Title
Investigating deep and shallow earth structures using different seismological approaches

Author
Chen, Li-Wei

Publication Date
2021

Peer reviewed|Thesis/dissertation

Investigating deep and shallow earth structures using different seismological approaches

by

Li-Wei Chen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Earth and Planetary Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Barbara Romanowicz, Chair
Professor Hans-Rudolf Wenk
Professor Harriet Lau

Fall 2021

Investigating deep and shallow earth structures using different seismological approaches

Abstract

Investigating deep and shallow earth structures using different seismological approaches

by

Li-Wei Chen

Doctor of Philosophy in Earth and Planetary Science

University of California, Berkeley

Professor Barbara Romanowicz, Chair

Investigating the elastic structure at multiple scales of the Earth is crucial for understanding the composition and dynamics of the Earth. Different seismological approaches to extract the information contained in seismic waveforms are presented here. The work consists of three topics: (I) applications of source stacking followed by cross-correlation, (II) modeling seismic anisotropy in the lower mantle and (III), temporal variation of near-surface seismic velocity and anisotropy in Hokkaido.

I. Accurate synthetic seismic wavefields can now be computed in 3D earth models using the spectral element method, which helps improve resolution in full waveform global tomography. However, computational costs are still a challenge. These costs can be reduced by implementing a source stacking method (Capdeville et al. 2003) , in which multiple earthquake sources are simultaneously triggered in only one teleseismic SEM simulation. One drawback of this approach is the perceived loss of resolution at depth, in particular because high-amplitude fundamental mode surface waves dominate the summed waveforms, without the possibility of windowing and weighting as in conventional waveform tomography. This can be addressed by redefining the cost-function and computing the cross-correlation wavefield between pairs of stations before each inversion iteration. While the Green's function between the two stations is not reconstructed as well as in the case of ambient noise tomography, where sources are distributed more uniformly around the globe, this is not a drawback, since the same processing is applied to the 3D synthetics and to the data, and the source parameters are known to a good approximation. By doing so, time windows with large energy arrivals corresponding to fundamental mode surface waves can be separated. Also, applying a weighting scheme can bring out the contribution of overtones and body waves and also can balance the contributions of frequently sampled paths. The approach is computationally very efficient and can help address such questions as model resolution in the presence of noise, and trade-offs between different physical parameters (anisotropy, attenuation, crustal structure etc..) that would be computationally very costly to address adequately, when using

conventional full waveform tomography based on single-event wavefield computations.

II. By assuming a geodynamical scenario of slab subduction and considering the seismic anisotropy in D" is dominated by crystal preferred orientation, the textural evolution and resultant elastic properties can be computed, which may help to explain the seismic anisotropy observations (e.g., McNamara et al. 2002, 2003; Wenk et al. 2011). In previous work, Cottaar et al. (2014) calculated the seismic anisotropy produced in a single mineral system, comparing bridgmanite ($MgSiO_3$ perovskite, Pv) to magnesium post-perovskite (pPv), based on elastic properties and slip systems determined from laboratory experiments and theoretical ab-initio computations. More realistic situations are considered in this study: (a) the polycrystal plasticity model (Lebensohn and Tomé 1993) which consists Pv/pPv, cubic calcium perovskite ($CaSiO_3$, CaPv), and cubic periclase (MgO), (b) the forward and reverse Pv-pPv phase transitions during the slabs' subduction and the subsequent upwelling, (c) partial melting in the deepest portions of the slab at the base of upwelling. To validate and compare the results with seismological observations, the spatial distribution of radial anisotropy described by the parameter $\xi = (V_{SH}/V_{SV})^2$, and shear wave splitting (SWS) directions and strengths are extracted.

III. Here an earthquake-based method (Chen et al. 2017) is applied through interferometry of earthquake coda waves to investigate the temporal changes of elastic properties by using borehole-surface station pairs. The technique provides comprehensive observations not only of S-wave velocity but, additionally, of P-wave and S-wave azimuthal anisotropy between two sensors. The 15-year temporal variations of elastic properties are monitored at the KiK-net station IBUH03 in southern Hokkaido, which experienced the 2003 Mw 8.3 Tokachi-Oki (PGA= $\sim 350 cm/s^2$) and the 2018 Mw 6.7 Hokkaido Eastern Iburi (PGA= $\sim 500 cm/s^2$) earthquakes. The preliminary results show sudden reductions of V_{Siso} and V_P after the two major earthquakes considered, and these were subsequently recovered with short-term and long-term recovery rates. Changes in the fast S-wave polarization direction and strength of anisotropy were also detected during the coseismic periods. Furthermore, a clear seasonal trend of elastic parameters can be observed. The results are compared with temporal variations in equivalent water thickness, precipitation, and temperature. A Hudson-Crampin anisotropic effective model is implemented to understand the seasonal crack behavior. The modeling result indicates that crack density and the aspect ratio could play an important role in our observations

Contents

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank my advisor Barbara Romanowicz for her guidance, encouragement, and support over the last five years. She has always been enthusiastic and optimistic in advising on my research topics. I would also like to thank the other two dissertation committee members, Rudy Wenk and Harriet Lau, for providing valuable feedback for this dissertation. Much of the works were benefited from their thorough reviews. I would like to acknowledge all my collaborators, Scott French, Chase Chandler, Mingming Li, Taka'aki Taira, and Hiroe Miyake, for providing their insights and assistance. I am grateful to all the past and present members of the Global Seismology Group, especially Yann Capdeville, for his vision of the source stacking method and for providing us with the cSEM code.

I would like to thank all the staff in EPS and BSL for maintaining administrative services and providing assistance with computing resources. These works were financially supported by a UC Berkeley Graduate Fellowship, grant EAR-1758198 from the National Science Foundation, and Taiwan government scholarship to study abroad (GSSA). Much of the computation works were carried out at the National Energy Research Scientific Computing Center and Texas Advanced Computing Center. Therefore, I would like to thank them for granting us access to the resources and maintaining the facilities.

It has been a wonderful experience to interact with the members in BSL and EPS. Thank you all for creating a friendly and inspiring community. It's been a pleasure to have been involved in part of this community. Special thanks to Dana Lapides and her partner Michael Pinkard for their warm friendship and companionship.

I have to acknowledge all my family and friends in Taiwan whose constant encouragement and support, even from the other side of the Pacific Ocean, has been essential for me to complete this journey. Last but not least, this journey would not have been possible without the companionship of my fiancee Yu-Hsiang Chien, with whom I feel so lucky to have shared my life for the past five years.

Chapter 1

Introduction

The dissertation consists of three topics: (1) applications of source stacking followed by cross-correlation (2) modeling seismic anisotropy in the lower mantle and (3) temporal variation of near-surface seismic velocity and anisotropy in Hokkaido.

1.1 Applications of source stacking followed by cross-correlation for full-waveform tomography

Global seismic tomography traditionally relies on constraints from traveltimes of a limited number of seismic 'phases', which allow very fast forward computations, even when finite frequency kernels are considered. The seismic phases considered are those that are easy to isolate on seismograms, such as fundamental mode surface waves or teleseismic P and S waves, and in some studies, their surface and core reflections, leaving a large part of the information contained in seismic records unexploited.

On the other hand, full waveform inversion should ultimately make it possible to exploit every wiggle in a seismogram. It was first introduced by Tarantola (1984) in the world of exploration geophysics and by Woodhouse and Dziewonski (1984) for the study of global earth structure. Full waveform inversion requires the computation of synthetic seismograms in the current 3-D earth model, at each iteration of the model. Woodhouse and Dziewonski (1984) derived this in the context of the simplest asymptotic approximation with first order normal mode perturbation theory, that was shown to be equivalent to the 'path-average' approximation (PAVA) used in surface wave dispersion studies (Mochizuki 1986; Park 1987; Romanowicz 1987). Li and Tanimoto (1993) showed how the sensitivity of body waves to structure may be accounted for more accurately by including coupling across mode branches (not included in PAVA), while Li and Romanowicz (1995) combined this with the PAVA to develop the NACT (non-linear asymptotic coupling theory). This allowed these authors to construct whole mantle global shear velocity models, based entirely on long period waveforms, including surface waves, overtones as well as body waves down to 30 s period (e.g. Li and Romanowicz 1996; Mégnin and Romanowicz 2000; Panning and Romanowicz 2006).

Recently, with the introduction to global seismology of the spectral element method (SEM, e.g. Komatitsch and Vilotte 1998; Komatitsch and Tromp 1999, 2002), an efficient numerical algorithm allows the computation of accurate regional and teleseismic wavefields in 3-D earth models containing realistic wavelengths and strengths of heterogeneity. While efficient compared to other numerical methods, the computational cost is much higher than ray theory or PAVA/NACT, in particular as it increases roughly as the cube of the highest frequency considered.

These costs can be reduced by implementing a source stacking method, in which multiple earthquake sources are simultaneously triggered in only one teleseismic SEM simulation (Capdeville et al. 2005), which can reduce computational time by at least an order of magnitude. In this approach, many different sources are first aligned at their origin time, normalized by their seismic moment. The resulting summed wavefield is computed only once per iteration and compared with the corresponding observed summed wavefield at each station. Capdeville et al. (2005) showed that this could work with a relatively small number of events and stations, for the retrieval of very long wavelength whole mantle structure. One drawback of this approach is the perceived loss of resolution at depth, in particular because high-amplitude fundamental mode surface waves dominate the summed waveforms, without

the possibility of windowing and weighting as in conventional waveform tomography. This can be addressed by redefining the cost-function and computing the cross-correlation wavefield between pairs of stations before each inversion iteration. While the Green's function between the two stations is not reconstructed as well as in the case of ambient noise tomography, where sources are distributed more uniformly around the globe, this is not a drawback, since the same processing is applied to the 3D synthetics and to the data, and the source parameters are known to a good approximation. By doing so, time windows with large energy arrivals corresponding to fundamental mode surface waves can be separated. Also, applying a weighting scheme can bring out the contribution of overtones and body waves and also balance the contributions of frequently sampled paths versus rarely sampled ones, as in more conventional tomography. While application to real data requires additional steps to account for such issues as missing data, this methodology can help address such questions as model resolution in the presence of noise, and trade-offs between different physical parameters (anisotropy, attenuation, crustal structure etc..) that would be computationally very costly to address adequately, when using conventional full waveform tomography based on single-event wavefield computations.

In chapters 2-4, accelerating full-waveform inversion (FWI) by using source stacking followed by cross-correlation is presented. Also, demonstrations of the applications from proof-of-concept, to experiments of realistic synthetic long-period data at the global scale are shown. In chapter 2, the results of testing such an approach for a synthetic 3-component long period waveform dataset (periods longer than 60 s) are presented, computed for 273 globally distributed events in a toy 3D radially anisotropic upper mantle model (the "BLOB") which contains shear wave anomalies at different scales. The comparisons of results of inversion of 10,000 s stacked time series, starting from a 1D model, using source stacked waveforms and station-pair cross-correlations of these stacked waveforms in the defiinition of the cost function are shown. The effects of crust and uppermost mantle on teleseismic waveforms must be carefully addressed when performing full-waveform inversion for global mantle structure using the spectral element method for wavefield computations.

In chapter 3, the merits of constraining the structure using surface wave dispersion data in the period range 20-150s are evaluated. A group-velocity dispersion-based models extending down to two different depths (171 and 471 km) are constructed using the concept of homogenization, and compare the waveform fits that these models provide to those of model SEMUCB-WM1, the crust and upper-mantle part of which was constructed using FWI with added constraints from dispersion, at each iteration. An improved, potentially versatile approach is proposed to taking into account uppermost mantle structure in full-waveform tomography while limiting computational costs.

In chapter 4, a synthetic test based on a realistic synthetic dataset derived from SEMUCB-WM1 (French and Romanowicz, 2014) is conducted for the first time, which includes a 3D crustal model and mantle structure down to scales of 800 km. to evaluate the capability of resolving the shear wave velocity and its anisotropic structure in the upper mantle. The results show structure can be well recovered down to at least 1,500 km depth both by straight source stacking and by stacking followed by cross-correlation. Dealing with

the issue of missing data, two different strategies are combined: (1) using a cluster analysis for grouping events and stations in order to minimize missing waveforms in each group and (2) replacing missing waveforms by synthetics computed in the current 3D model using an efficient approximate method such as normal mode perturbation theory.

1.2 Modeling seismic anisotropy in the lower mantle

Anisotropy in the deepest 200-300 kilometers of the Earth's mantle, called D" (Bullen 1949), is still poorly understood due to lack of sampling by seismic waves and signal contamination from anisotropy in the upper mantle. Shear wave splitting (SWS) measurements on the core-mantle boundary generally show that $V_{SH} > V_{SV}$ is found in regions of faster than average isotropic shear wave velocity (Vs_{iso}) in seismic imaging results (the graveyard of slabs); $V_{SH} < V_{SV}$ or no significant splitting is found in regions of slower than average Vs_{iso} such as the large low shear velocity provinces (LLSVPs) beneath the central Pacific and Africa. Also, there are uncertainties on the deformation properties of pPv in lower mantle, where laboratory experiments (e.g., Miyagi et al. 2010; Wu et al. 2017) favor a (001) slip plane but theoretical computations (e.g., Goryaeva et al. 2016) disagree and prefer a (010) slip plane in pPv.

By assuming a geodynamical scenario of slab subduction and considering the seismic anisotropy in D" is dominated by crystal preferred orientation, the textural evolution and resultant elastic properties can be computed. The following situations are considered in chapter 5: (a) the polycrystal plasticity model (Lebensohn and Tomé 1993) which consists Pv/pPv, cubic calcium perovskite ($CaSiO_3$, CaPv), and cubic periclase (MgO), (b) the forward and reverse Pv-pPv phase transitions during the slabs' subduction and the subsequent upwelling, (c) partial melting in the deepest portions of the slab at the base of upwelling. The radial anisotropy described by the parameter $\xi = (V_{SH}/V_{SV})^2$ is computed. The result shows that the 001 model has a strong but complicated anisotropy, with $V_{SH} > V_{SV}$ in regions of faster than average isotropic V_s. At the base of the upwelling, 001 model presents the patterns of $V_{SH} < V_{SV}$ which is consistent with available seismic observations (e.g., Long 2009; Romanowicz and Wenk 2017) while the pattern is not obvious in the 010 model. In addition, the SWS polarization direction and the strength of splitting are extracted. For Model 010, the fast axis direction remains horizontal (∼1.0–3.0 % SWS strength) after the reverse transition and the trend continues during the upwelling; the model 001 shows a complex pattern of mixed horizontal and vertical fast axis directions (∼2.0 % SWS strength) after the reverse pPv to Pv transition, and a tilted fast axis (∼1.0–2.0 % SWS strength) near the borders of the upwelling segment also in agreement with Cottaar and Romanowicz (2013).

1.3 Temporal variation of near-surface seismic velocity and anisotropy in Hokkaido

Temporal variations of elastic properties of the crustal structure provide a unique constraint on its behaviors such as change of dynamic stress after earthquakes, seasonality from the responses to hydrological phenomena, healing process of fractures, etc, which could lead to better understand the near-surface structure and could provide information for site classification. In the past decade, noise-based monitoring has become a very well-developed and efficient procedure for probing the S-wave changes. An earthquake-based method (Chen et al. 2017) is proposed through interferometry of earthquake coda waves, which will allow to extract the near-surface temporal variations on not only S-wave velocity but additionally P-wave and S-wave azimuthal anisotropy (strength of anisotropy and the polarization direction of the fast S wave). Only observations in the southern Hokkaido area are discussed which experienced strong ground motion from the 2003 Mw 8.3 Tokachi-Oki and the 2018 Mw 6.7 Hokkaido Eastern Iburi earthquakes.

In chapter 6, the decade-long temporal variation of near-surface seismic velocity and Vs azimuthal anisotropy in Southern Hokkaido is presented. The data from all available KiK-net seismic waveforms ($\sim$800 events) is then analyzed in a frequency band of 1-15 Hz around the southern Hokkaido area, following the recipe from Chen et al. (2017). Evaluation of the temporal robustness of the reconstructed cross-correlation functions (CCFs) in each time segment ensures the reliability of observations. The preliminary result shows that IBUH03 has sudden reductions of V_{Siso} and V_P after the two events and then recover at short-term and long-term recovery rates; changes in the fast S-wave polarization direction and strength of anisotropy are also observable during the coseismic periods. The results indicate that the dynamic stress changes after the events create the opening of fluid filled cracks. It is also responsible for the reductions of seismic velocities and the changes of V_S azimuthal anisotropy. Furthermore, seasonality is notable from some of the observations which imply the changes of elastic properties might also be affected by hydrological process. The results are compared with equivalent water thickness, precipitation, and temperature to understand the possible mechanism of the temporal variation.

Chapter 2

Applications of source stacking and followed by cross-correlation for full-waveform tomography

---

modified from "Accelerating full waveform inversion via source stacking and cross correlations" (Romanowicz et al., 2020, GJI). Results of computation and analysis are contributed by the author (Chen).

2.1 Introduction

This study presents and illustrates a computation saving approach, based on source stacking, as first introduced by Capdeville et al. (2005), which can reduce computational time by at least an order of magnitude. In this approach, many different sources are first aligned at their origin time, normalized by their seismic moment. The resulting summed wavefield is computed only once per iteration and compared with the corresponding observed summed wavefield at each station. Capdeville et al. (2005) showed that this could work with a relatively small number of events and stations (84 events and 174 stations), for the retrieval of very long wavelength whole mantle structure. This approach, however was not pursued further at least in global seismology, due to concerns that information was lost in the process of summation, and that the summed records are dominated by large amplitude fundamental mode surface waves, without the possibility of relative weighting of body waves, overtone and fundamental mode surface waves, nor weighting of redundant paths, as can be done when working with single records (e.g. Li and Romanowicz 1996; Maggi et al. 2009).

In exploration geophysics, such an approach was perfected and is still an active area of research. It was shown in particular that 'source encoding', that is the use of multiple sums over the same group of events to which a random perturbation to the source phase is applied, can help improve the resolution of the inverted model (e.g. Krebs et al. 2009; Schuster et al. 2011)—particularly by reducing 'cross-talk', that is, the appearance of interference terms in the computation of the gradient by the adjoint method, which introduces noise in the solution. This is not an issue when computing the gradient and Hessian using NACT, so no source encoding should be necessary.

As noted above, another key concern when dealing with summed traces is the risk of information loss. However, information needs not be lost when using source encoding, as illustrated in Zhang et al. (2018) where the authors use a source encoding method that allows them to decode single records from the sum, albeit at the particular frequency used for encoding. Recovering the complete record requires multiple runs with different frequency encoding, which can be done, but at the expense of increasing computational time.

An new idea is taken from the now widely popular field of ambient noise tomography (ANT, e.g. Shapiro and Campillo 2004) due to the fact that surface waves cannot be down-weighted in the summed seismograms. Indeed, it has been known since the work of Aki (1957) that the Green's function between two stations can be retrieved from the cross-correlation of the noise wavefield at these two stations, provided the noise sources are uniformly distributed. The resulting methodology, which involves stacking of time-series over significant time periods, has been particularly successful for the retrieval of crustal structure, owing to the strong prevalent seismic noise in the microseismic band (3–25 s), made up primarily of short period surface waves generated in the oceans. Attempts at retrieving deeper mantle structure, requiring the use of weaker, longer period noise, have shown potential at retrieving the very longest wavelengths of upper mantle structure (Nishida 2013; Haned et al. 2016). Body wave energy has been identified in these cross-correlation stacks (e.g. Gerstoft et al.

2008; Lin et al. 2013). However many issues remain before an approach based solely on ANT can be used confidently for global tomography (e.g. Tsai 2009; Pham et al. 2018).

The idea presented here is adding a second step after source stacking, namely computing the cross-correlation of the source-stacked wavefield between pairs of stations and using these cross-correlations as data in the definition of the cost function for the inversion. This step is computationally negligible. Because the distribution of earthquake sources is not uniform around the globe, the Green's function between two given stations will not be retrieved perfectly. However, this is not an issue for this approach, since in this case, unlike in ANT, the location and source mechanism of the sources are well known, so that the synthetic sums and following cross-correlations can be formed in exactly the same way as the observed ones. An additional advantage of using earthquake sources is that the depth distribution of earthquakes is broader than that of the primary noise sources, which are confined to the ocean-seafloor interface, thus providing stronger excitation of overtones and body waves in a wide frequency band. Stated differently, recovery of the Green's function is not the objective of the technique described here although motivated by ANT, the cross-correlation is simply used as a derived data type that has specific advantages over plain summed traces. Namely: (1) the dominant energy corresponding to the fundamental mode surface waves is more localized in time, so that appropriate windowing and weighting may be applied in order to more effectively bring out the contribution of smaller amplitude overtones and body waves to the wavefield and (2) it is possible to introduce path weighting, to avoid the dominance of redundant paths in the inversion.

In what follows, an illustration shows how the additional cross-correlation step can help separate the large amplitude waveforms due to fundamental mode surface waves, thus mitigating the primary source of information loss associated with inversion of summed traces. Synthetic tests are presented here for first inversion using the stacked records themselves, and then for inversion using cross-correlations of stacked records. The computation of partial derivatives of the cross-correlations is outlined in the Appendix. These kernels have been computed using NACT, however, extending them to an adjoint or any other inversion approach is straightforward.

2.2 Methods

2.2.1 Separation of distinct wave trains by cross-correlation

Fig. 2.1 shows the distribution of events and stations used for the synthetic experiments presented in this paper. It consists of the 273 earthquake sources and 537 stations used for the development of SEMUCB-WM1. However, 22 stations are removed which were located at a distance smaller than 150 km from one or more events, in order to avoid very large amplitudes at the beginning of the time-series. The assumption considered here is that no records are missing which will be discussed later.

Fig. 2.2 shows an example of vertical component 10,000 s long traces calculated at station TATO in the 1-D average model of SEMUCB-WM1, for 38 different events, aligned to origin time and ordered as a function of epicentral distance. The resulting summed trace is shown at the bottom. This figure illustrates how long period surface waves dominate the individual records and therefore the sum, hiding the lower amplitude overtone energy, and making it difficult to apply the same kind of weighting scheme as is applied in conventional waveform tomography, where larger amplitude phases are weighted down compared to weaker ones, in order to bring out the information contained in the latter (e.g. Appendix in Li and Romanowicz 1996). This is not a major drawback for the simple tests at long periods presented here, but could become an issue when higher resolution lower mantle structure is targeted, which requires the inclusion of shorter period, relatively low amplitude, body waves.

In Figs 2.3-2.4, L–L, T–T and Z-Z station pair cross-correlations for the entire data set are shown. The cross-correlations have been ordered as a function of distance between stations, and stacked in bins of 100 km in distance. Fig. 2.3 upper panel and Fig.2.4 shows the emergence of fundamental mode Rayleigh wave energy on the L and Z components and Love wave energy on the T component in Fig. 2.3 lower panel, for the minor and major arcs. The strong energy in front of these surface waves contains information on earth's deeper structure. In particular, for the L–L component, reconstructed overtone energy is also distinguishable.

Notably, the main goal of computing the cross-correlations is not to extract the Green's function, for which the distribution of stacked events is clearly insufficient, but to enable the capability of windowing out the strong energy corresponding to fundamental mode surface waves, giving better access to the information contained in overtones. Because the same operations are performed on the synthetics and the actual data, forming Green's function would not be an issue.

There are two ways of bringing out the pseudo-Green's function by cross-correlation.The first one is presented in Figs 2.3-2.4, stacking over distance bins. The other one is obtained by stacking individual station-pair cross-correlations obtained in several realizations of summed randomly 1-bit source-encoded synthetics. For example, Fig. 2.5 shows the results of an experiment in which the Z–Z cross-correlations have been obtained from stacking 10 realizations of random 1-bit source-encoded synthetics (e.g. Krebs et al. 2009). Aligned by interstation distance, the energy corresponding to the fundamental mode Rayleigh wave is strongly localized. In this way of stacking, the computational cost is larger, since it needs to compute a number of realizations of the stacked seismograms. Because the stacking is performed over a given interstation path, it may result in better recovery of structure than when cross-correlations are stacked over specific distance bins, where station pairs may correspond to different locations on the globe. However, since the same operations are applied to the data and the synthetics, the inversion should recover structure in the correct geographical location also in the latter case.

2.2.2 Construction of the synthetic data sets and experiments

In order to develop the methodology and for the initial tests on a synthetic data set, a simple global shear velocity model is constructed, hereafter called "BLOB", which contains both large and small anomalies in shear wave velocity (Vs) distributed around the globe, and centred at different depths in the upper mantle and transition zone (Fig. 2.6). It also includes heterogeneities in the anisotropic parameter $\xi = (Vsh/Vsv)^2$, with a different distribution and size than for the isotropic anomalies, and centred at a depth of 150 km (Fig. 2.7). In the synthetic tests presented here, lateral variations in crustal structure are not considered. Three component summed synthetic seismograms of 10,000 s duration were computed in these models for the 273 events used in the development of SEMUCB-WM1 (French and Romanowicz 2014), and the 515 stations considered here. The sources were aligned at their origin time, normalized by the seismic moment of each source, with source parameters obtained from the CMT catalogue (Dziewonski et al. 1981; Ekström et al. 2012). Both minor and major arc Love and Rayleigh waves are included in the 10,000 s computation, although the corresponding cross-correlations are complete only up to $\pm$5000 s. Synthetics were computed using the CSEM (Capdeville et al. 2003), a version of a spectral element code which couples the SEM computation with 1-D normal mode computation in the core, originally developed for computational efficiency. In what follows, this synthetic data set will be referred to as 'data', while the predictions from successive iterations of modelling will be called 'synthetics'. These synthetics are also computed using CSEM, at each iteration.

The starting model for the inversion is the 1-D average model of "BLOB", which is the 1-D radially anisotropic part of model SEMUCB-WM1, a model which, contrary to PREM (Dziewonski and Anderson 1981) does not contain a discontinuity at 220 km depth. Spatial correlation lengths of 2000 km for Vs and 6000 km for ξ are used for the computation of the model covariance matrix in the inversion.

In the first experiment, the summed time domain seismograms were inverted directly, applying as many iterations as was necessary until convergence. In the second experiment, at each iteration of the inversion, cross-correlation was applied to all available pairs of stations, and inverted the cross-correlograms instead. In both experiments, NACT is used to compute the gradient and Hessian. The partial derivatives and Hessian in NACT depend non-linearly on the 3-D model, and thus are recomputed at every iteration, based on the evolving 1-D model and 3-D perturbations around it. The derivation of misfit function gradient and Jacobian, in the case of cross-correlations, is given in the Appendix, and is independent of the methodology used for the computation of gradient.

2.3 Inversion results

In what follows, synthetic tests of inversion using two different definitions of the cost function are shown here: one based directly on the source-stacked traces, and another based on the resulting station-pair cross-correlations.

2.3.1 Experiments on synthetic data without noise

Figs 2.8 and 2.9 show the results of inverting three component summed seismograms directly, for isotropic Vs and ξ, respectively, after six iterations. The residual variance does not evolve significantly after the 6th iteration (Fig. 2.10). The larger, shallow anomalies emerge first, constrained by the dominant fundamental mode energy, while the deeper, narrower anomalies come into focus in the later iterations. All anomalies are well recovered, with little distortion in shape, although with some loss in amplitude, as commonly observed in tomographic inversions, and due to the introduction of damping. As expected, and not unlike what is observed in conventional tomography, the anisotropic part is less well resolved, with larger amplitude loss and some vertical smearing.

Figs 2.11 and 2.12 show the corresponding results for the inversion of cross-correlations, with the variance reduction evolution shown in Fig. 2.10. The variance reduction tracks that of the summed seismogram inversion. It is slightly higher, due likely to the slight difference in the damping scheme used. There is no significant difference for the variance reduction in the results obtained by both methods.

2.3.2 Experiments on synthetic data with noise

2.3.2.1 Effect of noise on the signal to noise ratio of summed synthetics and cross-correlations

An important issue to consider is the effect of noise in the data. This is particularly important given the fact that much of the low frequency background noise is due to surface waves generated by ocean-solid earth interactions, the so-called 'seismic hum' (e.g. Suda et al. 1998; Nishida 2013), the energy of which is equivalent to about one M 5.8 earthquake per day. Since sources of seismic hum are not included in the computation of the synthetics, it is important to evaluate whether they might significantly contribute and bias the inversion. It is a priori unlikely, because months of data are usually stacked for the Green's function to emerge in cross-correlation ANT studies.

To assess the potential effect of background noise and estimate the signal to noise level in the cross-correlations, three stations in Eurasia are considered for which a large number of real records were available, stations ECH, HIA and MDJ. For these stations, 10,000 s of noise waveforms on the vertical component are collected before the origin time of each of the 179 events jointly recorded by both stations and compared them to 10,000 s of real data starting at the origin time of each event. The distribution of these events and location of the three stations is shown in Fig. 2.13. Fig. 2.14 (right-hand panels) also shows a comparison of the station pair cross-correlations of summed noise (red traces) and summed signals (black traces) for these data. Fig. 2.15 shows the cross-correlation spectra, filtered with corner frequencies of 20 and 800 s. The corresponding spectra for the summed traces are shown in Fig. 2.16. It is clear that in the period band considered and beyond (down to at least 20 s and up to 400 s), the signal-to-noise ratio is large. The background noise does not

significantly affect the recorded signals, at least in the frequency band considered here and up to frequencies that one would consider for whole mantle tomography including body-waveforms.

2.3.2.2 Inversion with realistic noise added to synthetic data

A synthetic data set constructed by adding realistic noise to each record are constructed for the inversion here. The real noise preceding each record can not be added in the previous experiment, because in the real data set considered for the development of SEMUCB-WM1, not all stations recorded all events. Instead, the probability density functions (PDFs) provided on the IRIS website (http://service.iris.edu/mustang/noise-pdf/1/), for vertical and horizontal components, are separately downloaded for the entire period of operation of each station. For the 178 stations for which this is not available, the average between the NLNM (new low noise model) and NHLM (new high noise model, Peterson (1993)) for both vertical and horizontal components is considered. The amplitude spectrum obtained at each station is converted back to the time domain with a random phase added. The filter is then used for the resultant time-series in all experiments (60-80-250-400s). This noise time-series N times is added to the synthetic summed data, where N is the number of events stacked (N = 273).

Figs 2.17 and 2.18 show the results of inversion of stacked seismograms after six iterations with this noise added. The anomalies are still recovered although their shape is a little distorted, and a stronger background level of spurious structure is introduced, as reflected in the much poorer variance reduction. Also, as expected, the effect of noise is stronger on the recovery of the anisotropic parameter ξ. The evolution of variance reduction as a function of iteration number is shown in Fig. 2.10. The variance reduction, as expected, is smaller than in the case without noise and stops evolving after the 4th iteration. The corresponding variance reduction for the cross-correlations after three iterations is also shown, which tracks that of the stacked-event inversion, as do the modelling results, as shown in Figs 2.19–2.20. To save computational time for more important tasks, no further iteration is obtained.

2.3.2.3 Evolution of model misfit as a function of iteration number

With access to more extensive synthetic tests than is possible in conventional full waveform tomography, one can follow the evolution of the model misfit as a function of iteration. The evolution model residual misfit for all the experiments described above, is defined as:

$$res_i = 100 \times \left[1 - \|m_0 - m_i\|_2^2 / \|m_0\|_2^2\right] \tag{2.1}$$

where $m_0 = dlnX_T$ represents the distribution of relative perturbations in the target model X_T with respect to the 1-D reference model X_0, with $X = V_s$ or ξ. The corresponding distribution of relative perturbations in the model obtained at iteration i is $m_i = dlnX_i$. A feature of the target "BLOB" model considered is that only a very small volume of the upper mantle contains non-zero 3-D perturbations. Fig. 2.21 a shows that when the model misfit is computed over the entire global volume of the model down to 550 km depth (below which

no 3-D structure exists in the target model), in the case of source stacking with noise, the residual misfit decreases until iteration 3, but stabilizes or starts slightly increasing at further iterations. The reason for this is that while the model continues to improve in the regions where 3-D structure actually exists, low amplitude spurious structure is introduced elsewhere because of the presence of noise, in particular due to the importance of the norm damping term of the regularization. Because the 3-D structure occupies only a small portion of the model space, the noise in the model plays an increasingly important role when quantifying the misfit. Indeed, Fig. 2.21b shows the evolution of the residual model misfit when only those parts of the model space that correspond to 3-D anomalies of amplitude larger than 1 percent are included in the calculation. In that case, the model misfit decreases continuously as a function of iteration. On the other hand, Fig. 2.21c shows that the increase in residual misfit after iteration 3 originates from the regions where weak or no 3-D structure exists in the target model. The same behaviour is found for $xi (Fig. 2.22).

The difference between the inverted Vs model and the target one is presented in Figs 2.23 and 2.24, for the cases without noise added of stacked-source inversion with and without adding cross-correlations, respectively, at iteration 6. It is also important to note a global distribution of the residual misfit with more or less uniform amplitude down to 300 km, while the amplitude recovery of the narrower anomalies at larger depth is poorer at least at this iteration, although slightly better in the case when cross-correlations are added (red ellipse in Fig. 2.24).

Figs 2.25 and 2.26 show the same as Figs 2.23 and 2.24 but for ξ. Note that the poorer recovery of the amplitudes of the target anomalies in ξ, and greater smearing with depth as already seen in Fig. 2.9, manifested here in the sidelobes on both sides of the target location of 3-D anomalies in ξ. Fig. 2.22bc indicate that the amplitude recovery of ξ in the regions where the target model has non-zero 3-D structure is slightly better with cross-correlation, but that more noise is introduced elsewhere. However, this is subtle and may be due to the difficulty of choosing equivalent damping schemes for the two types of models.

2.4 Discussion

Source stacking, as well as source stacking followed by cross-correlations between pairs of stations provides a promising approach for significantly reducing the number of SEM wavefield computations, at least for resolving structures at the scale presented in this paper. While path-weighting is already included in the implementation, the additional step of windowing allowed by stacking cross-correlations will help achieve sufficient resolution when shorter period body wave constraints on deep mantle structure are included. The uncertainties in source parameters will be reduced when using stacked cross-correlations, just as they are in ANT, and particularly so when using source encoding.

A particularly useful application of this approach is for exploring modelling uncertainties such as due to background noise, as illustrated here, in ways that cannot be addressed using conventional waveform tomography based on the inversion of single records, since perform-

ing full synthetic tests is too costly computationally and generally only the last iteration can be assessed using resolution matrix computations and bootstrapping (e.g. French and Romanowicz 2014) or by sampling the Hessian (e.g. Fichtner and Trampert 2011). This also opens the way to exploring various trade-offs, such as, for example, between crustal structure and radial anisotropy, which are known to be important (e.g. Ferreira et al. 2010; Lekić et al. 2010), as well as uncertainties in source parameters and other trade-offs due to choices of parametrization. Also, the choice of damping schemes generally involves some level of arbitrariness. This could be informed by tests performed on synthetic data sets, and involving several iterations, as illustrated here. At the very least, insights gained in such tests can inform uncertainties associated with conventional waveform inversions. The "BLOB" model considered here is very simple, and the frequency range rather low. In the case of conventional waveform tomography, one could envisage designing more realistic synthetic tests based on the methodology presented here.

A concern with the source stacking approach voiced in the applied geophysics community, and more generally in the community that relies on the computation of gradients using an adjoint formalism, is the introduction of cross-terms when computing the gradient based on the adjoint of the summed source seismograms. This is generally addressed by performing random source encoding which adds to the number of SEM runs but is still worthwhile computationally (e.g. Schuster et al. 2011). Recently, a method has been proposed whereby the encoding is such that it allows the decoding of single traces from the stack at given frequencies (e.g. Zhang et al. 2018), although computational trade-offs still need to be evaluated comprehensively.

In this approach so far, there is no cross-talk that reduces the quality of the partial derivatives, since adjoint formalism is not used for the inverse part of the modelling. Rather, the gradient and Hessian to quite a good approximation are computed using normal mode perturbation theory: as the tests show, the result can converge to a good approximation of the starting model with relatively few iterations starting from a 1-D model. The updates to the gradient and Hessian at each iteration can be rapidly computed.

In the case of real data, another obvious challenge is that not all stations, and not all three components of all stations are generally available. This is the case for the data set used for the construction of SEMUCB-WM1, and not only because of limited operation times of stations, but also because of the careful data selection that was performed, to eliminate noisy data, and in particular nodal stations. However, the data set assembled for the construction of SEMUCB-WM1 spanned an interval of $\sim$20 yrs (from 1992 to 2012) during which broadband networks expanded significantly, so that records for older events were missing at stations installed more recently. When focusing on the last ten years, assembling event and station groups in a way can minimize missing records in any particular event/station group. Once a set of generally high quality stations is selected for performing a particular stack, removing individual noisy records will not be necessary (due for example to a station being nodal for a particular event). Whether or not it is better to keep noisy records or replace them by synthetics computed in the current 3-D model can be assessed by synthetic tests in the near future. Particular care will need to be taken to adjust the model parametrization in

successive iterations (starting with a 1-D model), as well as the period range (progressing from longer periods to shorter periods), in order to avoid undesirable effects of cycle slipping, which will be difficult to detect explicitly in this approach.

2.5 Conclusions

Through a series of experiments based on synthetic waveform data with and without realistic noise added, some of the drawbacks encountered when trying to invert stacked-source seismograms directly may be overcome by redefining the cost-function using station pair cross-correlations of the summed traces. Indeed, large amplitude fundamental surface waves that dominate the latter can be, to a large extent, separated in time in the cross-correlations, allowing implementation of windowing and weighting to bring out overtone and body wave constraints to the inversion for deep mantle structure. Moreover, the use of cross-correlations allows straightforward implementation of path-weighting, an important step in balancing the available coverage in the inversion step, as well as wave-packet weighting (i.e. windowing), which is important to bring out the contributions of lower amplitude body waves.

2.6 Figures

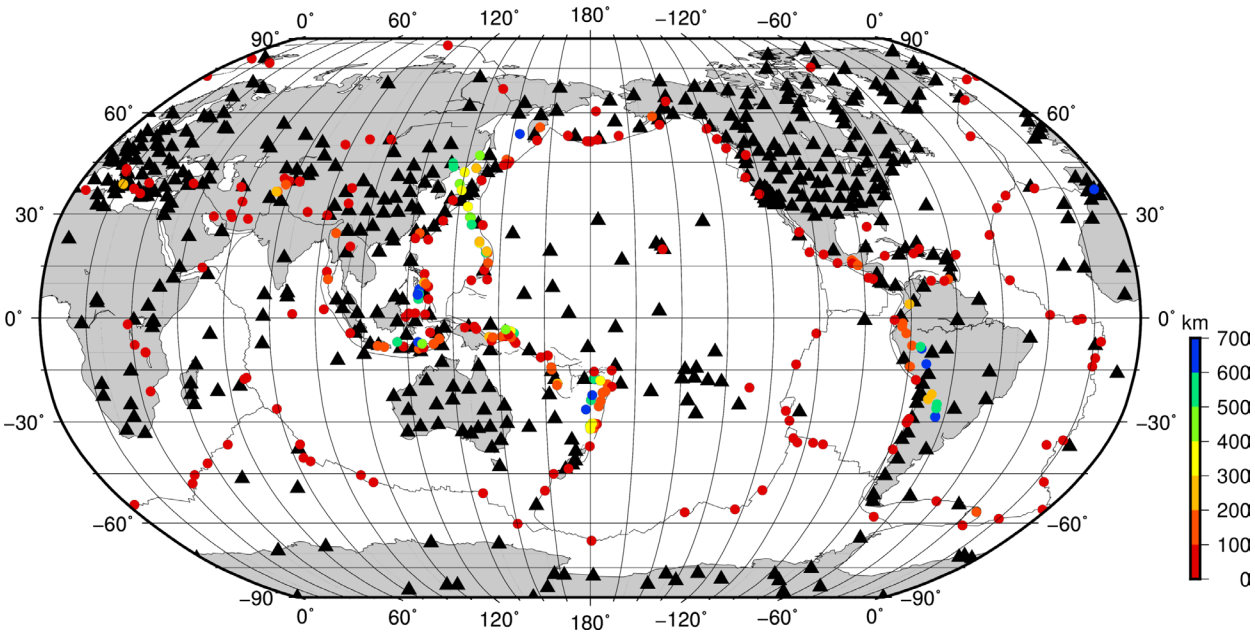

Figure 2.1: The distribution of events and stations used for the synthetic experiments presented in this paper. It consists of the 273 earthquake sources and 537 stations used for the development of global mantle model SEMUCB-WM1. The colours indicate the depths of sources.

