferred velocity heterogeneity. However, preliminary results from a nonlinear approach, involving 3-D ray tracing (Bijwaard et al., 1998b), indicate that for the mantle below Iceland, the linear and nonlinear solutions are very similar (see also Chapter 8).

#### *5.1.4 Interpretation of the Iceland anomaly*

Combining our tomographic inferences with the experimental evidence for plumes as the origins of hotspots and the apparent hotspot characteristics of the Iceland region, we arrive at the following interpretations.

The continuity of the anomaly from the CMB to the surface suggests a D'' origin for the Iceland plume. We note that the rather small anomaly amplitudes around 800 km depth and in D'' may be due to the limited vertical and lateral resolution, respectively. The initiation of plumes near the CMB is often observed in mantle convection studies (Davies, 1998) and recent geochemical evidence points to D'' as the source of the Hawaiian lava's (Brandon et al., 1998). On the basis of the inferred vertical resolution of 120-200 km we can, however, not exclude small-scale interruptions of the seemingly continuous anomaly. With the present resolution we can, for example, not distinguish between a continuous upwelling from the CMB to the surface or an upwelling from the upper-to-lower mantle discontinuity, induced by an upwelling from the CMB, as has been inferred from some numerical studies (Steinbach and Yuen, 1997).

The waviness of the Iceland anomaly, which is resolved laterally, suggests a non-stationary plume that may be advected by prevailing mantle currents. This is in agreement with recent kinematic calculations (Steinberger and O'Connell, 1998; Steinberger, 1998) that suggest that plume conduits may be displaced with respect to each other. The observed lateral low velocity branches appear to be connected with the plume-shaped anomaly. These connections are resolved on a scale of 300-500 km. In this respect, the mid-mantle branch stretching below hotspots in Europe and northern Africa is an intriguing feature.

The width of the Iceland anomaly suggests a plume diameter of maximally 400-700 km in the lower mantle, but probably less, due to the resolution length scale on the order of 300-500 km. On comparison of the sizes of synthetic input structures and their output (not all shown), we estimate the actual plume diameter to be  $\leq 500$  km. This is within the range of recent estimates for Iceland from numerical modelling (Ito et al., 1996; Thompson and Tackley, 1998). The broadening of the low velocity anomaly in the lower mantle (from approximately 2000 km depth to the CMB) is an interesting feature that seems to be resolved according to Figure A5h, but we cannot exclude that it may be caused by lateral smearing of smaller anomalies as is shown in Figure A4b between 2200 and 2400 km depth.

Above the transition zone the plume may be as broad as  $\sim 1200$  km, which is in agreement with laboratory estimates of the width of the top of a plume after spreading below the lithosphere (Griffiths and Campbell, 1991). There are, however, several uncertainties: Lateral resolution in the western Atlantic ocean is rather limited above the transition zone and the anomaly width is thus difficult to infer. Furthermore, laboratory experiments indicate that after spreading the top of the plume resembles a broad but thin (200 km) disc (Griffiths and

Campbell, 1991). However, the width of the imaged anomaly remains on the order of 1000-1500 km down to at least 400 km depth. This may be due to the limited vertical resolution, which is on the order of 200 km. A surface wave study (Zhang and Tanimoto, 1993) imaged such a broad anomaly up to only 250 km depth. Finally, Iceland is also part of the Mid-Atlantic Ridge and thus part of the imaged velocity anomaly at shallow depths is probably associated with structure below the mid-ocean ridge: It is impossible to distinguish between the heterogeneity induced by hotspot or mid-ocean ridge activity.

Both thermal and compositional anomalies as well as other effects may account for the observed low wavespeeds, but at present we have no means to discriminate between these contributions. However, if we assume that the anomalous velocities in the lower mantle are entirely caused by anomalous temperatures, we may apply the estimates of  $\partial V_p / \partial T$  of Karato (1993) to obtain maximum temperature anomalies on the order of 200 K. For the upper mantle below ~300 km we obtain similar estimates of 200-300 K. This is in excellent agreement with the observationally determined excess temperature (263 K) (Schilling, 1991). We must note, however, that the amplitudes in the model of Bijwaard et al. (1998a) (Chapter 4) may be underestimated due to model damping and as a result of non-ideal spatial resolution and that these bulk estimates of temperature anomalies may thus be underestimated as well. A recent study by Ji and Nataf (1998) has indicated much higher amplitudes for the Hawaiian region, which would argue for a non-thermal component.

### 5.1.5 Conclusions

Iceland is known to be both a hotspot and a part of the Mid-Atlantic Ridge. We have inferred from the global mantle model of Bijwaard et al. (1998a) (Chapter 4) a relatively narrow ( $\leq 500$  km) whole mantle upwelling, interpreted as a plume with a broad ( $>1000$  km) root zone near the CMB and culminating in a ~1200 km wide plume top above 400 km depth. The anomaly below Iceland has a complicated shape: it is deflected towards the east and it has several branches. This may indicate that the position of the Iceland hotspot is not stationary (Steinberger, 1998). Furthermore, if temperature is the only cause, the amplitude of the anomaly yields an estimate of the temperature difference (200-300 K), which agrees well with the excess temperature calculated by Schilling (1991).

**Acknowledgments.** We are grateful for support from the Netherlands Geosciences Foundation (GOA) with financial aid from the Netherlands Organization for Scientific Research (NWO). This work was conducted under the program of the Vening Meinesz School of Geodynamics.

## 5.2 Mesozoic subducted slabs under Siberia

The interpretations of tomography results provided in this section have been derived from Van der Voo et al. (1999a). Their study was based on the southeast Asia part of the global tomography model obtained in Chapter 4 (Bijwaard et al., 1998a). This section basically gives a summary of their results.

### 5.2.1 Paleogeographic background

Since Australia and the Americas are slowly approaching the Eurasian continent, this tectonic plate is likely to become the next supercontinent. However, the amalgamation of the Asian continent already started long ago by continental accretion around its Siberian core (Şengör and Natal'in, 1996). In late Jurassic (Parfenov, 1991) or early Cretaceous (Enkin et al., 1992) times, the combined Mongolia and North and South China blocks joined Siberia. In this continental collision the Mongol-Okhotsk (or Khangai-Khantey according to Şengör and Natal'in, 1996) ocean was destroyed. More or less simultaneously, the northeasternmost Siberian block (Omolon) accreted and thereby closed the Kular-Nera ocean (Şengör and Natal'in, 1996). In the present-day situation the sutures of the Mongol-Okhotsk and Kular-Nera oceans are still visible near Lake Baikal and the Verkhoyansk mountains, respectively. However, the deeper (lower mantle) structure below these sutures has, up to now, not thoroughly been investigated.

### 5.2.2 Tomography results

The tomography results in Chapter 4 (Bijwaard et al., 1998a), and also those obtained by Van der Hilst et al. (1997), clearly show high-velocity anomalies below Asia associated with the subduction of Pacific and Tethyan oceanic plate material. Figure A6 gives an overview of the tomography model of Chapter 4 for the (lower) mantle below Siberia. In the uppermost mantle and crust, the Mongol-Okhotsk and Kular-Nera sutures are marked by *P*-wave velocity contrasts that have also been observed in surface wave studies (e.g. Ritzwoller and Levshin, 1998; Curtis et al., 1999). In the deeper mantle, anomalies related to subduction of Pacific (P), Kula-Farallon (K-F), and Tethyan (T) oceanic material can easily be identified. Below approximately 1200 km a single high-velocity anomaly (M) with amplitudes of 0.3% and more, remains unidentified. This anomaly starts to form a hook below 1500 km, running in west-northwest direction from Mongolia and northward to the Arctic coast of Siberia. Between 1900 and 2300 km this hook slowly turns into a Z shape. At about 2500 km depth this Z-shaped anomaly starts to connect with a very broad high-velocity anomaly stretching below most of East Asia in the lowermost mantle. This anomaly has been interpreted as a slab graveyard by Wysession (1996).

In Figure A7, the unidentified anomaly (M) is shown in a cross-section through the Siberian mantle, which also depicts the high-velocities related to subduction of Pacific (P) material (clearly outlined by earthquakes in the upper mantle). Both high-velocity structures seem to end in the proposed slab graveyard (Wysession, 1996). Panels (b) and (c) show input and output of a synthetic layer-cake test, which demonstrates that these structures are both well-resolved in the vertical sense below 1500 km.

### 5.2.3 Interpretation

As the unidentified anomaly in Figures A6 and A7 has similar amplitude and structure as the subduction related anomalies of Pacific and Tethyan origin, it seems logical to interpret

this anomaly as a remnant of subduction. Although the anomaly is locally connected to upper mantle structure below Siberia, this connection shall not be discussed here, because it is not well-resolved.

Because traces of all Cenozoic subductions below Siberia have been identified in Figure A6, the unidentified structure must be of older (Mesozoic) age. The identification of older subductions is, however, complicated by the fact that the Asian continent may have moved significantly with respect to the underlying mantle. Unfortunately, indications of hotspot motion (Cande et al., 1995; Tarduno and Cottrell, 1997) disqualify the hotspot reference frame for the inference of motion of the Asian continent. An estimate of this motion can, however, be derived from the positions of Mesozoic subductions of the Pacific plate. The western subduction of the Pacific plate dips westward, but its lower mantle portion appears to be nearly vertical (see Figure A7). The eastern subduction of the Pacific plate, however, dips eastward below North America and is observed below the Great Lakes and as far east as New England (Grand et al., 1997, and Chapter 4). This indicates that North America has moved westward since the onset of these subductions and that Asia has remained rather fixed. This is supported by the fact that other anomalies below Asia related to subduction of Tethyan material do not seem to be shifted in E-W direction. As a consequence of the fixity of Eurasia, the main opening of the northern Atlantic ocean must have been compensated by westward drift of North America.

The position of the unidentified structure approximately below the Mongol-Okhotsk suture, and the inferred fixity of Eurasia since the closure of the Mongol-Okhotsk ocean, identify this ocean as the prime candidate for this anomaly. Since the subduction of the Mongol-Okhotsk ocean ceased some 150 Myr ago and the top of the anomaly is retrieved at approximately 1500 km depth, this implies an average sinking rate of 1 cm/yr since Jurassic time. This corresponds reasonably well with sinking rates estimated by Grand et al. (1997) and is slightly below rates inferred from geodynamic modelling (Lithgow-Bertelloni and Richards, 1998). Even if subduction occurred only on the Siberian side, the Mongol-Okhotsk ocean must have been at least 2000 km wide, 200 Myr ago. This is in agreement with plate tectonic reconstructions (Scotese and Golonka, 1992).

#### *5.2.4 Conclusions*

The high-velocity anomaly below central Siberia as observed in the tomographic results of Chapter 4, can be explained by subduction of the Mongol-Okhotsk ocean before the early Cretaceous. This implies that slab material of Jurassic age is still visible 150 Myr after the cessation of subduction. This material and that of Pacific subduction seems to feed an earlier inferred slab graveyard below east Asia (Wyssession, 1996). The interpretation of yet another lower mantle high-velocity anomaly in terms of subduction related material, indicates that most (if not all) significant high-velocity structures in the lower mantle are associated with subduction. This identifies seismic tomography as an important tool to test paleogeographic reconstructions (see also De Jonge et al., 1994). Furthermore, with respect to Eurasia, the many inferred downwellings may be characteristic of a growing supercontinent (Gurnis, 1988).

### 5.3 Tethyan subducted slabs under India

This section describes and interprets the high-velocity anomalies that are found in the model of Chapter 4 (Bijwaard et al., 1998a) below India and its surroundings. It represents a summary of Van der Voo et al. (1999b) who give a more elaborate treatise of this subject.

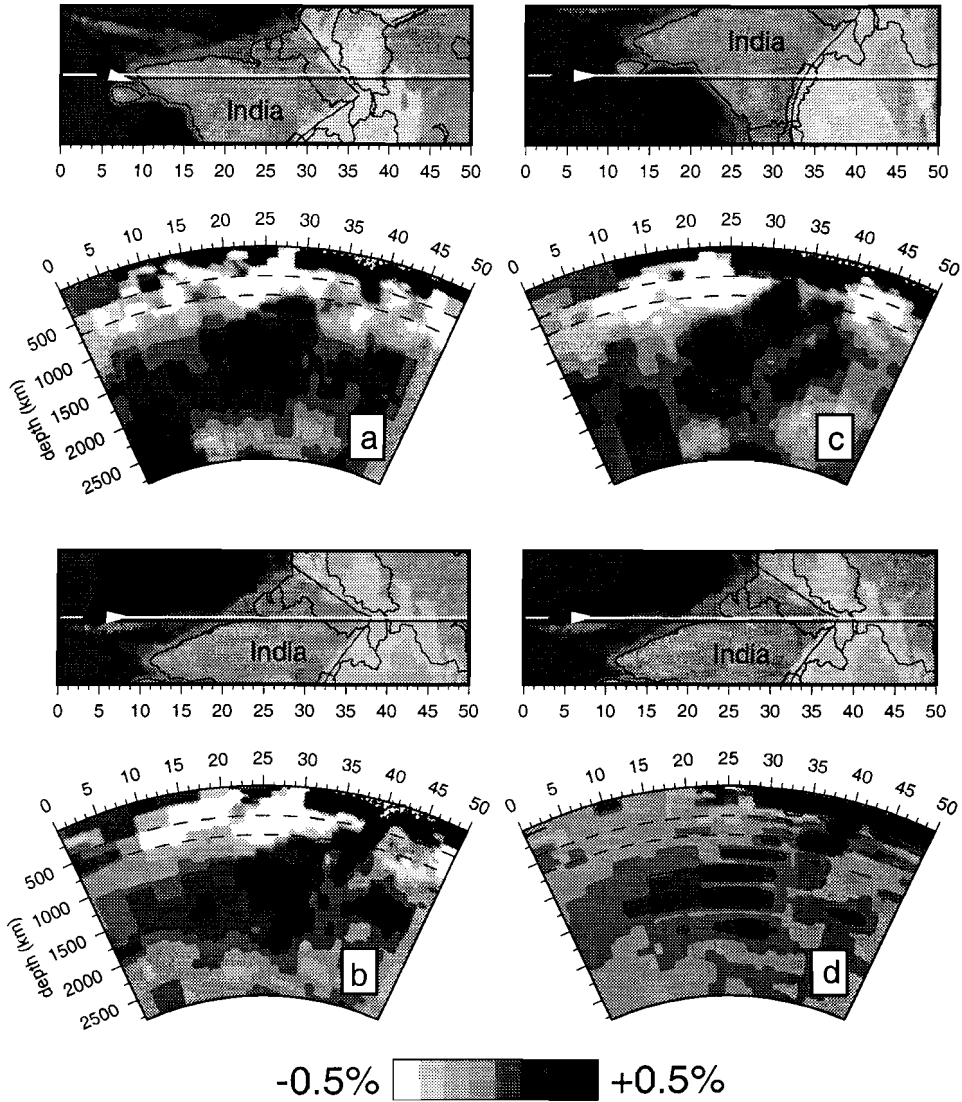
#### 5.3.1 Paleotectonic overview

In the early Tertiary (50-65 Myr ago), the subcontinent of India joined Asia (Patriat and Achache, 1984; Scotese and Golonka, 1992; Klootwijk et al., 1992; Lee and Lawver, 1995; Rowley, 1996). In this collision some 1000 km of Greater Indian crust was probably buried in the Tibetan lithosphere (Veevers et al., 1975; Klootwijk et al., 1985; Zhao and Morgan, 1987; Nelson et al., 1996). At that time, the Tibetan lithosphere already consisted of the Qiangtang and Lhasa blocks that joined Asia in Triassic and Jura/Cretaceous times, respectively (Dewey et al., 1988; Dercourt et al., 1993; Şengör and Natal'in, 1996). These blocks formed, together with parts of Turkey, Iran, and Indochina, the Cimmerian continent that separated from Gondwana in the early Mesozoic (Şengör, 1984). While the Cimmerian continent moved northward to join Asia, the Paleo- or Meso-Tethys ocean was subducted to its north and the Neo-Tethys ocean was formed to its south. At its peak the Neo-Tethys ocean stretched from India (still attached to Gondwana) to Tibet (some 35°-60°, see Besse and Courtillot, 1988; Scotese and Golonka, 1992; Klootwijk et al., 1992). Because it is very likely that some sea-floor spreading occurred within this ocean, it is estimated here that before India's collision with Tibet some 4000-7000 km Neo-Tethyan lithosphere was subducted.

#### 5.3.2 Tomographic results

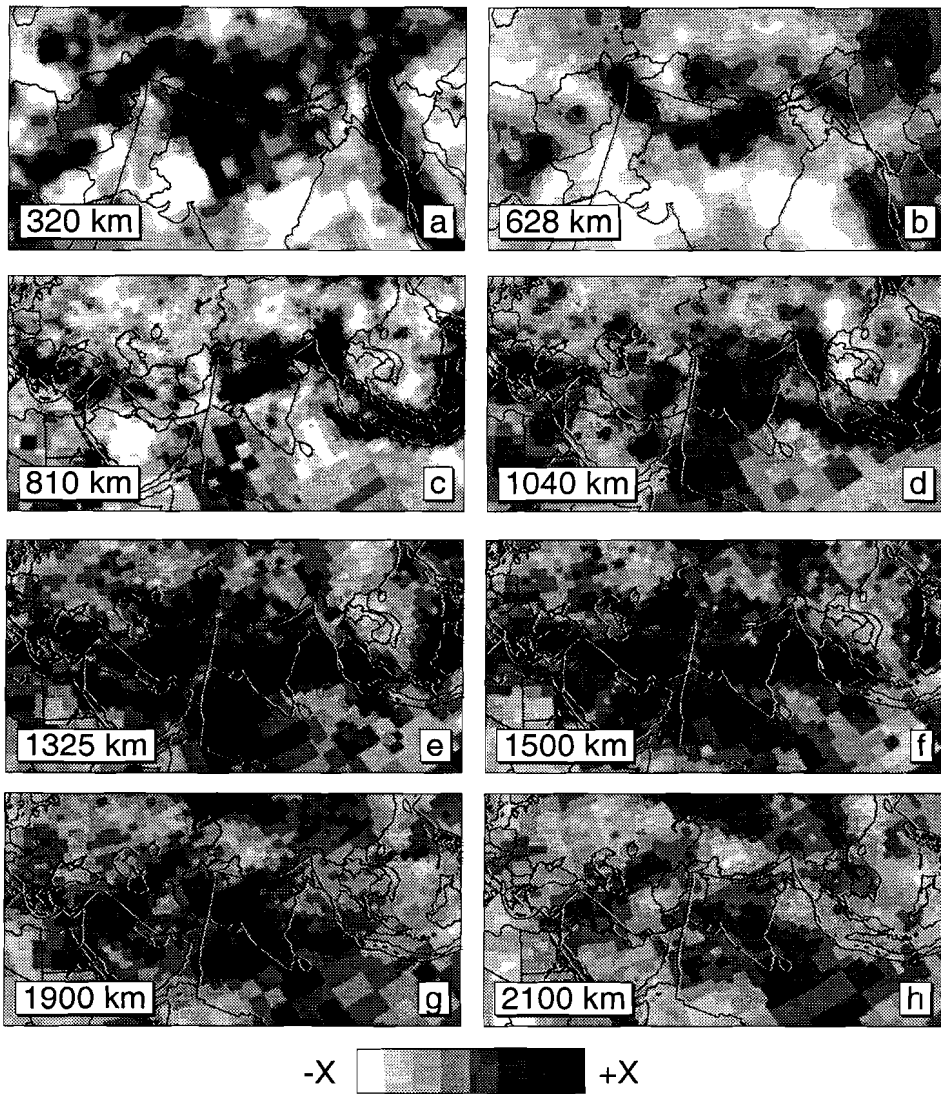
The tomography model of Chapter 4 (Bijwaard et al., 1998a) is depicted for India and its surroundings in Figures 5.1 and 5.2. In these well-resolved images (see Figure 5.3) the following structures can be identified. First of all, the Indian and Tibetan crust and uppermost mantle is characterized by high *P*-wave velocities with the exception of the area around the Deccan traps (see also Kennett and Widiyantoro, 1999). In the deeper upper mantle, a north dipping subduction related anomaly can be observed beneath northern Afghanistan (Hindu Kush) down to some 600 km depth (Figure 5.1a) as has previously been observed in regional tomography studies (Roecker, 1982; Mohan and Rai, 1995). More to the east a similar, but seemingly rolled over, structure can be observed below Pakistan (Figure 5.1b). This roll-over is confirmed by recently relocated hypocenters (Pegler and Das, 1998). Even further eastward, under Nepal, no clear connection between lithospheric and deeper (400-1000 km) anomalies seems to exist (Figure 5.1c). This may suggest a detachment of subducted material. In general, the map images show that the high-velocity anomalies between 300 and 810 km follow the curved trend of the Himalayas up to Yunnan (SW China), where they form a cusp and continue southward through Burma and Malaysia to connect with the Sunda arc subduction.

In the deeper lower mantle (below 1000 km) several high-velocity zones can be seen that



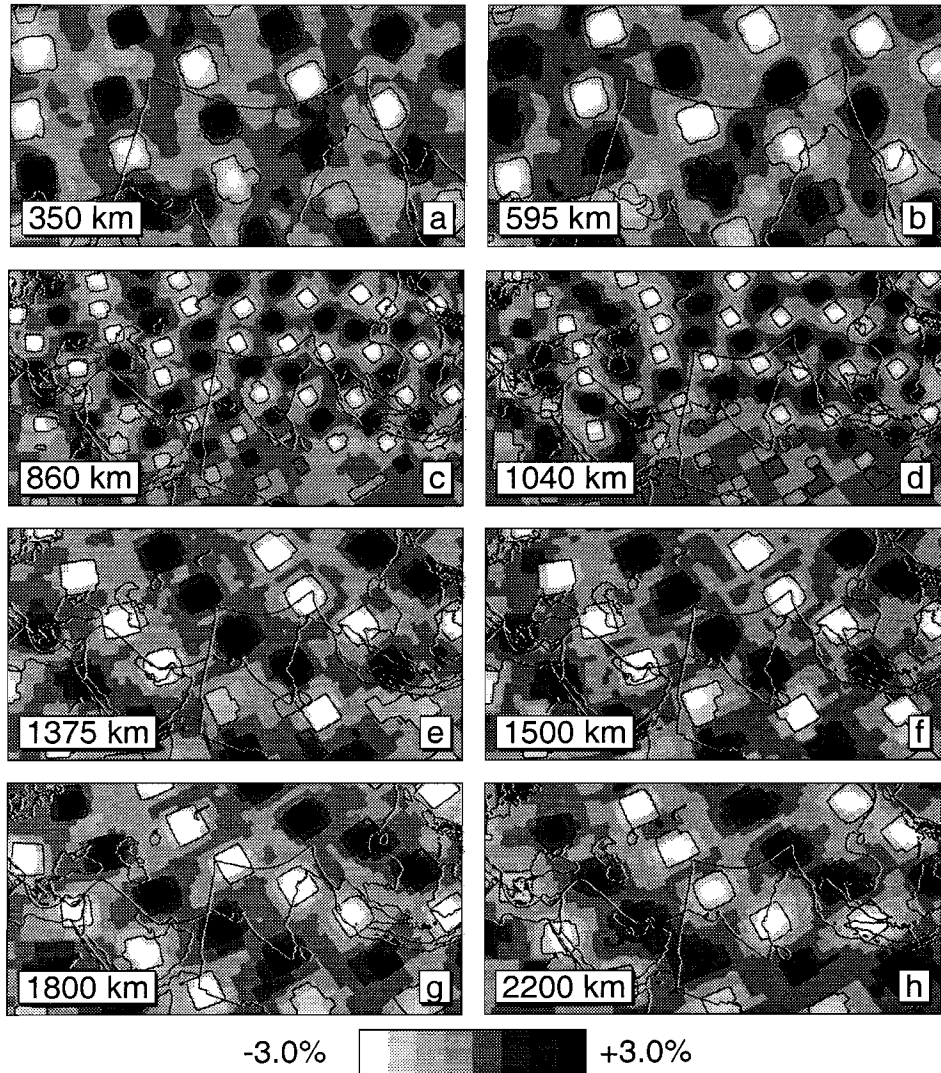
**Figure 5.1:** Whole mantle cross-sections for the Indian-Tibetan mantle and lithosphere. Velocity anomalies are displayed in percentages with respect to reference model ak135 (Kennett et al., 1995). (a) cross-section below northern Afghanistan (Hindu-Kush) showing a northward dipping slab and the large volume of lower mantle high-velocity anomalies; (b) cross-section further East below northern Pakistan showing a seemingly rolled-over slab; (c) cross-section further East below Nepal showing no clear connection between lithospheric and deeper high-velocity anomalies. Panel (d) depicts synthetic input (black contour lines) and output (color) of a layer-cake test for the same cross-section as (b).





**Figure 5.2:** Layer sections of the upper mantle ((a) and (b)) and lower mantle ((c) to (h)) below Greater India and the entire Tethyan region, respectively. For panels (a) and (b)  $X=1\%$ ; for panels (c) to (h)  $X=0.5\%$ . Notice the subdivision of the large Tethyan anomaly in three parallel, E-W striking zones, clearly visible in panels (d), (e), and (f).

are separate from each other and from the anomalies above 1000 km. Three parallel zones stretch WNW-ESE till some 1900 km depth (see Figure 5.2), with the middle one connected to high velocities below Indonesia. Together, the three zones form the Tethys anomaly as it has been identified by Van der Hilst et al. (1997), Grand et al. (1997), and also in Chapter



**Figure 5.3:** Input (black contour lines) and output (gray tones) of synthetic spike tests for the same layer sections as in Figure 5.2. In panels (a) to (d)  $3^\circ$  spikes are shown, and in panels (e) to (h)  $6^\circ$  spikes. All spikes have (absolute) input amplitudes of  $\pm 5\%$ .

4 (Bijwaard et al., 1998a) that has been related to subduction of Tethyan lithosphere. Resolution estimates for this region (shown in Figure 5.3) confirm the observed separation between the different anomalies. In general, the three zones (numbered 1, 2, and 3 from North to South) have the following characteristics. Zone 2, which is the most pronounced, runs almost straight from western India to Sumatra. Zone 3 runs parallel to zone 2 from

Saudi-Arabia to south of India, where the resolution degrades. According to the results of Vasco and Johnson (1998), who have better sampling of this region due to the inclusion of more later arrivals, this zone may stretch further southeastward to west of Australia. Zone 1, located below northern India and Tibet, appears to go slightly deeper than the others, perhaps as deep as the core-mantle boundary.

### 5.3.3 Interpretation

Clearly, the deep mantle high-velocity anomalies below the Mediterranean-Himalayas-Indonesia region are related to subduction of the Tethys ocean (see also Van der Hilst et al., 1997; Grand et al., 1997). However, a comparison of the model of Chapter 4 (Bijwaard et al., 1998a) with the outcome of the resolution tests in Figure 5.3, seems to justify a subdivision of the Tethyan structures in three zones. These three lower mantle zones lie south of the present day boundary between India and Tibet. This may be due to the continued northward motion of India after their subduction. During its collision with Tibet, India rotated some 35° counterclockwise with respect to Tibet (Van der Voo, 1993) which could explain why the upper mantle anomalies are rolled over beneath Pakistan and perhaps detached below Nepal. This rotation, which took place after collision, also identifies these anomalies as due to *continental* subduction (see also Matte et al., 1997; Mattauer et al., 1999), and explains why the cusp in high velocities below Burma is only observed in the upper 900 km. This continental subduction should be younger than zone 2, which is probably due to subduction of the Tethys ocean (see below), and hence it is suggested here that it may represent delaminated sub-continental lithosphere. This is supported by the fact that the amount of slab-related anomalies in the upper mantle approximately matches the amount of Indian continental lithosphere that has disappeared since the collision (some 1000 km).

The three lower mantle zones of high velocities can be explained in terms of Tethyan subductions in the following way. Zone 2 can be related to subduction of the Neo-Tethys ocean. It is located slightly north of the inferred position of this subduction (Westphal et al., 1983; Pozzi et al., 1984; Besse et al., 1984; Besse and Courtillot, 1988; Klootwijk et al., 1992; Patzelt et al., 1996), but this may well be due to a northward dip of the subducting slab. The interpretation of zone 2 in terms of Neo-Tethyan subduction leads to an interpretation of zone 1 as due to subduction of the last remnant of Paleo- or Meso-Tethys to the north of the Cimmerian continent (Dercourt et al., 1993). The remaining zone 3 can then speculatively be interpreted as due to intra-Tethys subduction (as in Reuber, 1986; Allègre, 1988; Honegger et al., 1989; Besse and Courtillot, 1988; Beck et al., 1996; Searle et al., 1997), for which timing (mid to late Cretaceous) and paleolatitude (near-equatorial) match the observations.

If the above interpretations are correct, then sinking rates can be estimated from the timing of the cessation of the different subductions: 130 Myr ago for zone 1, 55-65 Myr ago for zone 2, and 70-85 Myr ago for zone 3. This leads to average sinking rates of 1-2 cm/yr, which corresponds well with estimates in the previous section and in Grand et al. (1997), but is slightly less than inferred from geodynamics (Ricard et al., 1993).

Similarly, the total amount of subducted material can give an estimate of the total extent of the plates (Neo- and Paleo-/Meso-Tethys) that have been subducted. Zones 1, 2 and 3 are

together some 3000 km in vertical extent. Assuming that this represents thickened (and formerly horizontal) plate material, the amount of thickening determines the former extent of the plates. A lower estimate of this amount derived from Figures 5.1 and 5.2 yields a factor of *at least* 2.5. This leads to a total plate extent of 7500 km, which is in good agreement with the upper estimate of the width the Neo-Tethys combined with a remnant of Paleo/Meso-Tethys.

#### 5.3.4 Conclusions

All high-velocity anomalies in the mantle below India can be interpreted in terms of fossil subductions. In the upper mantle, they have been explained as continental subductions of possibly delaminated sub-continental lithosphere, and in the lower mantle as different phases of Tethyan ocean subduction represented by three parallel zones of high velocity.

If these interpretations are correct, the total amount of subducted material suggests a fairly large Tethys (Neo-Tethys plus a remnant of Paleo/Meso-Tethys) ocean plate of some 7500 km. Part of this plate may have been subducted in an intra-oceanic setting which would explain the most southerly located zone of high velocities. Locally, the subducted material may presently reach the core-mantle boundary.

The observed difference in configuration of upper and lower mantle high-velocity anomalies supports the model of indentation and rotation of India after its collision with Tibet. In particular this model would explain the observed slab roll-over below Pakistan, the possible detachment below Nepal, and the formation of a cusp in the slab geometry below Burma.

### 5.4 Tomographic constraints on the dynamics of southeast Asia

In this section a combined summary of Rangin et al. (1999), and Spakman and Bijwaard (1998a) is presented. These papers are based on the results obtained in Chapter 4 (Bijwaard et al., 1998a), and describe and interpret several interesting features observed in the tomographic results for the southeast Asia region.

#### 5.4.1 Introduction

Southeast Asia is a region with complex geodynamics due to the convergence of three major tectonic plates: Eurasia in the North, the Pacific plate (together with the Philippine and Carolina plates) in the East, and the combined India-Australia plate in the South. Several tectonic reconstructions have been proposed for the development of the region (e.g. Rangin et al., 1990; Lee and Lawver, 1995; Hall, 1996) and several tomographic investigations have been performed to obtain 3-D constraints on this development (e.g. Fukao et al., 1992; Puspito et al., 1993; Sakurai, 1996; Widiyantoro and Van der Hilst, 1996, 1997). Among these constraints are the identification of large amounts of subducted material below the Sunda and Banda arcs, Sulawesi, the Philippines and New-Guinea, often extending to lower mantle depths. Detailed images obtained by Puspito et al. (1993) have also shown smaller scale structures such as the two-sided subduction of the Molucca Sea below Sulawesi and

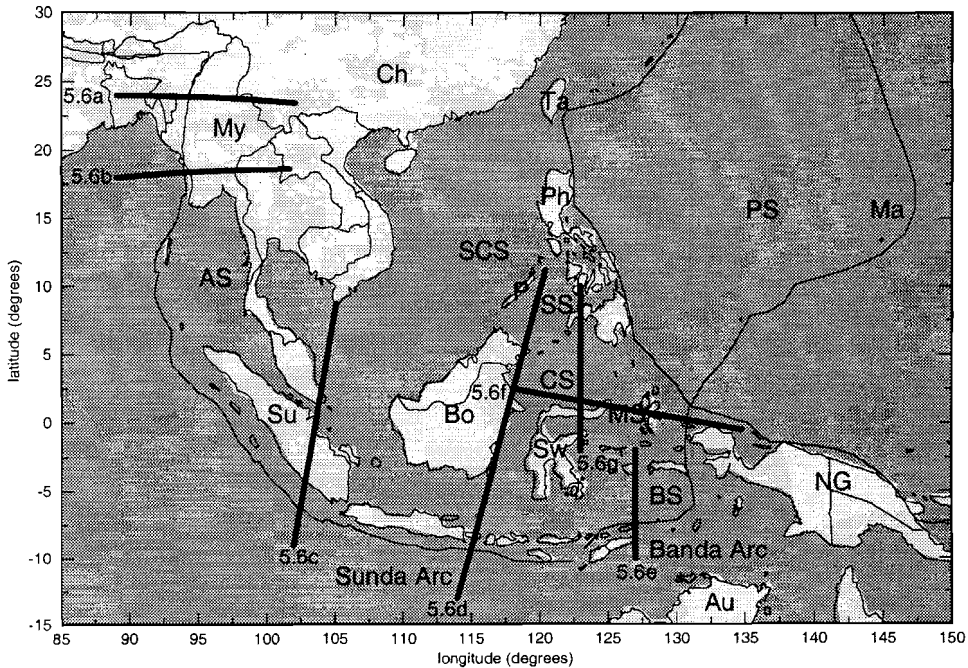
Halmahera. An interesting interpretation in the recent studies by Widiyantoro and Van der Hilst (1996, 1997), is the detachment of the subducting slab between 400 and 700 km depth below Sumatra and part of Java. In this section a description and interpretation is given of interesting structures observed in the results of Chapter 4 (Bijwaard et al., 1998a) for the southeast Asia region. These results show in several places a more detailed solution than has been obtained before.

#### *5.4.2 General description of the mantle below southeast Asia*

A map of southeast Asia and several layer and cross-sections through the model of Chapter 4 (Bijwaard et al., 1998a) are provided in Figures 5.4, 5.5, and 5.6 (for resolution aspects, the reader is referred to Chapter 4). In the upper mantle section at 200 km depth (Figure 5.5a), high velocity anomalies due to the different subductions can clearly be identified below Myanmar (Burma), the Sunda and Banda arcs, New-Guinea, the Molucca Sea, the Philippines, Taiwan, and the Marianas. This image is more detailed than obtained by Puspito et al. (1993) or Widiyantoro and Van der Hilst (1996, 1997). Interesting details are, for example, the gap in the Burma-Sunda subduction below the Andaman Sea (compare also Figures 5.6a and 5.6b), the kink in the curved Sunda arc subduction below West Sumatra, the strong curvature of the Banda arc subduction, and the small subduction patches below the Philippines. Going deeper into the mantle to 380 km depth (Figure 5.5b), the thin high velocity zones become wider, but the basic pattern remains unchanged, except for the Philippine and Sulawesi subductions which become connected, and the disappearance of subduction below the Northern Philippines (South China Sea subduction) and New-Guinea. At 560 km (Figure 5.5c), large high velocity zones have developed below the South China Sea and the Carolina Basin. Two major slabs (outlined by deep seismicity) have developed below central southeast Asia: One stretching from the Philippines to Sulawesi and the other from Sumatra to the Banda Sea. In the lower mantle (Figures 5.5d, 5.5e, and 5.5f) these slab-related anomalies merge into a single high-velocity zone. Although sensitivity estimates indicate that resolution below southeast Asia steadily decreases with depth from 100-200 km in most of the upper mantle to 300 km and larger in the lower mantle, this does not fully explain the broadening and merging of the high-velocity anomalies with depth.

#### *5.4.3 Detailed interpretations of anomalous structures*

Figure 5.6a shows a section across Myanmar (Burma) depicting an eastward dipping slab, down to 500 km depth. According to Lee and Lawver (1995) this subduction was initiated some 40 Myr ago during the collision of India with Tibet and its subsequent rotation (see previous section). Hence, an average convergence of 1.2 cm/yr can be deduced from the length of the slab. However, due to the continued northward motion of the India-Australia plate, this convergence must have been rather oblique. This may provide an explanation for the gap in subduction that can be observed below the Andaman Sea (Figure 5.6b). Although resolution for this area is rather limited, the lack of seismicity also indicates such a gap. The oblique convergence of India-Australia with the western part of the Sunda arc, may have lead to a large build-up of stresses in the subducting plate, causing a rupture and subsequent



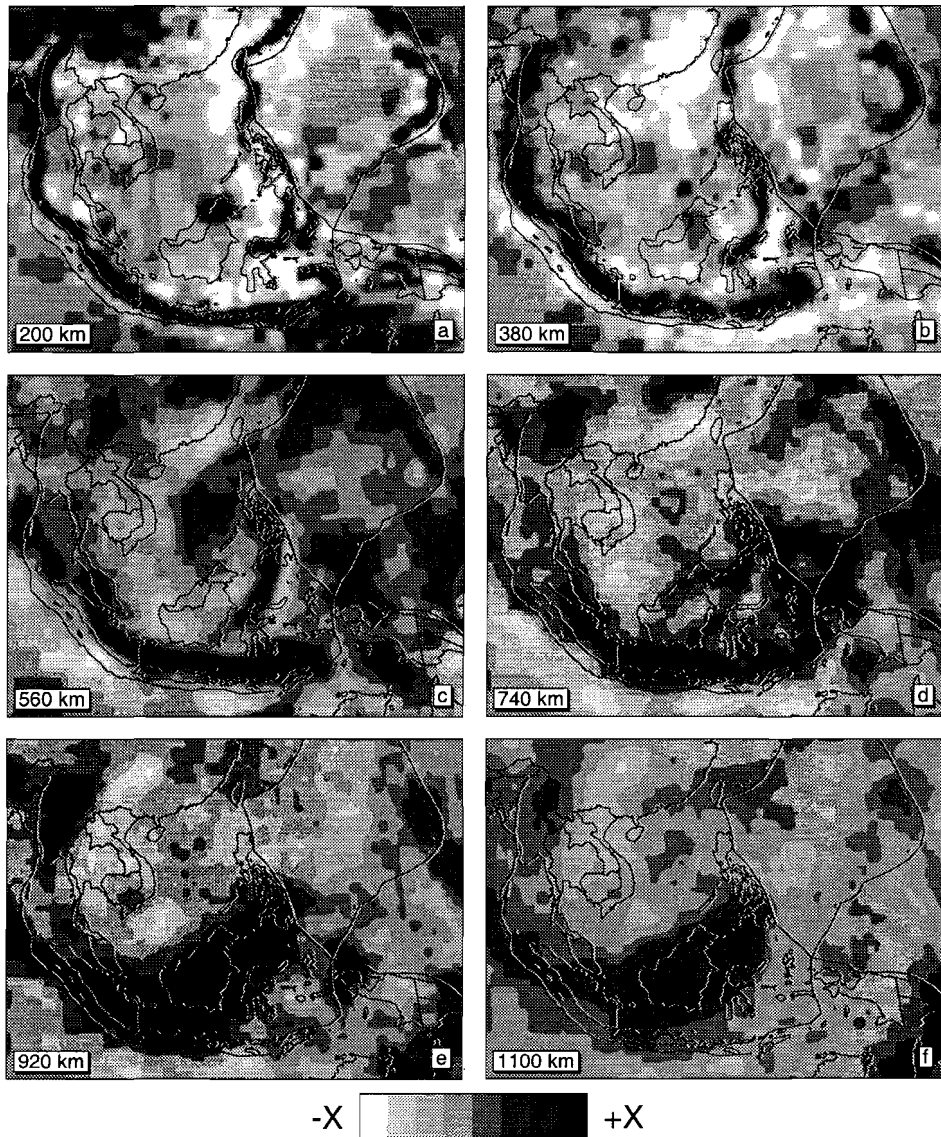
**Figure 5.4:** Geographic map of southeast Asia, indicating the locations of cross-sections in Figure 5.6, and the employed geographic terminology: AS is Andaman Sea, Au Australia, Bo Borneo, BS Banda Sea, Ch China, CS Celebes Sea, Ma Marianas, MS Molucca Sea, My Myanmar (Burma), NG New Guinea, P Palawan, Ph Philippines, PS Philippine Sea, SCS South China Sea, SS Sulu Sea, Sw Sulawesi, and Ta Taiwan.

northward displacement of part of the slab that is now observed below Myanmar. If (part of) Myanmar itself moved northward due to this displacement, this may explain the opening of the Andaman Sea.

Similarly, the kink in the curved Sunda arc subduction that is observed below western Sumatra (see Figures 5.5a, 5.5b, and 5.5c), which becomes more pronounced with depth (down to 560 km), may also have been caused by the oblique convergence. The sharpness of the kink suggests a failure in the slab that may be linked to a N-S oriented fracture zone that is present at 98 degrees East.

