UNIVERSITY OF TOKYO

A methodology for inversion of seismic waveforms for elastic and anelastic structure and its preliminary application to the mantle transition zone beneath the Northwestern Pacific

by

Nobuaki Fuji

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Solid Earth Science Group
Department of Earth and Planetary Science

January 2012

"Choisir! c'est l'éclair de l'intelligence.
Hésitez-vous?... tout est dit, vous vous trompez."

Honoré de Balzac

ABSTRACT

We develop a quantitative and objective methodology for inversion of body-wave waveform data for elastic and anelastic structure and apply it to the upper mantle and mantle transition zone (MTZ) beneath the Northwestern Pacific, using PREM as the starting model. We derive the Jacobian for elastic and anelastic perturbations to the initial model in order to efficiently calculate the partial derivatives. This study is the first attempt to invert body-wave waveforms to estimate the radial dependence of regional anelastic structure. As waveform inversion has often been thought to be a "black-box-like" procedure, we explicitly show that the model obtained by the inversion is just the projection of the data onto the model space using orthogonal vectors. We use the singular value decomposition (SVD) as the basis for our inversion and use Akaike's Information Criterion (AIC) to evaluate the statistical significance of the models and determine the appropriate size of the basis. Models are examined using several statistical tests and also confirmed by comparison to models obtained using the conjugate gradient (CG) method. We obtain an average positive S-wave velocity anomaly in the MTZ beneath the Philippine Sea Plate and the Japan Sea. Our inversions obtained a lower Q_μ, as compared to PREM, for all of the subregions. A slight sub-regional dependence of Q_μ structure is found. Beneath the Philippines Sea region, we find $Q_\mu \sim 50$ throughout the MTZ down to 660 km, while we find an increasing Q_μ at depths of 500-600 km beneath the Pacific. This may reflect the water rich content beneath the mantle wedge driven down to the lowermost part of the MTZ in the Philippines Sea.

ACKNOWLEDGMENTS

I thank Professor Robert J. Geller for his frequent advice and suggestions regarding my work during the doctoral course. He represented the way of living as a scientist and what is important and what is not. I am very proud to work with him. I also thank Drs. Tatsuhiko Hara, Nozomu Takeuchi, Hiromitsu Mizutani, and Kenji Kawai for their frequent advice. Kensuke Konishi has been also a powerful co-worker. A significant portion of this thesis was published in *Physics of the Earth and Planetary Interiors* (doi:10.1016/j.pepi.2009.10.004). I thank my co-authors Drs. Kenji Kawai and Robert J. Geller for their permission to use this material in this thesis.

Drs. Satoru Honda, Hitoshi Kawakatsu, Satoshi Ide, Takashi Iidaka, Takao Ominato, Daisuke Suetsugu kindly reviewed all or part of this thesis. Drs. Shutaro Sekine, Nobumasa Funamori, Fabio Cammarano, Paul J. Tackley, Sébastien Chevrot, David A. Yuen, Jean-Paul Montagner, How-Wei Chen, Shun'ichiro Karato, Takehiko Yagi, Hikaru Iwamori, Dimitri Komatitsch, Ataru Sakuraba, Fumiko Tajima, Dapeng Zhao, Rafaël Garcia, Lapo Boschi, Frédéric Deschamps, George Helffrich, Craig R. Bina, Tomoeki Nakakuki, Masanori Kameyama, Marie Calvet, Annie Souriau, Luis Rivera, Eric Debayle, Maarten de Hoop, Micheal Bergman, Martha Savage, Miguel Miranda, Francis Albarède, Stéphane Labrosse, Barbara Romanowicz, Kei Kurita, Satoru Tanaka, Kazunori Yoshizawa, Thorne Lay, Mitsuhiro Matsu'ura, Gaku Kimura, Hiromitsu Toriumi, and Satoshi Kaneshima commented on and offered many fruitful suggestions on my work. Shuko Tanaka, Kazuaki Ota, Aki Takigawa, Thibault Duretz, Wataru Fujiya, Yasushi Ogasawara, Chiho Kusuda, Manabu Morishige, Annemarie Baltay, Ming-Che Hsieh, Hsiao-Lao Wang, Pei-Ying Lin, Joana Ribero, Risa Sakai, Kenji Aoki, Taichi Nozu, Ryo Kimura, Yumiko Wada, Fumiko Fukuda, Sumiko Kakizawa, Drs. Marc Monnereau, Takashi Nakagawa, Katia Pinheiro, Guizhi Zhu, William Moore, Shunsaku Horiuchi, John Hernlund, Mei Xue, Zhenzhen Yan, Stephane Rella, Gabriele Morra, Kei Hiraki, Tomoko Sato and Takahiko Uchide encouraged me every day or every week during my graduate studies.

I am grateful to all the faculty members of the Department of the Earth and Planetary Science, Graduate School of Science, University of Tokyo. Especially I thank those who trained me and gave talks in the "Whole Solid Earth Colloquium" and "Geodynamic Modeling Seminar." Faculty members in Departement Erdwissenschaften in Eidgenössische Technische Hochschule Zürich, National Central University of Taiwan, Centre National de la Recherche Scientifique (délégation Midi-Pyrénées), Faculdade de Ciências da Universidade de Lisboa, Kyushu University, Hiroshima University, Institut de Physique du Globe de Paris, Institut de Physique du Globe de Strasbourg, Ludwig-Maximilians-Universität München kindly welcomed me with their best gemütlichkeit.

I also deeply thank all my friends including my friends and staff in the Orquestra de Câmara de Tóquio which I established in 2007 with M.B. Eiji Yamada for concert tours in Portugal and Taiwan, the Tomin Metropolitan Orchestra, and the Todai Orchestra, for their heartfelt and encouraging words. I especially thank Maestros Sanae and Hitomi Sugawara for their top-level instruction in music and for their warm monthly parties in their house.

I thank the staff of the National Research Institute for Earth Science and Disaster Prevention for providing the F-net seismic broadband data. I also thank the staff of Incorporated Research Institutions for Seismology for providing data.

I thank the Japan Society for the Promotion of Science for awarding me a Reserch Fellowership for Young Scientists.

I deeply thank my parents, Kazue and Takeshi, and my sister Haruka for their daily support. I am also grateful to my grandparents, Kiyomi and Takehiko, Kimiko and Masao, and my relatives (Fujis, Shohojis, Pedrosos, and Dobbelaeres).

Contents

Abstract **iv**

Acknowledgements **v**

1 General Introduction **1**
1.1 Global seismology – from travel-time tomography to body-wave waveform inversion – . . . 2
1.2 1-D elastic structure in the lowermost mantle obtained using body-wave waveform inversion . . . 4
1.3 Mantle transition zone . . . 6

2 Methodology **9**
2.1 Synthetics and partial derivatives . . . 10
2.1.1 Synthetics . . . 10
2.1.2 Partial derivatives with respect to μ_0 and Q_μ^{-1} . . . 11
2.2 Inverse problem . . . 13
2.3 SVD inversion . . . 18
2.4 Criteria for evaluating models . . . 19

3 Dataset, static corrections, and "Initial Q_μ model" **23**
3.1 Event and station selection . . . 23
3.2 Static corrections . . . 24
3.3 Data selection criteria . . . 25
3.4 "Initial Q_μ" model . . . 25

4 Synthetic tests **39**
4.1 "Checkerboard" test for Q_μ perturbation . . . 40
4.2 Sensitivity of eigenvectors to elastic and anelastic structure . . . 41

5 1-D Simultaneous inversion results: upper mantle and mantle transition zone structure beneath Northwestern Pacific **49**
5.1 Technical issues . . . 50
5.1.1 Variance reduction: Born vs. re-linearized . . . 50
5.1.2 Iterative inversion . . . 51

5.1.3 Statistical estimation 52
5.1.4 Quality control stacks 55
5.1.5 Comparison of data to synthetics 56
5.2 Philippines Sea Region 56
5.3 Pacific Sea Region 57

6 Inversion using the conjugate gradient (CG) method 71
6.1 Conjugate gradient (CG) method 71
6.2 Technical issues 74
6.2.1 Sensitivity study of conjugate vectors to elastic and anelastic structure 75
6.2.2 VAR and AIC analysis 75
6.2.3 Statistical estimation 76
6.2.4 Projection of CG vectors onto the SVD space 78
6.2.5 CG inversion results 79

7 Discussion 93
7.1 Problems arising when using AIC as the SVD truncation criterion . 93
7.1.1 Scaling of model parameters 94
7.2 Initial model dependence 95
7.3 Forward modeling experiments 96
7.3.1 Forward modeling experiment regarding Fig. 3.4 96
7.3.2 Forward modeling of shorter-period waveforms 97
7.4 Robustness of the models 99
7.4.1 Anelastic structures and its possible geophysical implications 99

8 Conclusions 115

Chapter 1

General Introduction

Global seismology has greatly contributed to improving our knowledge of the structure of the Earth's interior and to understanding the dynamic processes in the Earth (e.g., mantle convection, thermal history, etc.). Analyses of signals traveling through the Earth's interior other than seismic waves generated by earthquakes also are contributing to our knowledge. For example, cross-correlation analysis of seismic hum (the Earth's background free oscillation signals) has contributed to elastic global tomography (e.g, Rhie and Romanowicz, 2004; Nishida et al., 2009). Some groups have started to use neutrino radiography to obtain independent information on density (e.g., Gonzalez-Garcia et al., 2008; Borriello et al., 2009). Electromagnetic tomography is now also being used to extract information on water content or the thermal gradient in the deep Earth (e.g., Seama et al., 2007; Pinheiro and Jackson, 2008; Utada et al., 2009; Shimizu et al., 2009). Although such observations are significant contributions to all fields of solid Earth science (e.g., geodynamic modeling, mineral physics, geochemistry), seismic signals from earthquakes are and will continue to be our primary source of information on the Earth's interior. This dates back as far as the work of Jeffreys and Bullen [J-B model] (1940), which established travel-time tables and obtained a 1-D Earth model accurate enough to justify the linearization of the earthquake location computation (Geiger, 1912), and extends to the modern, detailed 3-D images of the Earth being obtained by several new techniques. We can still expect further

progress from analyses of the large volume of seismic waveform data which has already been recorded but has not yet been fully utilized due to technical problems such as heavy computational requirements and limitations of present data processing methods.

This thesis presents techniques to use waveform data to efficiently and accurately extract information on Earth structure and demonstrates the feasibility of this methodology by applying these techniques to real datasets. In this chapter, we review technical developments in global seismology and their application, theoretical background and our group's previous studies on the structure of the lowermost mantle. Since this study is our first attempt to simultaneously invert for elastic and anelastic structure and its target is the upper mantle and the mantle transition zone (MTZ) beneath the Northwestern Pacific, we also review issues concerning the MTZ and anelastic structure.

1.1 Global seismology – from travel-time tomography to body-wave waveform inversion –

One of the ultimate aims of the field of global seismology is to fully extract all of the information on the structure of the Earth's interior, including anisotropy and attenuation, contained in observed seismic waveforms. Recent deployments of dense broadband array networks such as IRIS, US-array, F-net and Hi-net have improved the dataset but further deployment of seismometers is still desirable for the rest of the globe. However, seismologists have not yet fully utilized the waveform data already obtained by present regional/global networks properly and efficiently, as they have lacked both the computing power and the necessary theoretical methods to treat them. They thus have used secondary information such as travel-times in inversion studies. This thesis is intended to further improve techniques for analysis of seismic waveform data.

One of the most powerful tools for extracting information on the Earth's interior is seismic tomography, which analyzes travel time data to determine 3-D Earth structure. Since the initial studies in the 1970s, seismic tomography (e.g., Sengupta and Toksöz, 1976; Aki et al., 1977; Dziewonski et al., 1977) has used increasingly large volumes of travel time data (e.g., the millions of arrival-time data available from the International Seismological Centre – ISC) to determine 3-D global fine structure (e.g., Fukao et al., 1992; Masters et al., 2000; Obayashi et al., 2006). The success of tomographic studies is partly due to improved methods developed over the past 30 years, but mainly because of the tremendous growth in the size of the datasets.

On the other hand, although there have been many studies which used body-wave waveforms to determine 1-D structure models for particular regions, almost all such studies have been conducted by forward modeling (e.g., Lay and Helmberger, 1983; Grand and Helmberger, 1984; Brudzinski et al., 1997; Tajima and Grand, 1998). These and other similar studies computed synthetic seismograms using either generalized ray theory or reflectivity, which use earth-flattening and other approximations. They then visually compare the observed waveform data to the synthetic seismograms and adjust the earth model by trial-and-error to improve the fit. Despite the many obvious problems, trial-and-error forward modeling studies have made many important contributions to our knowledge of the fine structure of the Earth's deep interior. However, we show in this thesis that automated inversion procedures using much larger datasets can make further significant advances.

There have already been relatively many waveform inversion studies using mainly long-period surface waves (e.g., Woodhouse and Dziewonski, 1984; Tanimoto, 1984; Nolet, 1990; Li and Romanowicz, 1996; Mégnin and Romanowicz, 2000; Hara, 2004; Takeuchi and Kobayashi, 2004; Takeuchi, 2007). Most such studies derived global models with a relatively low resolution. However, because of the efficient and accurate calculation of synthetic seismograms and partial derivatives, Hara (2004) was able to derive models with relatively better resolution.

One obstacle to performing waveform inversion of broad-band body-wave waveforms has been the need for accurate and efficient methods for computing synthetic seismograms and their partial derivatives. Our group has recently developed algorithms based on the Direct Solution Method (DSM: Geller and Ohminato, 1994) for computing highly accurate synthetics for spherically symmetric transversely isotropic Earth models (Kawai et al., 2006, and references cited therein). To date our applications of waveform inversion to Earth structure have only been for 1-D models, but our approach is formulated for application to 3-D problems (Geller and Hara, 1993; Cummins et al., 1997; Takeuchi et al., 2000). One previous study applied waveform inversion to short-period body waves (Gao et al., 2006). Gao et al. conducted waveform inversion of short-period regional array data (around 100 waveforms in total) for two events for the localized 1-D P- and S-wave velocity structure of the mantle transition zone. The reflectivity method was used to compute the synthetics and differential seismograms were used to approximate partial derivatives rather than exact theoretical results (e.g., Tarantola, 1984, Geller and Hara, 1993). Matzel and Grand (2004) also conducted waveform inversion for transversely anisotropic structure.

1.2 1-D elastic structure in the lowermost mantle obtained using body-wave waveform inversion

We have recently conducted waveform inversion of transverse component data for the localized 1-D S-wave velocity structure of D″ (Kawai et al., 2007a,b; Konishi et al., 2009; Kawai et al., 2009). The author of this thesis contributed as a co-author of the above studies and contributed in particular to the development of data processing procedures and the development of the inversion scheme, which is presented in this thesis as well. In this section, we briefly review the above four papers (see also Kawai and Geller, 2010).

The D″ layer was first defined as the lowermost 200 km of the mantle (Bullen, 1949). Since Lay and Helmberger (1983) detected the D″ discontinuity from a S-wave triplication, 1-D or quasi 2-D models for D″ beneath several regions have been obtained using methods such as stacking and forward modeling (e.g., Garnero and Lay, 2003; Thomas et al., 2004; Maupin et al., 2005; Lay et al, 2006). Both these models and the above models obtained using body-wave waveform inversion have contributed to our understanding of geodynamics (e.g., Hernlund et al., 2005; Nakagawa and Tackley, 2006; Nakagawa and Tackley, 2008; Kawai and Tsuchiya, 2009).

Global tomography using travel time data (e.g., Kennett et al., 1998; Masters et al., 2000) and waveform inversion (e.g., Panning and Romanowicz, 2004; Takeuchi, 2007) provides information on the average seismic velocity anomaly in D″ but lacks the resolution to reveal fine structure within D″. Kawai et al. (2007a) first conducted regional body-wave waveform inversion for the fine structure in the lowermost mantle beneath Central America and found higher S-velocities in the upper half of D″ and lower velocities in the lower half and the velocity decrease in the lower half. Kawai et al. (2007b) obtained similar results for D″ beneath the Arctic. Hernlund et al. (2005) suggested the possible existence of a "double crossing" phase transition in D″, but Kawai and Tsuchiya (2009) recently showed that thermal effects are sufficient to account for the observed fine velocity structure in D″, and that a double crossing phase transition would lead to significantly lower values of S velocity near the CMB than those actually observed.

Konishi et al. (2009) found a more complicated structure, which they call an "S-shaped" velocity model, for D″ beneath the Western Pacific. Their model suggests the existence of phase transitions at several depths within D″ and is consistent with the existence of depleted MORB in the lowermost mantle beneath the Western Pacific (e.g., Hirose et al., 1999; Xie and Tackley, 2004). The above studies show the important contribution to our understanding of mantle dynamics that can be made by seismic fine structure obtained by waveform inversion studies. In all of the waveform inversion studies cited above, the variance reduction for each

individual waveform only improves marginally because of the noisy nature of the data.

Waveform inversion is still widely considered to be a "black box" method. Some efforts have been made in the studies cited above (e.g., Konishi et al., 2009) to present visual confirmation of the waveform inversion results. This thesis makes a further contribution to validating the waveform inversion methodology by showing results in model space, in data space, and in the intermediate orthogonal space (SVD and CG space).

1.3 Mantle transition zone

The region we studied in this thesis, the mantle transition zone, is one of the most interesting regions in the Earth's interior. After Bullen (1949) named it the "C region," Birch (1952) predicted that some mineralogical phase transitions would occur in the depth range from 400 to 700 km. It later was confirmed that the major component in the upper mantle, olivine, transforms to wadsleyite at a pressure of about 12 GPa, corresponding to a depth of 410 km. Wadsleyite transforms to ringwoodite at a pressure of 15 GPa, corresponding to the 520 km discontinuity (e.g., Akimoto and Fujisawa, 1968; Ringwood and Major, 1970; Bina and Wood, 1987; Katsura and Ito, 1989). Ringwoodite transforms to (Mg,Fe)SiO_3 perovskite and (Mg,Fe)O magnesiowüstite at a pressure of 24 GPa, corresponding to the 670 km discontinuity. These discontinuities can be seismically detected by stacking waveforms (e.g., Shearer, 1990). Mantle convection is enhanced when the sign of the Clapeyron slope is positive, while convection is suppressed when the sign is negative (Schubert et al., 1975). While the Clapeyron slope is positive for the 410 and 520 km discontinuities, it is negative for the 670 km discontinuity. Therefore, the phase transition from ringwoodite to (Mg,Fe)SiO_3 perovskite and (Mg,Fe)O magnesiowüstite tends to suppress mantle convection. Due to the opposite signs of the Clapeyron slope for the 410 and 670 km discontinuities, the range between

these two discontinuities works as a thermostat (e.g., Suetsugu et al., 2004). Although the absolute value of the Clapeyron slope is still controversial, a negative slope may induce slab stagnation or two-layered convection, for which the 670 km discontinuity is a thermal boundary (e.g., Christensen and Yuen, 1984; Christensen and Yuen, 1985). Wen and Anderson (1996) support the two-layered convection model based on long-wavelength geoid and topographic studies. Geochemists also support two-layered convection due to evidence that MORB is depleted in incompatible elements (e.g., Hofmann, 1997; Albarède and van der Hilst, 2002).

The question of whether the mode of mantle convection is whole mantle or two-layered is important because it affects the earth's evolution. Recent studies on simulation of mantle convection support whole mantle convection (e.g., Tackley, 2000; Deschamps et al., 2009). Various seismic tomography studies have suggested both slab stagnation (Fukao et al., 1992) and slab penetration (Widiyantoro et al., 1999). Both tomographic models, however, support whole mantle convection to some extent. The stagnation and penetration scenarios can be explained by several mechanisms such as migration of trenches which might reduce the downward stress of the slab (e.g., Christensen, 1996; Christensen, 1997; Cizkova et al., 2002; Torii and Yoshioka, 2007), a viscosity jump at the 660 km phase transition which might support negative buoyancy of slabs and reduce the negative strength of the slab against the lower mantle (e.g., Christensen, 1996; Torii and Yoshioka, 2007), grain size reduction of the cold slab at the 410 km (e.g., Cizkova et al., 2007), or thermodynamic factors which might reduce the slab's negative buoyancy (e.g., Ita and King, 1998). The depleted incompatible elements in MORB can be explained by either water-filtering (Bercovici and Karato, 2003) or density cross-over (Irifune and Ringwood, 1987), which does not require a two-layered mantle model. The existence of water in the transition zone has been suggested based on seismological data (Song et al., 2004).

The mantle consists chiefly of olivine, which is well known to be anisotropic (e.g., Zhang et al., 2000; Jung and Karato, 2001; Katayama et al., 2004; Kneller et al., 2005). Because preferred orientations like LPO (lattice preferred orientation) and SPO (shaped preferred orientation), which grow under high shear stress, are

related to mantle flow (Park and Levin, 2002), seismic anisotropy can be interpreted as indicative of mantle flow. Many studies using S-wave splitting have been performed to detect anisotropic structure in the upper mantle and the MTZ (e.g., Silver and Savage, 1994; Vinnik and Montagner, 1996; Savage, 1999). But because only the delay time δt and the polarization ϕ are observed by shear wave splitting studies, the interpretation of these observations is uncertain (e.g., in terms of multiple isotropic layers which are aligned heterogeneously, aligned, fluid-filled fractures or cracks, aligned anisotropic minerals, water contrast, etc.). For long-period surface-wave-like waveforms, waveform inversion has been conducted for azimuthal anisotropy (e.g., Vinnik et al., 1989), but there have been few if any studies using body-wave waveform inversion to determine the anisotropic elastic parameters.

Another important question concerning the upper mantle and the MTZ is seismic anelastic attenuation. Accurate measurement of anelastic attenuation can help to constrain the physical and chemical state of the interior of the Earth, as anelastic attenuation reflects physical properties in the MTZ such as temperature, water content, and composition (e.g., Karato, 1993; Faul and Jackson, 2005). There have been several studies of attenuation structure, mainly using relatively long period (from 60 up to 1000 s) surface wave waveforms (e.g., Romanowicz, 1995, 1998; Gung and Romanowicz, 2004) or surface wave amplitudes (e.g., Romanowicz, 1994; Dalton et al., 2008). Mineral physics models can be tested against long-period seismic observations by direct forward modeling to find the best-fitting average 1-D thermal structures, and construct radial profiles of seismic velocity and seismic attenuation (Cammarano et al, 2008, 2009). Attenuation around Japan was also previously studied using stacking by Revenaugh and Jordan (1991) and Kato et al. (2001). Extracting information on attenuation from recorded seismograms requires careful treatment of elastic structure because focusing and defocusing effects or scattering effects can also affect the amplitudes. This thesis is the first attempt to utilize inversion of relatively short period (20-200 s) body-wave waveform data, to simultaneously invert for 1-D elastic and anelastic structure.

Chapter 2

Methodology

Fig. 2.1 shows the overall flowchart of our inversion procedures. First we choose a target region, select events and stations, and calculate synthetic seismograms and partial derivatives for the initial model. After correcting for the effect of local structure near the stations and sources by making static corrections (time shifts), we check whether each of the observed seismograms satisfies criteria for inclusion in the dataset.

As shown in Fig. 2.1, we perform waveform inversion to minimize the variance between the observed seismograms and synthetic seismograms, subject to regularization. Since this inversion is non-linear, iterative linearized inversion is desirable. However, due to the computational requirements we are not presently routinely performing iterative inversion. In this thesis we perform one iterative inversion as an example (see the subsection 5.1.2).

We have developed software which automatically performs the various steps in Fig. 2.1. Some parts of this software are publicly available from our website (http://www-solid.eps.s.u-tokyo.ac.jp/~dsm).

2.1 Synthetics and partial derivatives

In order to be proper enough to conduct waveform inversion, we are required to compute synthetics and their partial derivatives. Our group has developed accurate and efficient methods based on the Direct Solution Method (DSM: Geller and Ohminato, 1994; Cummins et al., 1994ab; Geller and Takeuchi, 1995; Takeuchi et al., 1996; Takeuchi et al., 2000; Kawai et al., 2006) for computing highly accurate synthetic seismograms in a spherically symmetric isotropic Earth model. Here we present the methodology for forward modeling synthetics and their partial derivatives. Because this thesis aims at simultaneous inversion for elastic and anelastic structure of the Earth's interior, we derive an efficient formulation for both elastic and anelastic perturbation using Jacobian in the frequency domain.

2.1.1 Synthetics

The DSM obtains the solution of the weak form of the equation of motion by directly solving the Galerkin weak form of the equation of motion:

$$(\omega^2\mathbf{T} - \mathbf{H} + \omega\mathbf{R})\mathbf{c} = -\mathbf{g}, \tag{2.1}$$

where $\mathbf{T}$ is the mass matrix, $\mathbf{H}$ is the stiffness matrix, $\mathbf{g}$ is the force vector, and $\mathbf{R}$ enforces continuity conditions at fluid-solid boundaries, which exist only for the spheroidal (P-SV) case. Our methods do not use geometrical optics or Earth-flattening approximations in computing the synthetics and their partial derivatives. The DSM efficiently computes highly accurate synthetic seismograms in a spherically symmetric transversely isotropic Earth model (Kawai et al., 2006). Although the publicly released software is only for 1-D media at present, this approach can also be applied to 3-D problems (e.g., Takeuchi et al., 2000).

2.1.2 Partial derivatives with respect to μ_0 and Q_μ^{-1}

When we compute the partial derivatives with respect to μ_0 (the real part of the elastic parameter at the reference frequency) and Q_μ (the anelastic attenuation parameter, which we assume to be frequency-independent within the frequency band being considered), we have to take anelastic attenuation into account. Azimi's law (Azimi et al., 1968) gives the frequency dependence of μ as follows:

$$\mu(\omega) \approx \mu_0 \left(1 + \frac{2q \ln(\omega/\omega_0)}{\pi}\right)(1 + iq) \tag{2.2}$$

where $i = \sqrt{-1}$, $q = Q_\mu^{-1}$, ω is the angular frequency, ω_0 is the reference angular frequency (a reference frequency of 1 Hz, i.e., $\omega_0 = 2\pi$, is used in this thesis), and $\mu_0 = \mathrm{Re}(\mu(\omega_0))$. We thus have:

$$\frac{\partial \mu(\omega)}{\partial \mu_0} = \left(1 + \frac{2q \ln(\omega/\omega_0)}{\pi}\right)(1 + iq) \tag{2.3}$$

$$\frac{\partial \mu(\omega)}{\partial q} = \mu_0 \left[\frac{2 \ln(\omega/\omega_0)}{\pi} + i\left(1 + \frac{4q \ln(\omega/\omega_0)}{\pi}\right)\right]. \tag{2.4}$$

We use the above results to obtain the perturbation to the shear modulus, $\Delta\mu(\omega)$, for perturbations to the shear modulus at the reference frequency or to the anelastic parameter, $\Delta\mu_0$ or Δq, respectively. We obtain:

$$\Delta\mu(\omega) = \frac{\partial \mu(\omega)}{\partial \mu_0}\Delta\mu_0 + \frac{\partial \mu(\omega)}{\partial q}\Delta q. \tag{2.5}$$

We use eq. (2.5) to compute the perturbation to the synthetic seismogram, $\Delta u(\omega)$, for perturbations to μ_0 or q. We have:

$$\Delta u(\omega) = \frac{\partial u(\omega)}{\partial \mu}\Delta\mu(\omega), \tag{2.6}$$

where

$$\frac{\partial u(\omega)}{\partial \mu}$$

is the partial derivative of the synthetic seismogram $u(\omega)$ with respect to $\mu(\omega)$.

In this thesis, we used $\Delta\mu_0 = 1$ GPa where μ_0 ranges from 20 to 100 GPa and $\Delta q = 1/1000$ where q is around $1/100$ in order to roughly normalize the amplitude of Δu, except that we make a few experiments to study the effect of varying the choice of scaling in Chapter 7.

The partial derivatives for a "voxel" perturbation to a 3-D starting model are obtained from the following result, which is based on the formulation of Geller and Hara (1993):

$$\left\{ \frac{\partial u_i^{(k)}(\omega)[\mathbf{r}^{(p)}]}{\partial m_l} \right\}^* = \int_V (\omega^*)^2 [u_j^{(k)}][\rho^{(l)}]^* v_j^{i(p)} dV - \int_V [u_{j,q}^{(k)}]^* [C_{jqrs}^{(l)}]^* v_{r,s}^{i(p)} dV, \quad (2.7)$$

where u_i is the i-component of the displacement, $v_i^{j(p)}$ is the i-component of the back-propagated displacement excited by a point force in the j-direction at the p-th station, summation over j, q, r, and s is implied, and

$$u_{i,j} = \frac{\partial u_i}{\partial x_j}$$

is the locally Cartesian derivative.

In this thesis, we compute partial derivatives for a spherically symmetric perturbation to a spherically symmetric, anelastic, starting model for the 1-D inversion. The perturbations to the density and the elastic constants at a particular depth $\mathbf{r}_l$ are written respectively as $\delta\rho^{(l)}\Psi_l(\mathbf{r})$ and $\delta C_{jqrs}^{(l)}\Psi_l(\mathbf{r})$, where $\Psi_l(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_l)$ in this study. Note that, as discussed above, we include the effect of physical dispersion in $C_{jqrs}^{(l)}$. In this thesis we invert only for toroidal (SH) S-wave velocity and anelastic attenuation Q_μ^{-1}, but the above formulation can be easily extended

to allow inversion for Q_κ^{-1} and P- and SV-wave velocities as well as an isotropic or 3-D structure.

2.2 Inverse problem

We define $\delta\mathbf{d}^{(kp)}$ to be the residual (the difference of the observed seismogram, $\mathbf{d}_{\mathrm{OBS}}^{(kp)}$, and the synthetic seismogram for the initial model, $\mathbf{d}_{\mathrm{INIT}}^{(kp)}$) for the k-th event and the p-th station:

$$\delta\mathbf{d}^{(kp)} = \mathbf{d}_{\mathrm{OBS}}^{(kp)} - \mathbf{d}_{\mathrm{INIT}}^{(kp)}. \tag{2.8}$$

For simplicity we assume here that we have K events and P stations with records available at all stations for all events, and we write the corresponding partial derivatives with respect to the l-th model parameter (in this study either μ_0 or Q_μ^{-1} for some depth) as

$$\frac{\partial\mathbf{u}^{(kp)}}{\partial m_l}$$

for all M model parameters. We also assume here that each time series has T points, and that data from all stations are available for all earthquakes; in this case the vector $\delta\mathbf{d}$ has dimension KPT and the matrix $\mathbf{A}$ has dimension $KPT \times M$. If, on the other hand, the various time series have different lengths and/or not all stations are available for all events, this can be handled by a straightforward extension of the above.

The matrix of partial derivatives $\mathbf{A}^{(kp)}$ for the k-th event and p-th station is written as follows:

$$\mathbf{A}^{(kp)} = \left[\frac{\partial\mathbf{u}^{(kp)}}{\partial m_1} \quad \frac{\partial\mathbf{u}^{(kp)}}{\partial m_2} \quad \cdots \frac{\partial\mathbf{u}^{(kp)}}{\partial m_l} \right]. \tag{2.9}$$

In order to conduct a simultaneous inversion for all event (k) and station (p) pairs, we combine the residuals $\delta\mathbf{d}^{(kp)}$ and partials $\mathbf{A}^{(kp)}$ into a single vector $\delta\mathbf{d}$ and a single matrix $\mathbf{A}$, respectively, as follows:

$$\delta\mathbf{d}^{\mathbf{T}} = \left[\delta\mathbf{d}^{\mathbf{T}(11)} \quad \cdots \delta\mathbf{d}^{\mathbf{T}(1P)} \delta\mathbf{d}^{\mathbf{T}(21)} \quad \cdots \delta\mathbf{d}^{\mathbf{T}(K1)} \quad \cdots \delta\mathbf{d}^{\mathbf{T}(KP)}\right] \tag{2.10}$$

$$\mathbf{A}^{\mathbf{T}} = \left[\mathbf{A}^{\mathbf{T}(11)} \quad \cdots \quad \mathbf{A}^{\mathbf{T}(1P)} \mathbf{A}^{\mathbf{T}(21)} \quad \cdots \quad \mathbf{A}^{\mathbf{T}(K1)} \quad \cdots \quad \mathbf{A}^{\mathbf{T}(KP)}\right]. \tag{2.11}$$

The inverse problem is usually written as:

$$\mathbf{A}\delta\mathbf{m} = \delta\mathbf{d} \tag{2.12}$$

where $\delta\mathbf{m}$ is the perturbation to the initial model. Since the number of unknowns (number of elements of $\delta\mathbf{m}$) is usually far smaller than the number of data points (number of elements of $\delta\mathbf{d}$), it is well known that eq. (2.12) implies the following minimization rather than strict equality:

$$|\mathbf{A}\delta\mathbf{m} - \delta\mathbf{d}|^2 = \text{minimum}. \tag{2.13}$$

Fig. 2.2 shows a schematic explanation of the inversion process. We use the SVD to decompose the matrix $\mathbf{A}$. The singular value decomposition (SVD) has been widely discussed in the geophysical literature (e.g., Wiggins, 1972; Menke, 1984). However, as discussed below, their results can be considerably simplified for the case when $\mathbf{A}^{\mathbf{T}}\mathbf{A}$ is non-singular. We obtain:

$$\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\mathbf{T}}, \tag{2.14}$$

where the non-zero eigenvalues are ordered as follows:

$$\lambda_1^2 \geq \lambda_2^2 \geq \cdots \geq \lambda_M^2 > 0. \tag{2.15}$$

Note that we are assuming that the problem is overdetermined (i.e., $M < KPT$), and that the first M eigenvalues of $\mathbf{A}$ (as defined in eq. 2.15) are non-zero, i.e., that $\mathbf{A^T A}$ is positive definite.

The unitary $M \times M$ matrix $\mathbf{V}$ contains the right eigenvectors of $\mathbf{A}$:

$$\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_M] \tag{2.16}$$

normalized to satisfy:

$$\mathbf{v}_j^{\mathbf{T}} \mathbf{v}_l = \delta_{jl}, \tag{2.17}$$

the unitary $KPT \times KPT$ matrix $\mathbf{U}$ contains the left eigenvectors of $\mathbf{A}$, normalized to satisfy:

$$\mathbf{u}_j^{\mathbf{T}} \mathbf{u}_l = \delta_{jl}, \tag{2.18}$$

and the $KPT \times M$ matrix $\mathbf{\Lambda}$ contains the eigenvalues:

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_M \\ & \mathbf{0} & \end{pmatrix}, \tag{2.19}$$

where δ_{jk} is a Kronecker delta. It follows from eqs. (2.14), (2.17) and (2.18) that:

$$\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{u}_j. \tag{2.20}$$

From eqs. (2.14) and (2.18) we have

$$\mathbf{A^T A} = \mathbf{V \Lambda^T U^T U \Lambda V^T} = \mathbf{V \Lambda^T \Lambda V^T}. \tag{2.21}$$

From eqs. (2.17) and (2.21) we have

$$\mathbf{v}_j^{\mathbf{T}} (\mathbf{A^T A}) \mathbf{v}_k = \lambda_k^2 \delta_{jk}. \tag{2.22}$$

We now use the above results to obtain the least squares solution of eq. (2.13) by solving the normal equations:

$$\mathbf{A^T A} \delta \mathbf{m} = \mathbf{A^T} \delta \mathbf{d}. \tag{2.23}$$

We write the solution of eq. (2.23) as an eigenfunction expansion:

$$\delta \mathbf{m} = \sum_{j=1}^{M} e_j \mathbf{v}_j = \mathbf{V e}, \tag{2.24}$$

where the expansion coefficients e_j are the unknowns, and the vector $\mathbf{e}$ is defined to be:

$$\mathbf{e^T} = (e_1 \; e_2 \; \cdots \; e_M). \tag{2.25}$$

We substitute eq. (2.24) into eq. (2.23), multiply both sides of the resulting equation on the left by $\mathbf{v}_j^{\mathbf{T}}$ and use eq. (2.22) to solve for each expansion coefficient of eq. (2.24) as follows:

$$e_j \;\; = \;\; \frac{\mathbf{v}_j^{\mathbf{T}} \mathbf{A^T} \delta \mathbf{d}}{\lambda_j^2}. \tag{2.26}$$

We can write eq. (2.26) in matrix form as follows:

$$\mathbf{e} = (\mathbf{\Lambda^T \Lambda})^{-1} \mathbf{V^T A^T} \delta \mathbf{d}. \tag{2.27}$$

The numerator of eq. (2.26) shows that the expansion coefficient of the j-th eigenvector will become larger as the vector $\mathbf{A^T}\delta\mathbf{d}$ becomes more nearly parallel (in the multi-dimensional vector space) to the eigenvector $\mathbf{v}_j$, while the denominator of eq. (2.26) shows that as the eigenvalues become smaller the expansion coefficients will tend to be magnified, and the solution of the inverse problem will thus tend to become unstable.