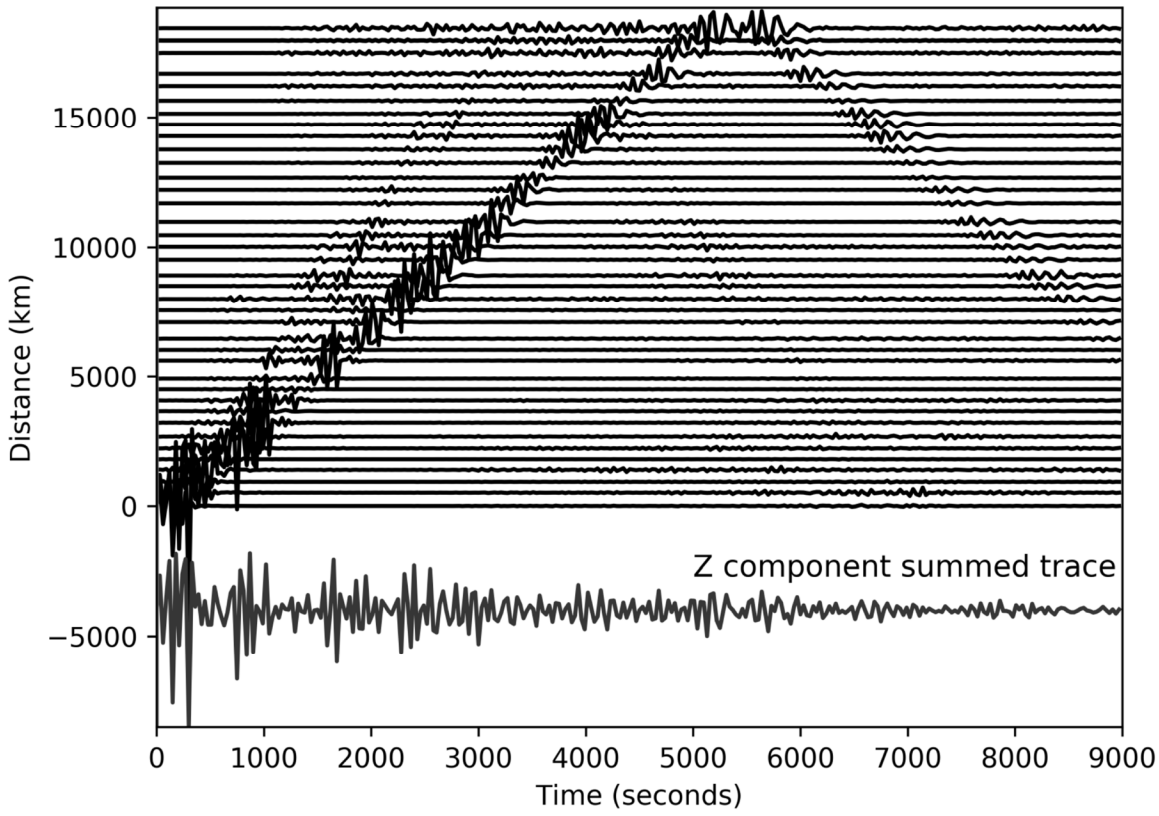

Figure 2.2: An example of vertical component 10 000-s long traces calculated at station TATO in the 1-D average model derived from SEMUCB-WM1 (French and Romanowicz 2014). 38 different events are shown here, aligned on origin time and ordered as a function of epicentral distance. The frequency band of the calculation is 400–60 s. The bottom trace is the sum of the 38 aligned traces.

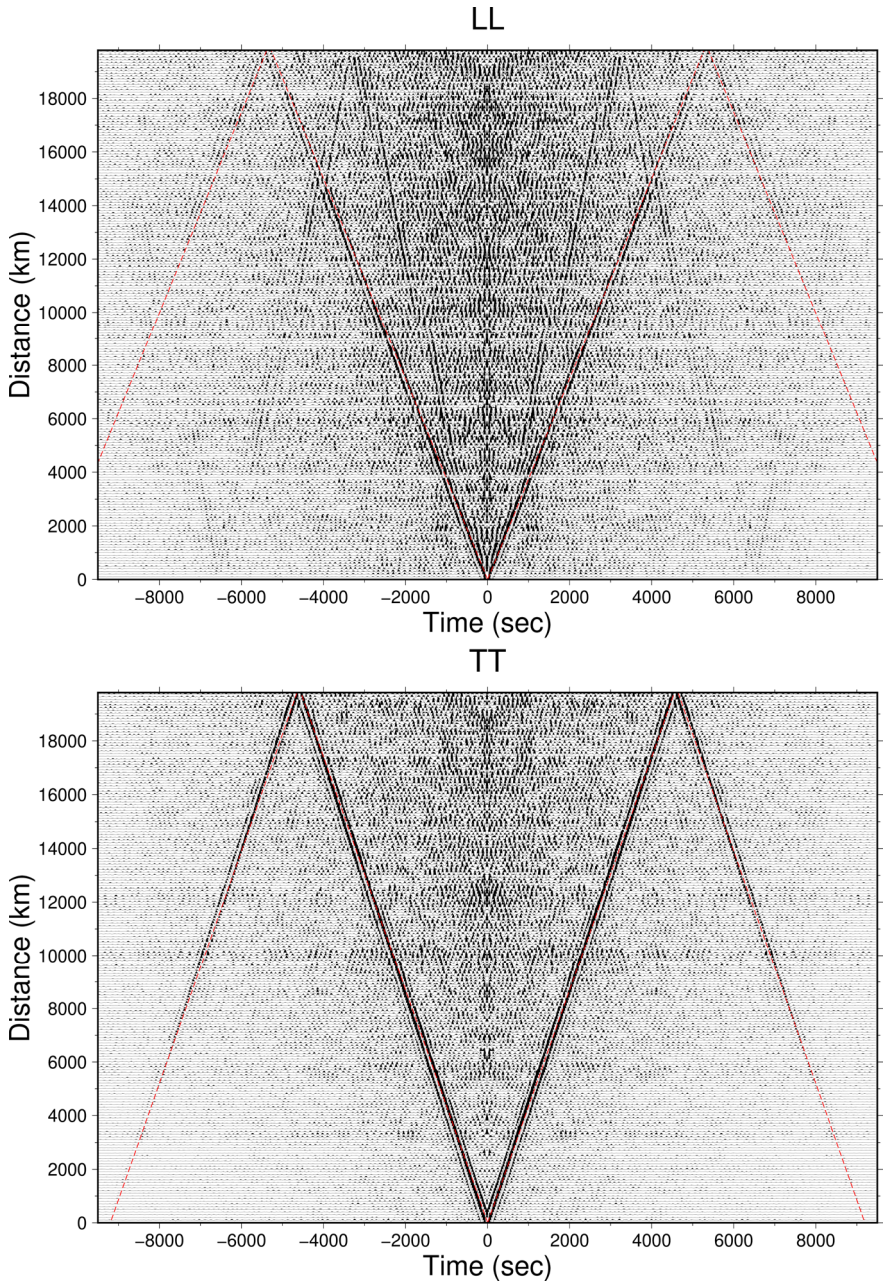

Figure 2.3: Stacked L–L (top panel) and T–T (bottom panel) cross-correlations between station pairs for the entire synthetic data set. The strongest energy corresponds to group velocities of 3.75 km s^{-1} for L–L and 4.35 km s^{-1} for T–T, which are in the middle range of group velocities for fundamental mode Rayleigh and Love waves, respectively.

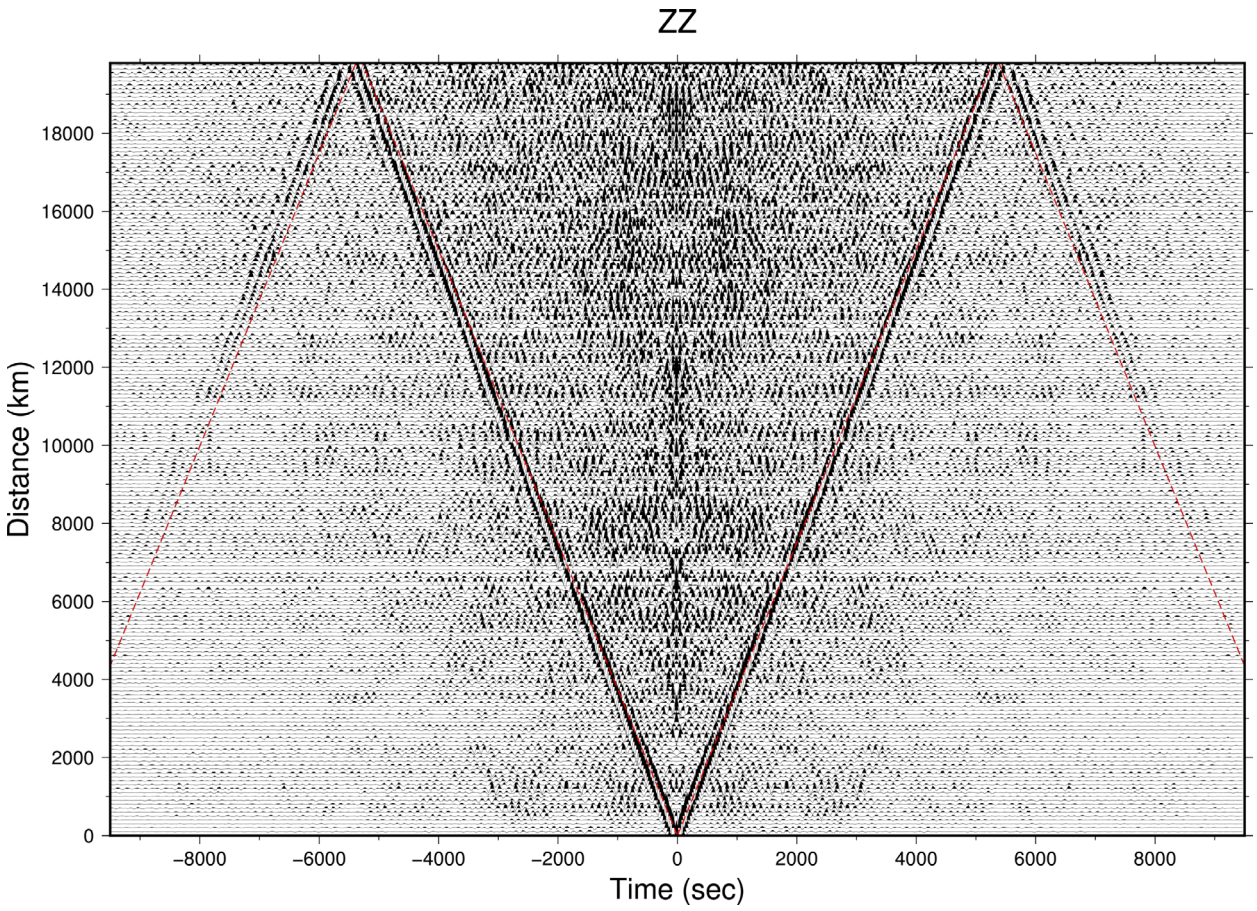

Figure 2.4: Z-Z component of stacked station pair cross-correlations for the entire dataset, arranged as a function of inter-station distance: The red lines indicate the group velocity of 3.75 km/s that corresponds to the middle of the range of Rayleigh wave group velocities.

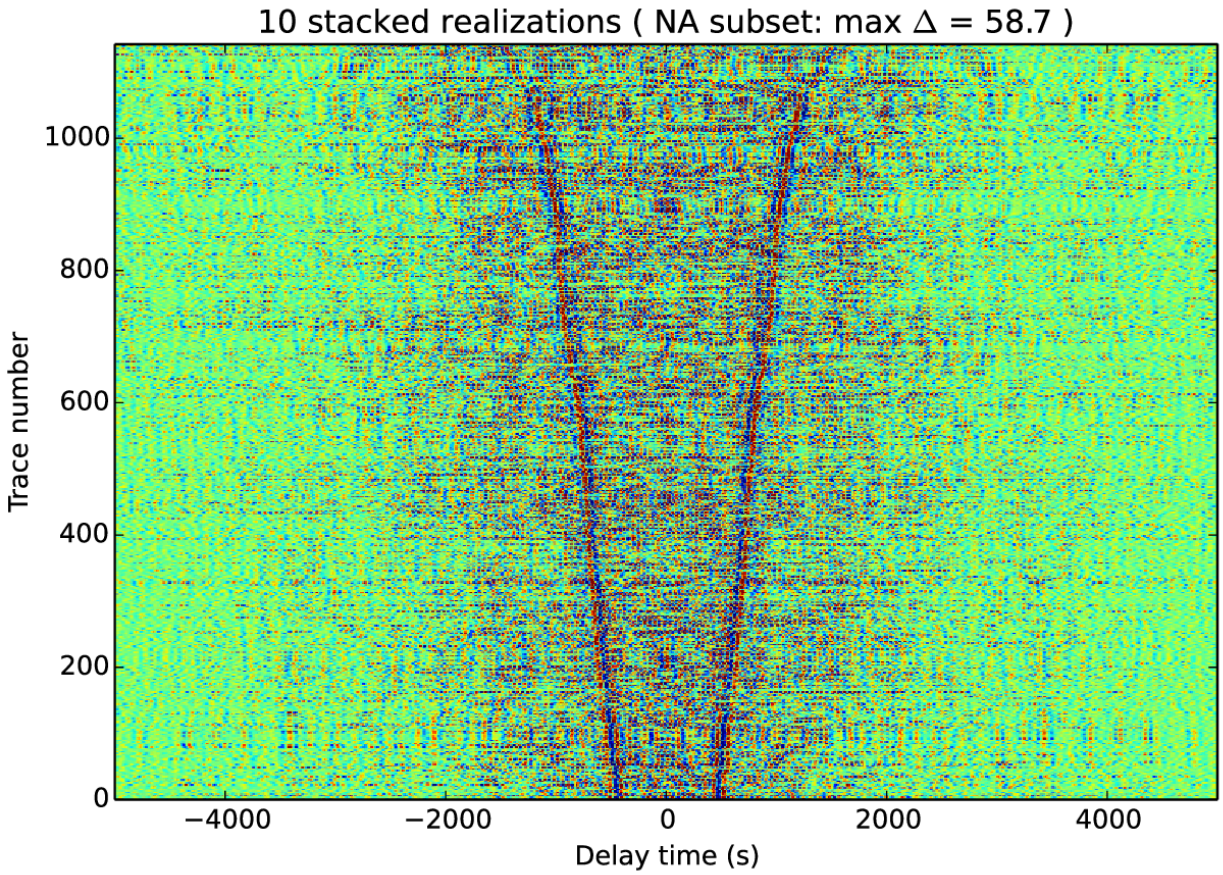

Figure 2.5: Record section of the Z-Z station pair cross-correlations, arranged as a function of distance, of 10 stacked realizations of source-encoded summed synthetics, for a subset of stations located on the north American continent. The maximum distance between stations is 58.7^o. The emergence of strong energy corresponding to the fundamental mode Rayleigh wave is clearly visible. The cross-correlations have been calculated for a time interval that includes the first orbit Rayleigh wave train (according to the theoretical Green's function). Figure courtesy of Dr. Scott French.

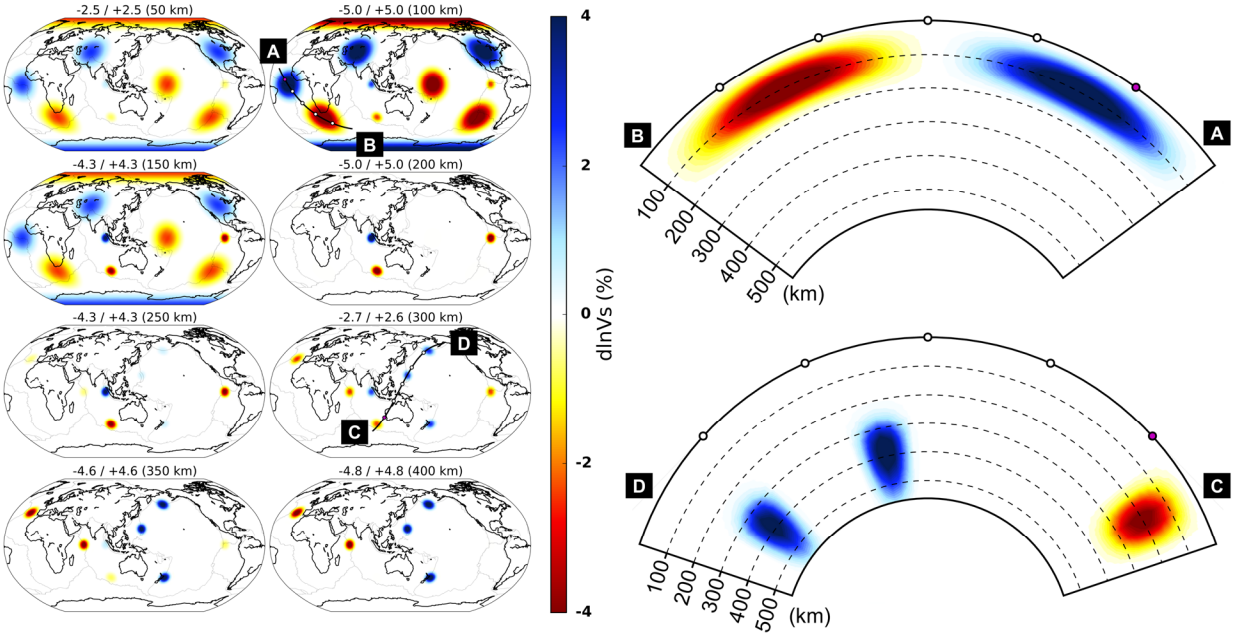

Figure 2.6: Isotropic Vs part of the synthetic "BLOB" upper mantle model used as input model for the synthetic experiments discussed in this paper. Left-hand panels show the isotropic Vs model at several depths, presented in percent with respect to the global mean at that depth. The extrema in Vs are shown above each map. Right-hand panels present depth cross-sections along two profiles $A - B$ and $C - D$ as indicated on the 100 and 300 km maps, respectively.

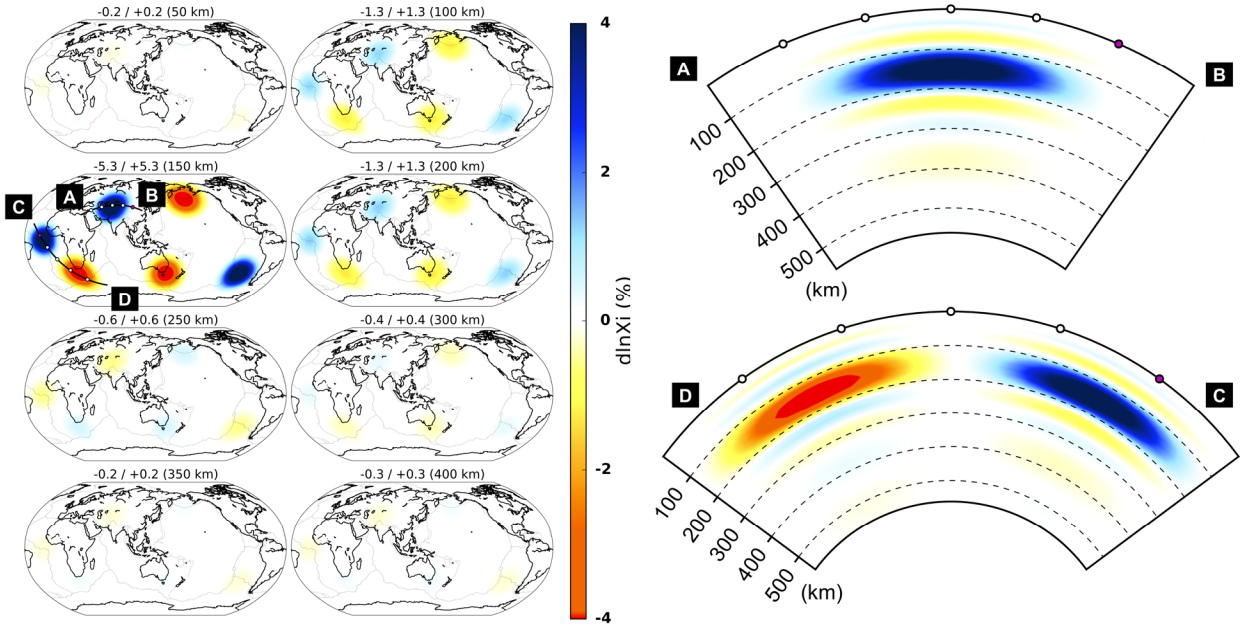

Figure 2.7: Anisotropic parameter ξ variations in the "BLOB" model, plotted in map view at several depth (left panel), in percent relative to the global mean (isotropic) at that depth. Extrema of each map are shown above the corresponding map. Right panels present depth cross-sections along two profiles $A - B$ and $C - D$ as indicated on the 150km map.

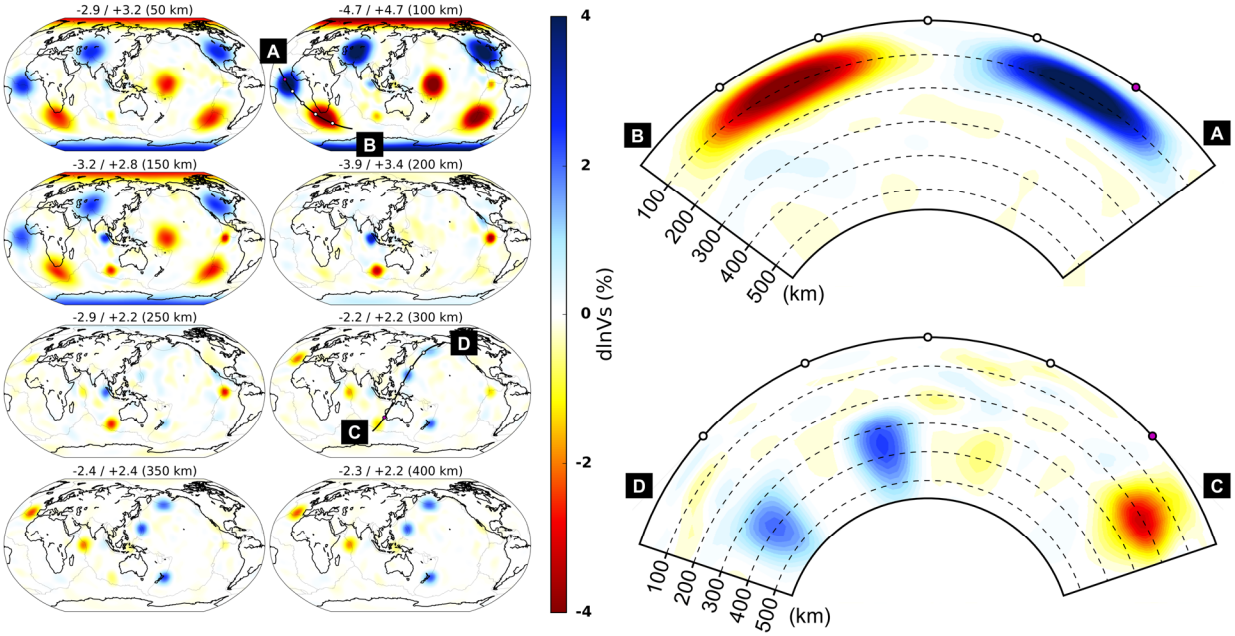

Figure 2.8: The recovered model in isotropic Vs after 6 iterations of direct inversion of 3 component summed seismograms, without noise. Plots are presented as in Figure 2.6.

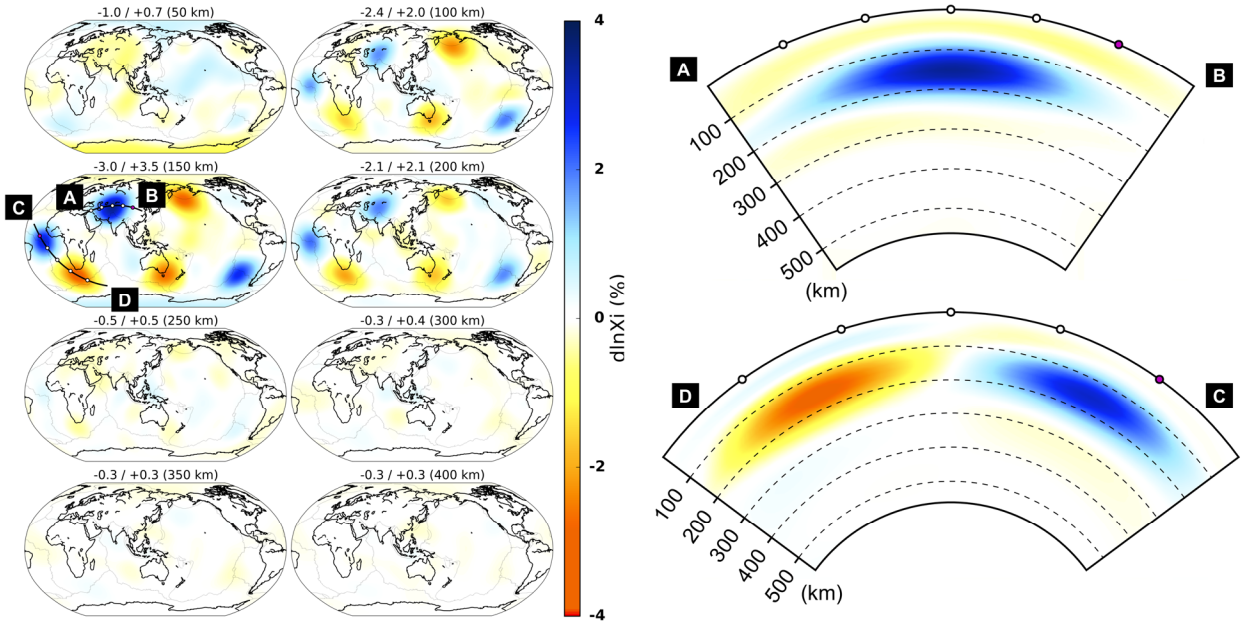

Figure 2.9: Same as Figure 2.8 for the parameter ξ.

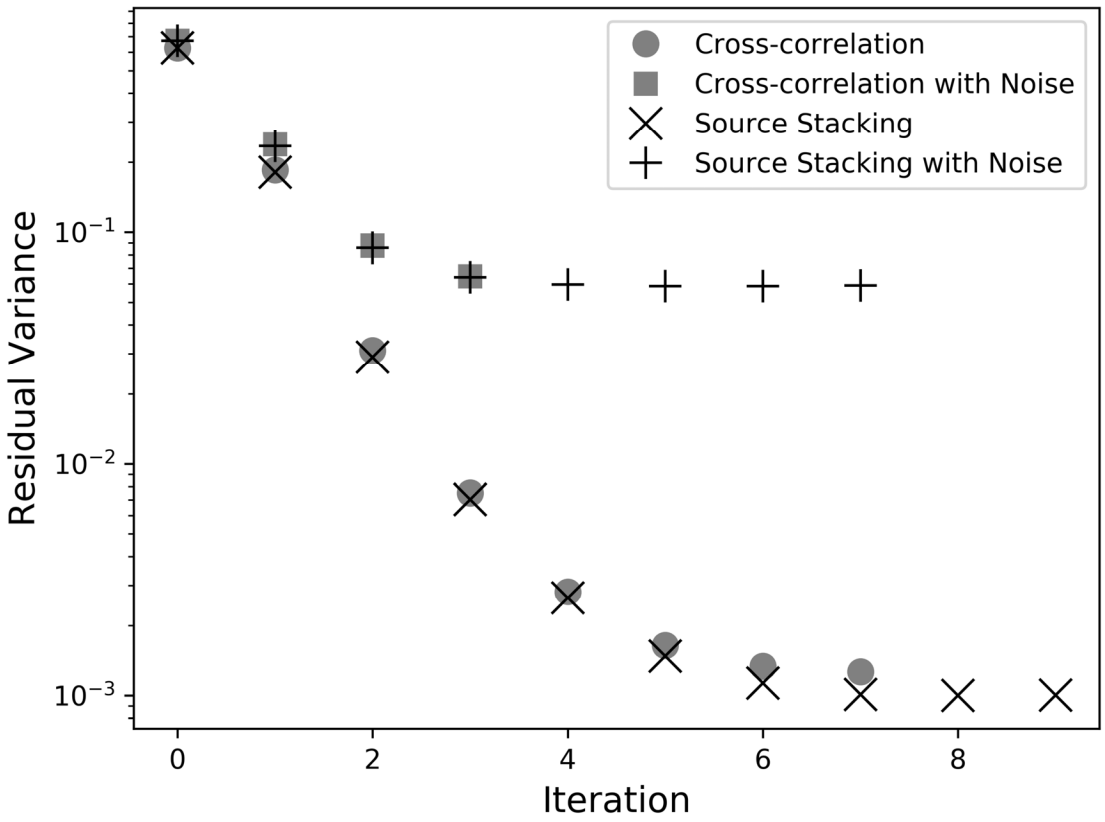

Figure 2.10: The evolution of variance reduction as a function of iteration for each of the experiments performed. Note that for iteration 1 with noise, the residual variances for source-stacking inversion and when the additional cross-correlation step is included, plot on top of each other.

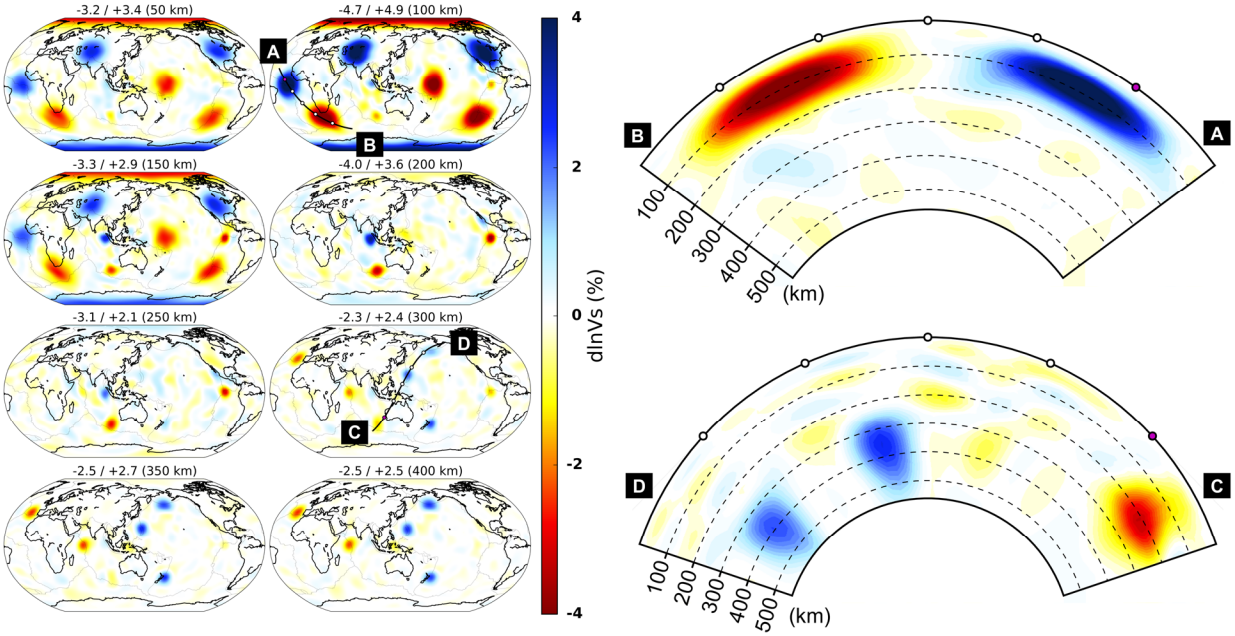

Figure 2.11: Isotropic Vs part of the model recovered after 6 iterations of inversion of station pair cross-correlations, without noise added.

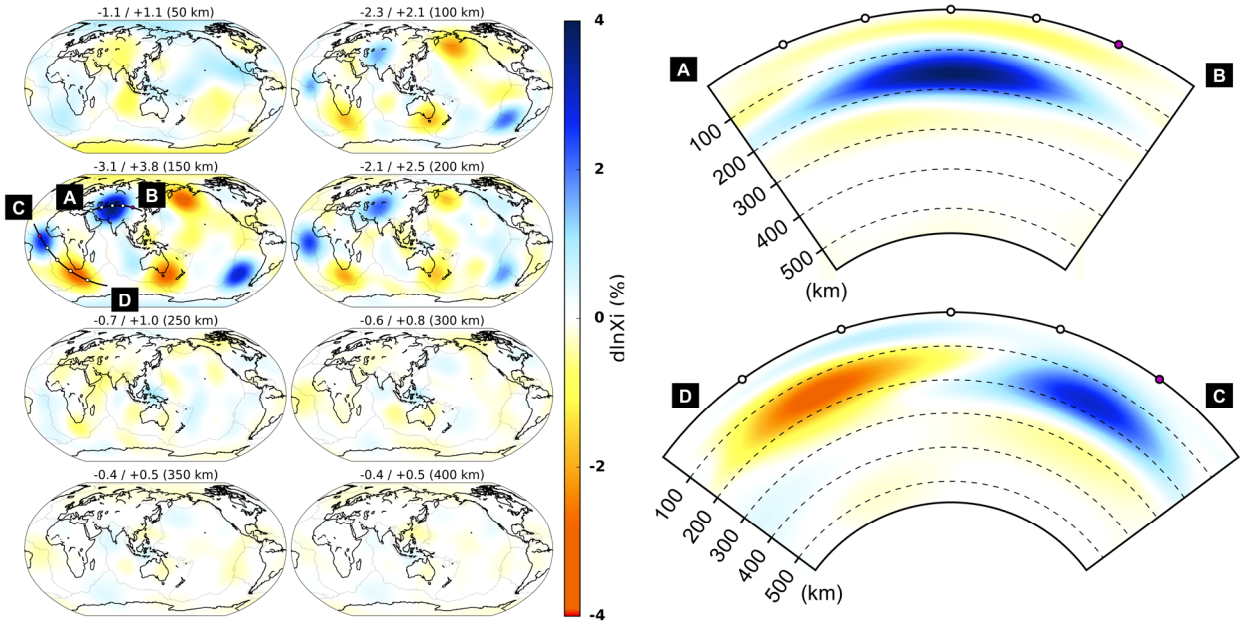

Figure 2.12: Same as Figure 2.11 for the parameter ξ.

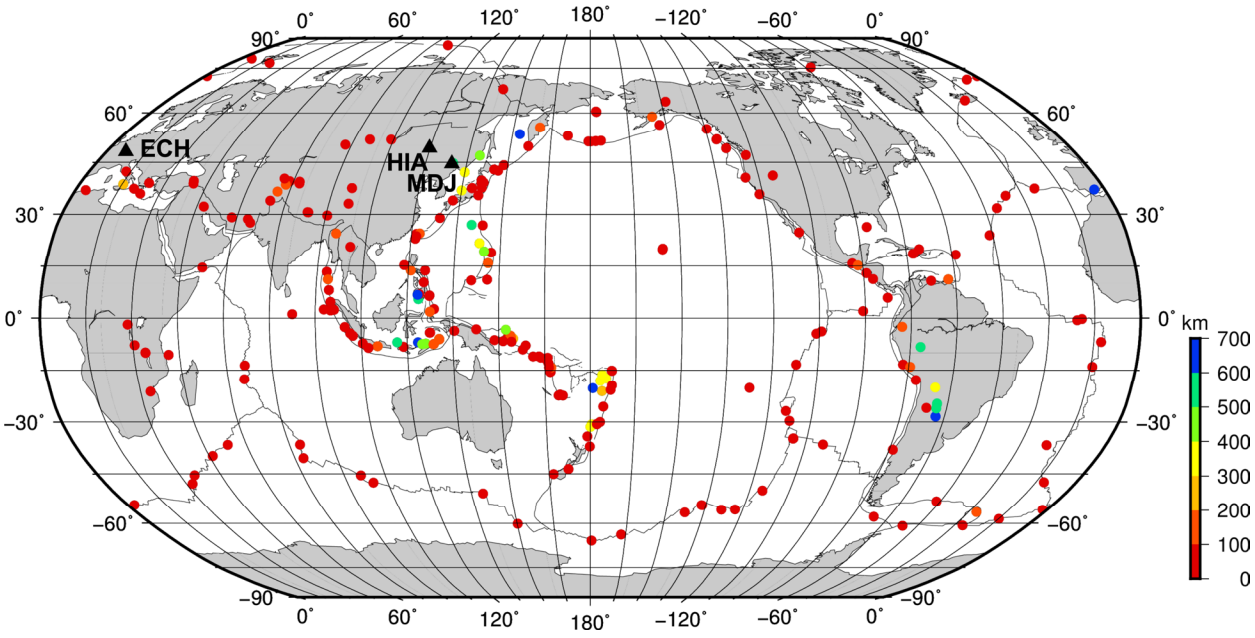

Figure 2.13: The distribution of 179 events and location of the three stations ECH, HIA and MDJ. For these stations, 10,000 s of noise waveforms on the vertical component are collected before the origin time of each of the 179 events jointly recorded by both stations and compared them to 10,000 s of real data starting at the origin time of each event.

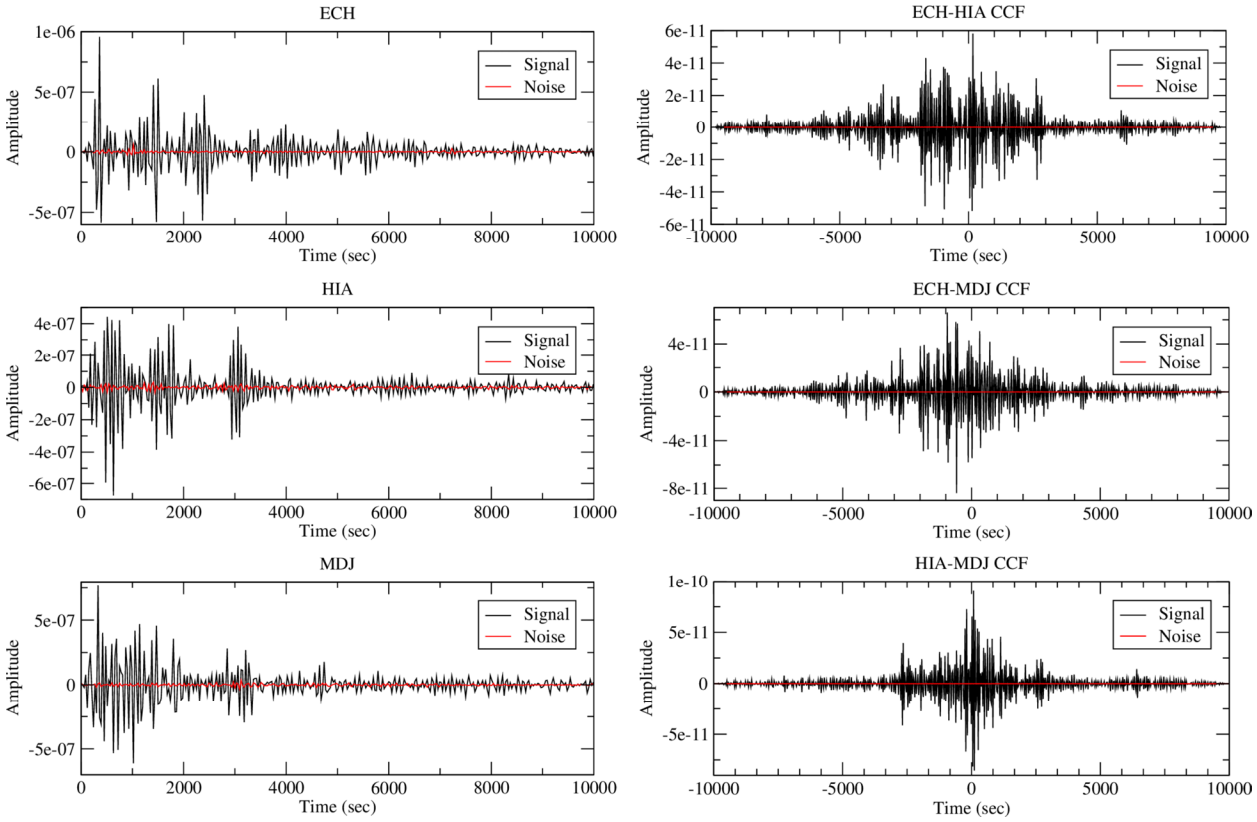

Figure 2.14: Left-hand panel: examples of comparisons of 10,000-s-long summed 'noise' and 'signal' records. Black traces show the summed signal. Red traces show summed noise. Right-hand panels: comparison of the station pair cross-correlations of summed noise (red traces) and signal (black traces) records for the three stations shown on the left-hand panels. The traces have been filtered with corner frequencies of 60 and 400 s. The distribution of the 179 events and location of the three stations for which records are shown is presented in Figure 2.3.