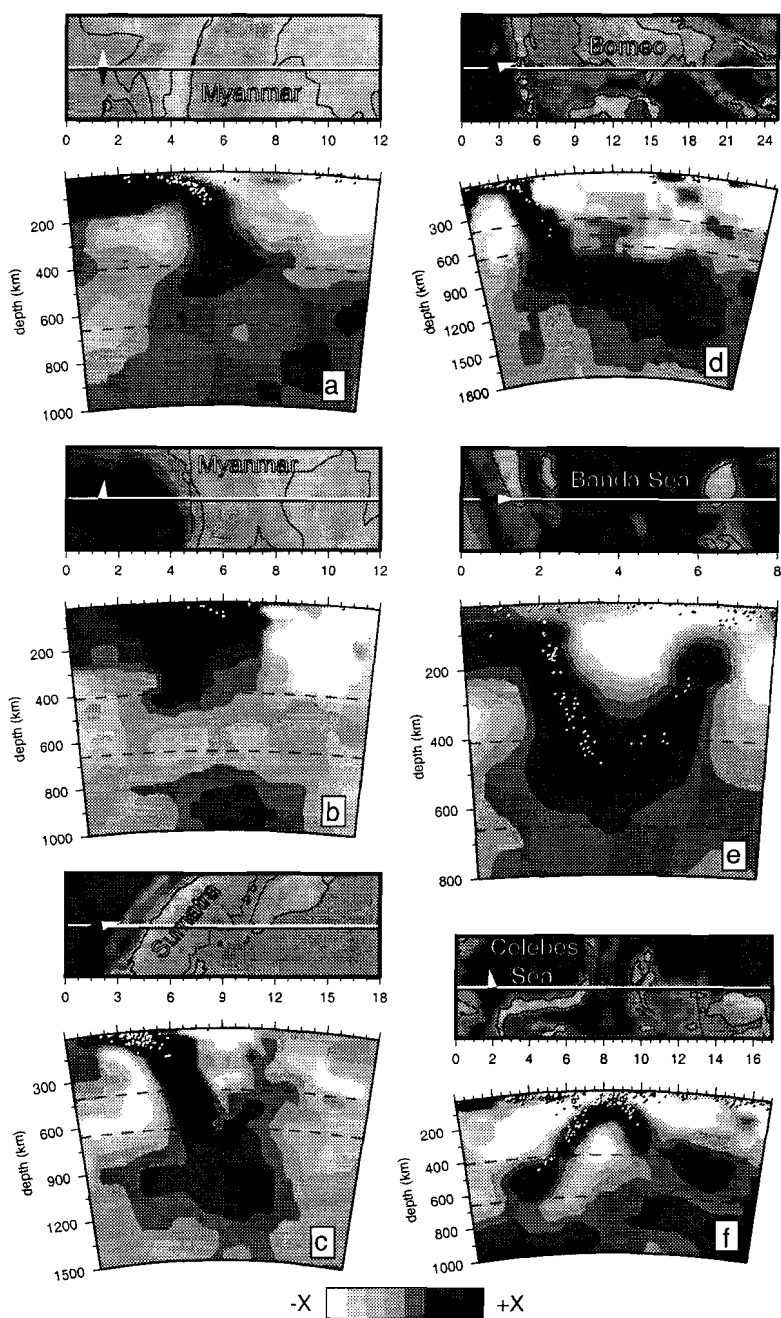
Contrary to Widiyantoro and Van der Hilst (1996, 1997), no indication of slab detachment is observed below Sumatra (see Figure 5.6c). Sensitivity tests for the vertical resolution below Sumatra indicate that if such a detachment existed between 400 and 700 km, it would probably have been imaged. The gap in the subducting slab observed by Widiyantoro and Van der Hilst (1996, 1997) may have been caused by horizontal smearing of low velocity anomalies that are present in their model on both sides of the subduction. The synthetic experiment they perform to verify the presence of this gap is in this respect inadequate, since their synthetic model also contains these low velocity anomalies.

The subduction below the Sunda and Banda arcs seems to be continuous from central Sumatra to the Banda Sea (see also Figure 5.6d), where a very strong curvature of the high-



**Figure 5.5:** Layer sections for southeast Asia through the model of Chapter 4 (Bijwaard et al., 1998a) at six depths: (a) 200 km ( $X=2\%$ ), (b) 380 km ( $X=1.5\%$ ), (c) 560 km ( $X=1.5\%$ ), (d) 740 km ( $X=1\%$ ), (e) 920 km ( $X=1\%$ ), and (f) 1100 km ( $X=1\%$ ). Velocity anomalies are displayed in percentages with respect to reference model ak135 (Kennett et al., 1995).

velocity anomaly is observed (Figures 5.5a and 5.5b). The origin of this curvature is enigmatic, but it may consist of two, opposing subductions (see Figure 5.6e) that are pushed together by the shape of the Australian continental margin (Cardwell and Isacks, 1978). This



**Figure 5.6:** Cross-sections for southeast Asia through the model of Chapter 4 (Bijwaard et al., 1998a). Locations are shown in Figure 5.4, top panels show detailed bathymetry and a compass needle pointing north. For panels (a) and (e)  $X=2\%$ ; for panels (b) to (d)  $X=1\%$ , and for panels (f) and (g)  $X=1.5\%$ .



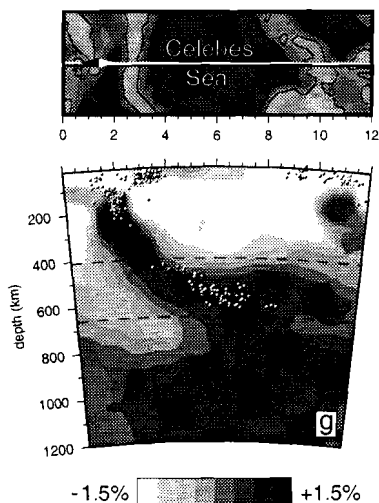


Figure 5.6: Continued.

would explain why the curvature is observed down to large upper mantle depths (see also Puspito et al., 1993; Widiyantoro and Van der Hilst, 1996, 1997). This curvature may have been accentuated by the rapid extension of the South Banda Basin due to the arrival of Timor (Australia).

The relatively simple geometry of the subductions on the western and southern sides of southeast Asia are contrasted by complex subductions on the eastern side beneath the Philippines, Sulawesi, and Halmahera (see Figure 5.5a). At 200 km depth the South China Sea slab subducts eastward below the Philippine Sea plate from Taiwan to South Luzon (northern Philippines), and the Philippine Sea plate subducts westward under the central Philippines. Below the southern Philippines (Mindanao) the Sangihe slab subducts to the west, and this subduction is continuous to Sulawesi where it opposes eastward Molucca Sea subduction below Halmahera (Figure 5.6f). This pattern of small-scale subduction patches changes at 380 km depth: The South China Sea subduction starts to disappear, and the subductions below the Philippines form a single slab that continues to the bottom of the upper mantle and is outlined by the seismicity. The large extent of this slab below Sulawesi (Figure 5.6g) suggests that it is the Sangihe slab, and not the very recent Celebes Sea subduction that is observed here. The configurations at 380 and 560 km depth suggest that the Sangihe slab represents the major subduction zone before the arrival of the Banda block (15 Myr ago). The total length of the Sangihe slab together with that due to the subduction below Halmahera (Figure 5.6f) argue for a large Molucca Sea, which may have formed a connection between the Indian and Pacific oceans.

The interpretation of the high-velocity anomaly in the deep upper mantle below the Philippines and Sulawesi as the Sangihe slab, and the formerly major subduction zone on the eastern side of southeast Asia, allows for the interpretation of the Celebes, Sulu, and South China Seas as the result of back-arc extension (Rangin et al., 1990). It has been

suggested that during the opening of the South China Sea (32-15 Myr ago), the Proto South China Sea was subducted (Holloway, 1982; Taylor and Hayes, 1983). The existence of the Proto South China Sea can be inferred from an ophiolitic suture running across Palawan and Borneo (Hutchison, 1989). Its subduction may explain the large high-velocity anomaly that is observed in the transition zone below the South China Sea, although one might have expected to find this anomaly further south.

In the upper part of the lower mantle, the major subductions of the Sunda and Sangihe slabs seem to merge into a single, curved high-velocity anomaly. As mentioned above, this is not (entirely) due to the steady decrease in resolution with depth; it is also caused by a (resolved) thickening of the slab-related anomalies. It may well be that this observed thickening of the slabs with depth is caused by the increase in viscosity and especially the viscosity jump at the 660 km discontinuity.

#### 5.4.4 Conclusions

The global model that has been derived in Chapter 4 (Bijwaard et al., 1998a), shows well-resolved small-scale details in the mantle below southeast Asia that have not been observed before. Observations include detailed structure of all (long-lived) subductions in this region and allow for new insights in the regional tectonic development. This has resulted in the following interpretations: (1) The potential gap between the Myanmar (Burma) and Sunda subductions below the Andaman Sea, together with the observed kink in the slab below West Sumatra, may result from oblique subduction. This could have led to the rupture and displacement of slab material northward. (2) The complex pattern of subductions on the eastern side of southeast Asia has created many small patches of subducted material in the uppermost mantle. These patches combine into a single major subduction of the Sangihe slab below 300 km depth. This suggests that the Celebes, Sulu, and South China Seas have opened in a back-arc setting. (3) The total length of subductions of the Molucca Sea argue for a formerly large basin that may have connected the Indian and Pacific oceans. (4) The two main subductions of the Sunda and Sangihe slabs combine into a rather well-resolved reservoir of high velocities in the lower mantle.

#### References

- Albers, M., and U.R. Christensen, The excess temperature of plumes rising from the core-mantle boundary, *Geophys. Res. Lett.*, 23, 3567-3570, 1996.
- Allègre, C., *The behavior of the Earth*, 272 pp., Harvard Univ. Press, 1988.
- Beck, R.A., D.W. Burbank, W.J. Sercombe, A.M. Khan, and R.D. Lawrence, Late Cretaceous ophiolite obduction and Paleocene India-Asia collision in the westernmost Himalaya, *Geodinamica Acta*, 9, 114-144, 1996.
- Besse, J., and V. Courtillot, Paleogeographic maps of the continents bordering the Indian ocean since the Early Jurassic, *J. Geophys. Res.*, 93, 11791-11808, 1988.
- Besse, J., V. Courtillot, J.P. Pozzi, M. Westphal, and Y.-X. Zhou, Paleomagnetic estimates of crustal shortening in the Himalayan thrusts and Zangbo suture, *Nature*, 311, 621-626, 1984.

- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998a.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Preliminary results from nonlinear global travel-time tomography (abstract), *Ann. Geophys.*, 16, suppl. 1, 33, 1998b.
- Brandon, A.D., R.J. Walker, J.W. Morgan, M.D. Norman, and H.M. Prichard, Coupled  $^{186}\text{Os}$  and  $^{187}\text{Os}$  evidence for core-mantle interaction, *Science*, 280, 1570-1573, 1998.
- Cande, S.C., C.A. Raymond, J. Stock, and W.F. Haxby, Geophysics of the Pitman Fracture Zone and Pacific-Antarctic plate motions during the Cenozoic, *Science*, 270, 947-953, 1995.
- Cardwell, R.K., and B.L. Isacks, Geometry of the subducted lithosphere beneath the Banda Sea in eastern Indonesia from seismicity and fault plane patterns, *J. Geophys. Res.*, 83, 2825-2838, 1978.
- Crough, S.T., Hotspot swells, *Annu. Rev. Earth Planet. Sci.*, 11, 165-193, 1983.
- Curtis, C., B. Dost, J. Trampert, and R.K. Snieder, Eurasian fundamental mode surface wave velocities and their relationship with tectonic structures, *J. Geophys. Res.*, 103, 26919-26947, 1998.
- Davies, G.F., Plates, plumes mantle, convection, and mantle evolution, in: *The Earth's mantle, composition structure and evolution*, edited by I. Jackson, pp. 228-258, Cambridge university press, 1998.
- De Jonge, M.R., M.J.R. Wortel, and W. Spakman, Regional scale tectonic evolution and the seismic velocity structure of the lithosphere and upper mantle: the Mediterranean region, *J. Geophys. Res.*, 99, 12091-12108, 1994.
- Dercourt, J., L.E. Ricou, and B. Vrielynck, *Atlas of Tethys Palaeoenvironmental Maps*, 14 maps, 1 pl., Gauthier-Villars, Paris, 307 pp., 1993.
- Dewey, J.F., R.M. Shackleton, C. Chang, and Y. Sun, The tectonic evolution of the Tibetan Plateau, *Phil. Trans. R. Soc. London A*, 327, 379-413, 1988.
- Duncan, R.A., and M.A. Richards, Hotspots, mantle plumes, flood basalts, and true polar wander, *Rev. Geophys.*, 29, 31-50, 1991.
- Engdahl, E.R., R.D. van der Hilst, and R.P. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Amer.*, 88, 722-743, 1998.
- Enkin, R., Z.Y. Yang, Y. Chen, and V. Courtillot, Paleomagnetic constraints on the geodynamic history of the major blocks of China from the Permian to the Present, *J. Geophys. Res.*, 97, 13953-13989, 1992.
- Fitton, J.G., A.D. Saunders, M.J. Norry, B.S. Hardarson, and R.N. Taylor, Thermal and chemical structure of the Iceland plume, *Earth Planet. Sci. Lett.*, 153, 197-208, 1997.
- Fukao, Y., M. Obayashi, H. Inoue, and M. Nenbai, Subducting slabs stagnant in the mantle transition zone, *J. Geophys. Res.*, 97, 4809-4822, 1992.
- Grand, S.P., R.D. van der Hilst, and S. Widiyantoro, Global seismic tomography: a snapshot of convection in the Earth, *GSA Today*, 7, 4, 1-7, 1997.
- Griffiths, R.W., and I.H. Campbell, Stirring and structure in mantle plumes, *Earth Planet. Sci. Lett.*, 99, 66-78, 1990.
- Griffiths, R.W., and I.H. Campbell, The interaction of plumes with the Earth's surface and onset of small-scale convection, *J. Geophys. Res.*, 96, 18295-18310, 1991.
- Gurnis, M., Large-scale mantle convection and the aggregation and dispersion of

- supercontinents, *Nature*, 232, 695-699, 1988.
- Hall, H., *Reconstructing Cenozoic SE Asia*, in: *Tectonic evolution of southeast Asia*, Geological Society Special Publication No. 106, edited by R. Hall and D. Blundell, pp. 153-184, 1996.
- Holloway, N.H., *The North Palawan block, Philippines: its relation to the Asian mainland and its role in the evolution of the South China Sea*, *Am. Assoc. Petrol. Geol. Bull.*, 66, 1355-1383, 1982.
- Honegger, K., P. Le Fort, G. Mascle, and J.L. Zimmermann, *The blueschists along the Indus Suture Zone in Ladakh, NW Himalaya*, *J. Metamorphic Geol.*, 7, 57-72, 1989.
- Hutchison, C.S., *Geological evolution of South-East Asia*, Oxford, 1989.
- Ito, G., J. Lin, and C.W. Gable, *Dynamics of mantle flow and melting at a ridge-centered hotspot: Iceland and the Mid-Atlantic Ridge*, *Earth Planet. Sci. Lett.*, 144, 53-74, 1996.
- Ji, Y., and H.-C. Nataf, *Detection of mantle plumes in the lower mantle by diffraction tomography: Hawaii*, *Earth Planet. Sci. Lett.*, 159, 99-115, 1998.
- Karato, S.-I., *Importance of anelasticity in the interpretation of seismic tomography*, *Geophys. Res. Lett.*, 20, 1623-1626, 1993.
- Kennett, B.L.N., and S. Widiyantoro, *A low seismic wavespeed anomaly beneath northwestern India - a seismic signature of the Deccan traps*, *Earth Planet. Sci. Lett.*, 165, 145-156, 1999.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, *Constraints on seismic velocities in the Earth from traveltimes*, *Geophys. J. Int.*, 122, 108-124, 1995.
- Klein, E.M., and C.H. Langmuir, *Global correlations of ocean ridge basalt chemistry with axial depth and crustal thickness*, *J. Geophys. Res.*, 92, 8089-8115, 1987.
- Klootwijk, C.T., P.J. Conaghan, and C.M. Powell, *The Himalayan arc: Large-scale continental subduction, oroclinal bending and back-arc spreading*, *Earth Planet. Sci. Lett.*, 75, 167-183, 1985.
- Klootwijk, C.T., J.S. Gee, H.W. Peirce, G.M. Smith, and P.L. McFadden, *An early India-Asia contact: Paleomagnetic constraints from Ninetyeast Ridge*, *ODP Leg 121, Geology*, 20, 395-398, 1992.
- Lee, T.-Y., and L.A. Lawver, *Cenozoic plate reconstruction of Southeast Asia*, *Tectonophysics*, 251, 85-138, 1995.
- Lithgow-Bertelloni, C., and M.A. Richards, *The dynamics of Cenozoic and Mesozoic plate motions*, *Rev. Geophys.*, 36, 27-78, 1998.
- Loper, D.E., and F.D. Stacey, *The dynamical and thermal structure of deep mantle plumes*, *Phys. Earth Planet. Inter.*, 33, 304-317, 1983.
- Mattauer, M., Ph. Matte, J.-L. Olivet, *A 3D model of the India-Asia collision at plate scale*, *C.R. Acad. Sci. Paris (Earth Planet. Sci.)*, 328, 499-508, 1999.
- Matte, Ph., M. Mattauer, J.-M. Olivet, and D.A. Griot, *Continental subduction beneath Tibet and the Himalayan orogeny: A review*, *Terra Nova*, 9, 264-270, 1997.
- Mohan, G., and S.S. Rai, *Large-scale three-dimensional seismic tomography of the Zagros and Pamir-Hindukush regions*, *Tectonophysics*, 242, 255-265, 1995.
- Morgan, W.J., *Convection plumes in the lower mantle*, *Nature*, 230, 42-43, 1971.
- Nataf, H.-C., and J. VanDecar, *Seismological detection of a mantle plume?*, *Nature*, 364, 115-120, 1993.
- Nelson, K.O., W. Zhao, L.D. Brown, *Partially molten middle crust beneath southern Tibet:*

- Synthesis of project INDEPTH results, *Science*, 274, 1684-1687, 1996.
- Parfenov, L.M., Tectonics of the Verkhoyansk-Kolyma Mesozoides in the context of plate tectonics, *Tectonophysics*, 199, 319-342, 1991.
- Patriat, Ph., and J. Achache, India-Eurasia collision chronology has implications for crustal shortening and driving mechanism of plates, *Nature*, 311, 615-621, 1984.
- Patzelt, A., H. Li, J. Wang, and E. Appel, Paleomagnetism of Cretaceous to Tertiary sediments from Southern Tibet: Evidence for the extent of the northern margin of India prior to the collision with Eurasia, *Tectonophysics*, 259, 259-284, 1996.
- Pegler, G., and S. Das, An enhanced image of the Pamir-Hindu Kush seismic zone from relocated earthquake hypocenters, *Geophys. J. Int.*, 134, 573-595, 1998.
- Pozzi, J.P., M. Westphal, J. Girardeau, J.P. Besse, and Y.-X. Zhou, Paleomagnetism of the Xigaze ophiolite and flysch: Latitude and direction of spreading, *Earth Planet. Sci. Lett.*, 70, 383-394, 1984.
- Pritchard, M.J., B.R. Julian, J. Fyen, and G.R. Foulger, Evidence for a deep low-velocity feature beneath Iceland from *P*-wave azimuth and slowness anomalies, *EOS Trans. AGU*, 77, 480, 1996.
- Puspito, N.T., Y. Yamanaka, T. Miyataka, K. Shimazaki, and K. Hirahara, Three-dimensional *P*-wave velocity structure beneath the Indonesian region, *Tectonophysics*, 220, 175-192, 1993.
- Rangin, C., L. Jolivet, and M. Pubellier, A simple model for the tectonic evolution of Southeast Asia and Indonesia region for the past 43 m.y., *Bull. Soc. Géol. France*, 8, 907-913, 1990.
- Rangin, C., W. Spakman, M. Pubellier, and H. Bijwaard, Tomographic and geological constraints on subduction along the eastern Sundaland continental margin (South-East Asia), *Soc. Geol. Fr.* (submitted), 1999.
- Reuber, I, Geometry of accretion and oceanic thrusting of the Spontang Ophiolite, Ladakh-Himalaya, *Nature*, 321, 592-596, 1986.
- Ribe, N., U.R. Christensen, and J. Theissing, The dynamics of plume-ridge interaction, 1: Ridge-centered plumes, *Earth Planet. Sci. Lett.*, 134, 155-168, 1995.
- Ricard, Y., M.A. Richards, C. Lithgow-Bertelloni, and Y. Le Stunff, A geodynamic model of mantle density heterogeneity, *J. Geophys. Res.*, 98, 21895-21909, 1993.
- Ritzwoller, M.H., and A.L. Levshin, Eurasian surface wave tomography: Group velocities, *J. Geophys. Res.*, 103, 4839-4878, 1998.
- Roecker, S.W., Velocity structure of the Pamir-Hindu Kush region: Possible evidence of subducted crust, *J. Geophys. Res.*, 87, 945-959, 1982.
- Rowley, D.B., Age of initiation of collision between India and Asia: A review of stratigraphic data, *Earth Planet. Sci. Lett.*, 145, 1-13, 1996.
- Sakurai, T., Whole mantle *P* wave tomography and differential *PP-P* time measurement, M.Sc. thesis, University of Tokyo, 1996.
- Schilling, J.-G., Iceland mantle plume: Geochemical study of Reykjanes Ridge, *Nature*, 242, 565-571, 1973.
- Schilling, J.-G., Fluxes and excess temperatures of mantle plumes inferred from their interaction with migrating mid-ocean ridges, *Nature*, 352, 397-403, 1991.
- Scotese, C.R., and J. Golonka, Paleomap Paleogeographic Atlas, Paleomap project, Univ. Texas, Arlington, 1992.
- Searle, M.P., R.I. Corfield, B. Stephenson, and J. McCarron, Structure of the north Indian

- continental margin in the Ladakh-Zaskar Himalayas: Implications for the timing of obduction of the Spontang ophiolite, India-Asia collision and deformation events in the Himalaya, *Geol. Magazine*, 134, 297-316, 1997.
- Şengör, A.M.C., The Cimmeride orogenic system and the tectonics of Eurasia, *Geol. Soc. America Spec. Paper* 195, 88 pp., 1984.
- Şengör, A.M.C., and B.A. Natal'in, Paleotectonics of Asia: fragments of a synthesis, in: *The Tectonic Evolution of Asia*, edited by A. Yin and M. Harrison, pp. 486-640, Cambridge Univ. Press, 1996.
- Shen, Y., S.C. Solomon, I.Th. Bjarnason, and C.J. Wolfe, Seismic evidence for a lower-mantle origin of the Iceland plume, *Nature*, 395, 62-65, 1998.
- Skilbeck, J.N., and J.A. Whitehead, Formation of discrete islands in linear island chains, *Nature*, 272, 499-501, 1978.
- Spakman, W., and H. Bijwaard, Mantle structure and large scale dynamics of South-East Asia, in: *The GEODYNAMICS of S and SE Asia (GEODYSSSEA) Project*, Scientific Technical Report STR98/14, edited by P. Wilson and G.W. Michel, GeoForschungsZentrum Potsdam, pp. 313-339, 1998a.
- Spakman W., and H. Bijwaard, Irregular cell parameterization of tomographic problems (abstract), *Ann. Geophys.*, 16, suppl. 1, 28, 1998b.
- Spakman W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, 155-187, Reidel, Dordrecht, the Netherlands, 1988.
- Steinbach, V., and D.A. Yuen, Dynamical effects of a temperature- and pressure-dependent lower-mantle rheology on the interaction of upwellings with the transition zone, *Phys. Earth Planet. Inter.*, 103, 85-100, 1997.
- Steinberger, B., Plumes in a convecting mantle: Models and observations for individual hotspots, *J. Geophys. Res.* (submitted), 1998.
- Steinberger, B., and R.J. O'Connell, Advection of plumes in mantle flow: implications for hotspot motion, mantle viscosity and plume distribution, *Geophys. J. Int.*, 132, 412-434, 1998.
- Su, W., R.L. Woodward, and A.M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, 99, 6945-6980, 1994.
- Tarduno, J.A., and R.D. Cottrell, Paleomagnetic evidence for motion of the Hawaiian hotspot during formation of the Emperor Seamounts, *Earth Planet. Sci. Lett.*, 153, 171-180, 1997.
- Taylor, B., and D.E. Hayes, Origin and history of the South China Sea basin, in: *The tectonic evolution of southeast Asian seas and islands*, Am. Geophys. Union Geophysical Monograph 2, edited by D.E. Hayes, pp. 23-56, 1983.
- Thompson, P.F., and P.J. Tackley, Generation of mega-plumes from the core-mantle boundary in a compressible mantle with temperature-dependent viscosity, *Geophys. Res. Lett.*, 25, 1999-2002, 1998.
- Van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl, Evidence for deep-mantle circulation from global tomography, *Nature*, 386, 578-584, 1997.
- Van der Voo, R., Paleomagnetism of the Atlantic, Tethys and Iapetus Ocean, 411 pp., Cambridge Univ Press, Cambridge, U.K., 1993.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Mesozoic subducted slabs under Siberia, *Nature*, 397, 246-249, 1999a.

- Van der Voo, R., W. Spakman, and H. Bijwaard, Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.* (in press), 1999b.
- Vasco, D.W., and L.R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, 103, 2633-2671, 1998.
- Veevers, J.J., C.M. Powell, and B.D. Johnson, Greater India's place in Gondwanaland and in Asia, *Earth Planet. Sci. Lett.*, 27, 383-387, 1975.
- Westphal, M., J.P. Pozzi, Y.-X. Zhou, L.S. Xing, and X.Y. Chen, Paleomagnetic data about southern Tibet (Xizang). 1. The Cretaceous formations of the Lhasa block, *Geophys. J. R. Astron. Soc.*, 73, 507-521, 1983.
- White, R., and D. McKenzie, Magmatism at rift zones: The generation of volcanic continental margins and flood basalts, *J. Geophys. Res.*, 94, 7685-7729, 1989.
- Whitehead, J.A., and D.S. Luther, Dynamics of laboratory diapir and plume models, *J. Geophys. Res.*, 80, 705-717, 1975.
- Widiyantoro, S., and R.D. Van der Hilst, Structure and evolution of lithosphere slab beneath the Sunda arc, *Science*, 271, 1566-1570, 1996.
- Widiyantoro, S., and R.D. Van der Hilst, Mantle structure beneath Indonesia inferred from high-resolution tomographic imaging, *Geophys. J. Int.*, 130, 167-182, 1997.
- Wilson, J.T., A possible origin of the Hawaiian Islands, *Can. J. Phys.*, 41, 863-868, 1963.
- Wolfe, C.J., I.Th. Bjarnason, J.C. VanDecar, and S.C. Solomon, Seismic structure of the Iceland mantle plume, *Nature*, 385, 245-247, 1997.
- Wyssession, M.E., Imaging cold rock at the base of the mantle; the sometimes fate of slabs?, in: *Subduction*, edited by G.E. Bebout, D.W. Scholl, S.H. Kirby, and J.P. Platt, pp. 369-384, *Am. Geophys. Union Monogr.* 96, Washington DC, 1996.
- Zhang, Y.-S., and T. Tanimoto, High-resolution global upper mantle structure and plate tectonics, *J. Geophys. Res.*, 98, 9793-9823, 1993.
- Zhao, W.-L., and W.J. Morgan, Injection of Indian crust into Tibetan lower crust: A two-dimensional finite element model study, *Tectonics*, 6, 489-504, 1987.
- Zhou, H., A high resolution *P* wave model for the top 1200 km of the mantle, *J. Geophys. Res.*, 101, 27791-27810, 1996.

# Chapter 6<sup>1</sup>

## Advancing global *S*-wave tomography toward the regional scale

**Summary.** We have used an irregular parameterization technique and an updated collection of nearly 1 million *S* phases to obtain a high-resolution model of *S*-velocity heterogeneity for the Earth's lithosphere and mantle. Although *S* data are generally less accurate and less abundant than *P* data, sensitivity tests indicate resolution on scales as small as  $0.6^\circ$  in some very well-sampled regions in the shallowest mantle. The most prominent features of the solution include thin high-velocity anomalies associated with subduction zones in the upper mantle, often continuing into the lower mantle and in several cases connected with large high velocity anomalies associated with past subduction. In this and other respects the *S* model is very similar to the *P* model of Bijwaard et al. (1998) (Chapter 4) with correlation coefficients reaching 0.85 for the best sampled cells. The amplitudes of the *S* anomalies are, however, up to twice as large in the lower mantle. Detailed comparisons show that there are very few discrepancies between *P* and *S* anomalies in the well-resolved parts of the mantle, but notable exceptions occur in the upper mantle below the western Mediterranean and the Timor Sea, in the lower mantle below Europe and the Coral Sea, and in the D'' layer.

### 6.1 Introduction

Recent global models of *S*-velocity heterogeneity (Masters et al., 1996; Grand et al., 1997; Widiyantoro et al., 1998; Vasco and Johnson, 1998) have not provided images as detailed as recent *P*-wave tomography studies (Zhou, 1996; Bijwaard et al., 1998). This has several reasons: First, *S* phases arrive in the coda of *P* phases and thus cannot be picked as accurately. Second, they are usually of longer wavelength and thus have less resolving power. Third, the total number of observed *S* phases is far below the number of *P* phases leading to a poorer sampling of the mantle. Finally, due to their limited coverage *S* phases are often supplemented by other data (*SS* phases, surface waves and normal modes) that are usually less sensitive to small-scale structure. Consequently, the parameterizations applied in recent global models of *S* heterogeneity are coarser than in recent models for *P* heterogeneity, ranging from  $2^\circ$  cells (Widiyantoro et al., 1998) up to degree 16 spherical harmonics (Masters et al., 1996), which is comparable to  $11^\circ$  cells. In addition, the estimates of spatial resolution range, when provided, from scales of  $10^\circ$  (Widiyantoro et al., 1998) to  $20^\circ$ - $25^\circ$  (Masters et al., 1996; Vasco and Johnson, 1998).

With the shortcomings of *S* data in mind, one cannot expect to derive an *S* model with a similar resolution as for *P*. Locally, however, it may be possible to image small-scale detail in upper mantle areas densely sampled by *S* phases. This would allow a detailed comparison between *P* and *S* heterogeneity in, for example, subduction zones. To construct such a high-

---

<sup>1</sup> This chapter has been submitted to *Geophys. J. Int.* as: Bijwaard, H., and W. Spakman, Advancing global *S*-wave tomography toward the regional scale.



resolution model of  $S$ -velocity heterogeneity, we apply in this paper the irregular parameterization technique of Spakman and Bijwaard (1998), which allows for the implementation of small cells (down to  $0.6^\circ$ ) in well-sampled model parts, in combination with the  $S$  phases from the data set of Engdahl et al. (1998).

## 6.2 Data, model parameterization and inversion

### 6.2.1 Data

We employ the data set of Engdahl et al. (1998), which is a reprocessed version of a combined ISC/NEIC data set supplemented with data from several regional deployments of seismic stations. The entire data set consists of more than 82000 well-constrained earthquakes and some 12 million first and later arriving seismic phases. From this data set we select regional  $S$  phases with residuals between -12 and 12 s, and teleseismic  $S$  phases with residuals between -9 and 9 s. These boundaries are chosen rather large to include nearly all  $S$  phases with the exception of very large residuals which we regard as outliers. The total amount of selected  $S$  phases is 953240 or 98.7 % of the total number of  $S$  phases in the original data set of Engdahl et al. (1998). These single  $S$  rays are combined into 697925 composite rays (Spakman and Nolet, 1988) from event cluster volumes of  $30 \times 30 \times 30 \text{ km}^3$  to single stations. The variance reduction obtained from the clustering of single rays amounts to 13.6%.

### 6.2.2 Parameterization

From the uneven ray sampling, an irregular model parameterization is derived in a similar fashion as in Bijwaard et al. (1998) (Chapter 4). Based upon hitcount, small cells are implemented in regions with a dense data coverage and large cells in regions with poor sampling (Spakman and Bijwaard, 1998). The irregular cell grid is constructed in the following way: First, regular cell grids and corresponding hitcount maps are determined for every cell size that is to be implemented. These cell sizes are all constructed from an integer number of the smallest (basic,  $0.6^\circ$ ) cells. Second, throughout the model volume the largest cells ( $9^\circ$ ) are implemented. Third, in an iterative process larger cells are replaced by smaller ones when the hitcount for the larger cell is above a certain threshold value. This leads to a mantle grid of 274981 non-overlapping, irregular cells varying in size from  $0.6^\circ \times 0.6^\circ \times 35 \text{ km}^3$  to  $9^\circ \times 9^\circ \times 400 \text{ km}^3$ . Lower mantle cells are restricted to dimensions larger than 100 km. With the exception of extremely poorly sampled regions (mainly the shallow mantle below the oceans), all cells are traversed by more than 50 rays. However, in very high hitcount regions the smallest ( $0.6^\circ$ ) cells may still contain several thousands of rays.

### 6.2.3 Inversion

The nearly 700000 composite residuals lead, together with 400000 unknowns (275000 cell slownesses,  $4 \times 30000$  event clusters relocation parameters, and 3000 station parameters), to

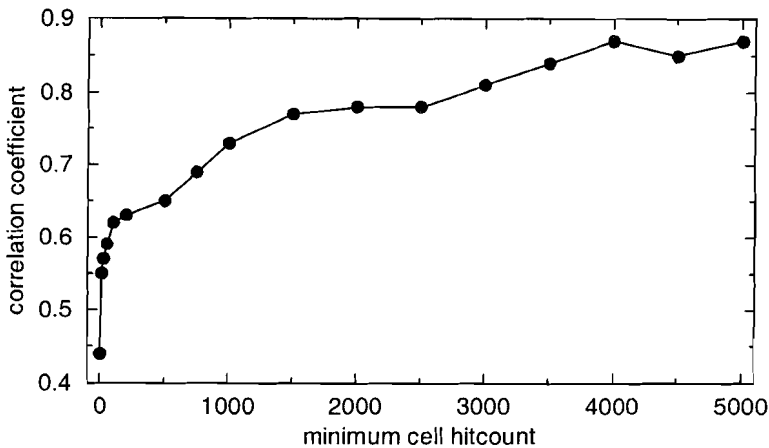
a large matrix-vector equation (identical to Equation (2) of Bijwaard et al., 1998 and Chapter 4) for which an approximate solution is calculated in 200 iterations of the least squares solver LSQR (Paige and Saunders, 1982). The inversion is regularized by penalizing the second derivative of the solution (with respect to the background velocity model ak135 (Kennett et al., 1995)). Due to poorer coverage and larger errors in the *S* data, the relative weight of the damping equations was larger than for the *P* inversion of Bijwaard et al. (1998) (Chapter 4). Contrary to the *P* inversion, for which *pP* and *pwP* phases were used, we have no depth phases for the *S* inversion to constrain the focal depths of the event clusters. Furthermore, the number of *S* arrivals per event is not always sufficient to constrain the hypocenters. To prevent large deviations from the accurate locations derived by Engdahl et al. (1998), and simultaneously allow for the absorption of event-specific signal, we apply a scaling to the relocation parameters such that these can attain only small values.

### 6.3 Results and resolution

#### 6.3.1 Overview of the results

In the inversion we obtain a variance reduction of 60.4%, leading to a total data variance reduction (including the reduction obtained by ray clustering) of 65.7%, which is comparable to the 64.3% obtained by Bijwaard et al. (1998) (Chapter 4) for their *P* model, but which is larger than obtained in previous studies of *S* heterogeneity. Widiyantoro et al. (1998), for example, obtained 33% variance reduction from the inversion of a very similar data set and a similar amount of parameters. In their study the regular parameterization introduces many underdetermined model parameters in poorly sampled regions, whereas in our irregular approach the small ( $0.6^\circ$ ,  $1.2^\circ$ ) cells are concentrated in the well-sampled areas and may thus contribute to the fitting of the residuals.

Our solution to the inverse problem consists of station corrections, event cluster relocations and cell slownesses. As for the *P* inversion of Bijwaard et al. (1998) (Chapter 4), the calculated station corrections mainly reflect average station residuals, whereas the average event residual is mainly absorbed by an origin time correction. Several layers of the resulting *S*-velocity anomalies are shown in Figure A8 and these can be compared with the *P* solution shown in Plate 1 of Bijwaard et al. (1998) (or Figure A1). Before considering the resolution of the *S* model, we note that for regions that are sampled adequately by both *P* and *S* phases the models are very similar in structure down to rather small scales in subduction zones. This similarity between *P* and *S* models has been observed before in models aimed at larger scale structure. Correlations between global *P* and *S* models have ranged in these models from 0.4 (Vasco and Johnson, 1998) to 0.6 (Robertson and Woodhouse, 1996; Su and Dziewonski, 1993). The correlation between our *S* model and the *P* model of Bijwaard et al. (1998) (Chapter 4) is shown in Figure 6.1, and increases rapidly from 0.45 for all  $0.6^\circ$  basic cells (including large parts of the mantle that are not sampled by seismic rays) in the models to 0.85 for the best sampled cells. Both numbers are highly significant at the 95% level, but more importantly Figure 6.1 shows that *P* and *S* heterogeneity become more and more similar when the ray sampling increases. Because large cells sampled by more than 50 rays have been replaced in the irregular parameterization by smaller cells down to the  $0.6^\circ$



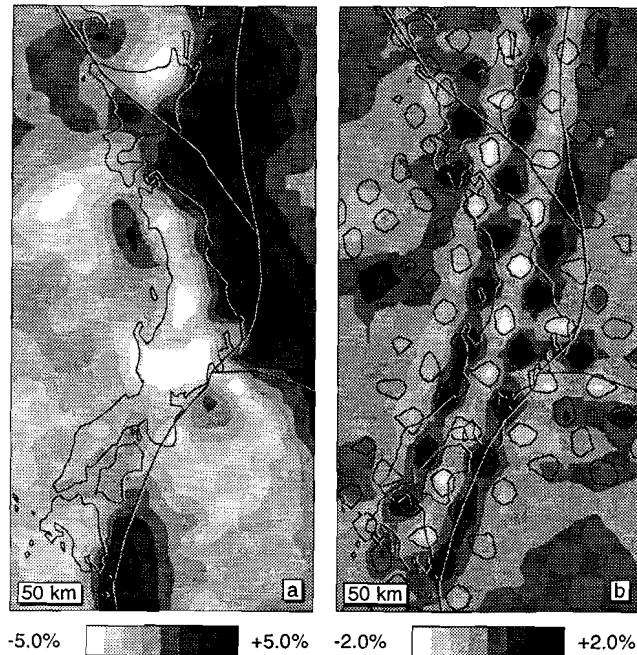
**Figure 6.1:** Linear correlation coefficient between the  $P$  model of Bijwaard et al. (1998) (Chapter 4) and the presented  $S$  model for regular  $0.6^\circ$  cells with increasing ray sampling.

cells, cells with hitcounts larger than 50 in the irregular parameterization have to be the smallest ones ( $0.6^\circ$ ). Hence, correlation seems to increase when considering more detailed parts of the models. In general, the best sampled parts of the model are the subduction zones and hence the highest correlations occur between subduction zone regions in our  $S$  model and the  $P$  model of Bijwaard et al. (1998).

An important difference between  $P$  and  $S$  heterogeneity lies in their relative amplitudes.  $S$  amplitudes are usually up to 2-2.5 times larger. We find  $\partial \ln V_s / \partial \ln V_p$  ratios increasing from 1.2 in the upper mantle to 2.3 in the lower mantle. These are systematically smaller numbers than obtained by Robertson and Woodhouse (1996) (1.7-2.6 in the lower mantle only), but they are more similar to values calculated by Vasco and Johnson (1998) (1.2-2.0). We suspect, however, that the smaller values are significantly biased by the stronger damping that has been applied to the  $S$  inversion in both our study as well as in Vasco and Johnson (1998).