Assuming that the data are uncorrelated and all have equal variance σ_d^2, we write the covariance matrix for $\mathbf{e}$ as follows in the eigenspace of $\mathbf{A^T A}$:

$$\begin{aligned} \text{covar. } (\mathbf{e}) &= \sigma_d^2 \, (\mathbf{\Lambda^T \Lambda})^{-1} \mathbf{V^T A^T A V} (\mathbf{\Lambda^T \Lambda})^{-1} \\ &= \sigma_d^2 \, (\mathbf{\Lambda^T \Lambda})^{-1} \mathbf{V^T} (\mathbf{V \Lambda^T \Lambda V^T}) \mathbf{V} (\mathbf{\Lambda^T \Lambda})^{-1} \\ &= \sigma_d^2 \, (\mathbf{\Lambda^T \Lambda})^{-1}. \end{aligned} \tag{2.28}$$

If we transform eq. (2.28) from the eigenspace to the model space, the covariance matrix is written as:

$$\begin{aligned} \text{covar. } (\delta\mathbf{m}) &= \sigma_d^2 \, (\mathbf{A^T A})^{-1} \\ &= \sigma_d^2 \sum_{j=1}^{M} \frac{\mathbf{v}_j \mathbf{v}_j^\mathbf{T}}{\lambda_j^2} \\ &= \sigma_d^2 \, \mathbf{V} (\mathbf{\Lambda^T \Lambda})^{-1} \mathbf{V^T}. \end{aligned} \tag{2.29}$$

Note that as has been pointed out by several previous works (e.g., Okal and Geller, 1979), the errors due to parameters not included in the inversion cannot be readily quantified. For example, the above formal covariance estimates cannot directly be taken as the "real" standard deviation, because the model does not include 3-D heterogeneity, anisotropy, CMT re-determination, etc. The real uncertainty can thus be several times the value of the nominal standard deviation.

2.3 SVD inversion

If we were to use eq. (2.24) as written, taking the sum over all eigenvectors, we would obtain the least squares solution of eq. (2.13) and eq. (2.23). However, in the actual inversions we sum over only the first n eigenvectors (which correspond to the n eigenvalues with the largest magnitudes — see eq. 2.15) as follows:

$$\delta\mathbf{m} = \sum_{j=1}^{n} e_j \mathbf{v}_j. \tag{2.30}$$

We discuss the choice of n below. For this case the covariance matrix in the model space is given by:

$$\text{covar. } (\delta\mathbf{m}) = \sigma_d^2 \sum_{j=1}^{n} \frac{\mathbf{v}_j \mathbf{v}_j^{\mathbf{T}}}{\lambda_j^2} \tag{2.31}$$

For each seismogram $\delta\mathbf{d}^{(kp)}$ we define the contribution of that seismogram to the perturbation of the j-th model parameter to be:

$$e_j^{(kp)} = \frac{\mathbf{v}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}(kp)} \delta\mathbf{d}^{(kp)}}{\lambda_j^2}, \tag{2.32}$$

where the final model is obtained by summing the contributions from each record:

$$e_j = \sum_{k,p} e_j^{(kp)}. \tag{2.33}$$

In the same way, we define the contribution from the k-th event as follows:

$$e_j^k = \sum_{p} e_j^{(kp)} \tag{2.34}$$

and the contribution from the records at the p-th station as follows:

$$e_j^p = \sum_k e_j^{(kp)}. \tag{2.35}$$

By analyzing the values of $e_j^{(kp)}$, e_j^k, and e_j^p we can see which data affect which model components, and which eigenvectors of the model space ($\mathbf{v}_1,\ \mathbf{v}_2,\ \cdots,\ \mathbf{v}_M$) are most sensitive to the data.

2.4 Criteria for evaluating models

We truncate the eigenvector expansion (eq. 2.30), using as a guide the AIC (Akaike Information Criterion: Akaike, 1977). The variance VAR_n and AIC_n for the model SVD_n (the model obtained by using n in eq. 2.30) are written respectively as follows:

$$\mathrm{VAR}_n = \frac{|\mathbf{d}_{\mathrm{OBS}} - \mathbf{d}_{\mathrm{SYN}}|^2}{|\mathbf{d}_{\mathrm{OBS}}|^2} \tag{2.36}$$

$$\mathrm{AIC}_n = \mathrm{ND}\ln 2\pi + \mathrm{ND}\ln(\mathrm{VAR}_n) + \mathrm{ND} + 2(n+1), \tag{2.37}$$

where ND is the number of independent data points. We discuss the evaluation of ND below. In the following we introduce the empirical parameter α, which we define as follows:

$$\mathrm{ND}^\alpha = \frac{1}{\alpha} KPT/20. \tag{2.38}$$

In this study the upper bound on the value of ND is $\mathrm{ND}^1 = KPT/20$ because we use 20-200 s bandpass-filtered waveforms, interpolated at a 1 Hz sampling frequency (the meaning of the superscript is explained below). It is reasonable to consider that the data are, to some extent, redundant, but at present we have no rigorous way to quantify the extent of the redundancy. The meaning of α is that the data are assumed to have α-fold redundancy. AIC_n, where ND^α is used in place of ND in eq. (2.37), is hereafter written as AIC_n^α.

Eq. (2.37) shows that AIC decreases as the variance decreases but increases as the number of model parameters, n, is increased. The minimum value of AIC thus formally represents the optimum balance between model complexity and variance reduction. In evaluating eqs. (2.36) and (2.37) in this study we use the variance reduction computed using the partial derivatives rather than re-computing synthetics for the models obtained by each inversion.

However, as discussed in Chapters 5 and 6, AIC analysis is not, strictly speaking, entirely appropriate for deciding where to truncate the SVD expansion. Although in practice reasonable results were obtained by SVD inversions for the cases considered in this thesis, this may be somewhat fortuitous. As discussed in Chapter 7, the use of a conjugate gradient (CG) expansion in conjunction with AIC is preferable from a strict theoretical point of view and also yields better results in practice (models highly similar to those from the SVD inversion but using a significantly smaller number of basis functions.)

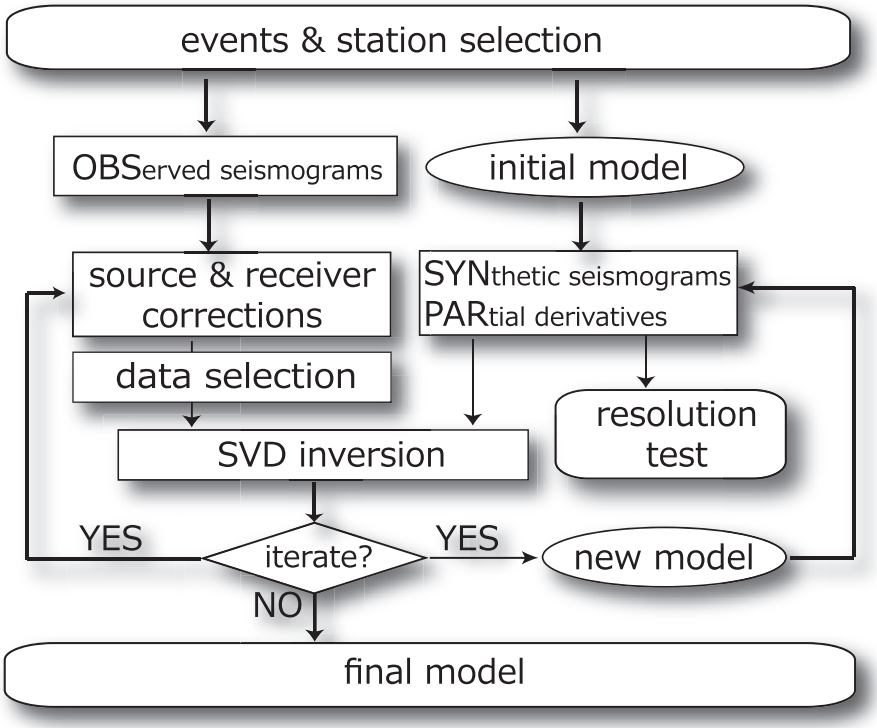


FIGURE 2.1: Our procedures for waveform inversion. We have developed software which automatically performs the above steps.

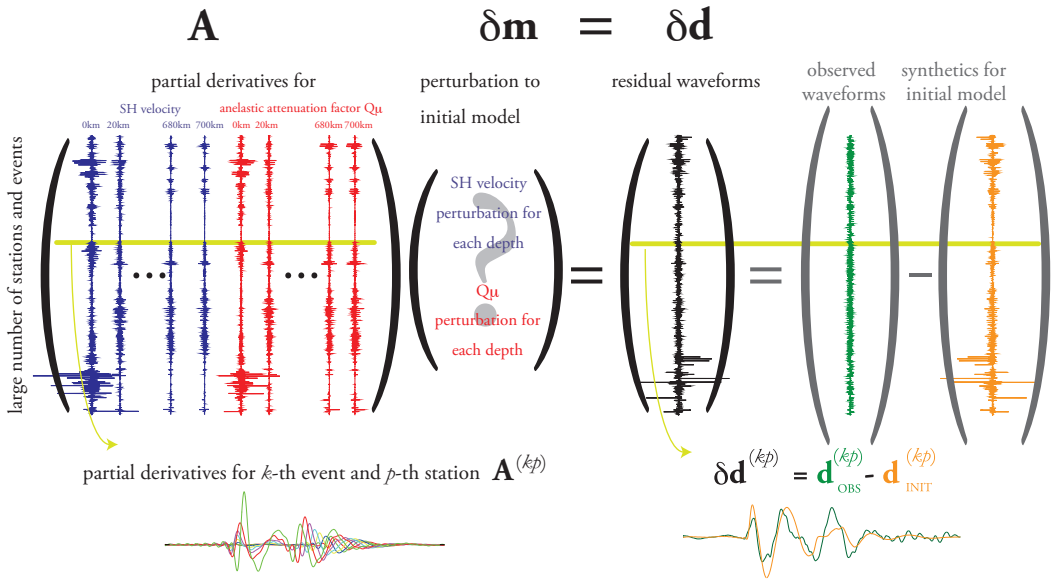


FIGURE 2.2: Schematic interpretation of the equation (eq. 2.12) being solved by waveform inversion. The residuals and partial derivatives for the various events are gathered in a single vector $\delta\mathbf{d}$ and a single matrix ($\mathbf{A}$) respectively. We then perform a simultaneous inversion for the perturbation to the model parameters ($\delta\mathbf{m}$) using the singular value decomposition (SVD).

Chapter 3

Dataset, static corrections, and "Initial Q_μ model"

We select the upper mantle and MTZ beneath the Northwestern Pacific as the target for this study. In this thesis we invert two datasets: (1) the entire dataset of waveforms which sample the region (Fig. 3.1, table 3.1): (2) the respective datasets for six subregions: the upper mantle and the transition zone beneath the Japan Sea; Philippines Sea West, Middle and East; Pacific West and East (Fig. 3.2).

As we show in this chapter, waveform data suggest that there may be lower Q_μ than the PREM model (Dziewonski and Anderson, 1981). By forward modeling and waveform inversion, we obtain more realistic 1-D Q_μ model for the whole region. We use a smoothed version of this model, which we call the "Initial Q_μ model" for the sub-regional inversion which we conduct in next chapter.

3.1 Event and station selection

We collect data for events with $M_w \geq 5.5$ with a half source duration time less than 10.0 s and a source depth greater than 100 km, using records from stations

at epicentral distances between 5° and 35° (based on Global CMT Solution parameters).

We use only transverse component (SH) data. The number of events from 1995 to 2007 is 161 and the number of waveforms is 6025. All of the seismograms are velocity seismograms and we deconvolve the instrumental response function. In this study we apply a Butterworth filter whose passband is between 20 s and 200 s to both the data and the synthetics; thus we can approximate the source time function (moment rate function) as a δ-function.

In this study we choose the direct S phase time window from 10 seconds before to 60 seconds after the first arrival (which contains overlapped triplication arrivals which are not individually identifiable because of the relatively long-period passband) as the portion of the data to be analyzed. We weight each waveform by the reciprocal of the maximum amplitude of the observed seismogram.

The partial derivatives for the moment tensor and centroid are given by Hara (1997). These results can be used to quantitatively evaluate the effect of uncertainty of the source parameters on the Earth structure model and also to invert for CMT parameters. However, as we have been careful not to use stations near nodes of the radiation pattern we do not think uncertainly of the CMT solution has a significant effect on the results presented in this thesis.

3.2 Static corrections

In order to correct for the effect of local structure near the stations and sources, we must make static corrections (Fig. 3.3). We align the onset times of the synthetic and the observed seismograms as follows. First we automatically pick the onset time of the synthetic seismogram and let the zero-to-peak time be Δt. We pick a time window from $4\Delta t$ before to Δt after the arrival time in the synthetic seismogram and also use this window for the observed seismogram. We then seek the time shift which gives the best cross-correlation coefficient in order to roughly

determine the observed arrival time, which we denote as $t_{\rm tmp}$. $t_{\rm tmp}$ represents the power centroid of the incoming S wave; to align the respective onset times we apply a causal filter. We finally search the observed waveform for a local zero near $t_{\rm tmp}$ by comparing the noise level of the time windows Δt before and after $t_{\rm tmp}$. We find a good correlation between the autopicked arrival times and manually picked arrival times. The procedures for determination of the time shifts are empirical, and there is probably room for further improvements. However, our present procedures yield reasonable results and our forward modeling tests suggest that our correction procedures are reasonable for the frequency band used in this analysis (20-200 s).

3.3 Data selection criteria

We consider pairs consisting of an observed record and the corresponding synthetic seismogram (Fig. 3.4). The amplitude ratios in Fig. 3.4 are calculated by dividing the maximum absolute value of the synthetics in the time window by the maximum absolute value of the observed seismograms. There are cases for which the observed and synthetic amplitudes differ greatly, and others for which the phases differ greatly even though the amplitudes are similar. For other waveforms the amplitudes differ greatly while the phases match quite well. For such waveforms the predominant period of the observed record is much shorter than that of the synthetic, perhaps because such seismograms contain waves that traveled through the slab, where there is a narrow high Q_μ zone. The above observations probably contain important information on 3-D structure, but we exclude them from the dataset for this thesis.

3.4 "Initial Q_μ" model

The amplitude ratios in Fig. 3.4a suggest that the actual Q_μ values may be systematically lower than the PREM Q_μ model. We first investigate the behavior of

our dataset by forward modeling with several Q_μ models (shown in Fig. 3.5). Figs. 3.4b and 3.4c illustrate the behavior for PREM1/2, which has half the Q_μ value of PREM, and Q19, which is modified from Romanowicz (1995) for this region. The first number above each panel of Fig. 3.4 shows the number of records which satisfy a criterion of an amplitude ratio between 0.3 and 3.0 and also have a cross correlation above 50%, which we regard as a good match. The second number is the total number of records in the dataset. "PREM1/2" improves the fit of the amplitude ratios but because of the broadening effect, the cross correlations drop significantly, while Q19 moderately changes the amplitude ratio but does not significantly degrade the cross-correlations. The above trade-offs suggest that we should simultaneously invert for elasticity and anelasticity.

We first invert for an initial Q_μ structure for our study region using the methods discussed in the previous section. The unknowns are μ_0 and Q_μ from the Earth's surface to a depth of 740 km at intervals of 10 km. We use 754 waveforms chosen randomly from those parts of our dataset which have amplitude ratios between 0.1 and 10.0 and cross correlation coefficients above 50%. The results obtained from SVD inversions (Fig. 3.5) using the n largest eigenvalues (up to $n = 20$) show a lower Q_μ model than PREM. We use a smoothed version of this model (Fig. 3.5c), which we call the "Initial Q_μ" model, as the Q_μ model in our subsequent inversions. Fig. 3.4 shows that when "Initial Q_μ" is used the number of waveforms satisfying both the amplitude ratio and cross-correlation criteria is 3629, as opposed to 3355, 3402, and 3077 for PREM, Q19, and PREM1/2 respectively.

Since the dataset used in this inversion is noisy and samples a complicated structure, a single S-wave velocity model cannot be meaningfully determined for the whole region. This is confirmed by the fluctuations of the S-wave velocity model in Fig. 3.5b. Fig. 3.5a shows AIC_n^1 and VAR_n for $n \leq 20$. AIC suggests that using a basis of 16-18 eigenvectors yields the most appropriate model.

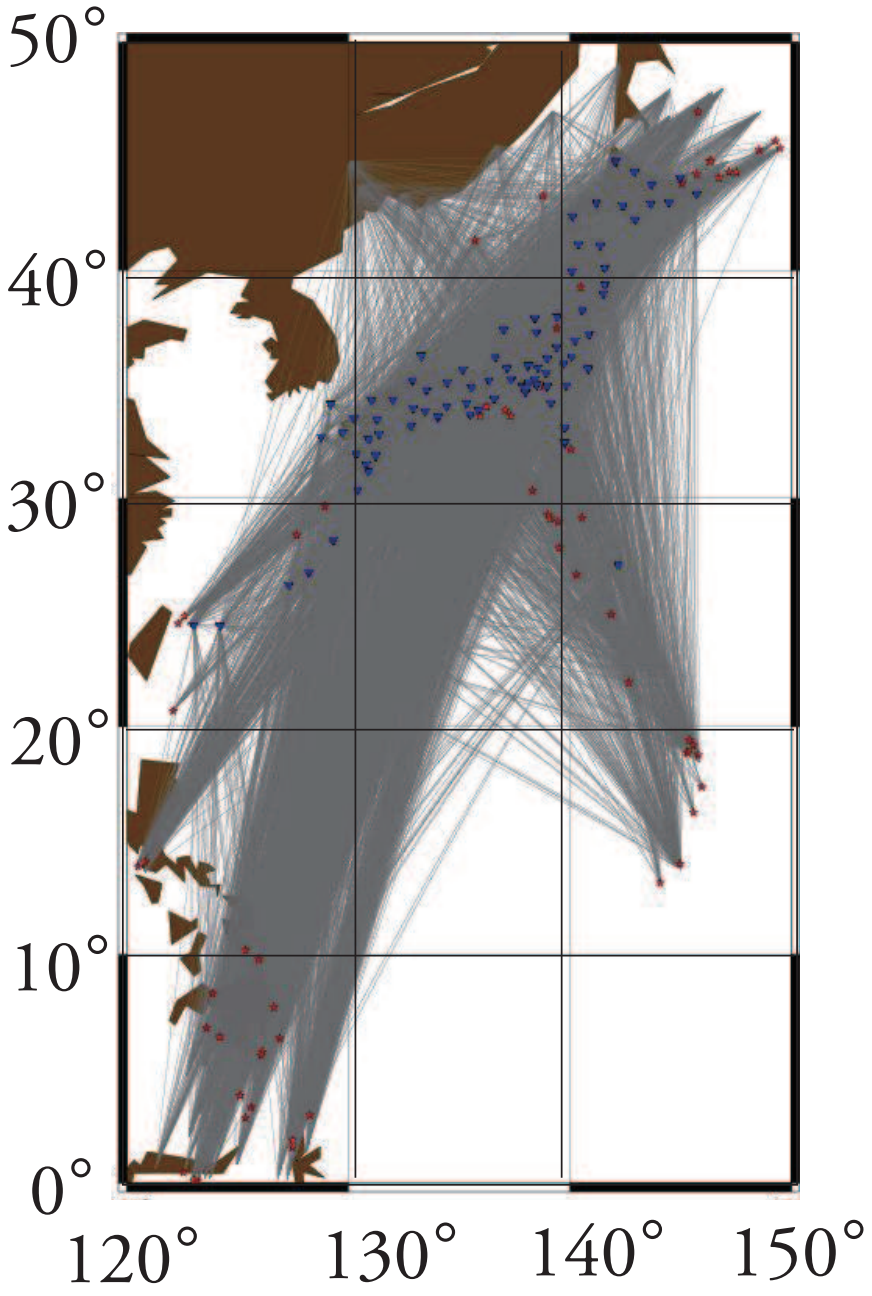


FIGURE 3.1: Dataset for this study. Blue dots denote F-net stations and red dots denote events with an epicentral distance between 5° and 40°. We use the entire dataset in the inversion for the "Initial 1-D Q_μ" model. We then divide the dataset into several subsets in order to invert for more localized structure.

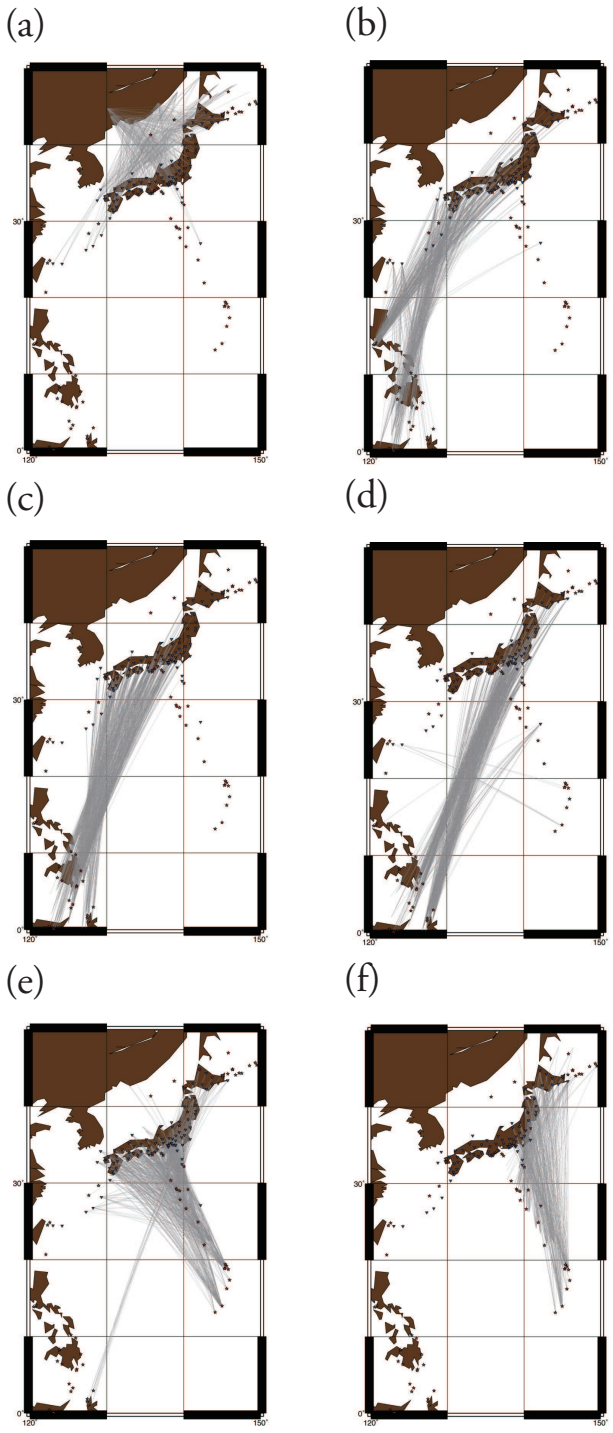


FIGURE 3.2: The six subregions for this study, and the number of waveforms used in the inversion for each subregion. (a) Japan Sea: 478 waveforms; (b) Philippines Sea West: 644; (c) Philippines Sea Middle: 437; (d) Philippines Sea East: 688; (e) Pacific West: 1176; (f) Pacific East: 525.

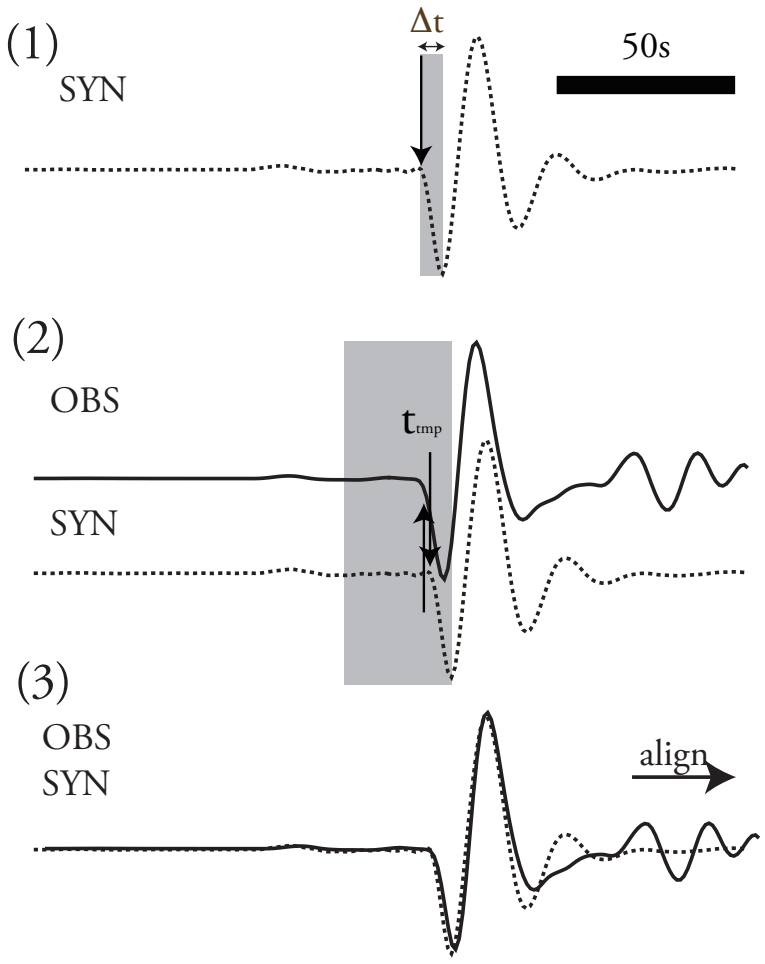


FIGURE 3.3: Procedure for static corrections. Step 1: Autopick the first arrival on the synthetics (the up arrow) and let the zero-to-peak time be Δt. Step 2: Use the time window from $4\Delta t$ before to Δt after the arrival) to seek the time shift $t_{\rm tmp}$ (the down arrow) which gives the best correlation coefficients. Step 3: Find the local zero around $t_{\rm tmp}$ in the observed seismogram and treat this as "the arrival time" with which we align the synthetic.

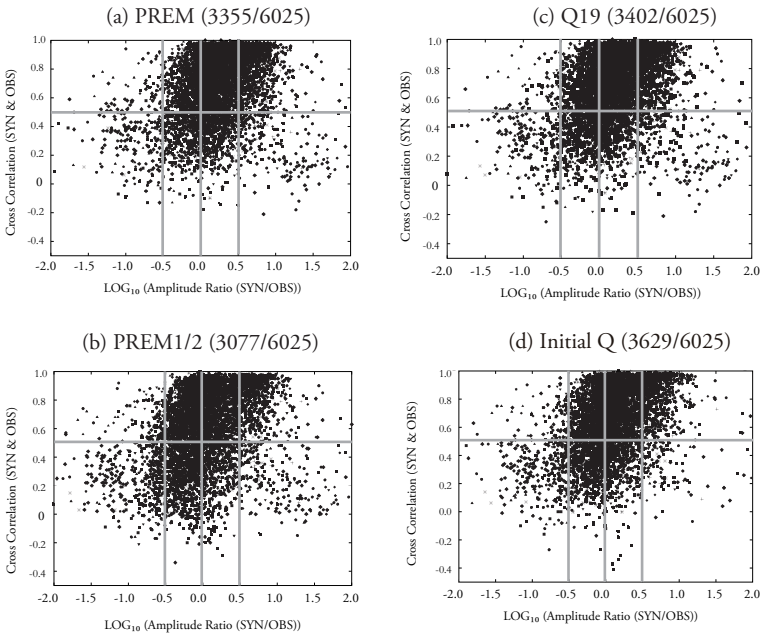


FIGURE 3.4: Plots of the amplitude ratios and cross correlations (the maximum values with a time lag in the range -10.0 s to 10.0 s) for pairs of observed and synthetic seismograms determined from 4 sets of synthetics, each calculated with a different Q_μ model. The different symbols (triangles, squares, etc.) denote individual events used in this study. The grey lines denote where the amplitude ratio is 3.0 ($\log_{10} 3.0 = 0.48$), 1.0, ($\log_{10} 1.0 = 0.0$), 0.33 ($\log_{10} 0.33 = -0.48$) and where the cross correlation coefficient is 50 %. The number in the parentheses above each panel show the number of waveforms which satisfy a criterion of 0.33 to 3.0 maximum amplitude ratio between the observed seismogram and the corresponding synthetic seismogram calculated for the the model, and also have a cross-correlation coefficient above 50 %. The second number is the total number of records in the dataset. As the first numbers in the parentheses show, the "Initial Q_μ" model is more suitable than the other 3 models in terms of the amplitude ratios and cross correlation coefficients.

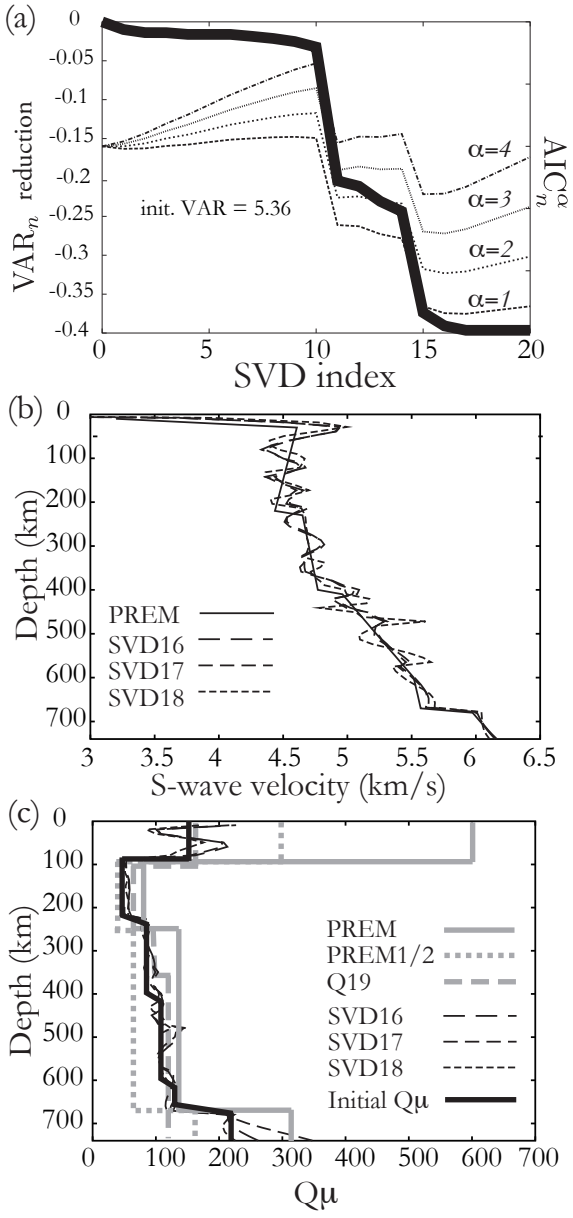


FIGURE 3.5: Inversion results for the dataset of all records with amplitude ratios (OBS/INIT) between 0.1 and 10.0 and cross correlation coefficients above 50 %. (a) Variance (thick curve) and AIC_n^1 value (thin curve). (b) S-velocity for PREM (the starting elastic model for the inversion) and the inversion results for $n = 16$, $n = 17$ and $n = 18$. (c) Q_μ model for PREM (the starting anelastic model for the inversion) as well as "PREM1/2" and "Q19" for reference, and the inversion results obtained using $n = 16$, $n = 17$ and $n = 18$. The smoothed model "Initial Q_μ" obtained from the inversion is used as the starting anelastic model for subsequent inversions.

TABLE 3.1: Events used in this study.