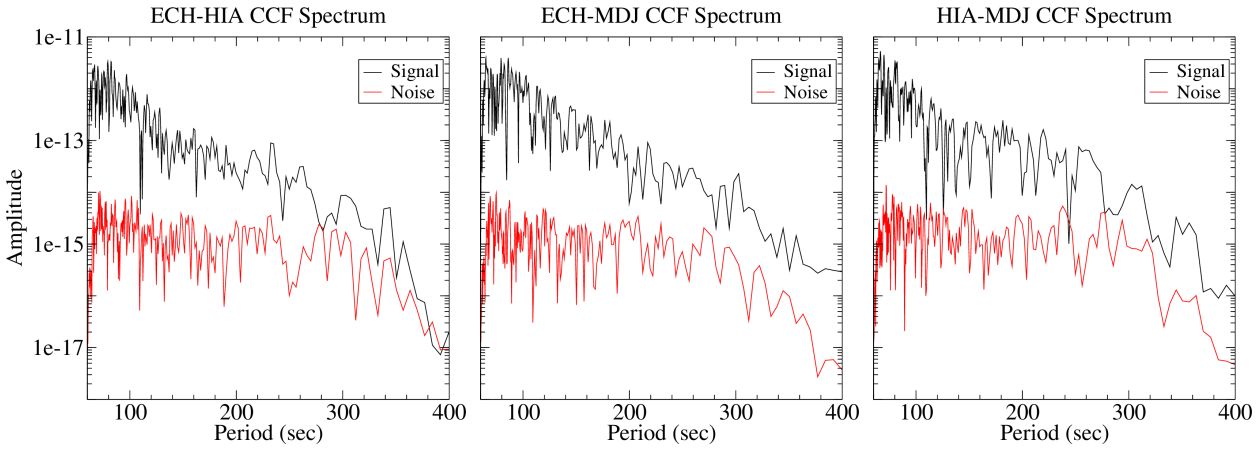

Figure 2.15: Comparison of the spectra of the cross-correlations traces shown in Figure 2.14. The traces have been filtered with corner frequencies of 20 and 800 s. Note the logarithmic scale on the y-axis.

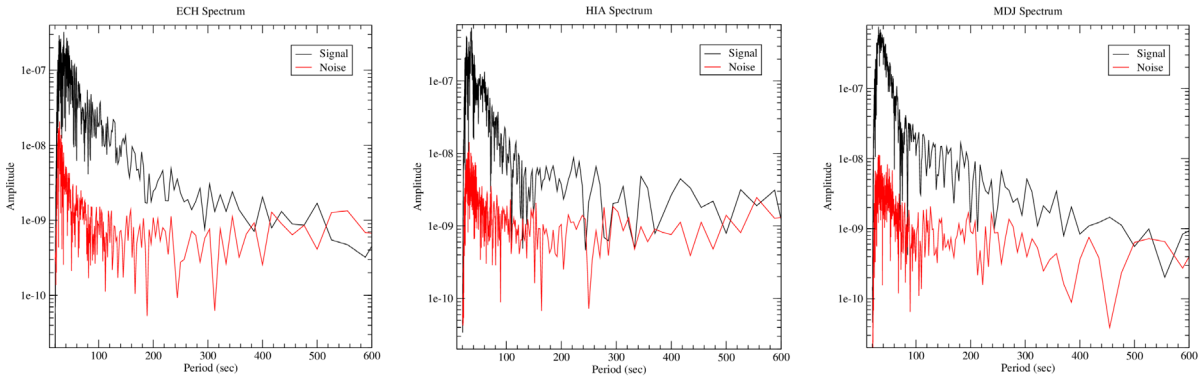

Figure 2.16: Comparison of signal (black) and noise (red) spectra for the summed synthetics shown in Figure 2.14.

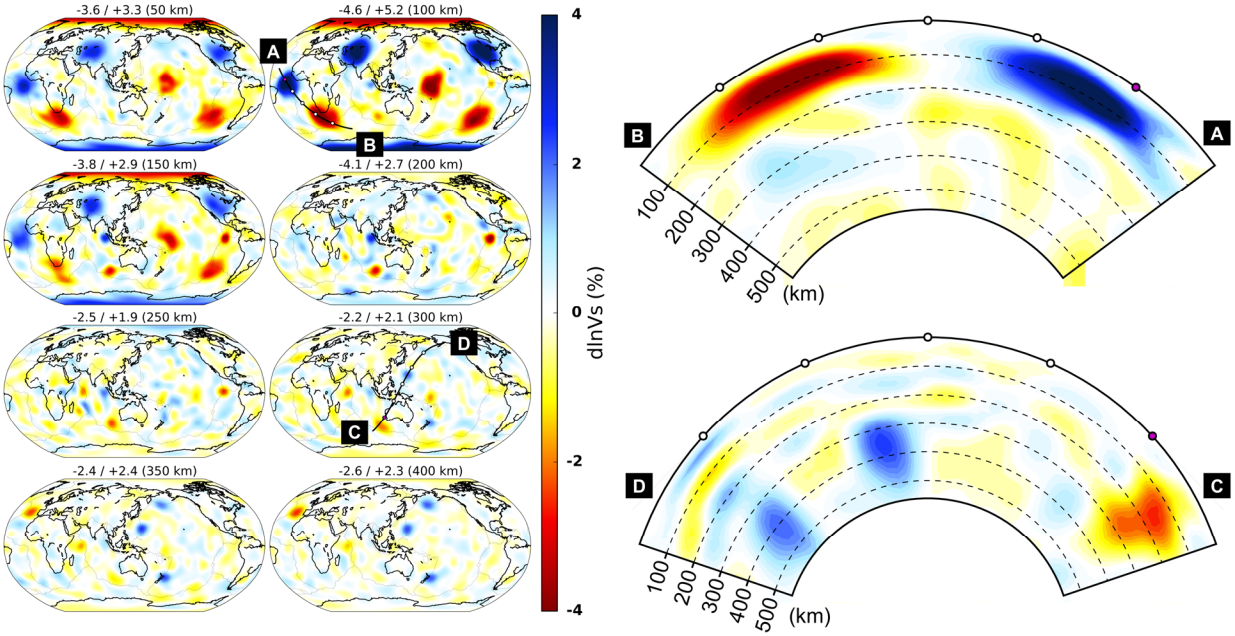

Figure 2.17: Vs isotropic part of the model recovered after six iterations of inversion of three component summed seismograms, with noise added. The target model is presented as in Figure 2.6.

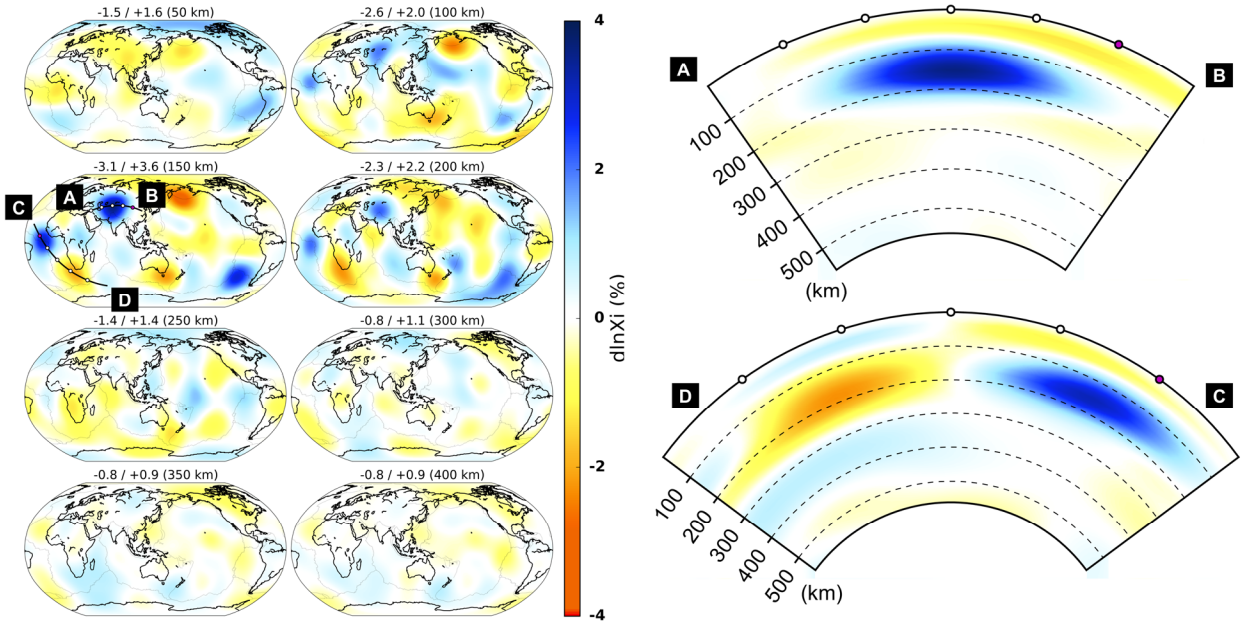

Figure 2.18: The ξ part of the model recovered from the same experiment as shown in Figure 2.17.

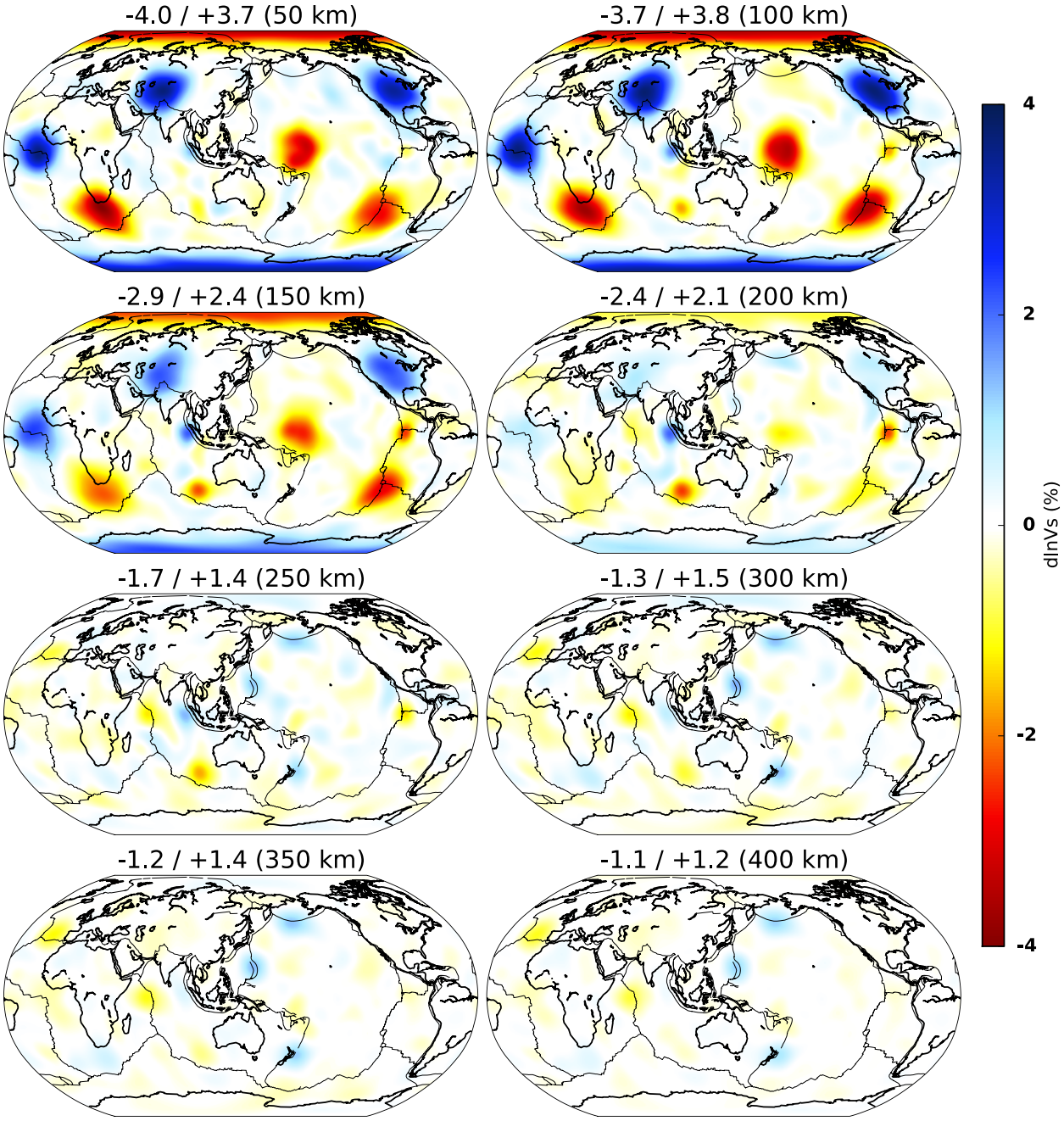

Figure 2.19: Results of three iterations of inversion of the summed synthetics with noise, using the cross- correlations to define the misfit function: recovered Vs model.

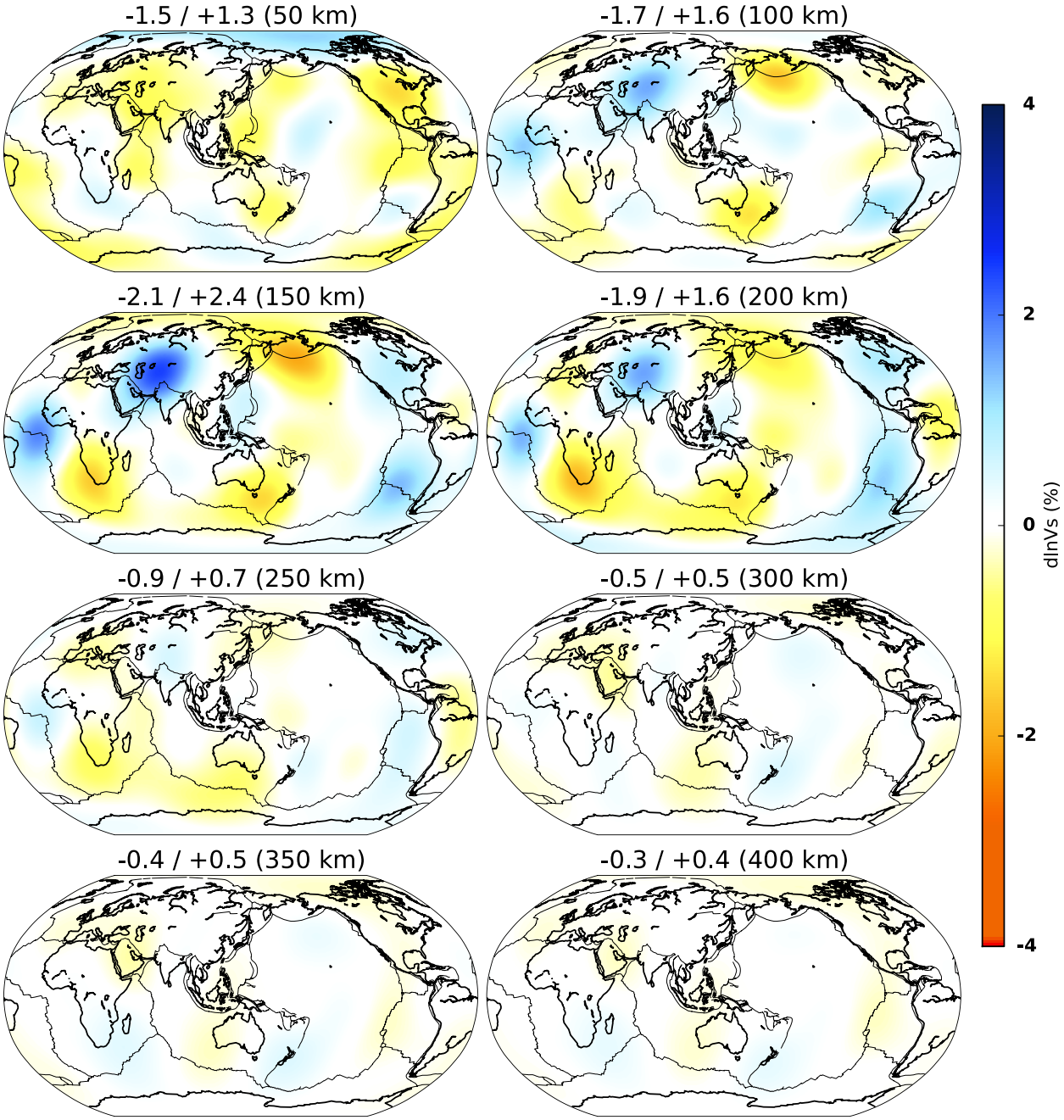

Figure 2.20: Same as Figure 2.19 for the anisotropic parameter ξ.

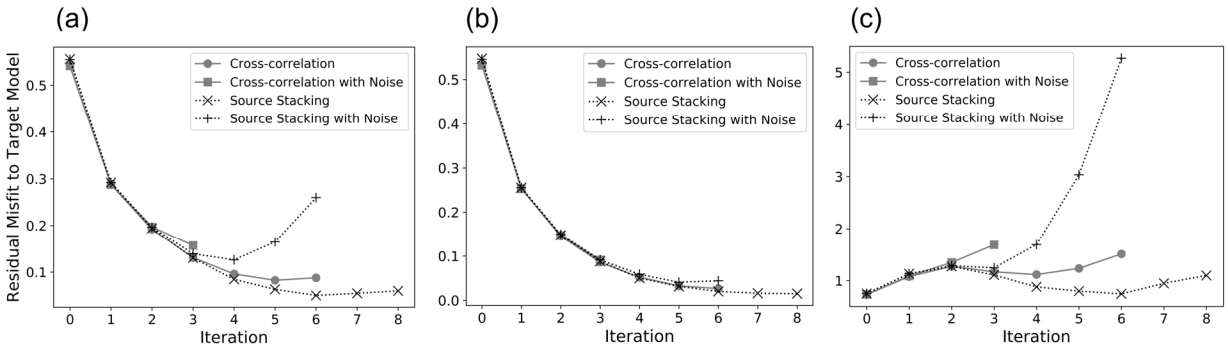

Figure 2.21: Residual Vs model misfit (normalized by variance in the target model), as a function of experiment and iteration when (a) the whole model volume down to 550 km depth is considered in the calculation of the misfit; (b) only the portion of the model space where dlnVs $>$ 1 per cent in the target model is considered in the calculation of the misfit (this corresponds to 7.9 per cent of the total volume); (c) only the portion of the model space where dlnVs $<$ 1 per cent is considered in the calculation of the misfit (this corresponds to 92.1 per cent of the total volume and is dominated by regions without 3-D structure in the "BLOB" model.

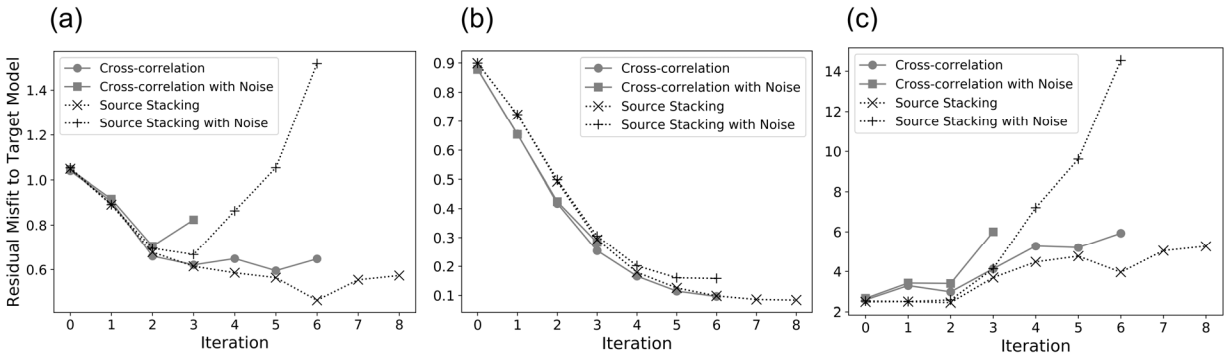

Figure 2.22: Residual ξ model misfit (normalized by variance in the target model), as a function of experiment and iteration when (a) the whole model volume down to 550 km depth is considered in the calculation of the misfit; (b) only the portion of the model space where dlnVs $> 1\%$ in the target model is considered in the calculation of the misfit (this corresponds to 2.1% of the total volume); (c) only the portion of the model space where dlnVs $< 1\%$ is considered in the calculation of the misfit (this corresponds to 97.9% of the total volume and is dominated by regions without 3D structure in the "BLOB" model.

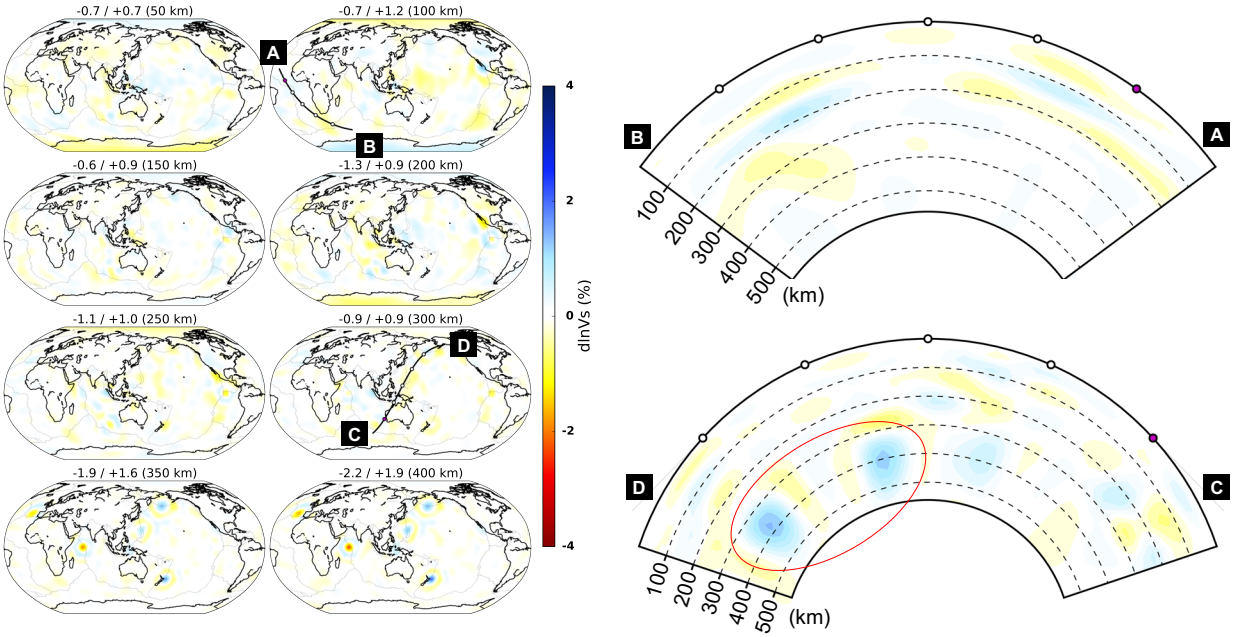

Figure 2.23: Difference between target "BLOB" Vs model and model obtained after 6 iterations of inversion of stacked waveforms without noise. Note that the color palette has been chosen for direct comparison with the recovered model as shown in Figure 2.20. The red ellipse in the cross-section shown on the bottom right highlights the region where smaller scale high velocity anomalies are present.

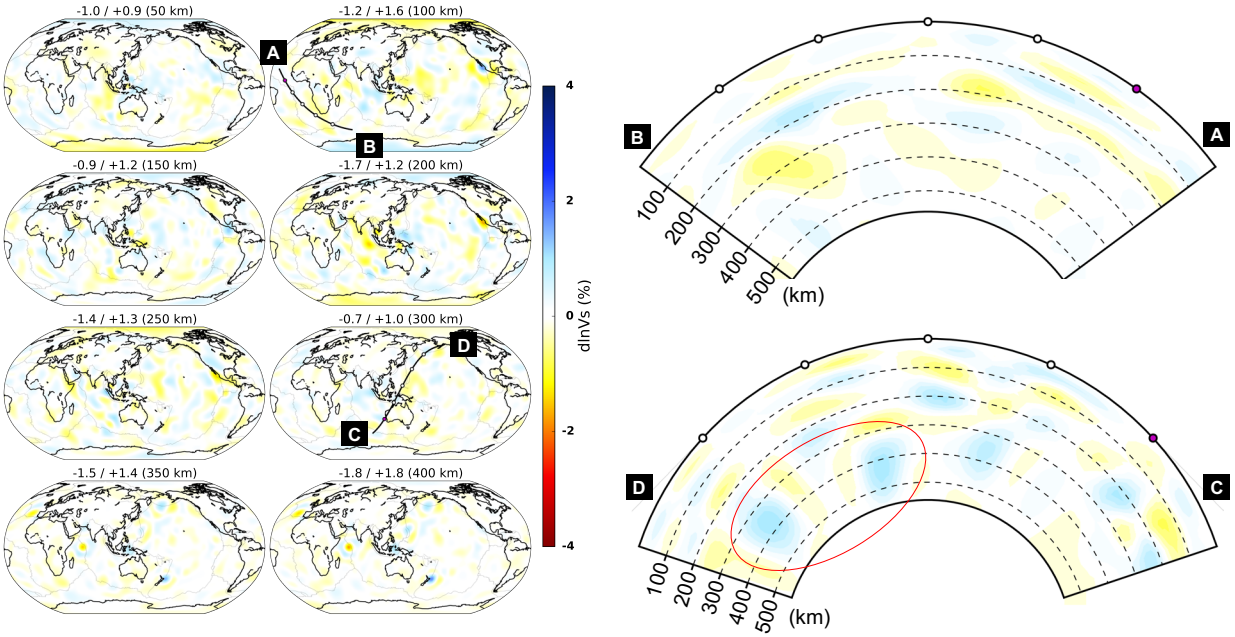

Figure 2.24: Same as Figure 2.23 for the inversion with cross-correlation based misfit function. Note that in the region highlighted by the red ellipse, the residual misfit is somewhat smaller than in the case shown in Figure 2.23.

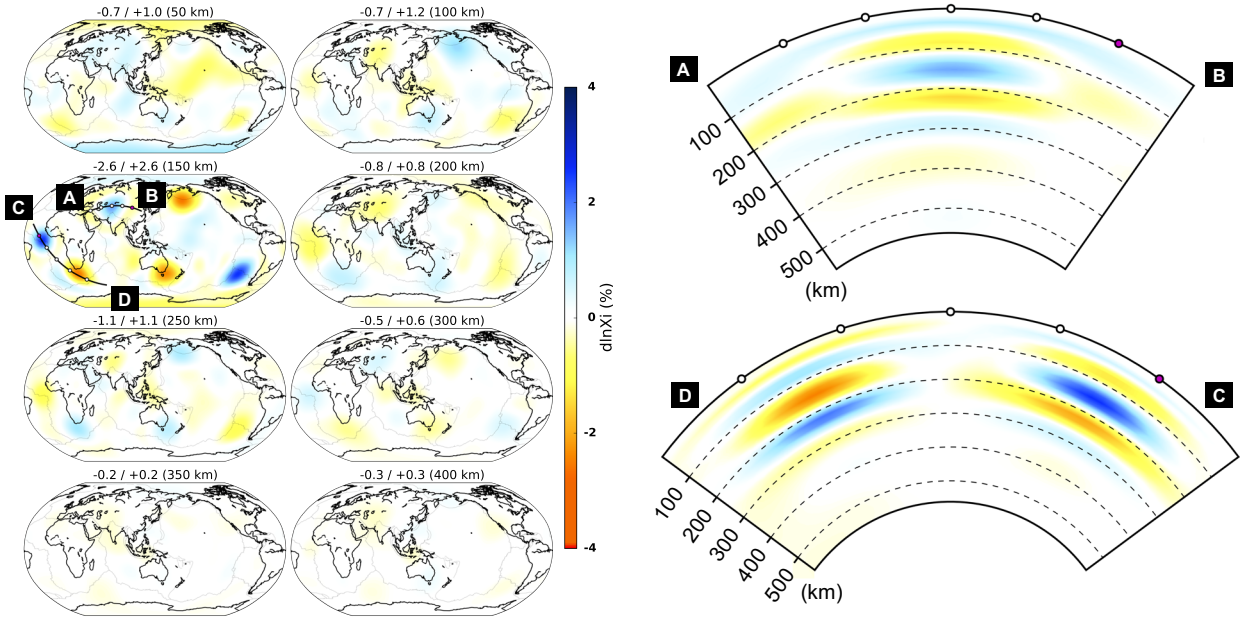

Figure 2.25: Same as Figure 2.23 for the anisotropic parameter ξ.

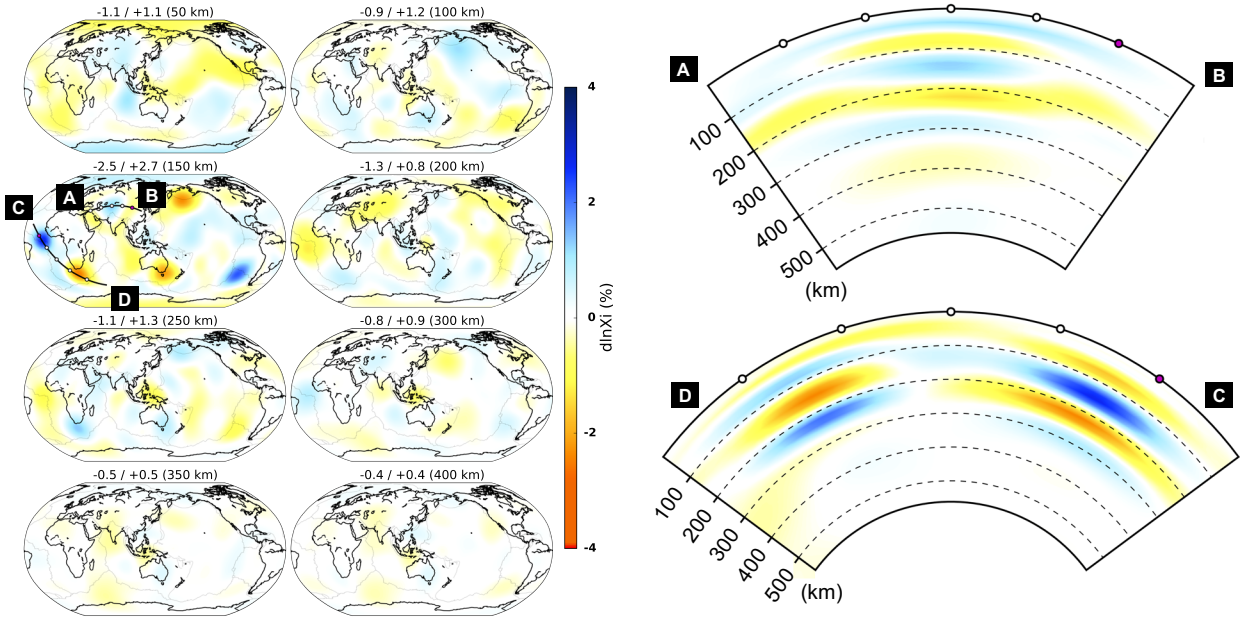

Figure 2.26: Same as Figure 2.24 for the anisotropic parameter ξ.

Chapter 3

Building an equivalent model of the crust and uppermost mantle constrained by surface wave dispersion for whole mantle full-waveform imaging

---

modified from "Building an equivalent model of the crust and uppermost mantle constrained by surface wave dispersion for whole mantle full-waveform imaging" (Chen and Romanowicz, 2021, submitted to GJI).

3.1 Introduction

Global seismic tomography has been developed for more than four decades and has revealed increasingly detailed features in the earth. From the earliest 3-D global images of the lower (e.g., Dziewonski et al. 1977) and upper mantle (e.g., Woodhouse and Dziewonski 1984; Nataf et al. 1986), several generations of global seismic models have been constructed with different approaches (e.g., travel times or waveforms) and resolution has steadily increased over time, but only long-wavelength structures (> 2500km) are consistent across models (e.g., Becker and Boschi 2002; Lekić and Romanowicz 2011). To obtain a high-resolution full-waveform tomographic model, the effects of crustal and uppermost mantle on the waveforms have to be carefully accounted for. Yet 3D crustal structure is not accurately known and this becomes a challenge when the Spectral Element Method (SEM; e.g., Komatitsch and Vilotte (1998)) is used for predicting the teleseismic wavefield in full-waveform inversion (FWI). It has been shown that even the most recent existing global crustal models, such as crust1.0, do not predict the teleseismic wavefield accurately, at least in some parts of the globe (e.g., Meier et al. 2007; Pasyanos and Nyblade 2007). Another issue is that the presence of thin slow layers (i.e., sediments) in the crust slows down already heavy computations. Another one approach to address crustal effects is to directly invert for the crust perturbations to an existing 3D crustal model using high-frequency short period waveforms (e.g., down to 17s period) by adjoint tomography (e.g., Fichtner et al. 2013; Rickers et al. 2013; Zhu et al. 2015; Bozdağ et al. 2016; Lei et al. 2020). Fully utilizing waveforms ensures the consistency of inversion between the crust and the mantle. But the inversion for of short-period high-frequency waveforms is computationally costly and the process requires proper pre-conditioning and smoothing for the misfit of the gradient. Notably, the presence of thin slow layers (i.e., sediments) in the crust slows down already heavy SEM computations. An alternative approach is to construct a smooth crustal model constrained by global dispersion data, using the concept of homogenization (e.g., Backus 1962; Capdeville and Marigo 2007), which avoids thin layers in the crust, accelerates the forward modeling computations, and may be more uniformly reliable across the globe than existing crustal models, at least in the relatively long period range considered for inversion. This approach has been used in the development of several recent models (e.g., SEMum, Lekić and Romanowicz (2011), SEMum2, French et al. (2013); SEMUCB-WM1, French and Romanowicz (2014)), where group velocity data in the period range 25-60s were used to constrain a smooth "equivalent" crustal model, valid for periods longer than 30s. In addition to constraining the crust, dispersion data can be used to help constrain uppermost mantle structure, to which long-period waveforms are also not sensitive enough, to provide stable and reliable 3D shear velocity models at depths shallower than 60-80 km. In order to address this, in the construction of upper mantle models SEMum and SEMum2, long period (T>60s) fundamental and overtone surface waveforms were used in combination with dispersion data in the period band 25-150s, to constrain the top 400 km of the earth (see details in French and Romanowicz (2014)). However, with this approach, sensitivity kernels of waveforms and

dispersion data are not computed in a consistent manner and finding an optimal weighting of the two datasets was not straightforward. Also, while the crustal model was updated at each iteration, it was found not to evolve much, while adding computational time. The mismatch of the sensitivities of waveform and dispersion data introduced bias and caused unrealistic variations of radial anisotropy at the top of the uppermost mantle (e.g., Fig. 3.1). One could thus ask whether these discrepancies could have contaminated the deeper portion of the resulting tomographic model.

In this study, the ability of model SEMUCB-WM1 (French and Romanowicz (2014)) to fit dispersion data is tested in a slightly broader frequency range than used in its development (20-150s instead of 25-150s), as well as different portions of the time-domain records that contain different seismic phases (i.e. first and second fundamental and overtone wavetrains down to 60 s period, different body waves such as S,SS, SSS sensitive to upper mantle structure down to 32 s period). The fits obtained for SEMUCB-WM1 are compared to those predicted for several models, either based entirely on fundamental mode dispersion measurements, or "hybrid" models as will be described below. This leads to propose a new workflow for full-waveform tomography, that can take advantage of constraints provided directly by independent data (in the present case, dispersion measurements) for the construction of a crust and uppermost mantle model, fix its upper portion and use that as a starting model for FWI at larger depths. The homogenization procedure used implies that the independent dataset needs to match the period band of the subsequent FWI. In the case considered here, the target frequency cut-off for FWI is 40 s for fundamental mode and overtone surface waves and 30 s for body waves. However, this can be extended to shorter periods, provided robust independent data are available globally at those periods to constrain the crustal model. The advantage of this approach is that the uppermost part of the model is constructed first without lengthy SEM computations, while the deeper portion of the model can then be iteratively constrained using FWI, saving computational time while improving the robustness of the mantle model.

In what follows, the construction of such a starting model based on dispersion data are presented. This also allows to evaluate how well a model based solely on fundamental mode dispersion data can fit teleseismic waveforms. The discussions of the dispersion data, the new workflow for constructing the dispersion-based models, as well as two models obtained by combining the latter with SEMUCB-WM1 in different depth ranges will be presented in section 3.2. Also, section 3.3 shows how to assess model performance, which includes not only comparison to dispersion data, but also comparison of the predicted wavefield computed using SEM to observed waveforms. In section 3.4, these results show that fundamental mode dispersion-based models provide acceptable fits to waveforms only when they are replaced by SEMUCB-WM1 at depths larger than $\sim$80 km. Section 3.5 concludes with proposing an improved, potentially versatile approach to taking into account uppermost mantle structure in full-waveform tomography while limiting computational costs.

3.2 Data and method

3.2.1 Dispersion data

Surface wave dispersion data are used to constrain the variations of crust and uppermost mantle shear wave velocity (V_s and radial anisotropy (ξ). The group velocity dispersion dataset is from the Colorado group (e.g., Shapiro and Ritzwoller 2002, Ritzwoller, personal communication 2009), provided in maps at a range of discrete periods between 16 and 200s. In particular, dispersion data at the shortest periods (e.g., 16-60s) are more sensitive to crustal structure. The group velocity data is preferred here because it is available at these short periods globally, and more sensitive to shallow depths than phase velocity data at the same frequency.

Following the same procedure for developing the homogenized crust as used in French and Romanowicz (2014), hereafter referred to as FR2014, the group velocity dispersion data is considered in the period range 20-60s (compared to 25-60s in FR2014) and then further the uppermost mantle down to 471 km depth is inverted using only dispersion data in the period range 20-150s (this step is different from FR2014).

3.2.2 Model parameterization

A transversely isotropic medium, which is described by five independent elastic moduli $(V_p, V_s, \xi = V_s v^2/V_s h^2, \phi = V_p v^2/V_p h^2, \eta)$, is considered to model the propagation of seismic waves. These are reduced to two parameters (V_s, ξ) by introducing the scaling relationships of Montagner and Anderson (1989) for the other three parameters and density, to which surface waves are less sensitive. Further details can be found in appendix A of Panning and Romanowicz (2006) and FR2014.

For the crust, and in order to reduce the cost of the future SEM computations in the FWI, as in FR2014, the geologically plausible layered crustal model Crust 2.0 is replaced by by a smooth anisotropic crustal model, in which the Moho depth is kept at that of Crust 2.0 if the crust is thicker than 30 km, and fixed to 30 km in regions where it is thinner (e.g., most ocean basins, and a few continental regions). The smooth V_s and ξ profiles within the crust are parametrized using Lagrange polynomials with 5 Gauss-Lobatto-Legendre (GLL) interpolation points, appropriate for implementation in the SEM. They are constrained at each geographical location of the spherical spline nodes by fitting to Rayleigh and Love group velocity dispersion data at that location. This approach is inherited from the crust construction of SEMUCB-WM1 (FR2014), but the period range of the dispersion data is extended down from 25s to 20s. As extensively discussed and demonstrated in the previous studies (e.g., Lekić and Romanowicz 2011, FR2014), for waveforms of periods longer than 30s, the seismic response of this model is "equivalent" to that for a realistic crust that would fit the same dispersion data at all locations.

In the mantle, the 3D perturbations of Voigt-average V_s and the radial anisotropy parameter ξ are referenced to a 1D global average model, modified from that of SEMUCB-WM1,

which has a crustal thickness of 30 km, and are described using cubic b-spline basis functions in depth (e.g., Mégnin and Romanowicz 2000) and spherical splines laterally (e.g., Wang and Dahlen 1995). The perturbations in elastic parameters $X \in \{V_s, \xi\}$ can then be described as a function of radius r, colatitude θ and longitude ϕ by a discrete set of coefficients C^*_{pq} as:

$$\delta lnX(r,\theta,\phi) = \sum_{p}^{Np} \sum_{q}^{Nq} C^*_{pq} \beta_p(\theta,\phi) v_q(r) \tag{3.1}$$

where $\beta_p(\theta,\phi)(p = 1, N_p)$ and $v_q(r)(q = 1, N_q)$ are elements of the lateral spherical-spline basis and radial cubic b-spline basis, respectively.