### 6.3.2 Resolution

To estimate the scales of heterogeneity that can be resolved in the  $S$  model, we have performed several sensitivity tests (Spakman and Nolet, 1988; Humphreys and Clayton, 1988). We have used artificial ‘spike’ models with 5% anomaly blocks of alternating sign spaced some distance apart from each other, for which artificial travel-time residuals were calculated and inverted. Random noise with a standard deviation of 0.5 s was added prior to inversion. The dimensions of the blocks in the different ‘spike’ models ranged from  $0.6^\circ$  to  $15^\circ$  (9 tests). The synthetic models are significantly less well retrieved for the  $S$  inversion than for the  $P$  inversion of Bijwaard et al. (1998) due the poorer ray coverage. However, the synthetics indicate that it is possible to resolve the smallest cells of  $0.6^\circ$  in some regions in the shallow upper mantle. An example is shown in Figure 6.2 for the shallow upper

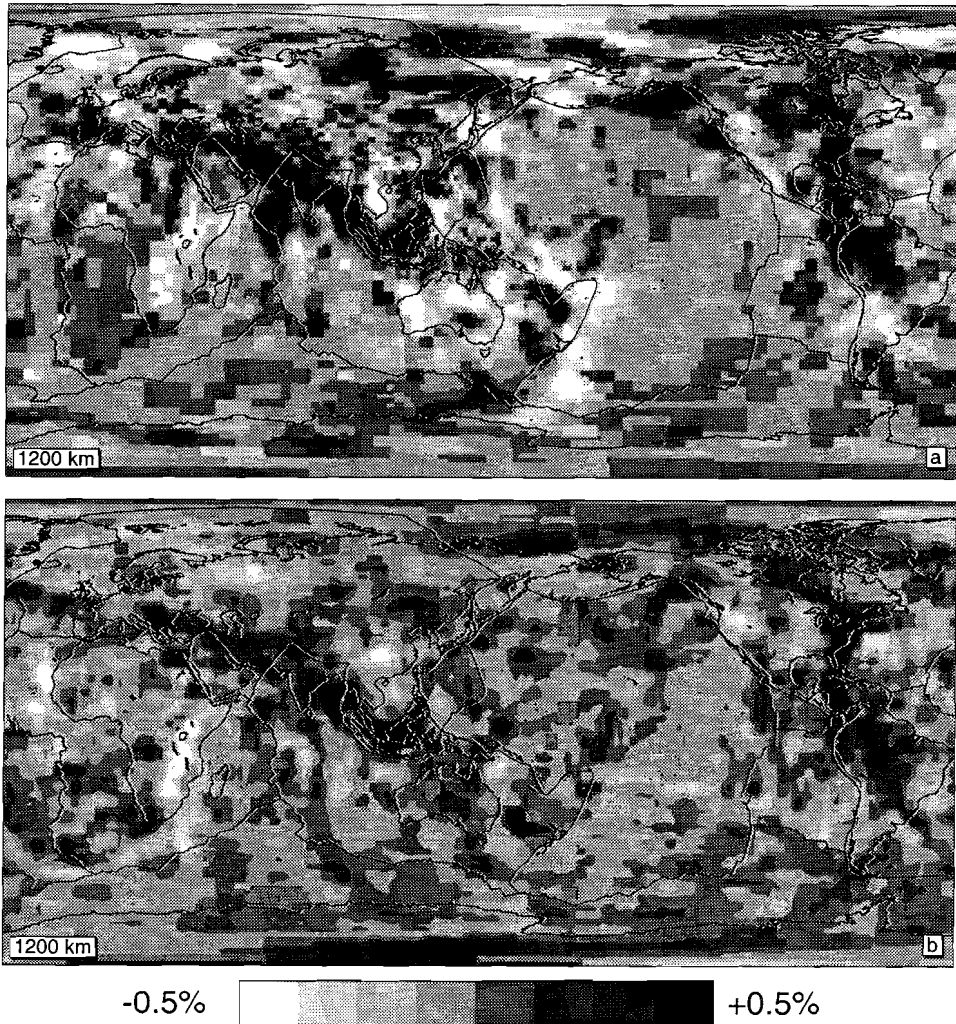


**Figure 6.2:** Solution (a) and sensitivity test for  $0.6^\circ$  spikes (b) at 50 km depth below Japan (projection centered around a great circle starting at  $(40^\circ\text{N}, 133^\circ\text{E})$  with azimuth  $120^\circ$  and length  $10^\circ$ ).

mantle below Japan. Significant parts of the upper mantle below Europe, Asia and in subduction zones are resolved on scales of  $1.8^\circ$  and  $2.4^\circ$ . Large parts of the relatively poorly sampled Americas are only retrieved on scales of  $4.2^\circ$  and larger. Most of Africa, western Australia, Antarctica and all of the oceanic regions cannot be resolved in the upper mantle. In the upper part of the lower mantle the synthetics indicate good resolution below Eurasia on a scale of  $3.0^\circ$ , below the America's on a scale of  $4.2^\circ$ - $6.0^\circ$ , and below the remainder of the northern hemisphere and parts of Africa, Australia and Antarctica on a scale of  $9.0^\circ$ . Around 1700 km the resolution decreases to  $9.0^\circ$  below the entire northern hemisphere and north and east Africa. In the lowermost mantle only Eurasia can be retrieved on scales of  $6.0^\circ$  to  $15.0^\circ$ .

### 6.3.3 Model description

In the shallow upper mantle, large regions are sampled very poorly and the inversion does not allow significant structure in these areas. In the regions that are sampled adequately we observe high velocities below cratonic areas such as northern Australia and eastern Europe, and low velocities below mountain belts in Europe and Asia Minor as well as beneath back arc basins in the northwest Pacific, southeast Asia, and west of Tonga-Kermadec. These features confirm observations from recent *P*- and *S*-tomography models (e.g. Widiyantoro,



**Figure 6.3:** Comparison between the presented model (a) and the model of Grand et al. (1997) (b) at 1200 km depth (centered around  $(0^{\circ}\text{N}, 145^{\circ}\text{E})$ ).

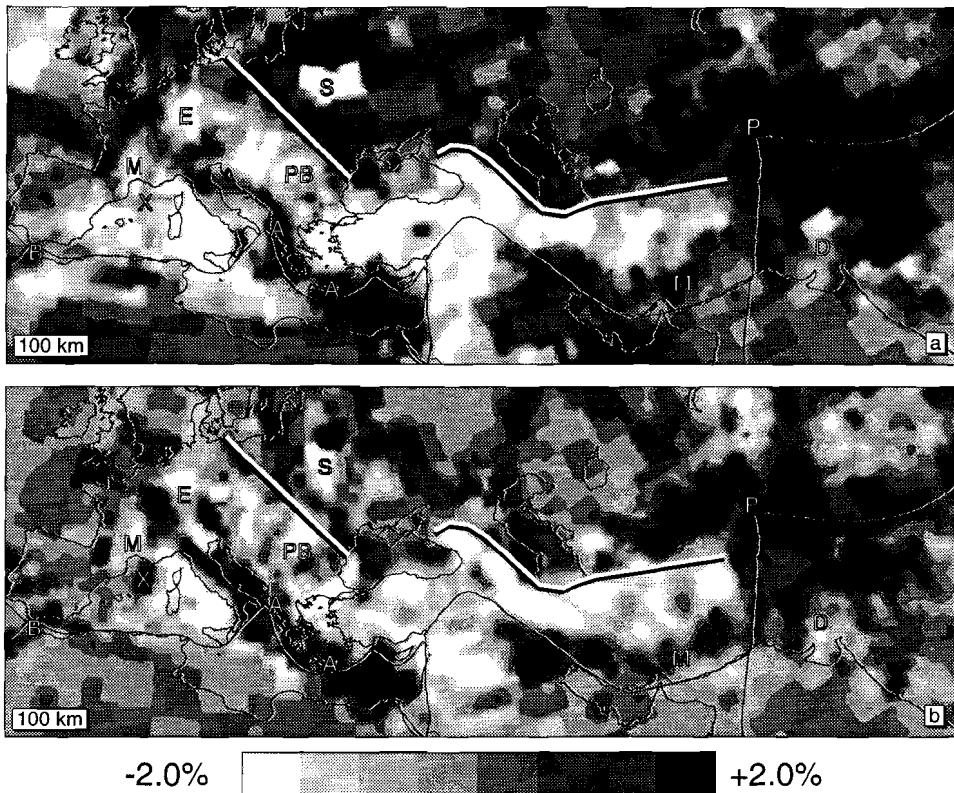
1997; Grand et al., 1997; Vasco and Johnson, 1998). The most prominent structures in the upper mantle are, however, the elongated high-velocity anomalies associated with major subduction zones in the Mediterranean (the Betic and Aegean slabs), central Asia (the Makran and Pamir zones), southeast Asia (below Japan, the Philippines, Indonesia, the Marianas, and New Guinea), the southwest Pacific (below Tonga-Kermadec, Vanuatu, and the Solomon Islands), and Central and South America (in the Caribbean, below the Andes and in the Scotia Sea). This corresponds very well with the  $P$  model of Bijwaard et al. (1998) (Chapter 4) and many of these subduction zones can also be observed in the  $P$

models of Zhou (1996) and Van der Hilst et al. (1997), but these anomalies are usually not resolved in global *S* tomography. Only the 2° cells in the *S* model of Widiyantoro (1997) depict some of these slab-related anomalies.

Although our sensitivity estimates indicate a poorer lower mantle resolution, the global *S* sampling improves in the lower mantle and our lower mantle *S* anomalies correlate very well with the *P* anomalies of Bijwaard et al. (1998). Most prominent are again the high velocity anomalies that continue from the subduction zones in the upper mantle into the lower mantle below the Mediterranean, southeast Asia, the western Pacific, and South America in accordance with observations from recent *P* and *S* models (Zhou 1996; Van der Hilst et al., 1997; Bijwaard et al., 1998; Widiyantoro et al., 1998; Grand et al., 1997). Our *S* model, however, has at this depth significantly higher amplitudes (1.3-2.3 times larger) than the *P* model of Bijwaard et al. (1998) (Chapter 4). A comparison between the model of Grand et al. (1997) and our solution at 1200 km depth is shown in Figure 6.3. The models are in general very similar, although Grand et al. (1997) seem to have a better global coverage, whereas our model is more focused with higher amplitudes. Some differences can be observed below Japan, Vanuatu/Tonga-Kermadec and Australia, but the anomalies in our model do match those of Widiyantoro et al. (1998) in these areas. These differences may reflect differences in the applied data sets. The most distinct anomalies at 1200 km depth are the large elongated high velocity anomalies associated with the subducted Tethys ocean and Farallon plate (Van der Hilst et al., 1997; Grand et al., 1997; Bijwaard et al., 1998), but smaller anomalies associated with subduction below Japan and Tonga can also be observed. Important low velocities can be observed at this depth below Iceland, Africa, Japan and the southwest Pacific, which can partly be correlated with hotspot activity at the surface. As we go deeper into the lower mantle, amplitudes slowly decrease and only large structures can be resolved. The data coverage decreases rapidly around 2000 km depth due to the small number of reliable *S* phases with regard to the *SKS* cross-over for epicentral distances associated with these depths. The Tethys and Farallon anomalies can be retrieved down to a depth of approximately 2000 km. At this depth large high velocity anomalies have appeared below Asia and northern North America. Together with low velocities below Africa and the western Pacific these start to form the well-known degree 2 pattern in the lowermost mantle (e.g. Su et al., 1994). Although the data coverage in *D''* is far from ideal, we do find discrepancies between our *S* heterogeneity and the *P* model of Bijwaard et al. (1998) (Chapter 4) below the Arctic region and Alaska (high  $V_s$ , low  $V_p$ ) that match the *S* models of Grand et al. (1997) and Widiyantoro et al. (1998), but not the *S* model of Vasco and Johnson (1998).

#### 6.3.4 Detailed comparisons of upper mantle regions

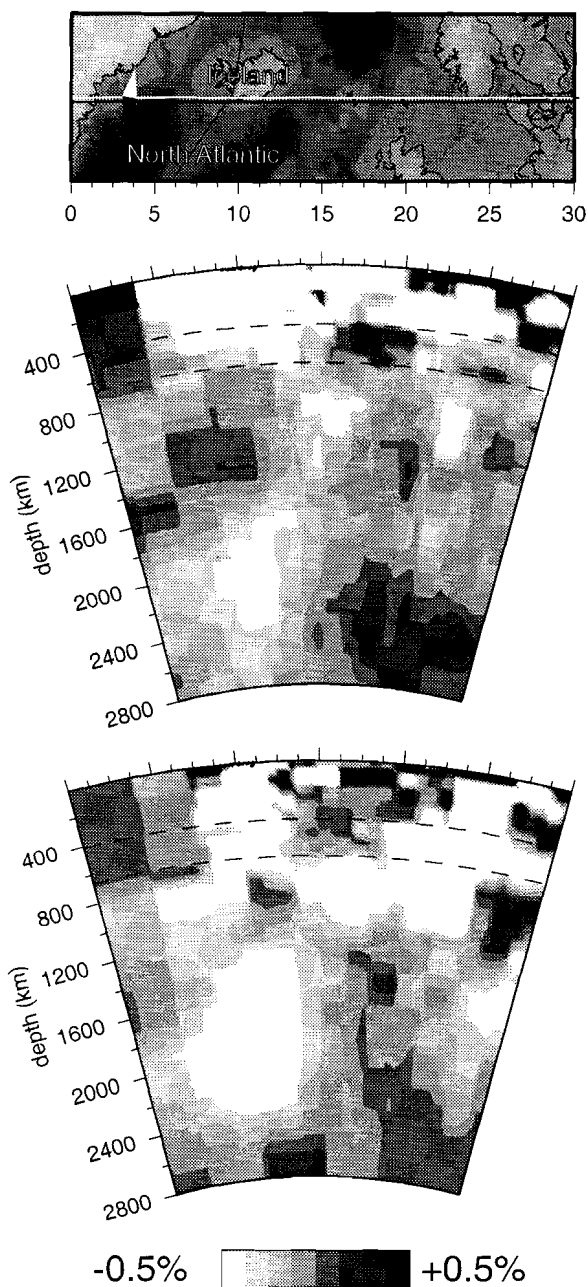
In this section we compare in detail the best sampled upper mantle regions of the presented *S* model with the *P* model of Bijwaard et al. (1998) (Chapter 4).



**Figure 6.4:** Comparison of upper mantle structure between the *P* model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented *S* model (b) for Europe to Central Asia at 100 km depth (projection centered around a great circle starting at (40°N, 10°W) with azimuth 70° and length 80°). The lettering refers to features discussed in the text: E stands for Eifel, M for Massif Central, S is a station mislocation, PB stands for Pannonian Basin, D denotes the Deccan traps, B stands for the Betic subduction, A is Aegean and Adriatic subduction, M is the Makran subduction, and P the Pamir. X denotes the discrepancy between *P* and *S* anomalies. The thick line denote the Torquist-Teisseyre zone and the margin of Paleotethys.

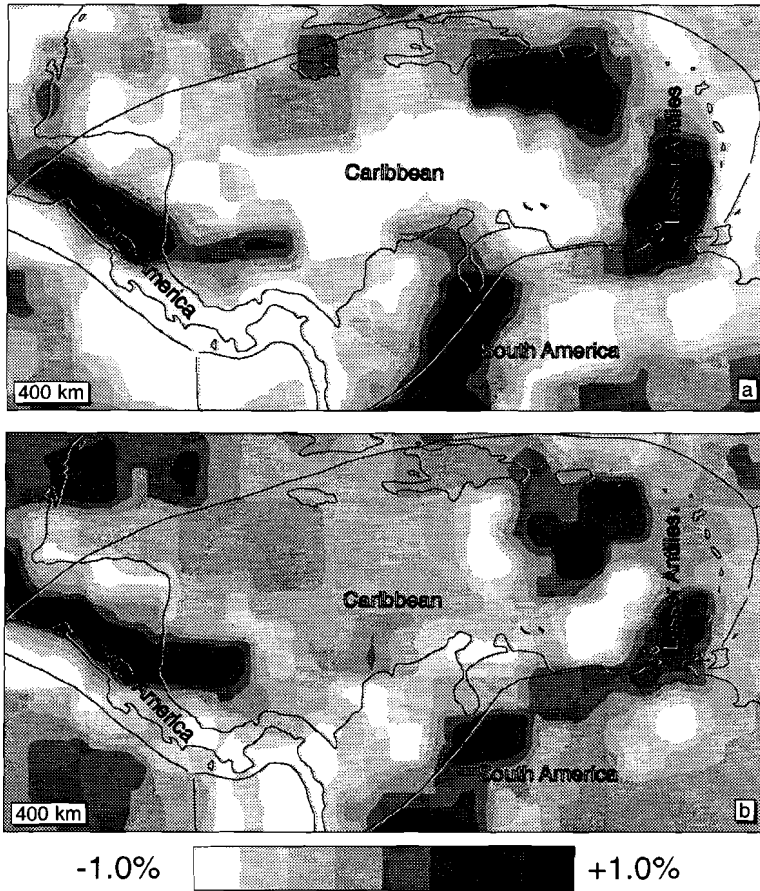
#### 6.3.4.1 Europe to Central Asia

In Figure 6.4 the *P* anomalies of Bijwaard et al. (1998) (Chapter 4) and our *S* anomalies at a depth of 100 km are shown for most of Europe, the Middle East and Central Asia. The central part of this region is sampled well by both *P* and *S* phases and it is resolved in our *S* model on scales of ~200 km. We observe a close correspondence between low *P* and *S* velocities associated with recent volcanic activity below the Massif Central, the Eifel and the Deccan traps, associated with back-arc basins in the Mediterranean, and associated with orogenic belts in Eastern Europe, Central Turkey and across Iraq, Iran and Afghanistan. Similarly, we find a good correlation between high *P* and *S* velocities in subduction zones (below southern Spain, the eastern Mediterranean, the Makran, and the Pamir), on the East-



**Figure 6.5:** Cross-sections along a great circle starting at (64°N, 44°W) with azimuth 80° and length 30° across Iceland and the North Atlantic for the *P* model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented *S* model (b).





**Figure 6.6:** Comparison of upper mantle structure between the  $P$  model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented  $S$  model (b) for the Caribbean region at 400 km depth (projection centered along a great circle starting at  $(13^{\circ}\text{N}, 90^{\circ}\text{W})$  with azimuth  $90^{\circ}$  and length  $30^{\circ}$ ).

European platform (although a low velocity artefact due to a station mislocation is obscuring the view) and below northern India and Tibet. The boundary between Phanerozoic and Precambrian Europe (the Tornquist-Teisseyre zone (Zielhuis and Nolet, 1994)) and the former Paleotethys margin are clearly visible as seismic velocity contrasts.

There seem to be hardly any clear discrepancies between  $P$  and  $S$  that are well-resolved. There is a small difference in the shape of the slab-related anomalies below the Adriatic and we observe a high  $S$ -velocity anomaly west of Corsica that is not as clearly present in  $P$ . We observe a general correspondence with the regional  $P$  tomography model of Spakman et al., (1993) which also shows the high velocity anomaly west of Corsica. Deeper in the mantle (not shown) we find only a few other discrepancies between  $P$  and  $S$ . At 600-700 km depth we observe high  $P$ , low  $S$  velocities below northern Italy and southeast France, around 800 km depth we find low  $P$ , high  $S$  velocities below England, and at 1300-1400 km depth we

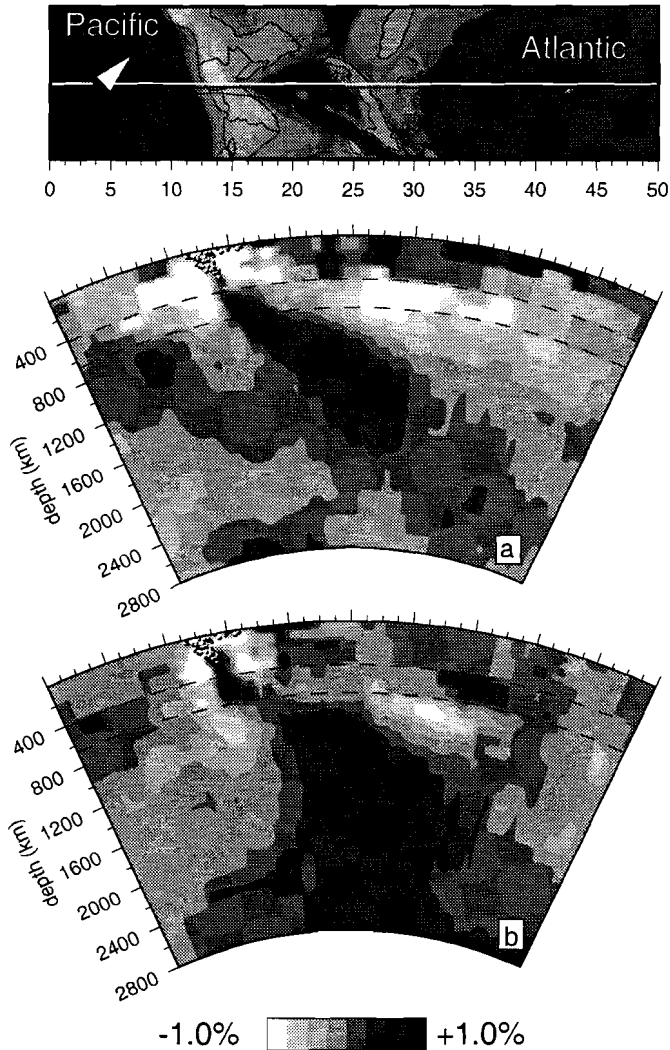
find low *P* and high *S* velocities in the western Mediterranean, below Sardinia. To further illustrate the general correlation between *P* and *S* heterogeneity we show two cross-sections in Figures A9 and 6.5. Figure A9 depicts a section across the Mediterranean and the Aegean from the *P* model of Bijwaard et al. (1998) (Chapter 4) and from our *S* model. Both the high slab-related velocities and the low back-arc related velocities correlate well. Figure 6.5 shows similar sections across Iceland and the North Atlantic. Bijwaard and Spakman (1998) (section 5.1) have, based upon the *P* velocities, inferred the existence of a whole mantle plume below Iceland. The lower panel in Figure 6.5 generally confirms their results in the presented *S* model.

#### 6.3.4.2 Southeast Asia to Tonga-Kermadec

In Figure A11 the *P* anomalies of Bijwaard et al. (1998) (Chapter 4) and our *S* anomalies at a depth of 350 km are shown for southeast Asia to Tonga-Kermadec. Our resolution estimates for this region range from 100-450 km for the *S* model, with generally good resolution in the subduction zones. We observe a striking correlation between high *P* and *S* wavespeeds below the subduction zones of Japan, Izu-Bonin, the Marianas, Taiwan, the Philippines, Sulawesi, the Sunda and Banda Islands, New Guinea, the Solomon Islands, Vanuatu, and Tonga-Kermadec. Furthermore, the low velocity anomalies surrounding many of these subduction zones also show up in both the *P* and *S* heterogeneity. Discrepancies between *P* and *S* below the Asian continent and parts of Australia can be attributed to poorer resolution for the *S* model. However, the discrepancy just north and northwest of Australia below the Timor Sea (high  $V_s$ , low  $V_p$ ) is, according to our sensitivity estimates, resolved. A similar discrepancy between *P* and *S* velocities occurs east of Australia below the Coral Sea at a depth of 900 km (not shown), where we observe low *P*, high *S* velocities. In Figures A10 and A12 we show two cross-section comparisons for Java and Borneo and for Sulawesi and Halmahera. Figure A10 clearly shows the subduction of the Indo-Australian plate below Java and the continuation of the related high *P* and *S* velocities into the lower mantle. Discrepancies in the uppermost mantle below Borneo are due to lack of *S* sampling. Figure A12 depicts the subductions in opposite directions below Sulawesi and Halmahera (also outlined by the seismicity). We again observe a very close correspondence between the *P* and *S* results.

#### 6.3.4.3 The Caribbean

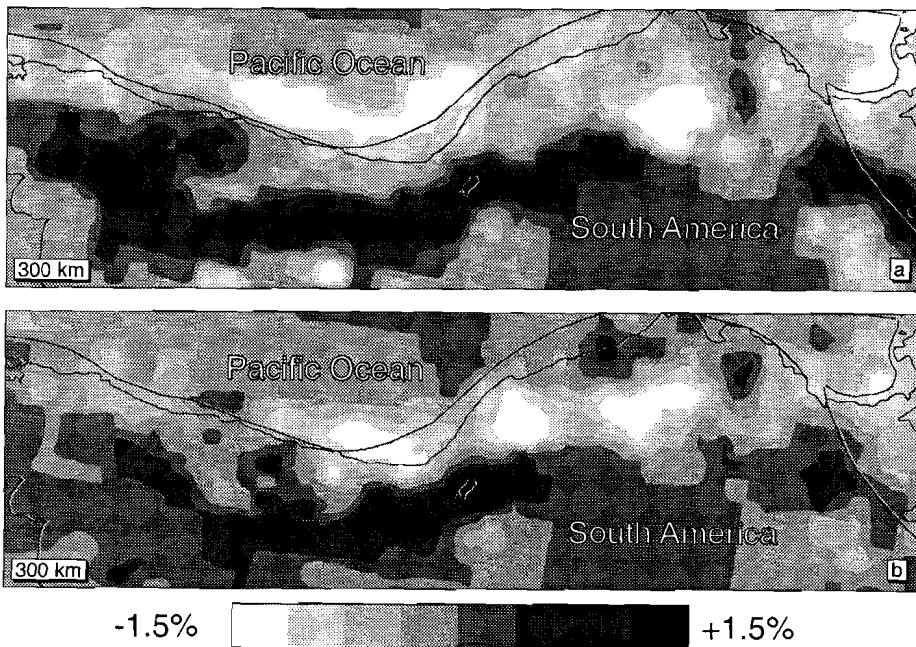
In Figure 6.6 the *P* anomalies of Bijwaard et al. (1998) (Chapter 4) and our *S* anomalies at a depth of 400 km are shown for the Caribbean region. Due to the poorer sampling by *S* phases of this region, the *S* anomaly amplitudes are suppressed by the regularization and hence somewhat smaller than the *P* amplitudes. Furthermore, the sensitivity tests only indicate resolution on scales of 400 km and larger. The pattern of heterogeneity is, however, down to small scales very similar for *P* and *S*. Most of the Caribbean Sea has both low *P* and *S* velocity, and it is surrounded by several patches of high *P* and *S* velocity related to subduction. Below Central America we observe high velocities related to the subduction of the Cocos or Farallon plates; below the Antilles the high velocities are due to subduction of



**Figure 6.7:** Cross-sections along a great circle starting at (5°N, 100°W) with azimuth 45° and length 50° across the Caribbean for the *P* model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented *S* model (b).

the Atlantic plate. The gaps in the high velocity, slab-related anomalies just west of the Lesser Antilles and in northern Venezuela may be due to roll-back of the southern Lesser Antilles part of the slab as has been suggested by Van der Hilst (1990). The high velocities below northern South America have been attributed to underthrusting of the Caribbean oceanic plateau below South America (Van der Hilst and Mann, 1994).

We do not find any major discrepancies between *P* and *S* heterogeneity. Figure 6.7 shows a cross-section for *P* and *S* through the Caribbean, which indicates slab-related high velocities in upper and lower mantle. Discrepancies in the upper mantle beneath the oceanic

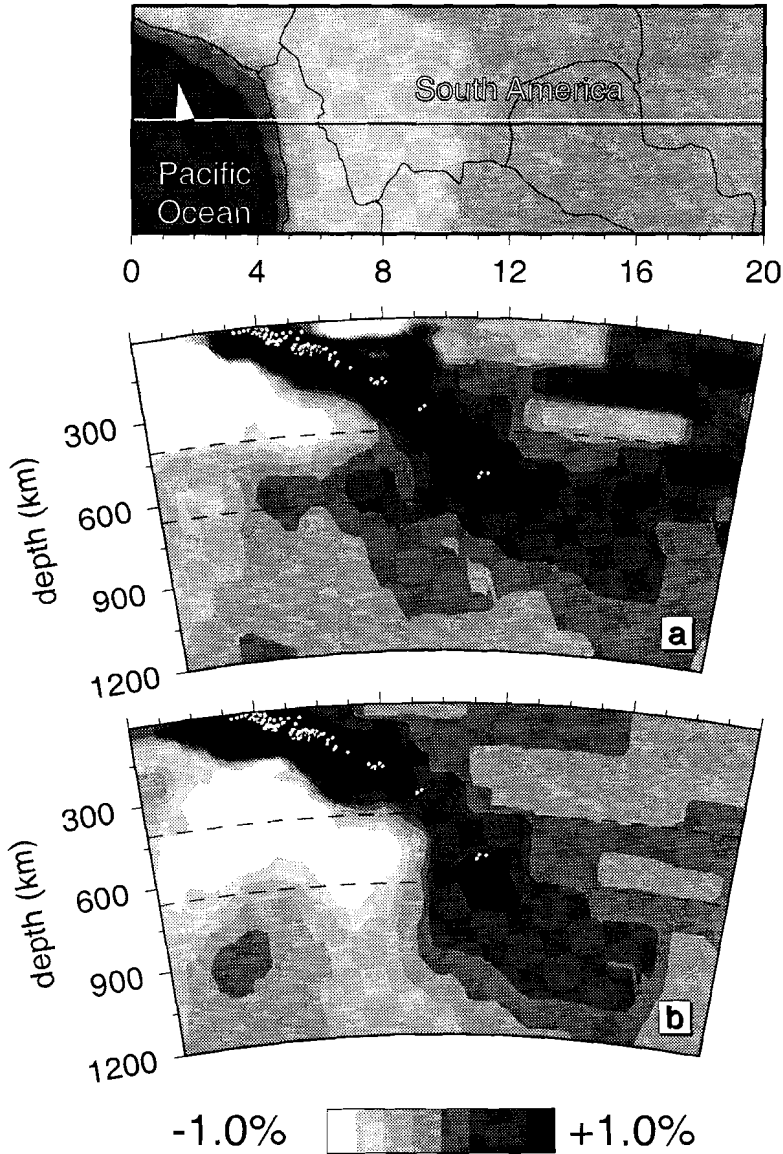


**Figure 6.8:** Comparison of upper mantle structure between the *P* model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented *S* model (b) for South America at 300 km depth (projection centered along a great circle starting at (42°S, 67°W) with azimuth 350° and length 52°).

regions are due to the absence of *S* sampling.

#### 6.3.4.4 South America

In Figure 6.8 the *P* anomalies of Bijwaard et al. (1998) (Chapter 4) and our *S* anomalies at a depth of 300 km are shown for the western part of South America. The central part of this region is sampled well by both *P* and *S* phases and resolved in our *S* model on scales of ~300 km. In this part we observe similar patterns of high (slab) and low *P* and *S* velocities. The *P* anomalies cover a larger area than the *S* anomalies because of the better *P*-wave coverage of South America. The slab-related high velocities agree well with a regional *P* tomography study by Engdahl et al. (1995), who argue that the downgoing Nazca plate must be continuous at this depth despite the absence of seismicity. Any observed discrepancies between *P* and *S* heterogeneity are usually of small scale and amplitude and not resolved. Figure 6.9 shows how the high slab velocities continue in both *P* and *S* from the surface into the uppermost lower mantle.



**Figure 6.9:** Cross-sections along a great circle starting at  $(19^{\circ}\text{S}, 75^{\circ}\text{W})$  with azimuth  $101^{\circ}$  and length  $20^{\circ}$  across South America for the  $P$  model of Bijwaard et al. (1998) (Chapter 4) (a) and the presented  $S$  model (b).

#### 6.4 Discussion

We have presented a new  $S$  tomography model with small-scale detail that is comparable in several well-sampled regions with regional tomography studies. Recently, two other global

*S*-wave tomography studies (Grand et al., 1997; Widiyantoro, 1997) have employed rather detailed ( $2^\circ$  and 275 km, respectively) parameterizations. Grand et al. (1997) have used a very different data set (including *ScS*, *SS*, *SSS*, and *SSSS* phases) with a more uniform but generally lower global sampling than we have. In the upper mantle their model shows large-scale heterogeneity beneath oceans and continents, but it lacks the small-scale subduction zone geometries we observe in our model. In the lower mantle their model is rather similar to our model (see for example Figure 6.3) with a better global sampling, but with somewhat smaller amplitudes and less focussed.

The model of Widiyantoro (1997) has been derived from a very similar data set, but on a  $2^\circ$ , regular cell grid. He has compared inversions with and without *SKS* data to investigate their influence on the retrieval of deepest mantle structure. His solution is in both upper and lower mantle similar to ours on scales of  $2^\circ$  and larger, but with a smaller amplitude. This may be due to a stronger regularization that is needed to constrain the solution in the (many) poorly sampled regular cells. In the upper mantle, major subduction zones are clearly visible in his results but seem to lack the small-scale detail that is often present in our model. In the lowermost mantle the inclusion of *SKS* data increases the ray sampling and probably leads to better determined structure below 2000 km (see also Widiyantoro et al., 1998). Since we have derived our *S* model in a similar fashion as the *P* model of Bijwaard et al. (1998) (Chapter 4), we have tried to compare *P* and *S* velocities. In general, seismic velocities depend on both temperature and composition. Temperature anomalies have a similar effect on *P* and *S* wavespeeds: Higher temperatures decrease *P* and *S* velocities (see e.g. Karato, 1993). The effect may be larger for *S* leading to larger *S* anomaly amplitudes, but we shall refrain from interpreting relative amplitudes of *P* and *S* heterogeneity because these are influenced by the different regularizations in their respective inversions and by the different cut-off applied to their residuals. Compositional anomalies can have both a similar and an opposite effect on *P* and *S* wavespeeds. Regions where *P* and *S* heterogeneity have a similar sign may thus be influenced by both temperature and compositional anomalies, but regions where they are of opposite sign are likely to result from compositional heterogeneity alone.

Recently, simultaneous global inversions of *P* and *S* phases (Robertson and Woodhouse, 1996; Su and Dziewonski, 1997; Kennett et al., 1998; Vasco and Johnson, 1998) have been performed. These studies all find the largest differences between either *P* and *S* or between bulk sound and *S* at the bottom of the mantle. Apart from *D''* these studies do not mention distinct differences between *P* and *S* anomalies (Kennett et al. (1998) do observe differences between bulk sound and *S*, but not between *P* and *S*). From separate inversions for *P* and *S* velocities, Widiyantoro (1997) concludes that these yield similar results and he does not discuss any differences.

As indicated above the anomalies in our *S* model correlate very well with the *P* anomaly model of Bijwaard et al. (1998) (Chapter 4). In only a few places we observe anticorrelated values of *P* and *S* heterogeneity. We find differences between *P* and *S* in the *D''* region at the bottom of the mantle, but due to the lack of *S* data this region is not well resolved in our model. The well-resolved discrepancies are situated below Europe (Corsica, northern Italy/southeast France, England, and Sardinia) and below the Timor and Coral Seas. These discrepancies were not mentioned by Widiyantoro (1997), but some of them (most notably the Timor Sea) are also present in his *P* and *S* models. In these regions we may expect compositional heterogeneity.

---

## 6.5 Conclusions

We have used the irregular parameterization technique of Spakman and Bijwaard (1998) and the  $S$  data of Engdahl et al. (1998) to obtain an  $S$  model with unprecedented detail in well-sampled regions. Since the  $S$  model is derived in a similar fashion as the high-resolution  $P$  model of Bijwaard et al. (1998) (Chapter 4) we have compared the  $P$  and  $S$  results. Although the quality and number of  $S$  data is far less than of  $P$  data, the  $P$  and  $S$  solutions are very similar in structure in parts of the models that are sampled adequately (correlation coefficients up to 0.85). The  $S$  model does exhibit generally higher amplitudes mainly in the lower mantle. The ratio  $\partial \ln V_s / \partial \ln V_p$  increases from 1.2 in the upper mantle to 2.3 in the lower mantle but may be underestimated due to the relatively high regularization applied in the  $S$  inversion. Sensitivity tests indicate that the resolution for the  $S$  model is generally poorer than for the  $P$  model, but small-scale resolution (down to  $0.6^\circ$ ) may be obtained in parts of the shallow upper mantle. In these parts we can observe the detailed structure of high  $S$ -velocity anomalies associated with subduction which are very similar to the  $P$ -velocity anomalies. In the lower mantle the imaged velocity heterogeneity resembles the models of Grand et al. (1997) and Widiyantoro et al. (1998), but is more focused and has a higher amplitude.

The most interesting parts of the model may be the few places where  $P$  and  $S$  heterogeneity differ. Discrepancies between  $P$  and  $S$  wavespeeds occur in the uppermost mantle below the western Mediterranean, in the lowest part of the upper mantle below the Timor Sea, from 600-900 km depth in several places below Europe and the Coral Sea, and in the  $D''$  layer below the Arctic region and Alaska. In these regions the velocity heterogeneity may have a (partly) compositional origin.

**Acknowledgments.** We thank E. R. Engdahl for providing the  $S$  data set. We are grateful for support from the Netherlands Geosciences Foundation (GOA) with financial aid from the Netherlands Organization for Scientific Research (NWO). This work was conducted under the program of the Vening Meinesz School of Geodynamics.

## References

- Bijwaard, H. and W. Spakman, Tomographic evidence for a narrow whole mantle plume below Iceland, *Earth Planet. Sc. Lett.*, 166, 121-126, 1999.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Engdahl, E.R., R.D. van der Hilst, and J. Berrocal, Imaging of subducted lithosphere beneath South America, *Geophys. Res. Lett.*, 22, 2317-2320, 1995.
- Engdahl, E.R., R.D. van der Hilst, and R.P. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Amer.*, 88, 722-743, 1998.
- Grand, S.P., R.D. van der Hilst, and S. Widiyantoro, Global seismic tomography: a snapshot of convection in the Earth, *GSA Today*, 7, 4, 1-7, 1997.
- Humphreys, E., and R.W. Clayton, Adaptation of back projection tomography to seismic

- travel time problems, *J. Geophys. Res.*, 93, 1073-1085, 1988.
- Karato, S., Importance of anelasticity in the interpretation of seismic tomography, *Geophys. Res. Lett.*, 20, 1623-1626, 1993.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, 122, 108-124, 1995.
- Kennett, B.L.N., S. Widiyantoro, and R.D. Van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, 103, 12469-12493, 1998.
- Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Phil. Trans. R. Soc. Lond. A*, 354, 1385-1411, 1996.
- Paige, C.C., and M.A. Saunders, LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.*, 8, 43-71, 1982.
- Robertson, G.S., and J.H. Woodhouse, Ratio of relative *S* to *P* velocity heterogeneity in the lower mantle, *J. Geophys. Res.*, 101, 20041-20052, 1996.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic problems, *Ann. Geophys.*, 16, suppl. 1, 28, 1998.
- Spakman, W., and G. Nolet, Imaging algorithms, accuracy and resolution in delay time tomography, in: *Mathematical Geophysics*, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, pp. 155-187, Reidel, Dordrecht, The Netherlands, 1988.
- Spakman, W., S. Van der Lee, and R.D. Van der Hilst, Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, 79, 3-74, 1993.
- Su, W.-J., and A.M. Dziewonski, Joint 3-D inversion for *P*- and *S*-velocity in the mantle, *EOS, Trans. Am. Geophys. Union*, 74(43), Fall Meeting Suppl., 557, 1993.
- Su, W.-J., and A.M. Dziewonski, Simultaneous inversion for 3-D variations in shear and bulk velocity in the mantle, *Phys. Earth Planet. Inter.*, 100, 135-156, 1997.
- Su, W.-J., R.L. Woodward, and A.M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, 99, 6945-6980, 1994.
- Van der Hilst, R.D., Tomography with *P*, *PP* and *pP* delay-time data and the three-dimensional mantle structure below the Caribbean region, Ph.D. thesis, Utrecht Univ., 250 pp., 1990.
- Van der Hilst, R.D., and P. Mann, Tectonic implications of tomographic images of subducted lithosphere beneath northwestern South America, *Geology*, 22, 451-454, 1994.
- Van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, 386, 578-584, 1997.
- Vasco, D.W., and L.R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, 103, 2633-2671, 1998.
- Widiyantoro, S., Studies of seismic tomography on regional and global scale, Ph.D. thesis, Australian National Univ., 256 pp., 1997.
- Widiyantoro, S., B.L.N. Kennett, and R.D. van der Hilst, Extending shear-wave tomography for the lower mantle using *S* and *SKS* arrival-time data, *Earth Planets Space*, 50, 999-1012, 1998.
- Zhou, H., A high-resolution *P* wave model for the top 1200 km of the mantle, *J.*



Geophys. Res., 101, 27791-27810, 1996.

Zielhuis, A., and G. Nolet, Deep seismic expression of an ancient plate boundary in Europe, *Science*, 265, 79-81, 1994.

# Chapter 7<sup>1</sup>

## Fast kinematic ray tracing of first and later arriving global seismic phases

**Summary.** We have developed a ray tracing algorithm that traces first and later arriving global seismic phases precisely (travel-time errors on the order of 0.1 s), and with great computational efficiency (15 rays/s). To achieve this, we have extended and adapted two existing ray tracing techniques: A graph method and a perturbation method. The two resulting algorithms are able to trace (critically) refracted, (multiply) reflected, some diffracted (*Pdiff*), and (multiply) converted seismic phases in 3-D, spherical geometry, thus including the largest part of seismic phases that are commonly observed on seismograms. We have tested and compared the two methods in 2-D and 3-D, Cartesian and spherical models for which both algorithms have yielded precise paths and travel times. These tests indicate that only the perturbation method is computationally efficient enough to perform 3-D ray tracing on global data sets of several millions of phases. To demonstrate its potential for nonlinear tomography, we have applied the ray perturbation algorithm to a data set of 7.6 million *P* and *pP* phases used by Bijwaard et al. (1998) (Chapter 4) for linearized tomography. This showed that the expected heterogeneity within the Earth's mantle yields significant nonlinear effects on travel times for 10% of the applied phases.

### 7.1 Introduction

In several branches of seismology and exploration seismics a need exists for fast and precise ray tracing in 3-D heterogeneous Earth models to find the paths seismic waves travel from source to receiver. Especially earthquake location determination (e.g. Virieux et al., 1988) and nonlinear seismic tomography (e.g. Sambridge and Snieder, 1993) require extensive ray tracing for travel times and ray geometries. Numerous algorithms have therefore been developed (for an overview see Červený, 1987; Virieux, 1996) which can be categorized as: shooting (e.g. Julian and Gubbins, 1977), bending (e.g. Um and Thurber, 1987; Moser et al., 1992; Koketsu and Sekine, 1998), finite difference (e.g. Pereyra et al., 1980; Vidale, 1988, 1990; Podvin and Lecomte, 1991), network or graph (e.g. Moser, 1991; Klimeš and Kvasnička, 1994) and perturbation methods (e.g. Farra and Madariaga, 1987; Virieux, 1991; Snieder and Sambridge, 1992, 1993).