Event Name	YY/MM/DD	Centroid Time	Depth (km)	Latitude, N	Longitude, E	Mw	half duration (s)
19950218	1995/02/18	13:29:09.2	353.0	46.61°	145.77°	5.5	1.4
19950310	1995/03/10	05:22:25.9	354.8	46.10°	143.31°	5.6	1.5
19950330	1995/03/30	22:15:54.9	329.8	44.79°	137.32°	5.8	2.0
19950331	1995/03/31	14:01:44.1	366.5	38.16°	135.11°	6.1	2.8
19950404	1995/04/04	17:21:08.7	108.9	6.13°	126.97°	5.7	1.9
19950408	1995/04/08	17:45:19.5	281.3	21.93°	142.68°	6.2	3.2
19950420	1995/04/20	08:45:17.2	102.5	6.33°	126.89°	6.4	4.4
19950423	1995/04/23	06:38:20.4	534.6	6.39°	124.17°	6.1	2.8
19950518	1995/05/18	14:31:15.7	100.6	44.29°	147.55°	5.5	1.4
19950707	1995/07/07	21:15:23.9	348.4	33.89°	137.11°	5.9	2.2
19950729	1995/07/29	16:18:48.1	445.2	30.35°	138.32°	5.5	1.4
19950823	1995/08/23	07:06:09.4	599.2	18.88°	145.30°	7.0	8.1
19950824	1995/08/24	01:55:39.3	594.0	18.93°	145.19°	6.2	3.1
19950824	1995/08/24	06:28:58.5	617.0	18.86°	145.24°	5.8	2.0
19950824	1995/08/24	07:54:40.0	611.8	18.79°	145.33°	6.1	2.8
19950824	1995/08/24	07:55:37.3	609.8	18.85°	145.35°	6.2	3.1
19950901	1995/09/01	06:30:39.2	137.9	0.12°	123.19°	5.9	2.2
19951001	1995/10/01	17:06:08.7	448.0	29.30°	139.02°	6.1	2.7
19951004	1995/10/04	01:31:10.4	127.6	1.87°	127.47°	5.6	1.6
19951020	1995/10/20	19:21:37.9	207.2	19.19°	145.58°	6.0	2.7
19951130	1995/11/30	15:09:24.9	146.9	44.23°	145.74°	5.9	2.4
19951205	1995/12/05	14:54:50.6	134.1	1.53°	127.45°	5.8	2.0
19960201	1996/02/01	07:18:07.8	170.2	44.79°	146.33°	6.2	3.0
19960214	1996/02/14	21:26:57.1	144.9	29.16°	140.55°	5.6	1.7
19960218	1996/02/18	09:57:21.6	243.1	14.07°	120.80°	5.6	1.6
19960222	1996/02/22	14:59:12.3	130.3	45.29°	148.56°	6.3	3.7

19960304	1996/03/04	15:59:10.6	152.7	2.84°	125.34°	5.9	2.2
19960316	1996/03/16	22:04:11.9	477.9	29.12°	139.12°	6.6	5.0
19960422	1996/04/22	11:27:59.0	205.7	29.67°	128.93°	5.6	1.6
19960609	1996/06/09	01:12:22.6	159.6	17.36°	145.98°	6.5	4.6
19960626	1996/06/26	03:22:09.0	478.6	27.82°	139.85°	6.2	3.1
19960626	1996/06/26	03:55:13.5	489.9	27.83°	139.53°	5.5	1.4
19960706	1996/07/06	21:36:34.8	252.5	22.02°	142.99°	6.2	3.2
19960715	1996/07/15	16:51:28.4	189.0	18.76°	145.82°	6.2	3.3
19960720	1996/07/20	09:14:07.4	119.4	13.87°	120.48°	5.9	2.2
19960801	1996/08/01	04:08:26.0	133.3	0.07°	123.17°	5.5	1.4
19960914	1996/09/14	02:53:28.7	177.6	0.11°	123.01°	5.8	2.0
19960928	1996/09/28	14:10:46.6	237.6	10.19°	125.33°	6.3	3.5
19961017	1996/10/17	16:15:28.5	113.6	5.74°	126.09°	5.9	2.3
19961018	1996/10/18	16:44:50.3	339.9	33.63°	137.33°	5.5	1.5
19961203	1996/12/03	15:49:15.9	143.6	37.48°	139.43°	5.6	1.6
19961222	1996/12/22	14:53:32.5	244.7	43.29°	138.78°	6.5	4.3
19970326	1997/03/26	18:18:39.4	131.4	2.94°	128.25°	5.8	2.0
19970527	1997/05/27	15:09:07.8	537.0	16.24°	145.61°	5.6	1.5
19970715	1997/07/15	11:05:33.7	108.7	24.52°	122.28°	5.5	1.4
19971005	1997/10/05	19:00:09.1	204.7	5.58°	126.05°	5.8	1.9
19971006	1997/10/06	12:30:11.4	114.4	9.79°	125.95°	6.4	4.2
19971011	1997/10/11	18:24:26.2	147.1	24.86°	122.57°	5.5	1.4
19971022	1997/10/22	09:55:52.5	161.1	44.83°	146.40°	5.5	1.5
19971026	1997/10/26	19:06:46.4	126.3	39.27°	140.50°	5.5	1.5
19971115	1997/11/15	07:05:18.8	157.0	43.82°	145.07°	6.0	2.7
19971128	1997/11/28	06:10:50.8	401.0	46.96°	145.81°	5.5	1.4
19971202	1997/12/02	11:47:48.8	591.9	6.79°	123.58°	5.5	1.2
19980207	1998/02/07	01:13:42.5	540.3	24.47°	141.72°	5.9	2.1

19980207	1998/02/07	01:19:06.2	533.6	24.92°	141.87°	6.4	4.1
19980228	1998/02/28	17:38:52.1	293.9	33.40°	138.18°	5.7	1.8
19980308	1998/03/08	00:35:46.7	158.9	20.70°	122.06°	5.7	1.8
19980515	1998/05/15	05:58:12.0	167.6	13.99°	144.99°	6.0	2.6
19980523	1998/05/23	17:44:50.0	628.8	8.29°	123.86°	6.0	2.5
19980601	1998/06/01	07:35:53.3	405.0	34.04°	136.24°	5.8	1.9
19980707	1998/07/07	14:32:02.9	100.9	32.16°	140.05°	5.5	1.0
19980818	1998/08/18	18:00:16.1	125.2	45.68°	149.31°	5.5	1.4
19980820	1998/08/20	06:41:10.5	425.5	28.99°	139.47°	7.0	4.7
19980908	1998/09/08	09:10:07.4	142.6	13.15°	144.10°	5.8	2.0
19980921	1998/09/21	06:52:45.1	144.1	0.46°	122.53°	5.9	2.2
19981003	1998/10/03	11:15:46.1	219.0	28.40°	127.66°	6.1	2.9
19981018	1998/10/18	08:33:58.0	151.9	19.43°	145.41°	5.5	1.4
19990109	1999/01/09	03:05:40.0	124.7	44.33°	147.21°	5.6	1.5
19990112	1999/01/12	02:32:31.1	459.6	26.67°	140.31°	5.9	2.3
19990205	1999/02/05	14:37:56.5	402.4	47.47°	147.73°	5.8	2.1
19990408	1999/04/08	13:10:40.2	575.4	43.66°	130.47°	7.1	8.3
19990602	1999/06/02	00:23:21.3	135.3	0.18°	123.72°	5.7	1.7
19990703	1999/07/03	05:30:14.1	433.7	26.29°	140.55°	6.0	2.6
19990726	1999/07/26	02:45:52.2	122.6	41.14°	142.02°	5.7	1.2
19991205	1999/12/05	22:00:34.8	427.0	29.89°	138.75°	5.7	1.7
19991231	1999/12/31	13:09:16.0	383.8	37.32°	134.61°	5.5	1.4
20000110	2000/01/10	16:40:46.1	454.4	27.33°	140.29°	5.7	1.7
20000128	2000/01/28	16:39:24.5	200.8	25.92°	124.56°	6.0	2.5
20000213	2000/02/13	02:57:12.7	524.3	42.77°	131.64°	6.0	2.4
20000213	2000/02/13	18:19:42.2	133.3	5.73°	126.88°	5.6	1.7
20000215	2000/02/15	02:05:05.1	525.5	17.67°	145.54°	5.8	2.2
20000226	2000/02/26	08:11:53.2	125.9	13.79°	144.98°	6.1	3.1

20000403	2000/04/03	15:20:03.9	139.6	4.10°	125.75°	6.1	3.0	
20000421	2000/04/21	11:06:19.6	357.7	35.73°	135.33°	5.5	1.4	
20000609	2000/06/09	22:35:17.0	484.7	30.41°	137.96°	5.7	1.8	
20000609	2000/06/09	23:31:50.4	492.1	30.47°	137.79°	6.2	3.2	
20000614	2000/06/14	17:00:54.2	111.4	4.75°	127.61°	6.3	3.6	
20000619	2000/06/19	09:34:59.5	106.7	13.88°	120.38°	5.6	1.6	
20000621	2000/06/21	16:25:11.7	121.9	13.92°	144.98°	5.9	2.4	
20000710	2000/07/10	09:58:22.9	361.3	46.83°	145.59°	5.8	1.9	
20000803	2000/08/03	08:53:24.7	104.2	0.65°	121.27°	5.5	1.4	
20000806	2000/08/06	07:27:23.6	411.3	28.89°	139.68°	7.3	10.3	
20000820	2000/08/20	07:55:56.2	174.4	7.96°	126.05°	5.6	1.6	
20000827	2000/08/27	17:19:09.0	105.7	22.27°	143.84°	5.6	1.6	
20000831	2000/08/31	08:24:12.7	578.7	5.05°	123.49°	5.6	1.7	
20000926	2000/09/26	16:49:38.2	147.9	1.18°	127.61°	5.9	2.5	
20001014	2000/10/14	20:02:52.4	123.6	23.41°	142.14°	5.5	1.6	
20001021	2000/10/21	22:30:36.5	152.5	13.74°	120.81°	5.8	2.4	
20001027	2000/10/27	04:21:56.4	387.6	26.31°	140.69°	6.0	2.8	
20001104	2000/11/04	12:38:42.2	116.4	5.83°	126.05°	5.7	2.0	
20001129	2000/11/29	01:48:46.9	109.4	13.67°	120.76°	5.5	1.3	
20001210	2000/12/10	18:58:37.8	120.9	0.16°	127.22°	5.6	1.6	
20001222	2000/12/22	10:13:03.4	140.8	44.76°	147.14°	6.1	3.0	
20010114	2001/01/14	08:58:29.2	112.5	22.19°	143.93°	5.9	2.2	
20010226	2001/02/26	05:58:28.8	396.4	46.79°	144.54°	6.0	2.7	
20010314	2001/03/14	18:56:23.8	121.8	0.65°	122.01°	5.9	2.3	
20010703	2001/07/03	13:10:49.3	307.7	21.66°	142.99°	6.4	4.0	
20010725	2001/07/25	21:56:41.9	165.5	0.16°	123.23°	5.6	1.6	
20011003	2001/10/03	17:25:16.7	288.1	46.95°	148.62°	5.9	2.3	
20011202	2001/12/02	13:01:57.8	123.7	39.49°	141.09°	6.4	4.3	

20020101	2002/01/01	11:29:25.5	130.6	6.35°	125.87°	6.3	3.4
20020107	2002/01/07	13:26:31.4	609.2	19.03°	145.08°	5.9	2.1
20020201	2002/02/01	21:55:25.3	369.6	45.50°	136.66°	5.8	2.0
20020212	2002/02/12	15:39:58.5	126.4	13.77°	145.02°	5.8	1.9
20020307	2002/03/07	00:07:12.4	453.8	47.97°	146.93°	5.6	1.4
20020323	2002/03/23	05:15:54.7	124.5	1.36°	128.09°	5.7	1.8
20020509	2002/05/09	23:41:37.9	200.7	2.85°	128.38°	5.5	1.4
20020521	2002/05/21	20:04:17.0	146.0	44.56°	146.71°	5.5	1.3
20020526	2002/05/26	00:10:25.3	114.1	1.99°	127.25°	5.9	2.3
20020603	2002/06/03	09:15:04.8	491.2	27.50°	139.86°	5.8	2.1
20020628	2002/06/28	17:19:40.1	581.5	43.74°	130.45°	7.3	10.9
20020802	2002/08/02	23:11:44.8	441.5	29.35°	139.25°	6.2	3.2
20020915	2002/09/15	08:39:36.3	589.4	44.77°	130.04°	6.4	3.7
20021105	2002/11/05	08:47:28.7	604.0	49.06°	142.28°	5.6	1.6
20021221	2002/12/21	17:41:17.5	574.3	5.13°	123.38°	5.7	1.6
20030109	2003/01/09	03:40:32.4	120.6	0.22°	123.50°	5.5	1.3
20030117	2003/01/17	23:46:10.6	459.9	2.01°	122.77°	5.8	2.0
20030219	2003/02/19	05:01:42.2	226.6	44.19°	141.73°	5.9	2.2
20030310	2003/03/10	02:09:41.8	100.2	1.74°	127.23°	6.3	3.6
20030505	2003/05/05	15:50:14.3	121.9	0.44°	127.38°	6.3	3.7
20030526	2003/05/26	23:13:35.7	579.7	6.90°	123.85°	6.8	6.3
20030701	2003/07/01	05:52:26.7	604.6	4.69°	122.67°	6.0	2.3
20030727	2003/07/27	06:25:36.1	477.2	46.99°	139.23°	6.7	5.5
20030831	2003/08/31	23:08:04.3	493.0	43.38°	132.37°	6.1	2.7
20031109	2003/11/09	19:23:31.6	137.8	1.82°	127.52°	5.8	1.9
20031111	2003/11/11	18:48:26.5	107.1	22.39°	143.48°	5.9	2.2
20031112	2003/11/12	08:26:47.2	381.8	33.31°	137.09°	6.3	3.6
20040129	2004/01/29	09:48:43.4	209.6	6.35°	126.98°	5.5	1.3

20040222	2004/02/22	23:17:18.7	194.9	18.48°	145.85°	5.6	1.6
20040630	2004/06/30	23:37:28.1	101.2	0.79°	124.95°	6.2	3.1
20040915	2004/09/15	19:10:51.2	106.4	14.39°	120.30°	5.9	2.3
20040915	2004/09/15	19:45:46.5	101.3	5.88°	125.28°	5.5	1.4
20041008	2004/10/08	14:36:09.9	110.9	13.91°	120.50°	6.4	4.0
20041015	2004/10/15	04:08:54.9	102.1	24.48°	122.74°	6.6	4.7
20041107	2004/11/07	02:02:29.6	493.0	47.93°	144.52°	6.1	2.8
20041219	2004/12/19	15:14:55.5	588.8	7.44°	123.79°	5.7	1.7
20041226	2004/12/26	20:50:37.1	124.0	6.20°	126.85°	5.5	1.4
20050121	2005/01/21	17:54:36.8	178.7	5.14°	125.49°	5.8	2.0
20050202	2005/02/02	02:30:29.8	159.4	13.86°	144.81°	6.3	3.3
20050205	2005/02/05	03:34:32.0	132.0	15.88°	146.19°	6.6	4.8
20050205	2005/02/05	12:23:23.6	530.6	5.47°	123.67°	7.1	8.5
20050222	2005/02/22	11:20:26.2	372.2	33.15°	137.16°	5.6	1.6
20050403	2005/04/03	01:41:32.1	118.6	13.62°	120.79°	5.6	1.6
20050419	2005/04/19	01:46:57.9	417.9	29.66°	139.03°	5.9	2.1
20050701	2005/07/01	09:36:17.6	199.5	31.90°	139.30°	5.5	1.4
20050904	2005/09/04	23:58:37.2	454.4	3.05°	123.32°	6.0	2.4
20050921	2005/09/21	02:25:10.3	107.2	43.90°	146.41°	6.0	2.5
20051015	2005/10/15	15:51:11.2	194.6	25.29°	123.43°	6.4	4.0
20051023	2005/10/23	10:08:16.4	393.5	37.36°	134.61°	5.9	2.3
20051121	2005/11/21	15:36:33.5	155.0	30.97°	130.31°	6.2	2.9
20051211	2005/12/11	16:23:34.5	199.2	14.25°	120.57°	5.6	1.6
20060217	2006/02/17	02:13:33.5	174.2	24.10°	141.58°	5.6	1.6
20060328	2006/03/28	13:32:49.7	411.6	31.72°	137.79°	5.9	2.2
20060416	2006/04/16	11:48:57.6	423.6	30.29°	138.60°	5.7	1.8
20060416	2006/04/16	17:46:18.0	186.7	4.54°	125.70°	5.7	1.7
20060611	2006/06/11	20:01:31.1	144.5	33.15°	131.34°	6.4	3.7

20060622	2006/06/22	10:53:14.4	104.8	45.37°	149.52°	5.9	2.3
20060821	2006/08/21	22:20:51.8	419.1	33.65°	135.94°	5.5	1.3
20060905	2006/09/05	04:53:03.2	120.4	7.68°	126.62°	5.8	1.9
20060916	2006/09/16	02:22:53.2	382.2	41.33°	135.71°	5.9	2.1
20061122	2006/11/22	11:15:13.0	105.5	44.08°	146.75°	5.6	1.6
20061212	2006/12/12	15:48:06.5	215.3	3.83°	125.10°	6.3	3.5
20061222	2006/12/22	11:06:08.7	183.7	3.30°	125.59°	5.5	1.4
20070115	2007/01/15	18:18:01.9	169.8	34.94°	138.81°	5.9	2.1
20070308	2007/03/08	05:03:34.5	135.2	29.89°	140.52°	6.1	2.7
20070309	2007/03/09	03:22:44.1	451.1	43.22°	133.65°	6.0	2.5
20070330	2007/03/30	09:05:07.6	102.9	44.11°	146.19°	5.5	1.4
20070401	2007/04/01	02:51:06.2	378.4	32.31°	137.67°	5.7	1.8
20070418	2007/04/18	15:07:33.0	125.6	42.67°	141.96°	5.5	1.4
20070509	2007/05/09	09:01:48.4	569.2	6.62°	123.91°	5.6	1.6
20070529	2007/05/29	20:12:41.5	172.8	0.01°	122.89°	5.5	1.3
20070701	2007/07/01	04:12:09.4	130.2	43.64°	144.81°	5.8	1.9
20070716	2007/07/16	14:17:42.0	374.9	36.84°	135.03°	6.8	6.2
20070822	2007/08/22	07:26:25.8	126.9	42.07°	140.80°	5.6	1.5

Chapter 4

Synthetic tests

We conduct several tests using synthetic data to study the resolution of our inversions and the potential that the static corrections could cause systematic errors. As shown in Fig. 4.1 we use an input model with alternating slow and fast layers of 100 km thickness, with an amplitude of ± 1.5 % relative to the initial model (PREM + "Initial Q_μ"). We compute synthetics for a small number of stations and events (only 50 records), and conduct inversions varying n. The unknowns in this case are the perturbations to the elastic constant μ_0 from the Earth's surface to a depth of 700 km at intervals of 20 km. Perturbations to the anelastic model parameters are not included as unknowns in this case. As shown in Fig. 4.1, the inversions for reasonable ranges of n (e.g., 17-23) are successful in reproducing the main features of the input model.

We next conduct synthetic tests to study the effect that hypothetical errors in the static corrections might have on the inversion results. Figs. 4.2a through 4.2c show synthetic results for systematically shifted synthetic data with shifts of 2, 3, and 4 s later than the actual travel times, respectively. The models in Fig. 4.2a recover the input model reasonably well, but those in Figs. 4.2b and 4.2c do not, due to the non-linearity of the inversion. Fig. 4.2d shows that when the data are randomly shifted by up to ±10 s the synthetic inversion can still reasonably recover the input model.

Fig. 4.3 examines synthetic inversion results for the case of synthetic seismograms systematically shifted to be 1 s late to simulate hypothetical errors in the static corrections. The left panel shows that the basic shape of the input model is recovered, but with a shift of about 0.2% towards lower velocities. The right panel shows $\theta(z)$, the angle between the partial derivatives with respect to the time shift and those with respect to μ_0 for each depth z, defined as follows:

$$\cos\theta(z) = \frac{\sum \frac{\partial u}{\partial t}\frac{\partial u}{\partial \mu_0(z)}}{\sqrt{\sum \left(\frac{\partial u}{\partial t}\right)^2}\sqrt{\sum \left(\frac{\partial u}{\partial \mu_0(z)}\right)^2}}. \tag{4.1}$$

On the basis of the results in Figs. 4.2 and 4.3 we conclude that hypothetical small systematic errors or reasonably large but random errors in the static corrections would cause only small errors in the absolute velocities while still preserving the basic character of the model, but that systematic errors of several seconds in most or all of the static corrections could be problematical. Such large systematic errors do not appear to be present in our actual data analyses. Nevertheless the subject of static corrections remains a subject for further investigation. Note that this applies not only to our work in particular, but to all studies of Earth structure.

4.1 "Checkerboard" test for Q_μ perturbation

In order to study the resolving power of a simultaneous inversion for elastic and anelastic parameters we used two input models for a synthetic inversion. The first has alternating slow and fast layers of 100 km thickness, with a perturbation of ± 2.0 % relative to the initial velocity model (PREM) but with no perturbation to the Q_μ model (Fig. 4.4). The second has the same elastic perturbation but also has a perturbation to the Q_μ model (Fig. 4.5). In our test inversions, the unknowns are μ_0 and Q_μ from the Earth's surface to a depth of 700 km at intervals of 20 km; we calculate 6000 synthetic "records" for each of the two perturbed models.

In Figs. 4.4a and 4.5a, we plot the variance (VAR) and the values of AIC_n^α (for $\alpha = 1, 10, 50, 100$ respectively) as a function of n and find AIC minima around SVD22 and SVD24 respectively if we take $\alpha \geq 50$. Fig. 4.4a shows that there is a drop in AIC at $n = 22$ regardless of the value of α. However this is only a local minimum for $\alpha = 1$ or $\alpha = 10$, but is a global minimum for $\alpha = 50$ or $\alpha = 100$. Since the absolute value of AIC_n^α is not important (only the relative values for a given α as n varies are significant) we have adjusted all the AIC curves in Fig. 4.4a (as well as similar later figures) to have the same AIC values for $n = 0$ and normalized by multiplying by α.

Figs. 4.4b and 4.5b show that even for a small number of eigenvectors (15 or 13, respectively), the main S-wave velocity features of the input model can be resolved, except for the layer from 600 to 700 km depth, for which the resolving power of this basis set is low. From Fig. 4.4c we see that the Q_μ structure is not changed significantly for the input model with no Q_μ perturbation even if we use a larger basis ($n = 22$).

Fig. 4.4 shows that including the partial derivatives for anelasticity in the inversion for a purely elastic model perturbation does not result in a significant perturbation to Q_μ. On the other hand, Fig. 4.5 shows that a larger number of eigenvectors ($n = 24$) is necessary to reproduce the anelastic structure, whereas we cannot reproduce it with a smaller number of eigenvectors ($n = 13$).

4.2 Sensitivity of eigenvectors to elastic and anelastic structure

Fig. 4.6a shows the magnitudes of the eigenvalues, while Fig. 4.6b shows the relative contribution to the eigenvectors of the elastic and anelastic parts of the model, defined respectively as P_j^{elast}, and P_j^{anel} as follows:

$$P_j^{\mathrm{elast}} = \frac{|\mathbf{v}_j^{\mathrm{elast}}|^2}{|\mathbf{v}_j|^2} \tag{4.2}$$

$$P_j^{\mathrm{anel}} = \frac{|\mathbf{v}_j^{\mathrm{anel}}|^2}{|\mathbf{v}_j|^2}, \tag{4.3}$$

where $\mathbf{v}_j^{\mathrm{elast}}$ and $\mathbf{v}_j^{\mathrm{anel}}$ are the elastic and anelastic components of each eigenvector $\mathbf{v}_j$. In Fig. 4.6b, we normalize the value of P_j^{elast} and P_j^{anel} so that:

$$\sum_{j=1}^{M} P_j^{\mathrm{elast}} = \sum_{j=1}^{M} P_j^{\mathrm{anel}}. \tag{4.4}$$

Fig. 4.6b shows that the eigenvectors for the modes with the largest eigenvalues reflect primarily elastic structure, while the eigenvectors for modes with smaller eigenvalues tend to be dominated by anelastic structure. This suggests that we cannot determine the Q_μ structure reliably and stably unless we first obtain a reliable elastic model, which is consistent with the results of previous surface wave attenuation tomographic studies (e.g., Romanowicz 1995, 1998; Gung and Romanowicz, 2004; Dalton et al., 2008).

As shown in the next section, the results obtained for the real dataset require a large number of eigenvectors to describe the Q_μ structure, and there are some fluctuations in the depth-dependence of the S-wave velocity model. This occurs because the actual anelastic structure is far from the starting model while the starting elastic model is much closer to the final model.

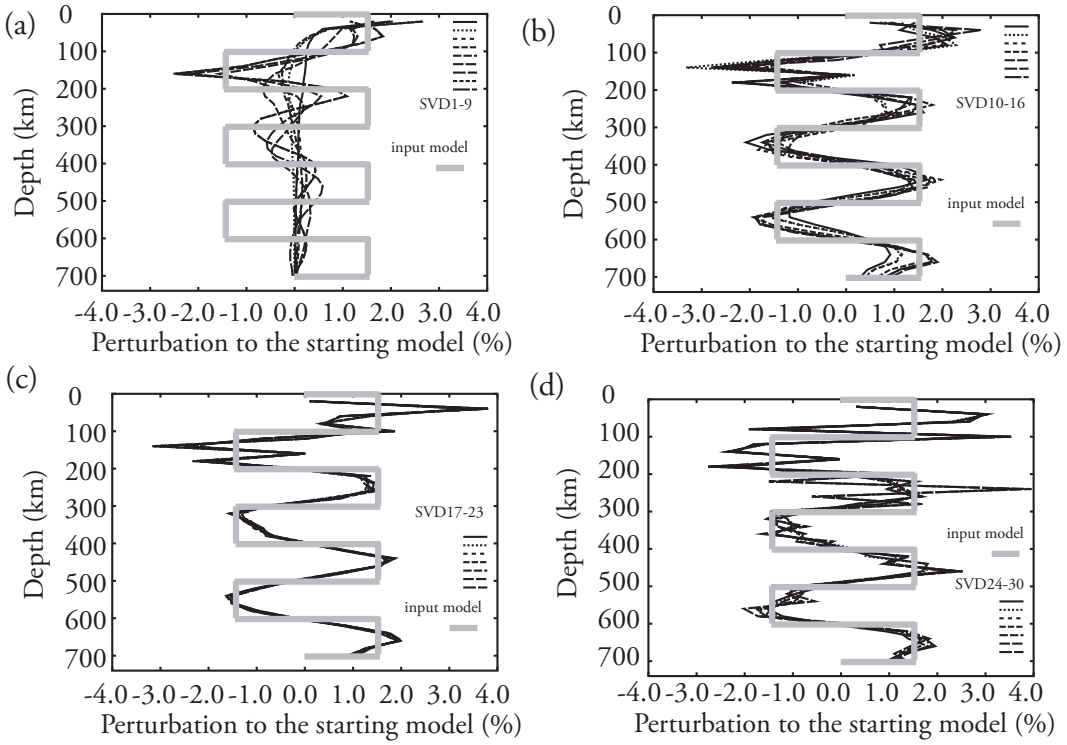


FIGURE 4.1: Synthetic inversion results for a dataset perturbed only with respect to the velocity structure (with no perturbation to Q_μ for this figure). Gray line denotes the input model and black lines are inversion results. Panel (a) shows models for $n = 1$ through 9, panel (b) for $n = 10$ through 16, panel (c) for $n = 17$ through 23, and panel (d) for $n = 24$ through 30. Panel (b) appears best, and panel (c) appears reasonable, while panel (a) cannot satisfactorily reproduce the input model and panel (d) fluctuates too much.

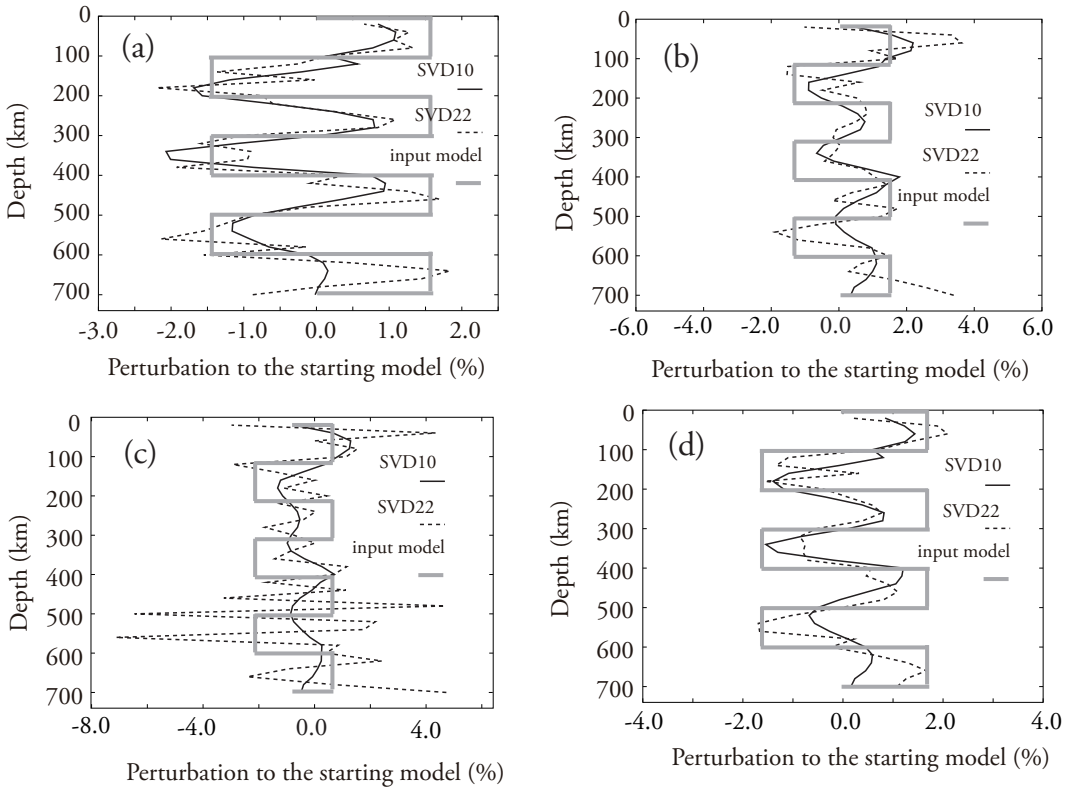


FIGURE 4.2: Synthetic inversion results for several perturbed datasets. Details are as same as Fig. 4.1 (only SVD10 and SVD22 are shown) except: (a) the input synthetic data are shifted 2 seconds; (b) shifted 3 seconds; (c) shifted 4 seconds; (d) randomly shifted (up to 10 seconds back and forward)

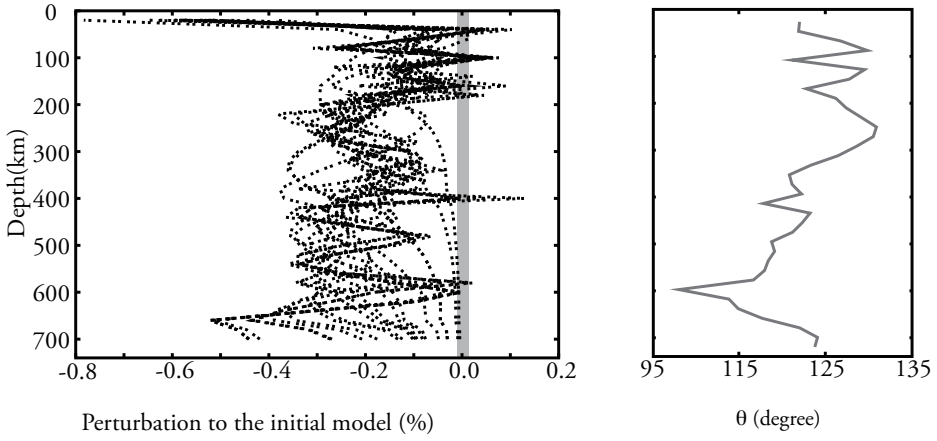


FIGURE 4.3: Synthetic inversion results for the systematically 1 s shifted seismograms (left). The time shift does not significantly affect the inversion results because the inner product of the partial derivatives with respect to the depth dependence of SH velocity and the partial derivatives for the time shift is small, albeit non-zero (right).

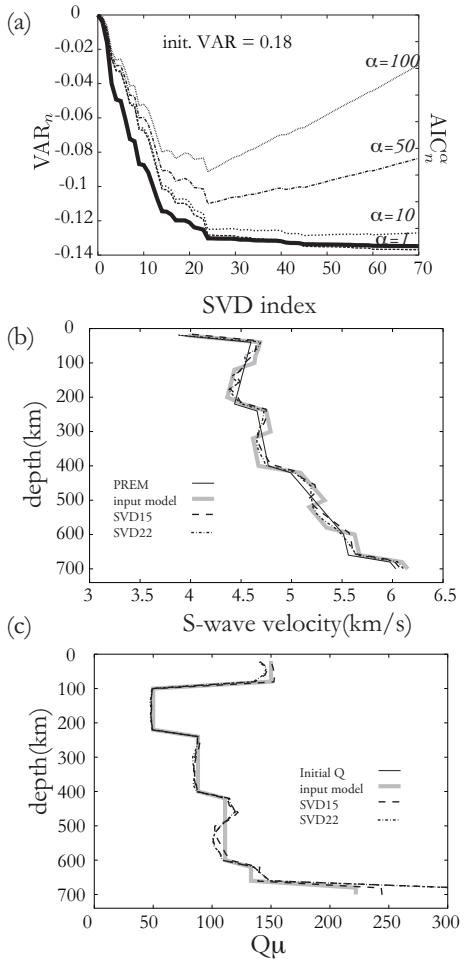


FIGURE 4.4: Results of inversion of synthetic data for a perturbation to only the elastic part of the model. (a) Variance (thick curve) and AIC_n^α values (thin curves). Note that the AIC_n^α curves have been adjusted to have the same value for $n = 1$, as only the relative values of AIC_n^α for a given α are meaningful. Note that the scale for the variance and AIC_n^α are respectively on the left and right hand sides of the box. (b) S-velocity for PREM (the starting elastic model for the inversion), the perturbed ("checkerboard test") input model used to computed the synthetic "observed data" for the test inversion, and the inversion results for $n = 15$ and $n = 22$. (c) "Initial Q_μ" model (starting point for the inversion as well as the Q model used to compute the synthetic "observed data," as Q_μ was not perturbed in the input model, and the Q_μ models obtained by the inversions for $n = 15$ and $n = 22$ (basically unchanged from the starting model for the synthetic inversion).

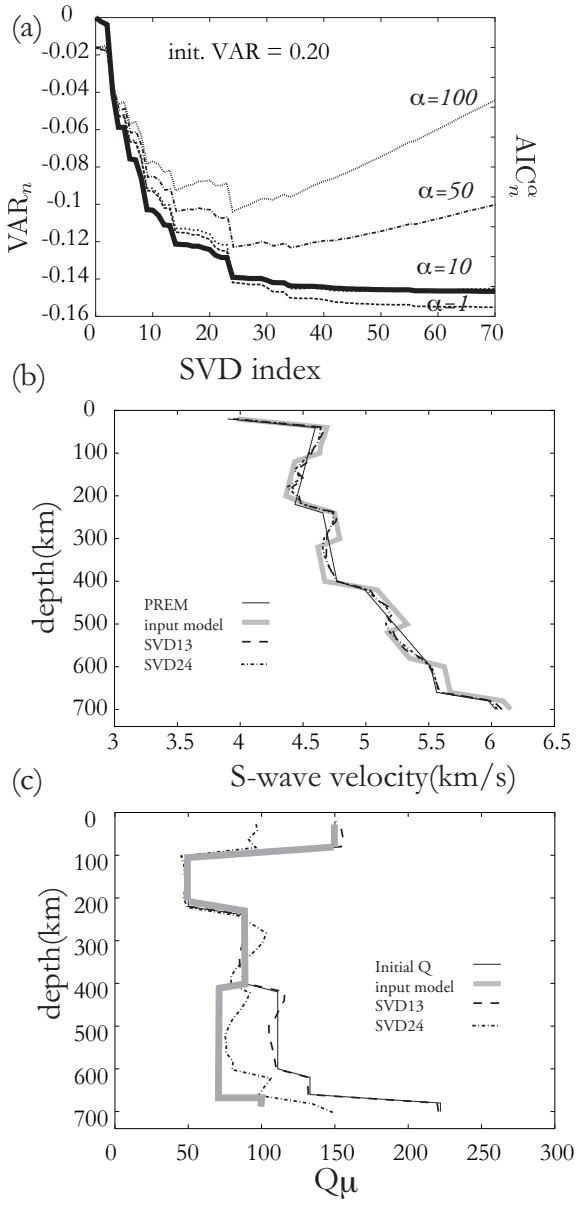


FIGURE 4.5: Same as Fig. 4.4, except that the model used to compute the synthetic "observed data" included perturbations to both the anelastic and elastic parts of the starting model for the synthetic inversion (PREM and "Initial Q" respectively).

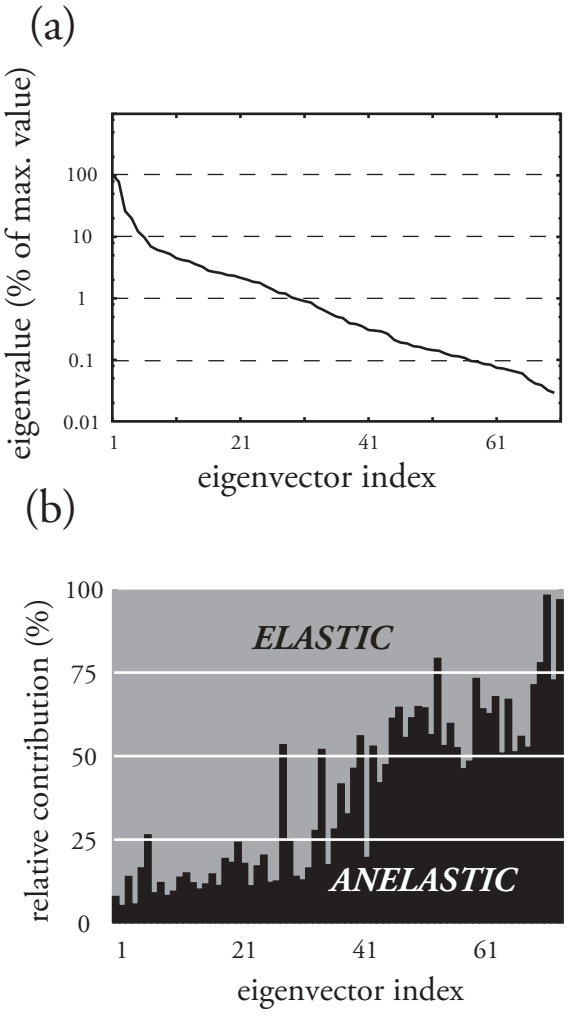


FIGURE 4.6: (a) the absolute value of the eigenvalues for the SVD inversion (plotted on a logarithmic scale); (b) the relative contribution of the elastic and anelastic components of the eigenvectors. This is an example figure for the "large dataset" for the Philippines West sub-region. However, results for other datasets are similar.

Chapter 5

1-D Simultaneous inversion results: upper mantle and mantle transition zone structure beneath Northwestern Pacific

On the basis of the success of these synthetic tests we now proceed to to conduct inversions of observed data for elastic and anelastic structure.

In Fig. 5.1 we show the inversion results for the Japan Sea region using a "small dataset" (104 records) with a criterion of amplitude ratio factors between 0.5 and 2.0 and cross correlations above 80 %, while in Fig. 5.2 we show results for the same region using a "large dataset" (284 records) with a criterion of an amplitude ratio factor between 0.25 and 4.0 and cross correlations above 40 %.

In Fig. 5.1a the thick curve is the variance reduction and the thin curves (labeled 1, 2, 3, 4 respectively) are the AIC_n^α values (see eqs. 2.37 and 2.38) for $\alpha = 1, 2, 3, 4$ respectively. Regardless of the value of α, there is a local minimum at $n = 3$ and another at $n = 21$. However, for $\alpha = 1$ it is just barely clear in Fig. 5.1a that $n = 21$ is a global minimum, whereas for $\alpha = 4$, $n = 21$ is a local minimum but

has a higher AIC value than for $n = 3$. If $\alpha = 2$ or $\alpha = 3$ is used, then $n = 21$ gives a clear global minimum.

In this thesis, as the number of records in a dataset increases we use higher values of α. This seems reasonable, because as the size of the dataset increases the redundancy of the dataset will increase. Our methods for selecting α are still empirical, and we will investigate more objective methods for choosing α in our future work.

Using $\alpha \geq 3$ for the small dataset and $\alpha \geq 5$ for the large dataset, we conclude that SVD21 for the small dataset or SVD30 for the large dataset are the most preferable models for this region. All of the preferred models show a positive S-wave velocity anomaly in the MTZ and in the layer from 300 to 400 km. A kink around 400-500 km suggests the possibility of a velocity discontinuity. Because the large dataset has a larger number of scattered data, the results fluctuate more but still show the same feature (a positive velocity anomaly). As for Q_μ, we do not see any significant perturbation when we choose SVD3 for either dataset but we obtain a lower Q_μ if we take a larger number of eigenvectors.