The following radii of cubic b-spline knots $v_q(r)$ are considered in all inversions presented here: 5900, 6050, 6100, 6150, 6200, 6250, 6270, 6290, 6305, 6320, 6331, and 6341 km, as shown in Fig. 3.2(a). Close to the Moho, the distribution of knots is denser than the parametrization of SEMUCB-WM1, to better accommodate the transition between the crust and the uppermost mantle, where there are significant lateral variations. Fig. 3.2(b) shows the 10,242 spherical-spline nodes ($\sim 2^\circ$ lateral spacing, spherical spline "level" 6) used for both perturbations of V_s and ξ laterally. The Vs parametrization is the same as in the development of SEMUCB-WM1, but for ξ it is denser (in SEMUCB-WM1: 642 nodes, $\sim 7.9^\circ$ lateral spacing, spherical spline "level" 4). Due to the denser distribution of horizontal splines, smaller-scale ξ variations are expected to be present in the model.

the inversion starts from a 1D global average model based on that of SEMUCB-WM1. However, the 1D average profile of model SEMUCB-WM1 presents unrealistically large values of ξ right below the Moho. To try and correct this feature, which could influence the inversion results deeper in the mantle, the starting 1D model SEMUCB-WM1 is modified by replacing the large ξ variations in the depth range of 30- 60 km by a constant value, picked at 60 km depth, as shown in Fig. 3.3. The V_s and ξ structure of the crust and uppermost mantle are inverted with both parametrization of 10,242 horizontal splines.

The uppermost mantle structure is constructed in two steps: (1) in the first 5 iterations, use the top eight cubic b-spline knots, ranging from the Moho to the depth of 171 km and dispersion data in the period range 20-150s; (2) from the 6th to 11th iteration, fix the structure in the top 6 b-splines, and use 6 cubic b-spline knots from 121 to 471km depth, and dispersion data in the period range 25-150s. A different value of the norm damping than for the first six iterations and additionally, stronger damping on the deepest 3 knots at 271, 321, 471km depth are applied, since the group velocity data considered have weak sensitivity at depths larger than 250 km.

3.2.3 Modeling workflow

A high-level workflow is shown in Fig. 3.4. It includes an iterative process, mainly consisting of 3 parts:

1. *Construction of the homogenized crust* :The crustal model is generated in two steps:
   (1.1) As in FR2014, in the first step, a non-linear model space enumeration approach

is used to obtain a starting isotropic crustal model, using the following search bounds:
1. shear-velocity: $3.0 \leq V_s \leq 4.6$ km/s
2. Realistic bathymetry: $0 \leq h_b \leq 6$ km
3. Restricted Moho depth: $30 \leq h_m \leq 60$ km. Beneath the Moho, the 1-D mantle structure is that obtained in the previous iteration of inversion for mantle structure. In the first iteration, this is the 1D global average starting model.
(1.2) In the second step, radial anisotropy is introduced into the inversion using an optimization approach to obtain a radially anisotropic crustal model that fits both Rayleigh and Love group velocity data (still as in FR2014).

2. *Inversion for the uppermost mantle* : After updating the crustal model, group velocity dispersion data from 25s to 150s are used to invert for uppermost mantle structure using a generalized least-squares scheme. The partial derivatives for this inversion are built at each geographical spherical spline node using a finite difference approach applied to predictions from the normal mode code MINEOS (e.g., Woodhouse 1998; Lekić and Romanowicz 2011) and only computed for the knots below the local Moho depth. To allow reasonable V_s and ξ structure variations, we gradually relax the damping constraint at each iteration.

3. Radial anisotropy for the crust (step 1.2) and for the uppermost mantle (step 2) are iteratively inverted until convergence (e.g., Fig. 3.5). Once this process has converged, the corresponding model will be treated as a starting model for further inverting (iterating step (2)) deeper structures from 121 km to 471 km depth.

3.2.4 Models combined with SEMUCB-WM1

To better understand the ability of the dispersion-based models to fit both dispersion data and waveforms, compared to SEMUCB-WM1, two hybrid models are constructed, by replacing the parts of model SEMUCB-WM1 above 80km and 220km depth, respectively, with the dispersion-based models. The two additional hybrid models are refered as "merged models". The specific choice of 80km and 220km, where the Vs and ξ are similar between 1D SEMUCB-WM1 and the 1D dispersion models (e.g., Fig. 3.6), is somewhat arbitrary: it makes it possible to avoid artificial discontinuities or strong gradients as a function of depth in the merged models. The performance of the following four models will be compared with that of SEMUCB-WM1, for fitting of dispersion data, on the one hand, and long period waveforms, on the other:
Dispersion models:

1. Disp_20s_iter5: consists of 3D structure above 171 km, and SEMUCB-WM1 1D structure below.

2. Disp_20s_iter11: consists of 3D structure above 471 km, and SEMUCB-WM1 1D structure below.

Merged models:

3. Merged_80km: full 3D structure, consists of Disp_20s_iter5 3D structure above 80km and SEMUCB-WM1 3D structure below.

4. Merged_220km: full 3D structure, consists of Disp_20s_iter11 3D structure above 220km and SEMUCB-WM1 3D structure below.

3.2.5 Evaluation of model performance

After finalizing the model construction as described above, model performance are compared by computing predicted group velocity dispersion data and waveforms. The synthetic dispersion is calculated between 16 and 150 s using MINEOS and the misfit to data is computed using L-1 norm. Since the group velocity data are themselves a product of tomographic inversions, which could be biased by the theoretical assumptions that are part of the path-average approximation, and could also be affected by parameterization or regularization choices, it is helpful to evaluate the performance of the model in terms of its ability to predict waveforms, using the SEM. Fig. 3.7 shows five carefully chosen globally distributed events (Table 3.1) and over 200 stations equipped with very broadband sensors around the globe. The SEM synthetic waveforms are then computed and the automated windowing approach (e.g., Li and Romanowicz 1996; Panning and Romanowicz 2006), and the same data selection criteria as in the development of SEMUCB-WM1, are applied to select a collection of body and surface wavepackets by comparing synthetic waveforms computed in SEMUCB-WM1 to the data. The same time windows for selecting wavepackets are applied to the synthetics computed in the 4 other models. Based on the starting time and duration of the selected wavepackets, for each wavepacket type, the real waveforms (d) are compared with the synthetic waveforms (g) generated by different models (m). For this, the residual variance ($\|d-g(m)\|_2^2/\|d\|_2^2$), maximum cross-correlation coefficients with the observed wavepackets, and the corresponding lag times (e.g., negative lag time means the phase in the synthetic waveform travels faster than is observed) are computed.

3.3 Results

3.3.1 Results of Inversion of Dispersion Data

The evolution of the L1 misfit is shown in Fig. 3.5. After 5 iterations for crust and uppermost mantle structure, the L1 misfit (e.g., Fig. 3.5 blue) has dropped below the level of the acceptable misfit as defined in the construction of SEMUCB-WM1 (e.g., Fig. 3.5 red). Since Rayleigh wave group velocities in the period range 100-150s are most sensitive to the structure around 100-200km, it is still possible to improve the misfit at longer periods. Starting with the sixth iteration, the structure above 101 km is fixed and the depth range of inversions extends down to 471 km, while damping the last 3 B-splines in depth. After another 6

iterations, the L1 misfit (e.g., Fig. 3.5 green) again falls below the level set for the construction of SEMUCB-WM1. The additional 12th iteration shows no improvement of L1 misfit, so the inversion stops at the 11th iteration.

The evolution of the 1D profile for iterations 0-5 is shown in Fig. 3.3. The V_s in the depth range of 0-20 km progressively decreases and the ξ in the whole homogenized crust is close to 1.0. For the uppermost mantle, the V_s at 30-50km depths progressively increases as does ξ between 30-80 km depth. Fig. 3.8 shows the evolution of the 1D profile for iterations 6-11. At depths of 100-250 km, V_s becomes faster and ξ decreases.

Fig. 3.6 shows the V_s (a) and ξ (b) profiles respectively for each model in section 3.2.4. The V_s of dispersion models is generally slower in the homogenized crust (0-30 km depth) and faster in the uppermost mantle (30-250 km depth) than SEMUCB-WM1. The ξ of dispersion models is lower than SEMUCB-WM1 in the depth ranges 0-30 km (the homogenized crust), 30-45 km, and 80-200 km and only larger in the 45-80 km depth range.

Fig. 3.9-3.10 present the maps of V_s variations in two different depth ranges: 40-100km and 110-200km, respectively, compared to SEMUCB-WM1. It is clear that V_s models in the depth range 40-100km (e.g., Fig. 3.9) are in good agreement. In the depth range 110-200km (e.g., Fig. 3.10), the strength of the dispersion-based models fade out as depth increases: the Disp_20s_iter5 model at depths of 140-170 km (e.g., Fig. 3.10(b)); and the Disp_20s_iter11 model at depths of 170-200 km (e.g., Fig. 3.10(c)).

Next, the maps of the global ξ structure in the same depth ranges are shown in Fig. 3.11 and Fig. 3.12. The dispersion models show smaller-scale variations because of the denser lateral parametrization compared to SEMUCB-WM1, and are weaker at the top of the uppermost mantle. The large-scale structure is generally consistent except for its strength, which is much reduced in the dispersion-based models in the 110-220 km depth range.

3.3.2 Model Performance

In what follows, the inversion results are evaluated using (1) the L1 misfit of dispersion data at different periods and (2) the waveform residual variance value, cross-correlation coefficients and lag times of certain phases (including body and surface waves) among all the models.

3.3.2.1 L1 misfit of dispersion data

Fig. 3.13 summarizes the mean absolute L1 misfit of Rayleigh and Love wave dispersion for the different models in different period ranges (L1 misfit for individual periods can be found in Fig. 3.14). All the Rayleigh wave L1 misfits are slightly lower than the Love wave ones, in the same period range. In the 25-150s period band, the Rayleigh and Love wave L1 misfits are similar for the dispersion and merged models and fall below the level of SEMUCB-WM1, lower than the target range (>50m/s). The L1 misfits are separately compared in two period bands: a shorter-period range (16-60s, noting that 16s and 18s were not included in this inversion), more sensitive to the crust and uppermost mantle, and a

longer-period range (80-150s), mainly affected by structure at depths larger than $\sim$ 100 km. In the longer-period range, the Rayleigh wave L1 misfits are comparable, with a 10% larger misfit for Love waves for SEMUCB-WM1. The misfits for SEMUCB-WM1 are somewhat larger, since that model was constructed to fit not only dispersion data but also waveforms, with a different theoretical approach for the latter. Importantly, SEMUCB-WM1 was not designed with the purpose of resolving the shallowest crustal layers (in particular in regions of thick sediments). In the shorter-period range, the misfits for Love and Rayleigh waves and for all models are rather high, compared to longer periods. This can be explained by the fact that the shortest periods (>20s), sensitive to shallowest structure were not used in the inversion (and for SEMUCB-WM1, it was >25s).

3.3.2.2 Selected wavepackets for model assessment

SEM simulations are performed for the 5 chosen events in two different passbands (as previously mentioned: 60-400s for surface waves and 32-300s for body waves) to validate the models by examining SEM-predicted waveforms. Three-component waveforms are considered for a record lengths of 10,000s for surface waves and 5,500 s for body waves, starting at the origin time, filtered using the same passband filters as for the simulations, described by cut-off and corner frequencies as follows: surface wave passbands: 60-80-250-400s (cut-off at 400 and 60s, corners at 250 and 80s); body wave passbands: 32-38-180-300s (cut-off at 300 and 32s, corners at 180 and 38s). Fig. 3.15(a) Fig. 3.15(b) show histograms of the number of vertical component wavepackets picked by using SEMUCB-WM1 synthetic waveforms in different passbands for surface and body waves, respectively. Fig. 3.16 and Fig. 3.17 show the histograms on longitudinal and transverse components. Note the many picks of fundamental mode first and second orbits (e.g., R1, R2, G1, G2..), overtone first, second and third orbit (e.g., XR1, XR2, XG1, XG2..), and mixed mode (e.g., R1+XR1, R2+XR1, XR1+XR2...) surface waves, as well as body wavepackets containing S-type phases (e.g., S, SS, SSS), multiple-ScS waves, and even P waves (e.g., P, Pdiff, SP). The different depth sensitivity of these phases can help to evaluate the models. For example, fundamental mode surface waves are most sensitive to the uppermost mantle structure down to 200 km depth, while surface wave overtones are sensitive to upper and mid-mantle structure, as are body waves such as SS or SSS.

3.3.2.3 Residual variance of waveforms

Fig. 3.18 shows residual variance comparisons for selected surface waves (Panel a) and body waves (Panel b) on the vertical component. A clear trend of increasing residual variance cab be found as follows: SEMUCB-WM1 has the lowest, then successively, Merged_80km, Merged_220km, Disp_20s_iter11, and Disp_20s_iter5. The residual variance of SEMUCB-WM1 is always lowest, the "Merged_80km" has a similar performance as SEMUCB-WM1 and the two dispersion models show significantly larger residual variance. The residual variances of surface waves are positively correlated with the distance travelled, (e.g., in increasing order

of residual variance: XR1, R1, XR2, R2, and XR3). The body waves are arranged by their number of surface bounce points, e.g., S, SS, and SSS. The residual variance for body waves also shows a positive trend with increasing number of bounce points. Fig. 3.19 and Fig. 3.20 show a similar trend for the longitudinal and transverse components, respectively. More complete residual variance results can be found in Fig. 3.21 for surface waves and Fig. 3.22 for body waves.

3.3.2.4 maximum cross-correlation coefficient and lag time of waveforms

The phase list which was established by picking wavepackets in SEMUCB-WM1, are used to find the maximum cross-correlation coefficient and its lag time between different synthetic datasets and real data, to further discuss the performance of these models. Figs 3.23-3.28 show the histograms and probability density curves by kernel density estimation (KDE, e.g., Silverman (2018)), for the vertical component, of the maximum cross-correlation coefficient and corresponding lag times for the selected first-orbit paths (e.g., R1, XR1, G1, XG1), the second-orbit paths (e.g., R2, XR2, G2, XG2), and multiple-bounce S waves (e.g., S, SS, SSS). The results show all the models have similar distributions of maximum cross-correlation coefficient, with a peak between 0.9 and 1.0, but with very different lag times. For SEMUCB-WM1 and the "Merged_80km" model, the lag time distributions have similar shapes, centered close to a zero lag-time, except for the slight offset for G2 (e.g., Fig. 3.23(b)) in the "Merged_80km" model indicating subtle accumulation of the bias on ξ structure. These distributions are different for the two models based only on dispersion data: for example, for R1 (e.g., Fig. 3.24(a)) and G1 (e.g., Fig. 3.23(a)), they are broader and shifted to the left (which indicates that the model is too fast). This broadening and shift to the left is larger for R2 (e.g., Fig. 3.24(b)) and G2 (e.g., Fig. 3.23(b)), which travels a longer distance, indicating accumulation of a systematic bias, part of which is most likely due to the global 1D average of the model at depths greater than 80 km. The peaks of the Love waves (G1, G2) lag time distribution are shifted further to the left compared to the Rayleigh waves (R1, R2). It couples imply that the ξ is too strong. However, it is hard to directly assess the ξ structure for two models since the depth sensitivities are different for Rayleigh and Love waves at the same period. Compared to fundamental mode surface waves, the distribution of lag times for overtones (e.g., Fig. 3.25 and Fig. 3.26) and body waves (e.g., Fig. 3.27 and Fig. 3.28) are more centered near 0 lag time and narrower, but broader than for SEMUCB-WM1 and the "Merged_80km" model, reflecting some sensitivity of these waveforms to uppermost mantle structure. In particular, the second-orbit overtones (e.g., XR2 in Fig. 3.25(b) and XG2 in Fig. 3.26(b)) and the multiple-bouncing S waves (e.g., SS, SSS in Fig. 3.27(b) and Fig. 3.28), exhibit more distortion and broadening of the lag-time distribution than is observed for XR1, XG1 and S, respectively (the latter have travelled shorter distances in the upper mantle). In addition, the level of distortion for model "Merged_220km", for all the wavepacket types considered, is in between that for the two dispersion-based models and model "Merged_80km", and not as good as for SEMUCB-WM1. In all cases, the lag time shift from 0 is always negative. Note that the correlation

coefficient distribution is much broader for overtones than for fundamental mode surface waves, for all models, which may indicate either errors in the source parameters, or the need for further improvements in the mid-mantle shear wave structure, compared to that of SEMUCB-WM1.

3.4 Discussion

3.4.1 Dispersion data fitting

The dispersion-based models and the merged models generally fit the dispersion data somewhat better than SEMUCB-WM1 (e.g., Fig. 3.13and Fig. 3.14), except for the 80-150s Rayleigh waves of "Disp_20s_iter5" which has 3D structure only above 171 km. Note that the significant improvement of Love wave fitting (e.g., Fig. 3.13, left panel) could be due to the denser ξ parameterization for the uppermost mantle (10,242 nodes versus 642 nodes in SEMUCB-WM1 and denser B-splines in the uppermost mantle). Also, including 20s dispersion data helps improve the fit to the even shorter period dispersion data (16s, 18s). Fig. 3.29 shows the spatial distribution of the 20-60s L1 dispersion misfits for model SEMUCB-WM1 and "Disp_20s_iter5". This period range is used for constructing the crust and uppermost mantle in the dispersion-based models. The model "Disp_20s_iter5" improves the fit to dispersion data in the period range 20-60 s particularly in the oceans (e.g., the Indian Ocean and southeastern Pacific Ocean for Love waves, Fig. 3.29, left panel) and also, both for Love and Rayleigh waves, in the off-shore areas around the continents (e.g., Fig. 3.29, right panel) i.e. in regions where the thickness of sediments is large. Rayleigh waves are fit better than Love waves in general, which can also be explained by higher sensitivity of Love waves to shallower structure. Further improvement of the fit to short period dispersion data will require taking better account of variations in the thickness of sediments, which translates into denser parametrization in the crust, i.e., increase in the number of GLL points used to represent depth variations in the crust. Clearly, this is necessary for waveform modeling at periods shorter than 30s.

3.4.2 Waveform modeling performance for different models

3.4.2.1 Residual variance

As discussed in section 3.3, the residual variance for fundamental mode surface wavepackets increases as a function of distance travelled (e.g., Fig. 3.18, Fig. 3.19 and Fig. 3.20). Notably, the residual variance for the two dispersion-based models increases from XR1 to XR3, much more than for SEMUCB-WM1, and the same trend is observed for body waves (e.g., S, SS, SSS). In contrast, the residual variance of SEMUCB-WM1 and model "Merged_80km" becomes stable after XR2 for overtone surface waves, while for body waves, the residual variance for SEMUCB-WM1 barely increases, while that of model "Merged_80km" slowly

grows. Interestingly, the trend of increasing residual variance for the "Merged_220km" model is close to the level of the two dispersion-based models and only slightly better than the deeper model "Disp_20s_iter11", but still worse than for model "Merged_80km". This implies that the accumulated model bias is mainly from the depth range of 80-220km and below, and that the contribution to it from the portion of the model above 80km depth is small.

3.4.2.2 Cross-correlation and lag time

As previous mentioned in Section 3.3.2.4, most of the models have similar distributions of maximum cross-correlation coefficient but different distributions of lag times, for surface wavepackets and body wavepackets (e.g., Fig. 3.23, Fig. 3.24, and Fig. 3.27), which suggests the difference in the residual variance (e.g., Fig. 3.18) among all the models could be due to accumulated phase-shifts, but not to waveform dissimilarity. So, only the distributions of lag times are discussed below. Since the dispersion models are not fully 3D for the whole mantle, to fairly assess which depth intervals of the models are biased the most, the lag-times computed for the merged models and SEMUCB-WM1 are compared.

The result for fundamental surface waves (e.g., R1, R2 in Fig. 3.23; G1, G2 in Fig. 3.24), which are most sensitive to the uppermost mantle and crust, shows that the "Merged_80km" model and SEMUCB-WM1 predict similar lag-times, except for G2 in "Merged_80km". This indicates that the structure in the first 80km of the dispersion-based models has a similar effect as that of SEMUCB-WM1, but with subtle radial anisotropy bias. The distributions of lag-time are compared for models "Merged_80km" and "Merged_220km" and the shapes for "Merged_220km" are distorted and left-shifted, which means the portion of the dispersion-based model in the depth range 80-220km is on average too fast. Finally, the two dispersion-based models are clearly offset from 0 lagtime and further distorted for the wavepackets that have travelled longer paths, like R2 or G2. Note that model "Disp_20s_iter11", which has deeper 3D structure (171-471km), performs only slightly better than "Disp_20s_iter5", which only has the 3D structure down to 171km. It could be due to the 3D structure at depths larger than ∼80 km in the dispersion-based models is inadequate for fitting waveforms in the same period-range, even when inversion of dispersion data is extended to depths in excess of 200 km.

For first orbit overtone surface waves (e.g., XR1 in Fig. 3.25(a) and XG1 in Fig. 3.26(a)) and direct body waves (e.g., S, in Fig. 3.27(a)), which are mostly sensitive to upper and mid-mantle structure, all the models give similar lag-time distributions, centered near 0 lag. This shows that these wavepackets are less affected by structure above 200 km. For waves that spend more time in the upper mantle (e.g., XR2, XG2, SS, SSS), the "Merged_80km" lag times are still quite similar to those for SEMUCB-WM1, but the differences between "Merged_80km" and "Merged_220km" models emerge, whereas lag-times for the two dispersion-only based models have a broader distribution and larger shifts to the left. These results confirm that it is reasonable to consider the depth range 0-80km of fundamental mode dispersion-based models for fitting teleseismic long period waveforms in the period ranges considered, but the structure below ∼80 km depth is biased in these models. There

could be two main reasons for this discrepancy: the first one is the diminishing sensitivity of fundamental modes to structure as depth increases, even at periods larger than 100s. This may, to some extent, be remedied by the inclusion of overtone dispersion data. However, the increasing bias with distance for overtone surface waves for the "Merged_220km" model indicates that there may also be a theoretical issue: dispersion-based tomography assumes that surface waves are sensitive to the 1D average structure between the source and the receiver (i.e. the path average approximation), which is a better approximation for fundamental modes, but breaks down for overtones, where more accurate kernels need to be considered, that is, at least 2D kernels in the vertical plane such as those provided by NACT (e.g., Li and Romanowicz 1995; Romanowicz et al. 2008).

To summarize this section, some model bias in fundamental mode dispersion-based models is present in the depth range of 80-220km (in addition to poorly resolved structure below 220 km), as shown by the results of cross-correlation coefficient and lag time. This also agrees with the results of residual variance (section 3.4.2.1). This is contrary to the common assumption that fundamental mode dispersion-based tomographic inversion accurately retrieves shear wave velocity structure down to at least 200 km depth.

3.4.2.3 Spatial distribution of the lag times

To understand path-dependent effects on the lag time, the spatial distribution of R1 lag time for event 4 (Table 3.1) is illustrated in Fig. 3.30. The background maps show lateral variations in Vs at 100km depth for each model, centered on the eastern Pacific. Source-station paths for event 4 are color-coded by the measured lag-time for R1. Although the Vs structures at 100km depth are similar, for model "Disp_20s_iter5", the lag-time distribution shows a significant structure-related pattern, with negative lag times (bluish) along oceanic paths and positive lag times (reddish) for paths across the North American continent. Furthermore, for the dispersion-based model "Disp_20s_iter11" and the "Merged_220km" model, the positive lag times on paths sampling the North American continent paths are much reduced (especially those traversing the Canadian craton), but this is not the case for the negative lag times in the Pacific Ocean paths, which indicates that the structure is still not slow enough. On the other hand, the "Merged_80km" and SEMUCB-WM1 do not show this structure-related pattern. If anything, the models are slightly too slow. As for the spatial distribution of G1 lag time (e.g., Fig. 3.31), it is similar to the R1 pattern for the dispersion-based models except for the continent paths, which are close to zero. The "Disp_20s_iter11" and the "Merged_220km" models predict negative lag times for the Pacific Ocean paths but stay close to zero for the continental paths. The "Merged_80km" and SEMUCB-WM1 lag times also do not reflect the structure-related pattern. Fig. 3.32 show the comparison of the R1 synthetic waveforms computed in the different models considered with the data for an oceanic path (station RAR) and a continental path (station ALE), which confirms the observations shown in Fig. 3.30: The model based on dispersion down to $\sim$100 km is too fast on the oceanic path and too slow on the continental path (e.g., Fig. 3.30 (b)). This is corrected for continental paths when inversion is extended to depths greater than 200 km,

but not for oceanic paths (e.g., Fig. 3.30). Only merging the dispersion-model with the FWI model SEMUCB-WM1 at relatively shallow depths (∼80 km) helps restore the phase match on the oceanic path. These observations can be interpreted as indicating that the dispersion data could not recover the strength of heterogeneity in oceanic regions in the depth range 80-220km and in particular, they do not seem to fully capture the strong velocity reduction in the oceanic low velocity zone. Because two thirds of the globe is covered by oceans, the fundamental mode dispersion-based models are overall biased fast, as reflected in the lag-time analysis. There could be a theoretical underpinning to this bias, related to the fact that dispersion interpretation in the context of PAVA is a ray-theoretical interpretation that is affected by wavefront healing (e.g., Nolet and Dahlen 2000).

3.4.2.4 The use of source stacking for model assessment

Source stacking (e.g., Capdeville et al. 2005; Romanowicz et al. 2020) is used for reducing the computational costs when performing SEM simulations. In this approach, a large number of sources, normalized by their seismic moment, can be triggered simultaneously for the computation of summed waveform synthetics at each station, and compared to the corresponding summed observed waveforms. Here, how this approach can be used to help assess model performance at reduced computational cost is illustrated. Fig. 3.33 (Panel a) shows the stacked observed vertical component waveforms for the selected 5 events at the station SSPA, normalized by their seismic moments. The panel b of Fig. 3.33 shows a comparison of summed data and synthetics waveforms for model SEMUCB-WM1, in the surface wave passband 60-80-250-400s, obtained with only one SEM simulation. Because the records for only 5 events are stacked in this experiment, in order to average the path-dependent effects from individual events, the other two summed waveforms at stations POHA and DWPF are additionally stacked. These three stations are the only three stations that recorded high quality waveforms in this bandpass for all five events.

Fig. 3.34 shows the vertical component results of the cross-correlation coefficient, lag time, and residual variance in the time windows (500-2800s, 2800-5100s, 5100-7400s, and 7400-9700s) by comparing the resulting summed waveforms with the corresponding synthetic summed waveforms for the different models considered in this study. For all the models, the cross-correlation coefficient and the lag time have a downward trend and the residual variance has upward trend, as time increases. The results clearly show two groups of models: the SEMUCB-WM1 and "Merged_80km" models have higher cross-correlation coefficients, smaller lag times, and lower residual variance than any of the other 3 models. The results of source stacking are consistent with the results described in Section 3.3.2.3 and Section 3.3.2.4. Interestingly, the results of "Merged_220km" are very similar to those of the dispersion models, which again, indicates that the accumulated model bias of the dispersion models mainly comes from the depth range 80-220km. The implication agrees with the statement mentioned in Section 3.4.2.1.

3.5 Conclusions

This study presents the constructions of global crust and uppermost mantle models by inverting fundamental mode Rayleigh and Love wave dispersion data down to two different depths (171 and 471 km), and assessed how well such models can fit body and teleseismic waveform data, compared to model SEMUCB-WM1 (FR2014), the upper mantle part of which was developed using constraints from both dispersion and full waveform data. Based on the comparison of predicted and observed waveforms collected from 5 globally distributed events observed at 200 very broadband stations, these dispersion-based models provide a significantly poorer fit to waveform data than SEMUCB-WM1. Two additional hybrid models are then constructed, by combining the shallower part of the dispersion-based models with the deeper part of SEMUCB-WM1 starting at 80 km and 220 km depth respectively. While SEMUCB-WM1 always performs better than any of the other models, only the hybrid model with SEMUCB-WM1 structure below 80 km has the closest performance to that of SEMUCB-WM1 based on variance reduction, cross-correlation coefficients and lag times computed over different wavepackets. In particular, the results show that most of the models have similar distributions of maximum cross-correlation coefficient but with very different lag times. Only the "Merged_80km" model and SEMUCB-WM1 predict similar lag-times. By analyzing the spatial distribution of the lag times, the dispersion models show patterns of lag-time distribution that are correlated with tectonic provinces sampled by the ray paths. This shows that fundamental mode dispersion-based models do not represent global shear velocity structure adequately at depths greater than 80 km, and notably, they cannot recover the strength of the velocity reduction in the oceanic low-velocity zone. This is contrary to the assumption that fundamental mode dispersion-based tomographic inversion has the capability to retrieve shear wave velocity structure accurately down to 200 km depth.

The results suggest that a group velocity dispersion-based model can be used down to 80 km depth in FWIs for mantle structure using fundamental and overtone surface waveforms at periods larger than 60s and body waves at periods larger than 32s. A workflow is proposed for global FWI, in which the top 60-80 km depth is first constrained by dispersion, using an approach based on the concept of effective medium, and the FWI is subsequently performed only at larger depths. This can significantly save the computation time involved in FWI. In the future, the top portion of the model can be improved progressively as better independent constraints are made available, either globally or regionally (for example, introducing shorter period dispersion data from ambient noise measurements). In addition, how the source-stacking method helps assess model performance at a reduced computational cost is illustrated.

3.6 Table

	Event ID	Location	Origin Time (UTC)	Latitude	Longitude	Depth	Magnitude
Event 1	C200805310437A	MID-INDIAN RIDGE	2008-05-31 04:38:01	-41.350 °	80.410 °	12 km	6.4 Mw
Event 2	C200810291132A	PAKISTAN	2008-10-29 11:32:48	30.290 °	67.570 °	12 km	6.4 Mw
Event 3	C200903061050A	NORTH OF SVALBARD	2009-03-06 10:50:32	80.330 °	-2.320 °	13 km	6.5 Mw
Event 4	C201604290133A	NORTHERN EAST PACIFIC RISE	2016-04-29 01:33:42	10.420 °	-103.900 °	15 km	6.6 Mw
Event 5	C201811152309A	SOUTHERN EAST PACIFIC RISE	2018-11-15 23:09:06	-56.230 °	-122.260 °	12 km	6.3 Mw

Table 3.1: Hypocentral parameters for the five selected events for the waveform comparison part of this study.

3.7 Figures

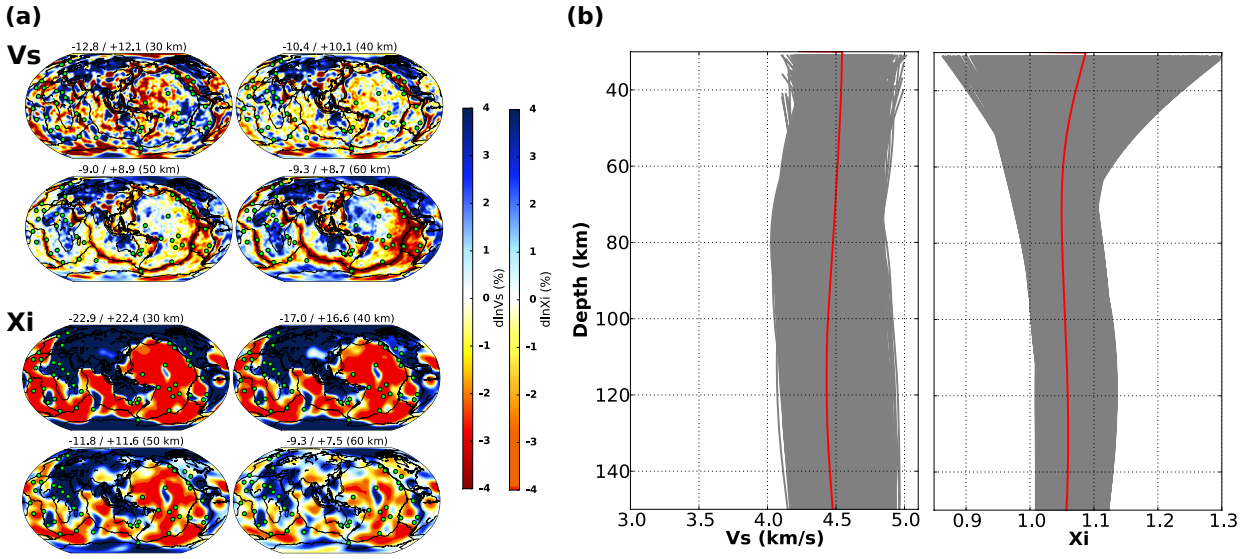

Figure 3.1: Panel (a): Map views of global maps of the distribution of Vs (upper panel) and ξ (lower panel) variations in the depth range 30-60km, in model SEMUCB-WM1 (French and Romanowicz 2014). Panel (b): Vertical profiles of V_s and ξ from 30 to 150 km depth. Red lines represent the global average 1D depth profile of SEMUCB-WM1 and gray lines represent the population of local depth profiles sampled in the 3D SEMUCB-WM1 model.

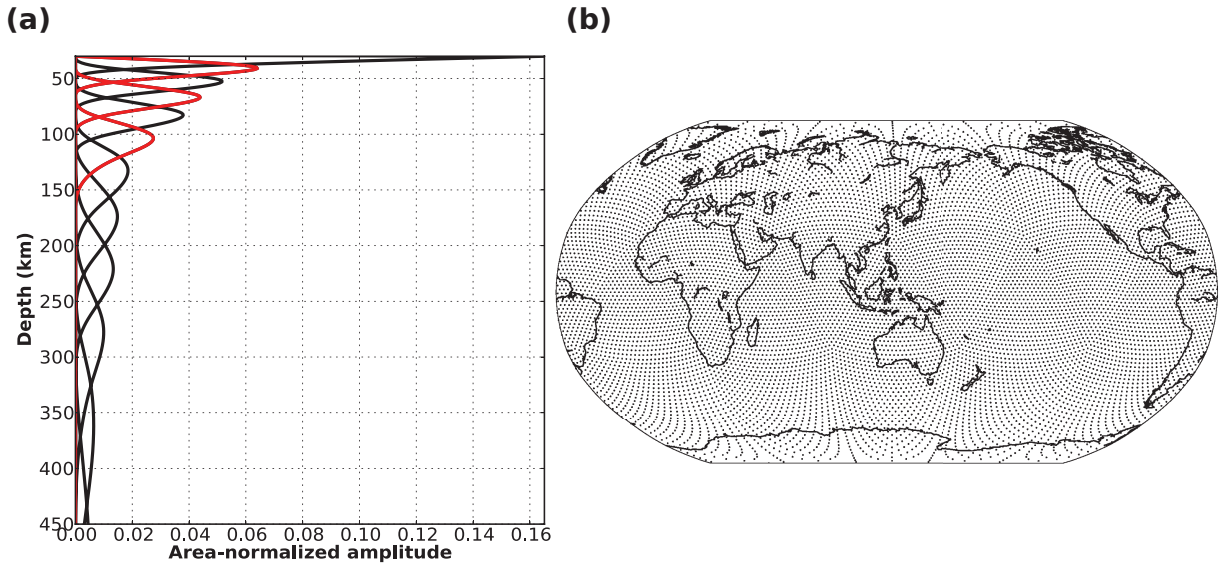

Figure 3.2: Parametrization of the dispersion-based models. Panel (a): the top 12 radial b-spline basis functions distributed between depths of 30 and 450km. The newly added 3 b-splines (compared to SEMUCB-WM1) are shown in red. Panel (b): "level 6" spherical spline knot distribution for both Vs and ξ (with an average spacing of $\sim 2^\circ$, corresponding approximately to spherical-harmonic degrees 96.)

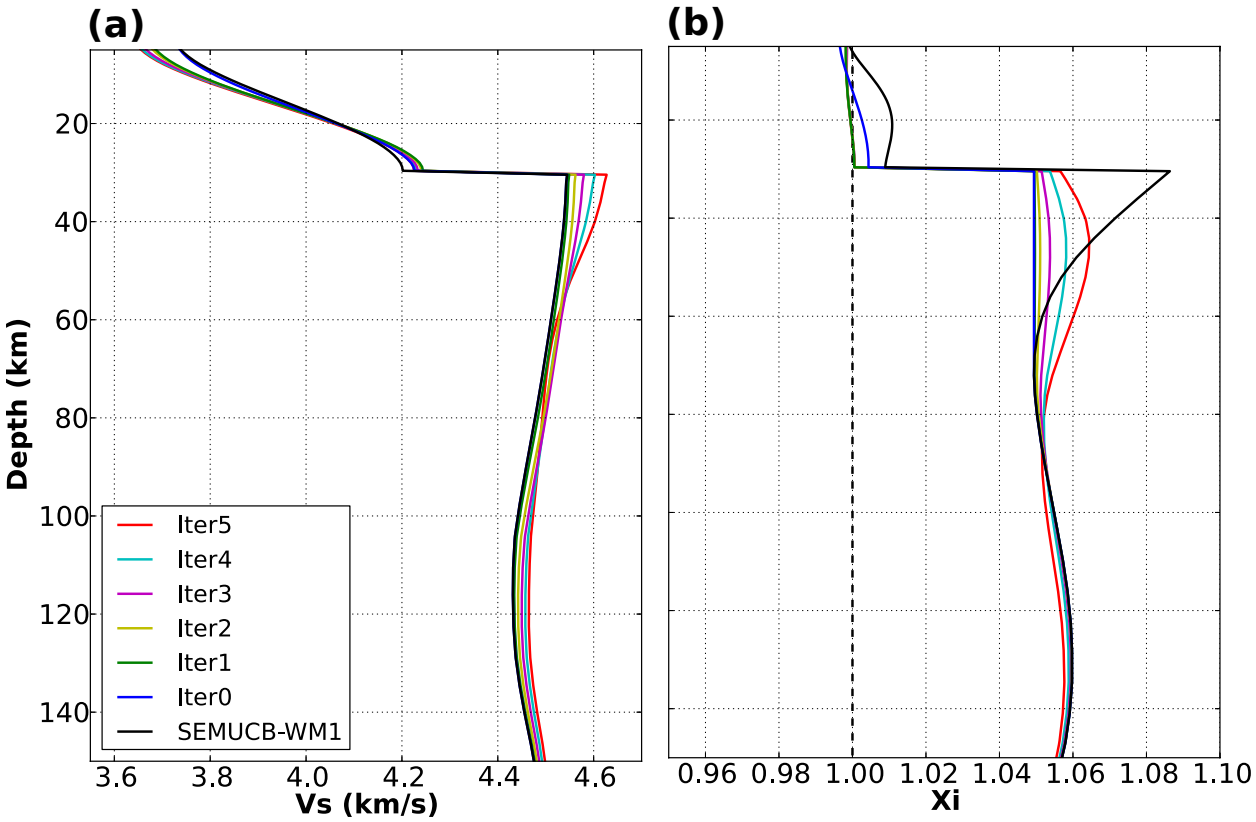

Figure 3.3: Evolution of the V_s (Panel a) and ξ (Panel b) 1D average depth profiles for iterations 0-6 and comparison with SEMUCB-WM1.