The complexity of the ray tracing problem is illustrated by the fact that all existing techniques have one or more specific shortcomings: Shooting methods cannot trace rays in shadow zones and are only guaranteed to find absolute travel-time minima at the cost of increased CPU time. Bending methods are generally fast, but may converge toward local travel-time minima. The effectiveness of finite difference and graph methods trades off with the size of the study volume as either whole wavefronts or complete shortest path trees are

---

<sup>1</sup> This chapter is in the press for publication in *Geophys. J. Int.* as: Bijwaard, H, and W. Spakman, *Fast kinematic ray tracing of first and later arriving global seismic phases.*

---

calculated. Finally, the developed perturbation theory is only useful in mildly perturbed media because the resulting ray paths are only precise to first order (see Sambridge and Snieder, 1993). More importantly, most of the implementations of these methods were designed to trace only direct  $P$  or  $S$  phases. Even though some existing 3-D algorithms are able to trace reflections and/or head waves, we are not aware of methods that can perform 3-D ray tracing for a large variety of seismic phases on a whole Earth scale efficiently enough to trace data sets of several million phases in a feasible amount of time.

Our motivation for the development of a very fast ray tracing algorithm stems from our goal to perform nonlinear global travel-time tomography. According to Kennett et al. (1998) the 3-D ray tracing of their 300000 rays will already be a major task. However, the extension of the linear global tomography model of Bijwaard et al. (1998) (Chapter 4) to nonlinear tomography requires the ray tracing of 7.6 million teleseismic  $P$  and  $pP$  phases. A (hypothetical) computation time of half a second per ray would already lead to 44 days of computing on a single processor.

Here we present two kinematic ray tracing techniques that are substantial extensions and adaptations of existing methods: A graph method, following the work of Nolet and Moser (1993), and a perturbation method based upon Snieder and Sambridge (1993) and Pulliam and Snieder (1996). Both methods are essentially used as bending methods and converge toward nearby stationary paths, although the graph method can also be employed to yield global minimum travel-time paths. We have extended both techniques to handle first order seismic discontinuities, at which rays may reflect, (critically) refract, in some cases diffract, and/or convert in 3-D, spherical geometry, thus enabling the ray tracing of the largest part of seismic phases that are commonly observed on seismograms.

## 7.2 General framework for ray tracing

In general, one needs a seismic velocity model to perform seismic ray tracing. Basic in our development of the ray tracing algorithms is that this is independent of the particular model geometry, i.e. Cartesian, spherical, 2-D or 3-D, used to obtain the seismic velocities and interfaces, and independent of the model parameterization used in, for example, seismic tomography. Although tying ray tracers to a particular model geometry or parameterization has the advantage of obtaining high precision results, it precludes the implementation of any other model geometries. Apart from that, high precision ray tracing does not ensure accurate results, i.e. with respect to the actual travel times in the real velocity medium, because generally model velocities are in error. For instance, models obtained from tomographic experiments have specific - often unknown - bounds on spatial resolution and model covariance. To be able to use different model geometries, we have separated all input required for ray tracing from the actual ray tracing algorithms. The required input, which consists (among other things) of the model geometry, seismic velocities, interfaces, an initial ray path, and tuning parameters, is supplied by the ray tracer framework. A selected ray tracing algorithm then iteratively updates the initial path and returns new ray coordinates and travel times. Iteration is necessary because in general initial rays may be far away from the final ray.

In order to trace complex ray paths that may refract, reflect and/or convert at seismic discontinuities that were defined by the input model geometry, we use a so-called 'phase

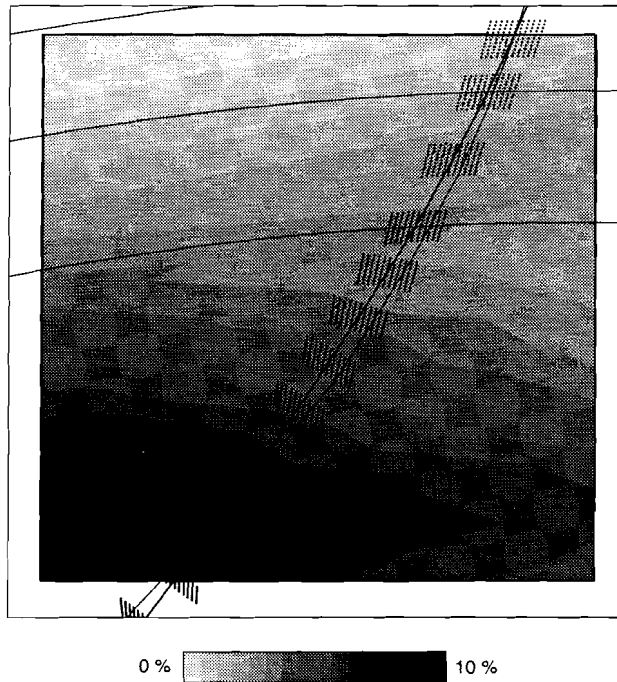
code'. This code describes how the initial path runs through the model volume, i.e. whether it is going up or down from the source, which interfaces can be crossed, and at which interfaces the ray is reflected and/or converted. An extra input parameter is used to indicate critically refracted phases (head waves) or diffracted phases. Because the 'phase code' cannot change during ray tracing, one can trace rays belonging to specific phases and particular travel-time branches. This considerably reduces the problem of multipathing that may arise because both methods are essentially bending techniques. Now, multipathing can only occur if the 3-D heterogeneity allows for more than one stationary path within the same phase branch. A 3-D ray tracing method that computes the entire wavefront sufficiently precise (e.g. Lucio et al., 1996) can in that case find all possible paths. However, multipathing is an observational problem as well: In case of multipathing it is, with respect to the uncertainty in the velocity model, uncertain which path belongs to the phase that has been identified on the seismogram.

With the 'phase code' an initial ray path is determined through a reference model such as the 1-D reference mantle model ak135 (Kennett et al., 1995) by a fast 1-D shooting algorithm. In general initial rays can also be simply defined by only giving initial ray coordinates at source and receiver and at any interface crossed by it (or a reflection point) which are then implicitly connected by straight ray segments. At the start of the ray tracing the ray tracer framework provides the ray tracing algorithm with the initial ray coordinates, which include all intersections with seismic interfaces (but not necessarily at the right location), and the 'phase code'. During iteration the ray tracer framework supplies seismic velocities and tuning directives when needed. The tuning directives are essentially adjustment schemes for node distances or graph attributes (graph plane distances, graph node density) for each iteration, which can be adapted to the user's needs.

The ray tracer framework described above serves both the graph method and the perturbation method. Our motivation for the development of two ray tracers is that it allows us to cross check the ray tracing results. In complex media both ray tracers need to compute similar ray paths and travel times (e.g. travel-time differences  $< 0.1\text{-}0.2$  s for teleseismic phases). Although both methods are essentially used as bending techniques, they are completely independent and there is no a priori reason why they should yield the same ray path. The perturbation method uses local velocity gradients to find a stationary path, whereas the graph method performs a slowness grid search for the minimum travel-time path. Linear slowness variation is assumed between ray nodes (for the perturbation method) or graph planes (for the graph method) which implicitly dictates maximum node or graph plane distances. For the graph method tuning parameters can be adjusted (by taking, for example, large graph planes) to search for absolute minimum travel-time paths.

### **7.3 The adjusted graph method**

In Moser's original implementation of the graph method (Moser, 1991), a grid of potential ray nodes is defined in the entire model volume, and shortest paths are calculated from a single source node to all other nodes simultaneously. The effectiveness of the graph method hence depends on the spatial dimensions of the study volume. Computing precise paths and travel times in very large models (like the entire Earth's mantle and crust) requires huge amounts of storage and computation time.



**Figure 7.1:** An example of a 3-D graph network consisting of planes of nodes perpendicular to the initial ray path (thick black line) or parallel to velocity discontinuities (interfaces are shown as parts of concentric circles). The first update of the ray in the graph is shown as a thin black line. A 2-D section from a 3-D anomaly model is inserted to approximately depict the heterogeneous velocities along the ray path.

To circumvent this computational problem, Nolet and Moser (1993) decided not to discretize the entire model space, but only a restricted volume around an initial ray path. Their ‘cigar-shaped’ grid consisted of equidistant graph planes containing the grid nodes, perpendicular to the ray. At interfaces the graph planes were, by exception, positioned in the plane of the discontinuity. The sum of all graph planes constituted the graph in which Moser’s original graph method (Moser, 1991) would then find the shortest path. To be able to cope with large ray displacements, this procedure was iterated: The ray update became the center of a new graph.

Our approach resembles the ‘Bended Graph’ of Nolet and Moser (1993), but there are a few subtle differences. We, for example, do not use a cigar-shaped graph: Our graph planes all have the same size. This approach takes slightly more nodes, but it has the advantage of a very simple geometry and a lucid node numbering (an example is shown in Figure 7.1). Furthermore, nodes need not be placed equidistantly in the graph planes: The distances can be scaled in the direction perpendicular to the ray such that the node density is higher in the immediate surroundings of the ray path. This is especially useful in later iterations to zoom in on the final solution. Finally, all graph tuning parameters, i.e. the distance between graph planes, the number of nodes in each plane, the node distances and their scaling, and the

forward star (Moser, 1991), can be adjusted between iterations. This gives us the opportunity to experiment with different sets of parameters to find the most efficient way to trace a specific collection of rays in a specific model.

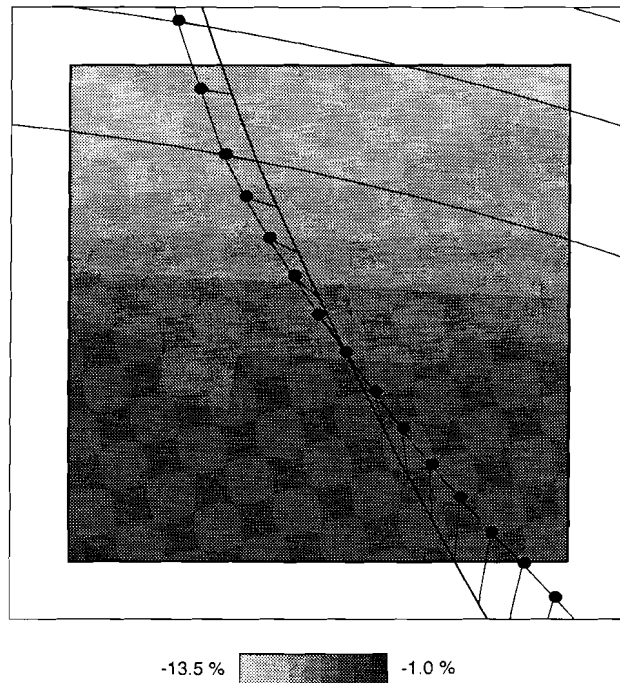
Apart from these subtle differences, we have extended the graph method to include (multiply) reflected and (multiply) converted phases, diffractions at specific model interfaces (to include *Pdiff* and *PKPdiff*), and head waves. This is done in the following way: The initial ray path of a reflected phase has a graph plane at the reflective interface. This plane can, during iteration, only move along the interface. The shortest path will then automatically satisfy the reflection law. In the calculation of a head wave, we proceed in a similar manner: The initial ray path is calculated from a reflection with the critical incidence angle at the refractor. The upgoing leg of this critical reflection is shifted to the receiver and the space in between downgoing and upgoing leg is filled by inserting an appropriate amount of extra planes at the refractor. For the updates of this initial path, the planes at the refractor can only move along the interface. The procedure is very similar for diffracted phases except for the fact that we start from a ray path that is grazing the appropriate discontinuity instead of a critical reflection. Finally, for converted phases we only need to assign *P* and *S* velocities to the right graph nodes: The shortest path method will then automatically find the path that obeys Snell's law.

The computation time of the graph method mainly depends on the number of nodes in the grid, and the applied forward star, but the anomaly model itself may also be of importance. For instance, a very heterogeneous model will require larger graph planes, a larger forward star, and in some cases more iterations. The precision of the method depends on the density of the grid (especially in the final iterations), and the applied forward star. The graph plane distance and node spacing should be chosen carefully as the slowness is supposed to vary linearly between nodes. For an optimal choice of parameters some tests should be performed with different tunings on a representative subset of the data in the specific model. We shall illustrate the attainable precision and computation speed in section 7.5.

#### 7.4 The extended ray perturbation method

The applied perturbation method is based on Pulliam and Snieder's (1996) implementation of the ray perturbation theory developed by Snieder and Sambridge (1993). They derived from the equation of kinematic ray tracing the first order perturbation of the ray geometry, and the second order perturbation of its travel time, due to either a perturbation of the slowness, a perturbation of the ray's endpoints or the reference curve not being a true ray. Pulliam and Snieder (1996) developed a very fast implementation of this theory for 2-D triangular-cell models. They use approximated perturbation equations, and assume small, linear slowness variations and straight reference ray segments in each cell. Their method essentially balances the curvature of the ray, and the component of the slowness gradient perpendicular to it. It does not explicitly minimize the travel time along the ray.

We extended the method to work in 3-D velocity models with discontinuities at which rays may (critically) refract, reflect, in some cases diffract (to include *Pdiff* and *PKPdiff*), and/or convert. The equations which were originally intended for rays with straight segments between triangular cell boundaries hold equally well for our rays with straight segments between ray nodes. The equations that describe our extension of the method are given in



**Figure 7.2:** Similar to Figure 7.1: An example of the perturbation method in a (part of a) 3-D, spherical medium. The thick black line represents the initial path, the thin line with black dots its first update.

appendix A. In short: For refracting and converting rays we have implemented Snell's law at the velocity discontinuities, and for reflecting rays the reflection law at the reflective interfaces. An example of the perturbation of a refracting ray in a heterogeneous model is shown in Figure 7.2.

Pulliam and Snieder's (1996) equations for the ray geometry, together with Snell's law, and the reflection law form a matrix-vector equation for all ray node positions. In Pulliam and Snieder's (1996) implementation, the matrix in this equation is tridiagonal, and each ray node coordinate can be inverted separately (resulting in two small matrices in 2-D). With the inclusion of velocity discontinuities, and corresponding equations this is, however, no longer possible (see also appendix A), since the dot products involving the normals on discontinuities couple the model coordinates. The inversion of a much larger matrix (including all ray node coordinates) with a broader bandwidth (14 in 3-D) is thus inevitable. This implies that the inclusion of velocity interfaces will reduce the computation speed of the perturbation method.

Another difference with the implementation of Pulliam and Snieder (1996) is that the solution is iterated instead of applying a single inversion for new ray node coordinates. This has proven to be useful for rays in very heterogeneous parts of models for which a single iteration is not guaranteed to give a precise result (see Sambridge and Snieder, 1993).

The main computational burden of the perturbation method consists of the determination of

(perpendicular) slowness gradients, and the matrix inversion of the equations for the ray nodes. In 3-D these take up 60% of the CPU time. Again, the number of ray nodes, the model discretization, and its heterogeneity level will influence the computational efficiency. The precision of the method depends on the applied ray node spacing, and the number of iterations. The ray node spacing should be chosen carefully as the slowness is supposed to vary linearly between nodes. Again, for an optimal choice of parameters some tests should be performed with different tunings on a representative subset of the data in the specific model.

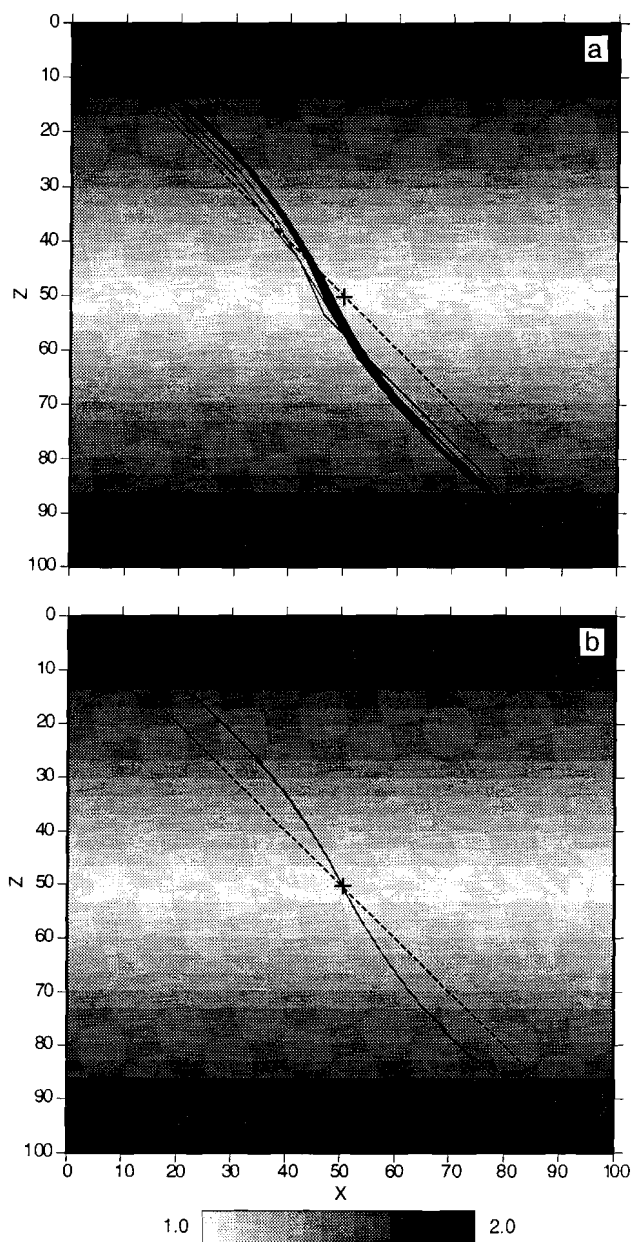
## 7.5 Tests for precision and convergence speed

In order to assess precision and required CPU times and to compare the results of both ray tracing methods, we have performed several tests. We begin with ray tracing in three 2-D, Cartesian models for which analytical paths and travel times exist, then we turn to a synthetic 2-D, spherical model, and finally we test both methods with a substantial part of the data set of Engdahl et al. (1998) in the 3-D, spherical global tomography model of Bijwaard et al. (1998) (Chapter 4). All computations in this section and in section 7.6 have been performed on an adequate but not very special mainframe computer, namely an SGI Power Challenge L that has four R8000 processors.

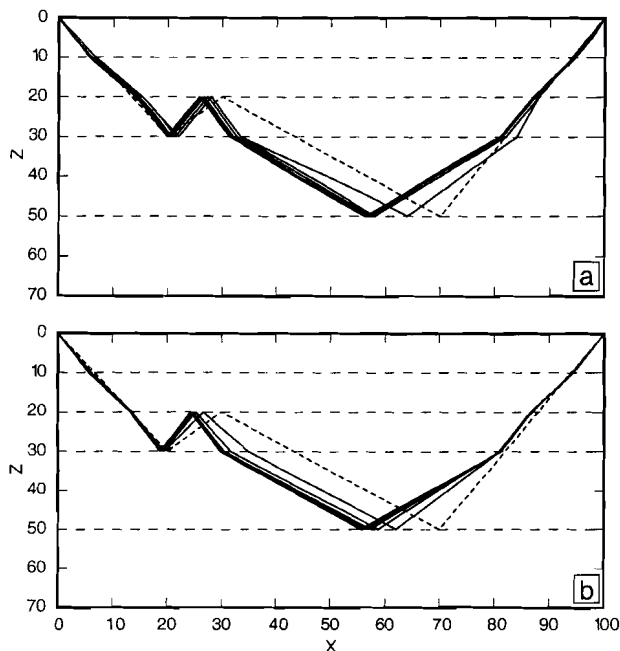
### 7.5.1 A comparison with analytical results

For models in which the velocity varies linearly with depth only, it is possible to calculate travel times and ray paths analytically. As a first example we take a dimensionless 2-D, Cartesian model with  $v(x,z)=1.0+0.02|z-50|$  ( $v$  stands for velocity,  $z$  is depth and  $x$  is the horizontal coordinate). This leads to the low velocity zone that is depicted in Figure 7.3. The analytically calculated (or exact) travel time for a ray starting at (0,0) and ending at (100,100) is 96.2424. Starting with a straight initial path, we arrive numerically at 96.2767 with the graph method (9 iterations and 0.04 s CPU time) and 96.2417 with the perturbation method (10 iterations and 0.07 s CPU) or relative errors of 1:3000 and 1:140000, respectively. For both calculations the ray trace parameters were chosen such as to obtain an accurate result irrespective of the required CPU time. For example, the ray node/graph plane distances decreased during iteration from 10 to 1. The ray path solutions (shown in Figure 7.3) first of all show that the initial ray is bent towards the higher velocities at the top and bottom of the model. Secondly, one can observe that the graph method takes much more iterations to converge towards a final path and that this final path is slightly asymmetric with respect to the center of the model volume. This asymmetry is due to the fact that the graph planes need not be positioned symmetrically with respect to the center (50,50) in subsequent iterations, and it explains why the relative travel-time error is unnecessarily large. The perturbation method traces the approximately correct path in only one iteration, but more iterations further improve the path and the obtained travel time. For a second experiment we use the layered 2-D, Cartesian model depicted in Figure 7.4. We now calculate a multiply reflected ray starting from the dashed initial path. The travel time obtained from bisection for the stationary path is  $32.2116 \pm 0.0001$  and the results for





**Figure 7.3:** A 2-D, Cartesian low velocity zone model [ $v(x,z)=1.0+0.02|z-50|$ ] with a straight initial path (dashed) that is updated with the graph method in (a) and with the perturbation method in (b). The center of the model is indicated by a +. See text for further details.

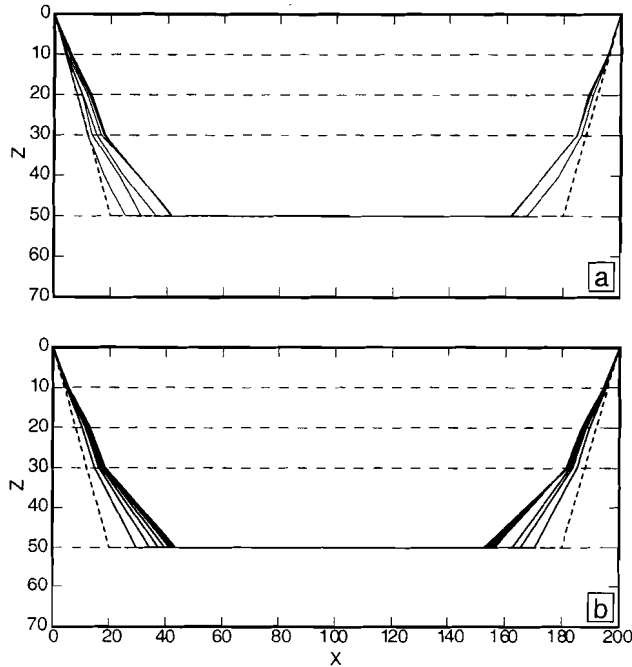


**Figure 7.4:** A layered 2-D, Cartesian model (the homogeneous layer velocities are 4.0, 5.0, 4.0, 6.5, and 8.3 from top to bottom) with a polygonal initial path (dashed) and the solution of the graph method in (a) and with the perturbation method in (b).

the graph and perturbation methods are 32.2122 (9 iterations and 0.04 s CPU time) and 32.2115 (8 iterations and 0.04 s CPU time) or relative errors of 1:54000 and 1:320000, respectively. The ray path solutions in Figure 7.4 show that the different reflection and refraction points quickly shift along the velocity discontinuities towards their final positions, which satisfy the reflection and refraction laws.

Finally, we again take the model of the second experiment, but now we attempt to trace a head wave that is critically refracted at the velocity discontinuity at  $z=50$ . The travel time obtained from bisection for the stationary path is  $39.8780 \pm 0.0001$  and the results for the graph and perturbation methods are 39.8798 (7 iterations and 0.03 s CPU time) and 39.8726 (10 iterations and 0.04 s CPU time) or relative errors of 1:22000 and 1:7400, respectively. The ray path solutions in Figure 7.5 show that both methods now need quite a few iterations to find the right path geometry.

These simple experiments demonstrate that both ray trace methods can accurately determine the travel times of direct, (critically) refracted or (multiply) reflected rays in 2-D, Cartesian media. Although no effort was taken in optimizing computational speed, both methods seem to require little CPU time in these (simple) models.



**Figure 7.5:** Calculation of a head wave in the model of Figure 7.4. The dashed line represents the initial path. Updates are calculated with the graph method in (a) and with the perturbation method in (b).

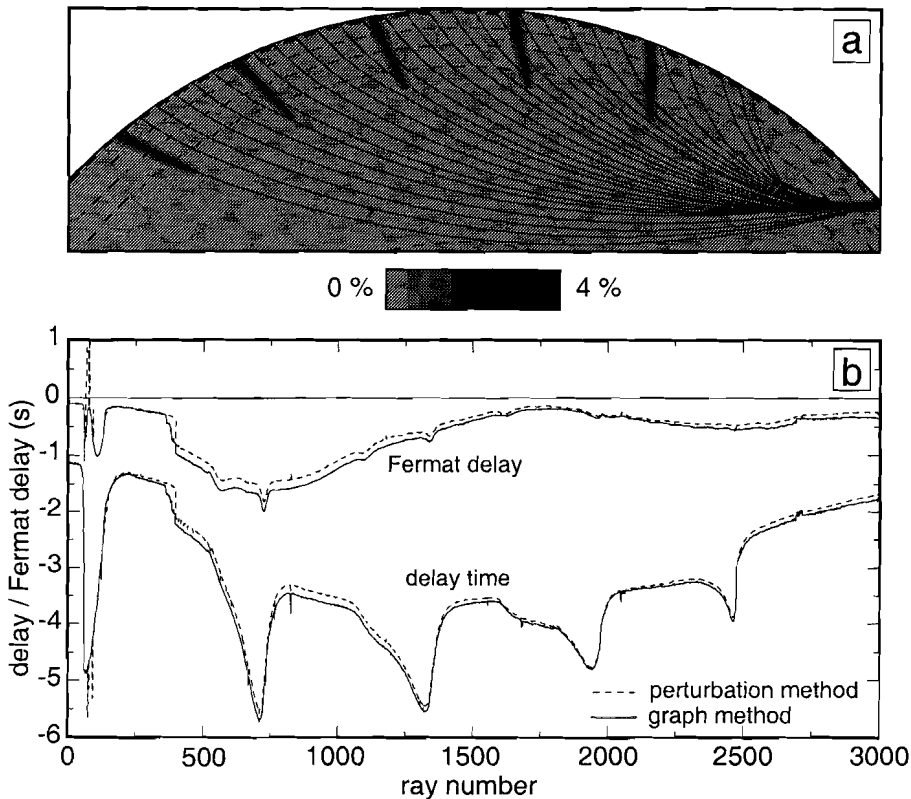
### 7.5.2 A synthetic experiment in a 2-D, spherical model

As a first step toward ray tracing on whole mantle scales, we perform an experiment in a synthetic, spherical model for which no analytical solutions exist. The purpose of this test is twofold: We want to compare the results of the two ray trace methods and simultaneously we wish to get some insight in the travel-time improvement we can obtain with (2-D) ray tracing with respect to simply using the initial (1-D reference model) rays.

A comparison between the two methods is now the only way to get some confidence in their results, because we do not know the real travel times and ray paths. Although both methods have been implemented in the same software framework, the two methods are completely independent, and there is a priori no reason why they should yield the same results.

The travel-time improvement along the final ray path over the travel time along the initial ray in the synthetic model is an indication of the importance of (2-D) ray tracing. Nolet and Moser (1993) named this difference the *Fermat delay* and it is defined as:

$$d_F = \int_L \frac{1}{v} dl - \int_{L_0} \frac{1}{v_0} dl \quad (7.1)$$



**Figure 7.6:** (a) a synthetic, 2-D, spherical model with 30 examples from a fan of 3000 rays. Six slablike velocity anomalies of maximally 4% higher velocity have been inserted. (b) resulting travel-time residuals and Fermat delays (see text for further explanation).

where  $d_F$  is the Fermat delay,  $L$  is the stationary path,  $L_0$  is the initial path, and  $v$  is the (in this case 2-D) seismic velocity in the model space. When the stationary path is a (local) minimum travel-time path (which is usually the case) then the Fermat delay is negative. If it becomes large (and negative) this will indicate that the ray tracing has improved the travel time considerably, which is an a posteriori justification of the computational investment in ray tracing.

The synthetic model we use for this experiment is shown in Figure 7.6a. It consists of the ak135 reference velocities with superposed six slablike anomalies of maximally 4% higher velocity. This particular model geometry is chosen because high velocity slabs pertain to what is actually inferred for the Earth (see for example Bijwaard et al., 1998 and Chapter 4). We shoot a fan of 3000  $P$  rays from an earthquake in one of the slabs, and investigate how the travel-time residual (with respect to ak135) and the Fermat delay vary as the epicentral distance decreases from  $\sim 90^\circ$  to  $\sim 20^\circ$  and the rays shift through 5 other slabs. The results are shown in Figure 7.6b. These first of all show that we obtain very similar results for both ray trace methods. The graph method, however, finds systematically slightly smaller

travel times, and for three rays the perturbation method does not seem to have converged toward a stationary path within the maximum number of iterations allowed (10). This behavior is caused by the fact that the scale and amplitude of the anomalies is at the limit of what can be handled with first order perturbation theory (see Sambridge and Snieder, 1993). These rays can be identified by the oscillating values of their travel times during iteration.

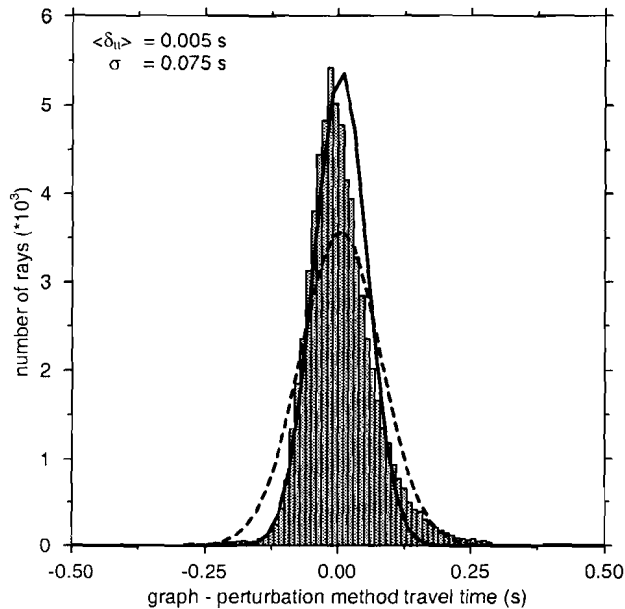
As there are only positive anomalies in the model space, all travel-time delays are negative (down to -6 s). Their pattern clearly shows deep troughs where rays travel through one of the five receiving slabs. Superposed on that is a longer wavelength depression (stretching approximately from ray 400 to 3000) that results from the rays progressively moving through the seismogenic slab. Also indicated are the Fermat delays which represent the amount of delay due to the migration of the ray paths during iteration. We again observe a long wavelength feature now with two lobes that is due to the seismogenic slab. These lobes result from rays initially travelling along the sides of the slab that are being drawn into the higher velocity slab anomaly. Upon close inspection we can also observe such lobes for the other slabs, although these become progressively smaller as the ray direction changes from parallel towards sub-parallel to those slabs.

The Fermat delays range approximately from -0.1 s to nearly -2 s, sometimes amounting to almost 50% of the total delay. This clearly indicates that heterogeneity with amplitudes comparable to those observed for the Earth, combined with a not unrealistic geometry can lead to considerable nonlinear effects on travel times. Hence, 2-D or 3-D ray tracing may significantly alter travel-time residuals and therefore also tomographic models derived from them.

### 7.5.3 Tracing rays in a 3-D, spherical Earth model

As a final test we applied both algorithms to a subset of nearly 70000  $P$  and  $pP$  data from the global data set of Engdahl et al. (1998). The  $P$  phases included nearly 100 critically refracted phases ( $P_n$  and  $P_g$ ) and the total number of reflected ( $pP$ ) phases amounted to nearly 1100. The ray tracing is performed in the global  $P$ -wave tomography model of Bijwaard et al. (1998) (with the ak135 reference model (Kennett et al., 1995)). The purpose of the experiment was threefold: We wanted to estimate the ray trace error by comparing the travel times of the different methods, we needed a reliable indication of the amount of CPU time that is required to trace the entire  $P$  and  $pP$  data set (7.6 million phases) of Engdahl et al. (1998), and we were interested in the size of the Fermat delays due to velocity heterogeneity such as obtained by mantle tomography.

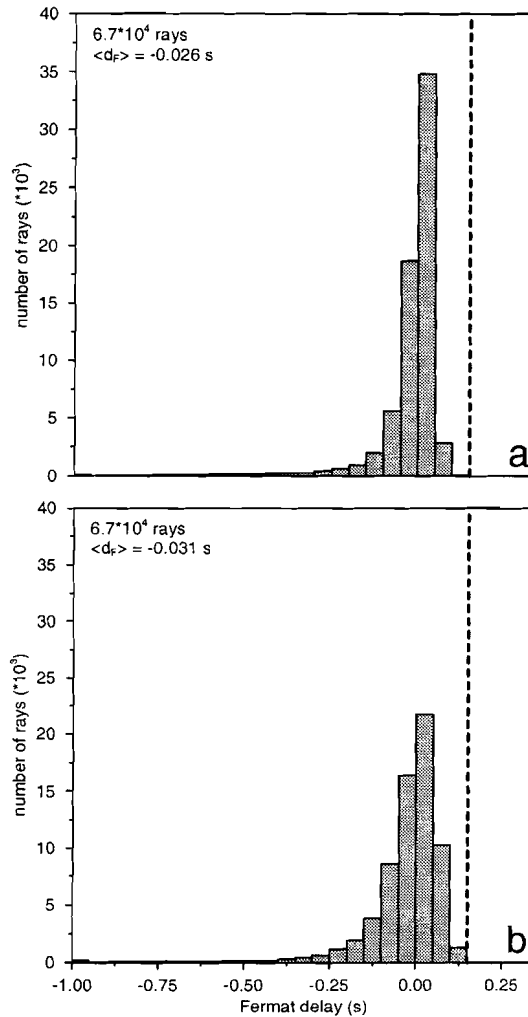
A histogram of the travel-time differences obtained after applying both methods is shown in Figure 7.7. The distribution is slightly asymmetric with more positive travel-time differences caused by slightly larger graph method travel times. This is due to rays that are inferred to have converged (because the convergence criterion of a small travel-time difference between two subsequent iterations is fulfilled), but that would gradually obtain smaller travel times if more iterations were performed. In general the distribution in Figure 7.7 appears to be very similar to a Gaussian with a standard deviation of 0.05 s, and we may use this standard deviation or the standard deviation of the histogram (0.075 s) to estimate the ray trace error. For a safe estimate we take twice the standard deviation or 0.1-0.15 s to



**Figure 7.7:** Histogram of travel-time differences between results obtained with the graph and perturbation methods. The average travel-time difference is denoted by  $\langle \delta_{it} \rangle$ ,  $\sigma$  represents the standard deviation. Two Gaussian distributions with the same surface area are indicated for comparison: The dashed distribution has the same standard deviation as the histogram; the solid distribution has a standard deviation of 0.05 s.

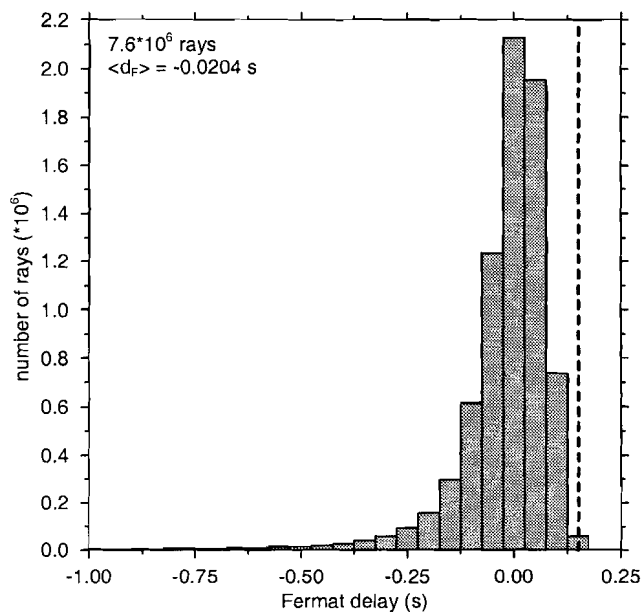
be an adequate measure of the travel-time error we obtain after ray tracing. This value is sufficient for seismological purposes as it is on the order of the error in picking phases from seismograms.

The required CPU time was reduced after experimenting with different sets of ray trace parameters. Mainly the ray node/graph plane distances were varied and for the graph method also the forward star (Moser, 1991). For the perturbation method, node distances were chosen that decreased gradually during iteration from 100 km to 10 km in the lower mantle, from 75 km to 7.5 km in the transition zone, and from 50 km to 5 km in the remainder of the upper mantle. For the graph method similar graph plane distances were taken and in addition the number of nodes in each plane and their spacing were adjusted during iteration from 121 to 25 and from 10 km to 2 km, respectively. In the final iterations the forward star was reduced from 2 to 1, and scaling was introduced to increase the node density in the vicinity of the ray path. For both methods a maximum number of 10 iterations could be performed, but iteration was stopped and convergence was supposed to have been reached as soon as the travel-time difference between two subsequent iterations became less than 0.05 s. It turned out that the CPU time consumed by the perturbation method is in 3-D only slightly larger than in 2-D (the number of ray nodes is similar, but the matrix bandwidth is larger) requiring only 0.07 s per ray. However, for the graph method the amount of CPU time increases dramatically in 3-D, because the number of nodes in the graph is much larger than in 2-D. Its CPU time of 1.8 s per ray makes the graph method too slow to apply to the



**Figure 7.8:** Histograms of the Fermat delays obtained with the graph method (a) and with the perturbation method (b). The thick dashed line is the inferred ray trace error,  $\langle d_f \rangle$  denotes the average Fermat delay.

entire  $P$  and  $pP$  data set. These 70000 rays already took more than 34 hours of computing. Figure 7.8 shows the Fermat delays associated with the 70000 rays for both the graph and perturbation methods. Some of the Fermat delays are positive, which is due to the inferred ray trace error of 0.1-0.15 s. We observe that for most rays the Fermat delay is rather small ( $-0.03$  s on average). This does, however, not imply that these rays hardly move away from their initial path. Rays may migrate over large distances without changing their travel times very much as is shown in Figure A13a. In this case a teleseismic  $P$  ray from Mexico to Scandinavia is drawn into a high velocity anomaly at the base of the mantle and it is shifted by nearly 100 km. The associated Fermat delay is, however, only  $-0.13$  s. In several cases



**Figure 7.9:** Histogram of all Fermat delays obtained with the perturbation method after tracing the entire data set employed by Bijwaard et al. (1998) (Chapter 4) for linearized global tomography (7.6 million rays). The thick dashed line represents the inferred ray trace error,  $\langle d_r \rangle$  again denotes the average Fermat delay.

we do find significant Fermat delays (8% is less than  $-0.15$  s) such as for the ray in Figure A13b, traveling from Samoa in the direction of the Fiji Islands. This ray is again displaced by nearly 100 km and has a Fermat delay larger than  $-2.0$  s. We must note, however, that large Fermat delays and large ray displacements seem to be rather exceptional (0.2% is larger than  $-1.0$  s) for the applied  $P$  and  $pP$  data in the model of Bijwaard et al. (1998) (Chapter 4). It requires a special geometry with respect to ray path orientation, and large amplitudes of the velocity anomalies to achieve a large nonlinear effect. This is for example the case in Figure A13b, where the final ray path is clearly avoiding the lowest velocities and is at the same time drawn into high velocity anomalies on both sides. However, the amplitudes in tomography models are generally slightly underestimated because of the rejection of large residuals that may be falsely regarded as outliers and the applied regularization in the inversion procedure. In the real Earth, Fermat delays and ray displacements may therefore be larger.

## 7.6 Ray tracing for global nonlinear travel-time tomography

For a nonlinear update of a 3-D model obtained in a linearized global tomography study one needs to ray trace all rays that were used in the linearized study. For the linearized tomography study of Bijwaard et al. (1998) (Chapter 4) 7.6 million  $P$  and  $pP$  phases were used from the data set of Engdahl et al. (1998). At a rate of 0.07 s CPU time per ray this



requires 148 hours of CPU time or a little more than 2 days on 3 processors, which is a feasible amount of computation time.

In the process of computing a first nonlinear update of the tomography model, we have performed this ray tracing with the perturbation algorithm using identical parameters as in section 7.5.3. The resulting Fermat delays are shown in Figure 7.9. This figure shows that 3-D ray tracing does not affect the travel times of the majority of global seismic  $P$  and  $pP$  data in the model of Bijwaard et al. (1998) (Chapter 4). However, for 10% of the data we find significant Fermat delays (larger than  $-0.15$  s), and ray displacements may be significant regardless of the size of the Fermat delay. It is therefore likely that a first nonlinear update of the model obtained by linearized tomography will not exhibit an overall dramatic change in amplitude and pattern of heterogeneity, but it may show a significant effect locally, for example in subduction zones.