5.1 Technical issues

5.1.1 Variance reduction: Born vs. re-linearized

In the above inversion we estimate the variance using the first order Born approximation to compute the variance (rather than re-linearizing with respect to each model obtained by the inversion, as was done by Kawai et al., 2007ab and Konishi et al., 2009). To estimate the errors due to using the first order Born approximation to estimate the variance, we re-compute the synthetics for each of models SVD1 through SVD30 for the small dataset. Fig. 5.1a shows the variance reduction and AIC values ($\alpha = 1, 2, 3, 4$) calculated using the first order Born approximation, while Fig. 5.1b shows the AIC values obtained using the re-computed synthetics up to $n = 30$. Due to the non-linearity of the inversion, the behavior of

the re-computed variance reduction differs from that for the Born approximation and the absolute value of the variance reduction is smaller. However, both methods of calculation show that we have large variance reductions when go from $n = 2$ to $n = 3$ and from $n = 15$ to $n = 16$. The "real" variance has a global minimum (between $n = 0$ and $n = 30$) for $n = 20$. Because we do not use re-linearization in our inversions when we compute the variance, we obtained a global minimum of AIC_n^1 for $n = 3$. We also plot $\mathrm{AIC}_n^{0.5}$ for reference in order to find the value of α which gives a global AIC minimum around $n = 21$. We re-compute synthetics for each SVD_n model only for the examples in this section because it is computationally intensive, but we can expect that the "real" variance will have a minimum near the AIC minimum calculated using the first order Born approximation for the models presented below.

5.1.2 Iterative inversion

As the large difference between the variance reductions computed based on the Born approximation and re-computed without linear approximation indicates, the inversions are non-linear. One might thus question whether the inversions have sufficiently converged after a single inversion. We present one example of iterative inversion for the Japan Sea region. We use the model obtained by the initial inversion for the "small dataset" as the starting model for the second iteration (Note that for simplicity we fix the starting model to PREM and the "Initial Q_μ" above a depth of 200 km).

Fig. 5.3a shows the variance reduction and AIC values ($\alpha = 1, 2, 3, 4$) calculated using the first order Born approximation. Fig. 5.3b shows the SVD1 model for the second iteration, which is suggested as the best by AIC analysis. For $\alpha = 1$ only, $n = 13$ seems to be a global minimum, whereas for $\alpha \geq 2$, $n = 1$ is a global minimum. As we can see, the second iteration suggests just a small S-velocity shift (in a positive sense) in the shallower part and no significant change in the anelastic model. If we use the $n = 13$ model, we obtain a lower Q_μ and some fluctuations

in the S-velocity model. Still, the overall features, compared to the starting model of the first iteration, seems stable for both elastic and anelastic structure.

We now consider why the second iteration produces such a small change. The inversion process is non-linear, but the data themselves have a large degree of noise due to scattering. Due to this noise, the first iteration seems to have resolved almost all observable information, and the second iteration does not significantly change the model obtained by the first iteration. Thus the non-linearity affects the estimates of the variance reduction, but does not significantly affect the models themselves.

5.1.3 Statistical estimation

From eq. (2.28), we can rewrite the variance for each SVD expansion coefficient e_j as follows:

$$\text{var. } (e_j) = \frac{\sigma_d^2}{\lambda_j^2}. \tag{5.1}$$

If we assume that σ_d is the variance of the data itself, we can estimate the nominal error bars δe_j for each of the 70 coefficients:

$$\delta e_j = \frac{\sigma_d}{\lambda_j}. \tag{5.2}$$

Fig. 5.4 shows that the nominal error bars are small compared to the expansion coefficients. Of course in any inversion of this type, it is well known that the real errors may be many times greater than the nominal error bars due to systematic errors (e.g., the use of a 1-D parametrization). Thus the nominal error bars should not be taken at face value.

In order to know how the residual for each record $\delta \mathbf{d}^{(kp)}$ contributes to the model $\delta \mathbf{m}$, we examine the distribution of angles (in the multi-dimensional vector space) between the SVD coefficient vector $\mathbf{e}_n$ and the expansion coefficients for each record $\mathbf{e}_n^{(kp)}$ truncated above n:

$$\cos\theta_n^{(kp)} = \frac{\mathbf{e}_n \cdot \mathbf{e}_n^{(kp)}}{|\mathbf{e}_n|\ |\mathbf{e}_n^{(kp)}|}, \tag{5.3}$$

where

$$\mathbf{e}_n = (e_1\ e_2\ \ \cdots\ e_n\ \ 0\ \ \cdots\ 0)^{\mathbf{T}} \tag{5.4}$$

and

$$\mathbf{e}_n^{(kp)} = (e_1^{(kp)}\ e_2^{(kp)}\ \ \cdots\ e_n^{(kp)}\ \ 0\ \ \cdots\ 0)^{\mathbf{T}}. \tag{5.5}$$

As shown by eq. (2.33), the final model vector $\delta\mathbf{m}$ in eq. (2.30) can be expressed as the sum of a vector for each station for each event. If all of these individual vectors were nearly parallel, each would yield the same model for each individual waveform for each station for each event. However, this is not the case for the actual inversions.

Figs. 5.5a-c show histograms for $\theta_3^{(kp)}$, $\theta_{21}^{(kp)}$, $\theta_{70}^{(kp)}$ for the small dataset, while Figs. 5.6a-c show histograms for $\theta_3^{(kp)}$, $\theta_{30}^{(kp)}$, $\theta_{70}^{(kp)}$ for the large dataset. We find that even for $\theta_3^{(kp)}$, the angles between the individual coefficients and the coefficients determined from eq. (2.27), are much closer to 90° than to 0° and that as we increase the value of n, the average value of $\theta_n^{(kp)}$ increases. This tells us that no single individual datum $\delta\mathbf{d}^{(kp)}$ has sufficient power to determine the structure by itself.

For the expansion coefficients for the k-th event, we calculate angles between the SVD coefficient vector $\mathbf{e}_n$ and the expansion coefficients $\mathbf{e}_n^k$ truncated above n:

$$\cos\theta_n^k = \frac{\mathbf{e}_n \cdot \mathbf{e}_n^k}{|\mathbf{e}_n|\ |\mathbf{e}_n^k|}, \tag{5.6}$$

where

$$\mathbf{e}_n^k = (e_1^k\ e_2^k\ \ \cdots\ e_n^k\ \ 0\ \ \cdots\ 0)^{\mathbf{T}}. \tag{5.7}$$

Figs. 5.5d-f show histograms with black bars for θ_3^k, θ_{21}^k, θ_{70}^k for the small dataset, while Figs. 5.6d-f show histograms with black bars for θ_3^k, θ_{30}^k, θ_{70}^k for the large dataset. We find that the angles are closer to 0° than those for individual waveforms. This means that the combined set of waveforms for each particular event has stronger information than any individual waveform for that event. Figs. 5.5d-f and Figs. 5.6d-f also show, with gray bars, the angles between $\mathbf{e}_n^{k_1}$ and $\mathbf{e}_n^{k_2}$ for all pairs of events ($k_1 \neq k_2$):

$$\cos\theta_n^{(k_1,k_2)} = \frac{\mathbf{e}_n^{k_1} \cdot \mathbf{e}_n^{k_2}}{|\mathbf{e}_n^{k_1}|\ |\mathbf{e}_n^{k_2}|}. \tag{5.8}$$

We find that there are some events which have similar information on the structure (angles close to 0°). This fact supports the above discussion concerning the need for the factor α in AIC_n^α, because the data are redundant to some extent. However, we also find that if we take larger n, there exist pairs of events which have very different information (angles larger than 90°, i.e., negative correlation of the respective model contributions). This shows that there is a discrepancy between some events for large n and that we therefore should not use all 70 coefficients in our models, which is confirmed by the increase in the AIC values for larger values of n. This discrepancy is probably due to 3-D structure.

Similarly we calculate angles between the SVD coefficient vector $\mathbf{e}_n$ and the expansion coefficients $\mathbf{e}_n^p$ for the p-th station truncated above n:

$$\cos\theta_n^p = \frac{\mathbf{e}_n \cdot \mathbf{e}_n^p}{|\mathbf{e}_n|\ |\mathbf{e}_n^p|}, \tag{5.9}$$

where

$$\mathbf{e}_n^p = (e_1^p\ e_2^p\ \ \cdots\ e_n^p\ \ 0\ \ \cdots\ 0)^{\mathbf{T}}. \tag{5.10}$$

Figs. 5.5g-i show histograms with black bars for θ_3^p, θ_{21}^p, θ_{70}^p for the small dataset, while Figs. 5.6g-i show histograms with black bars for θ_3^p, θ_{30}^p, θ_{70}^p for the large dataset. We find that the angles are slightly closer to 0° than for the individual

waveforms, but not as much as is observed for θ_n^k. This is because the independent information on the depth dependence of the structure is very limited in the waveforms for some stations, as the sources are very close to each other and the epicentral distances are similar to those for other nearby stations. Figs. 5.5g-i and Figs. 5.6g-i also show, with gray bars, the angles between $\mathbf{e}_n^{p_1}$ and $\mathbf{e}_n^{p_2}$ for all pairs of stations ($p_1 \neq p_2$):

$$\cos\theta_n^{(p_1,p_2)} = \frac{\mathbf{e}_n^{p_1} \cdot \mathbf{e}_n^{p_2}}{|\mathbf{e}_n^{p_1}|\ |\mathbf{e}_n^{p_2}|}. \tag{5.11}$$

We find (gray bars in Figs. 5.5d and 5.5g) that $\theta_3^{p_1,p_2}$ has one peak in the range 0-20° and another is around 60-80° for the small dataset, and that $\theta_3^{p_1,p_2}$ for the large dataset are widely distributed from 0° to 180°. The smaller angles $\theta_3^{p_1,p_2}$ are primarily for stations or events near to one another, suggesting the partial redundancy of the dataset. As n increases, θ_n^p is normally distributed around 90° because of the lack of resolving power for the larger basis set, which includes eigenvectors for smaller eigenvalues. We conclude that in general each record contributes weakly to the inversion results, and that a large number of waveforms must be analyzed systematically as a single dataset in order to extract reliable information on the dataset. Thus visual inspection of individual seismogram pairs (synthetic and observed) is unlikely to be edifying.

5.1.4 Quality control stacks

As waveform inversion is a "black box" statistical procedure, we desire some way to visually confirm the reasonableness of the inversion results. We therefore calculate "quality control stacks" of both the data and synthetics. We align the waveforms using the onset time after applying the same time shift (static correction) as used in the inversion. Note that these stacks are merely ancillary presentations for purposes of visualization of the results of the inversions, and are not used in the inversions. Also note that both the data and synthetics have been stacked in the same way. Fig. 5.7 shows the quality control stack for event 20041107 as an

example. We can see that the difference between the stack of the observed data and the stack of synthetics for the initial model is due primarily to the phase difference. The onset-to-peak time of the arrival packet for the synthetics for the initial model seems to be more longer than that of the observed data. By adding a positive S-wave velocity anomaly in the MTZ (SVD3), the stack has been modified moderately but still we see a large difference for the amplitude and the shape of the first packet: the amplitude should be smaller and the shape should be more broadened. If we apply the SVD21 model with a lower Q_μ, this misfit is going to be reduced obviously. Although the quality control stacks are just for confirmation of the inversion technique, we can visually realize how the model change governs the waveform change.

5.1.5 Comparison of data to synthetics

Fig. 5.8 shows record sections for the event 20041107. We compare observed data to synthetic waveforms for the initial model (left panel) and for the SVD21 model (right panel) for the small dataset. We apply a 20-200 s bandpass filter to both the observed and synthetic waveforms. We plot the entire dataset including records which do not satisfy the criteria for the small dataset. We see an improvement of the onset-to-peak time in the first branch because there is a positive velocity anomaly in the MTZ. From these record sections, we can conclude that we have extracted information on the structure from the combined dataset of 20-200 s bandpass filtered waveforms and that we can reasonably infer the average structure.

5.2 Philippines Sea Region

The sky-blue lines in Fig. 5.10 show the inversion results for the Philippines Sea West region using the “small dataset” (136 records) and the sky-blue lines in Fig. 5.10 show the inversion results for the same region using the “large dataset” (400

records). For the small dataset SVD15 is the preferred model for $\alpha = 2, 3, 4$, while for the large dataset SVD17 is the preferred model for $\alpha = 2, 5, 10$ (see Fig. 5.9a,f). Although the SH model obtained using the large dataset shows fluctuating features, both models have a positive anomaly in the MTZ; Q_μ is around 50 throughout the MTZ for both models.

The blue lines in Fig. 5.10 show the inversion results for the Philippines Sea Middle region for the "small dataset" (233 records) and the blue lines in Fig. 5.10 show the inversion results for the same region for the "large dataset" (471 records). For the small dataset SVD14 is the preferred model for $\alpha = 4$, while for the large dataset SVD13 is the preferred model for $\alpha = 10$ (see Fig. 5.9b,g). Although the S-wave velocity model shows fluctuating features which may be caused by laterally heterogeneous structure in the target region, the average velocity model is higher than PREM in the MTZ. The Q_μ model shows very low value (around 50) throughout the MTZ.

The purple lines in Fig. 5.10 show the inversion results for the Philippines Sea East region using the "small dataset" (156 records) and the purple lines in Fig. 5.10 show the inversion results for the same region using the "large dataset" (419 records). For the small dataset SVD14 is the preferred model for $\alpha = 3, 4$, while for the large dataset SVD17 is the preferred model for $\alpha = 5, 10$ (see Fig. 5.9c,h). The S-wave velocity in the MTZ shows a positive anomaly in the MTZ and in the upper mantle. Q_μ is around 50 throughout the MTZ.

5.3 Pacific Sea Region

The red lines in Fig. 5.11 show the inversion results for the Pacific Sea West region for the "small dataset" (163 records) and the red lines in Fig. 5.11 show the inversion results for the same region using the "large dataset" (645 records). For the small dataset SVD32 is the preferred model for $\alpha = 1, 2, 3, 4$, while for the large dataset SVD33 is the preferred model for $\alpha = 5, 10$ (see Fig. 5.9d,i). From both inversions we can consistently see a positive SH anomaly in the MTZ with

a kink above the 400 km discontinuity probably connected with a discontinuity depth change. Q_μ is again very low but in the MTZ is relatively larger (50-100) than in other subregions.

The orange lines in Fig. 5.11 show the inversion results for the Pacific Sea East region for the "small dataset" (87 records) and the orange lines in Figs. 5.11 show the inversion results for the same region for the "large dataset" (348 records). For the small dataset SVD20 is the preferred model for $\alpha = 2, 3, 4$, while for the large dataset SVD25 is the preferred model for $\alpha = 5, 10$ (see Fig. 5.9e,j). The inversion results show a fluctuating S-wave velocity model in the MTZ but the average velocity is not far from PREM. However, around 400 km depth we can see a slight negative S-wave velocity anomaly. Q_μ is very low but for the inversion using the large dataset, there might be a relatively large Q_μ (50-100) in the MTZ. In this subregion, the difference between the attenuation models obtained using the small and large datasets is the strongest. The reason for this discrepancy might be that the target subregion includes a strong lateral heterogeneity because of the complicated geometry beneath the Izu-Bonin-Mariana Arc.

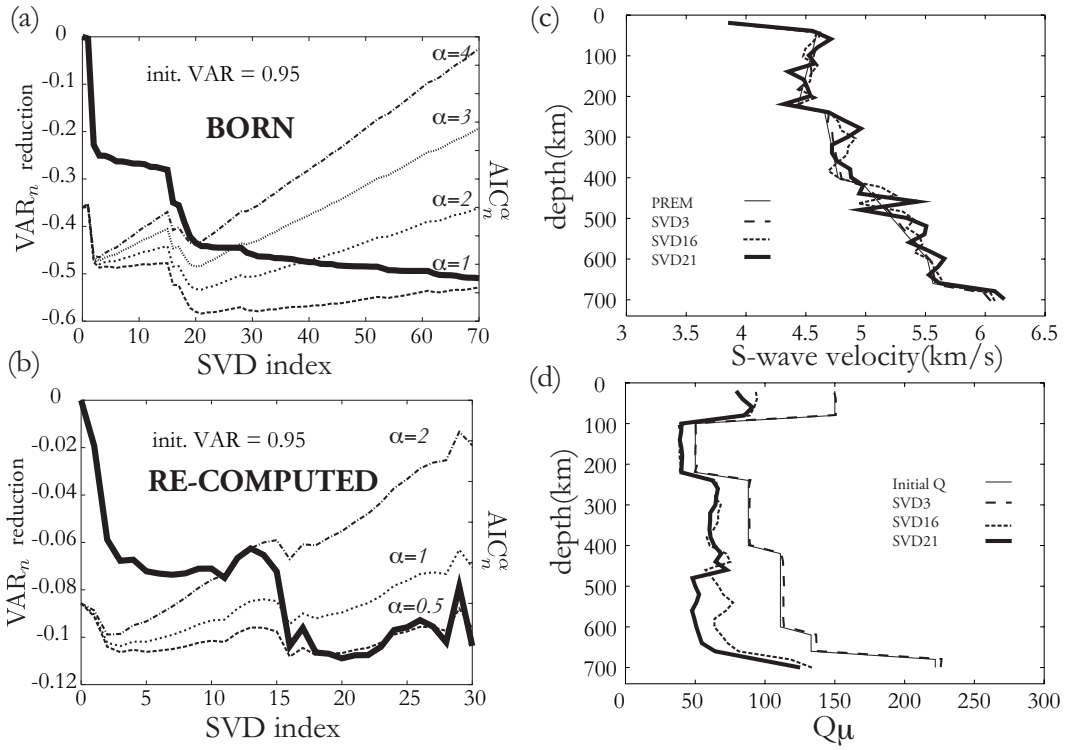


FIGURE 5.1: Results of inversion of the "small dataset" for the Japan Sea region. (a) Variance (thick curve) (computed using the 1st order Born approximation) and AIC^α_n values (thin curves). (b) Re-computed variance (thick curve) and AIC^α_n values (thin curves) up to $n \leq 30$. Note that we plot the re-computed variance and AICs only in this figure. $\mathrm{AIC}^{0.5}_n$ is also shown as a reference. (c) S-velocity for PREM (the starting elastic model for the inversion), and the inversion results for $n = 3$, $n = 16$ and $n = 21$. (d) "Initial Q_μ" model (starting point for the inversion) and the inversion results for $n = 3$, $n = 16$ and $n = 21$.

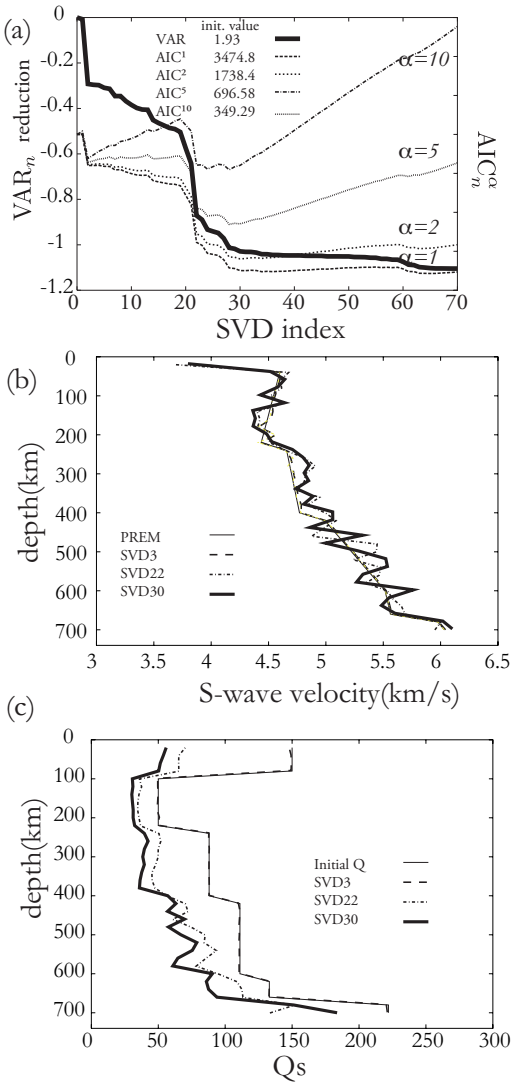


FIGURE 5.2: Inversion result for the Japan Sea Region for the "large dataset."

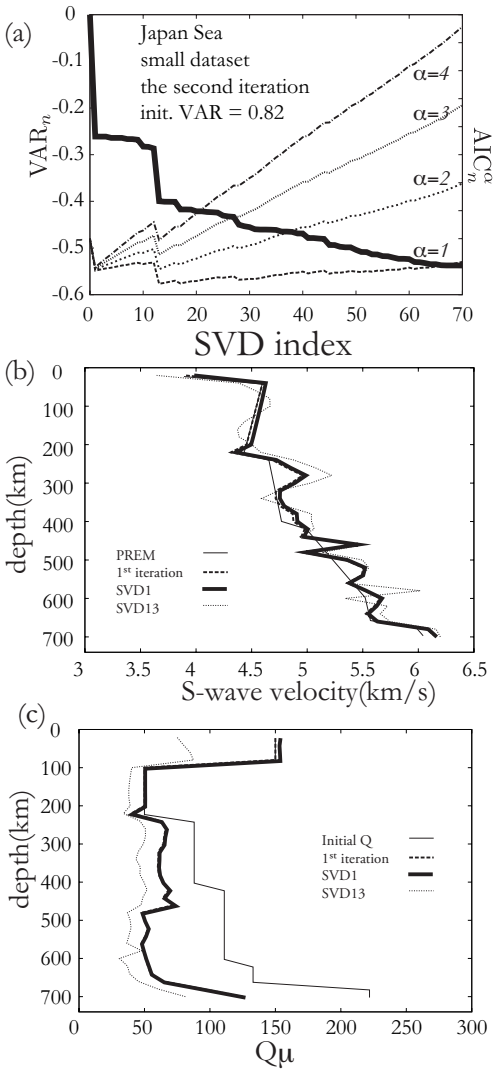


FIGURE 5.3: Results of the second iterative inversion using the "small dataset" for the Japan Sea region. (a) Variance (thick curve) (computed using the 1st order Born approximation) and AIC_n^{α} values (thin curves). (b) S-velocity for PREM (the starting elastic model for the first iteration), the starting elastic model for the second iteration which is the SVD21 model from the first iterative inversion below 200 km depth (above 200 km we use PREM), and the second iteration results for $n = 1$ and $n = 13$. (c) "Initial Q_μ" model (the starting anelastic model for the first iteration), the starting anelastic model for the second iteration which is the SVD21 model from the first iteration below 200 km depth (above 200 km we use PREM), and the second iteration results for $n = 1$ and $n = 13$.

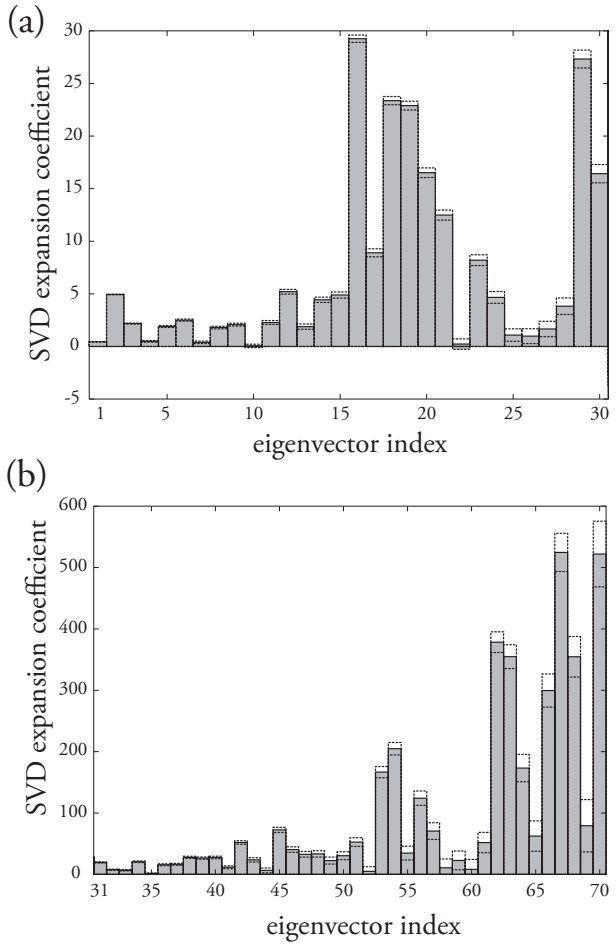


FIGURE 5.4: SVD expansion coefficients e_j for the inversion for the small dataset. Gray bars show the estimated values and two dotted bars show the error bars for each coefficient. Eq. (2.26) shows that e_j tends to increase with the eigenvector index j. We thus show this histogram in two columns: (a) from the first to thirtieth modes; (b) from the thirty-first to the seventieth modes.

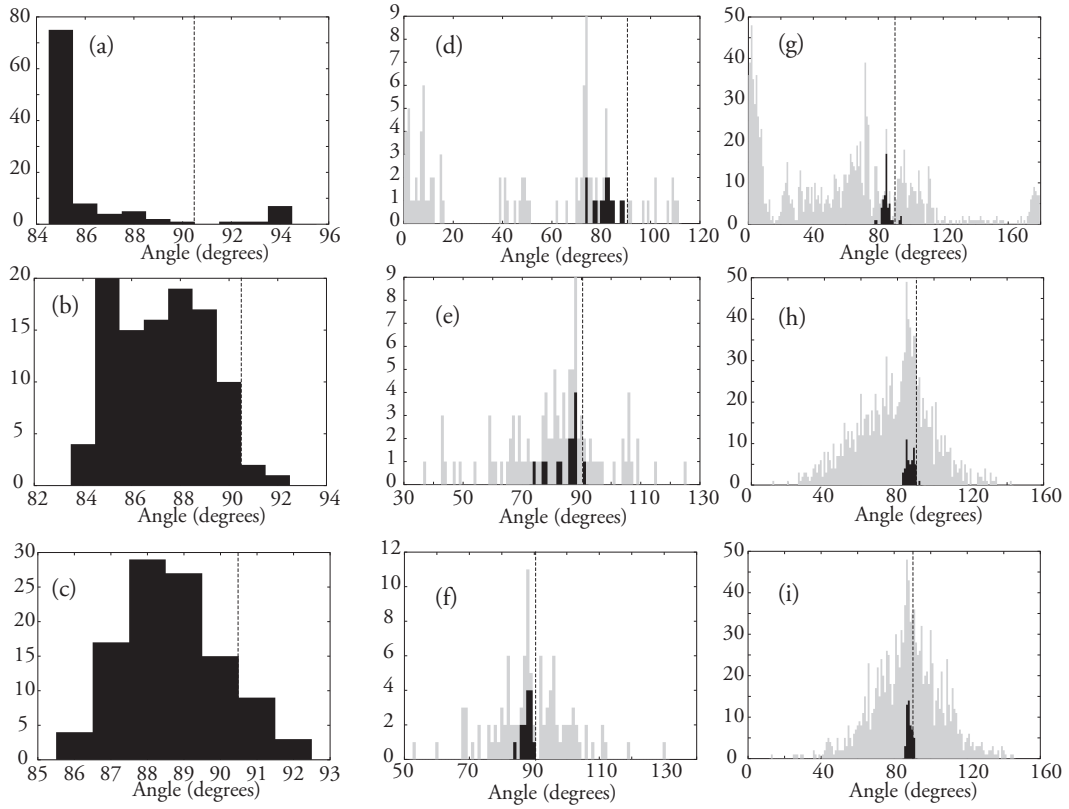


FIGURE 5.5: For the small dataset we plot histograms for the angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^{(kp)}$ for (a) $\theta_3^{(kp)}$, (b) $\theta_{21}^{(kp)}$, and (c) $\theta_{70}^{(kp)}$; angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^k$ for each event (black bars), and angle between the vector $\mathbf{e}_j^{k_1}$ and $\mathbf{e}_j^{k_2}$ (gray bars) for (e) $j = 3$, (e) $j = 21$ and (f) $j = 70$; angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^p$ for each station (black bars), and angle between the vector $\mathbf{e}_j^{p_1}$ and $\mathbf{e}_j^{p_2}$ (gray bars) for (g) $j = 3$, (h) $j = 21$ and (i) $j = 70$. Note that in each panel, the dotted line denotes an angle of 90°.

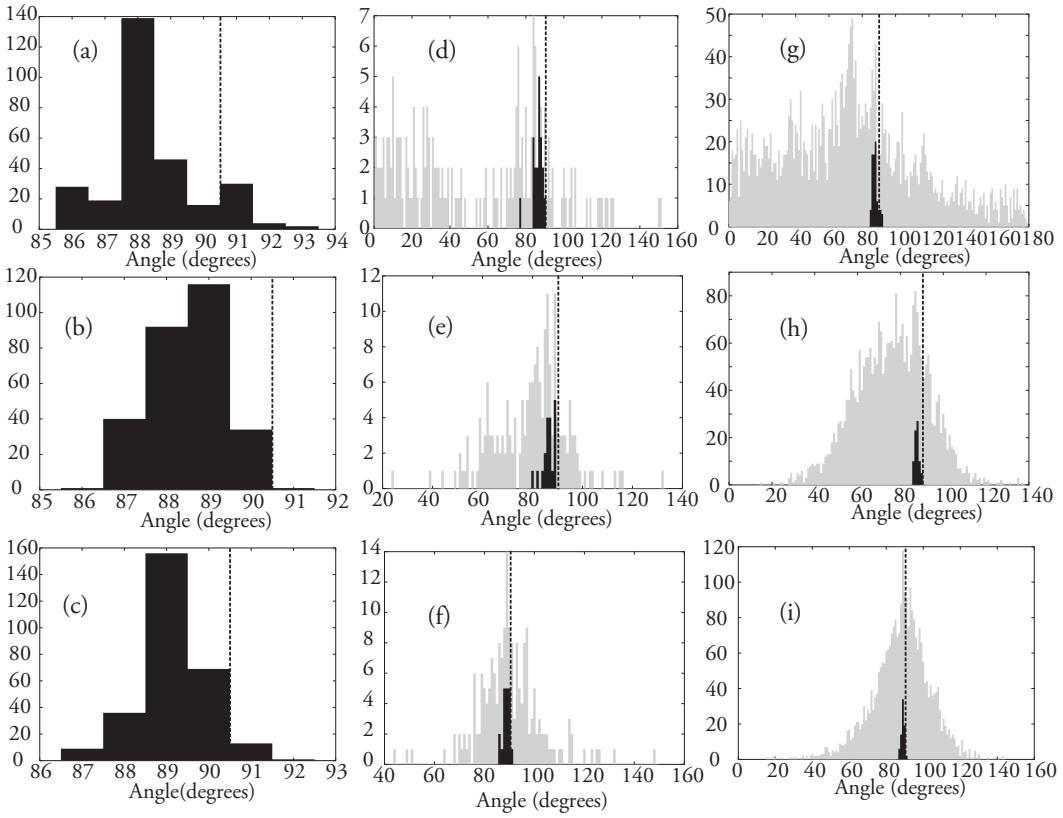


FIGURE 5.6: For the large dataset we plot histograms for the angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^{(kp)}$ for (a) $\theta_3^{(kp)}$, (b) $\theta_{30}^{(kp)}$, and (c) $\theta_{70}^{(kp)}$; angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^k$ for each event (black bars), and angle between the vector $\mathbf{e}_j^{k_1}$ and $\mathbf{e}_j^{k_2}$ (gray bars) for (e) $j = 3$, (e) $j = 30$ and (f) $j = 70$; angle between the vector $\mathbf{e}_j$ and $\mathbf{e}_j^p$ for each station (black bars), and angle between the vector $\mathbf{e}_j^{p_1}$ and $\mathbf{e}_j^{p_2}$ (gray bars) for (g) $j = 3$, (h) $j = 30$ and (i) $j = 70$. Note that in each panel, the dotted line denotes an angle of 90°.

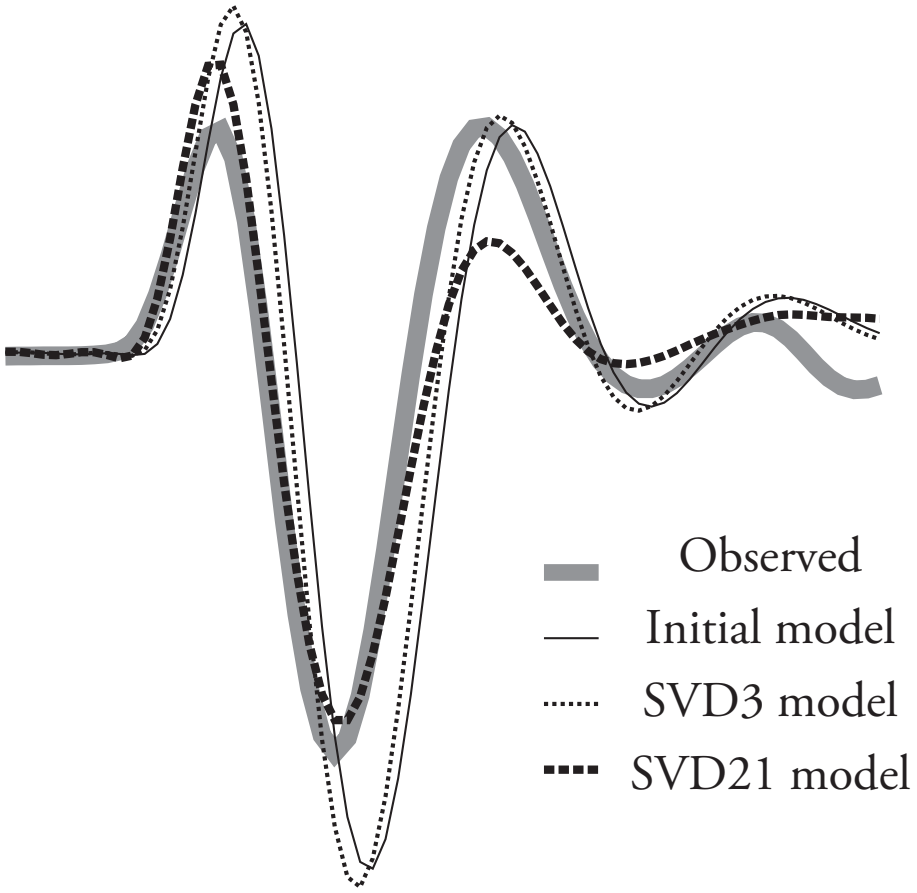


FIGURE 5.7: The "quality control stacks " for the event 20041107. The thick grey curve denotes the observed stack, the thin black solid curve the synthetic stack for the initial model, the thin dotted curve the synthetic stack for the SVD3 model with a positive S-wave velocity anomaly, and the thick dotted curve denotes the synthetic stack for the SVD21, the preferred model with a lower Q_μ. Synthetics waveforms are calculated using the 1st order Born approximation.

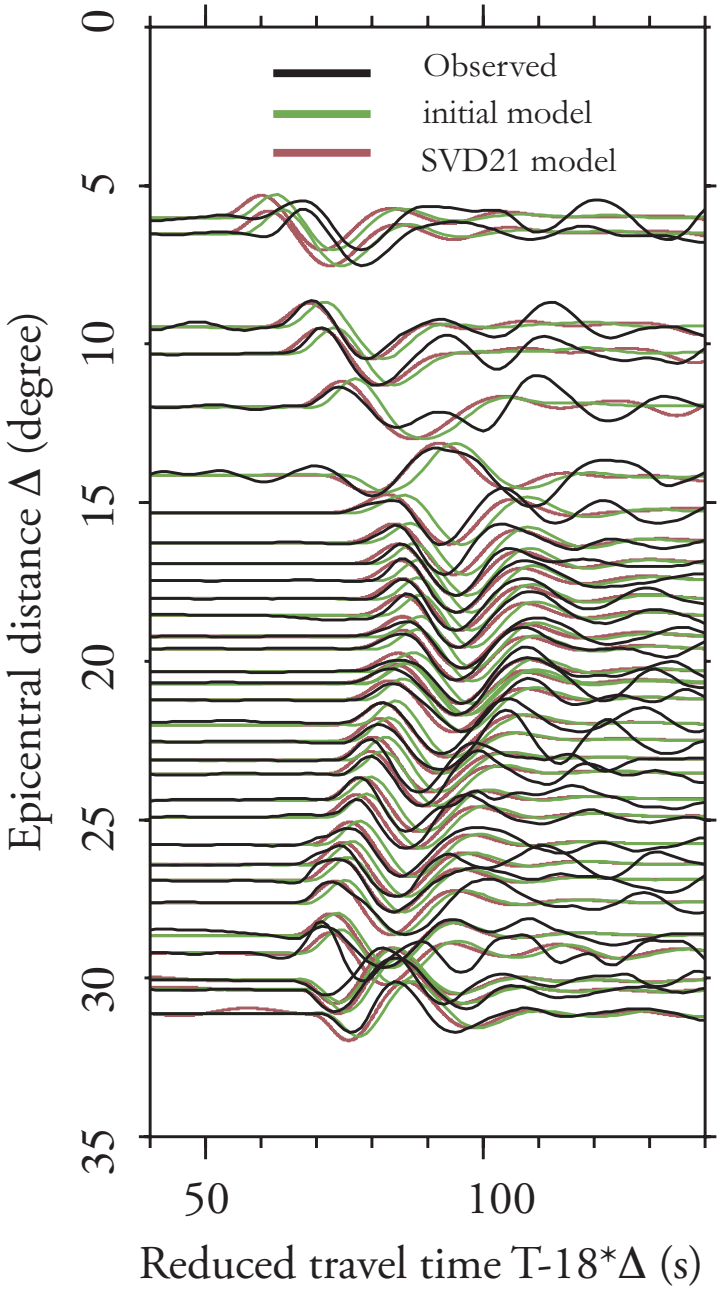


FIGURE 5.8: 20-200s bandpass filtered section record for the event 20041107. Black lines show the observed waveforms and overlapped green lines show the synthetic waveforms for the initial model and red lines show the synthetic waveforms for the SVD21 preferred model. The time duration is reduced to be that used for the inversion. The number of the waveforms is reduced so that the waveform comparison is visible.