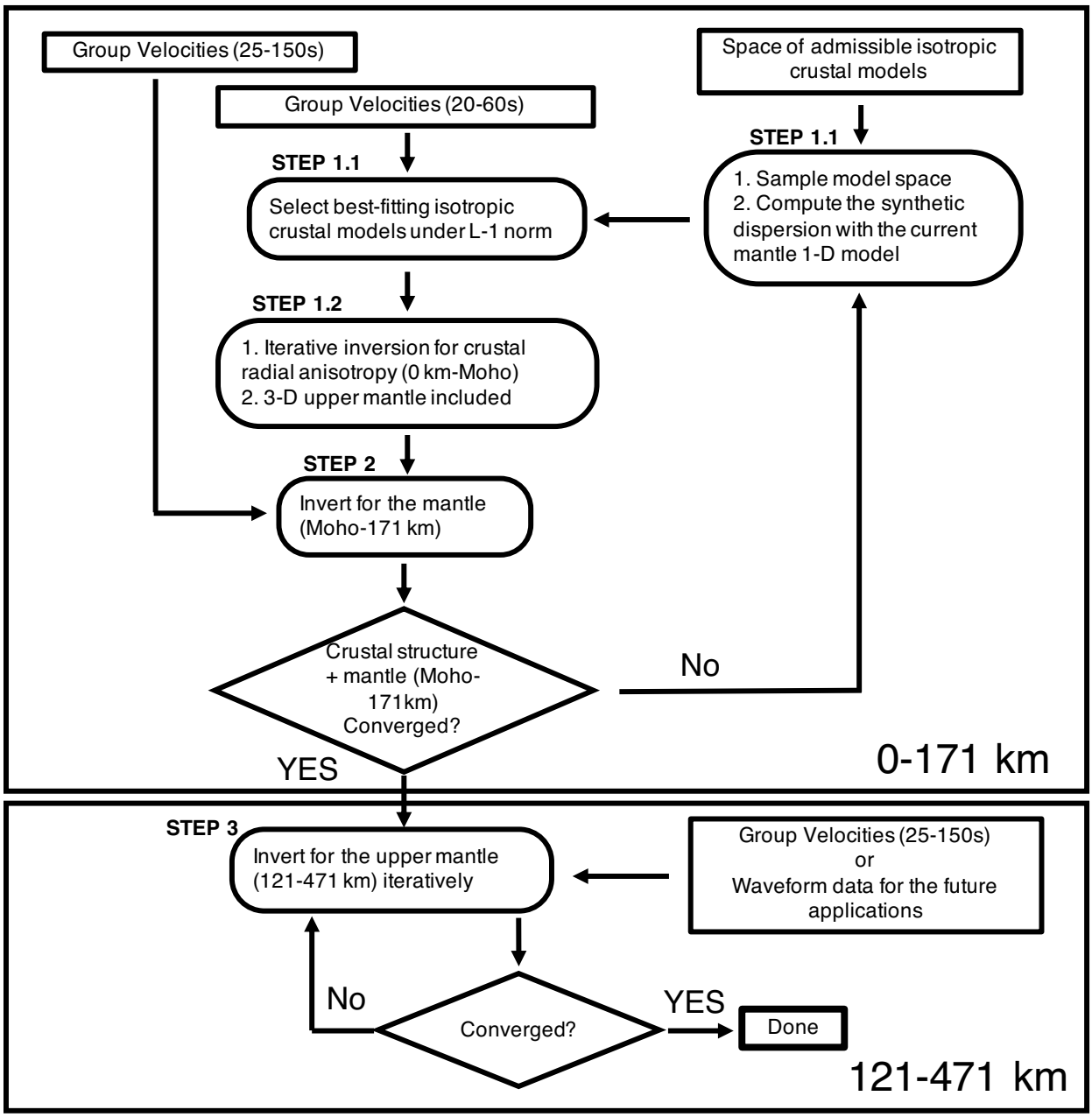

Figure 3.4: A high-level overview of the workflow considered in this study.

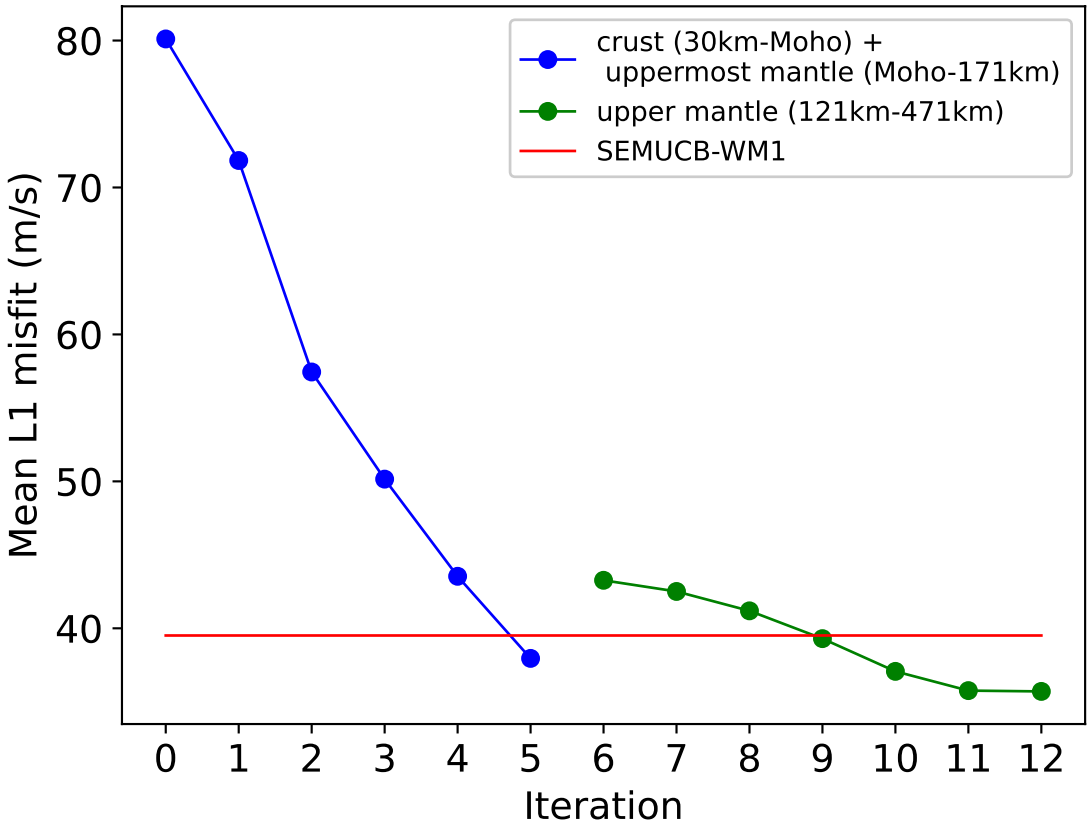

Figure 3.5: Evolution of mean absolute dispersion data L1 misfit (in m/s) as a function of iteration in the period range 25-150s. Iterations 0-5 are shown in blue and iterations 6-12 in green. Red line marks the maximum misfit level set for the construction of the crustal part of model SEMUCB-WM1.

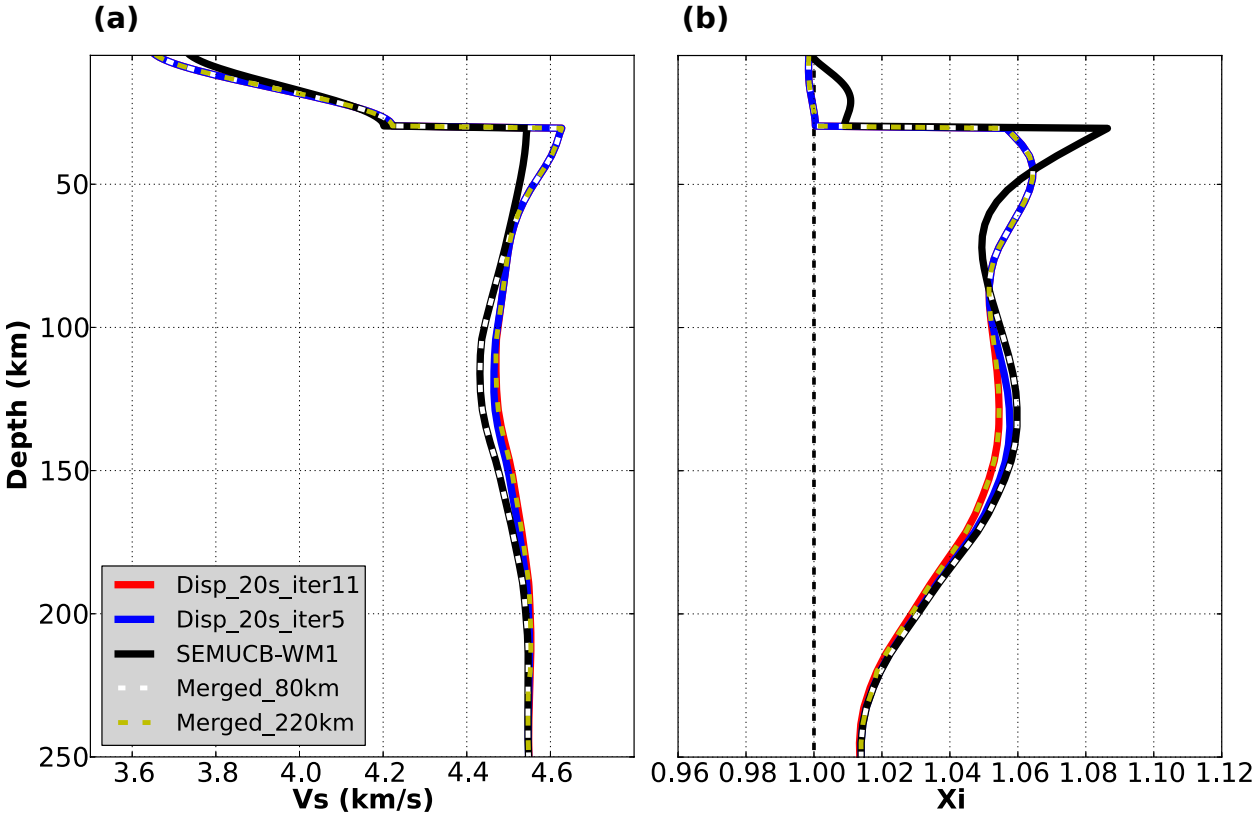

Figure 3.6: Comparison of V_s (Panel a) and ξ (Panel b) average depth profiles for each model discussed in section 3.2.4. Models are color-coded according to legend provided in the figure.

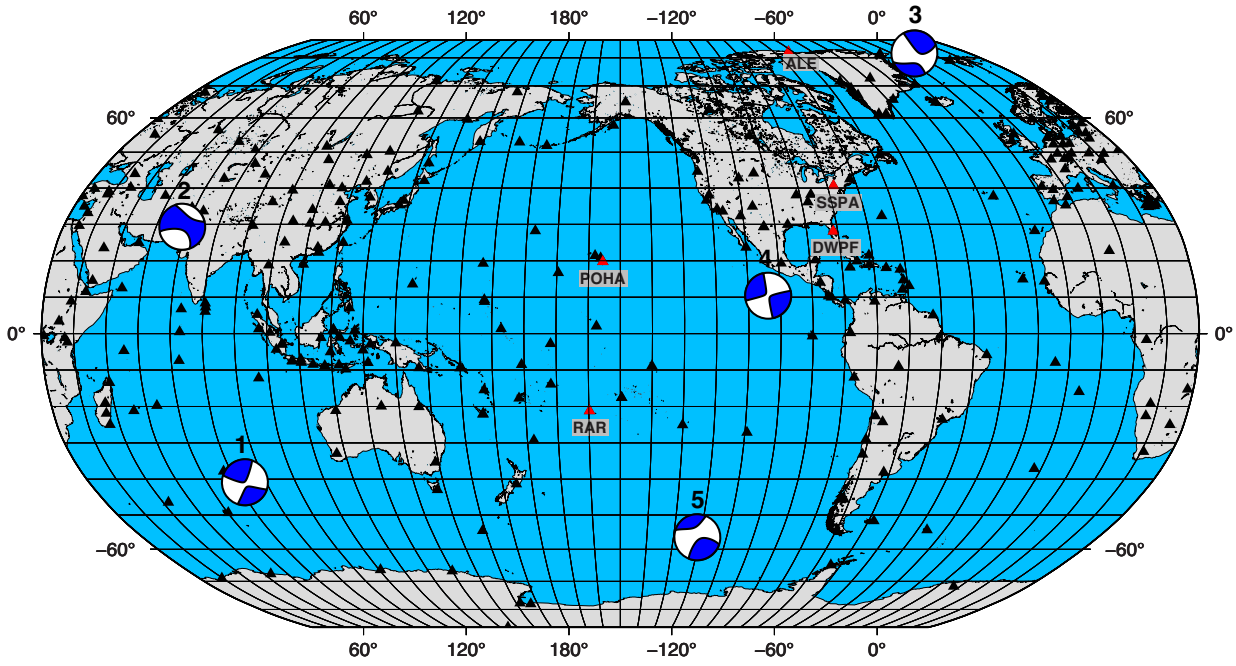

Figure 3.7: Location of the 5 selected seismic sources, with focal mechanisms indicated by beachballs and numbered as in Table 3.1, and of the 200+ seismic stations (black triangles) used in this study for waveform computations. The red triangles with station codes show the stations mentioned in Section 3.4.2.3 and Section 3.4.2.4.

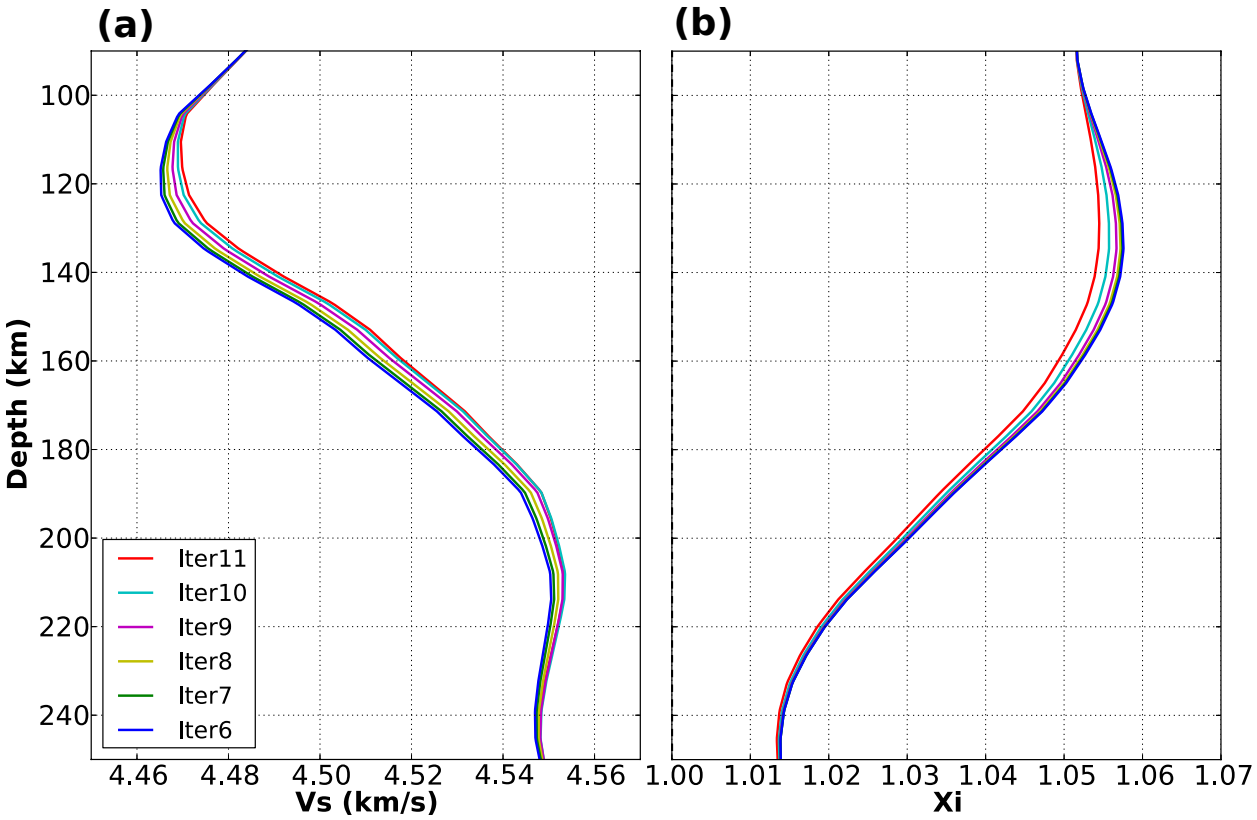

Figure 3.8: Evolution of the V_s (Panel a) and ξ (Panel b) 1D average depth profiles for iterations 6-11.

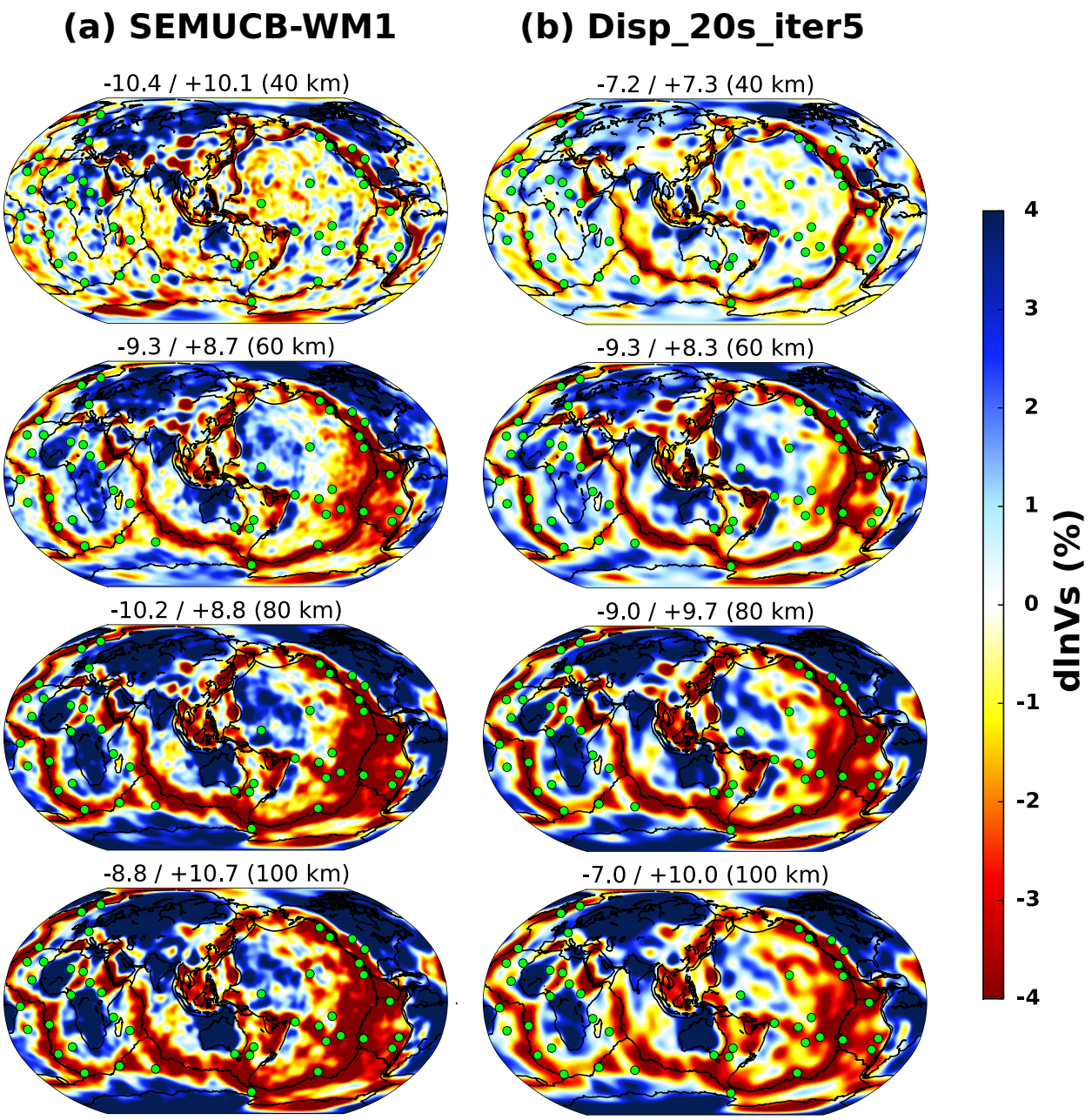

Figure 3.9: Global maps of lateral variations in Vs in the depth range 40-100km for models SEMUCB-WM1 (Panel a) and Disp_20s_iter5 (Panel b). Note that Disp_20s_iter11 is identical to Disp_20s_iter5 above the depth 100km. V_s variations are plotted in percent with respect to the corresponding average 1-D model. The maximum peak-to-peak variation for each model at the corresponding depth is shown at the top of each map. Green circles show hotspot locations from the catalog of Steinberger (2000).

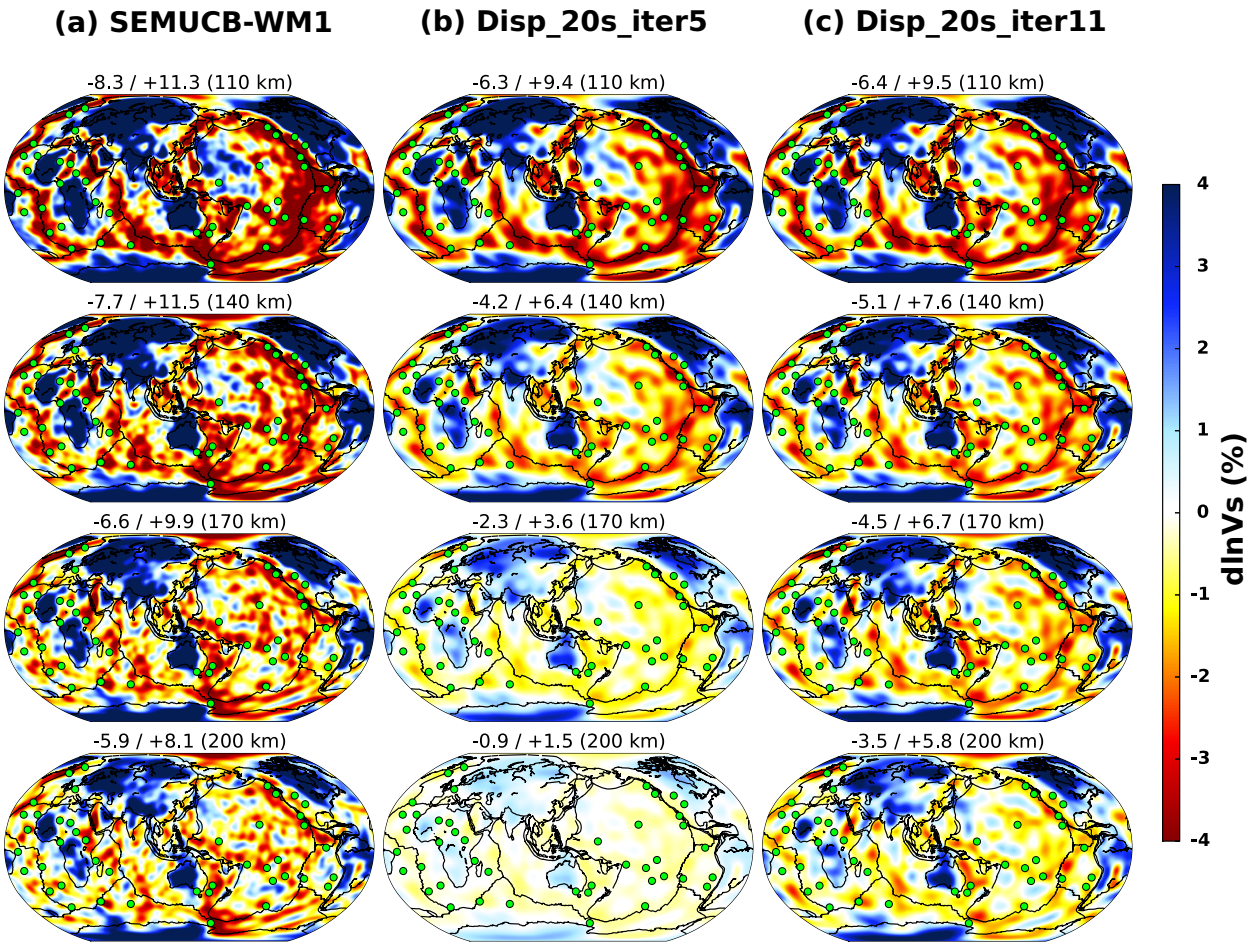

Figure 3.10: Global maps of lateral variations in V_s in the depth range 110-200km for models SEMUCB-WM1 (Panel a), Disp_20s_iter5 (Panel b), and Disp_20s_iter11 (Panel c).

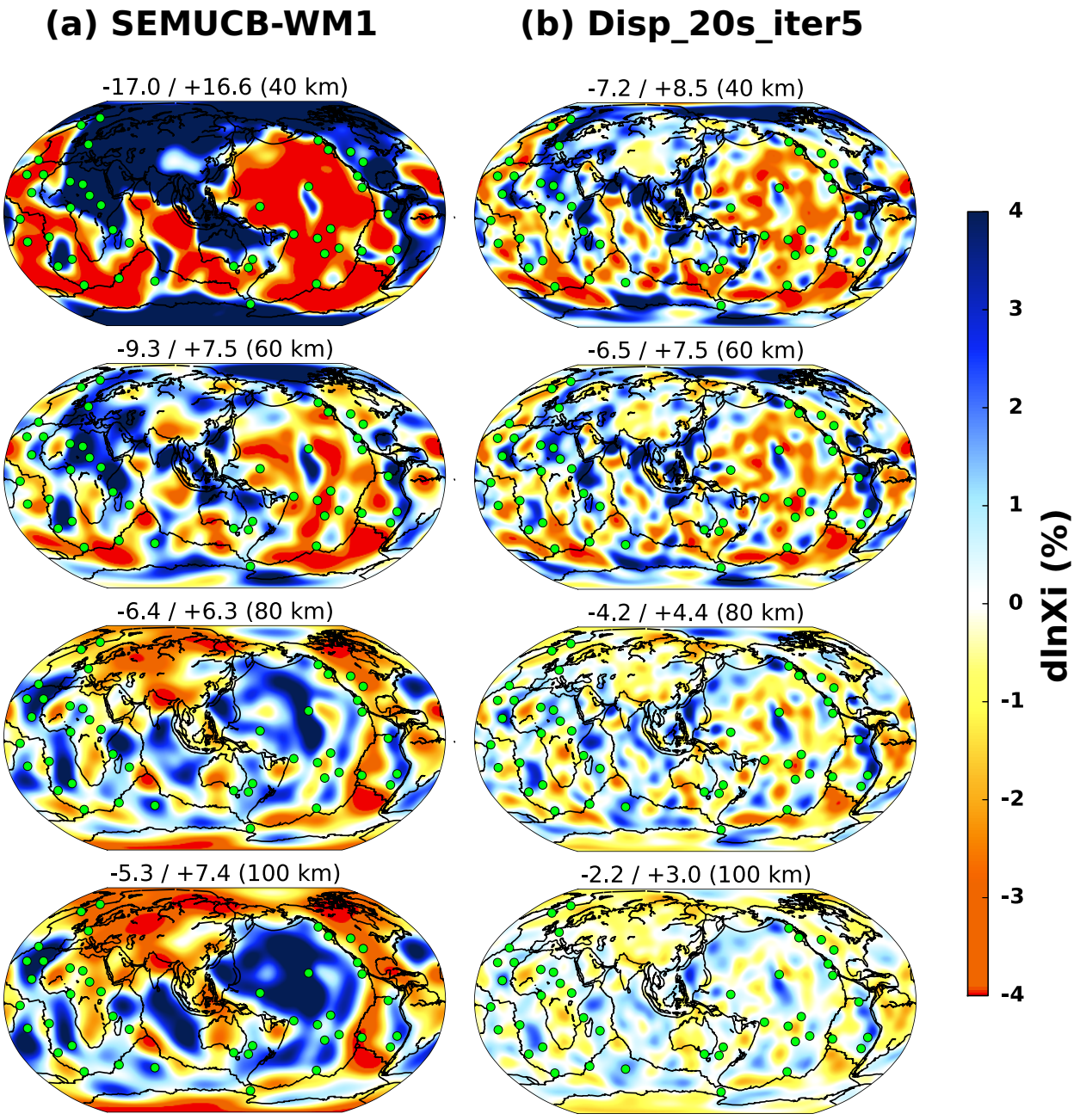

Figure 3.11: Same as Fig. 3.9 for ξ and models SEMUCB-WM1 (Panel a) and Disp_20s_iter5 (Panel b).

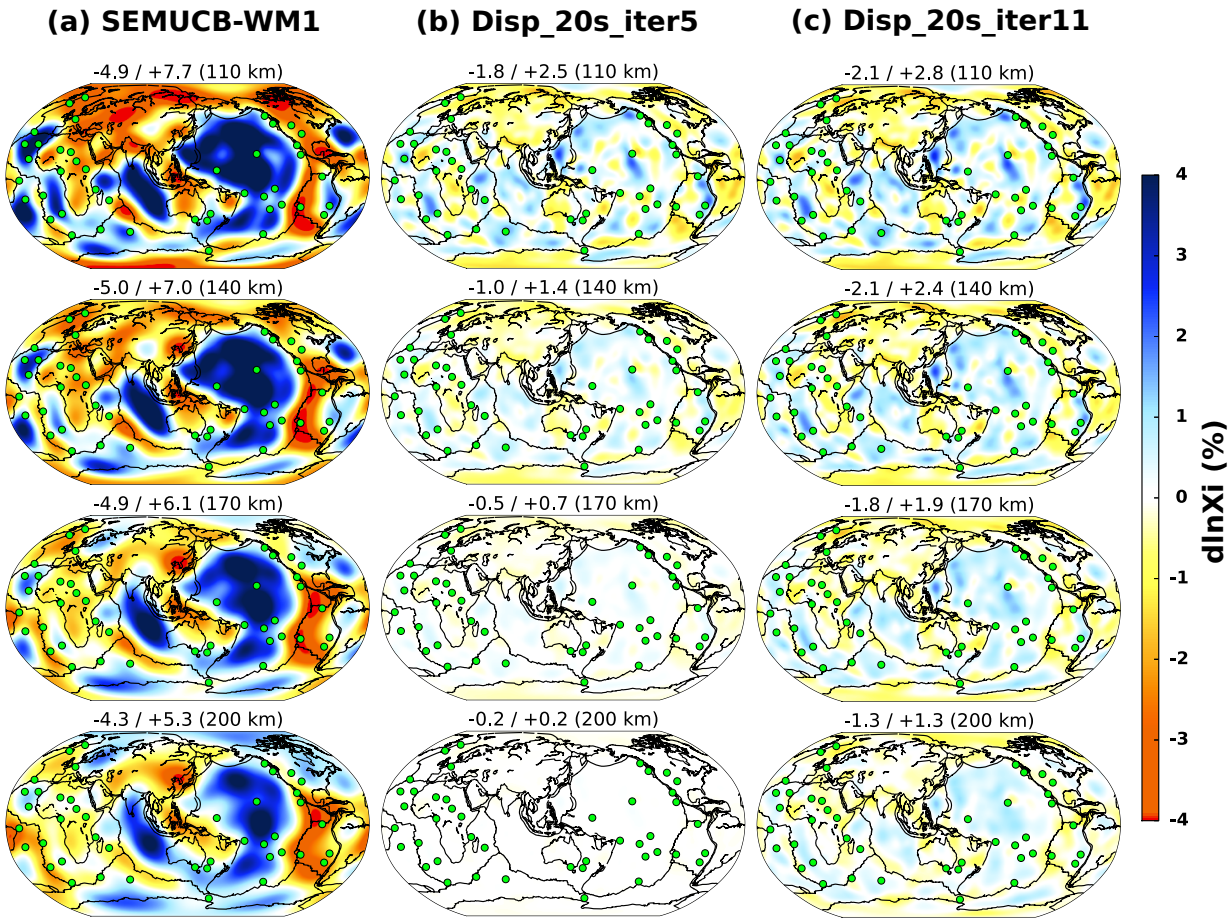

Figure 3.12: Same as Fig. 3.10 for ξ and models SEMUCB-WM1 (Panel a), Disp_20s_iter5 (Panel b), and Disp_20s_iter11 (Panel c)

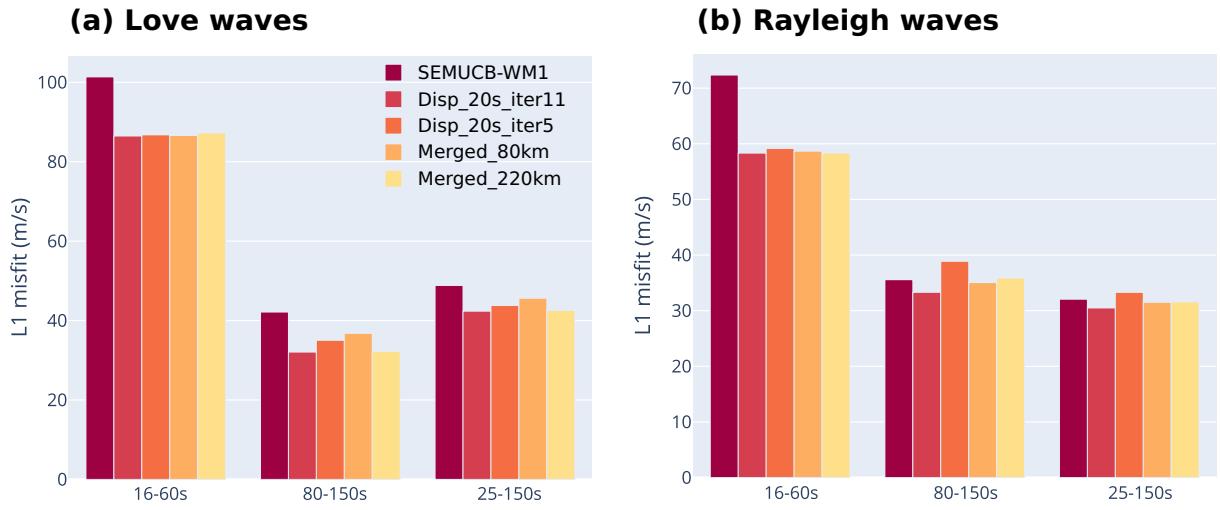

Figure 3.13: Comparison of mean absolute L1 misfit for Love (Panel a) and Rayleigh (Panel b) wave dispersion for all 5 models considered in the period ranges: 16-60s, 80-150s, and 25-150s.

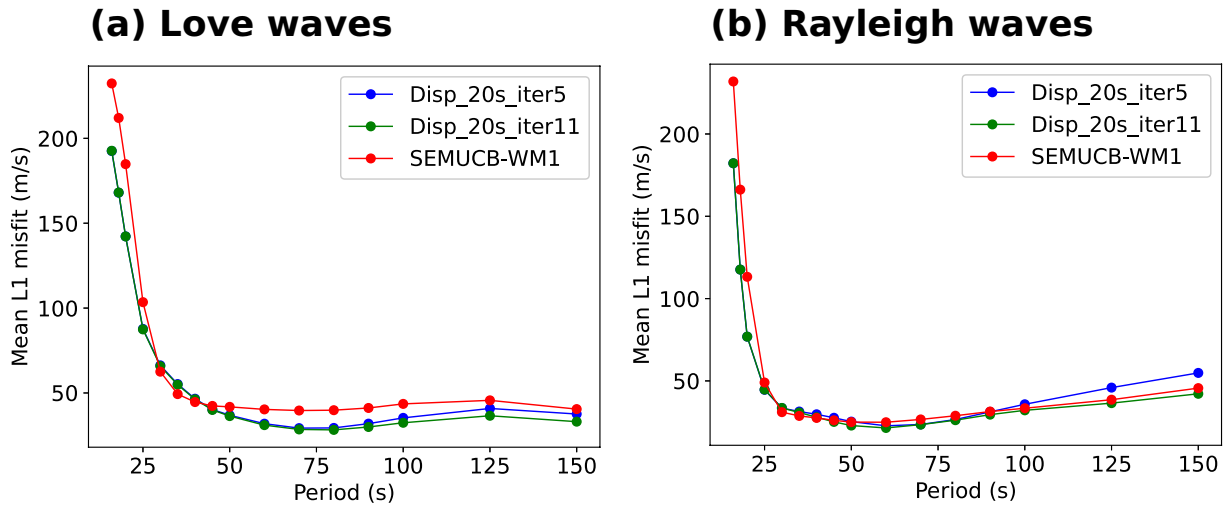

Figure 3.14: Absolute mean L1 misfit comparisons for Love (Panel a) and Rayleigh (Panel b) wave dispersion for selected periods.

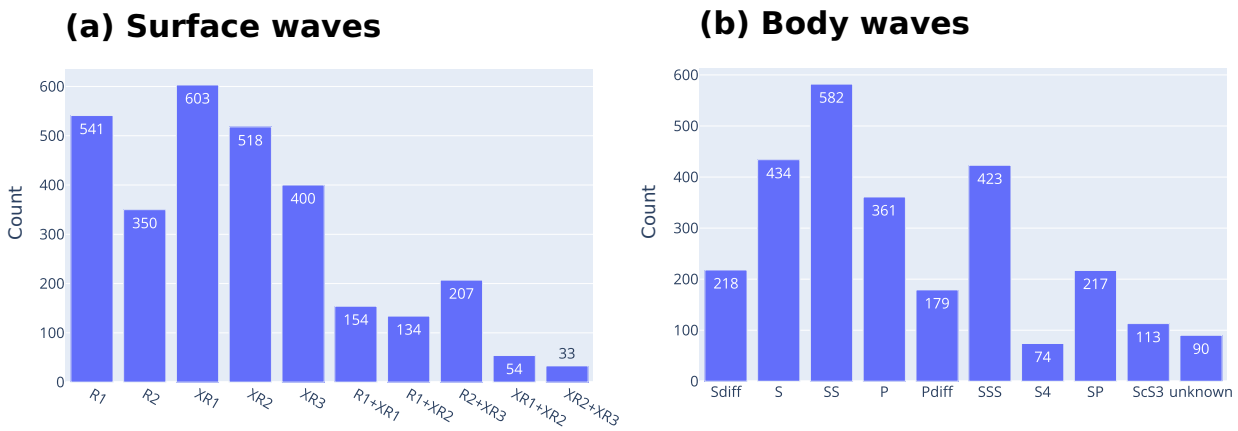

Figure 3.15: Number of vertical component wavepackets selected when using synthetic waveforms computed in SEMUCB-WM1 for comparison with observed waveforms (for the 5 events in Table 3.1), for surface (Panel a) and body waves (Panel b) in the passbands 60-80-250-400s and 32-38-180-300s, respectively.

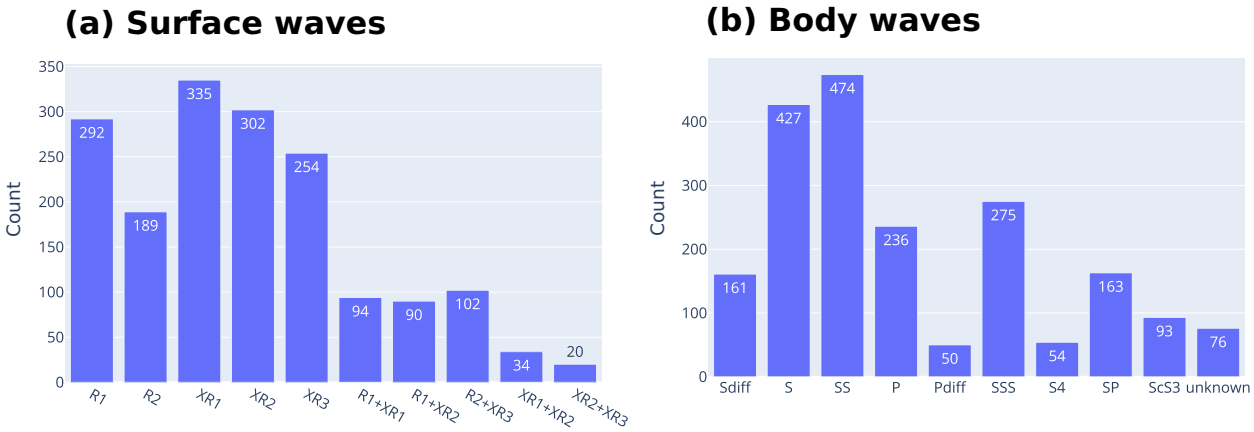

Figure 3.16: Histograms of the number of longitudinal component wavepackets selected by comparison of data with SEMUCB-WM1 synthetic waveforms for surface (Panel a) and body waves (Panel b) in the passbands 60-80-250-400s and 32-38-180-300s, respectively.

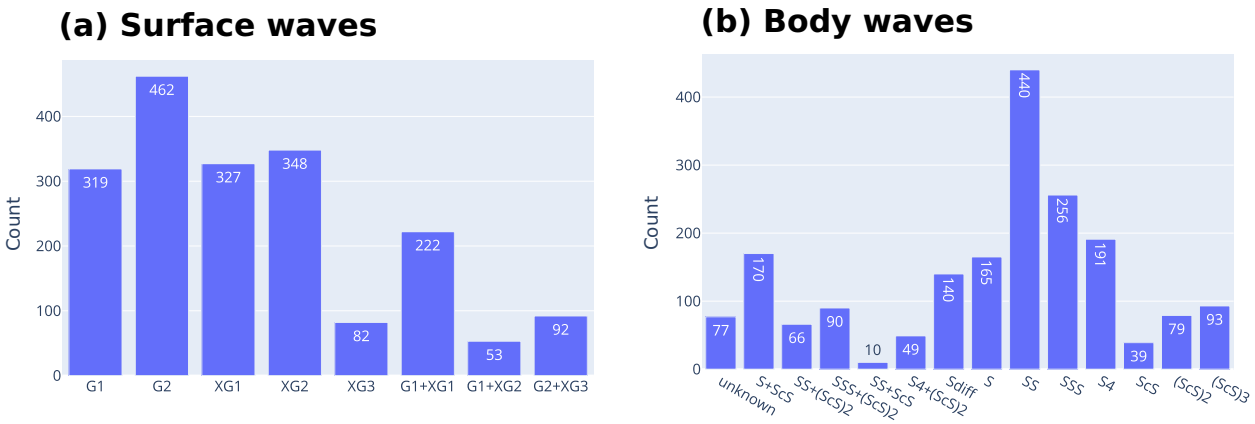

Figure 3.17: Histograms of the number of transverse component wavepackets picked by using SEMUCB-WM1 synthetic waveforms for surface (Panel a) and body waves (Panel b) in the passbands 60-80-250-400s and 32-38-180-300s, respectively.