## 7.7 Conclusions

The main objective of this research was the development of a fast and precise algorithm for the majority of seismic phases that are frequently observed on seismograms. For this purpose, an existing graph, and a perturbation method were extended and adjusted. The most important extension of both techniques consists of the inclusion of reflecting rays, converting phases, some diffracted phases, and head waves. In addition, the perturbation method was extended to 3-D. In our implementation both methods are essentially used as bending techniques that yield locally stationary paths instead of absolute minimum travel-time paths. The graph\_method can also be used to find global minimum travel-time paths. However, the need for complete and exact ray tracing only exists in case the subsurface seismic velocities are known up to high accuracy. In practice, the uncertainties in seismic velocities are not well known and hence complete and exact ray tracing may be precise but at the same time intrinsically inaccurate. In order to considerably reduce the associated problem of multipathing, both techniques have been extended to trace different travel-time branches of seismic phases. Multipathing within the same phase branch is still possible, but this is in practice an observational problem as well, because the uncertainties in seismic velocities prevent an unambiguous association between ray path and observed phase.

A comparison between analytical and numerical results has shown that both algorithms are able to trace different kinds of rays precisely and fast. The experiments in more complicated spherical models have demonstrated that both methods, although completely independent, yield similar results. This strongly argues for the reliability of the outcome of both methods. However, the perturbation method is, in 3-D models, up to 25 times faster than the graph method. Therefore, only the proposed perturbation technique makes the application of large data sets to, for instance, nonlinear tomography, feasible. The resulting error in the travel times for global models is expected to be on the order of 0.1 s. For highly heterogeneous models the graph method should be used because first order perturbation theory may then not be sufficiently precise (see also Sambridge and Snieder, 1993).

We have investigated the nonlinear effect of 3-D ray tracing by means of the Fermat delay. In a (probably underestimated) mantle tomography model the effect can be significant, but only for 10% of the global data set. We therefore expect a nonlinear global tomography model to differ significantly from the model obtained with linearized tomography only in

those places where there is a regionally systematic and large nonlinear effect.

**Acknowledgments.** We are grateful for support from the Netherlands Geosciences Foundation (GOA) with financial aid from the Netherlands Organization for Scientific Research (NWO). We thank Gilles Lambaré, Raúl Madariaga, and Ivan Pšenčík for constructive reviews, Roel Snieder and Hanneke Paulssen for helpful comments, and Jay Pulliam for supplying us with his ray perturbation software. This work was conducted under the program of the Vening Meinesz School of Geodynamics.

## References

- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Červený, V., Ray tracing algorithms in three-dimensional laterally varying layered structures, in: *Seismic tomography*, edited by G. Nolet, pp. 99-133, Reidel, Dordrecht, the Netherlands, 1987.
- Engdahl, E.R., R.D. Van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Amer.*, 88, 722-743, 1998.
- Farra, V., and R. Madariaga, Seismic waveform modeling in heterogeneous media by ray perturbation theory, *J. Geophys. Res.*, 92, 2697-2712, 1987.
- Julian, B.R., and D. Gubbins, Three-dimensional seismic ray tracing, *J. Geophys.*, 43, 95-113, 1977.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from traveltimes, *Geophys. J. Int.*, 122, 108-124, 1995.
- Kennett, B.L.N., S. Widiyantoro, and R.D. Van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, 103, 12469-12493, 1998.
- Koketsu, K., and S. Sekine, Pseudo-bending method for three-dimensional seismic ray tracing in a spherical earth with discontinuities, *Geophys. J. Int.*, 132, 339-346, 1998.
- Lucio, P.S., G. Lambaré, and A. Hanyga, 3D multivalued travel time and amplitude maps, *PAGEOPH*, 148, 449-479, 1996.
- Moser, T.J., Shortest path calculation of seismic rays, *Geophysics*, 56, 59-67, 1991.
- Moser, T.J., G. Nolet, and R. Snieder, Ray bending revisited, *Bull. Seism. Soc. Amer.*, 82, 259-288, 1992.
- Nolet, G., and T.J. Moser, Teleseismic delay times in a 3-D Earth and a new look at the *S* discrepancy, *Geophys. J. Int.*, 114, 185-195, 1993.
- Pereyra, V., W.H.K. Lee, and H.B. Keller, Solving two-point seismic-ray tracing problems in a heterogeneous medium, Part 1. A general adaptive finite difference method, *Bull. Seism. Soc. Amer.*, 70, 79-99, 1980.
- Podvin, P., and I. Lecomte, Finite difference computation of traveltimes in very contrasted velocity models: A massively parallel approach and its associated tools, *Geophys. J. Int.*, 105, 271-284, 1991.

- Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical recipes in FORTRAN, the art of scientific computing, 2nd ed., Cambridge University Press, Cambridge, U.S.A., 1992.
- Pulliam, J., and R.K. Snieder, Fast, efficient calculation of rays and travel times with ray perturbation theory, *J. Acoust. Soc. Amer.*, 99, 383-391, 1996.
- Sambridge, M., and R.K. Snieder, The applicability of ray perturbation theory to mantle tomography, *Geophys. Res. Lett.*, 20, 73-76, 1993.
- Snieder, R.K., and M. Sambridge, Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media, *Geophys. J. Int.*, 109, 294-322, 1992.
- Snieder, R.K., and M. Sambridge, The ambiguity in ray perturbation theory, *J. Geophys. Res.*, 98, 22021-22034, 1993.
- Um, J., and C.H. Thurber, A fast algorithm for two-point seismic ray tracing, *Bull. Seism. Soc. Amer.*, 77, 972-986, 1987.
- Vidale, J.E., Finite-difference calculation of travel times, *Bull. Seism. Soc. Amer.*, 78, 2062-2076, 1988.
- Vidale, J.E., Finite-difference calculation of travel-times in three dimensions, *Geophysics*, 55, 521-526, 1990.
- Virieux, J., Fast and accurate ray tracing by Hamiltonian perturbation, *J. Geophys. Res.*, 96, 579-594, 1991.
- Virieux, J., Seismic ray tracing, in: *Seismic modelling of Earth structure*, edited by E. Boschi, G. Ekström, and A. Morelli, pp. 223-304, Editrice Compositori, Bologna, Italy, 1996.
- Virieux, J., V. Farra, and R. Madariaga, Raytracing for earthquake location in laterally heterogeneous media, *J. Geophys. Res.*, 93, 6585-6599, 1988.

### Appendix A: Inclusion of velocity discontinuities in the perturbation method

In this appendix we formulate the boundary conditions needed to include refracting and reflecting rays in 2-D and 3-D media with velocity discontinuities in the ray perturbation method. Since our perturbation method is based on Pulliam and Snieder (1996), we start from their Equations (28) and (33) (in their notation):

$$r_1(w) \approx a + \frac{w}{W}(b-a) - \frac{w(W-w)}{W} \Gamma_{\perp} t \cdot (b-a) - \frac{\Gamma_{\perp}}{2} u_0(0) w(W-w) \quad (\text{A1})$$

$$\left[ \frac{dr_1(w)}{dw} \right]_{-}^{+} = -u_0(0) [t]_{-}^{+} \quad (\text{A2})$$

where  $r_1$  is the first order perturbation of the ray path,  $a$  and  $b$  are the endpoints of a

perturbed ray segment and  $W$  is its length. The variable position on the ray segment  $w$  is defined by Equation (9) of Pulliam and Snieder (1996),  $\Gamma_{\perp}$  is the component of the slowness gradient perpendicular to the reference path,  $t = dr_0/ds_0$  or the tangential unit vector of the unperturbed path, and  $u_0$  is the 3-D slowness in the medium along the reference path. It is assumed that the slowness varies linearly between  $\mathbf{a}$  and  $\mathbf{b}$  and hence  $\Gamma_{\perp}$  is a constant vector. The + and - signs indicate that Equation (A2) holds across the boundary from one segment to another.

Substituting (A1) into (A2) for segments  $n$  and  $n-1$  and generalizing  $\mathbf{a}$  and  $\mathbf{b}$  for segment  $n$  to  $\mathbf{a}^n$  and  $\mathbf{a}^{n+1}$  leads to:

$$\frac{1}{W_n} \mathbf{a}^{n+1} - \left( \frac{1}{W_n} + \frac{1}{W_{n-1}} \right) \mathbf{a}^n + \frac{1}{W_{n-1}} \mathbf{a}^{n-1} - [X^n(0, \mathbf{a}^n) - X^{n-1}(W_{n-1}, \mathbf{a}^{n-1})] = -u_0^n(0)[t^n - t^{n-1}] - [f^n(0) - f^{n-1}(W_{n-1})] \quad (\text{A3})$$

Where

$$X^n(w, \mathbf{a}^n) = \frac{W_n - 2w}{\Gamma_{\perp}^n W_n} \Gamma_{\perp}^n t^n \cdot (\mathbf{a}^{n+1} - \mathbf{a}^n)$$

$$f^n(w) = -\frac{\Gamma_{\perp}^n}{2} u_0^n(0) (W_n - 2w)$$

The  $X$ -terms between brackets on the left-hand side (l.h.s.) of Equation (A3) are small, and have been omitted by Pulliam and Snieder (1996) in their Equation (38). The l.h.s then reduces to a second order differential operator (in the finite difference sense) on  $\mathbf{a}^n$ . The right-hand side (r.h.s.) balances ray bending (directional change of the tangential vector) with perpendicular slowness gradients (included in the last term). Applying Equation (A3) to all ray segments results in a matrix-vector equation for the  $\mathbf{a}^n$ -vectors. Neglecting the  $X$ -terms, the matrix in this equation is tridiagonal, and we can invert each coordinate of the  $\mathbf{a}^n$  separately.

Although this equation has originally been derived with a cell model in mind, where the  $\mathbf{a}^n$  are displacements of the ray intersections at cell boundaries, it can easily be applied to rays that have been parameterized with nodes. Similar to the implementation for cells, in our implementation the  $\mathbf{a}^n$  displace the ray nodes which are the boundaries between ray segments.

Equation (A3) only holds for smoothly varying slownesses. For refractions at slowness discontinuities we have to implement Snell's law. In other words: The slowness  $u_0$  is no longer continuous across the boundary between two segments joining at an interface. This leads to a modification of Equation (A2):

$$\frac{dr_1}{dw}^+ = -[u_0(w)t]^- \quad (\text{A4})$$

where we have used a similar approximation as employed by Pulliam and Snieder (1996) to introduce the variable  $w$ , viz.  $u_0(w)/u_0(0) \approx 1$ , which implies that the velocity variation over a single segment should be small. From Equations (A4) and (A1) we can derive an equation for the  $\mathbf{a}^n$  at velocity discontinuities. However, we need to impose two extra conditions at interfaces: First,  $\mathbf{a}^n$ -vectors at interfaces must be parallel to them, because refraction nodes must always shift along the interface. This can be translated into (putting nodes  $n-1$  and  $n$  arbitrarily close to, but on different sides of the interface):

$$(\mathbf{n}_d \cdot \mathbf{a}^{n-1}) = (\mathbf{n}_d \cdot \mathbf{a}^n) = 0 \quad (\text{A5})$$

where  $\mathbf{n}_d$  denotes the unit normal of the discontinuity. Second, the  $\mathbf{a}^n$  must be continuous across the interface, leading to:

$$\mathbf{r}_0^n + \mathbf{a}^n = \mathbf{r}_0^{n-1} + \mathbf{a}^{n-1} \quad (\text{A6})$$

Equation (A6) renders one of the parts of Equation (A5) redundant. Furthermore Equation (A4) is now only needed for the tangential component of the displacement vectors and hence becomes:

$$\mathbf{t}_d \cdot \left[ \frac{dr_1}{dw} \right]^+ = -\mathbf{t}_d \cdot [u_0(w)t]^- \quad (\text{A7})$$

where  $\mathbf{t}_d$  denotes the unit tangential vector to the discontinuity (in 3-D there are of course 2 independent tangential vectors). Substituting Equation (A1) into (A7), generalizing  $\mathbf{a}$  and  $\mathbf{b}$  for segments  $n$  and  $n-2$  and neglecting segment  $n-1$  across the discontinuity (which is negligibly small) gives:

$$\begin{aligned}
\mathbf{t}_d \cdot \left[ \frac{1}{W_n} (\mathbf{a}^{n+1} - \mathbf{a}^n) - \frac{1}{W_{n-2}} (\mathbf{a}^{n-1} - \mathbf{a}^{n-2}) - (X^n(0, \mathbf{a}^n) - X^{n-2}(W_{n-2}, \mathbf{a}^{n-2})) \right] = \\
\mathbf{t}_d \cdot \left[ -(u_0^n(0) \mathbf{t}^n - u_0^{n-2}(W_{n-2}) \mathbf{t}^{n-2}) - (f^n(0) - f^{n-2}(W_{n-2})) \right]
\end{aligned} \tag{A8}$$

The dotproducts in Equations (A5) and (A8) couple the components of the  $\mathbf{a}^n$ -vectors. Therefore, the matrix that determines the  $\mathbf{a}^n$  is no longer tridiagonal and the components cannot be determined separately. This results in a three times larger matrix with a larger bandwidth when discontinuities are included in the model space.

With Equations (A5), (A6) and (A8) we can calculate refracting rays (including phase conversions), and head waves (that are critically refracted), but we also want to incorporate reflections. For this purpose we rewrite the familiar reflection law as:

$$\begin{aligned}
\left[ \mathbf{n}_d \cdot \frac{d\mathbf{r}}{ds} \right]^b &= - \left[ \mathbf{n}_d \cdot \frac{d\mathbf{r}}{ds} \right]^a \\
\left[ \mathbf{t}_d \cdot \frac{d\mathbf{r}}{ds} \right]^b &= + \left[ \mathbf{t}_d \cdot \frac{d\mathbf{r}}{ds} \right]^a
\end{aligned} \tag{A9}$$

Where the superscripts denote before (b) and after (a) reflection and  $\mathbf{r} = \mathbf{r}_o + \mathbf{r}_l$  is the ray position with arc length  $s$ .

Assuming the reflection to take place between nodes  $n-1$  and  $n$  (both slightly offset from the discontinuity for computational reasons), we can again apply Equations (A5) and (A6). This makes the first equation of (A9) redundant, because Equations (A5) and (A6) fully determine the perpendicular component. To incorporate phase conversions, like *PcS* phases, we can again not take the slowness continuous. This leads to an equation analogous to Equation (A7):

$$\mathbf{t}_d \cdot \left[ \frac{d\mathbf{r}}{dw} \right]_a^b = - \mathbf{t}_d \cdot [u_0(w) \mathbf{t}]_a^b \tag{A10}$$

After substitution of Equation (A1) (which is straightforward and leads to an equation very similar to (A8)), this equation, together with Equations (A5) and (A6), allows us to calculate reflecting ray paths.

In summary the procedure is as follows: For a specific initial ray path all refraction/reflection points are determined. For these nodes (two per point) Equations (A5), (A6) and (A8) (or a similar version of (A10)) are applied; for all ordinary nodes Equation (A3) is used. This leads to a matrix-vector equation for the node displacements  $\mathbf{a}^n$ , which can be solved with a banded matrix inversion (Press et al., 1992). Once the  $\mathbf{a}^n$ -vectors are known  $r_I$ -vectors between ray nodes can be calculated with Equation (A1). The travel time is then calculated by applying Equations (41), (42) and (43) from Pulliam and Snieder (1996).

# Chapter 8<sup>1</sup>

## Nonlinear global *P*-wave tomography

**Abstract.** Recent advances in global imaging have lead to tomographic mantle models with regional scale details. To further improve these models, we have extended the usual linearized approach to travel-time tomography to nonlinear tomography. Here ‘nonlinear’ means that seismic ray bending due to inferred velocity heterogeneity is taken into account in an iterative method in which inversion steps are alternated with 3-D ray tracing to update ray paths and travel times. As a starting point for our nonlinear inversion we have used the mantle model of Bijwaard et al. (1998) (Chapter 4) and ray tracing is performed following Bijwaard and Spakman (1999a) (Chapter 7).

The main results are the following. We observe no overall dramatic change in anomaly patterns, but subtle changes on the global mantle scale lead to a small increase in variance reduction and model amplitudes. These small changes together with very similar resolution estimates for the linear and nonlinear inversion, do not allow us to formally investigate possible model improvements. However, expected nonlinear effects such as focussing of structures and a baseline shift towards lower velocities indicate an improved solution. Apart from that, some very strong changes occur in distinct upper mantle regions, such as below Japan, Tibet, South America, Europe, and Tonga-Fiji, where 3-D ray bending effects are substantially larger than in the deeper mantle since model amplitudes fall-off rapidly with depth. In the lower mantle, however, increased focussing effects can be observed that may prove important for detailed interpretations.

### 8.1 Introduction

In the last few years significant progress has been made in globally imaging the Earth’s mantle with travel-time tomography. Models derived by Zhou (1996) and Van der Hilst et al. (1997) employ substantially more detailed parameterizations than before (e.g. Inoue et al., 1990; Zhang and Tanimoto, 1993; Su et al., 1994; Vasco et al., 1995; Masters et al., 1996; Trampert and Woodhouse, 1996). Particularly, the model of Bijwaard et al. (1998) (Chapter 4) (hereafter BSE) shows that for particular regions global imaging has reached the level of detail of regional tomographic mantle investigations (50-100 km) (e.g. Spakman et al., 1993; Van der Hilst et al., 1993).

When solving for ever smaller scale structure in the Earth’s mantle, the linearization of the tomographic problem, i.e. the use of reference ray paths and travel times calculated in an essentially 1-D background velocity medium, may become more and more inadequate because of the possible violation of the basic linearization assumptions. Up to now, all global tomography models have been derived in a linearized fashion, thereby neglecting the ray bending induced by the inferred 3-D velocity heterogeneity. However, in many local and regional crust and lithosphere studies the intrinsic nonlinear tomography problem was

---

<sup>1</sup>This chapter has been submitted to *Geophys. J. Int.* as: Bijwaard, H., and W. Spakman, *Nonlinear global P-wave tomography*.



(already in the early 1980s) tackled with an iterative scheme in which inversions of travel-time residuals for model updates were alternated with ray tracing steps, in which ray bending due to inferred heterogeneity was taken into account (e.g. Thomson and Gubbins, 1982; Thurber, 1983; Koch, 1985; Sambridge, 1990; Zhou, 1990; VanDecar, 1991; Zhao et al., 1992; Souriau and Granet, 1995; Papazachos and Nolet, 1997). A logical next step in improving global modelling would therefore be to adapt such a nonlinear inversion scheme to a whole mantle study (as recommended by Kennett et al., 1998). This would of course require in each iteration step the major computational task of 3-D ray tracing of the entire data set, which often consists of several millions of phases.

A global model obtained from a nonlinear approach is not expected to demonstrate a dramatic global change in the resolved heterogeneity with respect to previous linear inversions. However, systematic changes due to ray bending may occur in regions with a sufficient level of velocity heterogeneity (amplitude) and are therefore mainly expected in the upper mantle part of global models where amplitudes are generally significantly higher than in the lower mantle (e.g. Inoue et al., 1990; Gudmundsson et al., 1990; Vasco et al., 1995; Bijwaard et al., 1998).

In our approach we use the model obtained by BSE from linearized inversion as a starting point for a first, straightforward attempt to perform a nonlinear global inversion. The entire data set used by BSE, consisting of 7.6 million  $P$ ,  $pP$  and  $pwP$  phases, is traced through the BSE model using a ray perturbation approach (Bijwaard and Spakman, 1999a) (Chapter 7). New ray paths and travel-time residuals then form a new matrix-vector equation which is subsequently solved in a similar fashion as in BSE. This process is repeated once to obtain convergence estimates.

## 8.2 Data and ray tracing

As the nonlinear approach is based on the linearized inversion of BSE, we use exactly the same data set. The data consist of the  $P$ ,  $pP$  and  $pwP$  phases in the reprocessed data set of Engdahl et al. (1998) with travel-time residuals between -3.5 and 3.5 s for teleseismic rays and between -7.5 and 7.5 s for regional ray paths. This amounts to a total of 7.6 million phases that are subsequently combined into 4.7 million ray bundles which emerge from  $30^3$  km<sup>3</sup> source volumes and arrive at single stations (see BSE for details).

BSE use an irregular cell parameterization and invert simultaneously for cell slownesses (slowness being the reciprocal of velocity), station corrections and earthquake cluster relocations. They conclude from a set of ground truth events (for which locations are rather well known) that the average event location errors were reduced by 18% due to the calculated cluster relocations. It might therefore be argued that 3-D ray tracing should be performed from the relocated event positions. There are, however, several reasons not to do so. First of all, the relocation vectors have been determined for earthquake clusters and they may not be optimal for single events. The fact that they turn out to be reasonable for a small and perhaps not representative subset of ground truth events does not justify their application to the entire set of earthquakes. Second, these relocations have been determined using only  $P$ ,  $pP$ , and  $pwP$  phases and important phases for event positioning such as  $S$ ,  $PKiKP$  and  $PKPdf$  have not been included. Therefore, they may actually impair the accuracy of the Engdahl et al. (1998) locations (5-10 km). Finally, the relocation vectors are calculated from

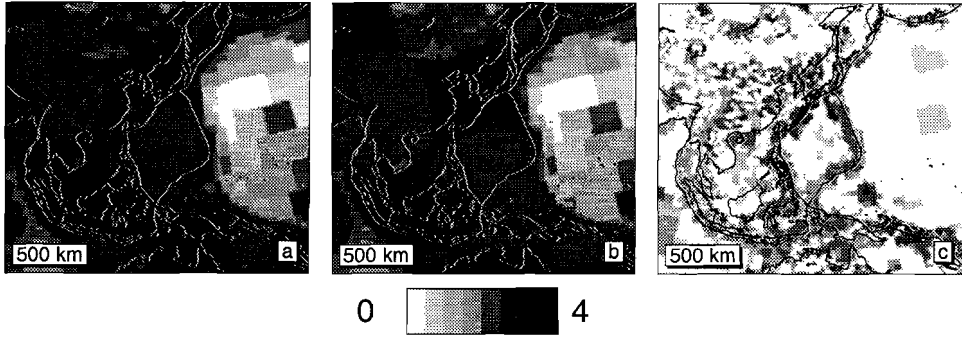
travel-time gradients, which are local, linear estimates that only hold in the direct vicinity of the hypocenter. However, in many cases we observe that earthquake locations are, for example, shifted across crustal discontinuities, for which these gradients are a poor approximation. For a theoretically more valid estimation of cluster relocation vectors these gradients should be calculated in a realistic 3-D Earth model. Multiple 3-D ray tracing evaluations per event cluster are then required for the computation of these gradients. We have therefore chosen to use the original hypocenter locations (Engdahl et al., 1998) as the starting points for 3-D ray tracing and to calculate new cluster relocation vectors in the next inversion. The latter will at least partially accommodate the approximation made, and still allows to solve for event specific signal in the inversion.

For the 3-D ray tracing the perturbation algorithm of Bijwaard and Spakman (1999a) (Chapter 7) is applied. This algorithm is based on the perturbation theory developed by Snieder and Sambridge (1992) and by Pulliam and Snieder (1996), and extends their approach to ray tracing in realistic global Earth models. The method essentially acts as a bending technique starting from the 1-D reference rays and searching for nearby minimum travel-time paths. The main advantage of the algorithm of Bijwaard and Spakman (1999a) is the ability to trace all sorts of seismic phases (i.e. including refracted, reflected, converted and diffracted rays at first order discontinuities) precisely (0.1-0.15 s error) and rather fast (15 rays/s on an SGI R8000 processor). The 3-D ray tracing is performed through a grid version of the BSE model, with nodes at the centers of a regular grid of  $0.6^\circ$  cells, and velocities that are linearly interpolated in between, thereby ignoring amplitude discontinuities between cells. The entire data set of 7.6 million phases was traced in less than 3 days (on 3 processors), and significant nonlinear travel-time effects were found for 10% of the data (see Bijwaard and Spakman, 1999a, and Chapter 7). The geometry of the majority of seismic ray paths is not severely altered, and the hitcount pattern of ray sampling in most of the Earth's mantle is similar to the pattern used for linearized tomography (see Figure 8.1). However, in the upper mantle locally strong differences exist.

### 8.3 Parameterization and inversion

In the linearized tomography study of BSE the Earth's mantle was parameterized with irregular cells, the sizes of which depended on the local ray sampling (Spakman and Bijwaard, 1998). Small cells were implemented in regions with a high ray density and large cells in regions with a low ray density. As the hitcount pattern of the 3-D ray paths is similar to the 1-D case, there is no reason to change the mantle parameterization for nonlinear inversion. Furthermore, an identical parameterization will facilitate a direct comparison between the results from linear and nonlinear inversion. We therefore employ the same irregular grid as BSE with smallest cells of  $0.6^\circ \times 0.6^\circ \times 35$  km in the upper mantle ranging up to largest cells of  $9^\circ \times 9^\circ \times 400$  km in the poorest sampled mantle regions.

The ray paths and travel times derived by 3-D ray tracing can be put into a matrix-vector equation, which needs to be inverted to obtain cell slownesses, relocation vectors, and station corrections. Independent of numerical least squares solvers (e.g. SVD or LSQR), it appears that different iterative inversion strategies can be used to solve such large size inverse problems (see review by Spakman, 1993). These mainly differ in the particular model being solved for in each iteration step, and in the reference model relative to which the solution



**Figure 8.1:** An example of the logarithmic hitcount pattern for southeast Asia at 500 km depth (determined on the parameterization that was used for inversion). Panel (a) shows hitcount for linearized inversion, (b) hitcount for nonlinear inversion, (c) change in hitcount ((b)-(a)). Notice that hitcount in (c) is generally increased in the subduction zone regions.

is regularized in the inversion. One can either solve for an incremental anomaly model that has to be added to the anomaly model from the previous inversion step (in our case the BSE anomaly model after one ray tracing step), or for the total anomaly with respect to the initial reference model (in our case 1-D reference model ak135 (Kennett et al., 1995) for all iteration steps), or even for the absolute slownesses (having zero reference model in all steps). In the inversion, one can combine most of these solution models with regularization of the increment, the total anomaly, or the absolute slownesses. What is being regularized effectively depends on the combination of the reference model for regularization with the particular model for which will be solved. This combination also determines how the original data have to be corrected before each iteration step.

All inversion strategies can be summarized in one equation for the cost function to be minimized in iteration step  $k$ :

$$\Phi(m_{k+1}) = (d_k - A_k m_{k+1})^T W^{-1} (d_k - A_k m_{k+1}) + \lambda^2 (m_{k+1} - m^0)^T C^T C (m_{k+1} - m^0) \quad (8.1)$$

where  $d_k$  is the data vector,  $A_k$  the tomographic coefficient matrix based on rays in model  $k$ ,  $m_{k+1}$  the solution vector ( $k=0$  corresponding to rays in ak135 and solution  $m_1 = \text{BSE}$ ),  $W$  the data covariance matrix,  $m^0$  the reference model for regularization (in our case ak135), and  $C^T C$  a damping operator. The constant  $\lambda$  controls the trade-off between minimizing the data misfit and the scaled model norm, and  $\lambda=0$  leads to normal least squares. The model  $m_{k+1}$  that minimizes  $\Phi(m_{k+1})$  can also be recognized as the least squares solution of the following matrix system:

$$\begin{pmatrix} W^{-1/2} A_k \\ \lambda C \end{pmatrix} m_{k+1} = \begin{pmatrix} W^{-1/2} d_k \\ \lambda C m^0 \end{pmatrix} \quad (8.2)$$

Spakman (1993) discusses six combinations of solution models and regularization reference models which have appeared in the geophysical literature, and which in theory prove to lead to three classes of different solutions. From synthetic tests on 2-D mantle slices (Spakman and Van Deelen, 1993), it proved that the two inversion strategies of Tarantola and Valette (1982) perform best in solving the nonlinear travel-time problem. One strategy (Equation (25) of Tarantola and Valette (1982), hereafter the 'incremental method') solves for the increment and regularizes the total anomaly with respect to the initial reference model. In this case the residual is corrected before every inversion step as follows (Spakman, 1993):

$$d_k = d_0 - (T_k(s_k) - T_0(s_0)) \quad (8.3)$$

where  $d_k$  is the delay for iteration  $k$ ,  $d_0$  the original delay,  $T_k(s_k)$  the travel time in the slowness model  $s_k$  obtained in iteration  $k$  along the (3-D) ray path in this model, and  $T_0(s_0)$  the travel time in reference model  $s_0$  (ak135) along the reference ray path.

The other strategy of Tarantola and Valette (1982) (their Equation (24)), hereafter the 'total anomaly method', solves for and penalizes the total anomaly in each iteration. In this case the residual correction is:

$$d_k = d_0 - (T_k(s_0) - T_0(s_0)) \quad (8.4)$$

where  $T_k(s_0)$  represents the travel time computed through the reference slowness model but along the 3-D ray path of iteration  $k$ . Thus, only the change in initial reference model travel times due to path geometry is transferred to the new travel-time residual.

In theory (i.e. with ideal numerical solvers and also in the absence of ray trace errors), both strategies are equivalent and lead to the same solution. In their synthetic tests, conducted with LSQR, Spakman and Van Deelen (1993) found small differences between results of both approaches, although both proved superior with respect to other strategies. From our tests with the incremental method and the total anomaly method we obtained somewhat less rough models and stronger focussing effects with the total anomaly method, which we qualitatively attribute to the following practical reasons.

First, because of the large size of our inverse problem we can only compute an approximate solution to the system of equations. Therefore starting from ak135 (i.e. from a zero anomaly model) in each iteration step may reduce the nonlinear propagation of incorrigible solution artifacts from one step to another. An opposite effect could occur in case of the incremental method, where we would build solution increment on increment. Hence, to reduce nonlinear error propagation we only allow this to occur through 3-D ray tracing (i.e. through the ray path bending) and avoid this to occur through adding (approximate) incremental solutions. Second, before each iteration the data vector needs to be corrected for the 3-D solution obtained so far. In the incremental method the delay times are corrected by the difference between the travel time in the 3-D model acquired along the 3-D ray path and the ak135 travel time (Equation 8.3). Ideally, the corrected delay times would converge to a noise vector in the final stages of iteration and the signal-to-noise ratio (S/N) decreases during iteration. In case of the total anomaly strategy the correction applied to the ak135 residuals is the difference between the travel time in ak135 integrated along the 3-D ray path and the

ak135 travel-time (see Equation 8.4). This correction term does not change strongly during iteration and is generally smaller than the correction term in the incremental method, because it only depends on the smooth ray bending geometry and not on the much more spatially variable 3-D velocities encountered along the ray. Therefore, in the total anomaly method the S/N ratio does not change strongly between iteration steps and nonlinear error propagation is kept to a minimum (i.e. propagation occurs through smooth 3-D ray geometry and not through 3-D travel times).

Third, and tied to the previous reason, our approximate least squares solver (LSQR) constructs its search directions in the model space explicitly from information on the data vector. This is essentially different from full matrix solvers like SVD, which construct eigenvectors in the solution space independent of the data. With the total anomaly method we can construct solutions from data with comparable S/N ratio in every iteration, which are probably least (as compared to the incremental method) affected by nonlinear error propagation, and thereby we may avoid LSQR-search directions partially based on propagated errors.

Thus, although in theory both strategies of Tarantola and Valette (1982) lead to the same solution, we believe that in practice (using approximate solvers) we can reduce nonlinear error propagation by adopting the total anomaly method.

The matrix-vector equation for nonlinear inversion contains 418000 unknowns (277000 cell slownesses together with cluster relocation parameters and station corrections) and 4.7 million composite ray residuals. This system of equations is augmented with second derivative damping equations identical to those of BSE (their Equation (3)), which act on the total anomaly model. Similar to BSE, we use 200 iterations of the LSQR algorithm of Paige and Saunders (1982) to calculate an approximate solution.

## 8.4 Results

The solution to the tomography problem discussed here is obtained after two nonlinear iteration steps (starting from the BSE model), which take approximately one week computation time each. We did not consider a third iteration step because of the very small additional change in the second step, which indicates that the iteration has sufficiently converged (formal correlation between the results from first and second iteration amounts to 0.99, whereas it is 0.89 between BSE and the first iteration). We shall in the following refer to the solution after two nonlinear steps as BSE\_NL.

### 8.4.1 Data fit

The variance reduction that is obtained after two nonlinear steps amounts to 57.2 %, which is only marginally higher than obtained after the linearized inversion of BSE (57.1 %). This small additional variance reduction is probably negatively influenced by the 3-D ray tracing error of  $-0.1-0.15$  s (see Bijwaard and Spakman, 1999a) (Chapter 7), which is 11-17% of the final root-mean-square residual and is of the same size in each nonlinear iteration. Apart from this, one should realize that small changes in a bulk parameter such as variance

reduction can still result from systematic changes in distinct parts of the model (such as the uppermost mantle). To illustrate this: The total variance reduction (i.e. both from ray clustering and inversion) obtained from the same data set by Van der Hilst et al. (1997) is only slightly less than obtained by BSE. However, very important global differences exist in both the small-scale structures in the upper mantle (see BSE) and some of the larger scale structures in the lower mantle. Apart from that, the sensitivity of travel times for small-scale global structure is small. For instance, synthetic delays computed from a whole mantle model consisting of regularly spaced  $2.4^\circ$  spikes (spike spacing  $3^\circ$ - $4^\circ$ ) with alternating amplitudes of  $\pm 1\%$  have a root-mean-square (rms) residual of only 0.27 s. For a similar model constructed with  $1.2^\circ$  spikes this is 0.18 s. For  $\pm 0.5\%$  spike amplitudes (corresponding to peak values imaged in most of the lower mantle) these rms values would be 0.13 and 0.09 seconds, respectively. These values are computed from the linearized system (Equation 8.2) and hence do not include the effects of 3-D ray tracing through the spike model. This implies that resolving small-scale structure (or model changes in nonlinear tomography) on global scales will only marginally improve the variance reduction.

Apart from a small change in variance reduction there is a relative increase of 7.0% in the overall anomaly amplitudes in the nonlinear solution with respect to BSE. This represents an important change regarding the fact that the BSE model is a damped, minimum norm solution that may underestimate the actual heterogeneity amplitudes in the Earth's mantle. Sensitivity estimates for the BSE model indicate that in well-imaged regions the model amplitudes may be underestimated by 20-30%. This is worse in other regions and hence will generally lead to underestimated ray bending and associated travel-time changes.

Basically, we found a new model that fits the data comparably well as for the BSE model. Further, formal co-variance and spatial resolution cannot be computed for such large inverse problems. Estimates derived from permuted data tests and sensitivity analyses indicate that the model variance and resolution are highly comparable to that of BSE (see below). Hence, we have no formal means to investigate the quality of the new model or the differences with the BSE model.

#### 8.4.2 Resolution

Before turning to detailed interpretations of the anomalous velocity structure obtained from the nonlinear tomography, we need to address the resolution aspects of our modeling. Similarly to BSE, the resolving power of the (now 3-D) ray geometry is tested against synthetic spike and layer-cake models, because formal resolution cannot be computed for such a huge inversion problem. Travel times and residuals are calculated for these synthetic models and Gaussian distributed noise (0.5 s standard deviation) is added to mimic observational errors in the real data. The comparison of the output of an inversion of these synthetic data with the original synthetic input model yields an idealized estimate of the formal resolution in the model. Spike tests have been performed with spikes of nearly all cell sizes in the parameterization, ranging from  $0.6^\circ$  to  $6.0^\circ$ , such as to give a reliable indication of the scale of heterogeneity structure that can be retrieved in different parts of the Earth's mantle. In general, the obtained sensitivity results are very similar to those of BSE. We do not regard the small differences discriminative between the models. An example of the input and output of a sensitivity test and a comparison with BSE is shown

in Figure 8.2. In the following we shall only interpret structural anomalies for which synthetic tests at the appropriate scale length indicate that they are well-retrieved, unless indicated otherwise.

#### *8.4.3 Cluster relocation vectors and station corrections*

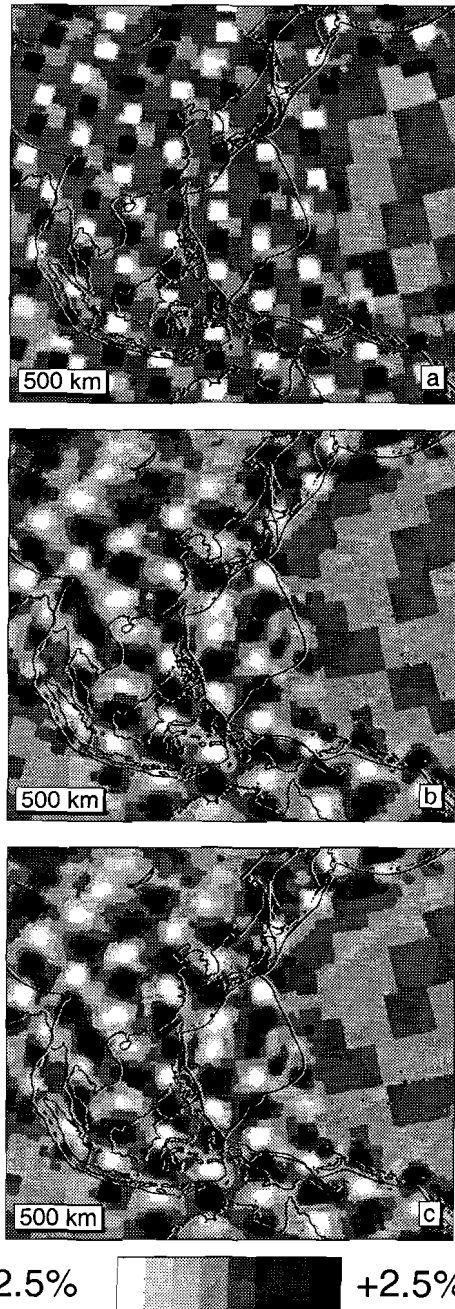
The solution to the tomographic problem consists of cell slownesses together with cluster relocation vectors and station corrections. The obtained cluster relocation vectors are very similar to the ones obtained from the linearized inversion (BSE). They are regionally systematic and serve to take out data errors and errors associated with hypocenter locations, which can amount to 10 km or more. As was mentioned above they are probably of limited value, because they have been derived from travel-time gradients determined in ak135, represent whole clusters of earthquakes, and have been determined without several important seismic phases.

The computed station corrections correlate well with the results of BSE and with average station residuals before inversion, and hence mainly serve to remove the latter from the slowness part of the solution. For stations with good azimuthal coverage they may represent shallow (crustal) structure directly beneath the station.

#### *8.4.4 The 3-D model obtained from nonlinear tomography*

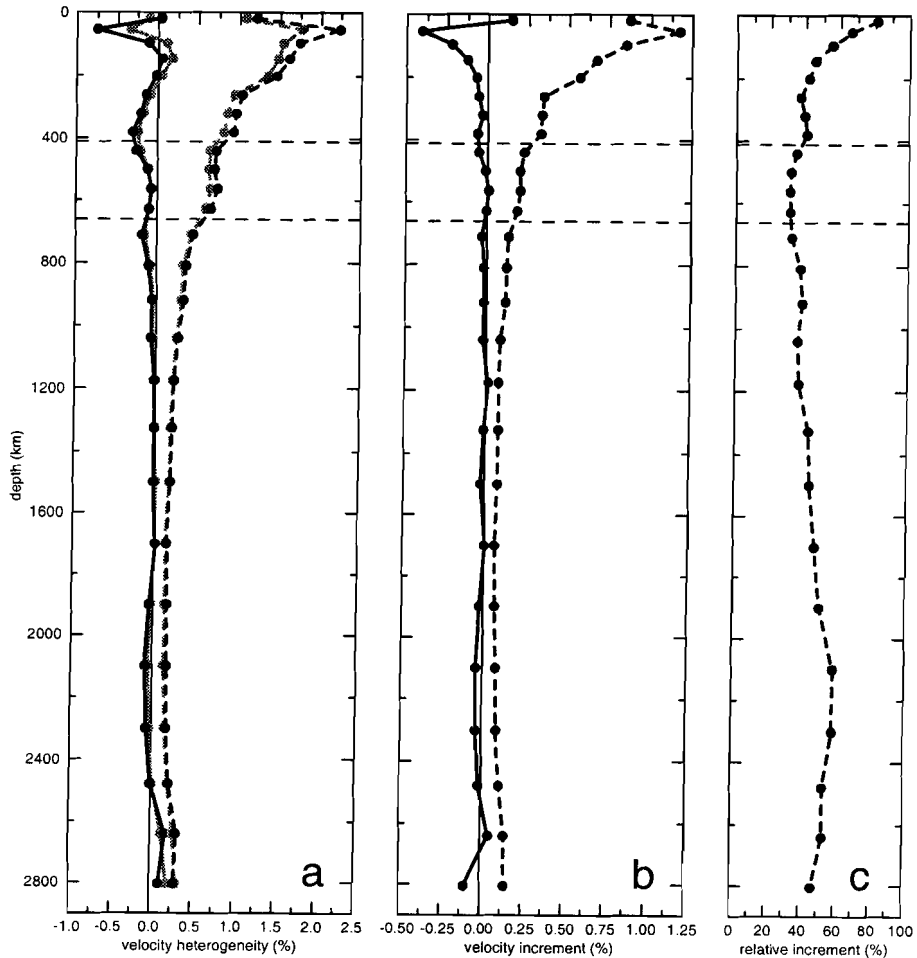
In the velocity part of the solution we observe a distinct change with respect to BSE (see Figure 8.3). The difference between BSE\_NL and BSE (hereafter the ‘increment’) in Figure 8.3b shows that the most significant changes in root-mean-square anomaly amplitudes take place in the upper mantle part of the model. However, the relative change in amplitudes (Figure 8.3c) lies between 30% and 80% for the entire mantle with lower mantle values reaching 60%. The observed change consists of (1) an overall shift towards more negative anomalies for nearly all depths in the model, but strongest for the uppermost 200 km of the mantle, and (2) incremental heterogeneity with a large standard deviation in the crust and mantle decreasing rapidly above the transition zone. Effect (1) partially corrects the bias that was introduced by the use of 1-D reference ray paths in the linearized inversion. These reference paths lead to relatively overestimated sizes of positive anomalies, because 3-D ray bending actually focusses ray paths in high-velocity zones (Wielandt, 1987; Nolet and Moser, 1993). The correction toward lower velocities can be clearly observed in three regions where strong negative anomalies are imaged in the nonlinear steps: Tibet, Japan, and South America (see Figure 8.4). We shall discuss these regions in more detail below. Effect (2) is due to the fact that ray bending and hence the nonlinearity of the tomographic problem depends heavily on the anomaly amplitudes. Because the amplitude of imaged heterogeneity in BSE and other global models based on travel times falls off rapidly with depth (see e.g. Van der Hilst et al., 1997; Vasco et al., 1998) to maximum values on the order of 0.5% in the mid-lower mantle, 3-D ray path geometry is much less affected below the uppermost mantle regions.