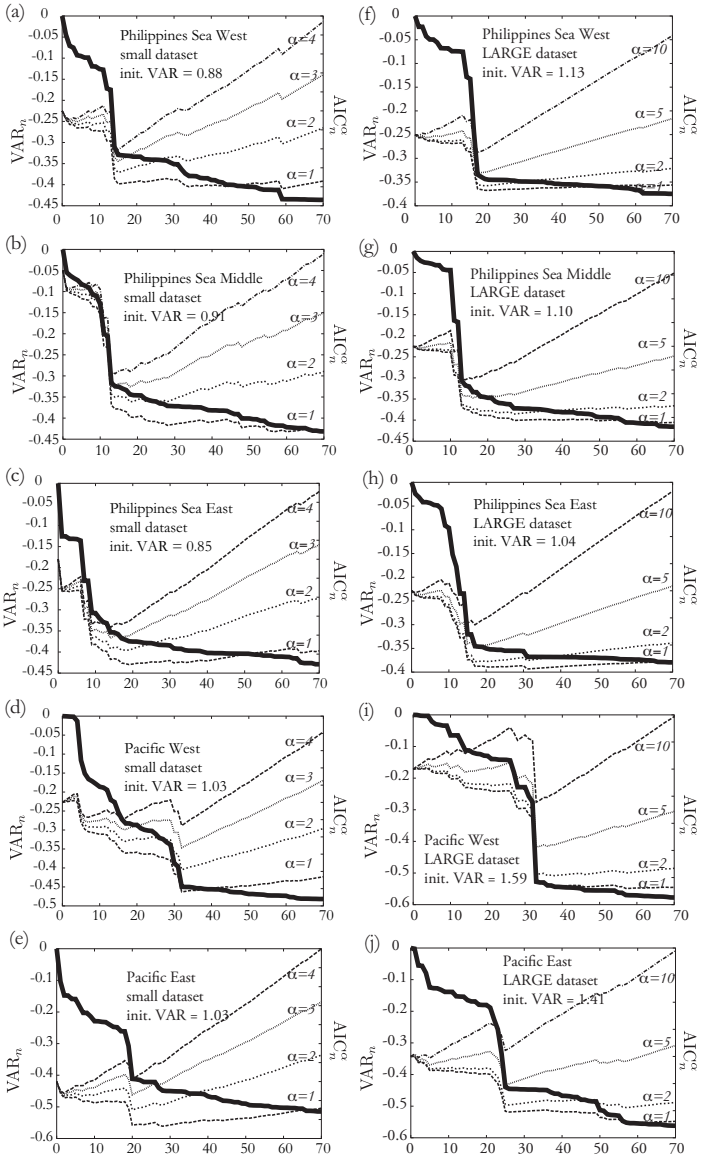


FIGURE 5.9: Variance (thick curve) and AIC_n^α values (thin curves) for inversion of the "small dataset" (a-e) and "large dataset" (f-j) for five sub-regions: (a,f) show the results of the Philippines Sea West Region, (b,g) show the results of the Philippines Sea Middle Region, (c, h) show the results of the Philippines Sea East Region, (d,i) show the results of the Pacific West Region, (e, j) show the results of the Pacific East Region.

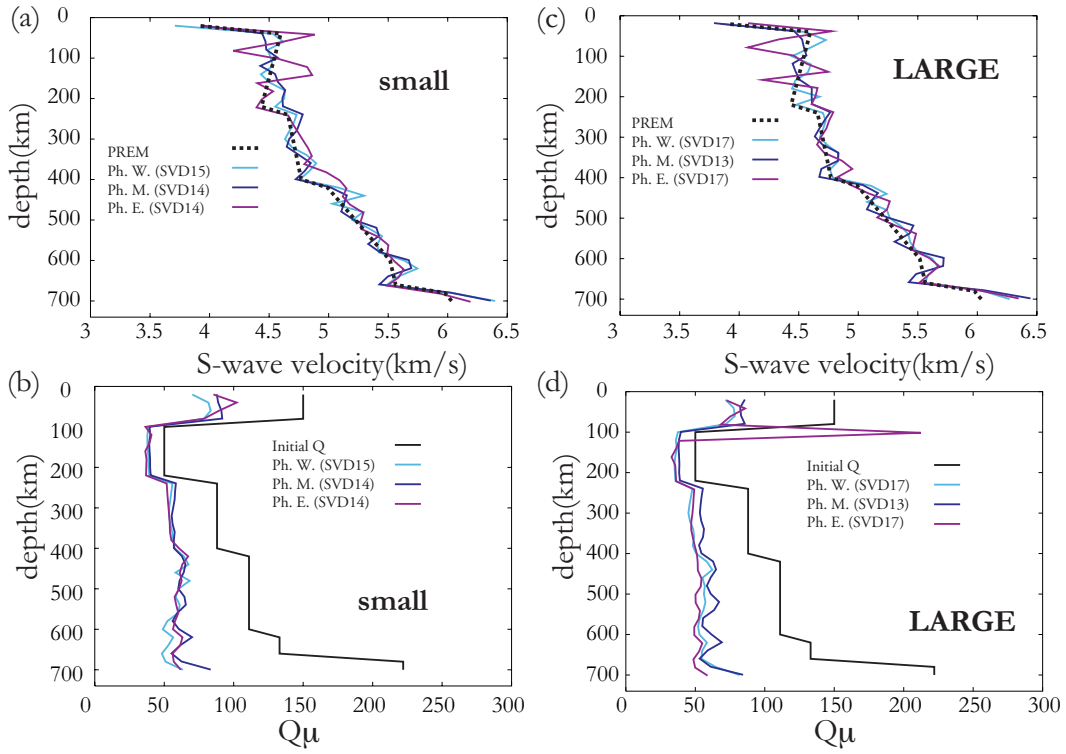


FIGURE 5.10: Results of inversion of the "small dataset" (a, b) and "large dataset" (c, d) for the three sub-regions in the Philippines. (a, c) S-velocity for PREM (black curve: the starting elastic model for the inversion), and the preferred inversion result (colored curve). (b, d) "Initial Q" model (thin curve: starting point for the inversion) and the preferred inversion result (colored curve). Sky blue curves are for the Philippines Sea West Region, blue curves for the Philippines Sea Middle Region, and purple curves for the Philippines Sea East Region.

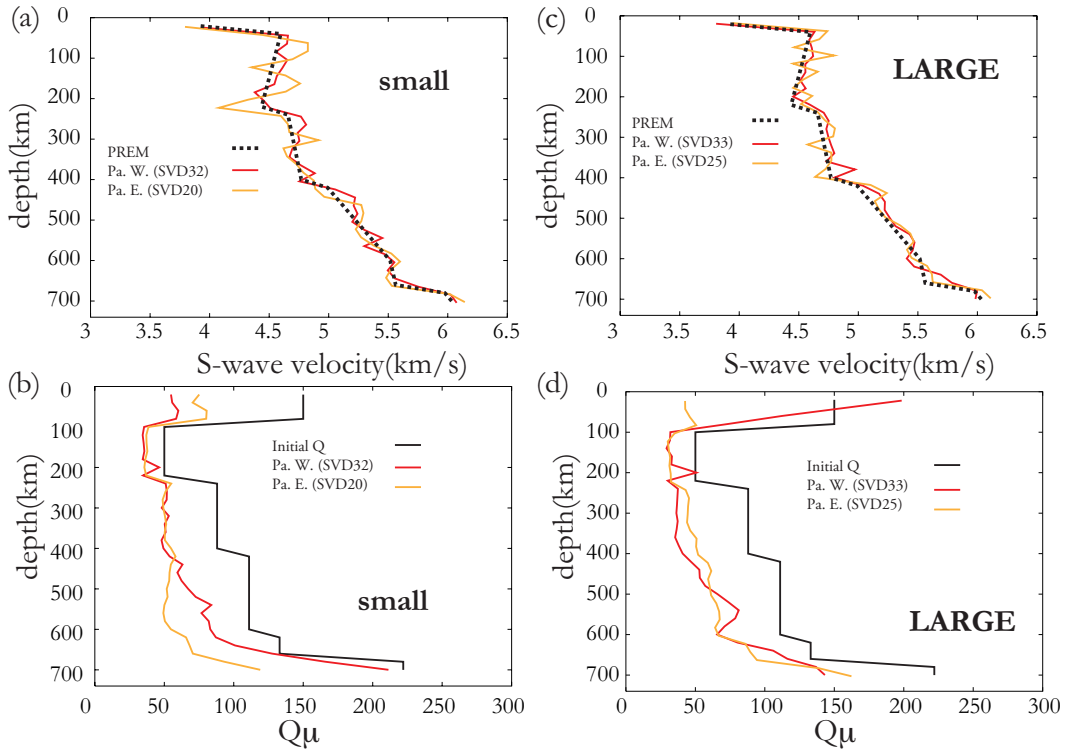


FIGURE 5.11: Results of inversion of the "small dataset" (a, b) and "large dataset" (c, d) for the two sub-regions in the Pacific. (a, c) S-velocity for PREM (black curve: the starting elastic model for the inversion), and the preferred inversion result (colored curve). (b, d) "Initial Q" model (thin curve: starting point for the inversion) and the preferred inversion result (colored curve). Red curves are for the Pacific West Region, and orange curves for the Pacific East Region.

Chapter 6

Inversion using the conjugate gradient (CG) method

As discussed above, we have succeeded in developing and applying methods for simultaneous localized waveform inversion for 1-D elastic and anelastic structure. The logical continuation of this research is waveform inversion for 3-D elastic and anelastic structure. Because the number of unknown parameters in an inversion for 3-D structure will, roughly speaking, be the cube of that for a 1-D inversion, we should seek an inversion method which is more efficient than SVD inversion. Here we consider the use of the conjugate gradient (CG) method (Beckman, 1960). The CG method is well known in general as a useful tool in inversion studies. In this section, we apply the methods which we derived in Chapter 2 to visualize the results of waveform inversions using CG.

6.1 Conjugate gradient (CG) method

We rewrite the normal equations eq. (2.23) as follows:

$$\mathbf{A^T A}\delta\mathbf{m} = \mathbf{A^T}\delta\mathbf{d}. \tag{6.1}$$

Suppose that $\{\mathbf{p}_j\}$ is a sequence of M mutually orthogonal conjugate vectors with respect to $\mathbf{A^T A}$ (recall that $\mathbf{A^T A}$ is symmetric and positive definite), i.e. that:

$$\mathbf{p}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_k = 0 \qquad \text{for} \qquad j \neq k. \tag{6.2}$$

First we write the solution of eq. (6.1) as an expansion in terms of the conjugate vectors:

$$\delta \mathbf{m} = \sum_{j=1}^{M} a_j \mathbf{p}_j, \tag{6.3}$$

where the CG expansion coefficients a_j are the unknowns. We define the vector $\mathbf{a}$ to be:

$$\mathbf{a}^{\mathbf{T}} = (a_1 \; a_2 \; \cdots \; a_M). \tag{6.4}$$

Now eq. (6.3) can be written in matrix form as follows:

$$\delta \mathbf{m} = \mathbf{P a} \tag{6.5}$$

where we define the matrix $\mathbf{P}$ to be:

$$\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \; \cdots, \; \mathbf{p}_M). \tag{6.6}$$

Substituting eq. (6.3) into eq. (6.1), multiplying both sides on the left by the transpose of some particular $\mathbf{p}_k$, and using eq. (6.2), we obtain:

$$\mathbf{p}_k^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \delta \mathbf{d} = a_k \mathbf{p}_k^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_k. \tag{6.7}$$

Thus the CG expansion coefficients are given by:

$$a_j = \frac{\mathbf{p}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \delta \mathbf{d}}{\mathbf{p}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_j}. \tag{6.8}$$

Using eq. (6.4) we can write this in matrix form as follows:

$$\mathbf{a} = \mathbf{L}^{-1} \mathbf{P}^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \delta \mathbf{d} \tag{6.9}$$

where we define the $M \times M$ diagonal matrix $\mathbf{L}$ by analogy to $(\mathbf{\Lambda^T \Lambda})$ in eq. (2.27) as follows:

$$\mathbf{L} = \begin{pmatrix} \mathbf{p}_1^\mathbf{T}\mathbf{A}^\mathbf{T}\mathbf{A}\mathbf{p}_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{p}_M^\mathbf{T}\mathbf{A}^\mathbf{T}\mathbf{A}\mathbf{p}_M \end{pmatrix}. \tag{6.10}$$

Assuming that the data are uncorrelated and all have equal variance σ_d^2, we can write the covariance matrix for $\mathbf{a}$ as follows in the CG space of $\mathbf{A^T A}$:

$$\text{covar. } (\mathbf{a}) = \sigma_d^2 \mathbf{L}^{-1}, \tag{6.11}$$

while in the model space, the covariance matrix is written as:

$$\text{covar. } (\delta\mathbf{m}) = \sigma_d^2 (\mathbf{A^T A})^{-1} = \sigma_d^2 \sum_{j=1}^{M} \frac{\mathbf{p}_j \mathbf{p}_j^\mathbf{T}}{\mathbf{p}_j^\mathbf{T}\mathbf{A}^\mathbf{T}\mathbf{A}\mathbf{p}_j} = \sigma_d^2 \mathbf{P}\mathbf{L}^{-1}\mathbf{P}^\mathbf{T}. \tag{6.12}$$

In the actual inversions we sum over only the first n conjugate vectors as follows:

$$\delta\mathbf{m}_n = \sum_{j=1}^{n} a_j \mathbf{p}_j. \tag{6.13}$$

Thus the covariance matrix in the model space is given by:

$$\text{covar. } (\delta\mathbf{m}_n) = \sigma_d^2 \sum_{j=1}^{n} \frac{\mathbf{p}_j \mathbf{p}_j^\mathbf{T}}{\mathbf{p}_j^\mathbf{T}\mathbf{A}^\mathbf{T}\mathbf{A}\mathbf{p}_j}. \tag{6.14}$$

In practice, the conjugate vectors $\{\mathbf{p}_j\}$ are chosen by taking the first conjugate vector $\mathbf{p}_1$ to be $-\mathbf{A^T}\delta\mathbf{d}$, in order to minimize the data variance as rapidly as possible at the first step. At each successive iteration, we choose the next conjugate vector which is as close as possible to the gradient vector while still being conjugate. We follow Reeves (1964) below. Let $\mathbf{r}_k$ be the residual (in the model space) at the $k-$th iteration:

$$\mathbf{r}_k = \mathbf{A}^\mathbf{T}\delta\mathbf{d} - \mathbf{A}^\mathbf{T}\mathbf{A}\delta\mathbf{m}_k. \tag{6.15}$$

Then we choose the next basis vector $\mathbf{p}_{k+1}$ to be as close as possible to $\mathbf{r}_k$, while also being conjugate to all previous basis vectors:

$$\mathbf{p}_{k+1} = \mathbf{r}_k - \sum_{i \leq k} \frac{\mathbf{p}_i^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{r}_k}{\mathbf{p}_i^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_i} \mathbf{p}_i. \tag{6.16}$$

We then calculate the next CG expansion coefficient as follows:

$$a_{k+1} = \frac{\mathbf{p}_{k+1}^{\mathbf{T}} \mathbf{r}_k}{\mathbf{p}_{k+1}^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_{k+1}}. \tag{6.17}$$

As we choose $\mathbf{p}_0 = \mathbf{A}^{\mathbf{T}} \delta \mathbf{d}$, the CG method is very much problem specific, but at the same time eq. (6.16) shows that we can construct an orthogonal basis system in both the model space and the data space. We thus can expect that the CG method can reach a solution using a smaller number of independent parameters (CG expansion coefficients) than SVD inversion. However, we still can evaluate statistical parameters as we did in the SVD inversion, because the conjugate vectors are orthogonal with respect to $\mathbf{A}^{\mathbf{T}} \mathbf{A}$ as were the SVD vectors. (The covariance matrix is thus diagonal.)

6.2 Technical issues

We conduct inversions of the observed data used in Chapter 5 to invert for elastic and anelastic structure. In this section we compare results obtained by the SVD inversion and by the CG inversion. We consider as an example the "LARGE dataset" for the Philippines Sea West region, but note that every other dataset shows behavior similar to that for this dataset.

6.2.1 Sensitivity study of conjugate vectors to elastic and anelastic structure

Fig. 6.1 shows the relative contribution to the conjugate vectors of the elastic and anelastic parts of the model, defined respectively as P_j^{elast} and P_j^{anel}, as follows:

$$P_j^{\text{elast}} = \frac{|\mathbf{p}_j^{\text{elast}}|^2}{|\mathbf{p}_j|^2} \tag{6.18}$$

$$P_j^{\text{anel}} = \frac{|\mathbf{p}_j^{\text{anel}}|^2}{|\mathbf{p}_j|^2}, \tag{6.19}$$

where $\mathbf{p}_j^{\text{elast}}$ and $\mathbf{p}_j^{\text{anel}}$ are the elastic and anelastic components of each conjugate vector $\mathbf{p}_j$. In Fig. 6.1, we normalize the value of P_j^{elast} and P_j^{anel} so that:

$$\sum_{j=1}^{M} P_j^{\text{elast}} = \sum_{j=1}^{M} P_j^{\text{anel}}. \tag{6.20}$$

Comparing Fig. 4.6 to Fig. 6.1, it appears that the CG vectors significantly reflect anelastic structure for lower values of j than the SVD vectors. This suggests that CG will be significantly more efficient than SVD in obtaining similar models.

6.2.2 VAR and AIC analysis

We truncate the conjugate vector expansion (eq. 6.13), using AIC as a guide. The variance VAR_n and AIC_n for the model CG_n (the model obtained by using n in eq. 6.13) are written respectively as follows:

$$\text{VAR}_n = \frac{|\mathbf{d}_{\text{OBS}} - \mathbf{d}_{\text{SYN}}|^2}{|\mathbf{d}_{\text{OBS}}|^2} \tag{6.21}$$

$$\text{AIC}_n = \text{ND} \ln 2\pi + \text{ND} \ln(\text{VAR}_n) + \text{ND} + 2(n+1). \tag{6.22}$$

We again introduce the empirical redundancy parameter α in order to write the number of independent data points ND as follows:

$$\mathrm{ND}^{\alpha} = \frac{1}{\alpha} KPT/20 \tag{6.23}$$

where ND^1 is the upper bound on the value of ND (see the previous discussion in the subsection 2.4). Note that the denominator (20) in eq. (6.23) is the ratio of the shortest period in the synthetics (20 s) to the sampling interval of the data (1 s) for this particular inversion. Fig. 6.2 shows the results for both the CG and SVD inversions. In Figs. 6.2a and b, the colored curves are the variance reduction and the black curves are the AIC_n^{α} values for $\alpha = 1, 2, 5$, and 10, respectively. We can see that the variance reduction curve for the CG inversion shows a steeper gradient than that for the SVD inversion and that the CG7 model achieves nearly the same value of VAR as SVD17. There seems no need to use values of the empirical redundancy parameter other than $\alpha = 1$ in the CG inversion, as CG7 seems in any case to be the best model. Figs. 6.2c and d shows models obtained using CG and SVD inversions and we can see that these models, CG7 and SVD17, are similar both in elastic and anelastic model space. We can conclude from the fact that the CG and SVD models are consistent that both inversion techniques have been validated by this cross-checking process. In the following subsections, we discuss several statistical aspects of the CG inversion results.

6.2.3 Statistical estimation

From eq. (6.14) we can rewrite the variance for each CG expansion coefficient a_j as follows:

$$\text{var. } (a_j) = \frac{\sigma_d^2}{\mathbf{p}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_j}. \tag{6.24}$$

If we assume that σ_d is the variance of the data itself, we can estimate the nominal error bars δa_j for each of the 70 coefficients:

$$\delta a_j = \frac{\sigma_d}{\sqrt{\mathbf{p}_j^{\mathbf{T}} \mathbf{A}^{\mathbf{T}} \mathbf{A} \mathbf{p}_j}}. \tag{6.25}$$

Fig. 6.3 shows that the nominal error bars are small compared to the expansion coefficients. However, the real errors may be many times greater than the nominal error bars as was discussed above in section 5.1.3.

In order to know how the residual for each record $\delta\mathbf{d}^{(kp)}$ contributes to the model $\delta\mathbf{m}$, we examine the distribution of angles (in the multi-dimensional conjugate vector space) between the CG coefficient coefficient vector $\mathbf{a}_n$ and the expansion coefficients for each record $\mathbf{a}_n^{(kp)}$ truncated above n:

$$\cos\xi_n^{(kp)} = \frac{\mathbf{a}_n \cdot \mathbf{a}_n^{(kp)}}{|\mathbf{a}_n|\ |\mathbf{a}_n^{(kp)}|}, \tag{6.26}$$

where

$$\mathbf{a}_n = (a_1\ a_2\ \ \cdots\ a_n\ \ 0\ \ \cdots\ 0)^{\mathbf{T}} \tag{6.27}$$

and

$$\mathbf{a}_n^{(kp)} = (a_1^{(kp)}\ a_2^{(kp)}\ \ \cdots\ a_n^{(kp)}\ \ 0\ \ \cdots\ 0)^{\mathbf{T}}. \tag{6.28}$$

Similarly, we can calculate the angles between the CG coefficient vector $\mathbf{a}_n$ and the CG expansion coefficient vector $\mathbf{a}_n^k$ for the k-th event:

$$\cos\xi_n^k = \frac{\mathbf{a}_n \cdot \mathbf{a}_n^k}{|\mathbf{a}_n|\ |\mathbf{a}_n^k|}, \tag{6.29}$$

where

$$\mathbf{a}_n^k = (a_1^k\ a_2^k\ \ \cdots\ a_n^k\ \ 0\ \ \cdots\ 0)^{\mathbf{T}}, \tag{6.30}$$

and the angles between $\mathbf{a}_n^{k_1}$ and $\mathbf{a}_n^{k_2}$ for all pairs of events ($k_1 \neq k_2$):

$$\cos\xi_n^{(k_1,k_2)} = \frac{\mathbf{a}_n^{k_1} \cdot \mathbf{a}_n^{k_2}}{|\mathbf{a}_n^{k_1}|\ |\mathbf{a}_n^{k_2}|}. \tag{6.31}$$

We can also calculate the angles between the CG coefficient vector $\mathbf{a}_n$ and the CG expansion coefficient vector $\mathbf{a}_n^p$ for the p-th station truncated above n:

$$\cos \xi_n^p = \frac{\mathbf{a}_n \cdot \mathbf{a}_n^p}{|\mathbf{a}_n| \ |\mathbf{a}_n^p|}, \tag{6.32}$$

where

$$\mathbf{a}_n^p = (a_1^p \ a_2^p \ \ \cdots \ a_n^p \ \ 0 \ \ \cdots \ 0)^{\mathbf{T}}, \tag{6.33}$$

and the angles between $\mathbf{a}_n^{p_1}$ and $\mathbf{a}_n^{p_2}$ for all pairs of stations ($p_1 \neq p_2$):

$$\cos \xi_n^{(p_1,p_2)} = \frac{\mathbf{a}_n^{p_1} \cdot \mathbf{a}_n^{p_2}}{|\mathbf{a}_n^{p_1}| \ |\mathbf{a}_n^{p_2}|}. \tag{6.34}$$

For the preferred models CG7 and SVD17, we plot, in Fig. 6.4, (a) $\xi_7^{(kp)}$ and (b) $\theta_{17}^{(kp)}$ for each record; (c) ξ_7^k, $\xi_7^{(k_1,k_2)}$ and (d) θ_{17}^k, $\theta_{17}^{(k_1,k_2)}$ for each event (black bars); (e) ξ_7^p, $\xi_7^{(p_1,p_2)}$ and (f) θ_{17}^p, $\theta_{17}^{(p_1,p_2)}$ for each station. From Fig. 6.4, we can see that no single individual datum $\delta\mathbf{d}^{(kp)}$ has sufficient power to determine the structure by itself. As we compare the results for CG7 and SVD17, we can find that $\xi^{(kp)}$ are slightly smaller than $\theta^{(kp)}$, as $\xi^{(kp)}$ has a peak at an angle of 87° while $\theta^{(kp)}$ has its peak at an angle of 89° although $\xi^{(kp)}$ has some elements over 90°. For each event or for each record, we find that CG7 has somewhat more individual elements close to the final model than SVD17.

6.2.4 Projection of CG vectors onto the SVD space

In order to evaluate the CG vectors, we project each CG basis vector and each model vector obtained by the CG inversion onto the orthogonal SVD space, which is independent of the data residual vector $\delta\mathbf{d}$. We define the projection vector $\mathbf{x}_{\mathrm{SVD}}$ of an arbitrary vector $\mathbf{x}$ onto the SVD space as follows:

$$\mathbf{x}_{(\mathrm{SVD})} = \mathbf{V}^{\mathbf{T}}\mathbf{x} = (\mathbf{v}_1^{\mathbf{T}}\mathbf{x}, \mathbf{v}_2^{\mathbf{T}}\mathbf{x}, \ \ \cdots, \ \ \mathbf{v}_M^{\mathbf{T}}\mathbf{x})^{\mathbf{T}}. \tag{6.35}$$

In Fig. 6.5, we plot $\mathbf{p}_{j(\mathrm{SVD})}$ for $j = 1, \cdots, 7$. As we can see, all the conjugate vectors consist mainly of combinations of the first 17 SVD vectors (the grey bars denote SVD indices from 1 to 17). Because the residual vector has strong information on the structure such as $\mathbf{v}_{17}$, as we see in the variance reduction curve in Fig. 6.2b, $\mathbf{p}_j$ for low j values correspond quickly to those components of the structure and show small expansion coefficients for the higher-order SVD vectors.

We plot the projection of the CG7 model $\delta\mathbf{m}_7$ onto the SVD space. We plot $\mathbf{e}$ as grey bars, and plot the red bar in Fig. 6.6, the projection of $\delta\mathbf{m}_{7(\mathrm{SVD})}$:

$$\delta\mathbf{m}_{n(\mathrm{SVD})} = \mathbf{V^T}\delta\mathbf{m} = \sum_{j=1}^{n} a_j \mathbf{V^T}\mathbf{p}_j. \tag{6.36}$$

Without using a large number of independent conjugate vectors, CG7 obtains basically the same features as SVD17, because the CG inversion "knows" the value of $\delta\mathbf{d}$, real data while the SVD inversion does not. As CG requires a smaller number of independent parameters, the CG solution has fewer fluctuations in the model space (especially in the elastic part of the model).

6.2.5 CG inversion results

Figs. 6.9, 6.10 show the compiled version of the inversion results both for CG and SVD and Figs.6.7 and 6.7 show variance reduction and AIC curves. As we discussed in the previous subsection, CG obtains almost the same features as SVD with smaller number of independent parameters and the CG models are smoother than SVD models. It thus seems reasonable to use CG in our future inversions. At the same time it should be stressed that without the SVD results as a standard of comparison there would be grave doubts about the reliability of the data-dependent CG results.

We show the models obtained using the CG method (only) in Figs. 6.11 and 6.12 for the Philippines Sea regions and Pacific Sea regions, respectively.

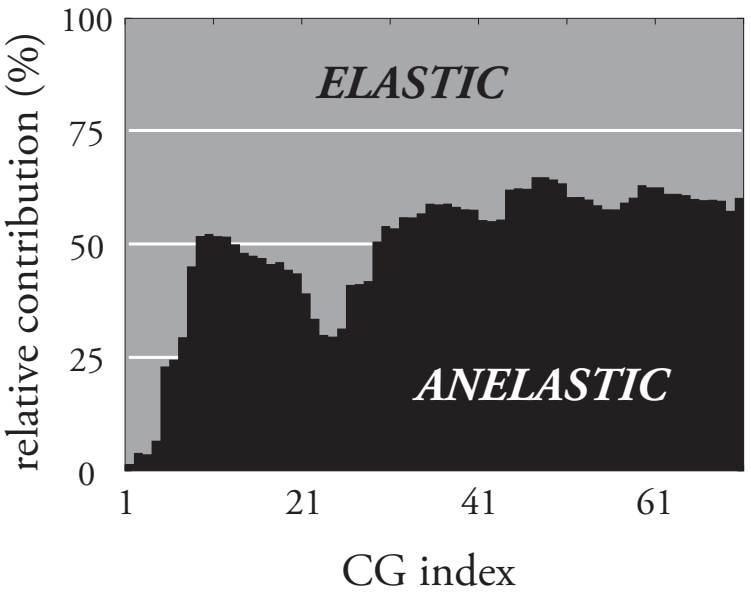


FIGURE 6.1: The relative contribution of the elastic and anelastic components of the conjugate vectors. This is an example figure for the "LARGE dataset" for the Philippines Sea West region. The CG vectors are determined by the residual vector $\mathbf{A^T}\delta\mathbf{d}$ and thus they are problem specific.

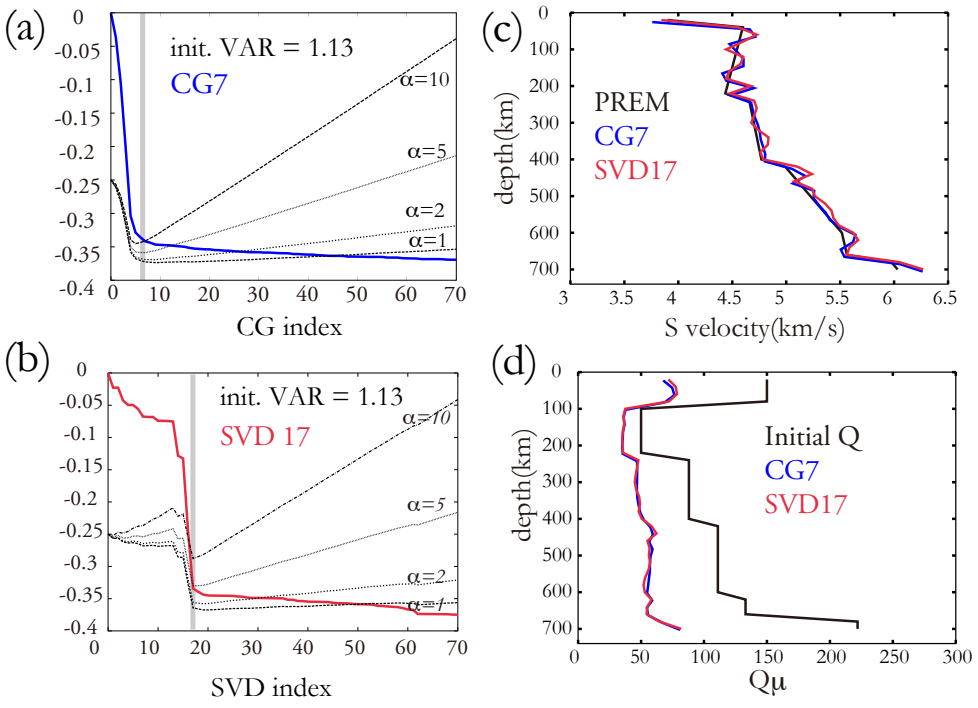


FIGURE 6.2: Results of CG/SVD inversion of the "LARGE dataset" for the Philippines Sea West region. (a) Variance for the CG inversion (blue curve) (computed using the 1st order Born approximation) and AIC_n^α values (black curves). (b) Variance for the SVD inversion (red curve) and AIC_n^α values (black curves). (c) S-velocity for PREM (black curve: the starting elastic model) and the preferred inversion results obtained using AIC analysis (blue curve for CG7 and red curve for SVD17). (d) "Initial Q_μ model (black curve: the starting anelastic model) and the preferred inversion results (blue curve for CG7 and red curve for SVD17).

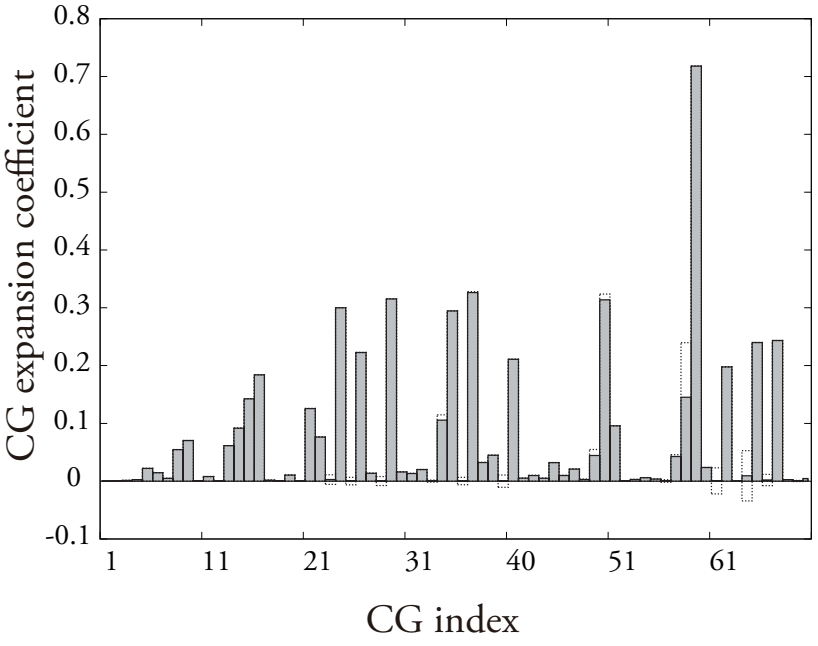


FIGURE 6.3: CG expansion coefficients a_j for the inversion for the large dataset. Gray bars show the estimated values and two dotted bars show the error bars for each coefficient.

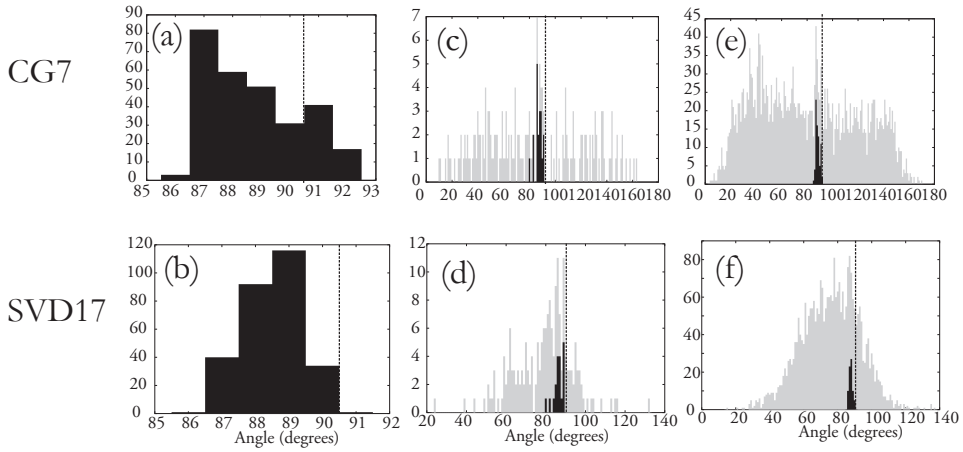


FIGURE 6.4: We plot histograms for (a) $\xi_7^{(kp)}$ and (b) $\theta_{17}^{(kp)}$ for each record; (c) ξ_7^k (black bars), $\xi_7^{(k_1,k_2)}$ (grey bars) and (d) θ_{17}^k (black bars), $\theta_{17}^{(k_1,k_2)}$ (grey bars) for each event; (e) ξ_7^p (black bars), $\xi_7^{(p_1,p_2)}$ (grey bars) and (f) θ_{17}^p (black bars), $\theta_{17}^{(p_1,p_2)}$ (grey bars) for each station. Note that in each panel, the dotted line denotes an angle of 90°.