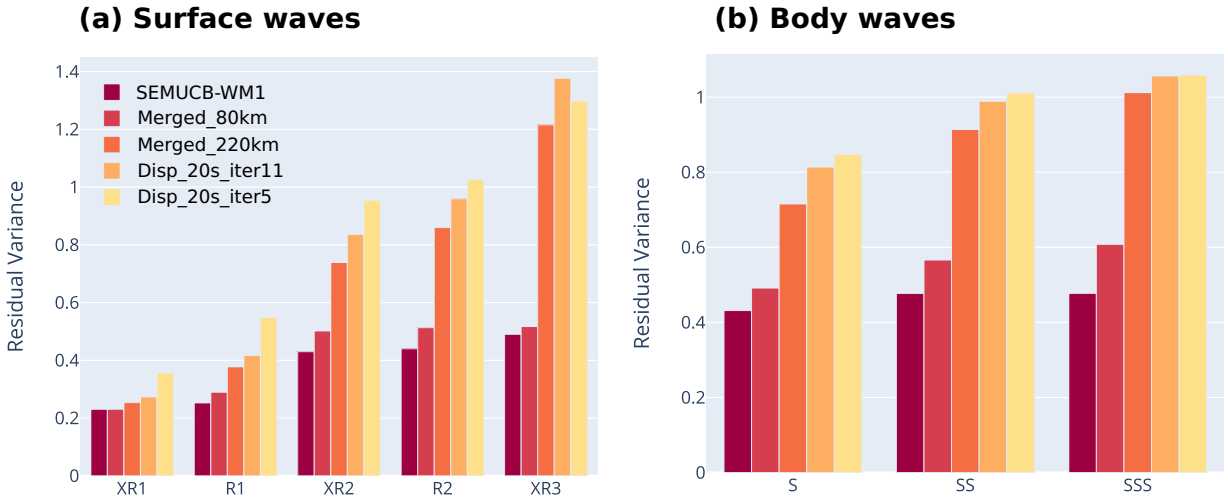

Figure 3.18: Waveform residual variance comparisons for selected surface wavepackets (Panel a) and body wavepackets (Panel b), on the vertical component.

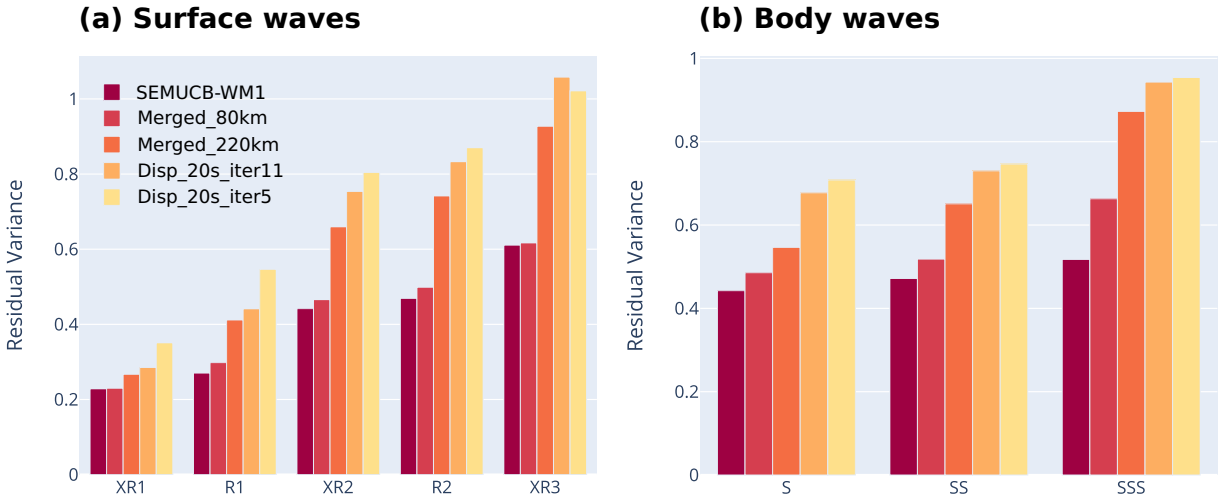

Figure 3.19: Waveform residual variance comparisons for selected surface waves (Panel a) and body waves (Panel b) on the longitudinal component.

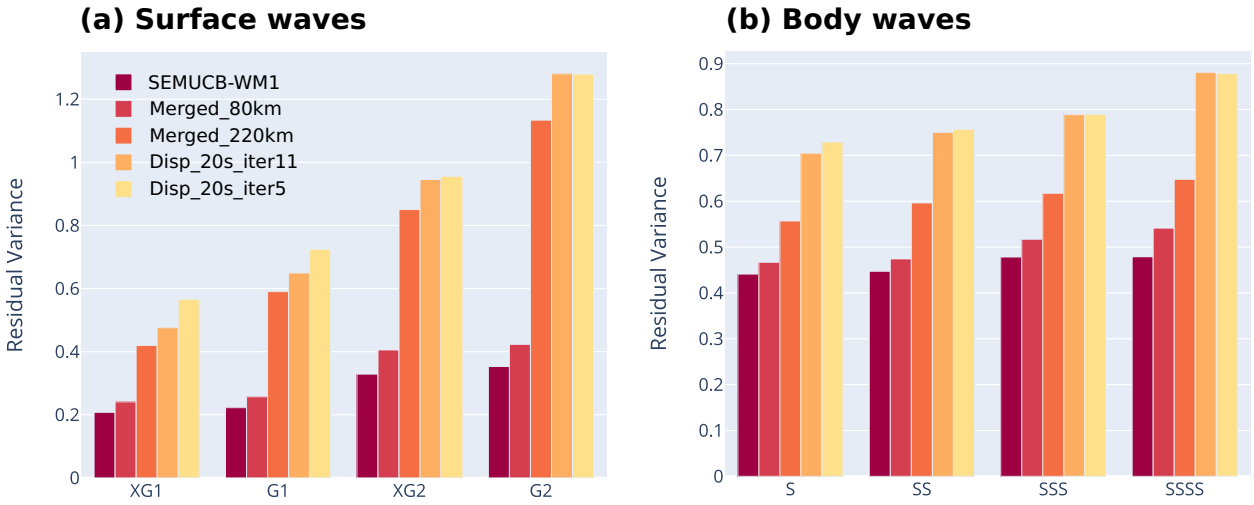

Figure 3.20: Waveform residual variance comparisons for selected surface waves (Panel a) and body waves (Panel b) on the transverse component.

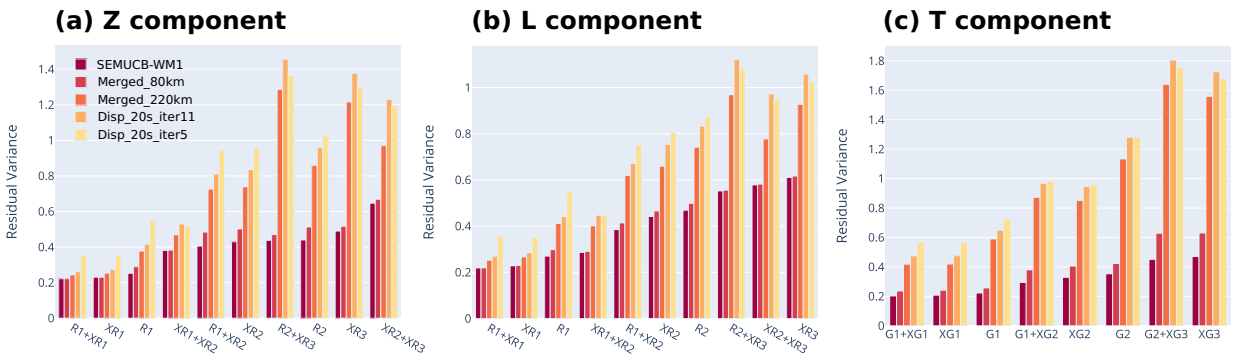

Figure 3.21: Waveform residual variance comparisons for all surface waves on the vertical (Panel a), longitudinal (Panel b) and transverse (Panel c) components.

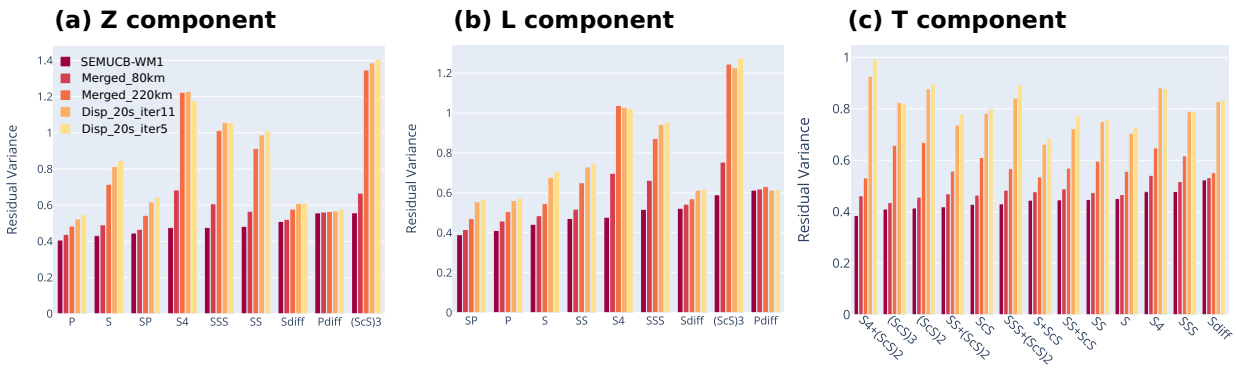

Figure 3.22: Waveform residual variance comparisons for all body waves on the vertical (Panel a), longitudinal (Panel b) and transverse (Panel c) components.

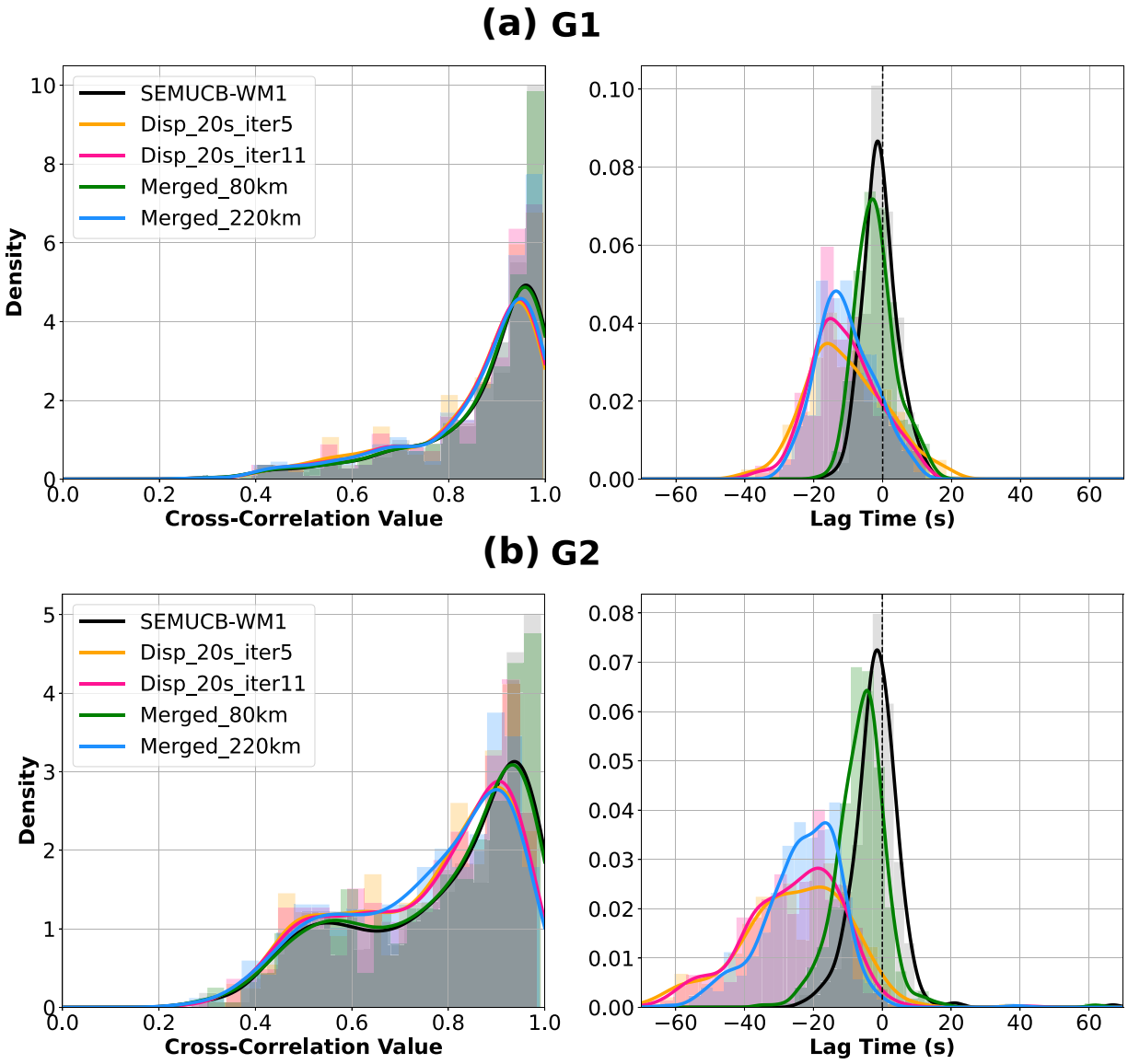

Figure 3.23: Comparisons of histograms and probability density curves for the maximum cross-correlation coefficient (left panels) and corresponding lag times (right panels), for first orbit fundamental mode Love waves (Panel a) and second-orbit Love waves (Panel b).

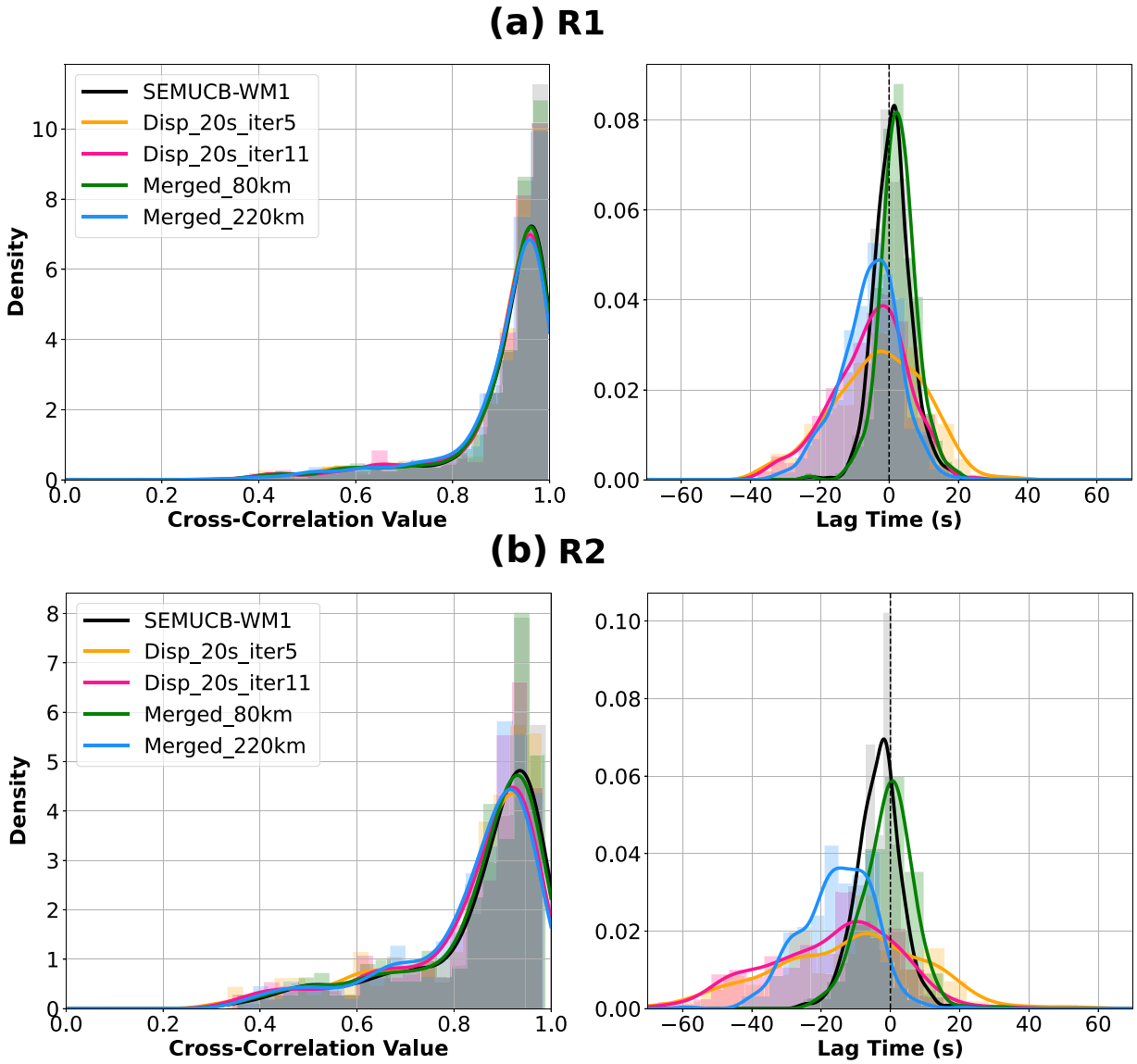

Figure 3.24: Same as Fig. 3.23 for first-orbit (Panel a) and second-orbit (Panel b) fundamental mode Rayleigh waves.

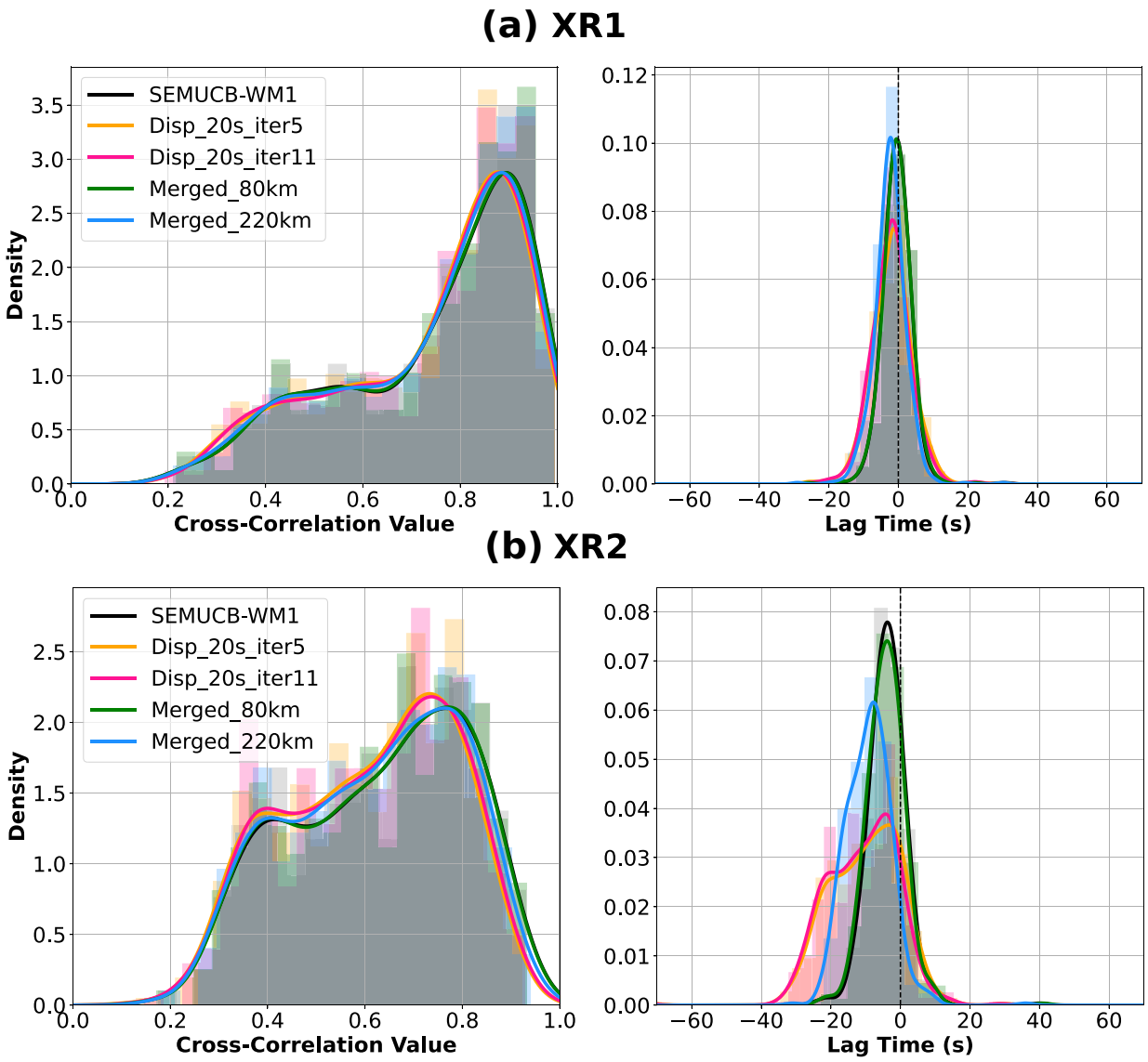

Figure 3.25: Comparison of histograms and probability density curves for the maximum cross-correlation coefficient (left panels) and corresponding lag times (right panels) for the 5 models considered. Panel (a) and (b) show the results for first-orbit XR1 and second-orbit XR2 wavepackets, respectively.

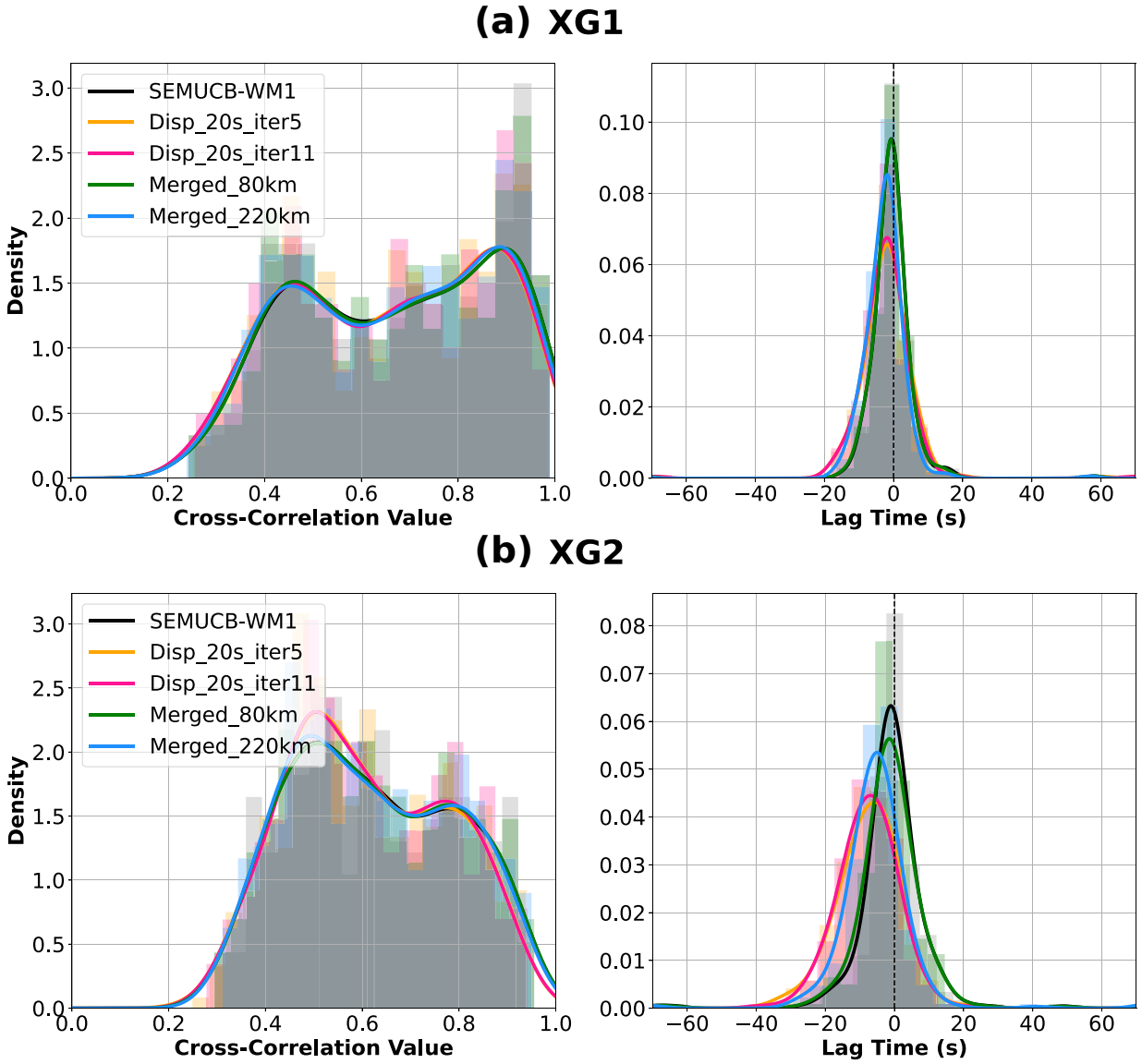

Figure 3.26: Comparisons of histograms and probability density curves for the maximum cross-correlation coefficient (left panels) and corresponding lag times (right panels). Panel (a) and (b) show the results for first-orbit XG1 and second-orbit XG2 wavepackets, respectively.

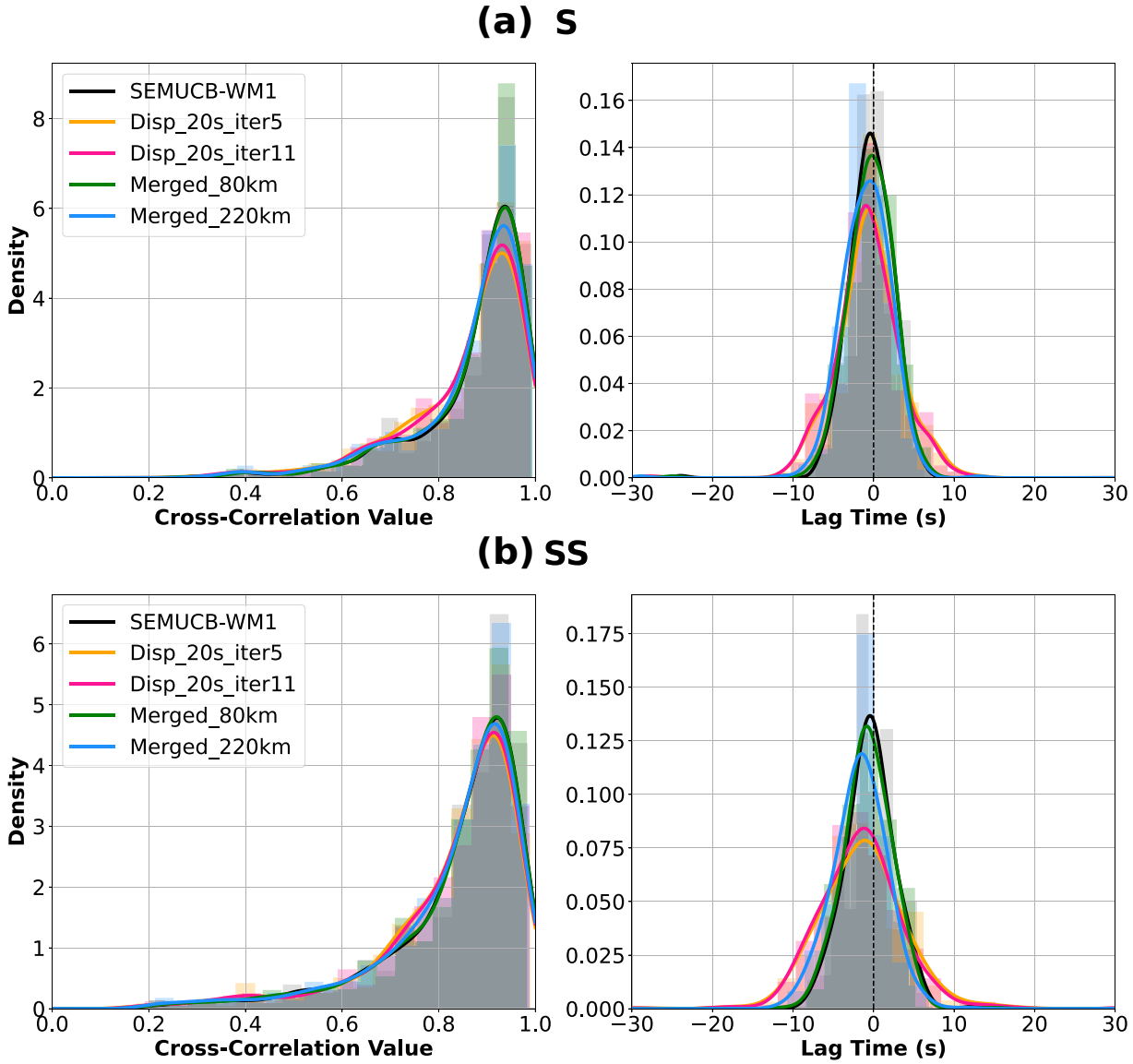

Figure 3.27: Same as Figures 3.14 and 3.15 for body waves: S waves are shown in Panel (a) and SS waves in Panel (b).

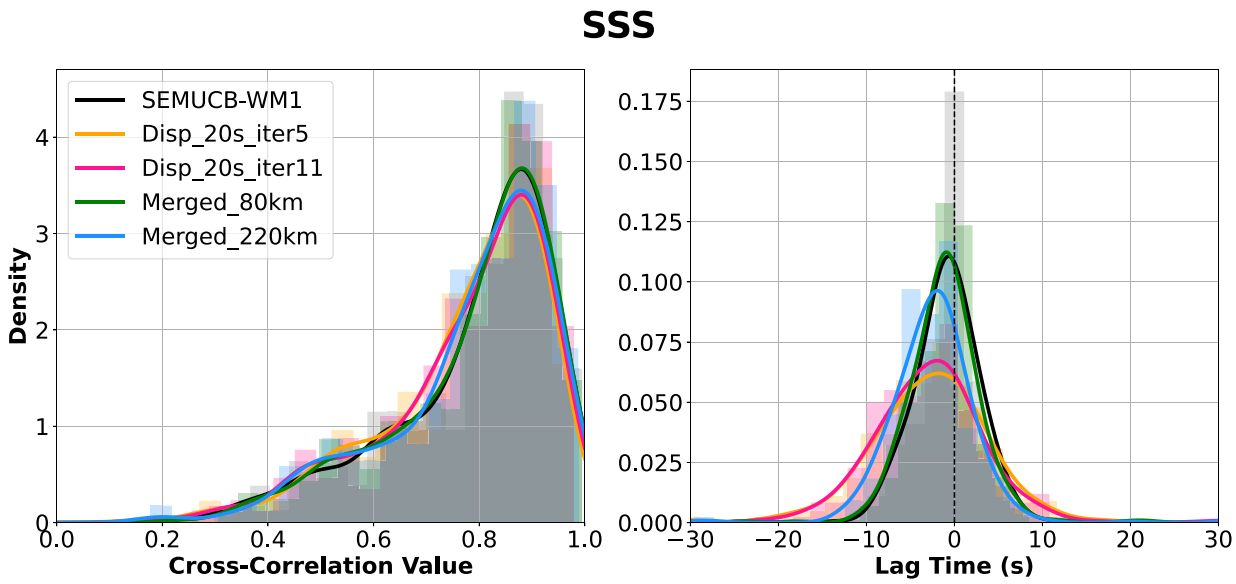

Figure 3.28: Comparison of histograms and probability density curves for the maximum cross-correlation coefficient (left panel) and corresponding lag times (right panel) for the body wavepackets containing SSS.

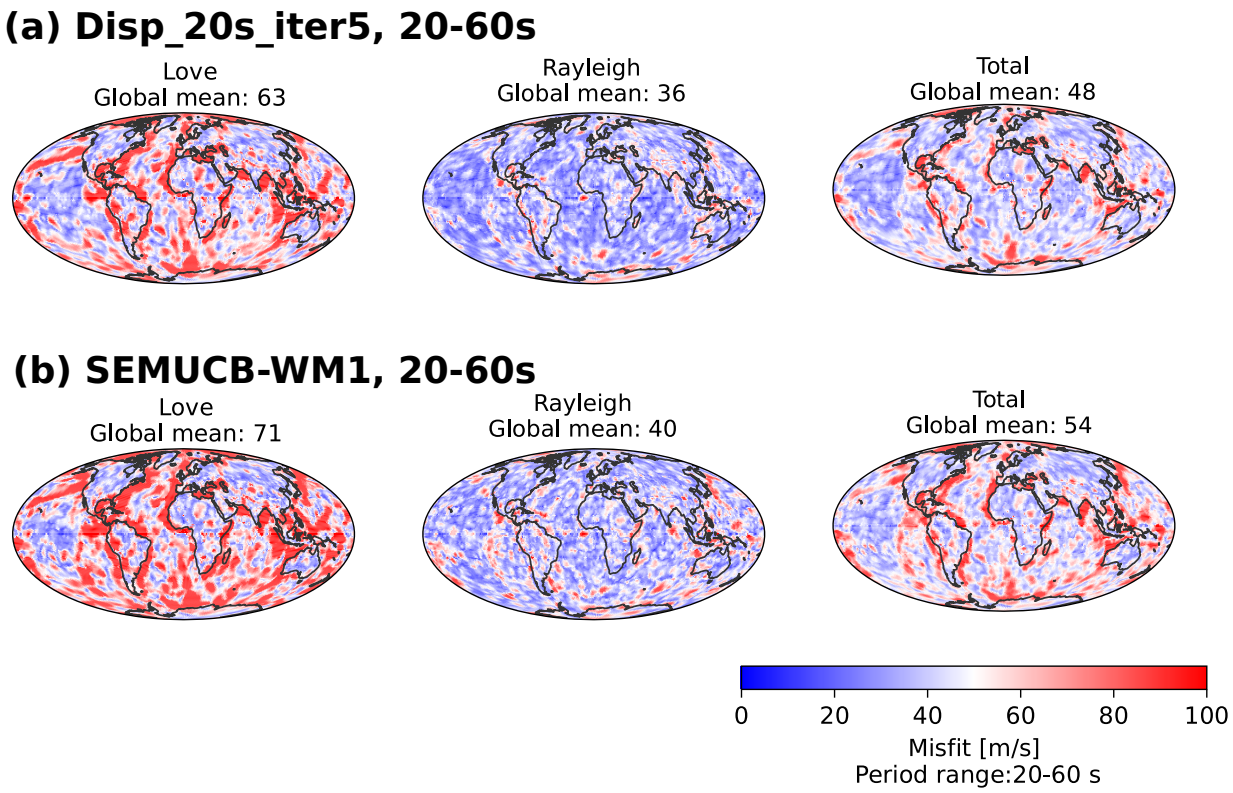

Figure 3.29: Spatial distribution of absolute mean L1 dispersion misfits for model Disp_20s_iter5 (Panel a) and SEMUCB-WM1 (Panel b) in the period range 20-60s.

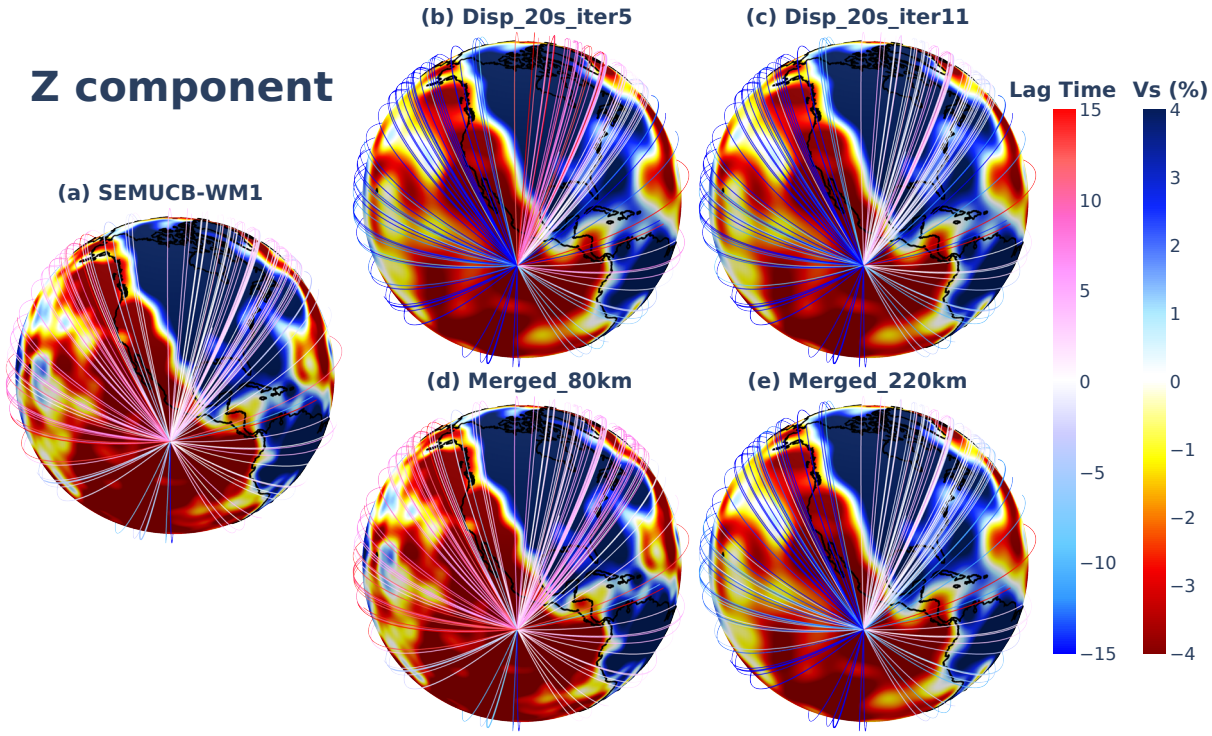

Figure 3.30: Source-station paths for event 4 (Table 3.1), color-coded by the measured lag-time for R1 for each of the 5 models considered, as indicated in the subtitles above each map. The background maps show lateral variations in Vs at 100km depth for each model, centered on the eastern Pacific.