One may argue that even though lower mantle heterogeneity itself hardly affects the seismic rays, the upper mantle structure may have such an influence on teleseismic paths, e.g. by



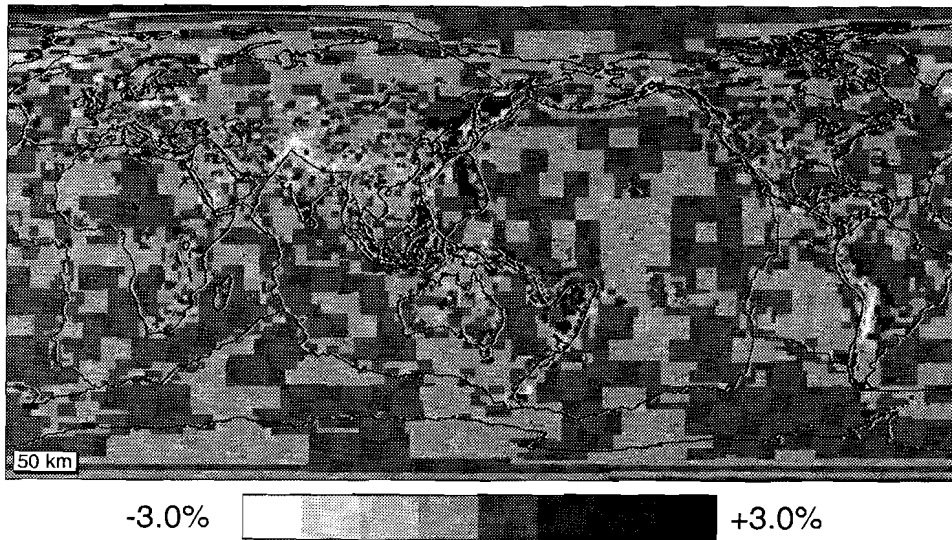
**Figure 8.2:** Example of input (a) and output for BSE (b) and BSE\_NL (c) of a synthetic spike test (3°, 5% spikes) at 500 km depth beneath southeast Asia. The fairly good retrieval of most of the input pattern below southeast Asia is considered to be indicative of good spatial resolution for this region. Note that the differences between the BSE and BSE\_NL solutions are very small.





**Figure 8.3:** Three panels showing the variation of linear (BSE) and nonlinear (BSE\_NL) solutions with depth for regions with hitcount values above 1000. Model averages are depicted by thick lines with dots and root-mean-square (rms) values by thick dashed lines with dots. Panel (a) shows the linear solution in gray and the nonlinear solution in black; panel (b) shows the increment (nonlinear result - linear result) at half the scale of panel (a); and panel (c) shows the ratio of the rms values of the increment over the rms values of BSE in percentages. Thin dashed lines denote mantle discontinuities at 410 and 660 km depth.

changing their position and angles near 660 km depth that these also deviate systematically from their 1-D equivalents in the lower mantle part of their trajectories. This does occur, but deviations are mostly on the order of a few tens of kilometers. Large lower mantle ray path deviations (on the order of 100 km) are rather exceptional (see Bijwaard and Spakman, 1999a, and Chapter 7). We shall therefore in the following concentrate on the solution in the upper mantle and briefly go into some of the more subtle changes we observe in the lower mantle part of the model.



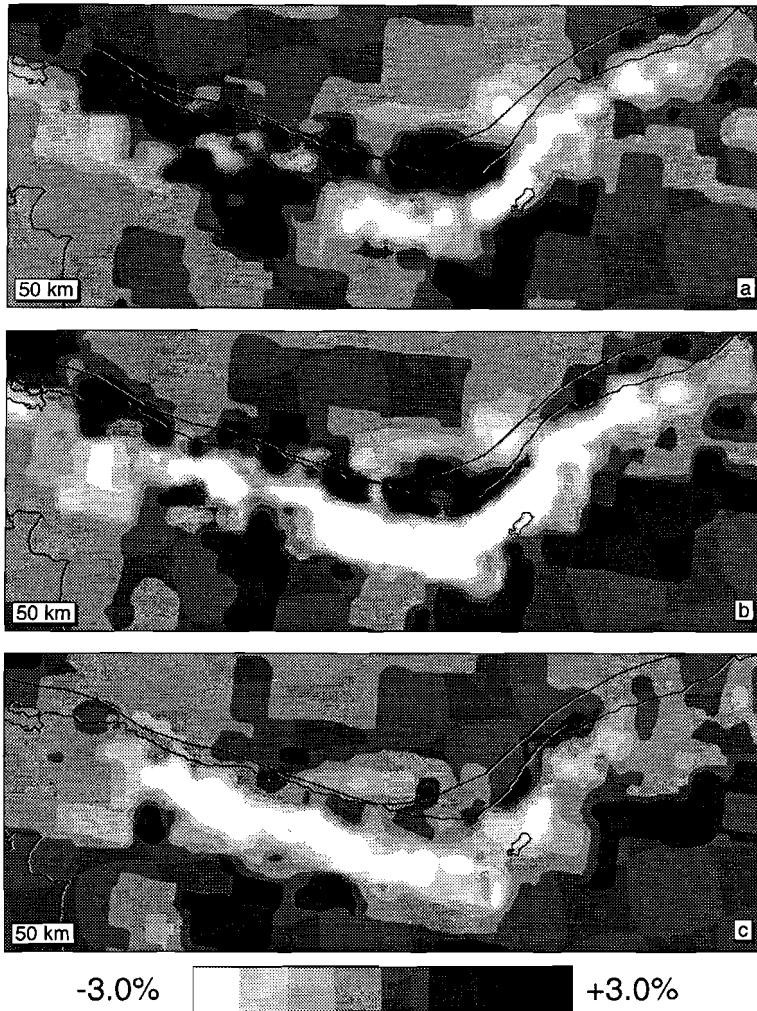
**Figure 8.4:** Whole Earth plot of the increment (BSE\_NL-BSE) at 50 km depth. At this depth most of the oceanic regions are poorly illuminated by seismic rays, but notice the strong heterogeneity below Tibet, Japan, Tonga-Fiji, and South America. Velocity anomalies are depicted in relative deviations from reference model ak135 (Kennett et al., 1995).

#### 8.4.5 Some comparisons between BSE and BSE\_NL

Figure 8.4 depicts the increment for the subcrustal lithosphere. We observe several regions where systematically large changes have resulted from the nonlinear approach (such as Tibet, Japan, and South America). Below we shall discuss some of these regions.

In the uppermost mantle beneath Japan and the surrounding region (depicted in Figure A14), we find two prominent structures in the increment. In the back-arc region of the Izu-Bonin trench a large high-velocity anomaly has appeared, and along the western side of the Ryukyu, Japan, and Kurile trenches increased low-velocity heterogeneity focusses the subduction related anomalies. It is unclear what causes the high-velocity anomaly, but it is located in the back-arc region, where ray sampling is rather poor.

The increased low-velocity heterogeneity along the trenches can be explained by 3-D ray focussing effects: For example, ray paths from (shallow) earthquakes that emerged at shallow depth from the subduction related high velocities in the 1-D case, are now (in the 3-D case) channeled within the high velocity structure to obtain smaller travel times. This results in a thinner and more sharply defined slab structure. Just to illustrate the effect this focussing can have on the appearance of subducted slabs, Figure A15 shows a cross-section through the eastern part of the Aleutian chain where the thickness of the slab in the BSE solution is reduced in the nonlinear result. These focussing effects are consistent with the expectations regarding the added value of 3-D ray tracing for travel-time tomography. Qualitatively, this points at a regionally improved image with respect to BSE.



**Figure 8.5:** Layer section for South America at 50 km depth: (a) the solution obtained by BSE, (b) the nonlinear solution, (c) the increment. This  $16^\circ$  wide slice was taken along a great circle starting at ( $43^\circ\text{S}$ ,  $67^\circ\text{W}$ ) with azimuth  $350^\circ$  and length  $40^\circ$ .

Another strong incremental anomaly can be observed for the subcrustal heterogeneity below South America, depicted in Figure 8.5. The main incremental anomaly is a large low-velocity structure extending approximately N-S below the Andes. The effect of this anomaly in the nonlinear solution is to extend the separation of high velocities associated with oceanic lithosphere of the Pacific plate from those associated with continental lithosphere of South America (note that most of the eastern part of South America is not well-resolved due to a lack of seismic stations and earthquakes). The band of low velocities overlies the Andean slab and may be related to raised temperatures and/or crustal thickening. Note that

a large part of this collision zone is also visible in the BSE solution.

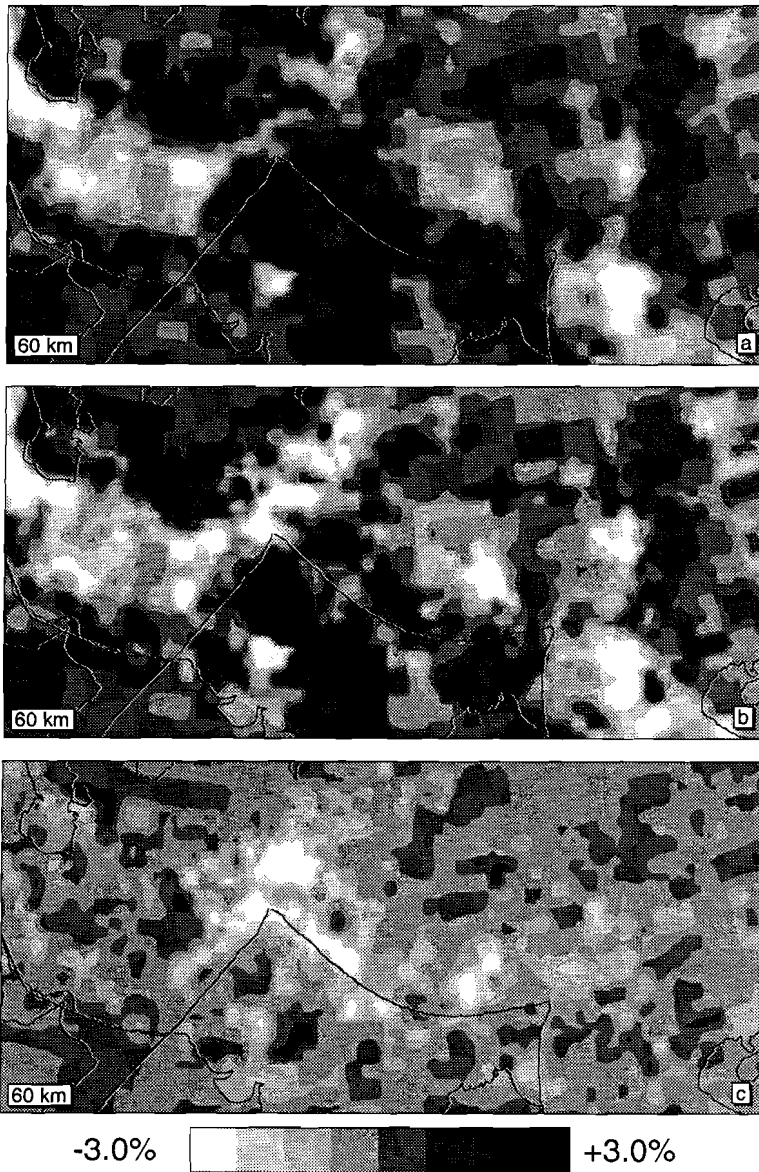
Figure 8.6 shows a result for the India-Asia collision zone. Below Tibet another significant low-velocity increment can be observed. In the BSE solution we find high velocities associated with thick Indian lithosphere and with the Eurasian continent, separated by a thin zone of lowered seismic wavespeed in the Pamir/Hindu-Kush region. This zone is considerably enhanced in the nonlinear solution and a second separation between high seismic velocities appears directly beneath the inferred plate boundary in the Himalayas. Generally, the nonlinear solution enhances the correlation (at depth) between the 3-D velocity pattern and the larger scale tectonic features in the area.

Figure 8.7 shows the shallow mantle structure below Europe at 150 km depth. Differences between linear and nonlinear solutions are spatially more subtle in this region but still of considerable amplitude. The subduction related anomalies below northern Africa, Italy, the Balkan, and Greece are clearly sharpened in the nonlinear solution. The clear(er) absence of high velocities below, for example, central Italy supports the possibility of detachment of slab material in this region (as suggested by Wortel and Spakman, 1992; Spakman et al., 1993). The low velocity increment below the Dinarides and central Greece focusses the high-velocity anomalies associated with Aegean subduction. The low velocity increment below the eastern Mediterranean attempts to reduce the thickness of the lithosphere.

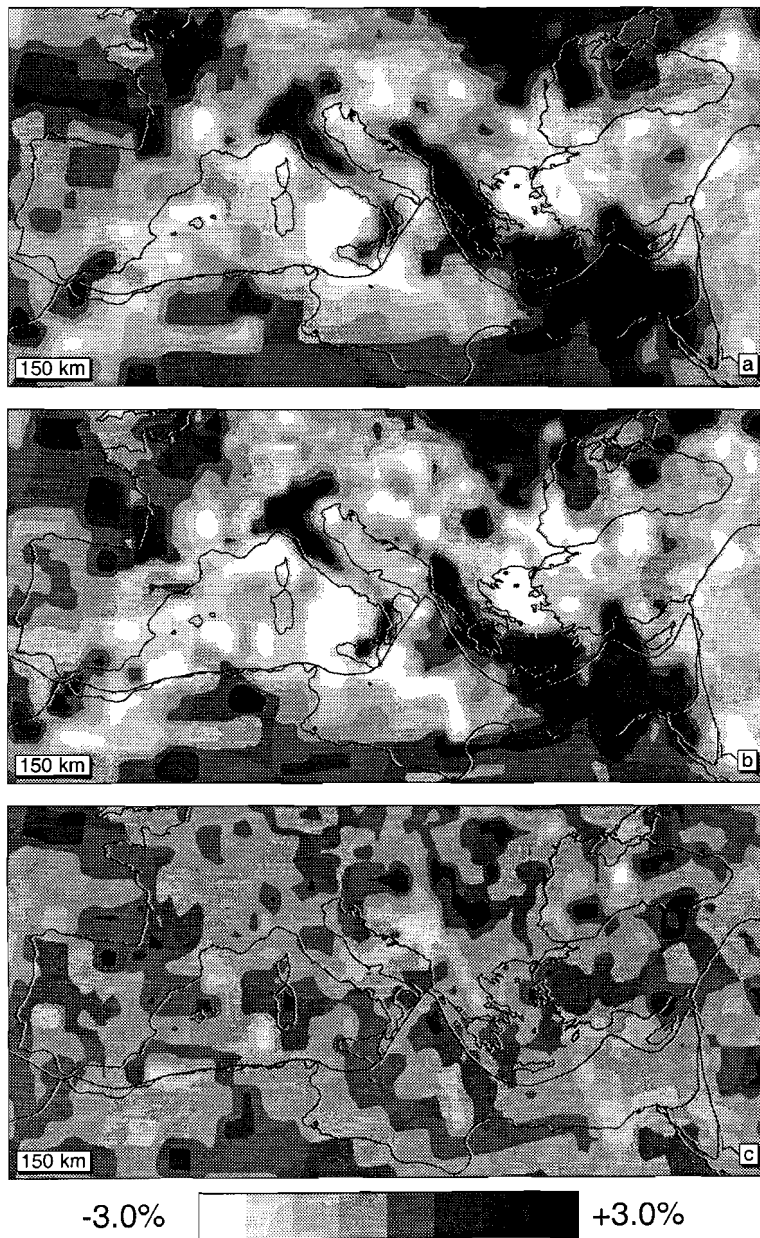
Finally, the Tonga-Fiji region also displays some significant incremental structure (see Figure 8.8). Here, we again observe lowered velocities alongside the subducting oceanic lithosphere, sharpening the contrast between slab and mantle, and some increased velocities in the back-arc region. However, in this case this does not lead to large differences between linear and nonlinear solution patterns.

The differences in the lower mantle are generally more subtle than in the upper mantle. Because of the small scales (with respect to the typical scale length of lower mantle anomalies in BSE) and amplitudes of these differences, it is more difficult to assess the improvement (if any) due to nonlinear inversion in this part of the model. We do, however, observe in many regions an increased, but subtle sharpening of the inferred structures. As an example, we show in Figure A16 the Tethys anomaly (see e.g. Bijwaard et al., 1998, and Chapter 4), the fine structure of which (three roughly parallel zones) has been interpreted by Van der Voo et al. (1999). A comparison of linear and nonlinear solutions shows that the different parts of this anomaly (which have been attributed to different subductions) become more clearly separated in the nonlinear result, which supports the interpretations of Van der Voo et al. (1999). In general, we also observe slightly raised amplitudes in this region.

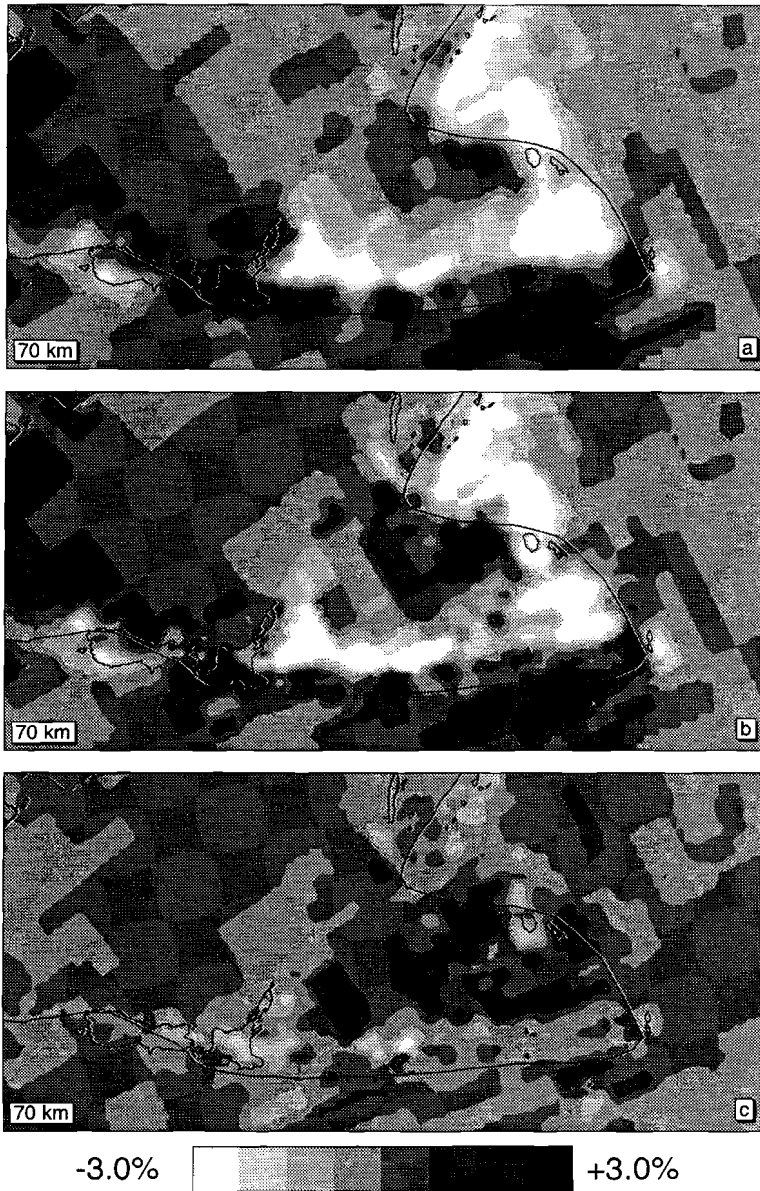
To further illustrate the differences between nonlinear effects in the upper and lower mantle, Figure 8.9 shows a whole mantle slice for the mantle below Iceland, which is vertically well-resolved in the lower mantle. Figure 8.9a was interpreted by Bijwaard and Spakman (1999b) (section 5.1) as the image of a whole mantle plume. Figure 8.9b depicts the nonlinear equivalent of Figure 8.9a, and the increment in Figure 8.9c clearly shows strong differences in the upper mantle (enhancing the low velocities) and only small changes in the lower mantle, the most interesting of which is the increase in velocities in  $D''$ .



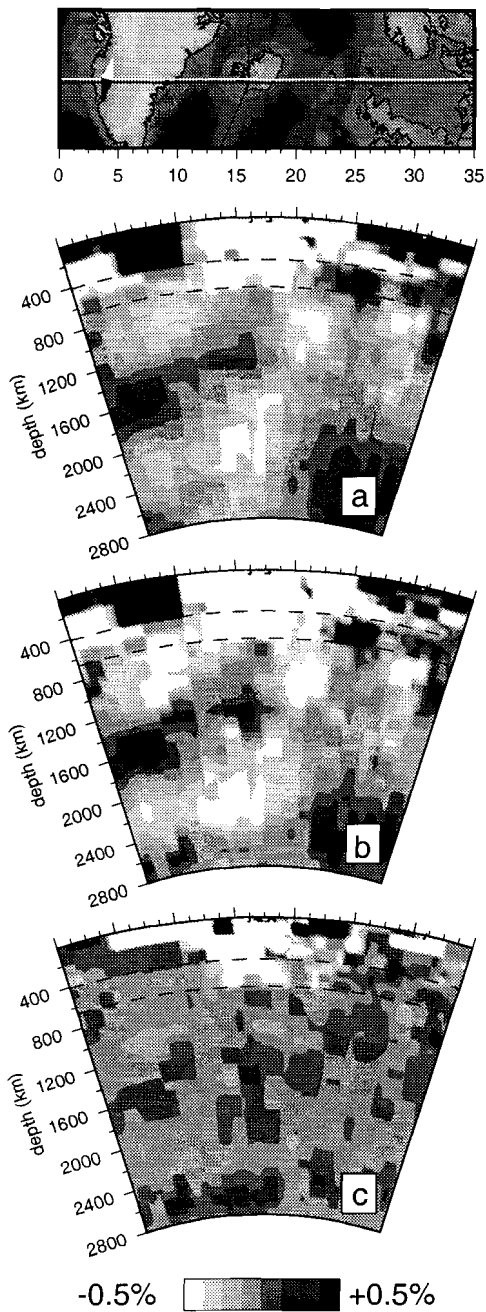
**Figure 8.6:** Layer section for Tibet and surroundings at 60 km depth: (a) the solution obtained by BSE, (b) the nonlinear solution, (c) the increment. This  $26^\circ$  wide slice was taken along a great circle starting at ( $30^\circ\text{N}$ ,  $50^\circ\text{E}$ ) with azimuth  $75^\circ$  and length  $55^\circ$ .



**Figure 8.7:** Layer section for Europe at 150 km depth: (a) the solution obtained by BSE, (b) the nonlinear solution, (c) the increment. This  $20^\circ$  wide slice was taken along a great circle starting at  $(40^\circ\text{N}, 10^\circ\text{W})$  with azimuth  $80^\circ$  and length  $40^\circ$



**Figure 8.8:** Layer section for the Tonga-Fiji region at 70 km depth: (a) the solution obtained by BSE, (b) the nonlinear solution, (c) the increment. This 24° wide slice was taken along a great circle starting at (48°S, 157°E) with azimuth 40° and length 50°



**Figure 8.9:** Cross-section across Iceland and the northern Atlantic Ocean: (a) the solution obtained by BSE, (b) the nonlinear solution, (c) the increment. This slice was taken along a great circle starting at (65.5°N, 60°W) with azimuth 75° and length 35°.



## 8.5 Conclusions

We have extended global travel-time imaging to nonlinear tomography in which seismic ray bending due to inferred velocity heterogeneity has been taken into account. The 3-D ray tracing algorithm of Bijwaard and Spakman (1999a) (Chapter 7) has been applied in updating the mantle model of Bijwaard et al. (1998) (BSE) (Chapter 4) that was obtained from linearized tomography. The final solution was computed in two nonlinear iteration steps (taking one week of computation time each).

The results consist of a small additional variance reduction and a general increase in model amplitudes, but no overall dramatic change in anomaly patterns. Basically we found a different model fitting the data equally well as the BSE model. Resolution and variance estimates are also highly comparable. Therefore, we could only investigate possible improvements by a comparison between BSE and the new model (BSE\_NL). Generally, we infer a more focussed image in accord with expectations about nonlinear effects and detailed mantle structure. We also find a small but systematic shift to lower velocities, which is in agreement with expectations. This indicates that the usual linearized approach to global mantle tomography with travel times is a reasonable approximation, at least for the mantle below 200-400 km. However, nonlinear inversion does lead to strong amplitude changes and more focussed images with exceptionally large changes occurring in several regions of the uppermost 200 km of the Earth's mantle (Japan, Tibet, South America, Europe, Tonga-Fiji). The depth dependence of nonlinear effects is due to the fact that these depend heavily on model amplitudes which are highest in the upper mantle and fall-off rapidly towards values on the order of 0.5% in the lower mantle. Although differences in the lower mantle between linear and nonlinear results are too small to be conclusively resolved, they do generally lead to a sharpening of observed structures (such as the Tethys anomaly). Usually, changes in the model are rather local, 100-500 km scale, which are, however, especially for the upper few hundred kilometers, important in the context of tectonic interpretation.

In general, the nonlinear effects that were found in this study will be somewhat underestimated, because the BSE model that served as a starting model for nonlinear inversion underestimates mantle amplitudes by 20-30% in well-resolved regions. Apart from this, the nonlinear effect of hypocenter mislocation on travel-time residuals was not fully taken into account, because of the limited value of the BSE relocation parameters. A promising next step would be to relocate all events in the 3-D Earth models obtained and next redo the tomographic analyses. Ongoing work (Spakman et al., 1999) shows that variance reductions of the Engdahl et al. (1998) data between 20%-60% can be achieved when events are located in our 3-D models. This allows then for a better separation between mislocation and 3-D heterogeneity signal in delay time data from which future tomographic models can benefit, as BSE and BSE\_NL did from the improvement of the Engdahl et al. (1998) data with respect to the ISC data.

## References

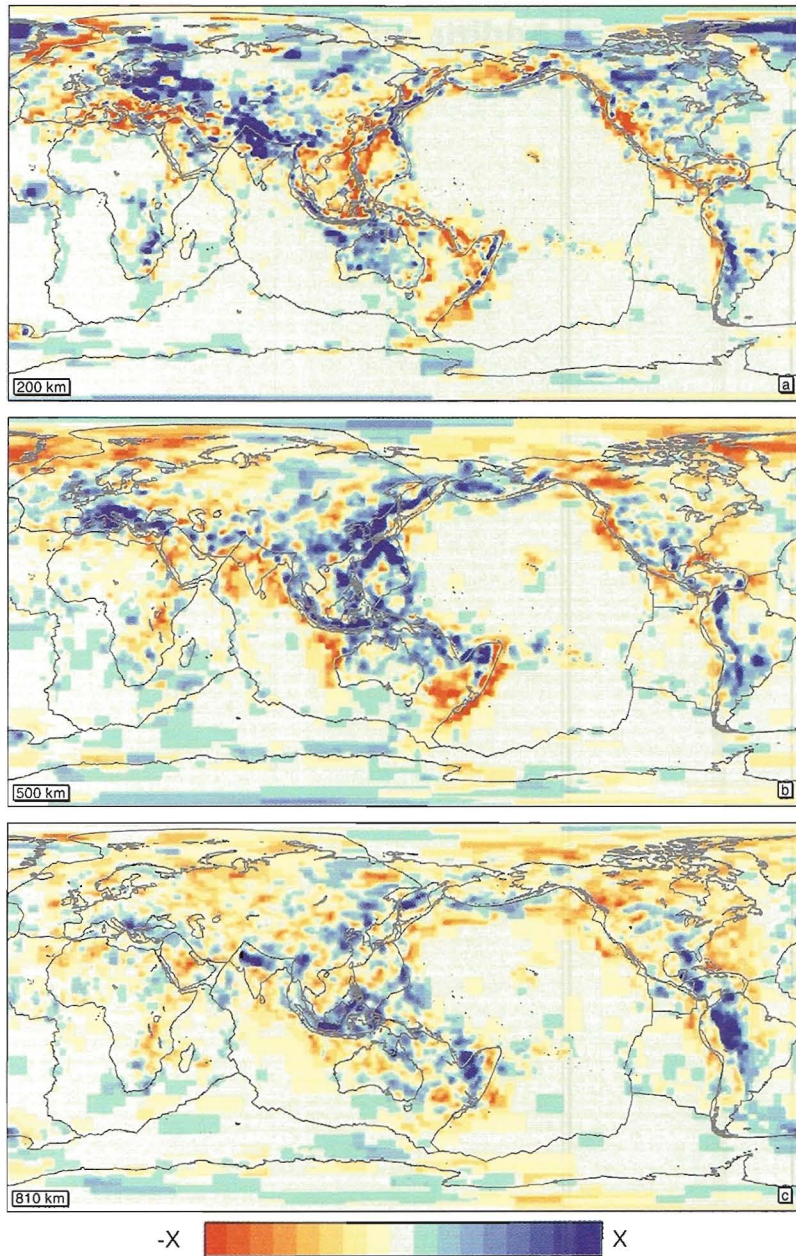
- Bijwaard, H., and W. Spakman, Fast kinematic ray tracing of first and later arriving seismic phases, *Geophys. J. Int.* (in press), 1999a.  
Bijwaard, H., and W. Spakman, Tomographic evidence for a narrow whole mantle plume

- below Iceland, *Earth Planet. Sci. Lett.*, 166, 121-126, 1999b.
- Bijwaard, H., W. Spakman, and E.R. Engdahl, Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, 103, 30055-30078, 1998.
- Engdahl, E.R., R.D. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seism. Soc. Amer.*, 88, 722-743, 1998.
- Gudmundsson, O., J.H. Davies, and R.W. Clayton, Stochastic analysis of global traveltime data: Mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, 102, 25-43, 1990.
- Inoue, H., Y. Fukao, K. Tanabe, and Y. Ogota, Whole mantle *P*-wave travel time tomography, *Phys. Earth Planet. Inter.*, 59, 294-328, 1990.
- Kennett, B.L.N., E.R. Engdahl, and R. Buland, Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.*, 122, 108-124, 1995.
- Kennett, B.L.N., S. Widiyantoro, and R.D. van der Hilst, Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle, *J. Geophys. Res.*, 103, 12469-12493, 1998.
- Koch, M., A numerical study on the determination of the 3-D structure of the lithosphere by linear and non-linear inversion of teleseismic travel times, *Geophys. J. R. Astron. Soc.*, 80, 73-93, 1985.
- Masters, G., S. Johnson, G. Laske, and H. Bolton, A shear-velocity model of the mantle, *Philos. Trans. R. Soc. London Ser. A*, 354, 1385-1411, 1996.
- Nolet, G., and T.-J. Moser, Teleseismic delay times in a 3-D Earth and a new look at the *S* discrepancy, *Geophys. J. Int.*, 114, 185-195, 1993.
- Paige, C.C., and M.A. Saunders, LSQR: An algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Soft.*, 8, 43-71, 1982.
- Papazachos, C., and G. Nolet, *P* and *S* deep velocity structure of the Hellenic area obtained by robust nonlinear inversion of traveltimes, *J. Geophys. Res.*, 102, 8349-8367, 1997.
- Pulliam, J., and R. Snieder, Fast, efficient calculation of rays and travel times with ray perturbation theory, *J. Acoust. Soc. Am.*, 99, 383-391, 1996.
- Sambridge, M.S., Non-linear arrival time inversion: constraining velocity anomalies by seeking smooth models in 3-D, *Geophys. J. Int.*, 102, 653-667, 1990.
- Snieder, R., and M. Sambridge, Ray perturbation theory for travel times and ray paths in 3-D heterogeneous media, *Geophys. J. Int.*, 109, 204-322, 1992.
- Spakman, W., Iterative strategies for non-linear travel time tomography using global earthquake data, in: *Seismic Tomography; theory and Practice*, edited by H.M. Iyer and K. Hirahara, pp. 190-226, Chapman and Hall, London, 1993.
- Spakman, W., and H. Bijwaard, Irregular cell parameterization of tomographic problems, *Ann. Geophys.*, 16, 28, 1998.
- Spakman, W., and M. van Deelen, Experiments for nonlinear travel-time tomography of the mantle: a search for the best iteration strategy, *Eos*, 74, 418, 1993.
- Spakman, W., S. van der Lee, and R.D. van der Hilst, Travel-time tomography of the European-Mediterranean mantle down to 1400 km, *Phys. Earth Planet. Inter.*, 79, 3-74, 1993.
- Spakman, W., R. Boevé, H. Bijwaard, and E.R. Engdahl, 3-D earthquake location with grid search methods: Application to nuclear explosions and earthquakes (manuscript in

- preparation), 1999.
- Souriau, A., and M. Granet, A tomographic study of the lithosphere beneath the Pyrenees from local and teleseismic data, *J. Geophys. Res.*, 100, 18117-18134, 1995.
- Su, W., R.L. Woodward, A.M. Dziewonski, Degree 12 model of shear velocity heterogeneity in the mantle, *J. Geophys. Res.*, 99, 6945-6980, 1994.
- Thomson, C.J., and D. Gubbins, Three-dimensional lithospheric modelling at NORSAR: linearity of the method and amplitude variations from the anomalies, *Geophys. J. R. Astron. Soc.*, 71, 1-36, 1982.
- Thurber, C.H., Earthquake locations and three-dimensional crustal structure in the Coyote Lake area, central California, *J. Geophys. Res.*, 88, 8226-8236, 1983.
- Trampert, J., and J.H. Woodhouse, High resolution global phase velocity distributions, *Geophys. Res. Lett.*, 23, 21-24, 1996.
- VanDecar, J.C., Upper mantle structure of the Cascadia subduction zone from non-linear teleseismic travel-time inversion, Ph.D. thesis, pp. 165, Univ. of Washington, Seattle, 1991.
- Van der Hilst, R.D., E.R. Engdahl, and W. Spakman, Tomographic inversion of  $P$  and  $pP$  data for aspherical mantle structure below the northwest Pacific region, *Geophys. J. Int.*, 115, 264-302, 1993.
- Van der Hilst, R.D., S. Widiyantoro, and E.R. Engdahl, Evidence for deep mantle circulation from global tomography, *Nature*, 386, 578-584, 1997.
- Van der Voo, R., W. Spakman, and H. Bijwaard, Tethyan subducted slabs under India, *Earth Planet. Sci. Lett.*, 171, 7-20, 1999.
- Vasco, D.W., and L.R. Johnson, Whole Earth structure estimated from seismic arrival times, *J. Geophys. Res.*, 103, 2633-2671, 1998.
- Vasco, D.W., L.R. Johnson, and R.J. Pulliam, Lateral variations in mantle velocity structure and discontinuities determined from,  $P$ ,  $PP$ ,  $S$ ,  $SS$ , and  $SS-SdS$  travel time residuals, *J. Geophys. Res.*, 100, 24307-24059, 1995.
- Wielandt, E., On the validity of the ray approximation for interpreting delay times, in: *Seismic Tomography*, edited by G. Nolet, pp. 85-98, Reidel, Dordrecht, the Netherlands, 1987.
- Wortel, M.J.R., and W. Spakman, Structure and dynamics of subducted lithosphere in the Mediterranean region, *Proc. K. Ned. Akad. Wet.*, 95, 325-347, 1992.
- Zhang, Y., and T. Tanimoto, High-resolution global upper mantle structure and plate tectonics, *J. Geophys. Res.*, 98, 9793-9823, 1993.
- Zhao, D., A. Hasegawa, and S. Horiuchi, Tomographic imaging of  $P$  and  $S$  wave velocity structure beneath northeastern Japan, *J. Geophys. Res.*, 97, 19909-19928, 1992.
- Zhou, H.-W., Mapping of  $P$ -wave slab anomalies beneath the Tonga, Kermadec and New Hebrides arcs, *Phys. Earth Planet. Int.*, 61, 199-229, 1990.
- Zhou, H.-W., A high-resolution  $P$  wave model for the top 1200 km of the mantle, *J. Geophys. Res.*, 101, 27791-27810, 1996.

# **Appendix**

## **Additional figures**



**Figure A1:** Selected whole Earth layer solutions from the *P*-wave model of Chapter 4 centered around ( $0^{\circ}\text{N}$ ,  $145^{\circ}\text{E}$ ) (slightly smoothed over distances of  $0.6^{\circ}$ ). Contour scales are saturated and range from  $-X$  to  $+X$  with respect to *ak135* (Kennett et al., 1995). Depths are (a) 200 km ( $X=2\%$ ), (b) 500 km ( $X=1.5\%$ ), (c) 810 km ( $X=1\%$ ), (d) 1325 km ( $X=0.5\%$ ), (e) 1900 km ( $X=0.5\%$ ), and (f) 2805 km ( $X=0.5\%$ ).



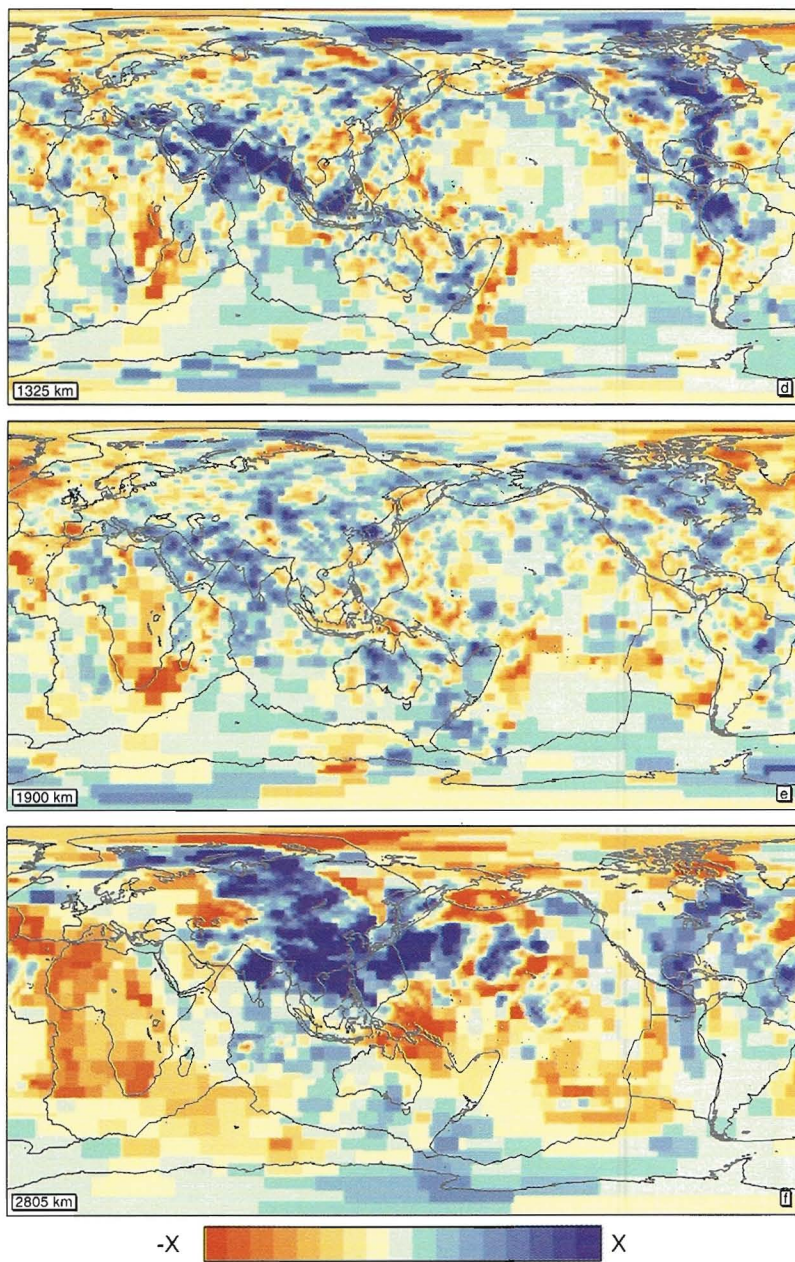
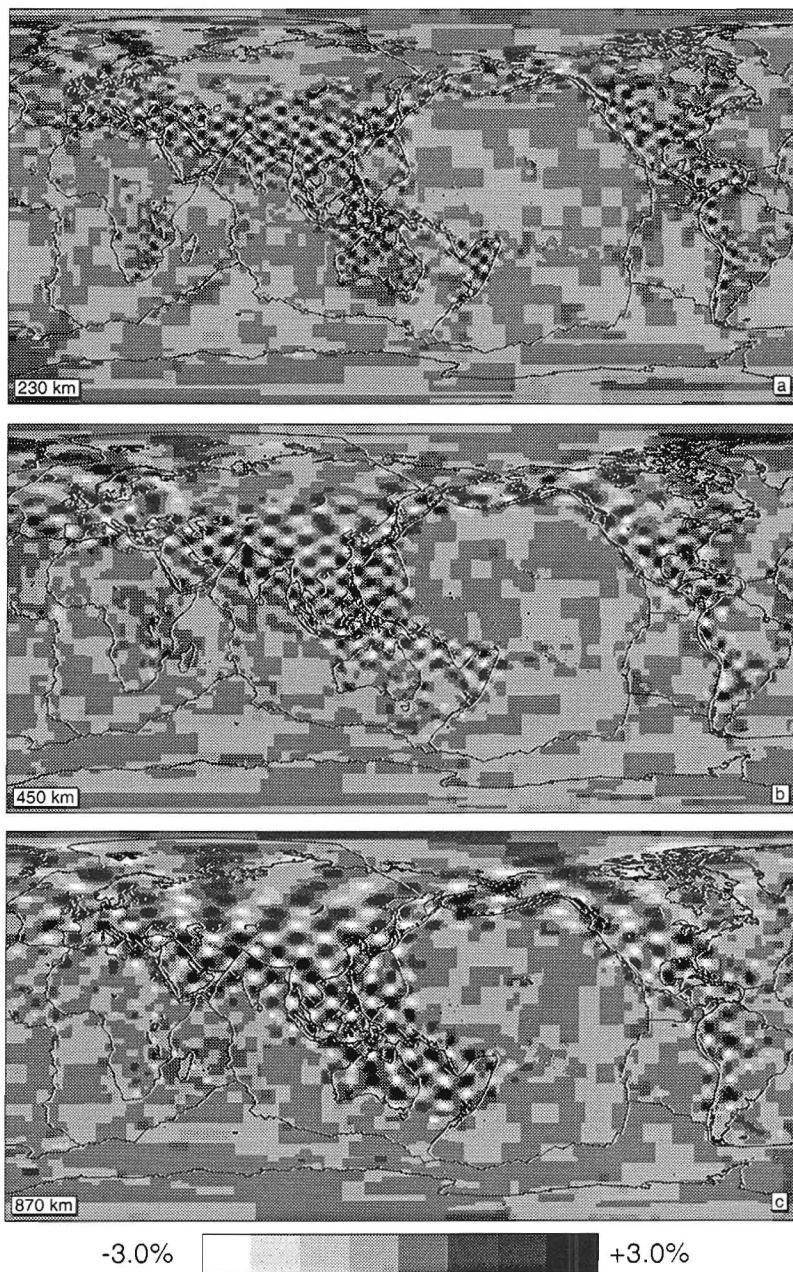


Figure A1: Continued.