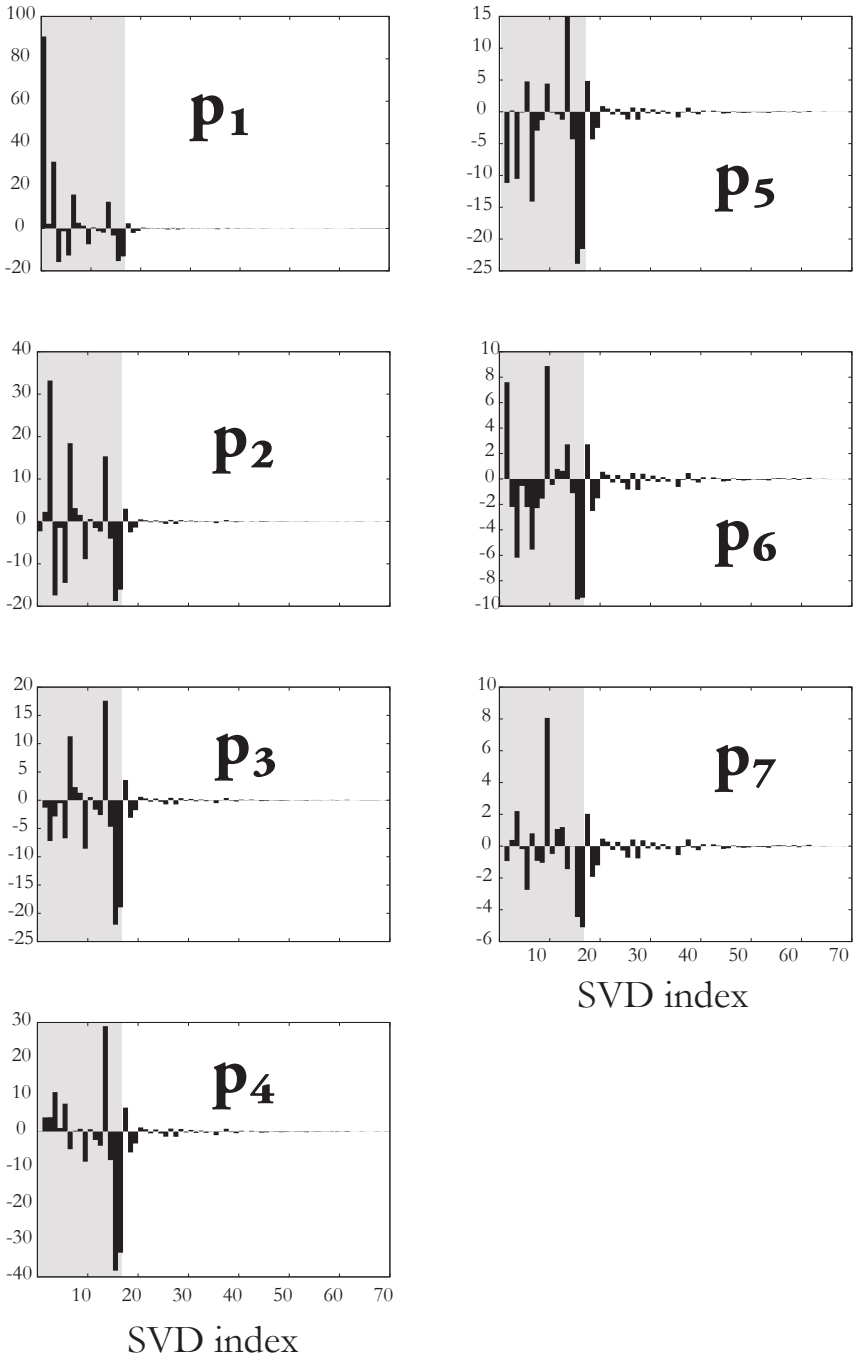


FIGURE 6.5: Projection of CG basis vectors $\mathbf{p}_{j(\mathrm{SVD})}$ $(j = 1, \cdots, 7)$ onto the SVD space. Grey bars denote the SVD index from 1 to 17 as a reference.

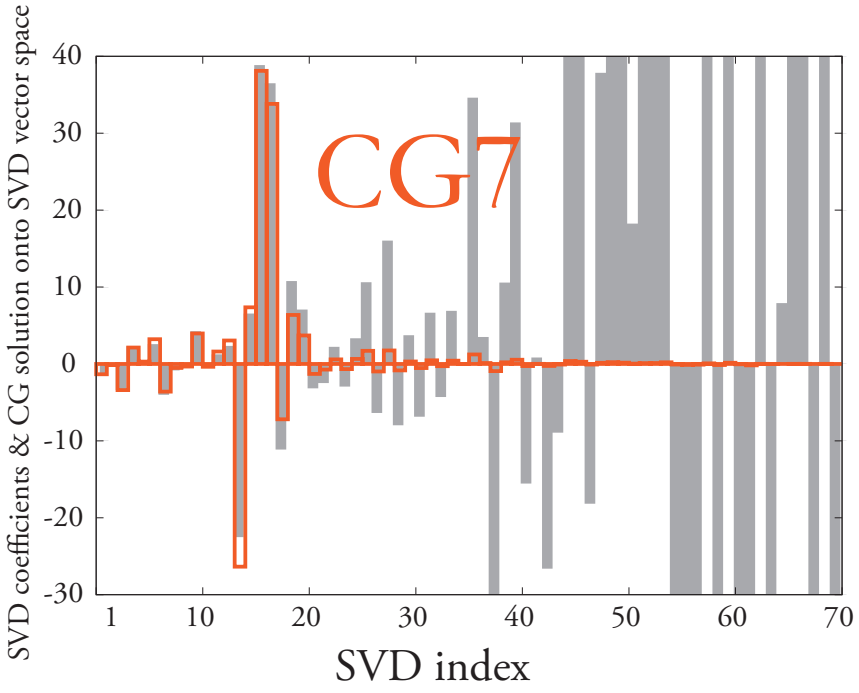


FIGURE 6.6: Projection of CG7 $\delta\mathbf{m}_{17(\mathrm{SVD})}$ onto the SVD space (red bars). Grey bars denote the SVD solution up to 70 as a reference.

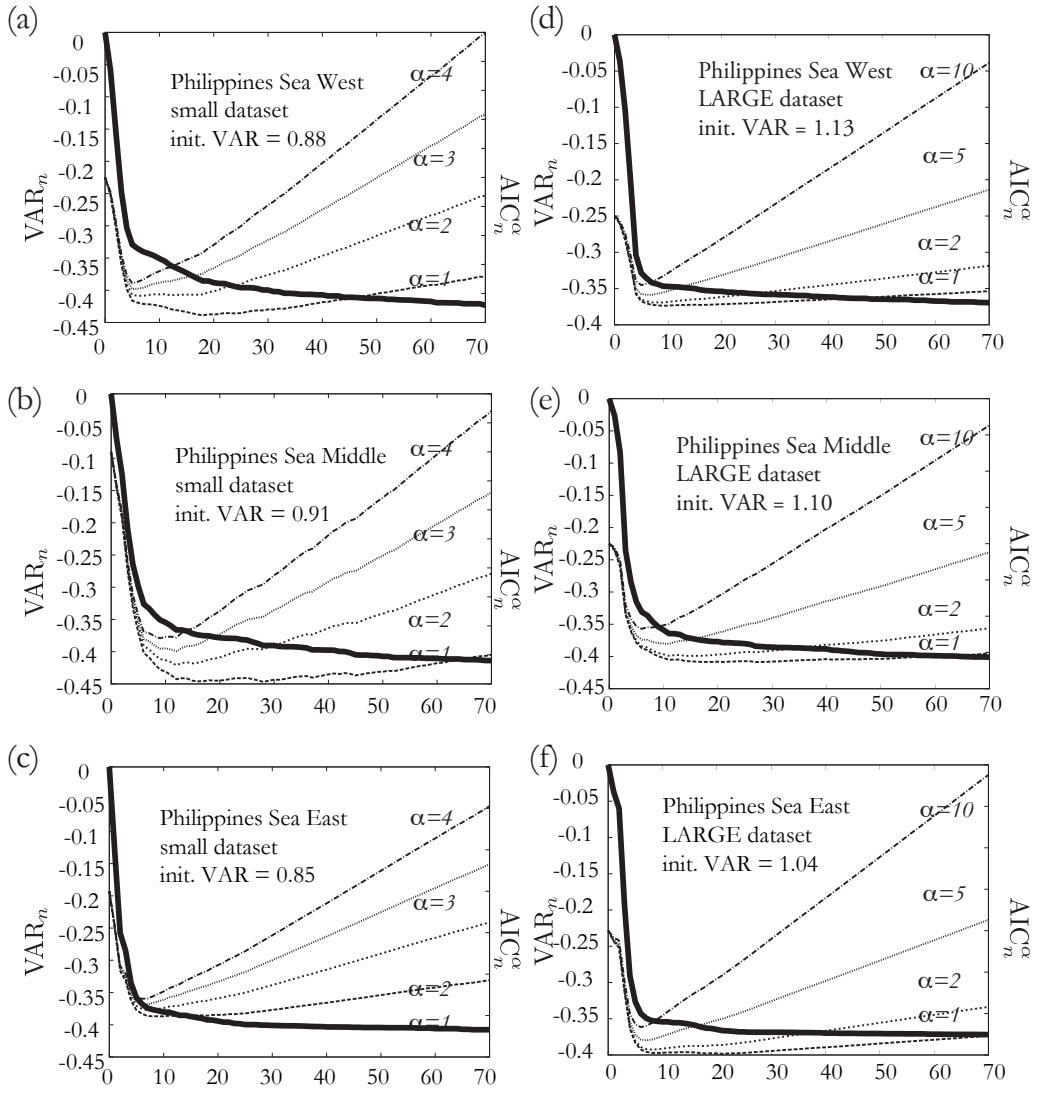


FIGURE 6.7: Variance (thick curve) and AIC_n^α values (thin curves) for the "small dataset" (a-c) and "large dataset" (d-f) for three sub-regions of the Philippines: (a and d) show the results of the Philippines Sea West Region, (b and e) show the results of the Philippines Sea Middle Region, (c and f) show the results of the Philippines Sea East Region.

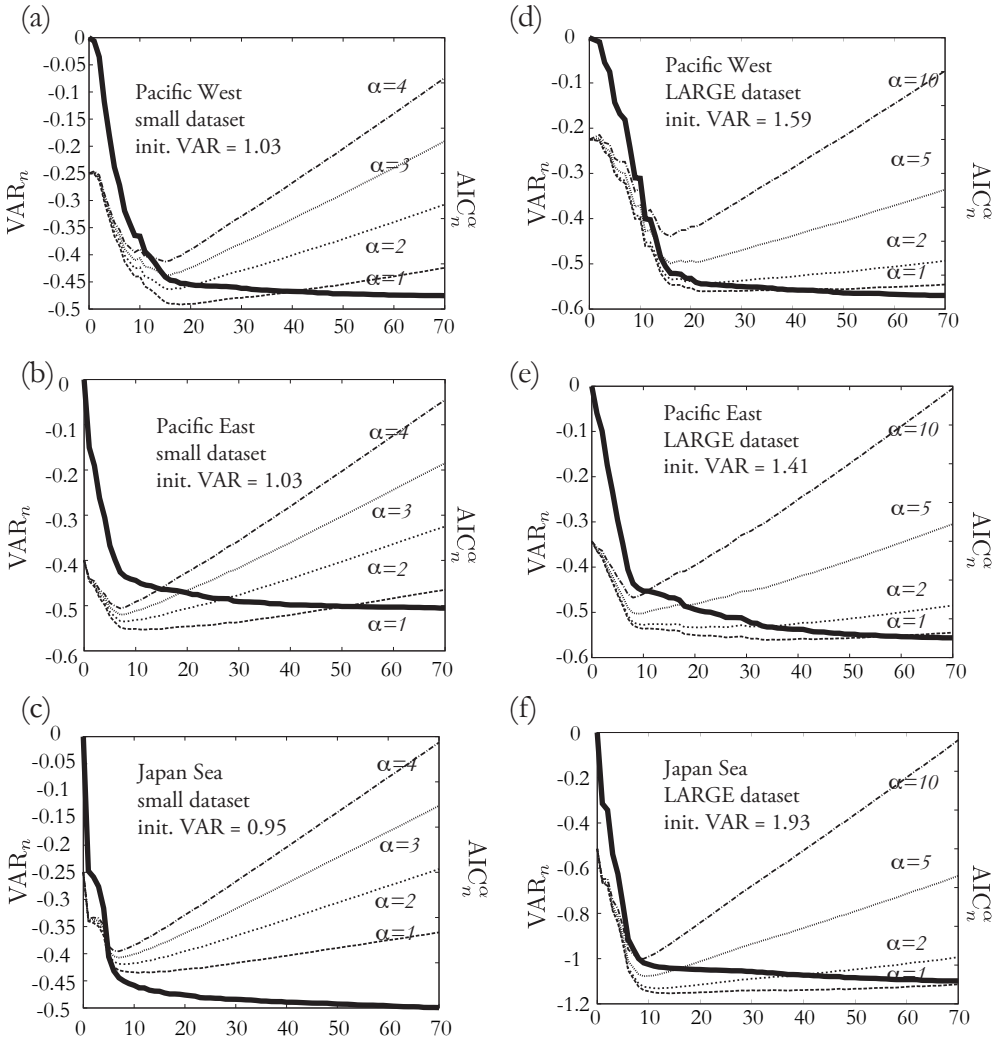


FIGURE 6.8: Variance (thick curve) and AIC_n^α values (thin curves) for the "small dataset" (a-c) and "large dataset" (d-f) for Japan Sea region and two sub-regions of the Pacific: (a and d) show the results of the Pacific West Region, (b and e) show the results of the Pacific East Region, (c and f) show the results of the Japan Sea Region.

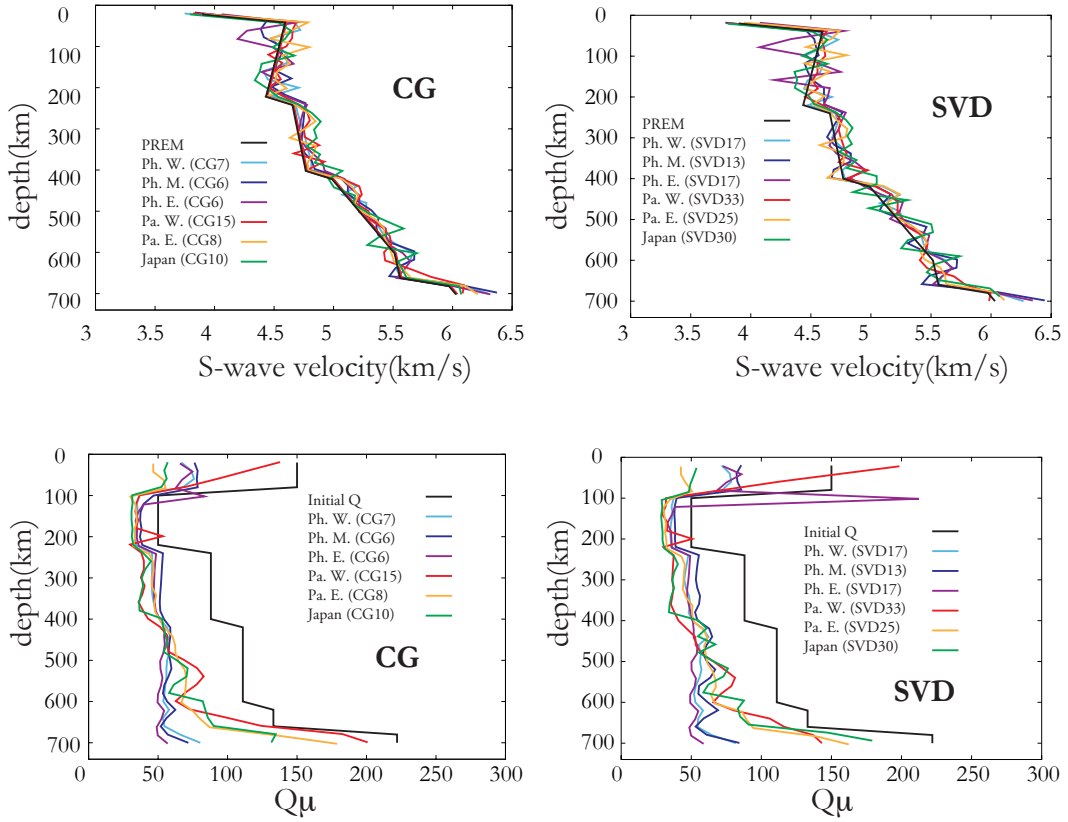


FIGURE 6.9: Compiled models for the "LARGE datasets" both for CG and SVD inversions. Black curves denotes the initial model for the inversion.

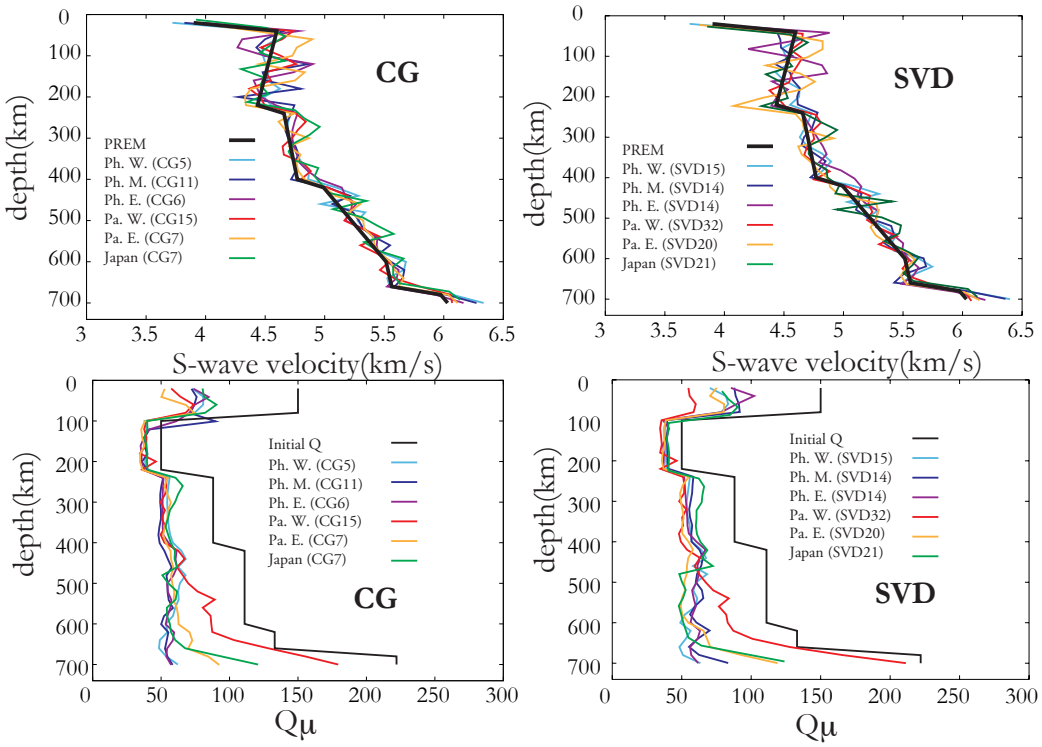


FIGURE 6.10: Compiled models for the "small datasets" both for CG and SVD inversions. Black curves denotes the initial model for the inversion.

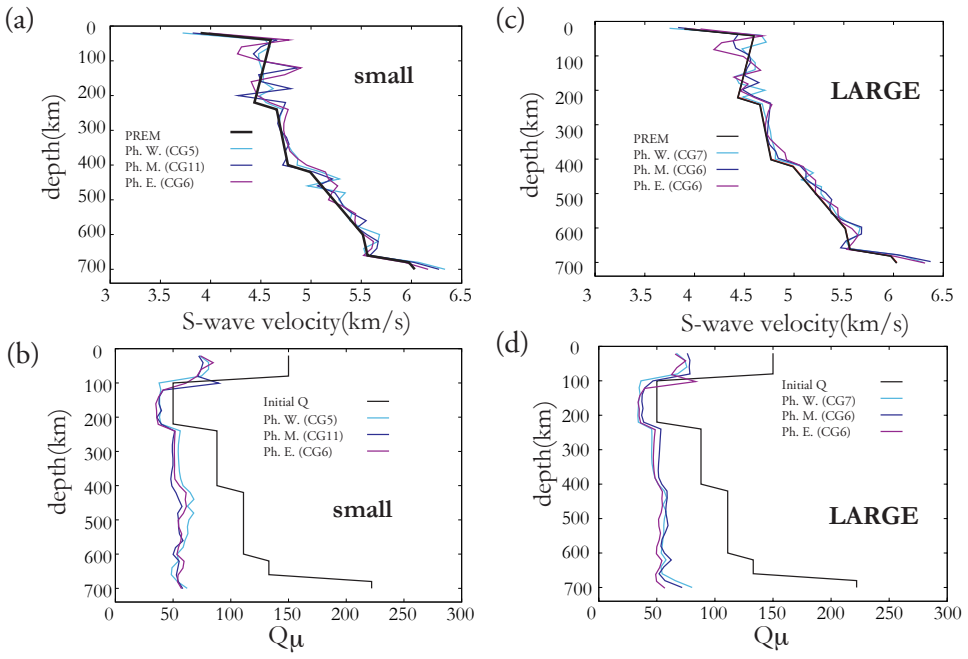


FIGURE 6.11: Results of inversion of the "small dataset" (a, b) and "large dataset" (c, d) for the three sub-regions in the Philippines. (a, c) S-velocity for PREM (black curve: the starting elastic model for the inversion), and the preferred inversion result (colored curve). (b, d) "Initial Q" model (thin curve: starting point for the inversion) and the preferred inversion result (colored curve). Sky blue curves are for the Philippines Sea West Region, blue curves for the Philippines Sea Middle Region, and purple curves for the Philippines Sea East Region.

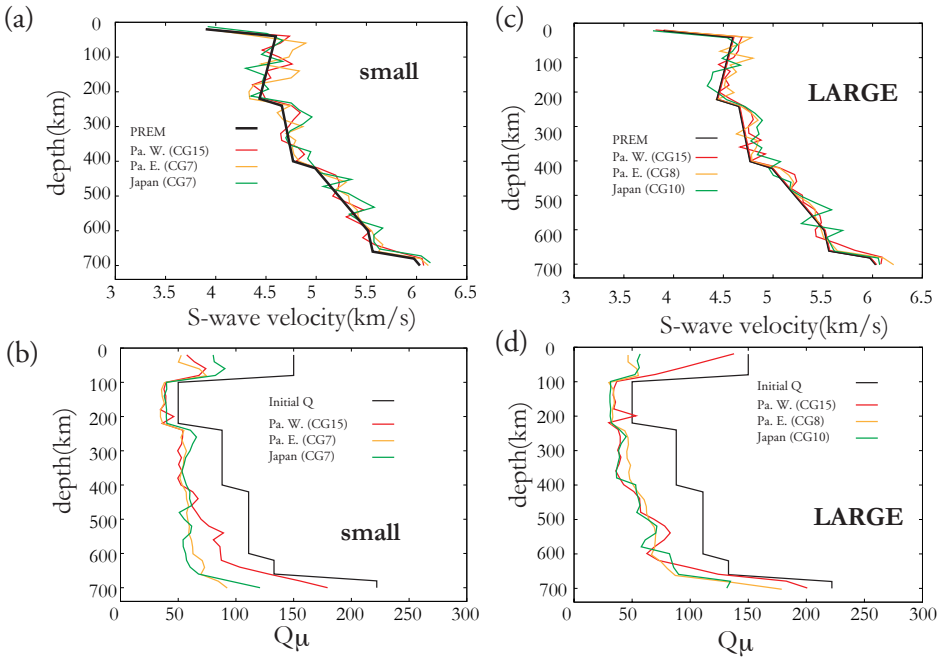


FIGURE 6.12: Results of inversion of the "small dataset" (a, b) and "large dataset" (c, d) for the two sub-regions in the Pacific and the Japan Sea sub-region. (a, c) S-velocity for PREM (black curve: the starting elastic model for the inversion), and the preferred inversion result (colored curve). (b, d) "Initial Q" model (thin curve: starting point for the inversion) and the preferred inversion result (colored curve). Red curves are for the Pacific West Region, orange curves for the Pacific East Region, and green curves for the Japan Sea Region.

Chapter 7

Discussion

In this chapter we discuss in general the methodology developed and the results obtained in this thesis. As discussed above in Chapter 6, SVD analysis using AIC as the truncation criterion seems to some extent to produce artifacts in the S-wave velocity model. We thus prefer the results obtained in Chapter 6 using CG with AIC truncation. In this chapter we discuss the robustness of our model and compare it to models obtained using surface-wave tomography. We also perform a forward-modeling experiment (up to 0.5 Hz) in order to study the validity and uncertainty of the value of $Q_\mu = 50$ obtained by our inversion which used 20 s and longer period waveforms. We find that $Q_\mu = 50$ agrees well with the long-period waveforms and also that such a low Q_μ value is consistent with the initial 50 s or so of the shorter-period waveforms.

7.1 Problems arising when using AIC as the SVD truncation criterion

If we use either SVD or CG, the model obtained when we take the sum over all of the orthogonal vectors in the basis will be equal to the least squares solution of eq. 2.23 (and thus the final variance reduction will be the same in either case). However, as shown in Chapters 5 and 6, in the actual inversions we truncate the

orthogonal vector expansion (using eq. 2.30 or 6.13, respectively). In this section, we consider the appropriateness of AIC-based truncation for SVD and CG. As a related topic, we discuss the effect of the scaling of the various model parameters, using the large dataset of the "Philippines Sea West Region" as an example.

Comparing the results in Chapter 5 and 6, we can see that CG obtains almost the same variance reduction as that obtained using SVD with a smaller number of orthogonal basis vectors than SVD. We can observe only one sharp variance drop close to the final variance reduction in the CG inversions, while we typically see two kinks in the variance curve in the SVD inversions. This is because CG uses conjugate vectors starting from the conjugate vector parallel to the residual vector $\mathbf{A^T}\delta\mathbf{d}$, while SVD uses eigenvectors starting from the eigenvector whose eigenvalue is the largest. The eigenvalues in the SVD method are determined by decomposing $\mathbf{A^T A}$ using its eigenvectors, regardless of the extent to which each eigenvector is parallel to, or orthogonal to, the residual vector $\mathbf{A^T}\delta\mathbf{d}$. In most cases some of the eigenvectors that correspond to the largest eigenvalues will be nearly perpendicular to the residual vector $\mathbf{A^T}\delta\mathbf{d}$, and thus the SVD inversion will be inefficient in reducing the variance. Therefore in general we have to use a larger number of model parameters in the SVD basis than in the CG basis to achieve the same variance reduction; also, in order to justify taking a larger number of eigenvectors than is suggested by the normal AIC using $\alpha = 1$, we had to introduce the empirical "redundancy" parameter α in deciding where to truncate the SVD expansion.

7.1.1 Scaling of model parameters

Elsewhere in this thesis we used $\Delta\mu_0 = 1$ GPa, where μ_0 ranges from 20 to 100 GPa, and $\Delta q = 1/1000$, where q is around $1/100$, in order to roughly normalize the amplitude of the various partial derivatives. In this subsection we study how the scaling of the elastic and anelastic parameters affects the inversion by testing three choices of scaling: (i) $\Delta\mu_0 = 1$ GPa and $\Delta q = 1/2000$, for which the apparent partial derivatives for the anelasticity are the smallest; (ii) $\Delta\mu_0 = 1$ GPa and

$\Delta q = 1/1000$, the values used elsewhere in this thesis; and (iii) $\Delta\mu_0 = 1$ GPa and $\Delta q = 1/500$, for which the apparent partial derivatives for the anelasticity are the strongest. We use the large dataset for the "Philippines Sea West Region" as the test case. Fig. 7.1 shows VAR and AIC curves for these three choices of scaling. Note that the variance reduction for the complete basis does not depend on the scaling or method (SVD or CG). We find that $\Delta q = 1/500$ scaling is the fastest (i.e., requires the smallest number of basis vectors) in reducing the variance. Still, the eigenvectors for the modes with the largest eigenvalues reflect primarily elastic structure when we use SVD inversion, as is clear from the relative contribution panels (see Fig. 7.2). This explains why SVD converges more slowly than CG for all choices of scaling. Fig. 7.3 shows the models obtained using the various choices of scaling. While the Q_μ structure seems stable, the S-wave velocity models fluctuate too much when we enhance the partial derivatives for elasticity even if we use CG inversion. This is because the dataset does not have as much power for elasticity as is expected. It is thus more difficult to extract a model vector nearly parallel to the residual vector $\mathbf{A^T}\delta\mathbf{d}$ if we ill-condition the matrix of the partial derivatives. Another conclusion from this experiment is that the eigenvalues of $\mathbf{A^T A}$ depend heavily on the scaling; that is, the eigenvalues do not necessarily reflect the sensitivity of the data, and hence AIC-based truncation for SVD is not appropriate.

7.2 Initial model dependence

To evaluate the dependence of the inversion on the initial model, we invert using iasp91 (Kennett and Engdahl, 1991) rather than PREM as the starting model for the "small dataset" for the Japan Sea Region. Because iasp91 is a purely elastic model we use PREM Q_μ values as the starting anelastic model. We refer to this combination as "iasp91" in the following. We calculate synthetic seismograms and partial derivatives for the "iasp91" model and first correct for the effect of local structure near the stations and sources by making static corrections (see section 3.2). Fig. 7.4 shows the results obtained for the starting "iasp91" model, and also,

for comparison, the results obtained above for the initial model of PREM and "Initial Q_μ." Although the "iasp91" model does not have a 200 km discontinuity and although the dataset has smaller resolving power at shallower depths, the preferred model seems to reproduce the 200 km discontinuity. At greater depths the consistency between inversion results for the two respective starting models improves. Although the initial "iasp91" Q_μ model has higher Q_μ values than "Initial Q_μ," the anelastic models obtained by inversions using the respective initial models seem to be generally consistent. Thus we see that the dependence on the initial model is not so great in this dataset.

7.3 Forward modeling experiments

We now present forward modeling results. We first show the data distribution in terms of cross correlation and amplitude ratio, which we plot in Fig. 3.4, taking the large dataset for the Pacific West Region as an example. Second, we show the forward modeling results up to shorter periods than used in the inversions and consider the validity of the low Q_μ values obtained by the inversion, using the Philippines event 20061212 as an example.

7.3.1 Forward modeling experiment regarding Fig. 3.4

Fig. 7.5 shows plots of the amplitude ratios and cross correlations for pairs of observed and synthetic seismograms determined from the initial model and from the modified model (CG15) for the large dataset for the Pacific West Region. From Fig. 7.5a, we can see that waveforms for the final model have smaller amplitudes than for the initial model because of the low Q_μ for the final model. As the inversion scheme attempts to reduce the variance, the cross correlation coefficients degrade. Consequently, the distribution of the data used in the inversion (Fig. 7.5a) is broadened in terms of both amplitude ratio and cross correlation and the number of waveforms which satisfy the criterion of an amplitude ratio 0.33 to 3.0

with a cross-correlation coefficient above 50 % is reduced from 501 to 454 out of the 645 data used. If we plot all of the data for the Pacific West Region (Fig. 7.5b), we can see that waveforms with a large amplitude anomaly have been improved and that the distribution has been shifted somewhat to the left. However, the number of waveforms meeting both criteria has slightly decreased (from 501 to 486 out of all the 1176 waveforms). The number of waveforms which satisfy the criterion of 0.5 to 2.0 amplitude ratio with a cross-correlation coefficient above 20 % has improved from 531 waveforms to 556 waveforms out of 1176 waveforms. We can observe that the outliers with a large amplitude ratio anomaly have been somewhat improved by the anelastic model obtained by the inversion.

7.3.2 Forward modeling of shorter-period waveforms

We perform forward modeling experiments for the Philippines event 20061212 to further evaluate the reasonableness of the Q_μ value of 50-100. Although data at epicentral distances greater than 35° were excluded in the inversion (Chapter 3), we include them here in order to examine the effect of the lower mantle, which we do not include in the inversion. We calculated synthetics up to 0.625 Hz and bandpass filter them up to 0.2 Hz for 4 models: the CG7 model for the large dataset for the Philippines West which we refer to as the "Ph.W." model in this subsection, the Q66 model, the Q100 model, and the Q150 model (Fig. 7.6). The "Ph.W." model has the lowest Q_μ (about 50 from a depth of about 200 km to the base of the MTZ), Q66 has 66, Q100 has 100, and Q150 has 150. All of the models have a Q_μ value of 312 below 700 km depth.

Fig. 7.7 shows six examples of observed and synthetic waveforms. It is interesting to compare the waveforms in the various passbands. For the longest period bandpass filtered waveforms (the leftmost column, 40-1000 s passband), the effect of varying the Q_μ value is so small that all of the Q_μ models satisfactorily fit the observations. The observed and theoretical waveforms are in reasonable agreement for the first 200 s of this passband. On the other hand, for the 20-200 s passband (center column) and 10-100 s passband (rightmost column) passbands

the theoretical waveforms for all Q_μ values have much lower amplitudes than the data for the latter part of the waveforms (after, say, the first 50 s). We interpret the later parts of the records in the higher frequency passband as noise generated by coda waves due to 3-D scattering in the Earth that is not present in the 1-D models.

Let us look at the first 50 s of the waveforms in the rightmost column (10-50 s) of Fig. 7.7, paying attention to the data at IGK (the top trace) in particular. We can see clearly that the theoretical waveforms for the 10-50 s passband for Q100 and Q150 have much larger amplitudes than the observed data in the first 50 s or so of the waveform, whereas the amplitudes for the "Ph.W." model are in good agreement with the first 50 s or so of the observed data. On the other hand, none of the synthetics, regardless of the Q_μ value, are in good agreement with the later part of the observed records, presumably because the later parts of the observed records consist primarily of coda waves which cannot be explained by any 1-D model. We can also see (in Fig. 7.8) that the $Q_\mu = 50$ model agrees better with the amplitudes of the initial part of the P-waves (right column) than the higher Q_μ models, although the differences between the various models are less pronounced than for the S-waves.

Figs. 7.9 - 7.13 show the overall results for the forward modeling experiment. The overlapped red curves show synthetics for each model. If we cannot see the black (observed) traces, this means that in general the synthetics (red traces) are in good agreement with the data. We can see a good fit at longer periods for all of the seismograms. The best fit is for "Ph.W." for the 20-200 s period waveforms. Waveforms with an epicentral distance greater than 35° cannot be well explained by any of the models. This is not too surprising in view of the fact that structure beneath the transition zone was not included in our inversions. We can also expect frequency-dependence of Q_μ and depth-dependence of absorption band (Anderson and Given, 1982).

7.4 Robustness of the models

In this section, we discuss the models and compare them with other global or regional models. The small-scale fluctuations in our S-wave velocity models seem to be due to the existence of 3-D structure that is not included in our parameterization and relatively poor resolving power for the depth dependence of the actual inversions as compared to the synthetic resolution tests. However, we can obtain sufficient information on the average velocity structure in general to suggest the existence of a positive velocity anomaly in the MTZ beneath all of the 6 subregions. Of course this feature should not necessarily be regarded as having been definitely established, due to the omission of 3-D structure in our inversions.

Figs. 6.11 and 6.12 tell that beneath all the regions except the Japan Sea region, there seems to be a larger jump at the 400 km discontinuity than PREM. A recent global model PREF (Cammarano et al., 2005a,b) suggests a possible strong jump at the 400 km discontinuity and also a large velocity gradient just a few tens of kilometers above the 400 km discontinuity. There seems to be a positive velocity anomaly beneath the Philippines Sea and Japan Sea in the MTZ (400-660 km). This is consistent with other long-period wave tomographic models (e.g., Li and Romanowicz, 1996; Mégnin and Romanowicz, 2000; Takeuchi, 2007). The magnitude of the positive anomaly is stronger beneath the Philippines Sea than the Pacific. A smaller positive velocity anomaly than the Philippines is found in the layer from 300 to 400 km depth and in the MTZ beneath the Pacific East region, which might be related to the negative P-wave velocity anomaly proposed by Zhao (2004) and Obayashi et al. (2006).