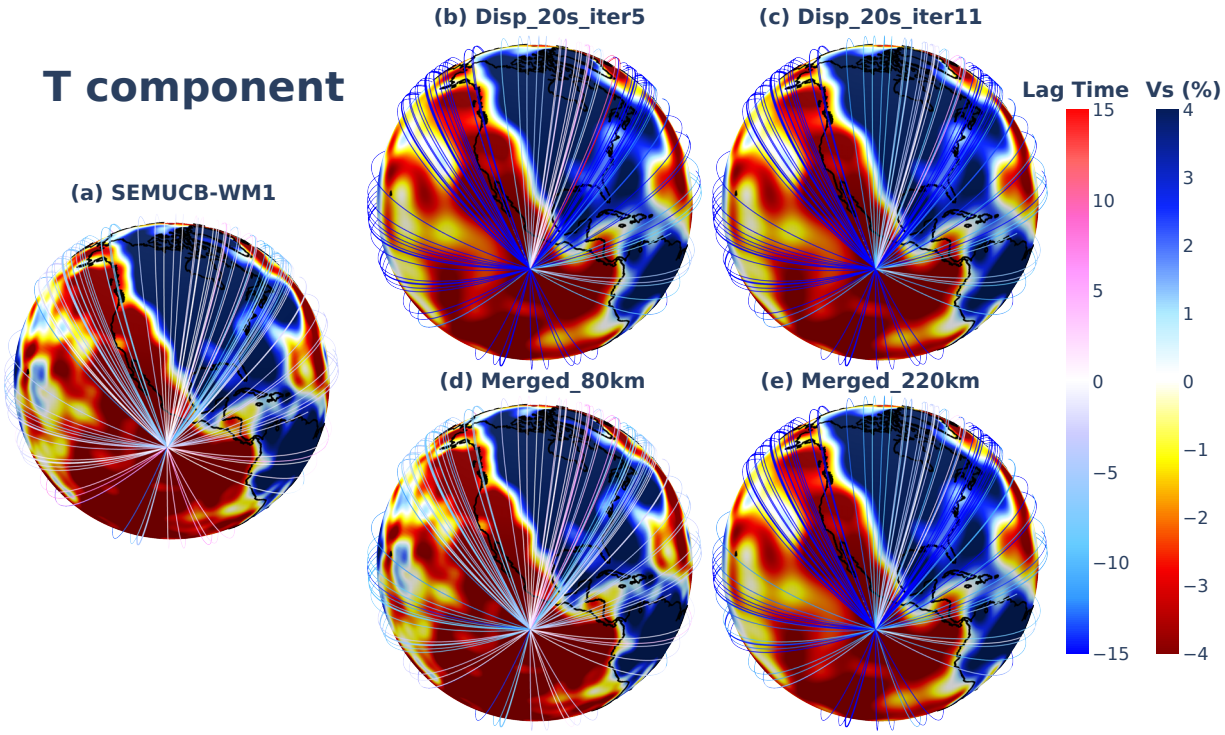

Figure 3.31: Source-station paths for event 4 (Table 3.1) are color-coded by the measured lag-time for G1. The background maps show lateral variations in Vs at 100km depth for each model (indicated as subtitle to each plot), centered on the eastern Pacific.

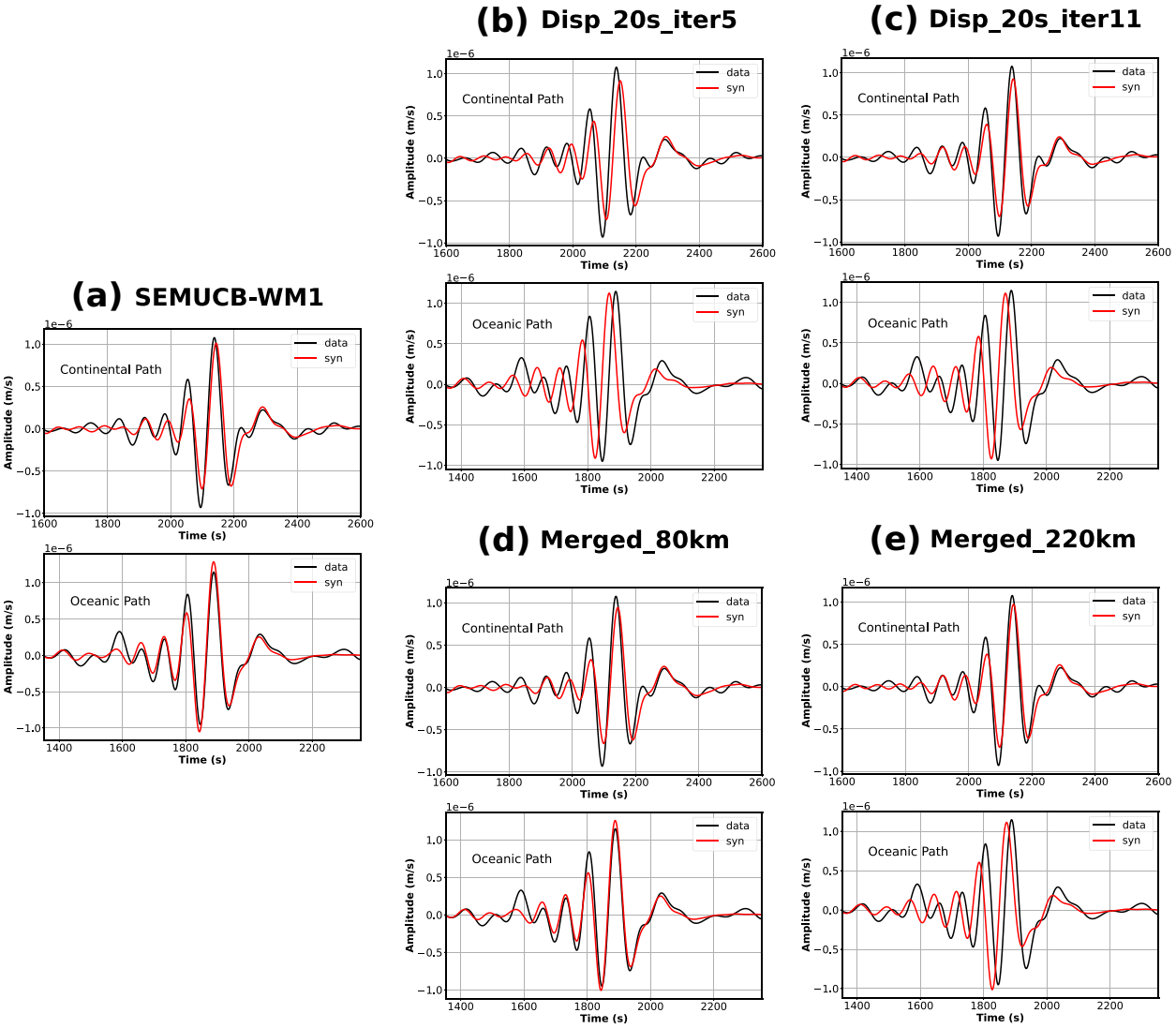

Figure 3.32: Comparison of R1 synthetic (red) and observed waveforms (black) computed in models (a)-(e) as labelled in Figure 3.17. For each of the five models, two panels are shown: the upper panels show the comparison for an oceanic path (event 4 to station RAR) and the lower panels for a continental path (event 4 to station ALE).

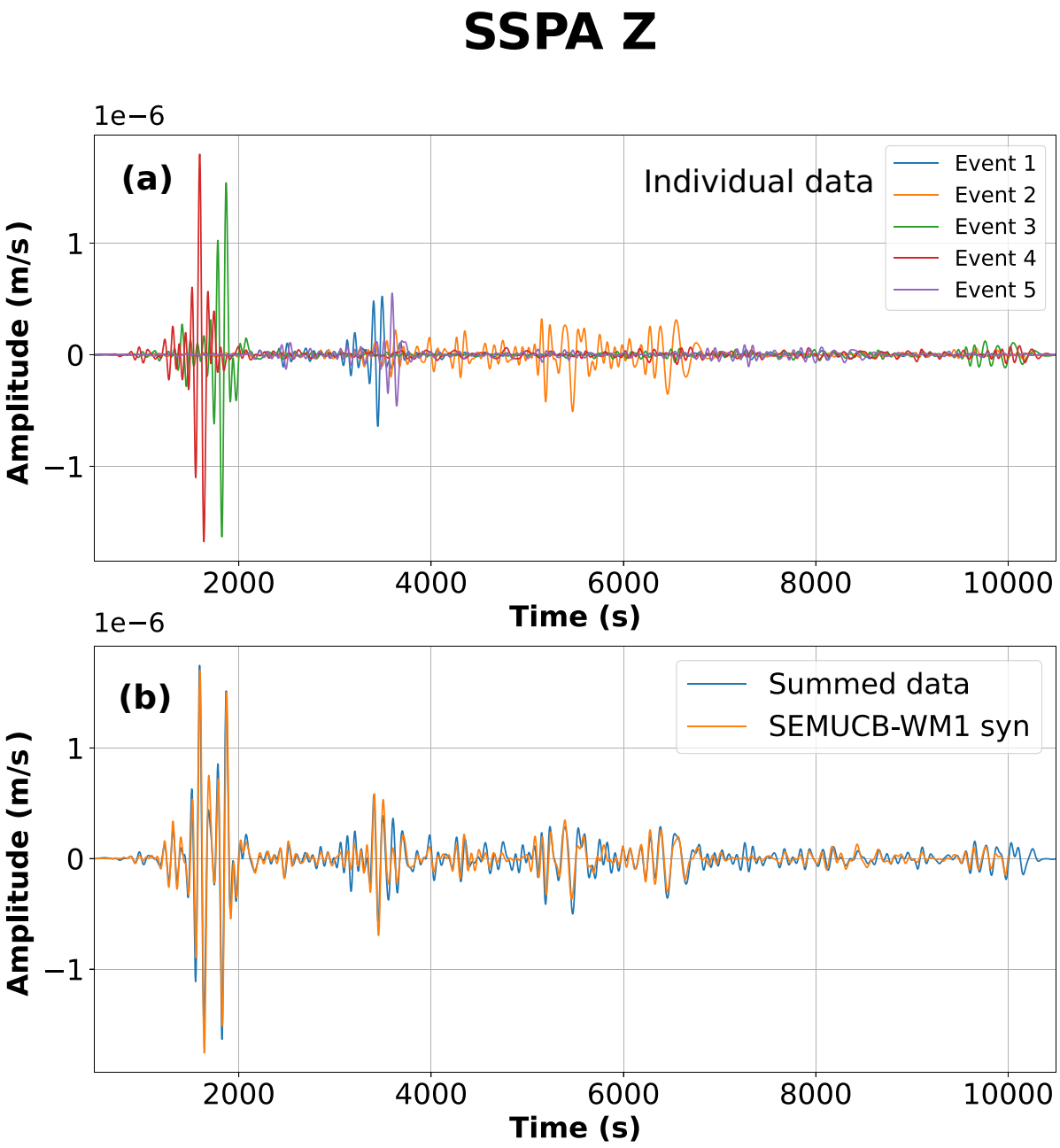

Figure 3.33: Panel (a): Examples of vertical component 10,000s long normalized waveforms at station SSPA. The traces for the 5 selected events used in the paper are shown, aligned at their origin time and normalized by their seismic moments. Panel (b): the corresponding summed waveform compared with source-stacked synthetic waveforms computed in model SEMUCB-WM1.

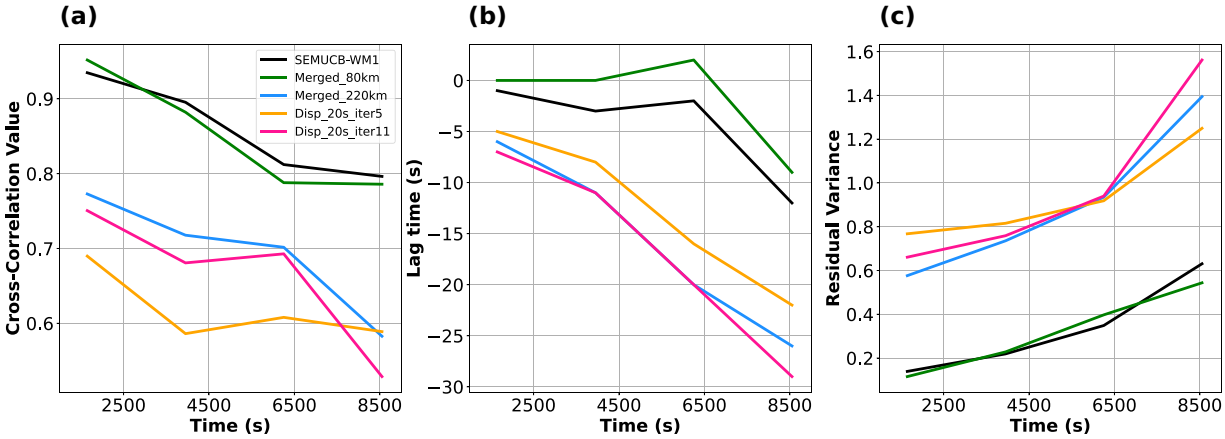

Figure 3.34: Cross-correlation coefficient, lag time, and residual variance in the time windows (500-2800s, 2800-5100s, 5100-7400s, and 7400-9700s) of waveforms from all 5 events (Table 3.1) summed over all 5 events and at 3 stations (DWPF, SSPA, POHA), on the vertical component for the different models. The traces are aligned on the origin time of the events before summation.

Chapter 4

Applications of source stacking followed by cross-correlation for full-waveform tomography II: the first application to realistic long period synthetic data at the global scale

---

The continuous work from chapter 2 and chapter 3 is presented here in modified form to fit coherently into this dissertation.

4.1 Motivation

In chapter 2, a series of experiments was conducted based on synthetic waveforms computed in a simple model "BLOB", with and without noise added. Also, a new cost-function using pair-of-stations cross-correlations of the summed traces was introduced to overcome the issue of dominance of high-amplitude fundamental mode surface wave in summed waveforms (Fig. 2.2), to make it possible to enhance the contribution of time windows containing overtones and body waves (Fig. 2.3 and 2.4), and to be able to apply path weighting to the data. However, in order to apply the methodology to real data, more tests on realistic 3D earth models need to be performed, in particular to investigate whether the methodology can handle the effects of complex crustal and uppermost mantle structure on the waveforms. In chapter 3, the capability of a dispersion-based model of the crust and uppermost mantle as starting model for global full-waveform inversion (FWI) was explored. A workflow based on the concept of effective medium, for constructing an equivalent model by inverting fundamental mode Rayleigh and Love wave dispersion data was developed. In the present chapter, the investigations will continue on the stacked-source methodology by performing tests in a realistic 3D earth model with an effective 3D crust. Also, the strategies are proposed and tested to overcome the "missing data" issue in real data application.

To evaluate the capability of resolving the shear wave velocity and its anisotropic structure in the upper mantle, synthetic tests are conducted based on a synthetic dataset derived from a realistic earth model, modified from model SEMUCB-WM1 (e.g., French and Romanowicz 2015), which includes both a 3D crust and 3D radially mantle structure. Since the original SEMUCB-WM1 presents strong radial anisotropy in the depth range from 30km to 50km in ocean basins, which is due to the way the crust was constructed and is likely unrealistically strong, the radial anisotropy profile in this depth range is replaced by a smoother one for the purpose of the present synthetic tests (Fig. 4.1b right panel). The modified SEMUCB-WM1 is parametrized horizontally in spherical splines (e.g., Wang and Dahlen 1995) of "level 6" in isotropic Vs (node spacing=2°) and "level 4" in ξ (node spacing=7.9°), to determine how well can long wavelength structure be retrieved (the relatively long period bandpass is considered here: cut off frequencies of 60 and 400 s and corner frequencies of 250 and 80s). However, in the inversion, the parametrization for Vs and ξ are limited to "level 4". Consequently, the target model is shown in the figures of this chapter filtered at spherical spline level 4 (node spacing=7.9°), to be compared to the results of inversion shown next.

4.2 Realistic synthetic data and method

4.2.1 Synthetic waveforms and dispersion data

The three-component summed synthetic seismograms of 10,000 s duration are computed using the modified SEMUCB-WM1 for the 273 events and the 515 stations, same as the synthetic experiments for "BLOB model" in chapter 2 (Fig. 2.1). Both minor and major arc

Love and Rayleigh waves are included. As mentioned previously in Section 2.2.2, synthetics are computed using the CSEM (e.g., Capdeville et al. 2003) and the sources are aligned at their origin time, normalized by the seismic moment of each source, with source parameters obtained from the CMT catalog (e.g., Dziewonski et al. 1981; Ekström et al. 2012). The synthetic group velocity dispersion data in the period range 25-150 s is calculated using the normal mode code MINEOS (e.g., Woodhouse 1998) based on the same model. The synthetic dispersion data are used for the recovery tests of 3D crustal model and the uppermost mantle structure. Note that this synthetic data set will be referred to as 'data', while the predictions from each iteration of inversion will be called 'synthetics'.

4.2.2 Inversion workflow

Two independent inversion experiments are conducted: (1) summed waveforms and (2) cross-correlation functions (CCFs) of the summed waveforms across pairs of stations. A fast converging Gauss-Newton optimization approach are used in the inversion. A high-level workflow can be found in Fig. 3.4 and the detailed information in Section 3.2.3. The following steps briefly summarize the process:

1. Starting from the 1-D average of the 3D target model, the structure is iteratively inverted using a model parametrized laterally in spherical splines of level 4 (node spacing=7.9°). Note the radii of cubic b-spline knot distribution (Fig. 3.2b) close to Moho is denser than the parametrization of SEMUCB-WM1 to better accommodate the transition between the crust and the uppermost mantle.

2. Rayleigh and Love group velocity dispersion data are then inverted for the construction of a 3D crustal model (0 km-Moho depth, step 1 in section 3.2.3) and uppermost mantle structure (Moho-170 km depth, step 2 in section 3.2.3).

3. When the dispersion inversion has converged, by following step 3 in section 3.2.3 and the summed waveforms (respectively, CCFs) are inverted to construct the upper mantle structure in the depth range 80-800 km, using the dispersion-derived model as starting model (and a 1D model at depths greater than 170km). A normal-mode perturbation method NACT (e.g., Li and Romanowicz 1995) is used for kernel computations. The kernels are updated in the new 3D model at each iteration, as mentioned in chapter 2. Also, we gradually relax the damping constraint at each iteration according to the residual variance to develop Vs and ξ structure variations.

4. If the inversion includes missing data, the synthetics are computed by NACT for replacements and will be updated at each iteration of the model.

4.3 Results and discussion

4.3.1 Recovery tests for the realistic synthetic data set

In what follows, recovery tests of inversion using (1) dispersion data in the depth range Moho to 170 km, (2) summed waveforms in the depth range 80-800 km and (3) whole segment CCFs and CCFs with different windows in the depth range 80-800 km are shown.

4.3.1.1 Inversion of dispersion data

The synthetic dispersion data is used to invert for crust and uppermost mantle structure above 170 km depth, using the same procedure as described in chapter 3. Fig. 4.1 shows the recovery results of inverting the synthetic dispersion data for isotropic Vs and ξ, after 10 iterations (Fig. 4.2) while the mean L1 misfit becomes stable after the 7th iteration. The Vs result shows good capability of retrieving structure in the depth range from 60 to 100 km, but not for the range of 30 to 60 km. That could be due to a 3D Moho depth was considered and the crust and the uppermost mantle were not computed in a consistent manner (step 1 and step 2 in section 3.2.3 and Fig. 3.4). On the other hand, the ξ result is less well resolved but, as expected, with larger amplitude loss and some vertical smearing. Note that there is a flipped heterogeneity pattern in ξ at 80 km depth compared to the target model (Fig. 4.1b right panel), which cannot be recovered by only using dispersion data alone.

4.3.1.2 Inversion of summed waveforms

After 10 iterations of the inversion for uppermost mantle structure using group velocity data, the inversion continues but for the depth range of 80-800 km, using summed waveforms. The resulting successive models are shown in Fig. 4.3 and the residual variance reduction evolution in Fig. 4.4. After only 6 iterations, the residual variance became stable and did not evolve significantly thereafter. The Vs result also shows good capability of retrieving the target model in the depth range from 100 to 300 km and promising capability in the depth range from 500 to 700 km, although with weak amplitudes. The ξ structure retrieving capabilities are good in the 100-300 km depth range, but less so in the 500-700 km depth range.

4.3.1.3 Inversions of CCFs across different windows

As discussed in chapter 2, the summed waveforms are dominated by high-amplitude fundamental mode surface waves. To effectively bring out the contribution of smaller amplitude overtones and body waves to the inversion, windowing is applied on cross-correlations of the summed waveforms between pairs of stations. Fig. 4.5 shows CCFs in longitudinal (L) component ordered as a function of distance between station-pairs, and stacked in bins of 100 km in distance. In particular, overtone energy is also distinguishable in the expected group velocity window. The entire segment of CCF is first inverted: a total length of 10,000 s on the casual portion (positive time) are taken into account because of the symmetry of CCF.

Three windows also on the casual portion (positive time) are defined: higher, fundamental, and mixed mode windows, shown in blue, red, and green colors, respectively, in Fig. 4.5. The fundamental mode window is based on the time segment of the emergence of the first-orbit fundamental mode in CCFs. The higher and mixed mode windows are then defined as the time segments before and after the fundamental mode window.

Fig. 4.6-4.8 show the corresponding results for the 4 different inversions: for the entire segment, and the higher, fundamental, and mixed mode windows, respectively. The corresponding residual variance reduction for CCFs can be found in Fig. 4.9. The residual variances have a decreasing trend and are still evolving after four iterations. No further iteration is performed to save computational time for more tasks. The residual variance for the mixed mode windowed CCFs is higher than for the other three inversions. It could be due to the fact that CCFs are complete only up to $\pm$5000 s and large portions of the mixed mode windowed CCFs extend to 10,000s (Fig. 4.5 green).

As concluded in chapter 2, inversions of whole segment CCFs (Fig. 4.6) show similar Vs resolving capability but relatively weaker for ξ compared to the results of summed waveforms, due likely to the slight difference in the damping scheme used or the fact that the CCF is not complete beyond 5,000s as previously mentioned. Fig. 4.7 presents a comparison of the results, for Vs, of inversion of the higher (a), fundamental (b), mixed (c) mode windowed CCFs with the target model. Interestingly, the three results show similar Vs retrieving capabilities with a slight distortion for the mixed mode result. However, differences in the ξ structure of different windowed CCFs are clearer. Only large long-wavelength ξ structure can be retrieved but it is further distorted especially for the mixed mode result (Fig. 4.8).

4.3.2 Model correlation with the modified SEMUCB-WM1

To evaluate the capability of resolving the Vs and ξ structures in the modified SEMUCB-WM1 model (level 6), the correlations are computed as a function of depth of the inverted models at level 4 with the target model, which is the filtered modified SEMUCB-WM1 model at level 4 to fairly compare. For all the correlation results (Fig. 4.10-4.12), slightly low correlation can be found around 80-100km for both Vs and ξ structures. It could possibly be due to the fact that the waveforms are used to invert for the structure below 80 km depth and the waveforms are relatively long period (60-400 s), compared to dispersion data (25-150s), and provide slightly less constraints on the uppermost mantle (80-100km depth).

In what follows, comparisons of the inversion results using different methods and different windowed CCFs are presented. Fig. 4.10 shows the correlations with the modified SEMUCB-WM1 as a function of depth with the methods using the entire cross-correlation (red), source stacking (black), and dispersion inversion (blue), respectively. Note that the dispersion results are only considered above a depth of 150 km, since, in this case, only the depth range of Moho-170km is inverted for. The correlations as a function of depth for Vs are about 70-75% in the upper 500 km with all the methods, and even higher in the uppermost 150km (>85%). As for the ξ parameter, the correlation result of source stacking is high above 350km depth, and gradually decreases with depth. The correlation of CCFs inversion is

lower and has a similar decreasing trend with depth as for source stacking. The correlation of dispersion is high ($> 75\%$) above 100km but quickly decays at increasing depth. To summarize, the methods of source-stacking and CCFs both have high resolution capabilities ($> 70\%$) for Vs, but the resolution of ξ is somewhat lower for the CCFs-based method. Also, the 25-150 s group velocity dispersion data can provide good constraints above 150 km depth except for the ξ structure in the depth range of 100-150 km. However, the inversions for CCFs are not yet converged and the inversion stopped at 4th iteration (Fig. 4.9). This could be the main reason that the correlations of CCFs inversion are lower than for the other two inversions. Also note that all the inversions have been performed without any noise added to the synthetic data. For the real situation with noise, the correlation coefficient could be lower.

Fig. 4.11 shows the correlations as a function of depth with the individual inversions of "all" (blue), "fundamental" (red), "higher" (black), "mixed" (green) mode segments of windowed CCFs. The Vs correlations are all above 50% with all the windowed CCFs and a clear trend of increasing correlation can be found as follows: the result of "all" has the highest, then successively, "higher", "fundamental", and "mixed" modes. The result of ξ also shows a similar trend of increasing correlation with the exception that "fundamental" result is slightly higher than "high" between 340 and 500 km depth. However, only the ξ structure above the depth 380km can be retrieved with a correlation coefficient greater than 50% by using "all" CCFs. Although the result of "mixed" shows the lowest correlation compared to others, note that the later portion of CCFs (i.e. "mixed" mode portion) can still provide some resolving capability, at least for the Vs structure. This is likely because the Green's function is not fully reconstructed with the distribution of events considered (as would be the case for a uniform distribution of sources, e.g., Stehly et al. (2006)). Interestingly, the experiment shows that the different segments of CCFs individually provide different constraints on the earth structure. It will provide insights for applying a proper weighting scheme to combine their contributions for future inversions.

4.3.3 Preparation for real data application

4.3.3.1 Real data collection

Having in mind a first application to real data, we have collected a real global long-period three-component waveform dataset including (1) waveform data of 355 events and 340 stations distributed around the globe (Fig. 4.13) for a record length of 10,000 s in order to include first and second orbit surface waves, filtered in the period range 60-400 s, and (2) a global dataset of fundamental mode Rayleigh and Love group velocity dispersion data (M. Ritzwoller, personal communication, 2009) in the period range 25-150 s. Data selection is important here. In the conventional FWI workflow, synthetics in a reference 3D model for individual traces are usually computed for data quality control and selection. For the sake of saving computation time, individual synthetics will not be computed here. Instead, synthetic waveforms from ShakeMovie (global.shakemovie.princeton.edu), computed in an existing 3D

mantle model, S362ANI (e.g., Kustowski et al. 2008), are used for quality control, under the criteria below:

1. cross-correlation coefficient > 0.6 between synthetics and data
2. time shift should not greater than ± 30 s
3. SNR > 3, where SNR described in the section 2.3.2.1

Table 4.1 shows the availability of picked three-component waveforms picked in this study.

However, ShakeMovie only provides 6,000 s-long synthetic waveforms which is not sufficient for data quality control for the entire 10,000 s long records, starting at event origin time. To avoid keeping records with glitches or large-amplitude noise in the observed waveforms after 6,000 s, combining these with NACT to compute 10,000 s long synthetics will be implemented for further quality control in the near future. An example of comparison of data with Shakemovie synthetics is shown in Fig. 4.14.

4.3.3.2 Missing data

An obvious obstacle for application of source stacking to real data is that not all stations, and not all three components of all stations are usually available. The issue of missing data is under investigations by combining two different strategies: (1) replacing missing waveforms by synthetics computed in the current 3D model, using an efficient approximate method such as NACT (e.g., Li and Romanowicz 1995), and (2) grouping events and stations in order to minimize missing waveforms in each group.

To understand the effect of replacing missing waveforms, a series of synthetic experiments is conducted. two sets of synthetic summed waveforms, removing 25% and 55% of data, are respectively generated. The numbers represent two extreme conditions from the real experience of collecting data. In order to generate the synthetic data with missing data in a source stacking manner (triggering multiple sources simultaneously), the event dataset is divided into 25 subgroups of events and randomly removed 25% or 55% of stations in each event subgroup. 25 subgroups of summed waveforms are generated and then stacked. In the following section, the inversion results will be discussed.

Also, the possibility of grouping is investigated based on a real data picking result in previous section (section 4.3.3.1), to minimize the number of missing waveforms in each subgroup. First, a series of indexes of 1 or 0 is assigned to each source-station pair, to label the existing or missing data. Next, there are three steps for grouping according to the similarity of series of indexes for different events.

1. Consider 3 different groups, one for each of the 3-component data
2. Cluster into 3-4 event subgroups based on the similarity of the series of 1's and 0's for different events.

3. Cluster into 3-4 station subgroups based on the similarity of the series of recorded events.

This results in approximately 27-48 subgroups for which separate SEM simulations will need to be performed for each iteration. However, comparing this procedure to the conventional FWI, it is still a large computational saving.

4.3.3.3 Effect of missing data and grouping

While application to real data requires additional steps to account for the missing data issue, the inversion results for 0% (black), 25% (red) and, 55% (blue) of missing data (Fig. 4.12) are tested in order to investigate the impact of missing data and to know how well does the replaced waveform works. The correlation for both Vs and ξ is, not surprisingly, strongest with the result of 0% missing data and weakest with the result of 55% missing data. The experiment shows that, with 75% of available data, the correlation of Vs and Xi structures is about 50%-75%. When only 45% of data are available, only the Vs structure above the depth 250km can be fairly retrieved (correlation coefficient > 0.5). The results further indicate the acceptable data availability in each subgroup, which will be discussed in the next section.

Following the procedure of grouping, Fig. 4.15 shows the comparison of data availability before and after grouping. The original data availability before grouping for horizontal components is 47% and for the vertical component it is 56%. After grouping, the data availability increases and is around 60-66% with the vertical component slightly higher than the horizontal components. There are 36 total subgroups generated in this experiment. Although the number of computations is increased (compared to a single computation in the ideal case), higher data availability could enhance the capability of the stacked source methodology according to Fig. 4.12.

4.4 Summary

The synthetic tests show the capability of retrieving both Vs and ξ perturbations (at long wavelengths) in the upper mantle for a target model similar to SEMUCB-WM1, using source stacking and source stacking followed by cross-correlation. Similar resolution capabilities for Vs are shown using the 3 different time intervals of the CCFs considered. But the ξ results show that the different segments in CCFs provide varying resolving capabilities. This indicates that a weighting scheme can possibly be applied to increase the contribution of time windows containing overtones and body waves. A new global real dataset that includes 355 events and 340 very broadband recording stations is collected. The issue of missing data is under investigations in two ways (1) by replacing missing data by NACT-computed synthetics and on the other (2) forming smaller groups of data so as to minimize the number of missing waveforms in each group. The experiments show promising capability of the stacked event approach with 25% to 55% of missing data, further enhanced by applying event grouping.

4.5 Table

	Available Waveforms	Total Waveforms	Availability
E	19,197	40,600	47.3%
N	19,288	40,600	47.5%
Z	47,138	84,681	55.7%

Table 4.1: The available three-component waveform picked in the study.

4.6 Figures

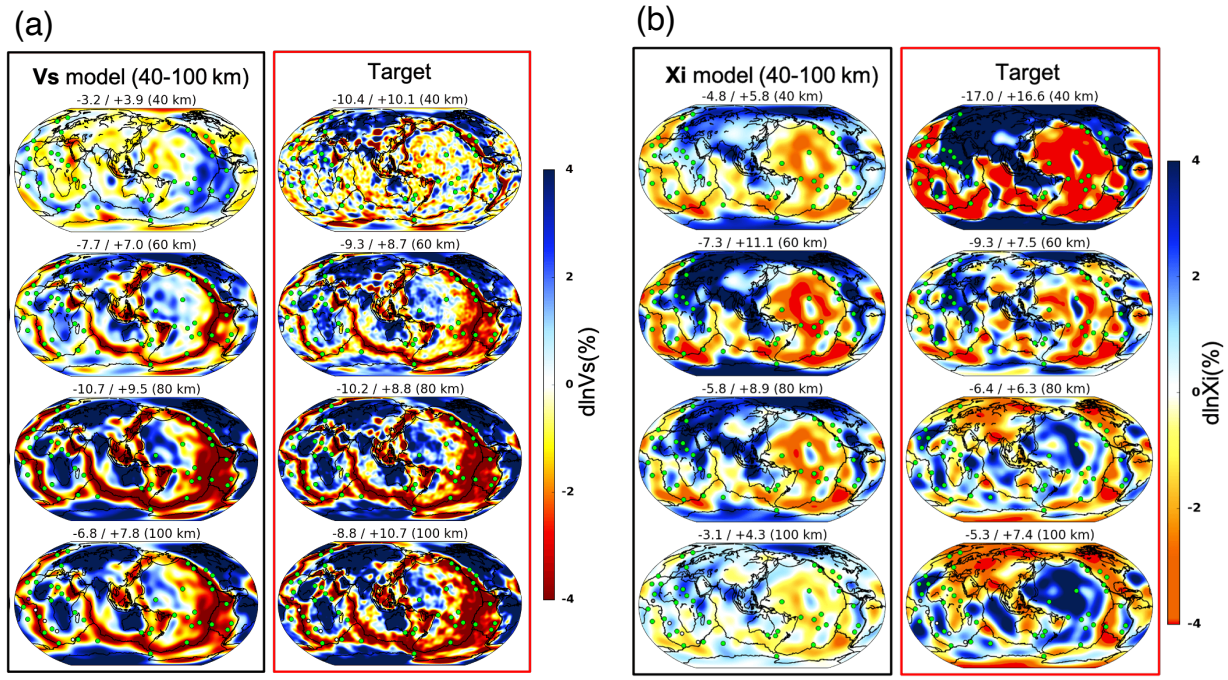

Figure 4.1: (a) The results of inversion for Vs in the depth range 40-100 km obtained by inversion of group velocity data, shown after 10 iterations. Vs variations are plotted in percent with respect to the global mean. Upper left corner of each map represents maximum lateral variations. (b) Same as in (a) for the anisotropic parameter ξ. The target model (modified SEMUCB- WM1) is shown in the rightmost column of each panel.

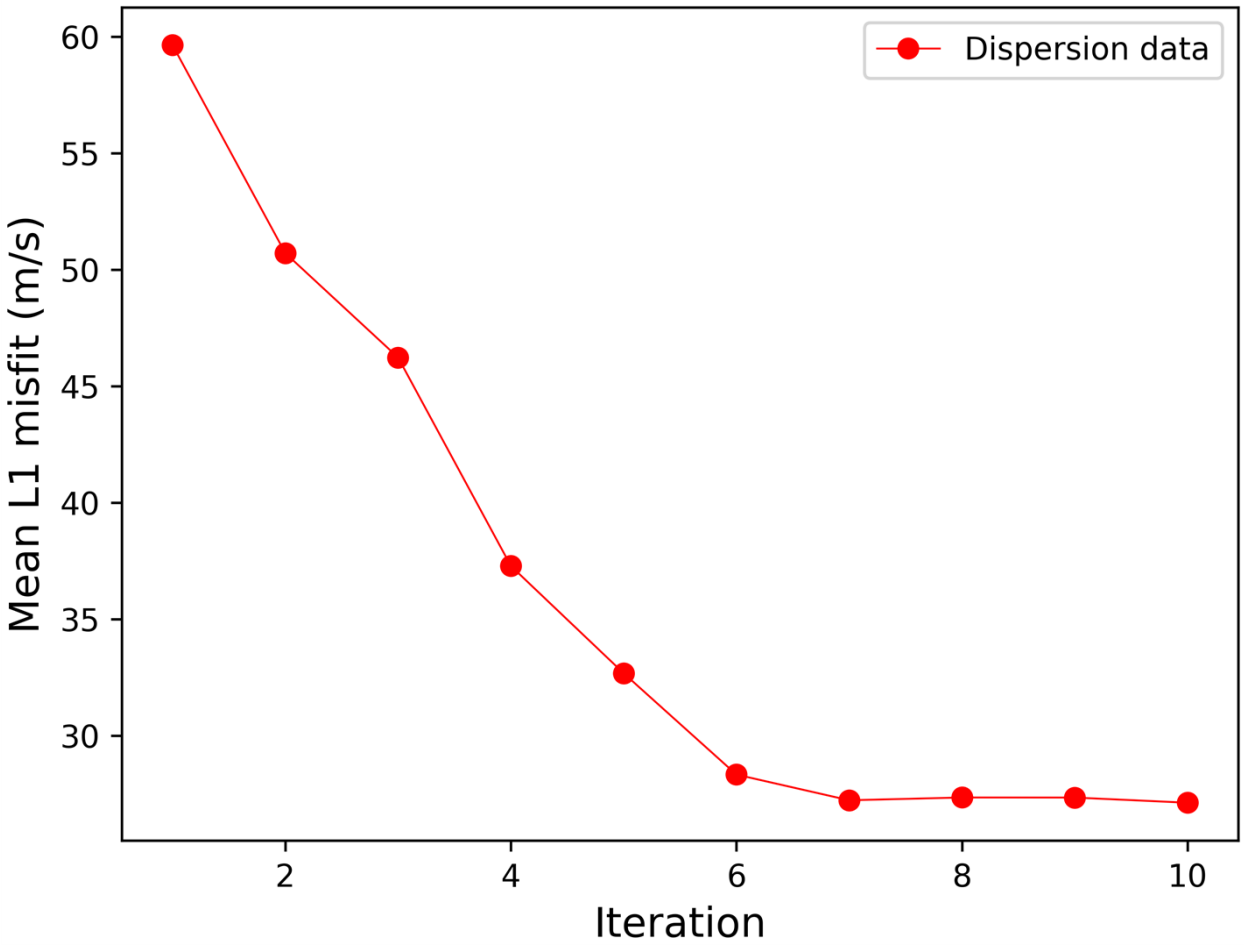

Figure 4.2: The evolution of mean L1 misfit of dispersion data as a function of iteration. The mean L1 misfit becomes stable after the first 7 iterations.

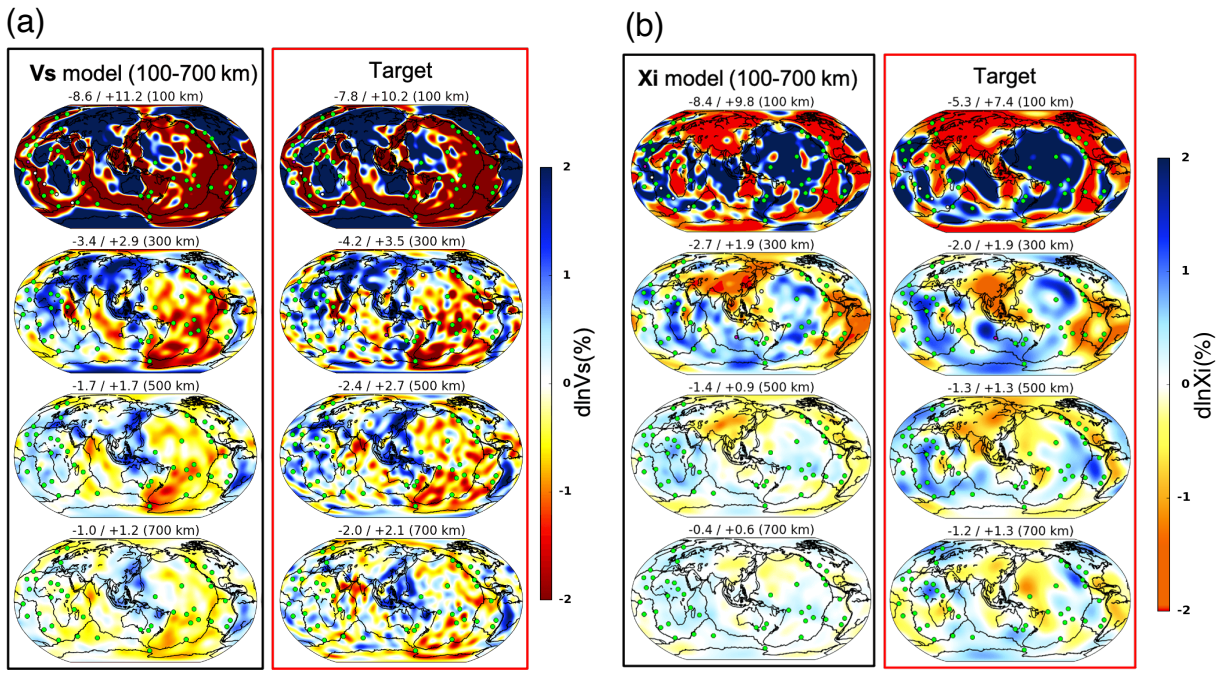

Figure 4.3: (a) The results of inversion for Vs in the depth range 100-700 km obtained by inversion of group velocity data after 10 iterations, followed by 6 iterations of summed three-component waveform data. Vs variations are plotted in percent with respect to the global mean. Upper left corner of each map represents maximum variation. (b) Same as in (a) for anisotropic parameter ξ. The target model (modified SEMUCB-WM1) is shown in the rightmost column of each panel.