**Figure B1:** Output from synthetic sensitivity tests (unsmoothed) for similar depths as shown in Figure A1. Spike sizes are  $1.8^\circ$ ,  $2.4^\circ$ ,  $3.0^\circ$ ,  $3.0^\circ$ ,  $4.2^\circ$ , and  $6.0^\circ$  for panels (a) to (f), respectively.



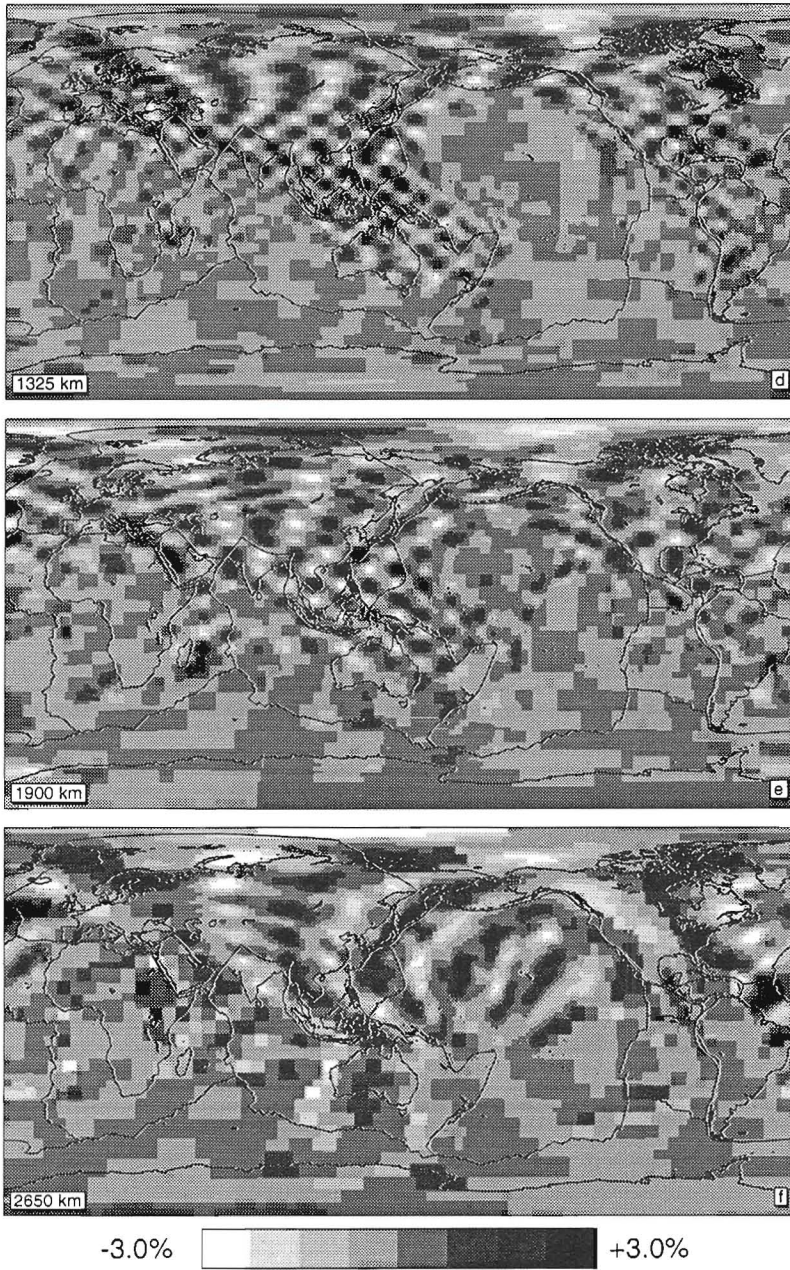
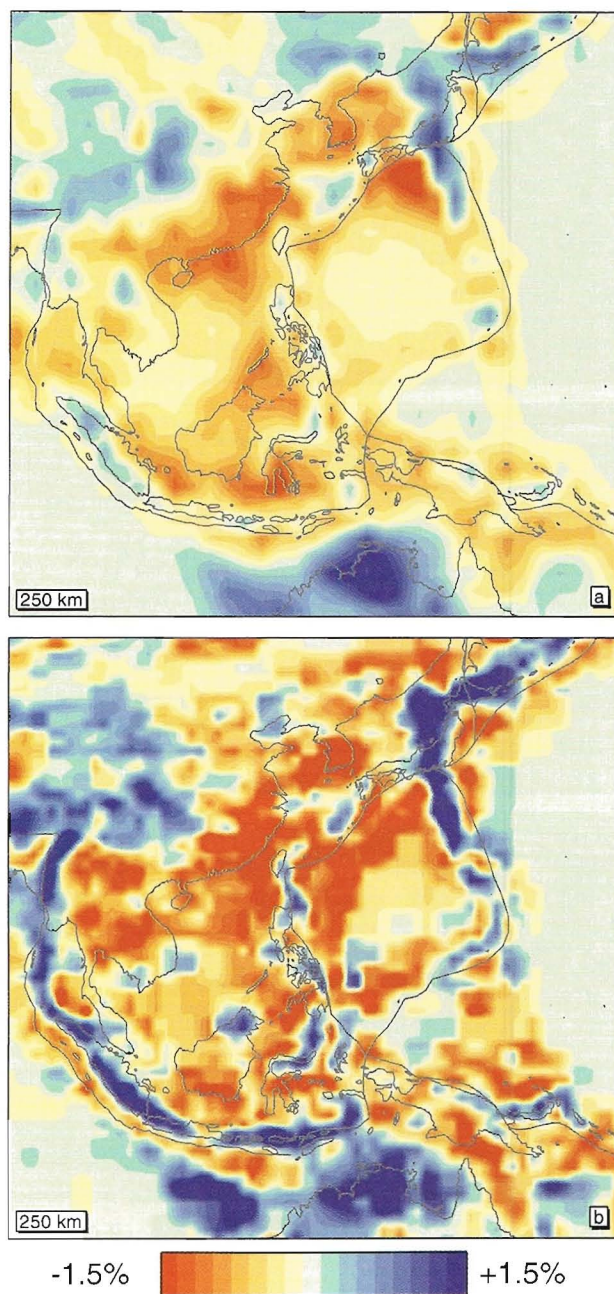
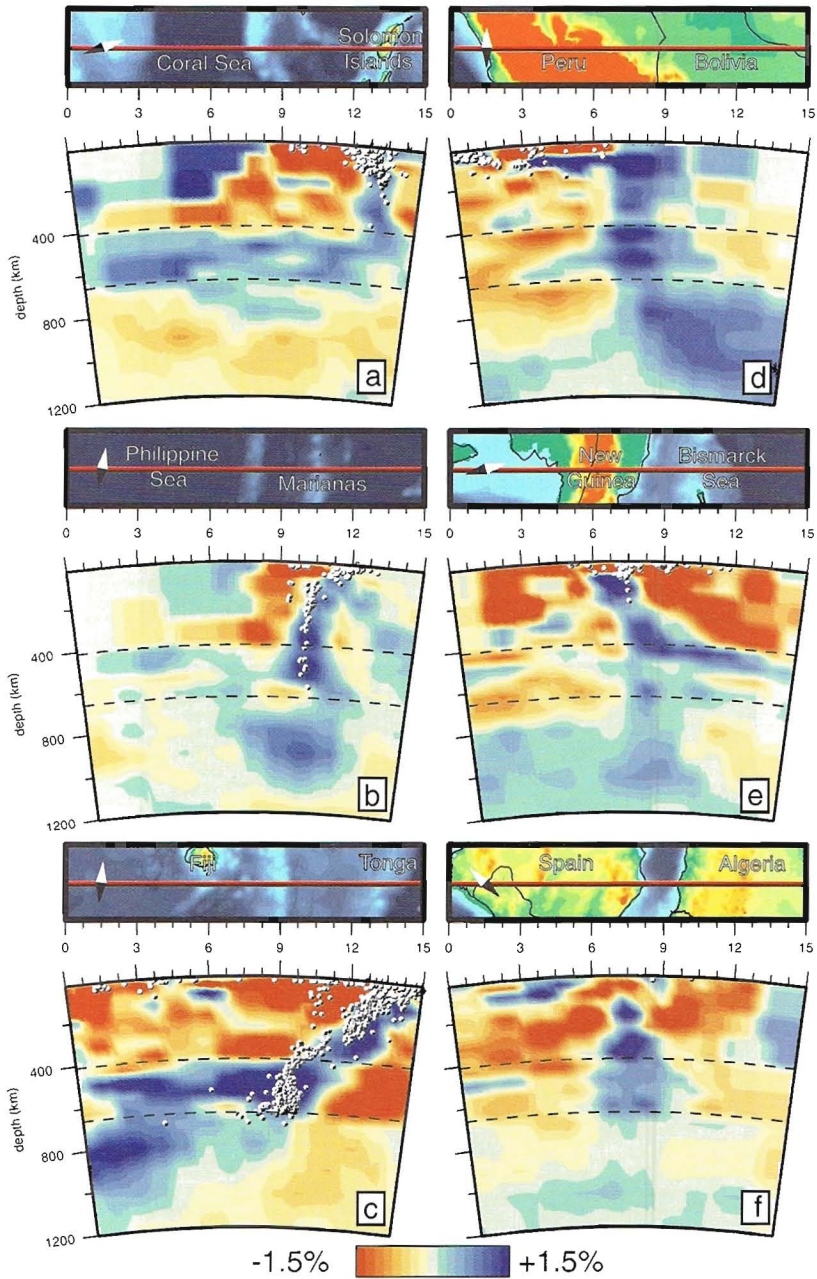


Figure B1: Continued.





**Figure A2:** Comparison of the solution for southeast Asia ( $90^{\circ}\text{E}$ - $160^{\circ}\text{E}$ ) ( $20^{\circ}\text{S}$ - $50^{\circ}\text{N}$ ) at 250 km depth between (a) the model of Van der Hilst et al. (1997) and (b) the model of Chapter 4 (unsmoothed).



**Figure A3:** Cross-sections displaying different types of subduction in the model of Chapter 4 below (a) the Solomon Islands (flattening on top of the 660 km discontinuity), (b) the Marianas (penetrating into the lower mantle), (c) Tonga-Fiji (penetrating after flattening), (d) Peru (kinking at shallow depth), (e) New Guinea (slab splitting (?) at the 660 km discontinuity), and (f) Spain (possibly detachment).



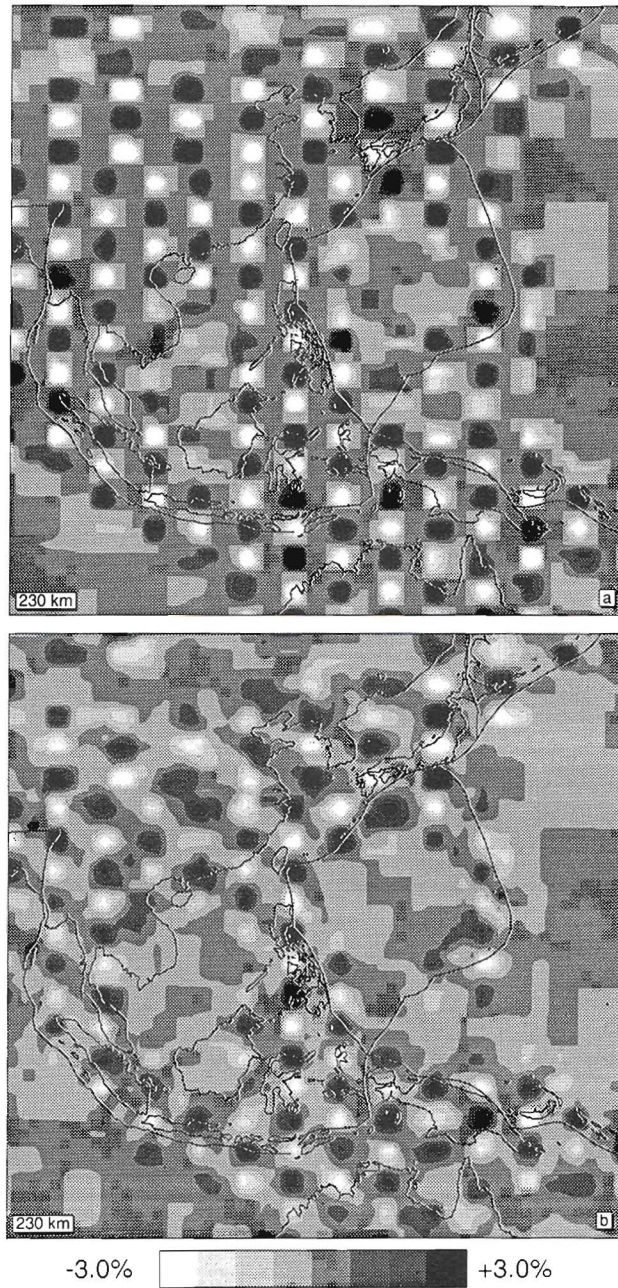
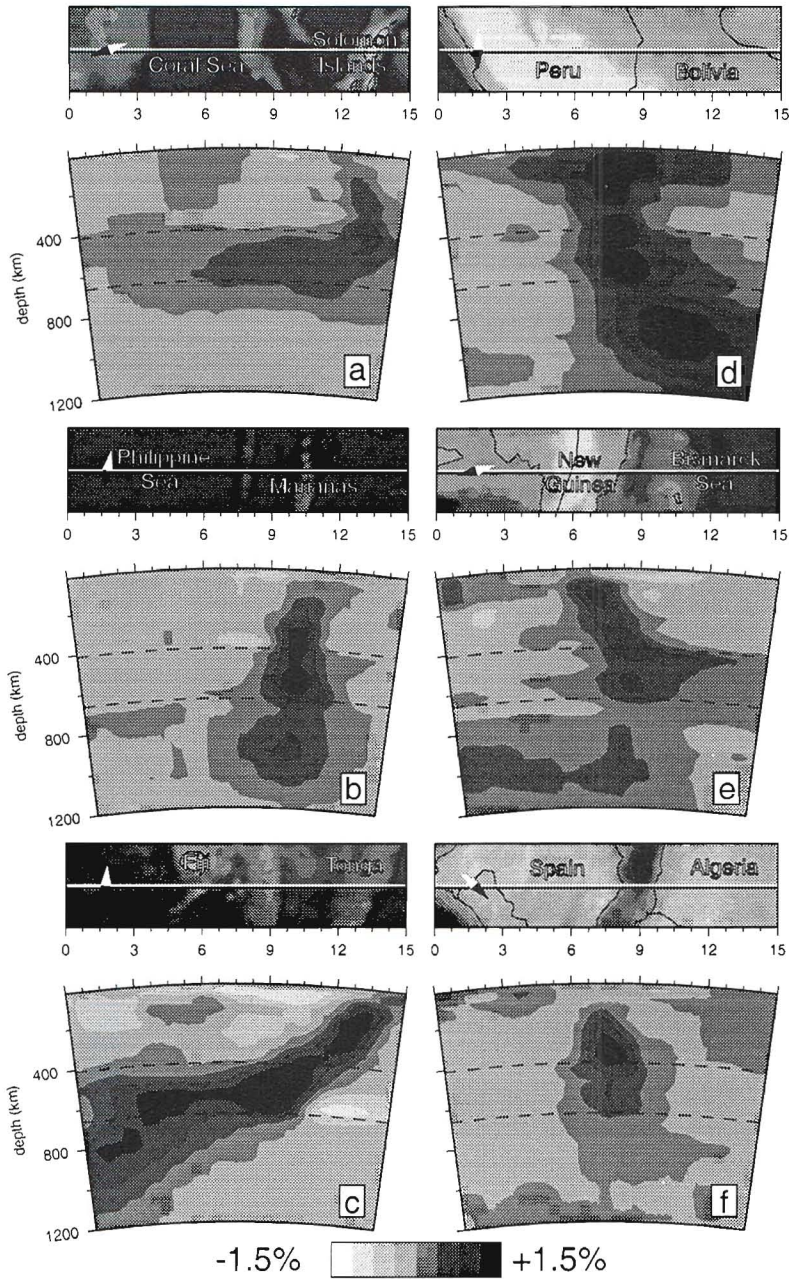
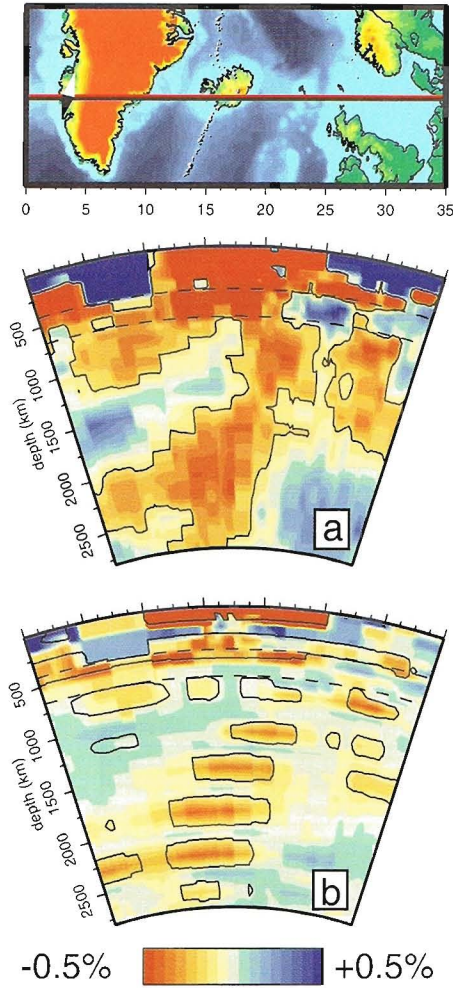


Figure B2: Input and output of a sensitivity test with  $1.8^\circ$  spikes for the layer section of Figure A2.

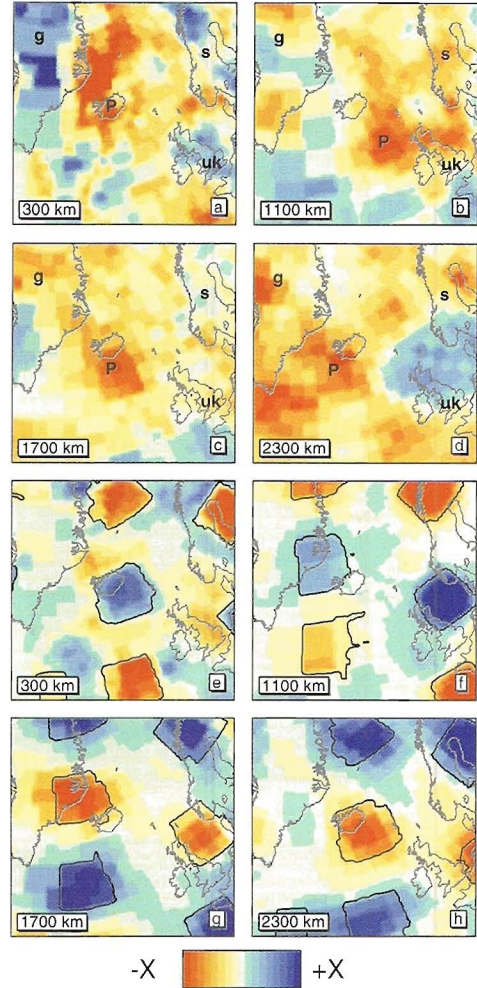


**Figure B3:** Output of a synthetic sensitivity experiment with slab structures on input for the cross-sections shown in Figure A3. All slabs are well-retrieved.

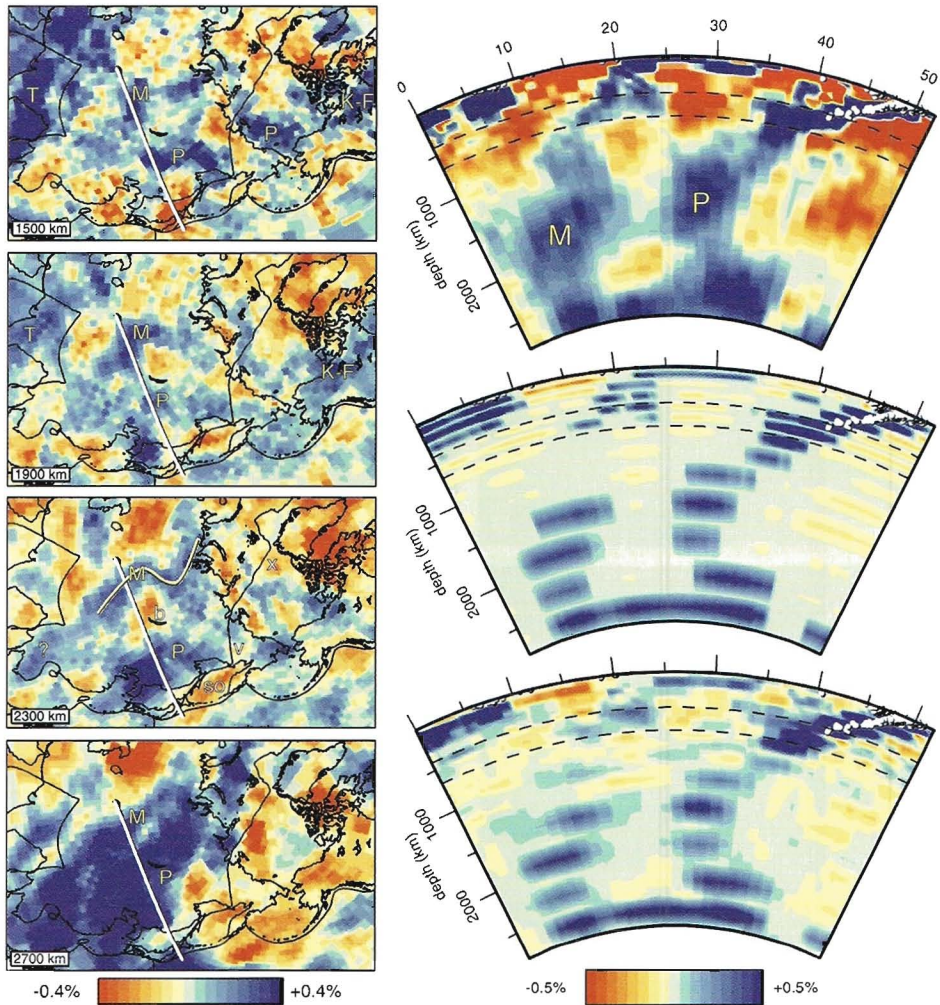




**Figure A4:** (a) Cross-section through the model of Bijwaard et al. (1998a) (Chapter 4) showing a large plume-shaped anomaly in the mantle below Iceland. A 0.15% contour is indicated in black for clarity. (b) Result from a synthetic layer-cake test. Superposed in black is the 0.15% contour of the synthetic input model. Dashed lines in both (a) and (b) indicate the 410 and 660 km discontinuities.



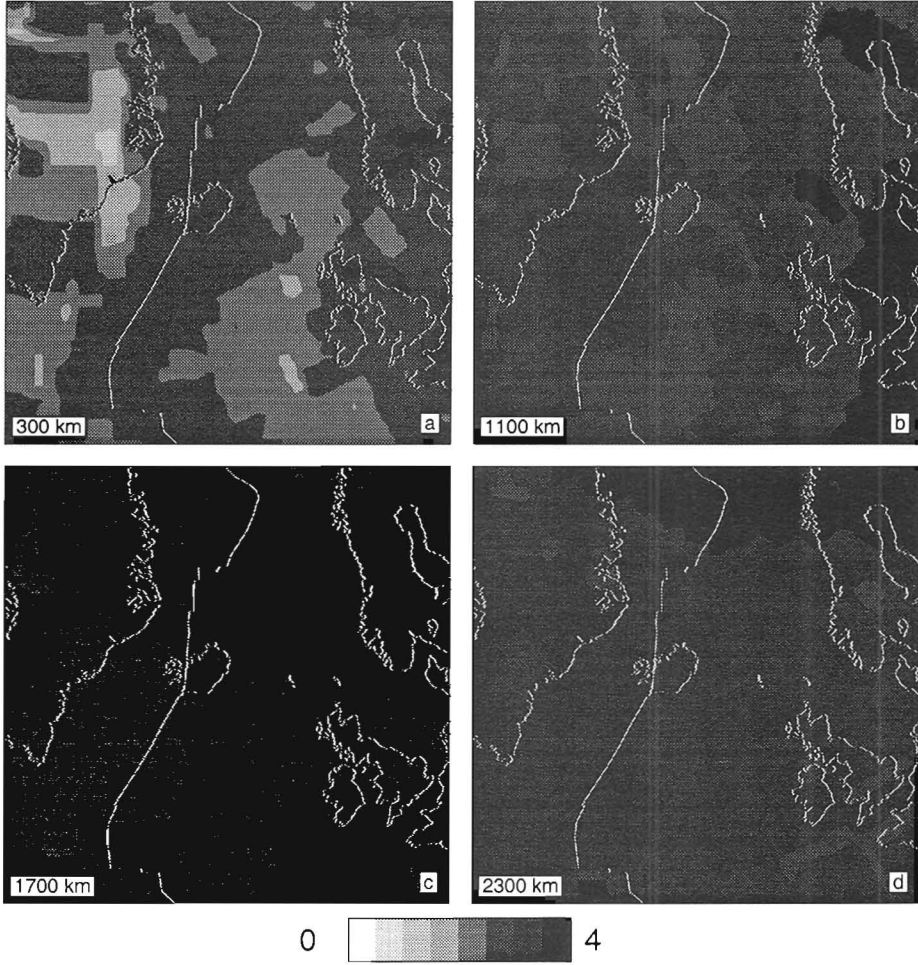
**Figure A5:** (a-d) layer sections through the model of Chapter 4 at 300 km ( $X=2\%$ ), 1100 km ( $X=0.5\%$ ), 1700 km ( $X=0.5\%$ ), and 2300 km ( $X=0.5\%$ ). Lower case letters are Greenland (g), Scandinavia (s), and the United Kingdom (uk); Upper case P stands for the position of the plume. (e-h) results of a 6° spike test for the same depths as in (a-d) ( $X=3\%$ ) with superposed in black the 1% contour of the input model.



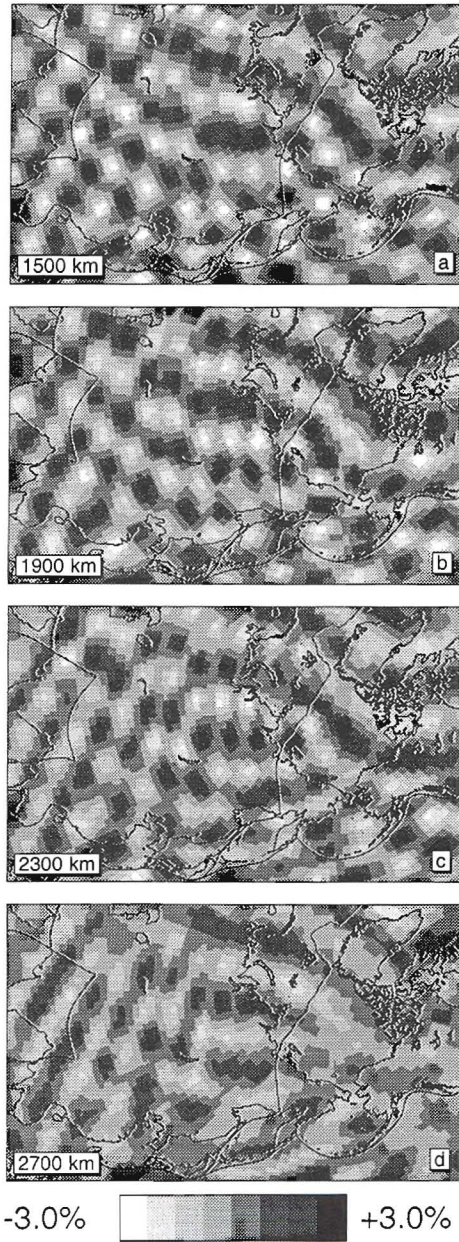
**Figure A6:** Layer sections from the model of Chapter 4 at four depths between 1500 and 2700 km below Asia. Capital yellow letters refer to interpretations as slabs derived from the Tethys (T), Pacific (P), and Kula-Farallon (K-F) subductions; M denotes the unidentified anomaly. Lower-case white letters in the 2300 km panel stand for Lake Baikal (b), the Sea of Okhotsk (so), the Verkhoyansk suture (v), and the North Pole (x). The white line indicates the location of the cross-section in Figure A7; the yellow line represents the Z shape of M at 2300 km depth.

**Figure A7:** Cross-section through the Asian mantle along the white line in Figure A6. The top panel shows the tomographic result with the Pacific subduction (P) on the right and the unidentified anomaly (M) on the left (white dots denote earthquakes). The middle panel is a synthetic layer-cake test model for the same cross-section. The bottom panel shows the output for the layer-cake test.



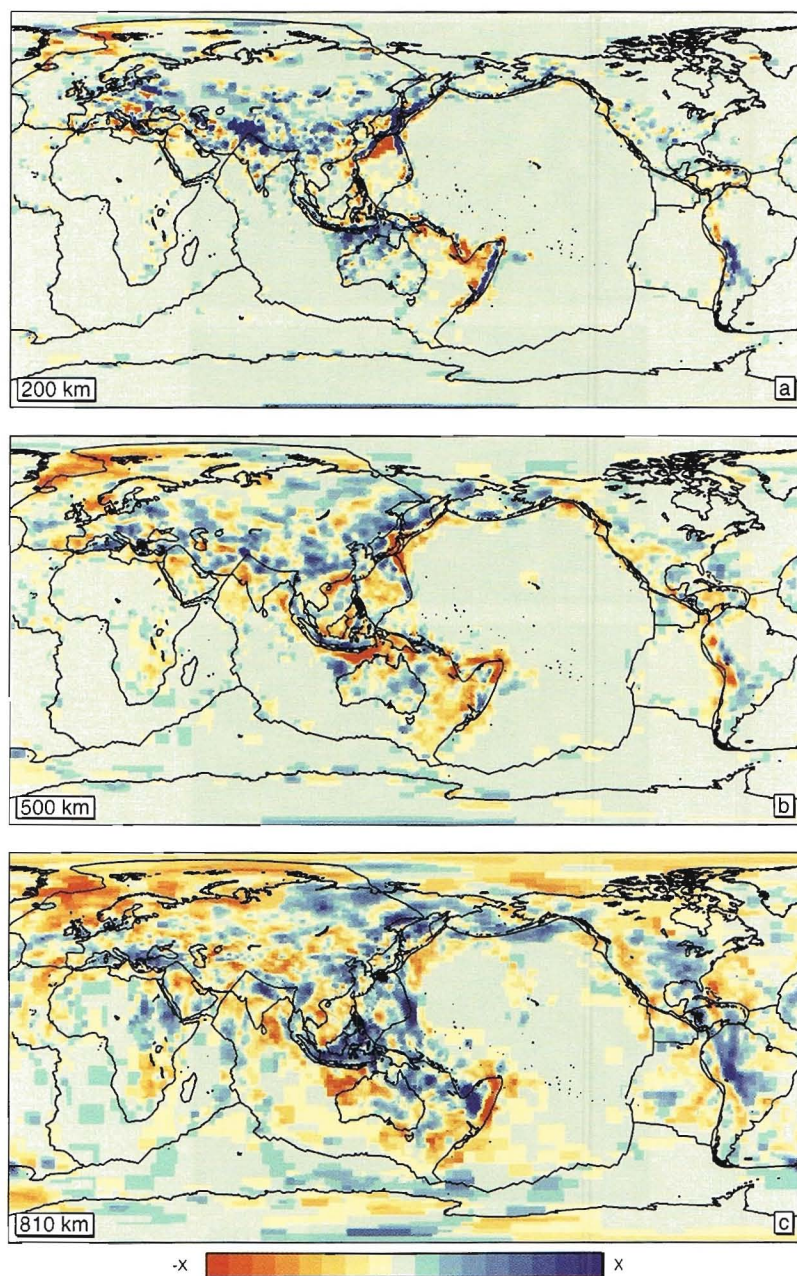


**Figure B5:** Logarithmic hitcount plots for the layer sections displayed in Figure A5.



**Figure B6:** Output of a synthetic sensitivity test with 3° spikes of 5% for the layer sections depicted in Figure A6.





**Figure A8:** Layer solutions for  $S$ -velocity heterogeneity (the model of Chapter 6) with respect to reference model ak135 (centered around  $(0^\circ\text{N}, 145^\circ\text{E})$ ). The (saturated) colour scale ranges from  $-X$  to  $X$ , where  $X=2\%$  for (a) (200 km),  $X=1.5\%$  for (b) (500 km),  $X=1.0\%$  for (c) (810 km), and  $0.5\%$  for (d), (e), and (f) (1325, 1900, and 2805 km, respectively).

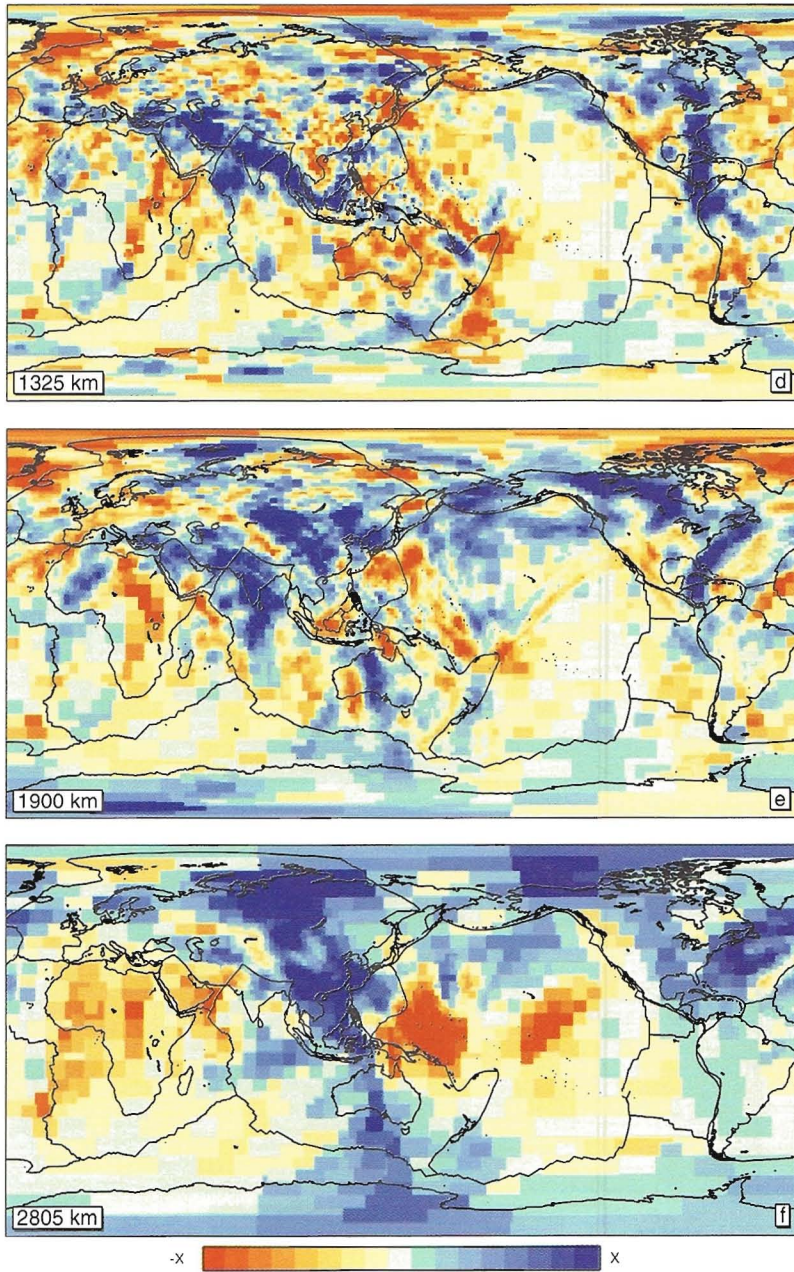
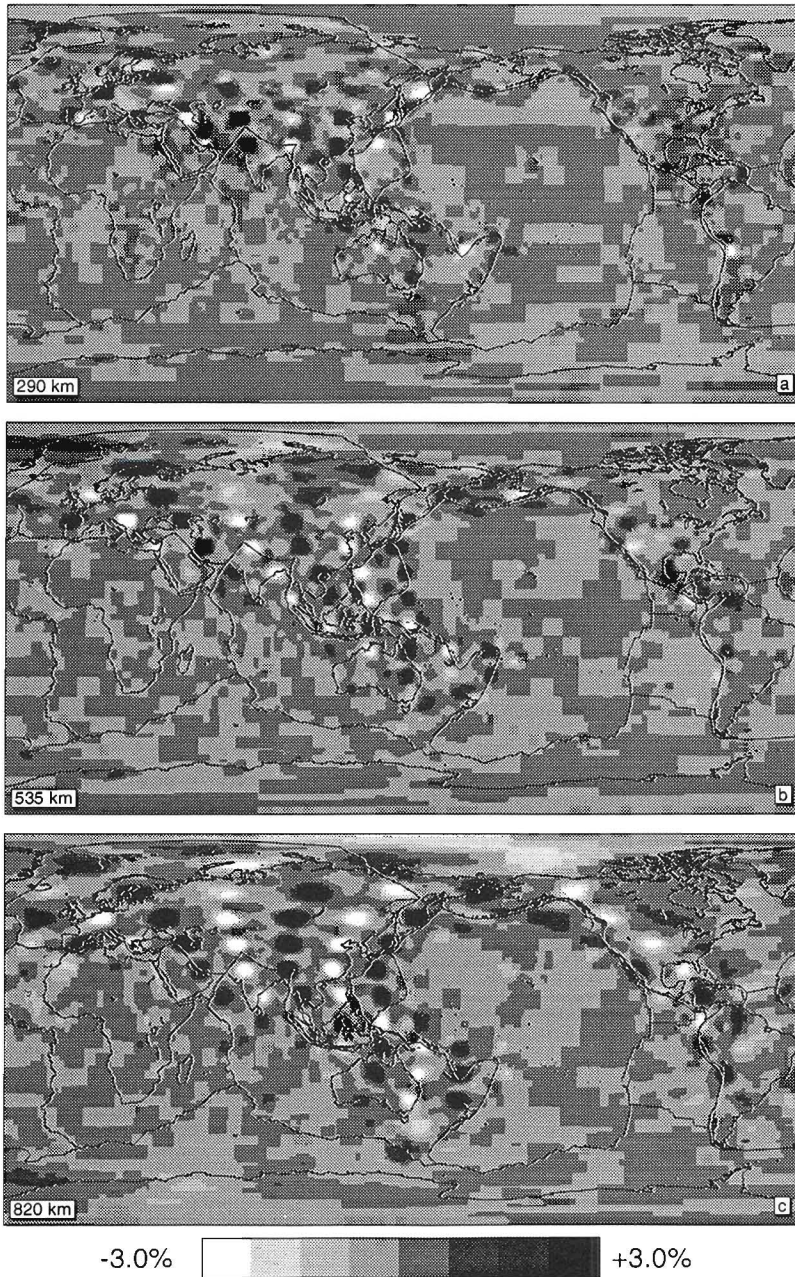


Figure A8: Continued.