7.4.1 Anelastic structures and its possible geophysical implications

A lower Q_μ (50-100) than PREM or the "Initial Q_μ" model itself, was obtained from the inversions for the 6 subregions. There are two possible reasons for caution about these low Q_μ values: (1) the fact that we assumed the source time function is

a δ-function; (2) the fact that we ignored the effects of 3-D elastic structure, which could also lessen the amplitude of the waveforms due to scattering and de-focusing effects, which are likely to occur during travel along a slab. Because both the apparent attenuation caused by heterogeneities in the medium of propagation and the intrinsic attenuation caused by internal friction cause a propagating seismic pulse to lose energy, they are difficult to discriminate (e.g., Menke and Dubendorff, 1985; Lerche and Menke, 1986). $Q_\mu = 50$ is thus probably a lower bound on for the average value in the MTZ beneath these regions (for comparison, previous studies, such as Gung and Romanowicz 2004, suggested a Q_μ of around 100-200 for the upper mantle or the MTZ, but they also found lower Q_μ for this region than other regions). On the other hand, as discussed above, the fits to the first 50 s or so of the 10-50 s passband data in Figs. 7.7 and 7.8 support $Q_\mu = 50$.

If $Q_\mu = 50$, the partial derivative of the seismic velocity with respect to temperature is double of that of PREM ($Q_\mu = 143$) in the MTZ (e.g., Karato, 1993). Beneath the Philippines Sea region, our inversions found $Q_\mu \sim 50$ throughout the MTZ down to 660 km, while we find an increasing Q_μ in the deep regions (500-660 km) beneath the Pacific. If taken at face value this might reflect water rich content beneath the mantle wedge driven down to the lowermost part of MTZ in the Philippines Sea, which is consistent with a scenario where water was transported from the downgoing slab into the deep mantle beneath the mantle wedge (Maruyama et al., 2009; Karato, 2010).

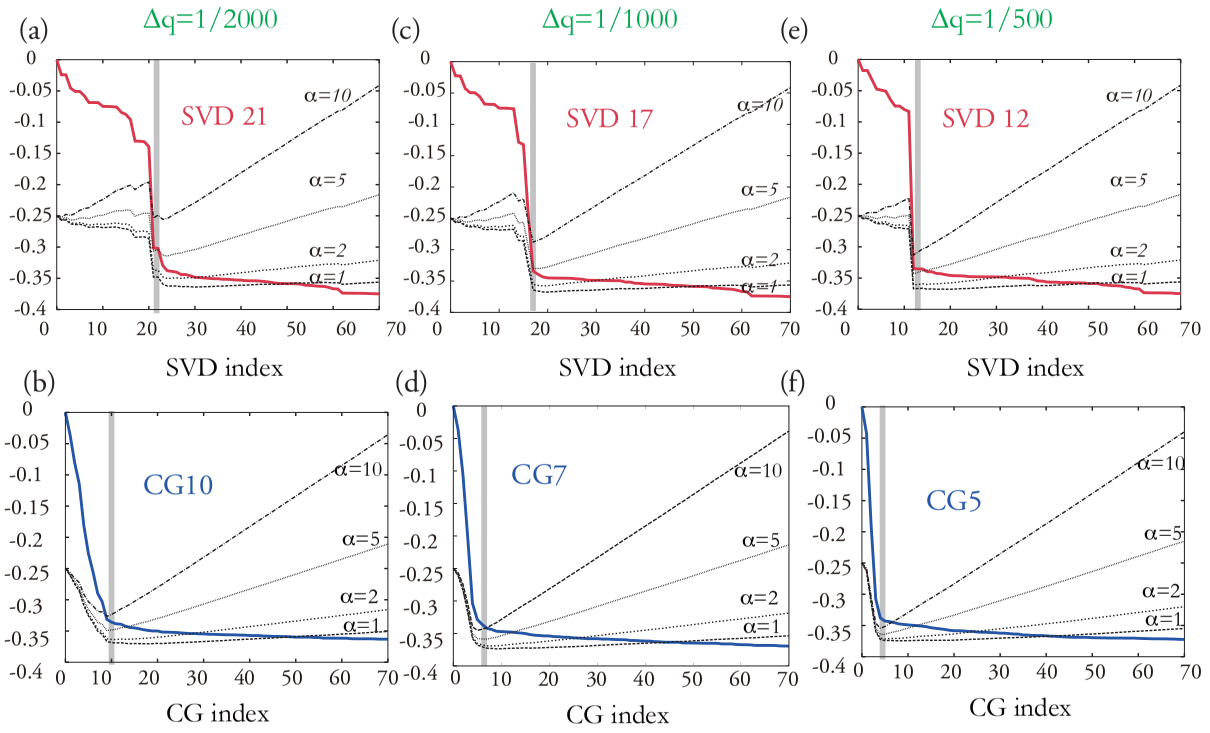


FIGURE 7.1: Variance (colored curves) and AIC_n^α values (thin curves) for inversion of the LARGE dataset for the Philippines Sea West Region for the three choices of scaling of model parameters discussed in section 7.1.1. (a) and (b) use the scaling of $\Delta q = 1/2000$; (c) and (d) use the scaling of $\Delta q = 1/1000$; (e) and (f) use the scaling of $\Delta q = 1/500$. (a), (c), (e) denote the SVD inversions, whereas (b), (d), (f) denote the CG inversions. When we use the strongest anelasticity scaling (panels e and f), the convergence of the variance is the fastest. However, SVD still requires a relatively larger number of basis vectors to reach its minimum value in all cases.

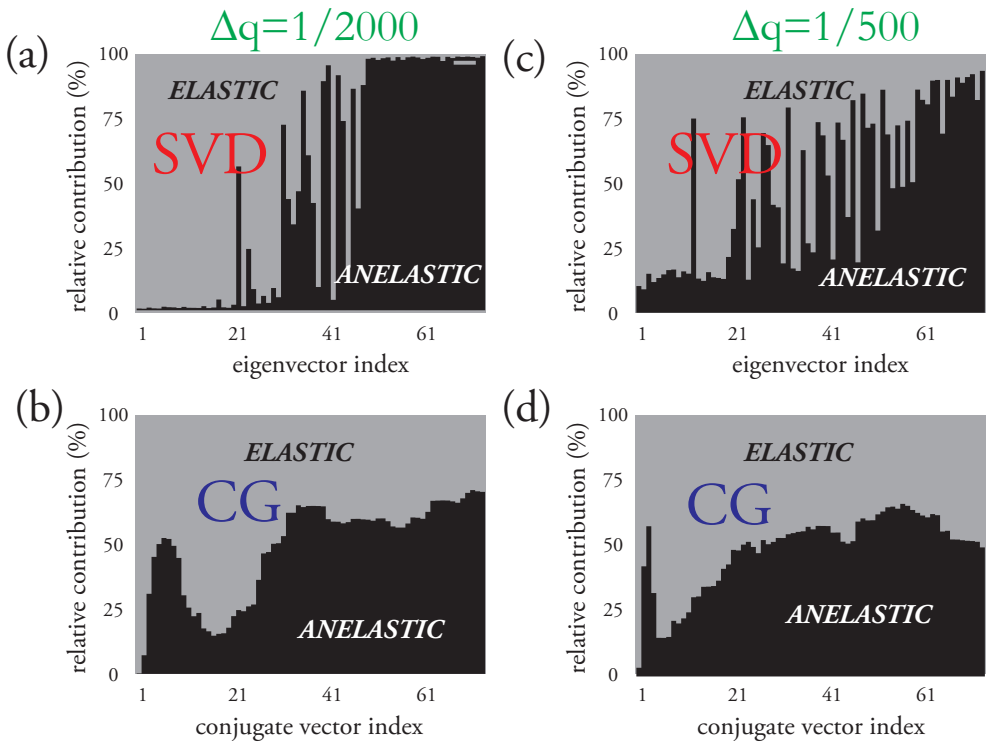


FIGURE 7.2: The relative contributions of the elastic and anelastic components of the SVD/CG vectors for the LARGE dataset for the Philippines Sea West Region. (a) and (b) use scaling of $\Delta q = 1/2000$ whereas (c) and (d) use the scaling of $\Delta q = 1/500$. The weakeer anelasticity scaling (left) reduces the anelastic contribution to the SVD basis.

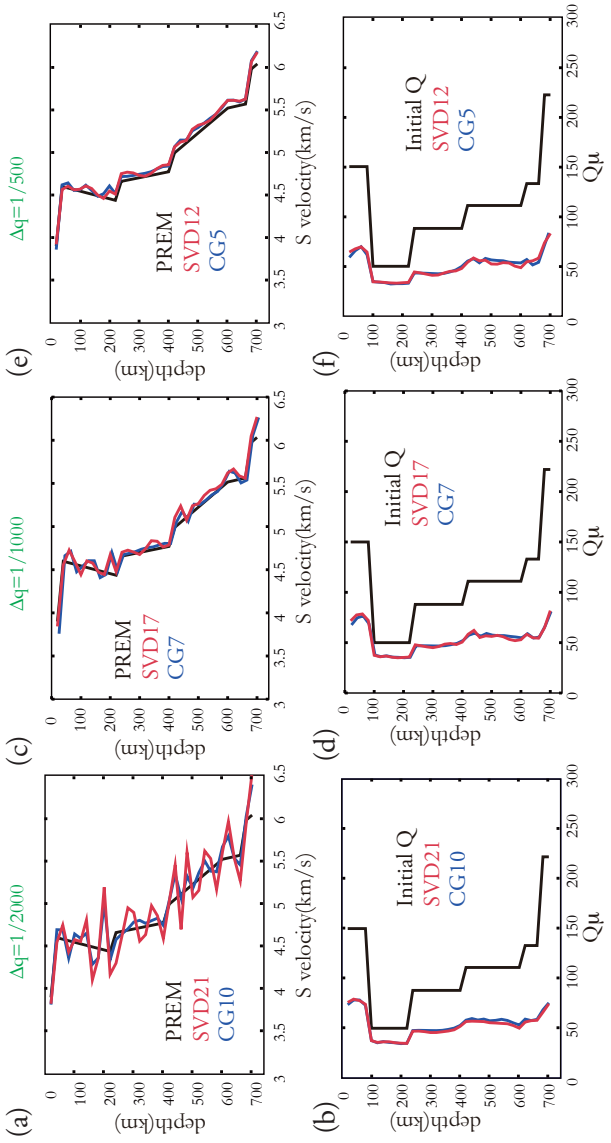


FIGURE 7.3: The models obtained for the inversion of the LARGE dataset for the Philippines Sea West Region. (a) and (b) use the scaling $\Delta q = 1/2000$; (c) and (d) use the scaling $\Delta q = 1/1000$; and (e) and (f) use the scaling $\Delta q = 1/500$. Red curves denote the models obtained using SVD inversion and blue curves denote the models obtained using CG inversion. The anelastic models are robust but we find fluctuations in S-wave velocity models when we use the weakest anelasticity scaling.

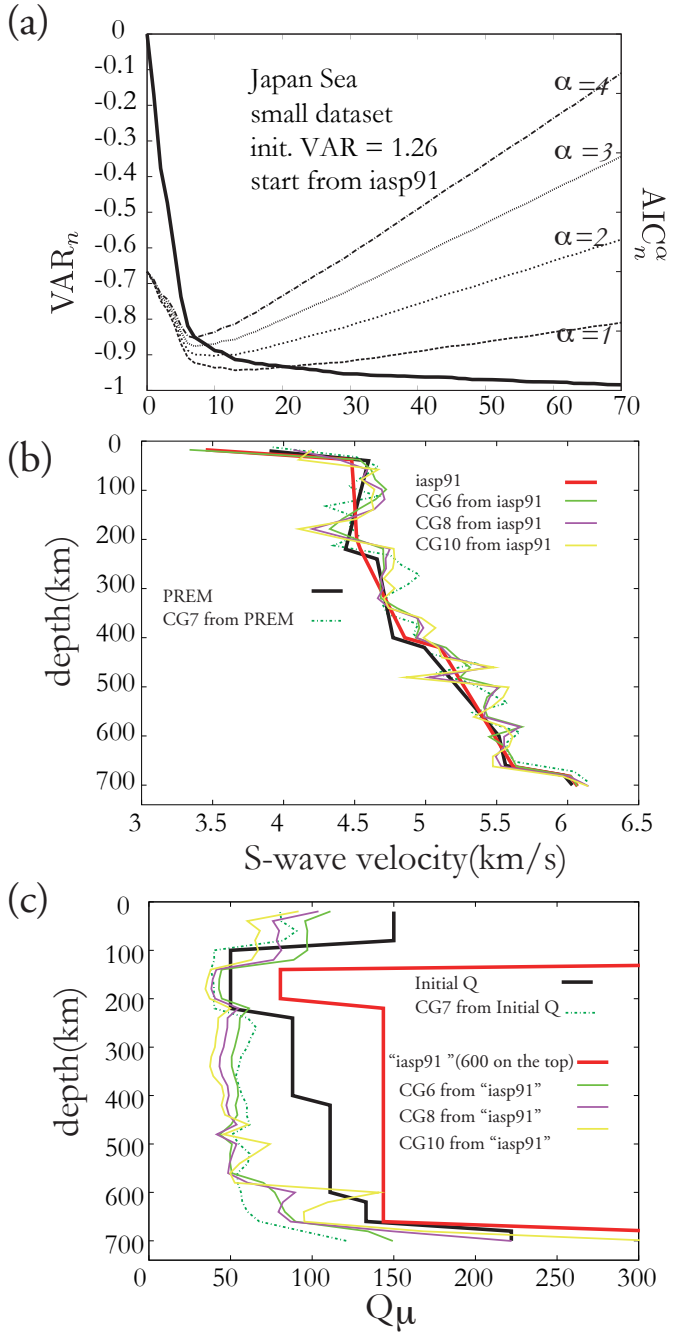


FIGURE 7.4: Results of inversion of the small dataset for the Japan Sea Region. (a) Variance (thick curve) and AIC (thin curves). (b) S-velocity for PREM (black curve) and iasp91 (red curve) as starting models, and preferred models (CG7 obtained from PREM as a starting model: dotted green curve; CG6, 8, 10 obtained from iasp91 as a starting model: dashed green, purple, yellow curve). (c) Q_μ model using "Initial Q" (black curve) and "iasp91" (red curve) as the respective starting models, and preferred models (CG7 obtained using Initial Q as a starting model: dotted green curve; CG6, 8, 10 obtained from "iasp91" as a starting model: dashed green, purple, yellow curve). Although the starting models are quite different, obtained models show consistency.

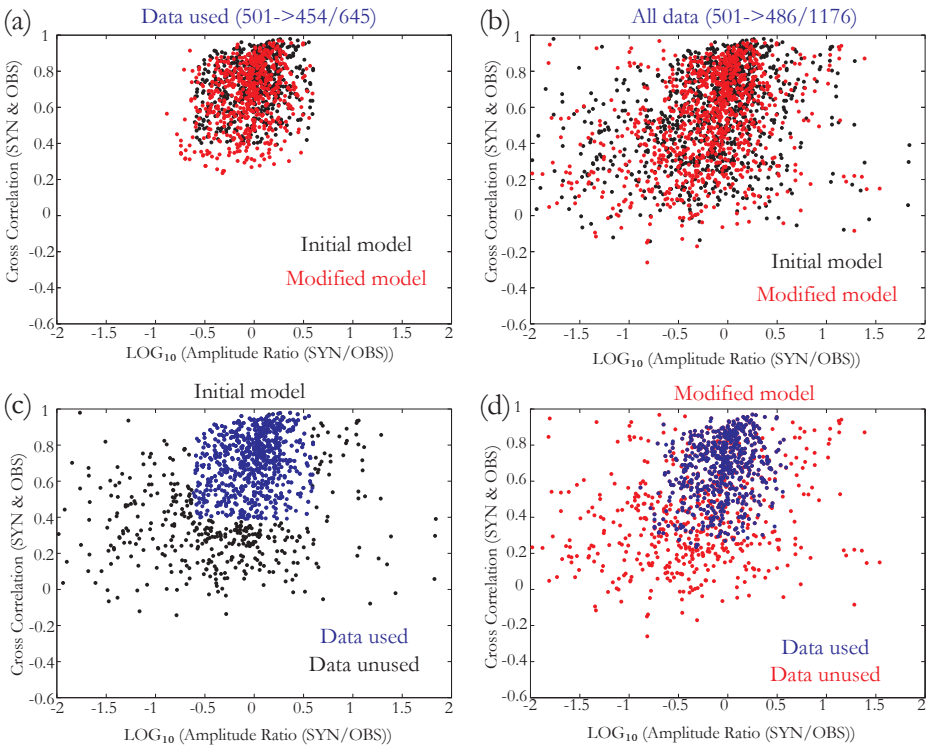


FIGURE 7.5: Plots of the amplitude ratios and cross correlations for pairs of observed and synthetic seismograms determined from the initial model and from the modified model (CG15) for the large dataset for the Pacific West Region. (a) shows the data used in the inversion and the black dots denote initial model and the red dots denote modified model. (b) shows all the data used and unused. (c) shows the distribution calculated for the initial model: blue dots denote the data used and black dots denote the data unused. (d) shows the distribution calculated for the modified model: blue dots denote the data used and red dots denote the data unused. The numbers in the parentheses in panels (a) and (b) show the number of waveforms which satisfy a criterion of 0.33 to 3.0 maximum amplitude ratio between the observed seismogram and the corresponding synthetic seismogram calculated for the the model, and also have a cross-correlation coefficient above 50 %. The second number is the total number of records in the dataset. We can see that the number in the parenthesis is not improved. However, the number of waveforms which satisfy the criterion of 0.5 to 2.0 maximum amplitude ratio with a cross-correlation coefficient above 20 % improves from 531 waveforms to 556 waveforms.

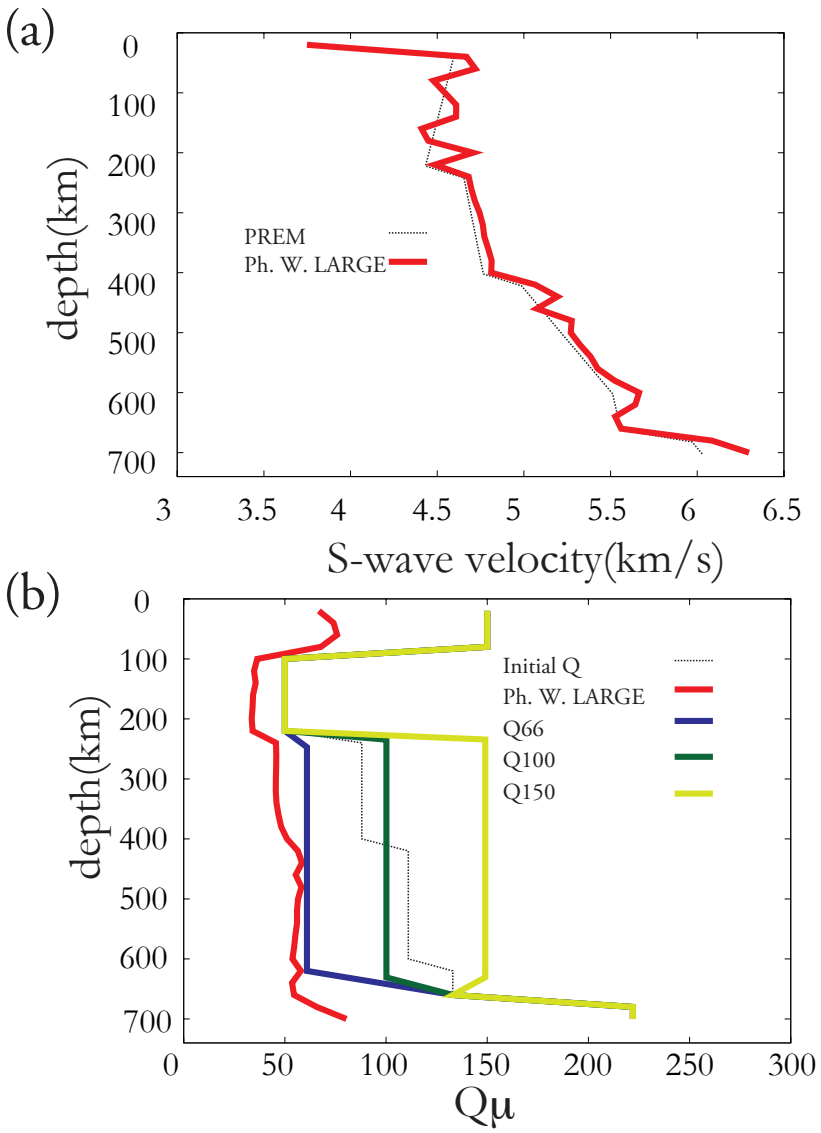


FIGURE 7.6: Models for the forward modeling experiment. Elastic structure is fixed to (a) CG7 model for the large dataset for the Philippines West. Four anelastic models are considered: (b) sky blue curve for CG7 model for the large dataset for the Philippines West, blue curve for Q66 model, green curve for Q100 model, light green curve for Q150 model.

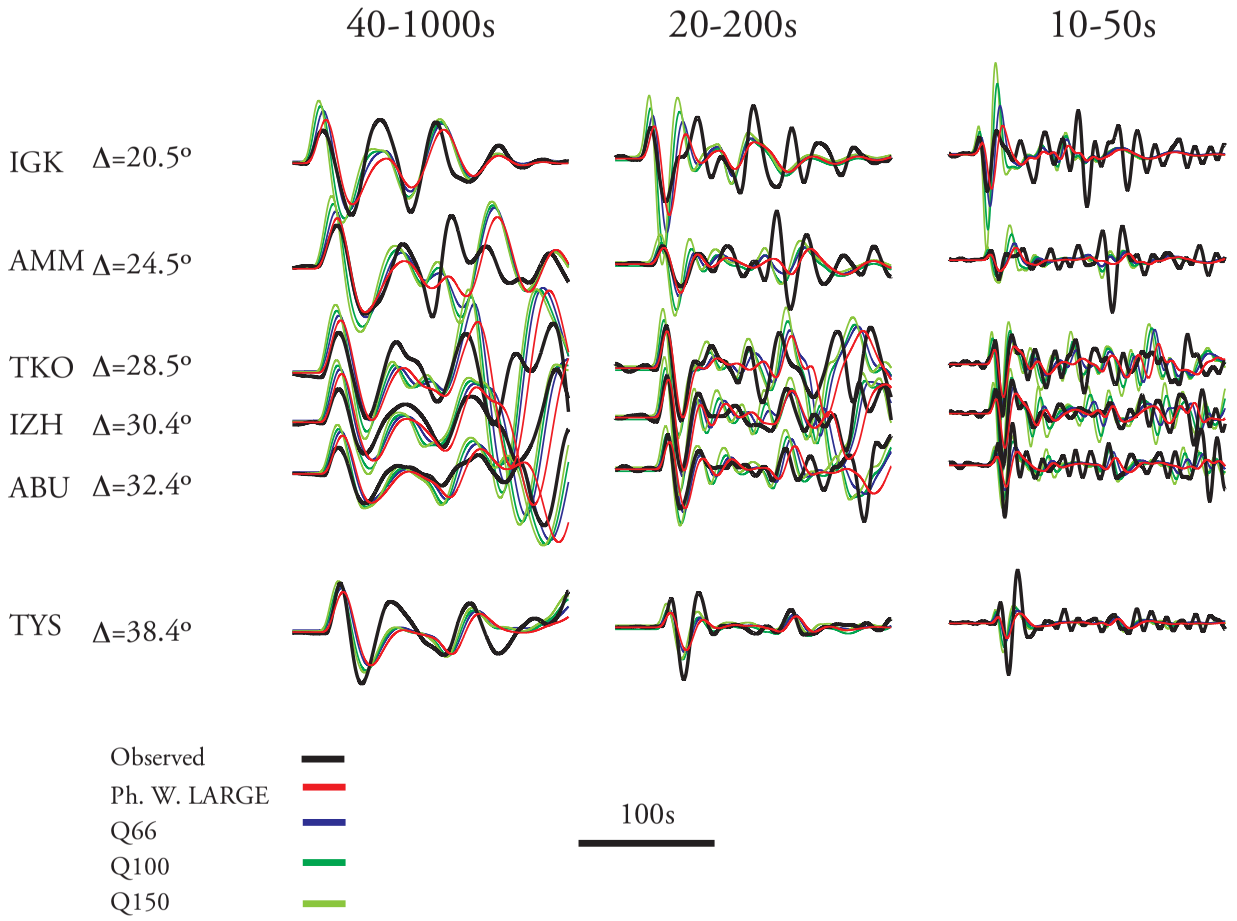


FIGURE 7.7: Six examples of transverse components of observed waveforms and synthetics calculated for "Ph.W.," Q66, Q100, Q150. Black curves show the observed data and colored curves show the synthetics for the four models. See the discussion in subsection 7.3.2 for details. The $Q_\mu = 50$ model agrees much better with the first 50 s or so of the observed data in the 10-50 s passband than the higher Q_μ models. The difference is especially pronounced at IGK (top trace). On the other hand the later part of the data in the 10-50 s passband is dominated by coda waves which cannot be explained by any of the 1-D models.

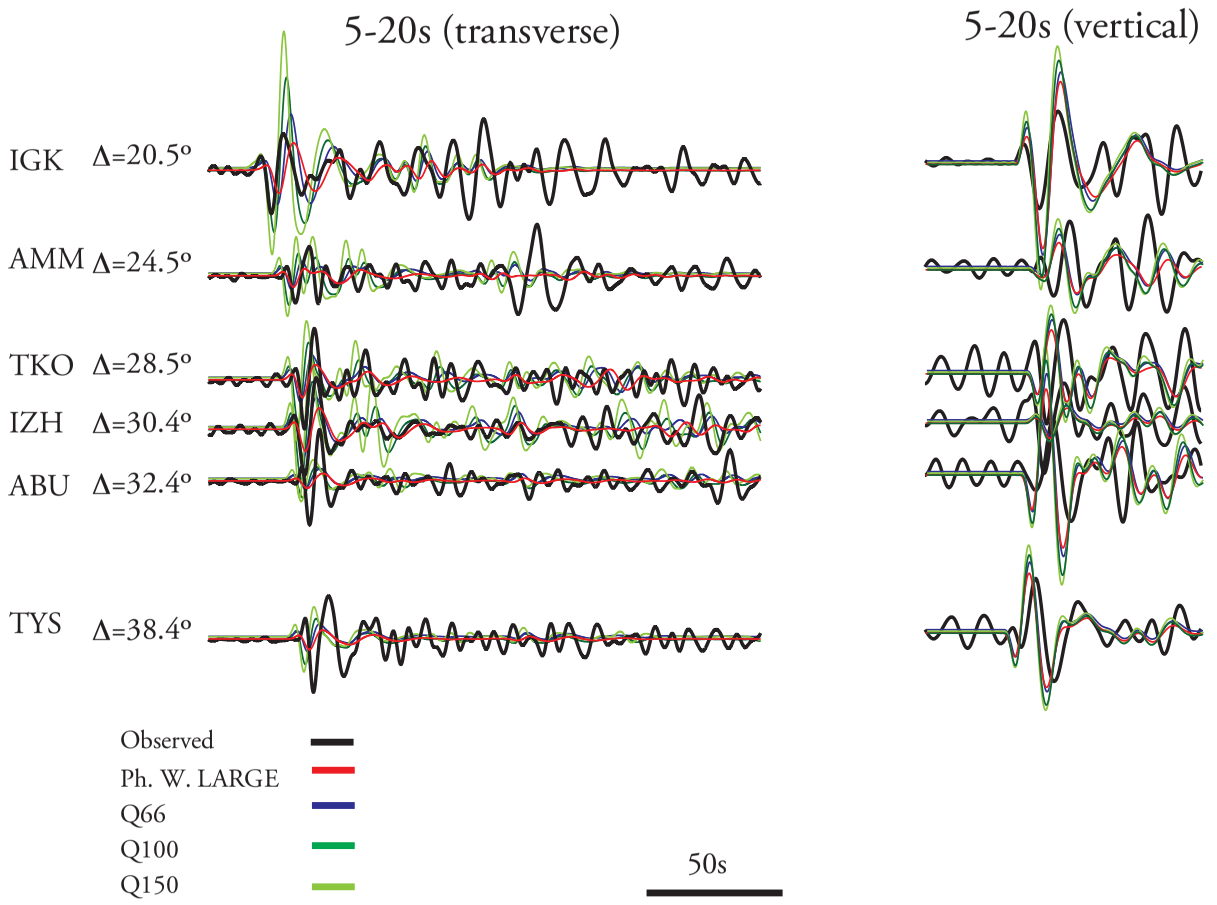


FIGURE 7.8: Six examples of transverse and vertical components of observed waveforms and synthetics calculated for "Ph.W.," Q66, Q100, Q150. Black curves for the observed data and colored curves for the synthetics for the four models. See the discussion in subsection 7.3.2 for details. Note that the time window in the left panel begins shortly before the S arrivals while that in the right panel begins shortly before the P-arrivals.

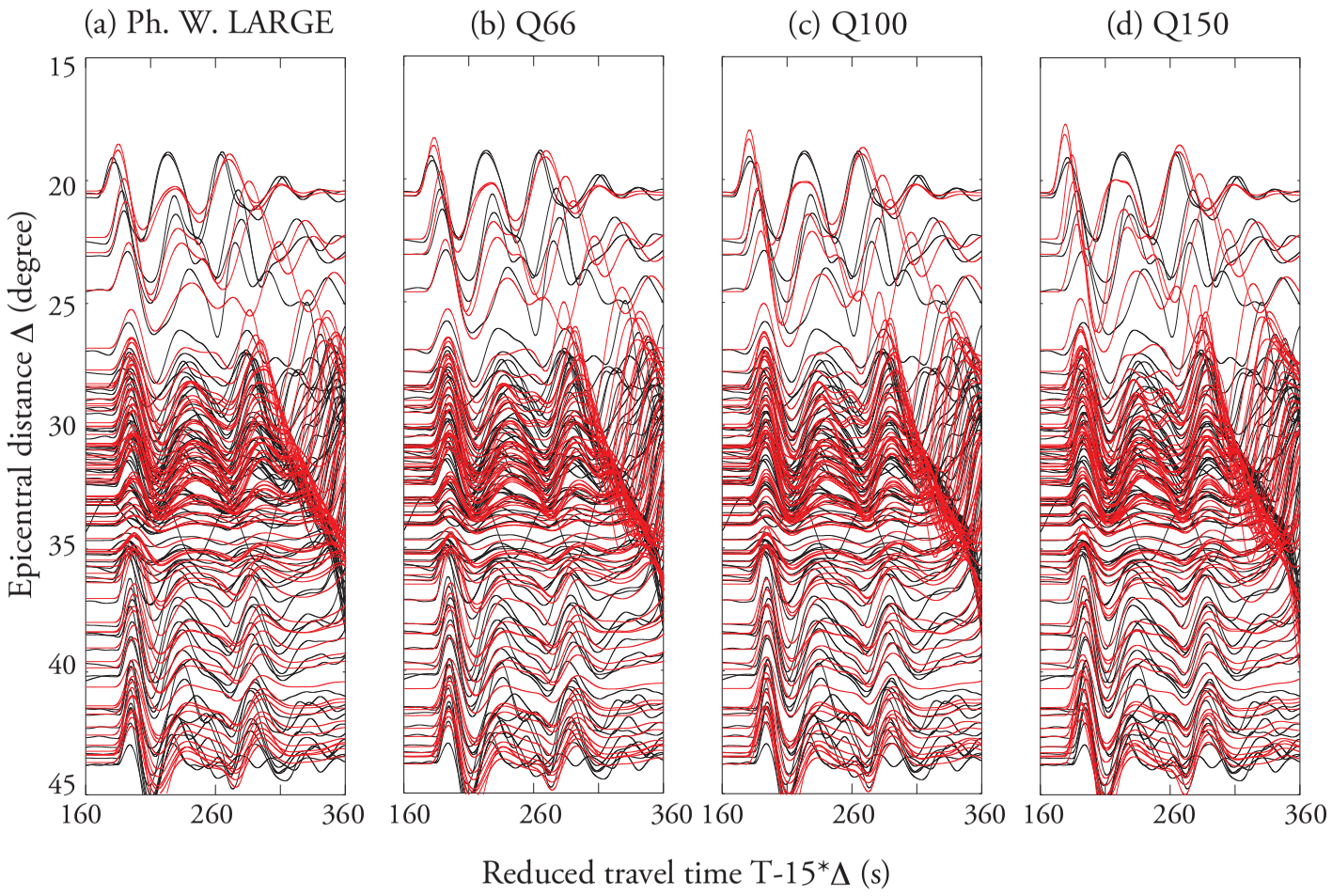


FIGURE 7.9: 40-1000 s bandpass filtered transverse section record for event 20061212. Black curves show the observed waveforms and overlapped red curves show the synthetic waveforms for (a) CG7 model for the large dataset for the Philippines West, (b) Q66 model, (c) Q100 model, (d) Q150 model. See the discussion in subsection 7.3.2 for details.

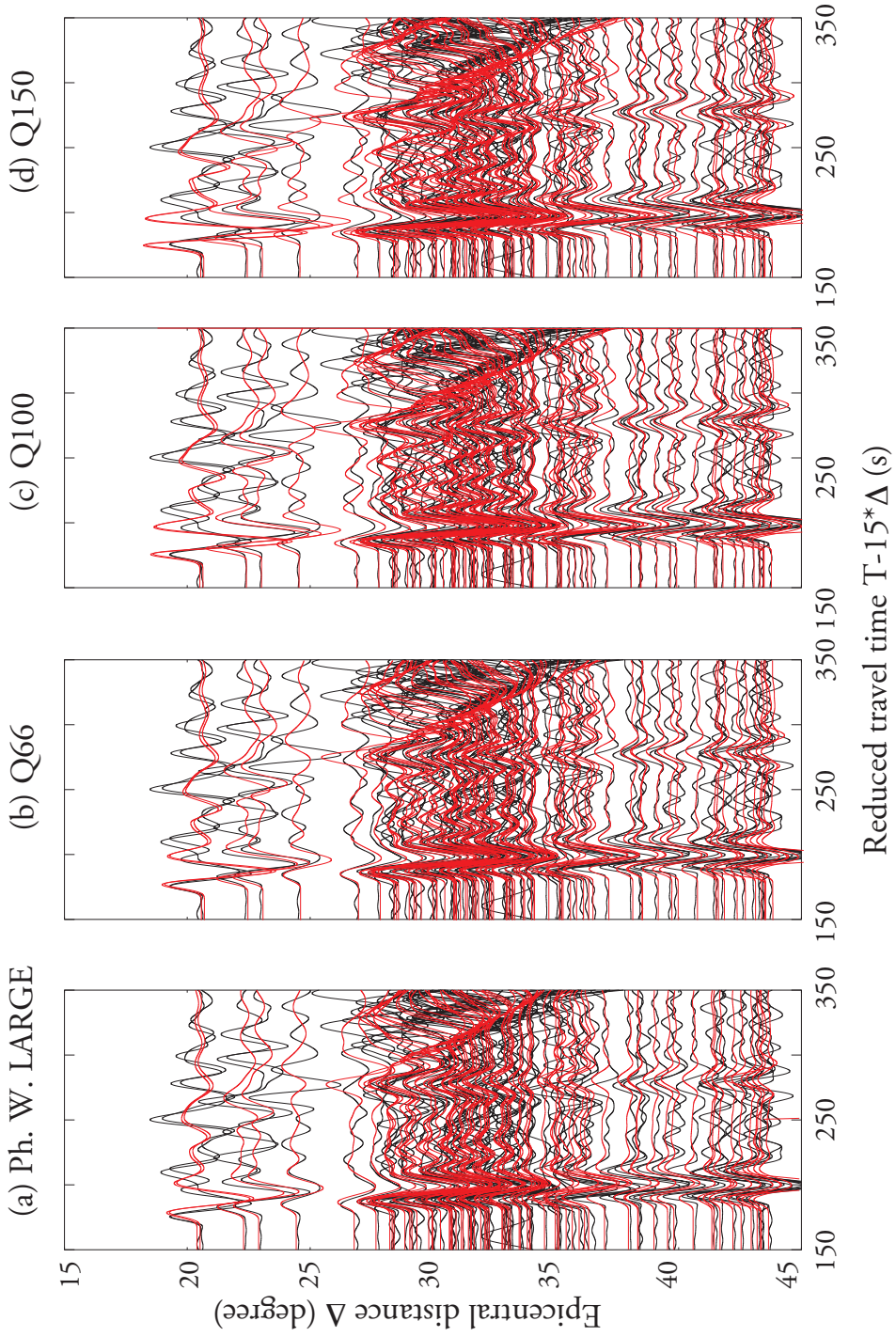


FIGURE 7.10: 20-200 s bandpass filtered transverse section record for event 20061212. Black curves show the observed waveforms and overlapped red curves show the synthetic waveforms for (a) CG7 model for the large dataset for the Philippines West, (b) Q66 model, (c) Q100 model, (d) Q150 model. See the discussion in subsection 7.3.2 for details.

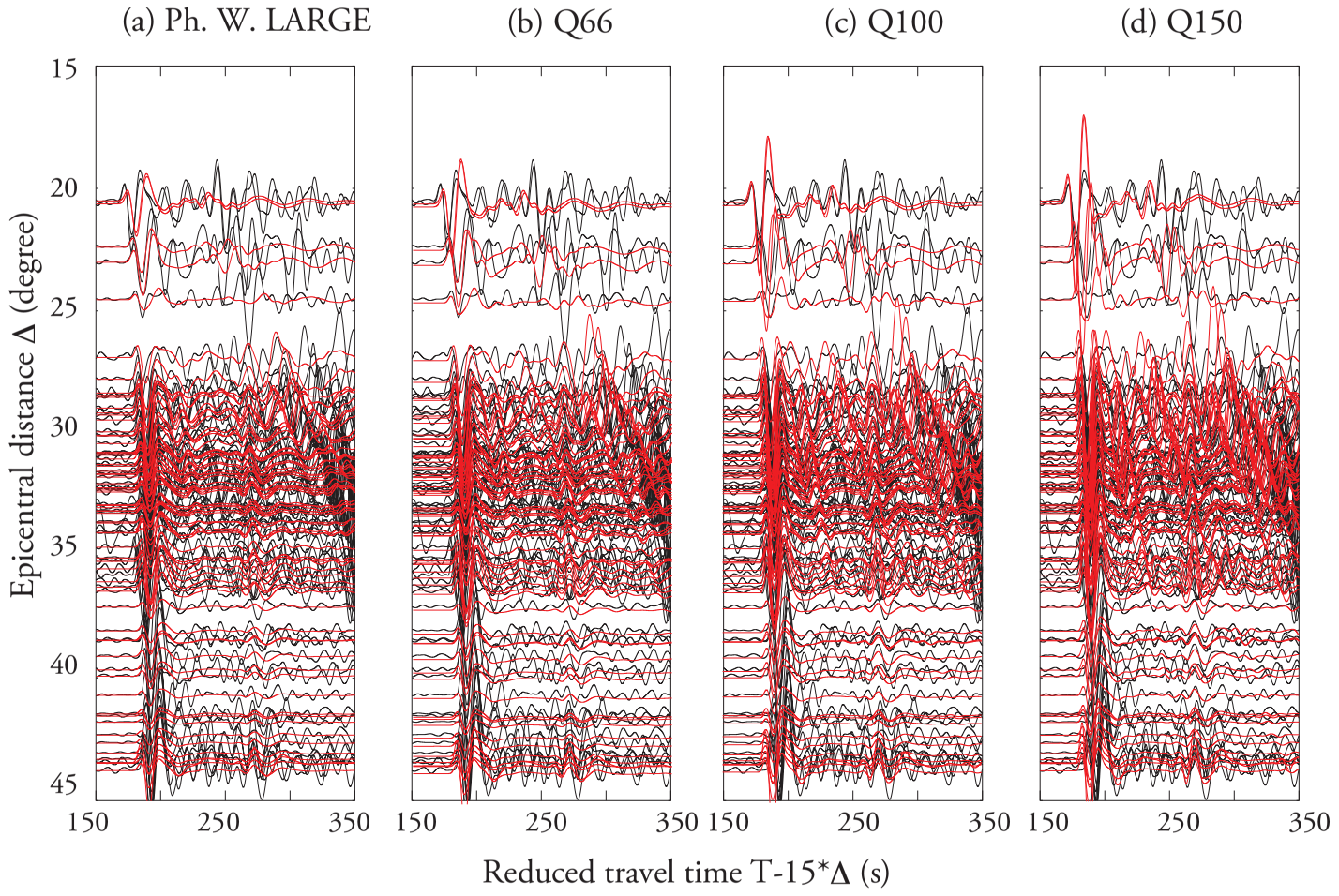


FIGURE 7.11: 10-50 s bandpass filtered transverse section record for event 20061212. Black curves show the observed waveforms and overlapped red curves show the synthetic waveforms for (a) CG7 model for the large dataset for the Philippines West, (b) Q66 model, (c) Q100 model, (d) Q150 model. See the discussion in subsection 7.3.2 for details.

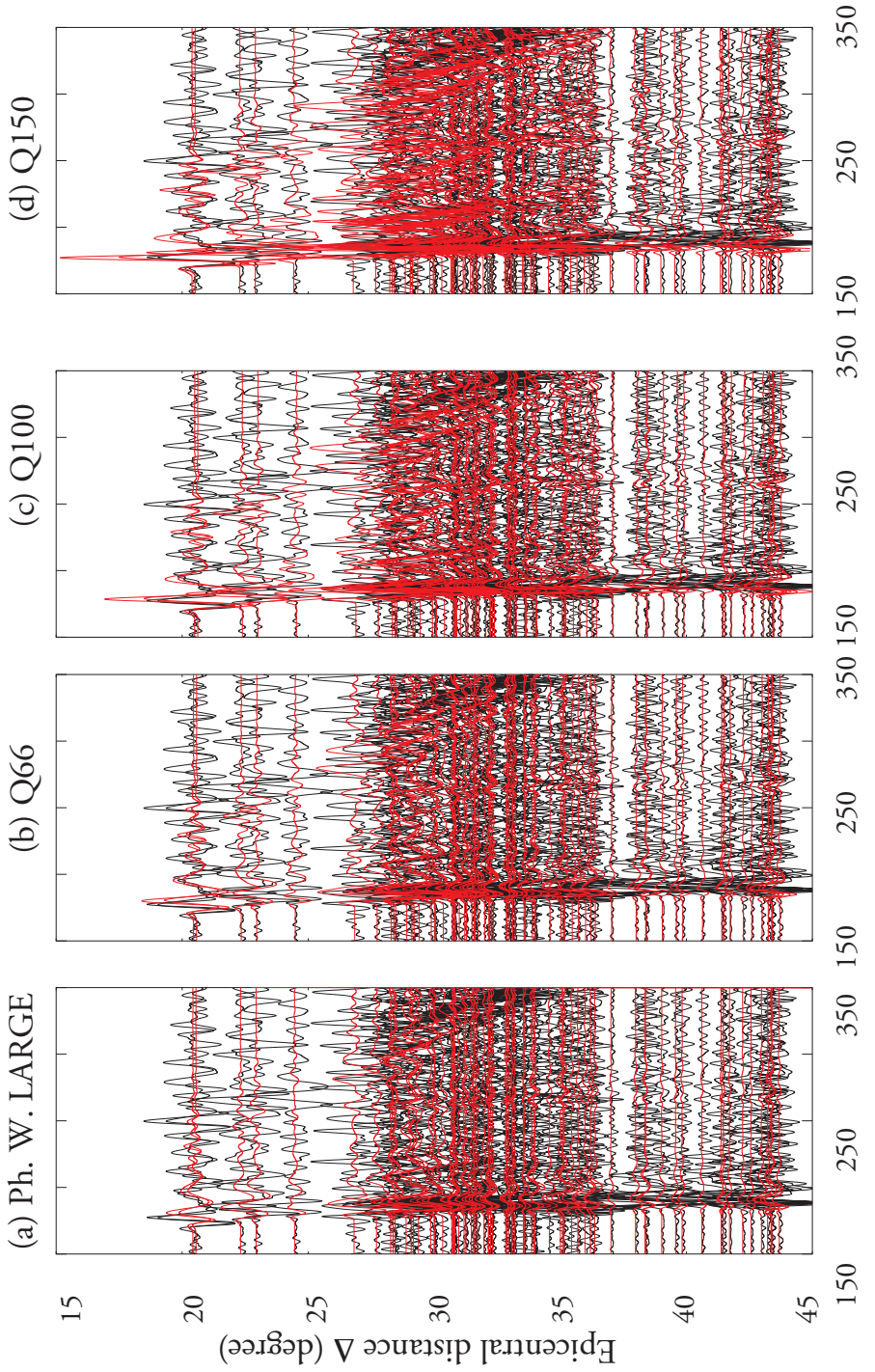


FIGURE 7.12: 5-20 s bandpass filtered transverse section record for event 20061212. Black curves show the observed waveforms and overlapped red curves show the synthetic waveforms for (a) CG7 model for the large dataset for the Philippines West, (b) Q66 model, (c) Q100 model, (d) Q150 model. See the discussion in subsection 7.3.2 for details.

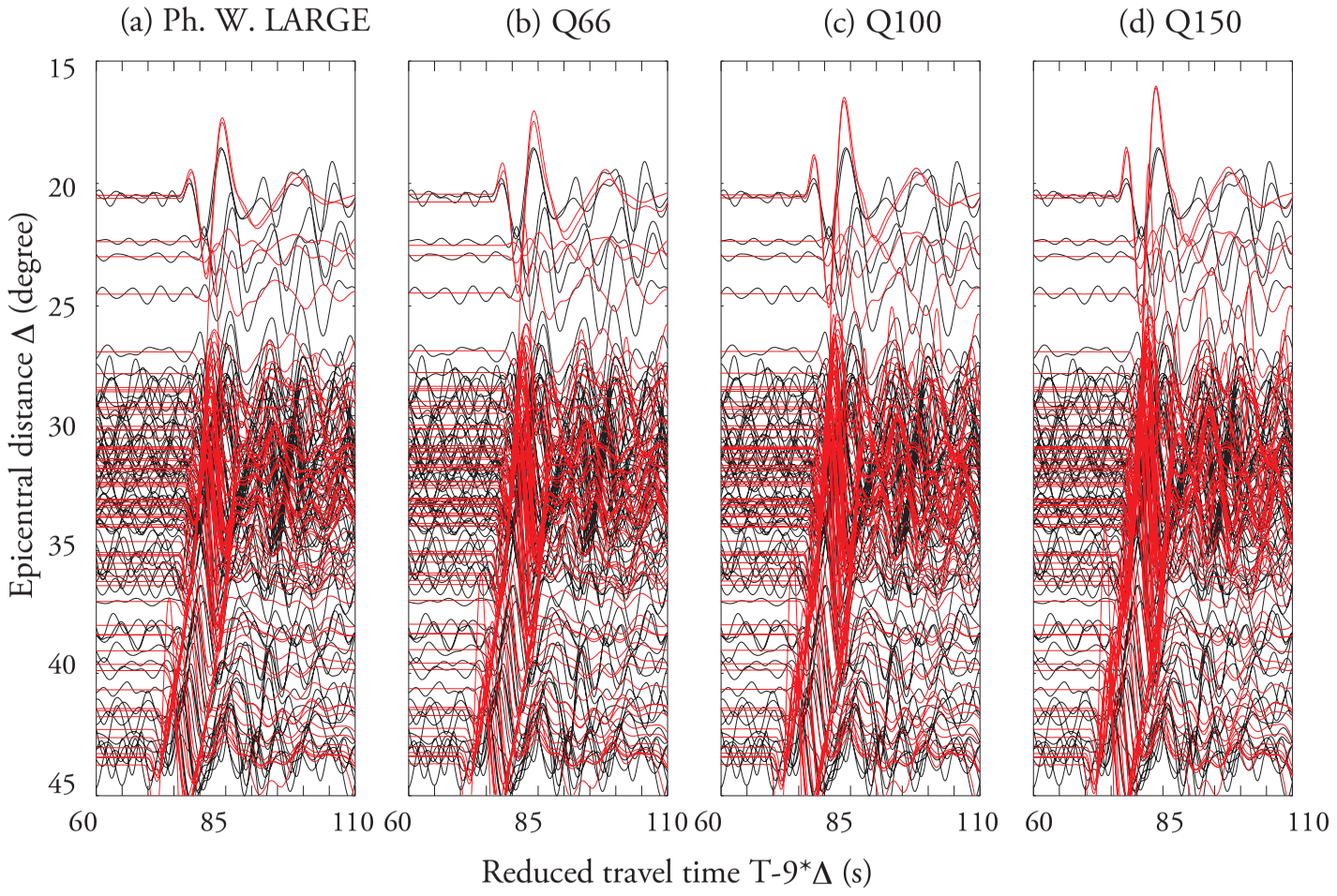


FIGURE 7.13: 5-20 s bandpass filtered vertical section record for event 20061212. Black curves show the observed waveforms and overlapped red curves show the synthetic waveforms for (a) CG7 model for the large dataset for the Philippines West, (b) Q66 model, (c) Q100 model, (d) Q150 model. See the discussion in subsection 7.3.2 for details.

Chapter 8

Conclusions

We have developed and applied methods for localized waveform inversion for elastic and anelastic structure. We use the conjugate gradient method (CG) as the basis for our inversion and use Akaike's Information Criterion (AIC) to evaluate the statistical significance of the models and determine the appropriate size of the basis. Waveform inversion has often been thought to be a black-box-like procedure, but we formally demonstrate show that the model obtained by the inversion is the projection of the data onto the model space using orthogonal vectors. Models are examined using several statistical tests. Because the CG inversion uses an orthogonal basis (and the covariance matrix is thus diagonal), we develop methodologies for evaluating the data distribution and results in the model space, in the data space, and in the intermediate orthogonal space to show the reliability and robustness of the obtained models.

As we observe a large difference between the variance reductions computed based on the Born approximation and re-computed without linear approximation, we perform one iterative inversion as an example. The second iteration produces only a small change in the model space. This is due to the fact that the data themselves have a large degree of noise due to scattering, etc. Even though the inversion process is non-linear, it appears that the noise level is sufficiently high that the first iteration obtains all information that can be extracted from the data. We also test the effect of initial model dependence which we find to be small.

The preliminary models obtained for the upper mantle and the mantle transition zone (MTZ) beneath the Northwestern Pacific show an average positive S-wave velocity anomaly in the MTZ and small Q_μ values and weak sub-regional dependence. For anelastic structure, our inversions obtained a lower Q_μ (50-100), as compared to PREM, for all of the subregions. We perform forward modeling experiments to study the validity of this model and found that our Q_μ model fit the initial part of high frequency waveforms better than models with higher Q_μ values. A weak sub-regional dependence of Q_μ structure was found, despite the large uncertainty about absolute values of Q_μ. Beneath the Philippines Sea region, we find $Q_\mu \sim 50$ throughout the MTZ down to 660 km, while we find increasing Q_μ in the deeper part of the MTZ (500-660 km) beneath the Pacific. This is consistent with a scenario where water is transported from the downgoing slab into the deep mantle beneath the mantle wedge.

Heretofore almost all studies of body-wave waveforms for regional upper mantle or transition zone structure have been based on forward modeling of individual profiles. Such studies have made many important contributions, but there are limits because the forward modeling approach cannot systematically analyze large amounts of data at one time. In contrast, statistical approaches based on waveform inversion, as shown in this thesis, can systematically analyze large volumes of data to invert not only for seismic velocities but also for anelastic structure. Even when there is considerable misfit between observed data and synthetics, the use of statistically-based waveform inversion makes it possible to obtain robust models. Further information can be expected through application of waveform inversion to 3-D Earth models. However, the forward modeling approach which was presented in Chapter 7, for example, can also help us to confirm the inversion results. Thus it seems likely that the Earth models will be updated and improved by continued progress in waveform inversion using forward modeling experiments as a validation tool.

Data and Resources

All seismograms used in this study come from the National Research Institute for Earth Science and Disaster Prevention F-net at http://www.hinet.bosai.go.jp/fnet/ (last accessed December 2008).

REFERENCES

Akaike, H., 1977. An extension of the method of maximum likelihood and the Stein's problem Annals of the Institute of Statistical Mathematics 29, 153-164.

Aki, K., Christofferssion, A., Husebye, E.S., 1977. Determination of 3-dimensional seismic structure of lithosphere. J. Geophys. Res. 82, 277-296.

Akimoto, S., and Fujisawa, H., 1968. Olivine-spinel solid solution equilibria in the system Mg_2SiO_4 – Fe_2SiO_4. J. Geophys. Res. 73, 1467-1479.

Albarède, F., van der Hilst, R.D., 2002. Zoned mantle convection. Phil. Trans. R. Soc. Lond. A 360, 2569-2592.

Anderson, D., Given, J.W., 1982. Absorption band Q model for the Earth. J. Geophys. Res. 87, 3893-3904.

Azimi, S.A., Kalinin, A.V., Kalinin, V.V., Pivovarov, B.B., 1968. Impulse and transient charactereristics of media with linear and quadratic absorption laws. Izv. Earth Physics. (English Translation) 2, 88 93.

Beckman, F.S., 1960. The solution of linear equations by the conjugate gradient method. Ralston, A. and Wilf H.S., Editors, Mathematical Methods for Digital Computers, 1, Wiley, New York, NY. pp. 62-72.

Bina, C.R., Wood, B.J., 1987. Olivine-spinel transitions: experimental and thermodynamic constraints and implications for the nature of the 400 km seismic discontinuity. J. Geophys. Res. 92. 4853-4866.

Birch, F. 1952. Elasticity and constitution of the Earth interior. J. Geophys. Res. 57, 227-286.

Bercovici, D., Karato, S., 2003. Whole-mantle convection and the transition-zone water filter. Nature 425, 39-44.

Borriello, E., Mangano, G., Marotta, A., Miele, G., Migliozzi, P., Moura, C.A., Pastor, S.,, Pisanti, O., Strolin, P., 2009. Sensitivity on earth core and mantle densities using atmospheric neutrinos. J. Cosmology and Astroparticle Phys. 6, 030.

Brudzinski, M.R., Chen, W.P., Nowack, R.L., Huang, B.S., 1997. Variations of P wave speeds in the mantle transition zone beneath the northern Philippine Sea. J. Geophys. Res. 102, 11815-11827.

Bullen, K.E., 1949. An Earth model based on a compressibility-pressure hypothesis. Monthly Notices of the Royal Astronomical Society 109, 720-720.

Cammarano, F., Goes, S., Deuss, A., Giardini, D., 2005a. Is a pyrolitic adiabatic mantle compatible with seismic data? Earth Planet. Sci. Lett. 232, 227-243.

Cammarano, F., Deuss, A., Goes, S., 2005b. One-dimensional physical reference model for the upper mantle and transition zone: combining seismic and mineral physics constraints. J. Geophys. Res. 110 (B1), B01306, doi:10.1029/2004JB003272.

Cammarano, F., Romanowicz, B., 2008. Radial profiles of seismic attenuation in the upper mantle based on physical models. Geophys. J. Int. 175, 116-134.

Cammarano, F., Romanowicz, B., Stixrude, L., Lithgow-Bertelloni, C., Xu, W., 2009. Inferring the thermochemical structure of the upper mantle from seismic data. Geophys. J. Int. 179, 1169-1185.

Christensen U.R., Yuen, D.A., 1984. The interaction of a subducting lithospheric slab with a chemical or phase-boundary. J. Geophys. Res. 89(NB6), 4389-4402.

Christensen U.R., Yuen, D.A., 1985. Layered convection induced by phase-transitions. J. Geophys. Res. 90(NB12), 291-300.

Christensen, U.R., 1996. The influence of trench migration on slab penetration into the lower mantle. Earth Planet. Sci. Lett. 140, 27-39.

Christensen, U.R., 1997. Influence of chemical buoyancy on the dynamics of slabs in the transition zone. J. Geophys. Res. 102(B10), 22435-22443.

Cizkova, H., van Hunen, J., van den Berg, A.P., Vlaar, N.J., 2002. The influence of rheological weakening and yield stress on the interaction of slabs with the 670 km discontinuity. Earth Planet. Sci. Lett. 199, 447-457.

Cummins, P.R., Geller, R.J., Hatori, T., Takeuchi, N., 1994a. DSM complete synthetic seismograms: SH, spherically symmetric, case. Geophys. Res. Lett. 21, 533-536.

Cummins, P.R., Geller, R.J., Takeuchi, N., 1994b. DSM complete synthetic seismograms: P-SV, spherically symmetric, case. Geophys. Res, Lett. 21, 1663-1666.

Cummins, P.R., Takeuchi, N., Geller, R.J., 1997. Computation of complete synthetic seismograms for laterally heterogeneous models using the Direct Solution Method. Geophys. J. Int. 130, 1-16.

Dalton, C.A., Ekström, G., Dziewonski, A.M., 2008. The global attenuation structure of the upper mantle. J. Geophys. Res. 113, B09303, doi:10.1029/2007JB005429.

Deschamps, F., Tackley, P.J., 2009. Searching for models of thermo-chemical convection that explain probabilistic tomography. II – Influence of physical and compositional parameters. Phys, Earth Planet. Inter. 176, 1-18.

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

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

Faul, U.H., Jackson, I., 2005. The seismological signature of temperature and grain size variations in the upper mantle. Earth Planet. Sci. Lett. 234, 119-134.

Fuji, N., Kawai, K., Geller, R.J., 2009. A methodology for inversion of broadband seismic waveforms for elastic and anelastic structure and its application to the mantle transition zone beneath the Northwestern Pacific. Phys. Earth Planet. Inter. doi:10.1016/j.pepi.2009.10.004.

Fukao, Y., Obayashi, M., Inoue, H., Nenbai, M., 1992. Subducting slabs stagnant in the mantle transition zone. J. Geophys. Res. 97. 4809-4822.

Gao, W., Matzel, E., Grand, S.P., 2006. Upper mantle seismic structure beneath eastern Mexico determined from P and S waveform inversion and its implications. J. Geophys. Res. 111, B08307. doi:10.1029/2006JB004304.

Garnero, E.J., Lay, T., 2003. D″ shear velocity heterogeneity, anisotropy and discontinuity structure beneath the Caribbean and Central America, Phys. Earth Planet. Inter. 140, 219-242.

Geiger, L., 1912. Probability method for the determination of earthquake epicenters from the arrival time only. Bull. St. Louis Univ. 8, 60-71.

Geller, R.J., Hara, T.,1993. Two efficient algorithms for iterative linearized inversion of seismic waveform data. Geophys. J. Int. 115, 699-710.

Geller, R.J., Ohminato, T., 1994. Computation of synthetic seismograms and their partial derivatives for heterogeneous media with arbitrary natural boundary-conditions using the Direct Solution Method. Geophys. J. Int. 116, 421-446.

Geller, R.J., Takeuchi, N., 1995. A new method for computing highly accurate DSM synthetic seismograms. Geophys. J. Int. 123, 449-470.

Gonzalez-Garcia, M.C., Halzen, F., Maltoni, M., Tanaka, H.K.M., 2008. Radiography of Earth's core and mantle with atmospheric neutrinos. Phys. Rev. Lett. 100, 061802.

Grand, S.P., Helmberger, D.V., 1984. Upper mantle shear structure beneath the Northwest Atlantic ocean. J. Geophys. Res. 89, 1465-1475.

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

Hara, T., 1997. Centroid moment tensor inversion of low-frequency seismic spectra using Green's functions for aspherical earth models. Geophys. J. Int. 130, 251-256.

Hara, T., 2004. Waveform inversion for 3-D earth structure using the Direct Solution Method implemented on vector-parallel supercomputer. Phys. Earth Planet. Inter. 146, 65-74.

Hernlund, J.W., Thomas, C., Tackley, P.J., 2005. A doubling of the post-perovskite phase boundary and structure of the Earth's lowermost mantle. Nature 434, 882-886.

Hirose, K., Fei, Y.W., Ma, Y.Z., Mao, H.K., 1999. The fate of subducted basaltic crust in the Earth's lower mantle. Nature 397, 53-56.

Hofmann, A.W., 1997. Mantle geochemistry: the message from oceanic volcanism. Nature 385, 219-229.

Irifune, T., Ringwood, A.E., 1986. Phase-transformations in a harzburgite composition to 26 GPa – Implications for dynamical behavior of the subducting slab. Earth Planet. Sci. Lett. 86, 365-376.

Ita, J., King, S.D., 1998. The influence of thermodynamic formulation on simulations of subduction zone geometry and history. Geophys. Res. Lett. 25, 1463-1466.

Jeffreys, H., Bullen, K.E., 1940. Seismological Tables. British Association for the Advancement of Science, London.

Jung, H., Karato, S., 2001. Water-induced fabric transition in olivine. Science 293, 1460-1463.

Karato, S., 1993. Importance of Anelasticity in the Interpretation of Seismic Tomography. Geophys. Res. Lett. 20, 1623-1626.

Karato, S.-I., 2010. Rheology of the deep upper mantle and its implications for the preservation of the continental roots: A review. Tectonophysics 481, 82-98.

Katayama, I., Jung, H., Karato, S., New type of olivine fabric from deformation experiments at modest water content and low stress. Geology 32, 1045-1048.

Kato, M., Misawa, M., Kawakatsu, H., 2001. Small subsidence of the 660- discontinuity beneath Japan probed by ScS reverberations. Geophys. Res. Lett. 28, 447-450.

Katsura, T., Ito, E., 1989. The system Mg_2SiO_4 – Fe_2SiO_4 at high pressures and temperatures: precise determination of stabilities of olivine, modified spinel and spinel. J. Geophys. Res. 94(B11), 15663-15670.

Kawai, K., Takeuchi, N., Geller, R.J., 2006. Complete synthetic seismograms up to 2 Hz for transversely isotropic spherically symmetric media. Geophys. J. Int. 164, 411-424.

Kawai, K., Takeuchi, N., Geller, R.J., Fuji, N., 2007a. Possible evidence for a double crossing phase transition in D″ beneath Central America from inversion of seismic waveforms. Geophys. Res. Lett. 34, L09314, doi:10.1029/2007GL029642.

Kawai, K., Geller, R.J., Fuji, N., 2007b. D″ beneath the Arctic from inversion of shear waveforms. Geophys. Res. Lett. 34, L21305, doi:10.1029/2007GL031517.

Kawai, K., Sekine, S., Fuji, N., Geller, R.J., 2009. Waveform inversion for D″ structure beneath northern Asia using Hi-net tiltmeter data. Geophys. Res. Lett., 36, L20314, doi:10.1029/2009GL039651.

Kawai, K., Tsuchiya T., 2009. Temperature profile in the lowermost mantle from seismological and mineral physics joint modeling. Proceedings of the National Academy of Sciences of the United States of America. doi:10.1073/pnas.0905920106.

Kawai, K., Geller, R.J., 2010. Waveform inversion for localized seismic structure and an application to D″ structure beneath the Pacific. J. Geophys. Res. 115, B01305, doi:10.1029/2009JB006503

Kennett, B.L.N., Engdahl, E.R., 1991. Traveltimes for global earthquake location and phase identification. Geophys. J. Int. 122, 108-124.

Kennet, B.L.N., Widiyantoro, S., van der Hilst, R.D., 1998. Joint seismic tomography for bulk sound and shear wave speed in the Earth's mantle. J. Geophys. Res. 103, 12469-1493.

Konishi, K., Kawai, K., Geller, R.J., Fuji, N., 2009. MORB in the lowermost mantle beneath the western Pacific: Evidence from waveform inversion. Earth Planet. Sci. Lett. 278, 219-225.

Kneller, E.A., van Keken, P.E., Karato, S.-I., Park, J., 2005. B-type olivine fabric in the mantle wedge: Insights from high-resolution non-Newtonian subduction zone models. Earth Planet. Sci. Lett. 237, 781-797.

Lay, T., Helmberger, D.V., 1983. A lower mantle S-wave triplication and the shear velocity structure of D″. Geophys. J. R. Astr. Soc. 75, 799-837.

Lay, T., Hernlund J., Garnero E.J., Thorne, M.S., 2006. A post-perovskite lens and D″ heat flux beneath the central Pacific. Science. 314, 1272-1276.

Lerche, I., Menke, W., 1986. An inversion method for separating apparent and intrinsic attenuation in layered media. Geophys. J. Royal Ast. Soc. 87, 333-347.

Li, X.D., Romanowicz, B., 1996. Global mantle shear velocity model developed using nonlinear asymptotic coupling theory. J. Geophys. Res. 101, 22245-22272.

Maruyama, S., Hasegawa, A., Santosh, M., Kogiso, T., Omori, S., Nakamura, H., Kawai, K., Zhao, D., 2009. The dynamics of big mantle wedge, magma factory, and metamorphic-metasomatic factory in subduction zones. Gondwana Res. 16, 414-430.

Masters, G., Laske, G., Bolton, H., Dziewonski, A.M., 2000. The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: Implications for chemical and thermal structure. Earth's Deep Interior: Mineral Physics and Tomography From the Atomic to the Global Scale, pp. 63-87. ed. Karato, S., Forte, A.M., Libermann, R.C., Masters, G., Stixrude, L., AGU, Washington D.C.

Matzel, E., Grand, S.P., 2004. The anisotropic seismic structure of the East European platform. J. Geophys. Res. 109, B01302. doi:10.1029/2001JB0000623.

Maupin, V., Garnero, E.J., Lay, T., Fouch, M.J., 2005. Azimuthal anisotropy in the D″ layer beneath the Caribbean. J. Geophys. Res. 110, doi:10.1029/2004JB003506.

Mégnin, C., Romanowicz, B., 2000. The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms. Geophys. J. Int. 143, 709-728.

Menke, W., 1984. Geophysical data analysis: discrete inverse theory. Academic Press, Inc., Orlando, 260 pp.

Menke, W., Dubendorff, B., 1985. Discriminating intrinsic and apparent attenuation in layered rock. Geophys. Res. Lett. 12, 721-724.

Nakagawa, T., Tackley, P.J., 2006. Three-dimensional structures and dynamics in the deep mantle: Effects of post-perovskite phase change and deep mantle layering. Geophys. Res. Lett. 33, L12S11, doi:10.1029/2006GL025719.

Nakagawa, T., Tackley, P.J., 2008. Lateral variations in CMB heat flux and deep mantle seismic velocity caused by a thermal-chemical-phase boundary layer in 3D spherical convection. Earth Planet. Sci. Lett. 271, 348-358.

Nishida, K., Montagner, J.P., Kawakatsu, H., 2009. Global surface wave tomography using seismic hum. Science. 326. 112-112.

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

Obayashi, M., Sugioka, H., Yoshimitsu, J., Fukao, Y., 2006. High temperature anomalies oceanward of subducting slabs at the 410-km discontinuity. Earth Planet. Sci. Lett. 243, 149-158.

Okal, E.A., Geller, R.J., 1979. On the observability of isotropic seismic sources: The July 31, 1970 Colombian earthquake. Phys. Earth Planet. Inter. 18, 176-196.

Panning, M., Romanowicz, B., 2004. Inferences on flow at the base of Earth's mantle based on seismic anisotropy. Science 303, 351-353.

Park, J., Levin, V., 2002. Geophysics - Seismic anisotropy: Tracing plate dynamics in the mantle. Science 296, 485-489.

Pinheiro, K., Jackson, A., 2008. Can a 1-D mantle electrical conductivity model generate magnetic jerk differential time delays? Geophys. J. Int. 173, 781-792.

Reeves, C.M., 1964. Algorithm-238 – Conjugate gradient method [F4]. Communications of the ACM. 7, 481-481.

Revenaugh, J., Jordan, T.H., 1991. Mantle layering from ScS reverberations, 3. The upper mantle. J. Geophys. Res. 96, 19781-19810.

Rhie J., Romanowicz, B., 2004. Excitation of Earth's continuous free oscillations by atmosphere-ocean-seafloor coupling. Nature. 431, 552-556.

Ringwood, A.E., and Major, A., 1970. The system Mg_2SiO_4 – Fe_2SiO_4 at high pressures and temperatures. Phys. Earth Planet. Inter. 3, 89-108.

Romanowicz, B., 1994. On the measurement of anelastic attenuation using amplitudes of low-frequency surface-waves. Phys. Earth Planet. Inter. 84, 179-191.

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

Romanowicz, B., 1998. Attenuation tomography of the earth's mantle: A review of current status. Pure Appl. Geophys. 153, 257-272.

Savage, M.K., 1999. Seismic anisotropy and mantle deformation: what have we learned from shear wave splitting? Rev. Geophys. 37, 65-106.

Schubert, G., Yuen, D.A., Turcotte, D.L., 1975. Role of phase-transitions in a dynamic mantle. Geophys. J. Royal Astr. Soc. 42, 705-735.

Seama, N., Baba, K., Utada, H., Toh, H., Tada, N., Ichiki, M., Matsuno, T., 2007. 1-D electrical conductivity structure beneath the Philippine Sea: Results from an ocean bottom magnetotelluric survey. Phys. Earth Planet. Int. 162, 2-12.

Sengupta, M.K., Toksöz, M.N., 1976. 3 Dimensional model of seismic velocity variation in Earth's mantle. Geophys. Res. Lett. 3, 84-86.

Shearer, P.M., 1990. Seismic imaging of upper-mantle structure with new evidence for a 520- km discontinuity. Nature 344, 121-126.

Shimizu, H., Koyama, H., Baba, T., Utada, H., 2009. Three-dimensional geomagnetic response functions for global and semi-global scale induction problems. Geophys. J. Int. 178, 123-144.

Silver, P.G., Savage, M.K., 1994. The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers. Geophys. J. Int 119, 949-963.

Song, T.R.A., Helmberger, D.V., Grand, S.P., 2004. Low-velocity zone atop the 410-km seismic discontinuity in the northwestern United States. Nature 427, 530-533.

Suetsugu, D., Saita, T., Takenaka, H., Niu, F., 2004. Thickness of the mantle transition zone beneath the South Pacific as inferred from analyses of ScS reverberated and Ps converted waves. Phys. Earth Planet. Inter.146, 35-46.

Tackley, P.J., 2000. Mantle convection and plate tectonics: Toward an integrated physical and chemical theory, Science. 288, 2002-2007.

Tajima, F., Grand, S.P., 1998. Variation of transition zone high-velocity anomalies and depression of 660 km discontinuity associated with subduction zones from the southern Kuriles to Izu-Bonin and Ryukyu. J. Geophys. Res. 103, 15015-15036.

Takeuchi, N., Geller, R.J., Cummins, P.R., 1996. Highly accurate P-SV complete synthetic seismograms using modified DSM operators. Geophys. Res. Lett. 23, 1175-1178.

Takeuchi, N., Geller, R.J., Cummins., P.R., 2000. Complete synthetic seismograms for 3-D heterogeneous Earth models computed using modified DSM operators and their applicability to inversion for Earth structure. Phys. Earth Planet. Inter. 119, 25-36.

Takeuchi, N., Kobayashi, M., 2004. Improvement of seismological earth models by using data weighting in waveform inversion. Geophys. J. Int. 158, 681-694.

Takeuchi, N., 2007. Whole mantle SH velocity model constrained by waveform inversion based on three-dimensional Born kernels. Geophys. J. Int. 169, 1153-1163.

Tanimoto, T., 1984. Waveform inversion of mantle Love waves – The born seismogram approach. Geophys. J. R. Astr. Soc. 78, 641-660.

Tarantola, A., 1984. Inversion of seismic reflection data in the accoustic approximation. Geophysics 49, 1259-1266.

Thomas, C., Garnero, E.J., Lay, T., 2004. High-resolution imaging of lowermost mantle structure under the Cocos plate. J. Geophys. Res. 109, B08307.

Torii, Y., Yoshioka, S., 2007. Physical conditions producing slab stagnation: Constraints of the Clapeyron slope, mantle viscosity, trench retreat, and dip angles. Tectonophysics. 445, 200-209.

Utada, H., Koyama, T., Obayashi, M., Fukao, Y., 2009. A joint interpretation of electromagnetic and seismic tomography models suggests the mantle transition zone below Europe is dry. Earth Planet. Sci. Lett. 281, 249-257.

Vinnik, L.P., Kind, R., Kosarev, G.L., Makeyeva, L.I., 1989, Azimuthal anisotropy in the lithosphere from observations of long-period S-waves, Geophys. J. Int. 99, 549-559.

Vinnik, L., Montagner, J.-P., 1996. Shear wave splitting in the mantle Ps phases, Geophys. Res. Lett. 23, 2449-2452.

Wen, L., Anderson, D.L., 1997. Layered mantle convection: A model for geoid and topography. Earth Planet. Sci. Lett. 146, 367-377.

Widiyantoro, S., Kennet, B.L.N., van der Hilst, R.D., 1999. Seismic tomography with P and S data reveals lateral variations in the rigidity of deep slabs. Earth Planet. Sci. Lett. 173, 91-100.

Wiggins, R.A., 1972. General linear inverse problem - implication of surface-waves and free oscillations for earth structure. Rev. Geophys. Space Phys. 10, 251-285.

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

Xie, S., Tackley, P.J., 2004. Evolution of U-Pb and Sm-Nd systems in numerical models of mantle convection and plate tectonics. J. Geophys. Res. 109, B11204. doi:10.1029/2004JB003176.

Zhao, D., 2004. Global tomographic images of mantle plumes and subducting slabs: insight into deep Earth dynamics. Phys. Earth Planet. Inter. 146, 3-34.

Zhang, S.Q., Karato, S., Fitz Gerald, J.D., Faul, U.H., Zhou, Y. 2000. Simple shear deformation of olivine aggregates. Tectonophysics 316, 133-152.