**Figure B8:** Output of synthetic experiments for similar layer sections as shown in Figure A8. Sizes of input spikes are  $6^\circ$  for (a) to (c) and  $9^\circ$  for (d) to (f).



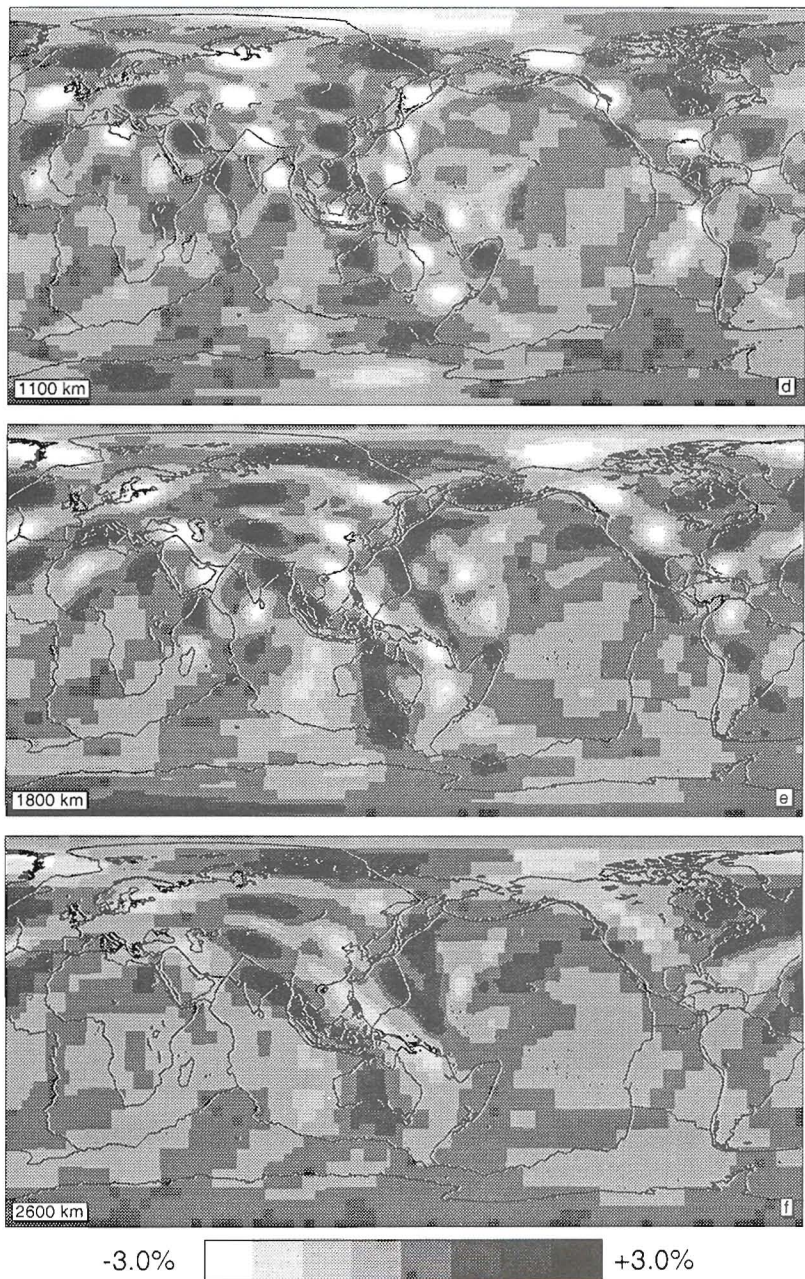
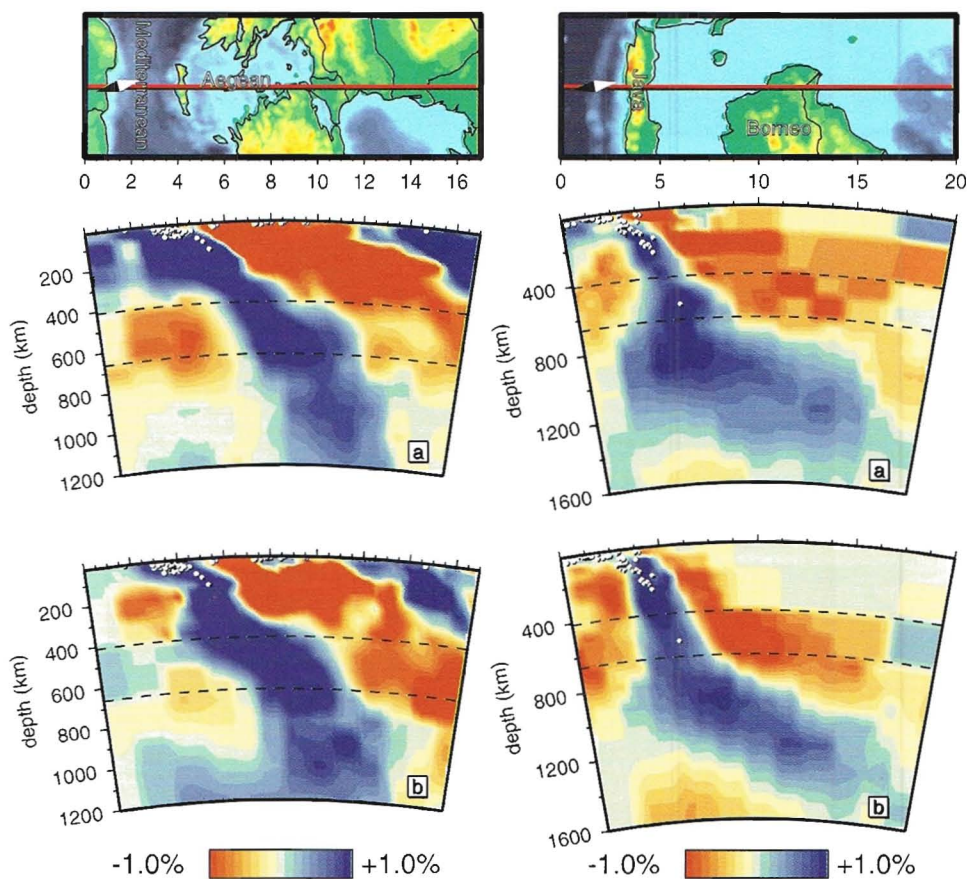


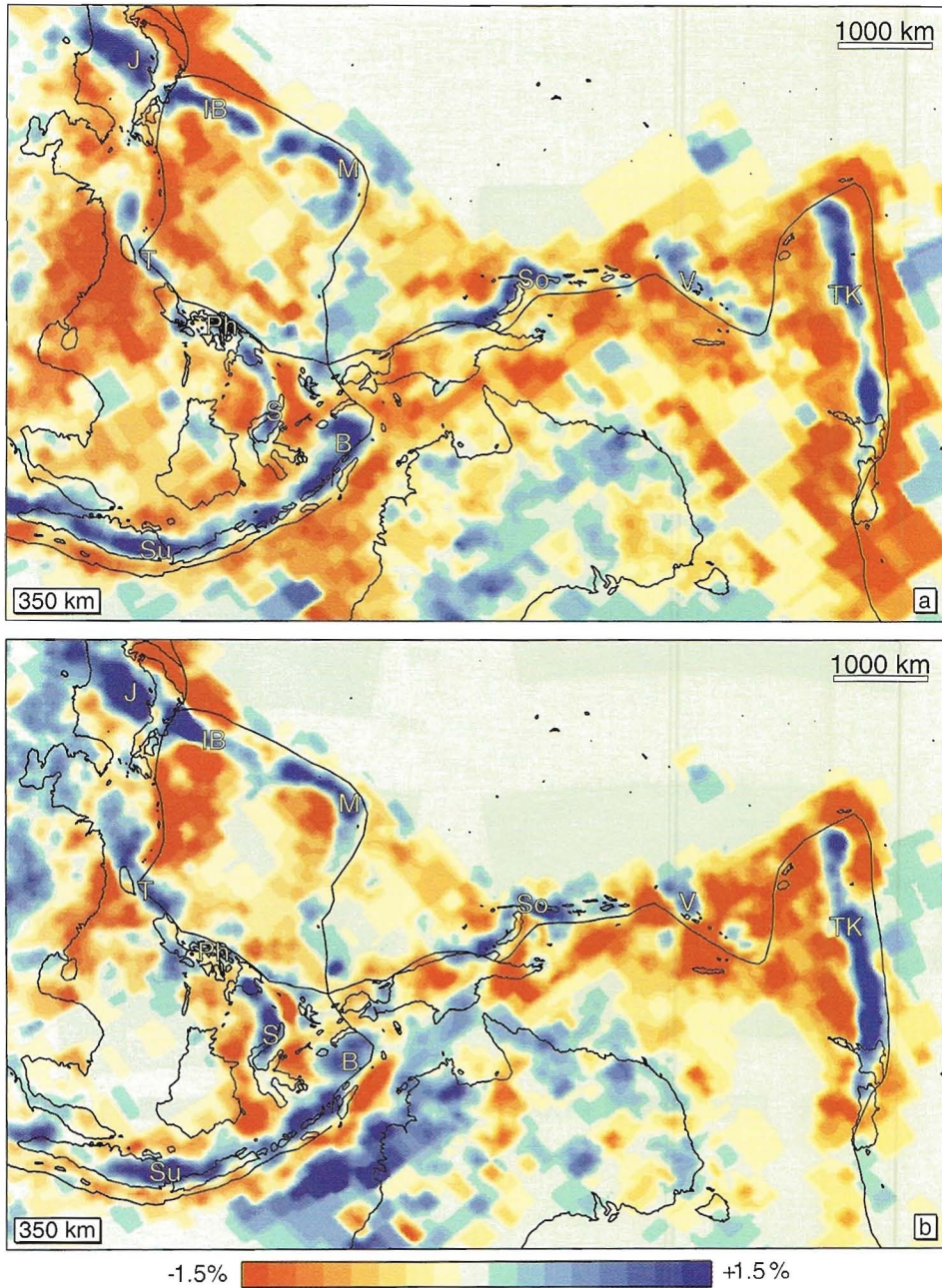
Figure B8: Continued.



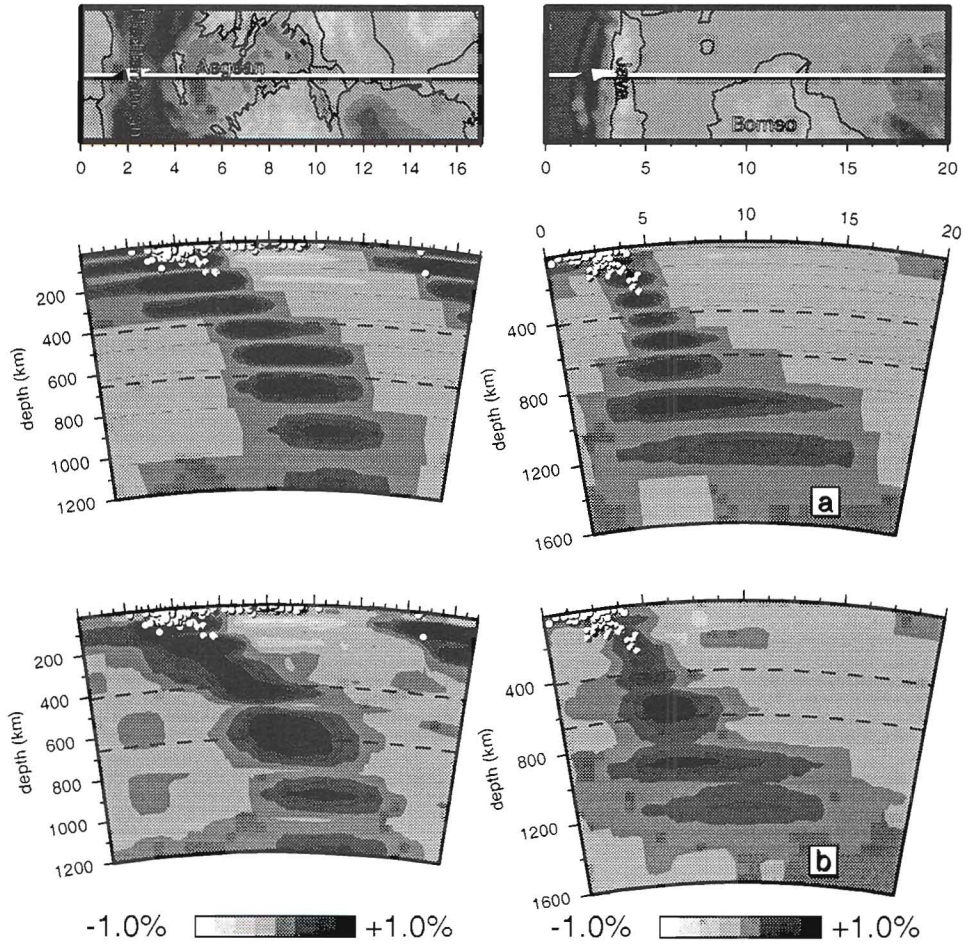
**Figure A9:** Cross-sections across the Mediterranean and Aegean Seas for the *P* model of Chapter 4 (a) and the *S* model of Chapter 6 (b).

**Figure A10:** Cross-sections across Java and Borneo for the *P* model of Bijwaard et al. (1998) (Chapter 4) (a) and the *S* model of Chapter 6 (b).





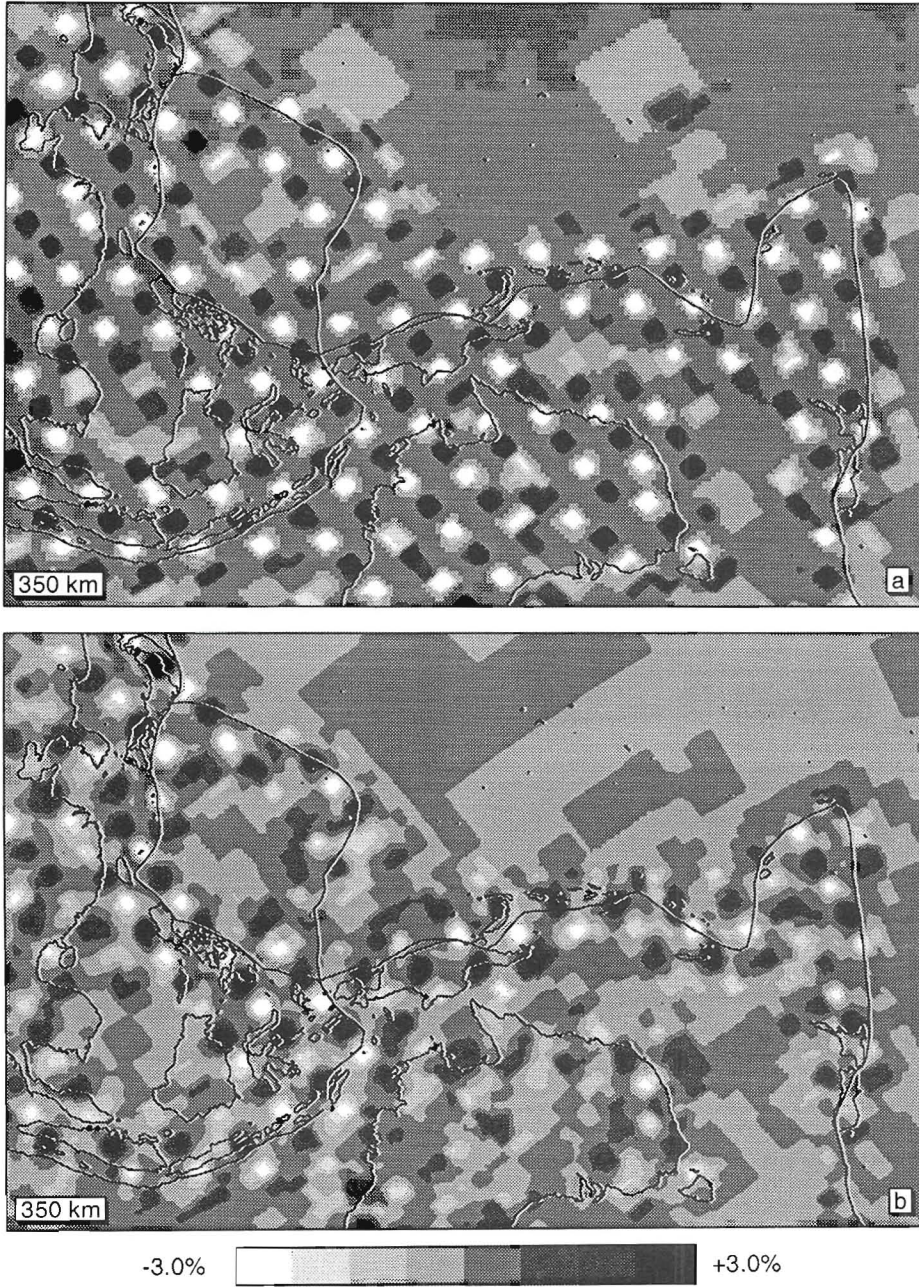
**Figure A11:** Comparison of upper mantle structure between the *P* model of Chapter 4 (a) and the *S* model of Chapter 6 (b) for southeast Asia to Tonga-Kermadec at 350 km depth. The lettering indicates the different subduction zones: J is Japan, IB is Izu-Bonin, M stands for Marianas, T is Taiwan, Ph is the Philippines, S is Sulawesi, Su is Sunda, B is Banda, So stands for Solomon Islands, V is Vanuatu, and TK is Tonga-Kermadec.



**Figure B9:** Input (top) and output (bottom) of a synthetic layer-cake test for the  $P$ -wave result along the same cross-section as depicted in Figure A9.

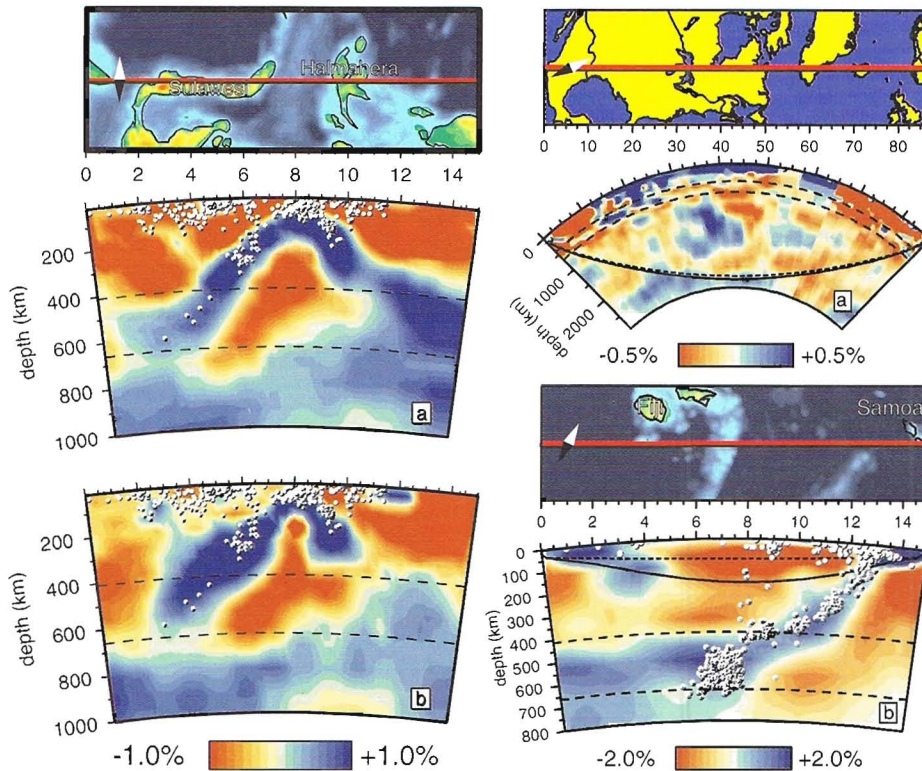
**Figure B10:** Input (top) and output (bottom) of a synthetic layer-cake test for the  $P$ -wave result along the same cross-section as depicted in Figure A10.





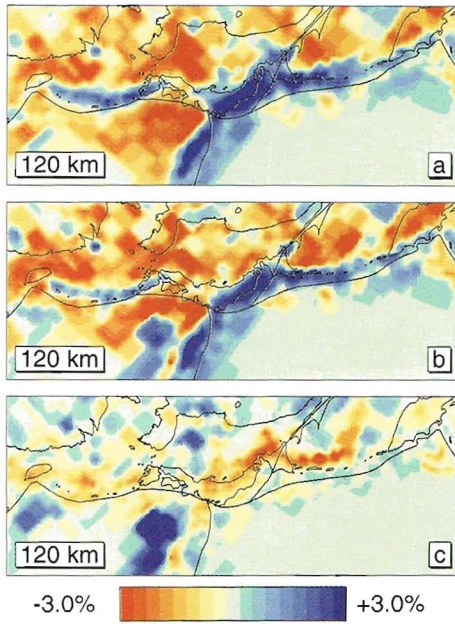
**Figure B11:** Input (top) and output (bottom) of a synthetic sensitivity experiment with 1.8° spikes of 5% for *P*-wave result shown in Figure A11



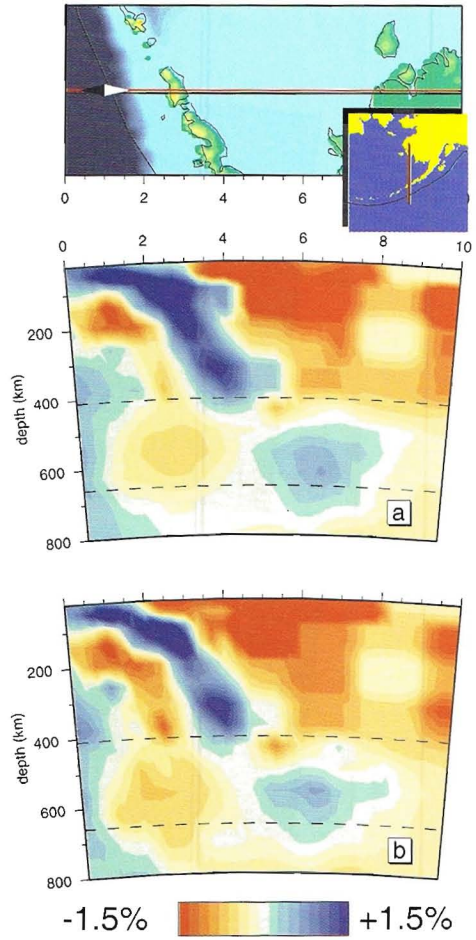


**Figure A12:** Cross-sections across Sulawesi and Halmahera in the  $P$  model of Bijwaard et al. (1998) (Chapter 4) (a) and the  $S$  model of Chapter 6 (b).

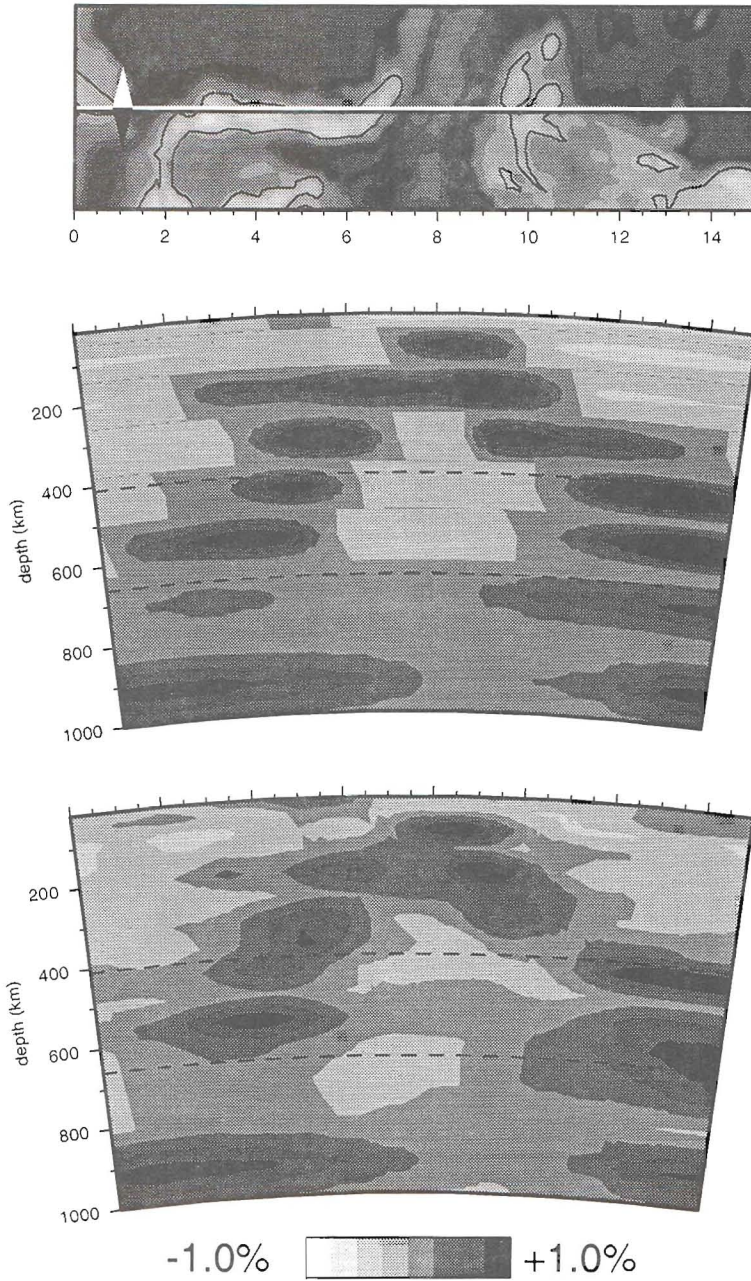
**Figure A13:** Two examples of (3-D) rays projected on 2-D cross-sections from the model of Chapter 4. Initial paths are dashed, final solutions are solid black. (a) cross-section from Mexico to Scandinavia; (b) cross-section from Samoa in the direction of the Fiji Islands. See Chapter 7 for comments.



**Figure A14:** Layer section for Japan and surroundings at 120 km depth: (a) the solution obtained in Chapter 4, (b) the nonlinear solution of Chapter 8, (c) the increment.

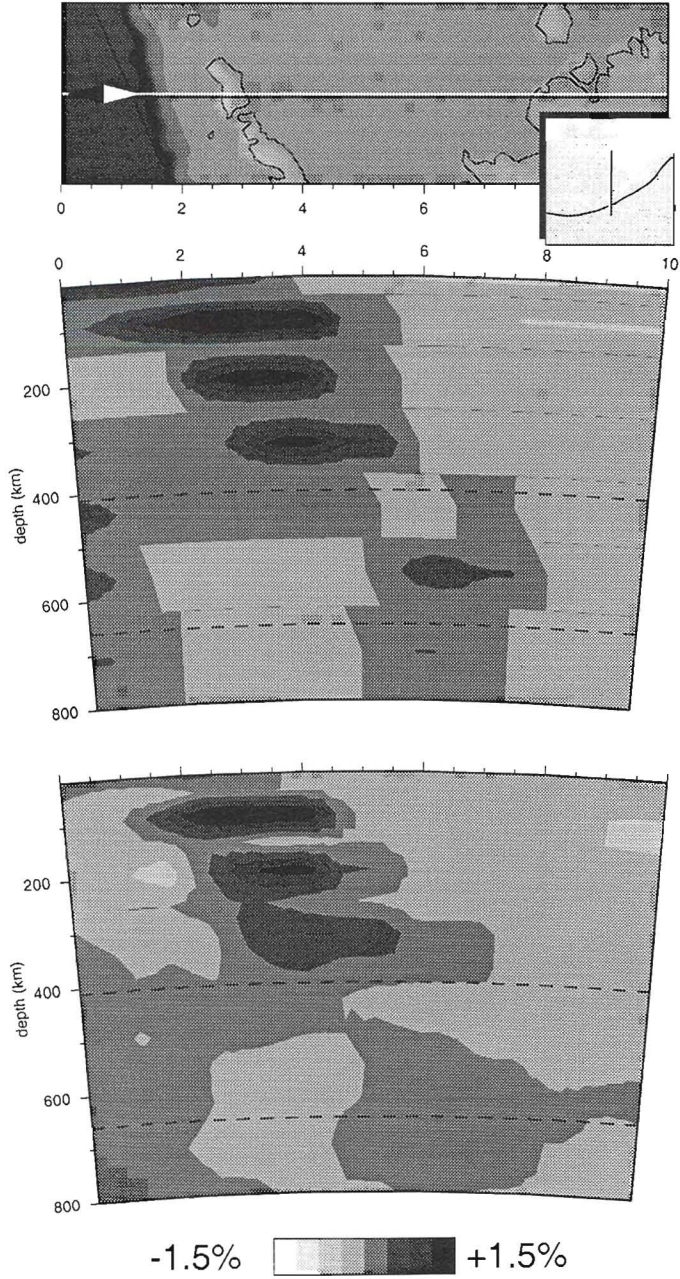


**Figure A15:** Cross-section across the Aleutian subduction zone: (a) the solution obtained in Chapter 4, (b) the nonlinear solution of Chapter 8. Notice the thinner slab structure in the nonlinear solution.

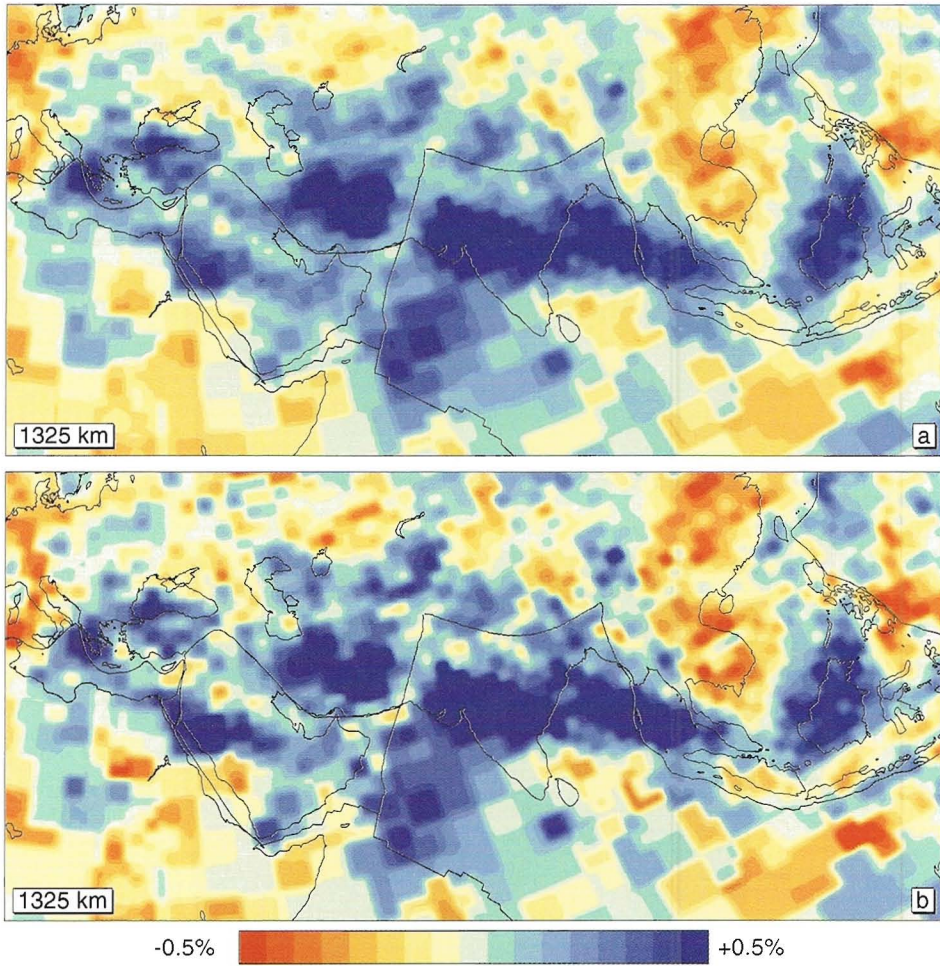


**Figure B12:** Input (top) and output (bottom) of a synthetic layer-cake test for the  $P$ -wave result along the same cross-section as depicted in Figure A12.





**Figure B15:** Input (top) and output (bottom) of a synthetic layer-cake test for the  $P$ -wave result along the same cross-section as depicted in Figure A15.



**Figure A16:** Layer section for the Tethys anomaly (slightly smoothed over distances of  $0.6^\circ$ ) at 1325 km depth: (a) the solution obtained in Chapter 4, (b) the nonlinear solution of Chapter 8.

## Samenvatting (Summary in Dutch)

In dit proefschrift worden nieuwe tomografische modellen van de hele aardmantel en -korst behandeld. Deze modellen tonen meer detail dan eerder behaalde globale tomografische resultaten en geven aanleiding tot nieuwe tektonische en geodynamische interpretaties. De verbetering van de bestaande modellen wordt bereikt in twee stappen. In de eerste stap (lineaire tomografie) wordt gebruik gemaakt van een nauwkeurige data set, die na selectie en middeling wordt geïnverteerd voor een onregelmatige parametrisatie in termen van cellen. In de tweede stap (niet-lineaire tomografie) wordt het model uit de eerste stap gebruikt om nauwkeurigere straalpaden van seismische golven te berekenen, looptijdresiduen te corrigeren en nieuwe inversies te plegen.

Kennis van het inwendige van de aarde wordt voornamelijk verkregen m.b.v. indirecte methoden. Het direct monstern van de diepe structuur is slechts technisch haalbaar tot een diepte van zo'n 12 km en beperkt zich aldus tot het bovenste deel van de aardkorst. Indirecte methoden gebaseerd op geologische of geochemische gegevens, metingen van het zwaartekrachts- of magnetische veld, de warmtestroom aan het aardoppervlak of de rotatie van de aarde leveren ofwel gedetailleerde informatie over de ondiepe (bovenste 50 km) structuur, ofwel grootschalige voorwaarden voor de diepe structuur. Hoogfrequente aardbevingsgolven zijn daarentegen gevoelig voor structuren van enkele tientallen km in de bovenmantel tot enkele honderden km in het diepste deel van de ondermantel (tot 3000 km diep). Zulke structuren zijn al aangetoond in de bovenmantel en bestaan waarschijnlijk ook in de ondermantel. Zij zijn inherent aan de convectie in de aardmantel waarbij relatief koud plaatmateriaal aan het aardoppervlak kan subduceren en zinken en hete materie van de kernmantelgrens kan opstijgen.

Seismische reistijdentomografie (de methode die in dit proefschrift wordt gebruikt) maakt gebruik van de reistijden van seismische golven om de variatie van de voortplantingssnelheden van deze golven in het inwendige van de aarde in kaart te brengen. Deze zgn.  $P$  en  $S$  snelheden zijn materiaaleigenschappen die afhankelijk zijn van temperatuur, druk en compositie. Door koud, gesubduceerd materiaal planten seismische golven zich sneller voort dan door warm materiaal. Variaties in seismische voortplantingssnelheden worden dan ook vaak geïnterpreteerd als temperatuurvariaties. Over de bijdragen van druk en compositie bestaat echter nog geen consensus en daarom wordt in dit proefschrift ook geen poging gedaan de verkregen modellen te interpreteren in termen van deze grootheden.

Seismische tomografie is afgeleid van een methode die in de geneeskunde wordt toegepast om het inwendige van het menselijk lichaam af te beelden, de zgn. computertomografie. Soortgelijke technieken zijn ontwikkeld in andere wetenschappen zoals de meteorologie, oceanografie, astronomie en helioseismologie. De verschillende vormen van seismische tomografie worden behandeld in hoofdstuk 2. Dit hoofdstuk geeft verder een overzicht van de belangrijkste ontwikkelingen en resultaten van dit vakgebied.

De opbouw van het resterende deel van dit proefschrift is als volgt. Hoofdstuk 3 is opgebouwd uit twee delen. Het eerste deel behandelt kort de theorie achter seismische reistijdentomografie en de benaderingen en aannames die daarbij gemaakt worden. Het tweede deel gaat in op de efficiënte verdeling van de aardmantel en -korst in onregelmatige cellen waarvan de afmetingen worden bepaald door de lokale dichtheid van seismische stralen. Deze dichtheid is in de aarde uiterst variabel t.g.v. de inhomogene verdeling van



aardbevingen en seismische stations aan het aardoppervlak.

Hoofdstuk 4 beschrijft de resultaten van de lineaire  $P$  tomografie die is uitgevoerd op de onregelmatige, globale parametrisatie behandeld in hoofdstuk 3. Het belangrijkste resultaat is de gedetailleerde structuur van het model in de dichtbemonsterde delen van het modelvolume, welke direct vergelijkbaar is met en soms een verbetering vormt van resultaten van regionale studies.

In hoofdstuk 5 worden vervolgens enkele goed opgeloste delen van het model uit hoofdstuk 4 verder geïnterpreteerd. De interpretaties omvatten o.a. de eerste afbeelding van een zgn. mantelpluim die oprijst vanaf de kern-mantelgrens, de ontdekking van de tot nog toe oudste gesubduceerde oceanplaat, de inwendige structuur van de gesubduceerde Tethys oceanplaat en gedetailleerde patronen van subductiezones in zuidoost Azië.

Hoofdstuk 6 behandelt de lineaire  $S$  tomografie die is uitgevoerd op gelijksoortige wijze als de  $P$  tomografie beschreven in hoofdstuk 4. De belangrijkste conclusie is dat de  $P$  en  $S$  resultaten zeer vergelijkbaar zijn in de goed opgeloste delen van beide modellen en dat verschillen (welke op afwijkende mantelcompositie zouden kunnen duiden) slechts bij hoge uitzondering voorkomen.

De laatste twee hoofdstukken tenslotte, zijn gewijd aan niet-lineaire tomografie. In hoofdstuk 7 worden twee verschillende methoden om straalpaden van seismische golven te berekenen verder ontwikkeld en met elkaar vergeleken. De meest efficiënte methode wordt vervolgens toegepast in hoofdstuk 8 bij de bepaling van een  $P$  model met niet-lineaire tomografie. Dit model wordt vervolgens vergeleken met het model van hoofdstuk 4 dat op lineaire wijze berekend is. De belangrijkste conclusie is dat de niet-lineaire aanpak niet leidt tot een heel ander model, maar vooral zorgt voor (subtiele) veranderingen in modeldelen met een sterke variatie in seismische  $P$  snelheid.

## Acknowledgments

Although I hope the scientific chapters form the most interesting part of this thesis, I am aware of the fact that the Acknowledgments section usually receives the first attention. This is due to the fact that this is the place where you might find out whose names also deserved to be on the cover of the booklet. Several people contributed either directly or indirectly to this thesis, some of whom have already been acknowledged in the previous chapters. The following people still need to be mentioned.

First of all, I would like to thank Roel Snieder for his willingness to be my promotor. His promptness in reading the different parts of the manuscript in a very short period of time enabled me to keep on schedule. For most of my scientific developments I am indebted to my co-promotor and daily advisor, Wim Spakman. He initiated this research and always believed in achieving its ambitious goals. More importantly, his door was always open, and there was always time for discussions about new results even though he usually had a busy schedule of his own. I am in this respect also indebted to his wife and children, who suffered the consequences of our enthusiasm. The only way I found to stop his continuous flow of new ideas and suggestions was a continuous return flow of written paper. This in turn led to interesting and often amusing debates about contents and context that have no doubt considerably improved the chapters of this thesis. My co-authors Bob Engdahl and Rob van der Voo are thanked for valuable discussions and explanations. This research could never have taken place without Bob's carefully processed data and it would never have become this successful without Rob's original interpretations. I am grateful to my colleagues at the Institute of Earth Sciences for their interest in my work and especially to Saskia Goes, Yvo S. Kok, and Axel Röhm for asking both difficult and simple questions.

The completion of this thesis was, however, not only a matter of scientific input. I would like to thank the 'AV-dienst' for their help with all sorts of visualizing problems and the system administrators, Joop Hoofd and Theo van Zessen for their support and advice in computational matters. The Geosciences Foundation (GOA) of the Netherlands Organization for Scientific Research (NWO) is thanked for their financial support and especially for the funding of fascinating 'field trips' to England, France, Greece, Israel, and the United States. I am grateful to David Brunet for his invitation to come to Toulouse and to Dave Yuen, Shigenori Maruyama, Daisuke Suetsugu, and the Building Research Institute in Tsukuba, Japan for the invitation for, the funding of, and the hospitality during my trip to Japan. I thank Rinus Wortel for the generous financial support for the publication of Chapter 4 and for a wonderful workshop on Cyprus.

Last, but certainly not least I would like to thank my parents for their support during all these years of study and their continued interest in whatever it was I was working on. And of course I am indebted to Suzanne for her patience and understanding and her ability to always get my mind off things.

Utrecht, 13 September 1999



# Curriculum Vitae

- 13 September 1971      Born in Eindhoven, the Netherlands
- 1983-1989              VWO (secondary school) 'het Rythoviuscollege' in Eersel,  
the Netherlands
- 1989-1994              Study of Geophysics (cum laude) at Utrecht University,  
the Netherlands
- 1994-1995              Research assistant at Utrecht University,  
the Netherlands
- 1995-1999              PhD student (OIO) in the Tectonophysics group,  
Institute of Earth Sciences, Utrecht University,  
the Netherlands