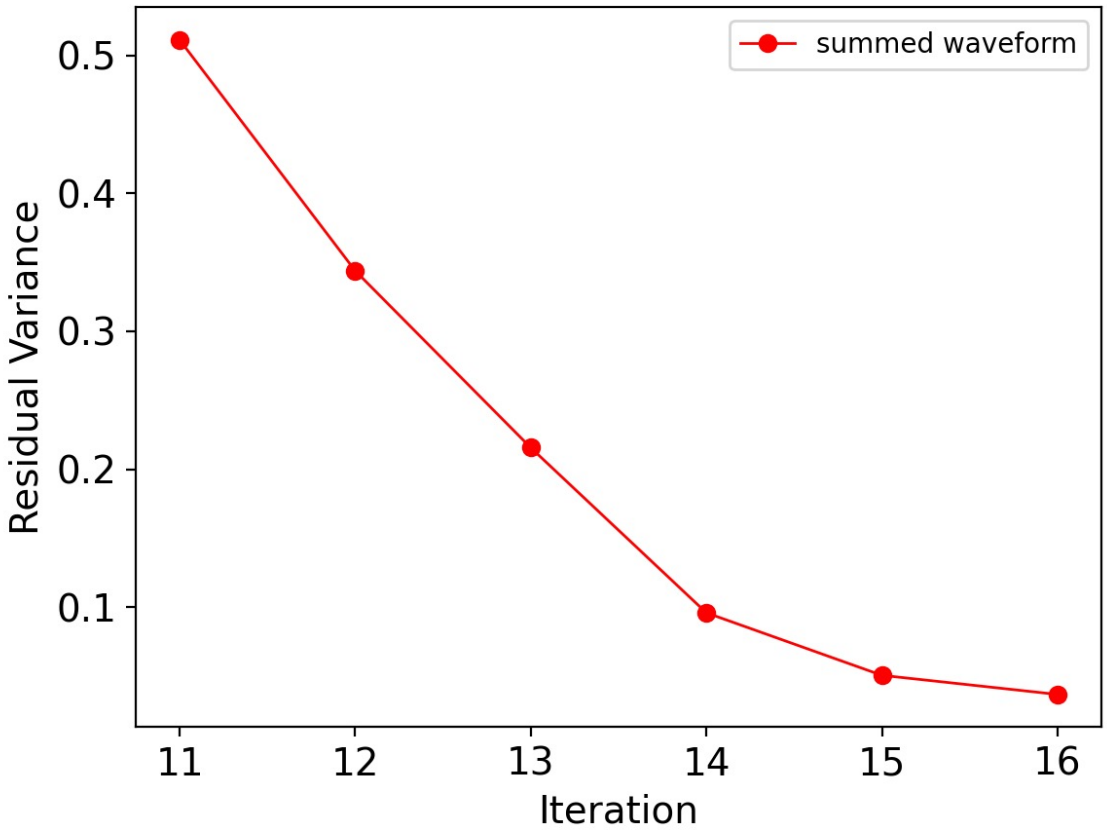

Figure 4.4: The evolution of residual variance of summed waveforms as a function of iteration.

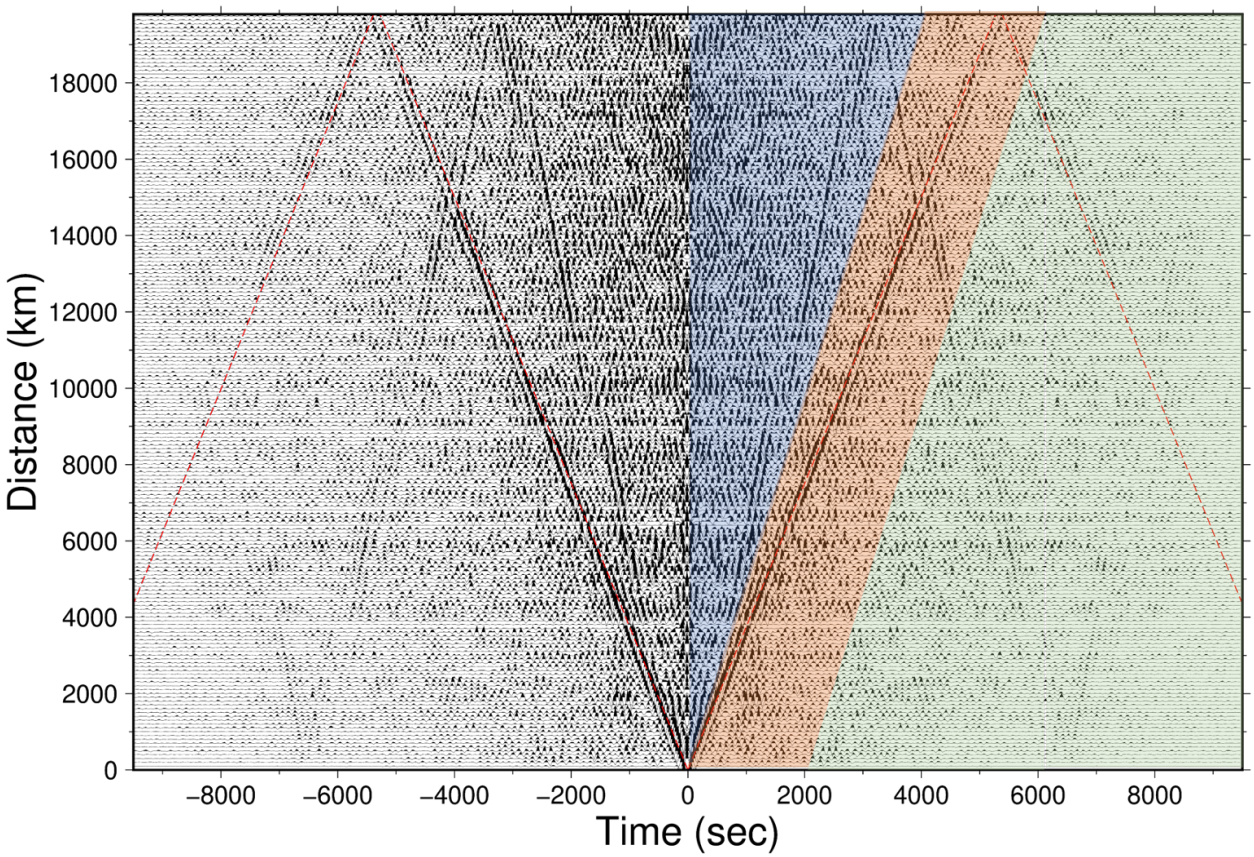

Figure 4.5: Cross-Correlations in longitudinal (L) component of stacked waveforms between station pairs. The traces are sorted by distance. Each trace is obtained by stacking the cross-correlations of the station pairs in bins of 100 km in distance. Blue, red, and green segments represent the higher, fundamental, and mixed mode segments, respectively, as discussed in the text. Note that clear overtone and fundamental mode energy can be found in CCFs.

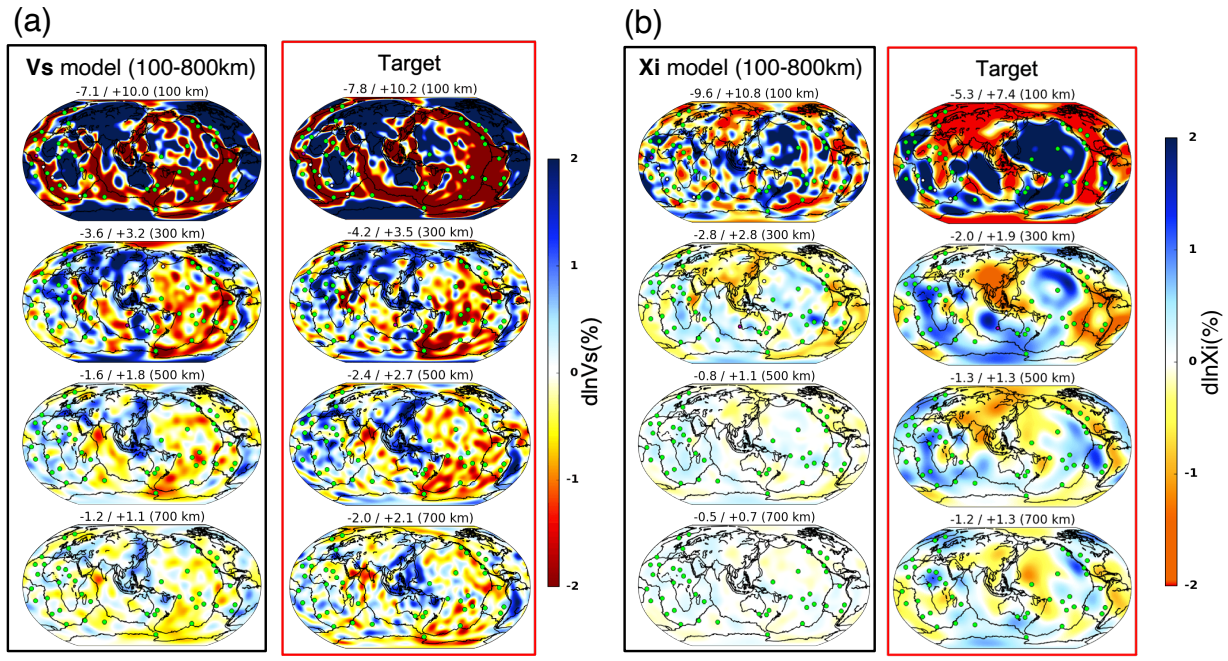

Figure 4.6: Same as Figure 4.1 and 4.3 for inversions performed using group velocity data (10 iterations) and cross-correlations of the summed waveforms (6 following iterations). (a) Results of inversion for Vs (b) Results of inversion for ξ. The target model (modified SEMUCB-WM1) is shown in the rightmost column of each panel.

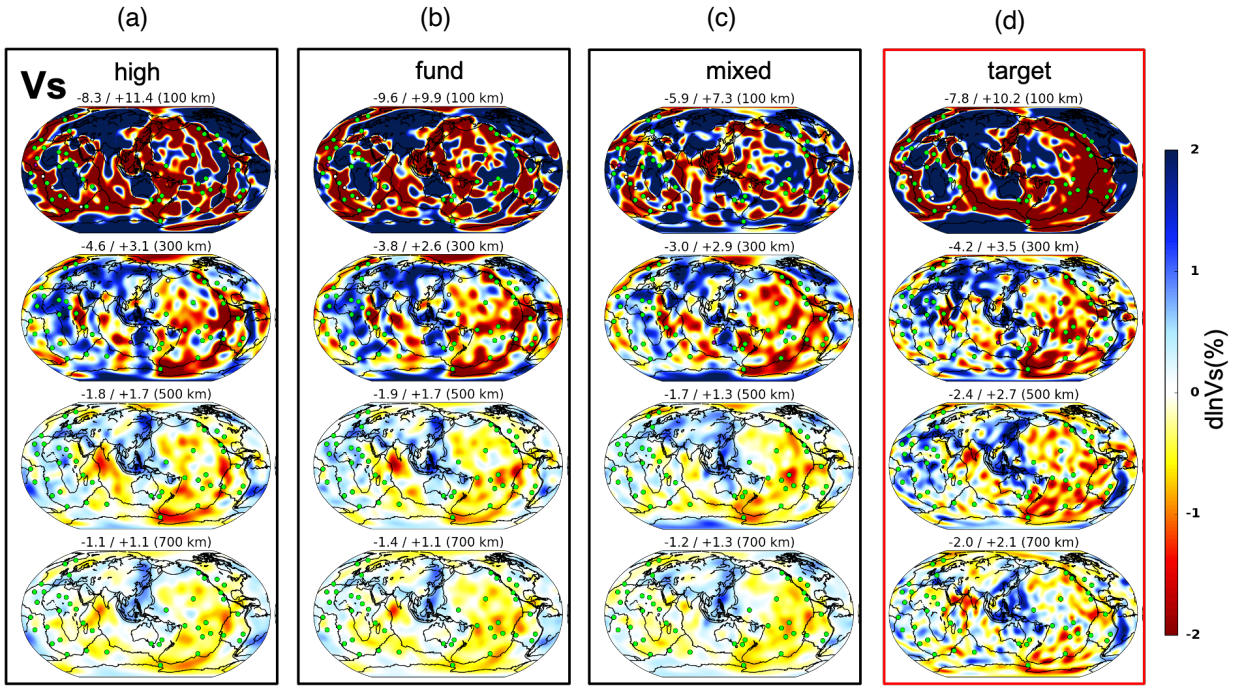

Figure 4.7: Same as Figure 4.1 and 4.3 for Vs inversions performed using group velocity data (10 iterations) and higher, fundamental, mixed mode windowed CCFs from left to right (4 following iterations). The target model (modified SEMUCB-WM1) is shown in the rightmost column. Vs variations are plotted in percent with respect to the global mean. Upper left corner of each map represents maximum variation. Interestingly, the result shows similar retrieving capabilities of the higher, fundamental, and mixed mode windowed CCFs.

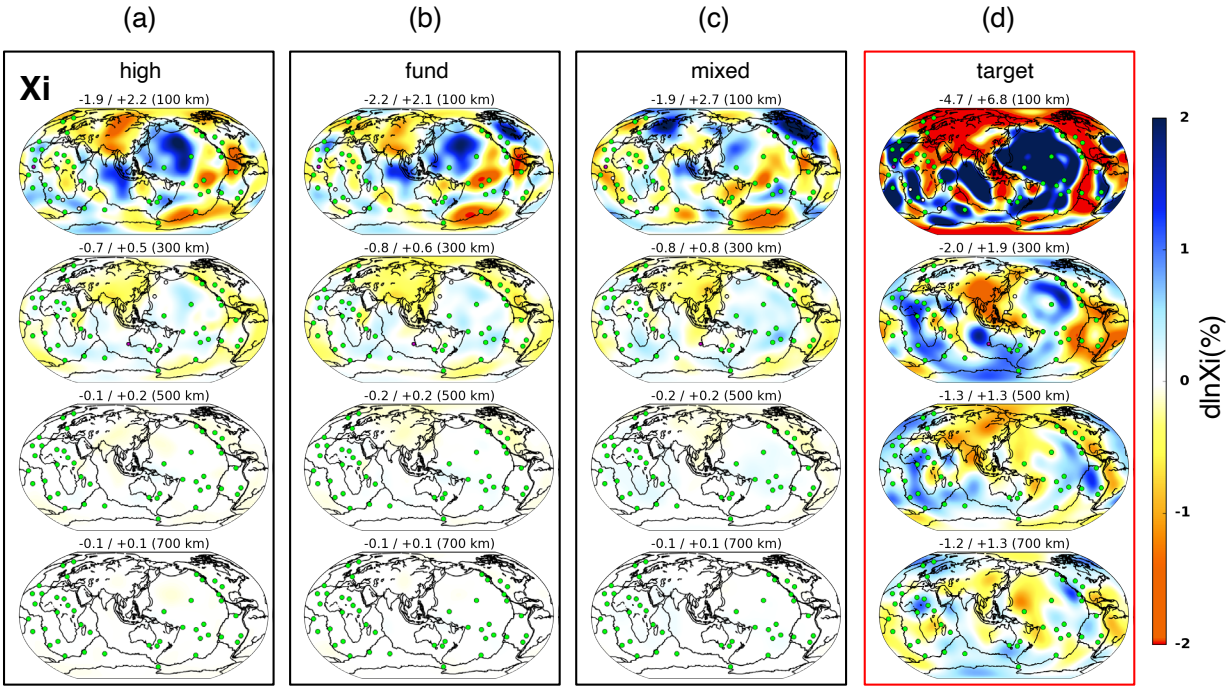

Figure 4.8: Same as Figure 4.1 and 4.3 for ξ inversions performed using group velocity data (10 iterations) and higher, fundamental, mixed mode windowed CCFs from left to right (4 following iterations). ξ variations are plotted in percent with respect to the global mean. Upper left corner of each map represents maximum variation. The target model (modified SEMUCB-WM1) is shown in the rightmost column.

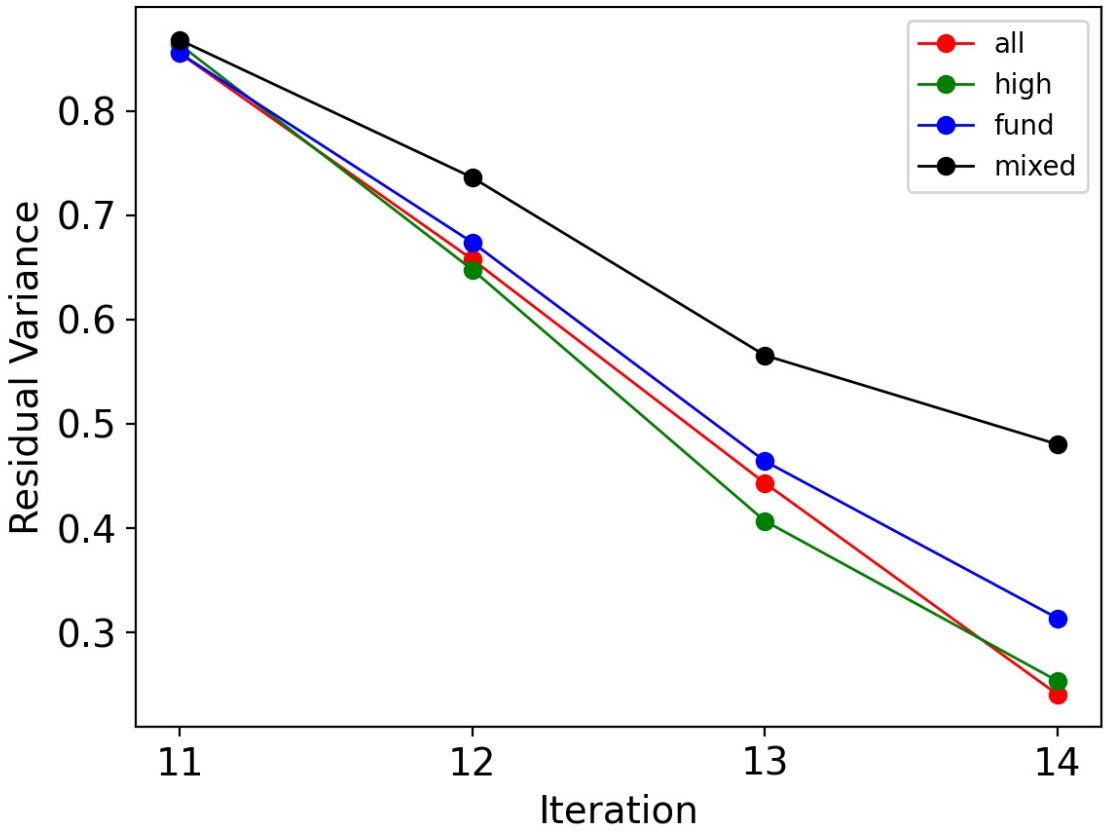

Figure 4.9: The evolution of residual variance of higher, fundamental, mixed mode windowed CCFs, and whole segment CCF as a function of iteration. The residual variances for the higher mode windowed CCFs are slightly lower than for the other three.

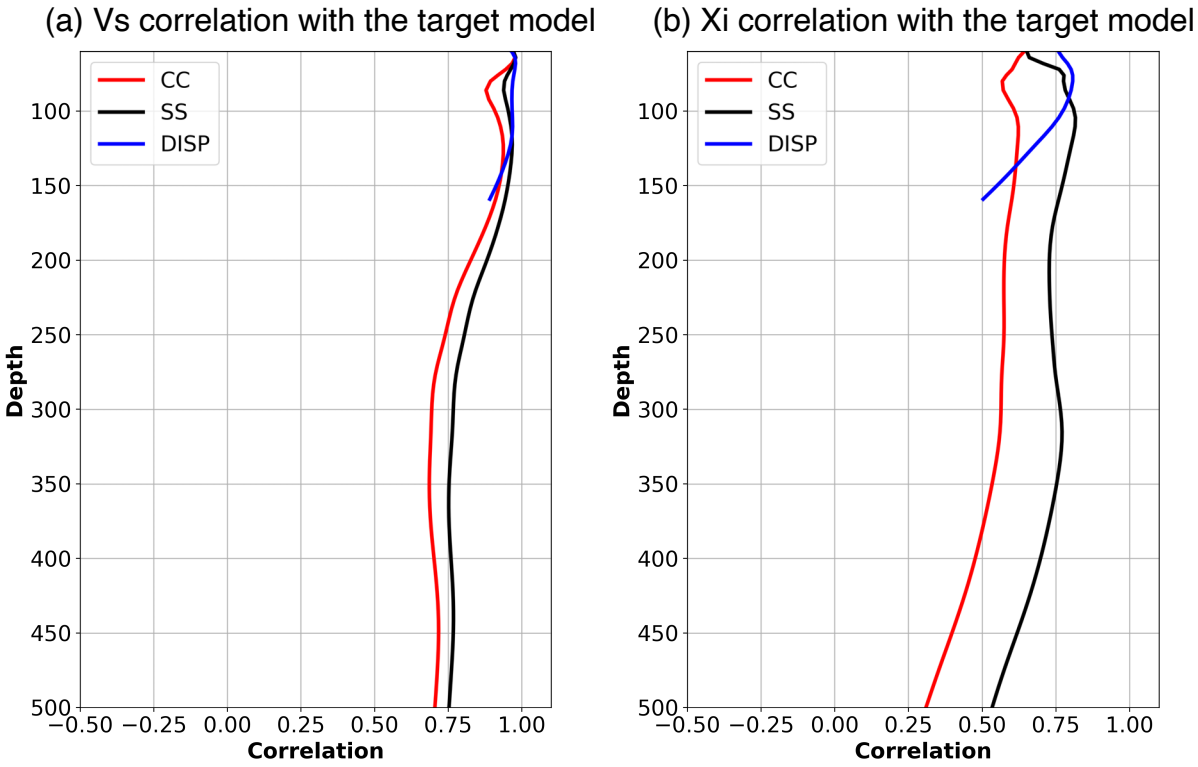

Figure 4.10: Model correlation to the target model (modified SEMUCB-WM1) of the Vs (a) and ξ (b) models as a function of depth, for the models inverted by using different methods. CC, SS, and DISP refer to the methods of cross-correlation, source stacking, and dispersion inversion, respectively. Note that the DISP model above 150 km is considered.

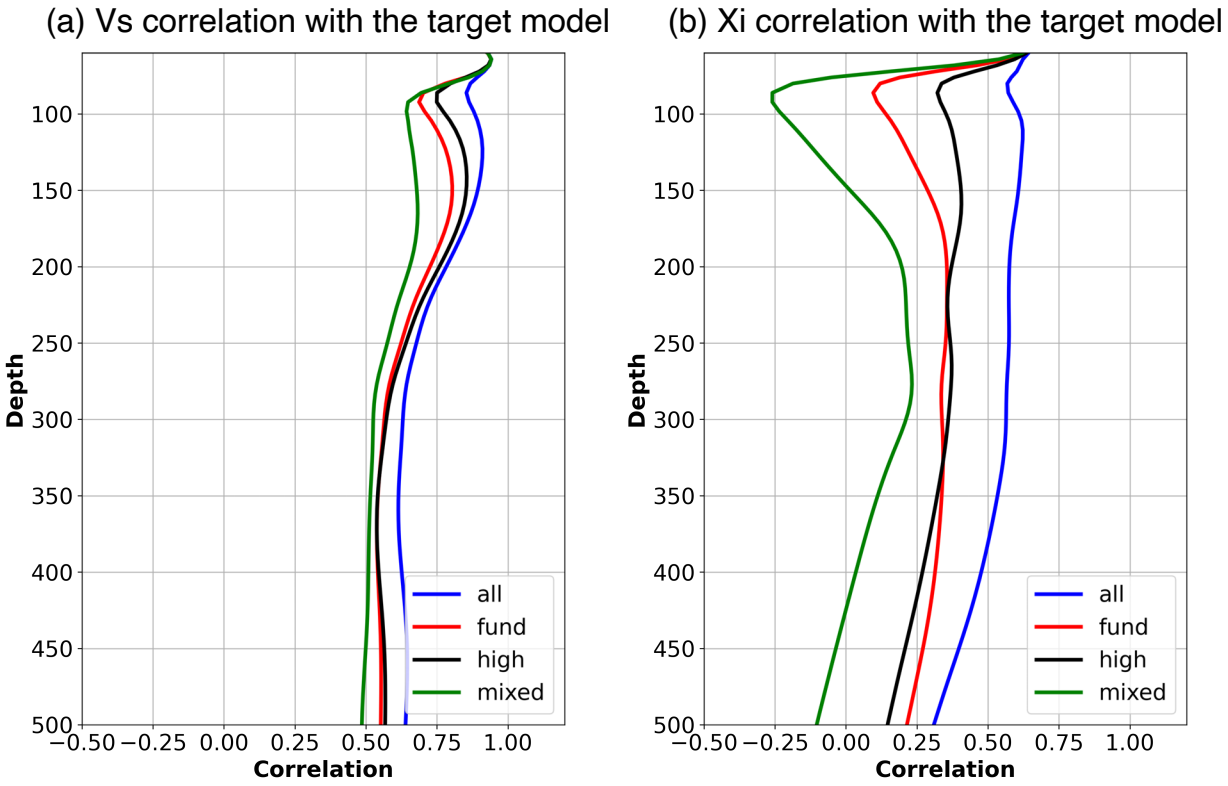

Figure 4.11: Model correlation to the target model (modified SEMUCB-WM1) as a function of depth of the Vs (a) and ξ (b) for the models inverted by using different segments of CCF. All, fund, high, mixed represent the models inverted by using all, fundamental, higher, mixed mode segments of CCFs.

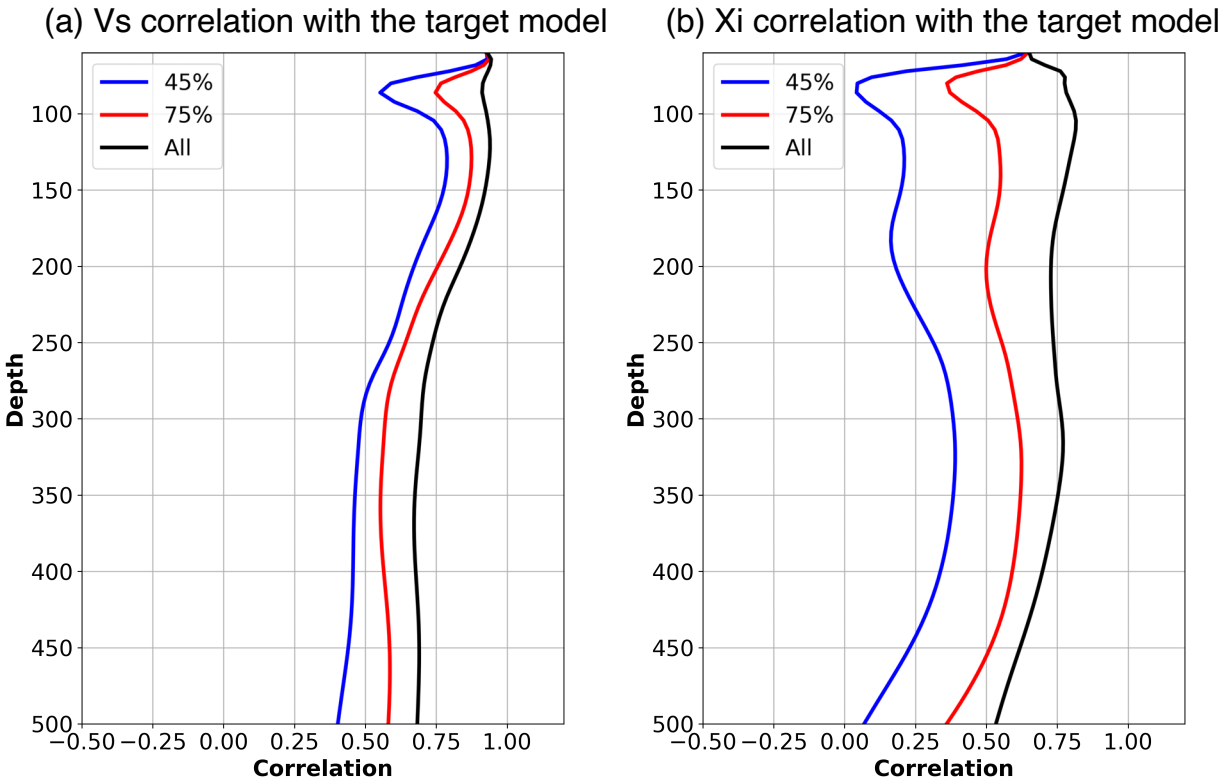

Figure 4.12: Model correlation to the target model (modified SEMUCB-WM1) of the Vs (a) and ξ (b) models for the inversions as a function of depth for the models inverted within 0%, 25% and, 55% of missing data, respectively denoted as all, 75%, and 45% in the figure.

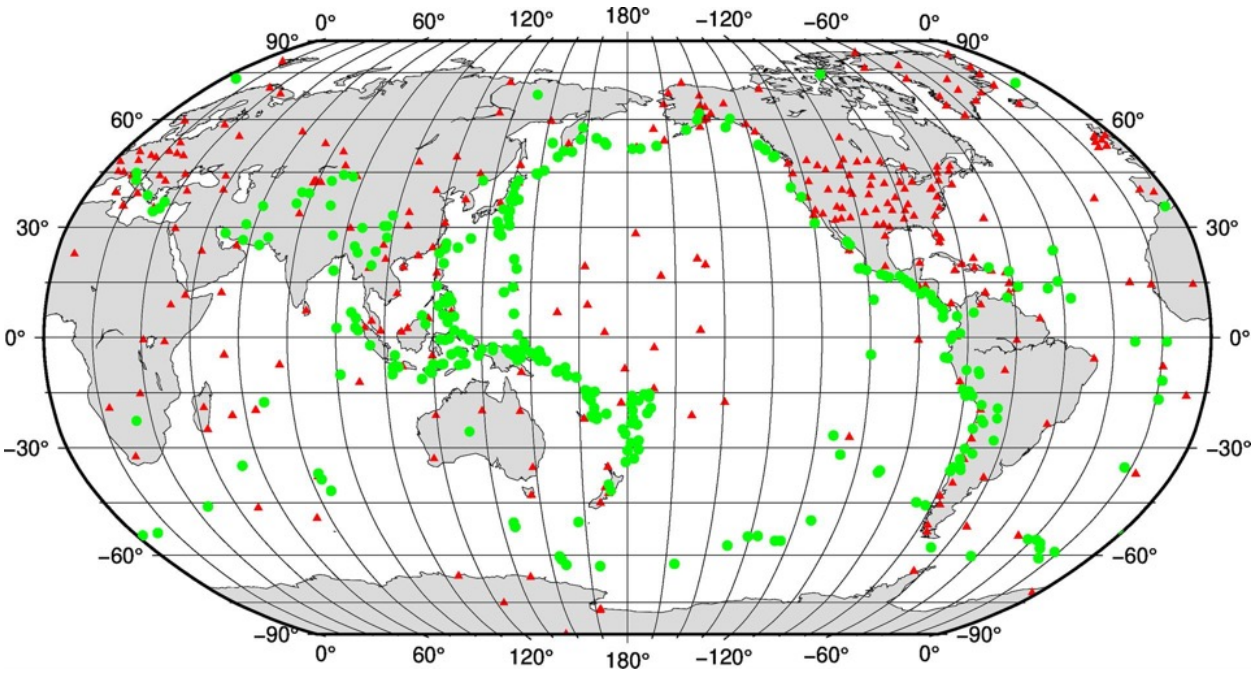

Figure 4.13: The distribution of 355 earthquakes (green) and 340 stations (red) used for constructing the source stacking model. Each station is equipped with a three-component set of very broadband seismometers (i.e. STS-1, KS54000, Trillium 240, CMG-1T, etc). M6-7 events from the period 2012-2017 are chosen.

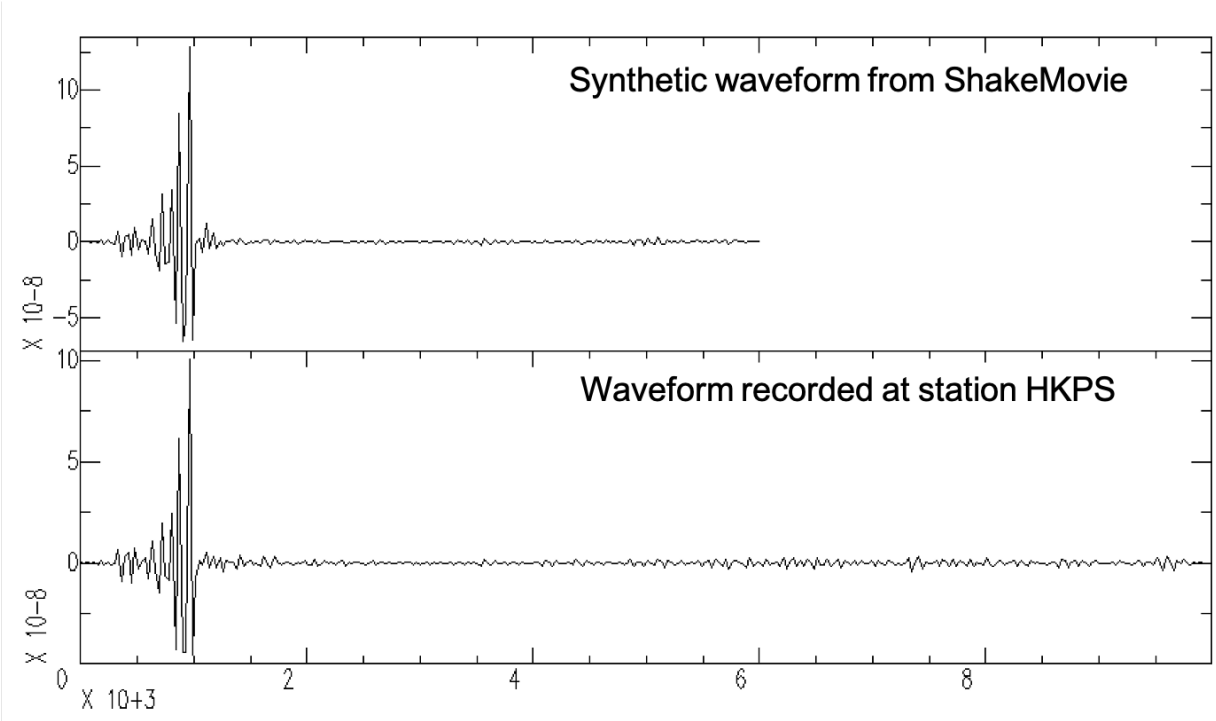

Figure 4.14: Comparison between the ShakeMovie vertical component synthetic waveform trace and the real waveform trace filtered with cut-off frequencies of 60 and 400 s, recorded at the station HKPS on the vertical component.

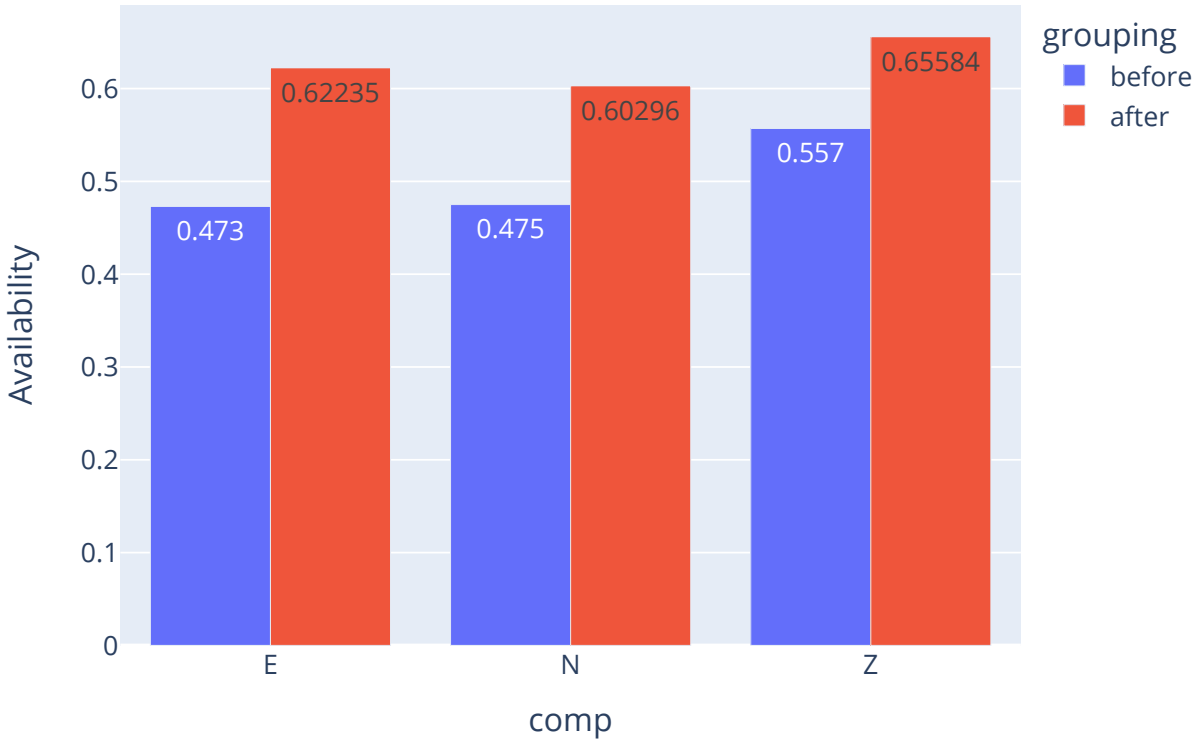

Figure 4.15: Comparison of 3-component data availability before (blue) and after (red) grouping.

Chapter 5

Modeling seismic anisotropy in the lowermost mantle

modified from "Seismic anisotropy, dominant slip systems and phase transitions in the lowermost mantle" (2021, Chandler et al., 2021, GJI). This is a multidisciplinary collaborative research between seismologists (Chen, Romanowicz), mineral physicists (Chandler, Wenk) and geodynamicist (Li). The author (Chen) provided the seismology computation for this study.

5.1 Motivation

Seismic anisotropy in the crust and the upper mantle has been broadly studied in different fields and has been well documented in the past several decades. However, anisotropy in the deepest 200-300 kilometers of the Earth's mantle, called D" (Bullen 1949), is still poorly understood due to lack of sampling by seismic waves and signal contamination from anisotropy in the upper mantle, as well as uncertainties in the deformation properties of lower mantle minerals. In contrast to the isotropic lower mantle, our current understanding indicates the presence of varying significant shear wave anisotropy in D" (Romanowicz and Wenk 2017). Shear wave splitting measurements of phases such as diffracted (Sdiff), reflected (ScS) waves on the core-mantle boundary (CMB) (Nowacki et al. 2011), and also core phases SKS/SKKS (Long 2009; Nowacki et al. 2011) generally show that $V_{SH} > V_{SV}$ is found in regions of faster than average isotropic shear wave velocity (Vs_{iso}) in seismic imaging results (the graveyard of slabs); $V_{SH} < V_{SV}$ or no significant splitting is found in regions of slower than average Vs_{iso} such as the large low shear velocity provinces (LLSVPs) beneath the central Pacific and Africa (e.g., Cottaar and Romanowicz 2013; Lynner and Long 2014).

It is tempting to attribute seismic anisotropy to crystal preferred orientation (CPO) of anisotropic minerals such as highly deformable post-perovskite (pPv) in colder areas of the D" region where large strains accumulated during flow in the mantle (e.g., McNamara et al. 2002, 2003; Wenk et al. 2011). However, there are uncertainties on the dominant slip systems in pPv, where laboratory experiments (e.g., Miyagi et al. 2010; Wu et al. 2017) favor a (001) slip plane but theoretical computations (e.g., Goryaeva et al. 2016) disagree and prefer a (010) slip plane in pPv. The dominant slip on (100) (Merkel et al. 2007; Walker et al. 2011) is excluded because it cannot match well to the observed seismic observations Wenk et al. (2011) and Cottaar et al. (2014).

By assuming a geodynamical scenario of slab subduction and considering the seismic anisotropy in D" is dominated by crystal preferred orientation, the textural evolution and resultant elastic properties can be computed, which will help to explain the seismic anisotropy observations (e.g., McNamara et al. 2002, 2003; Wenk et al. 2011). In previous work, Cottaar et al. (2014) calculated the seismic anisotropy produced in a single mineral system, comparing bridgmanite ($MgSiO_3$ perovskite, Pv) to magnesium post-perovskite (pPv), based on elastic properties and slip systems determined from laboratory experiments and theoretical ab-initio computations. More realistic situations are considered in this study: (a) the polycrystal plasticity model (Lebensohn and Tomé 1993) which consists Pv/pPv, cubic calcium perovskite ($CaSiO_3$, CaPv), and cubic periclase (MgO), (b) the forward and reverse Pv-pPv phase transitions during the slabs' subduction and the subsequent upwelling, (c) partial melting in the deepest portions of the slab at the base of upwelling. To validate and compare the results with seismological observations, the spatial distribution of radial anisotropy described by the parameter $\xi = (V_{SH}/V_{SV})^2$, and shear wave splitting (SWS) directions and strengths are extracted.

5.2 Geodynamic model setting and plastic deformation modeling

This section will briefly introduce the geodynamic model setting and mineral physics modeling. The detailed information can be found in Chandler et al. (2021), section 2.1 and 2.2.

Here a geodynamical scenario of slabs impinging onto the CMB and the subsequent upwelling are considered. The 3D spherical geodynamical model is developed by using a modified version of CitcomCU under the standard Boussinesq approximation (Zhong 2006). All the physical settings in the model are listed in Chandler et al. (2021) Table 1. Passive Lagrangian tracers record the location, non-dimensional temperature and velocity gradient along the slab for the later deformation computations. 25 path-lines of individual tracers, which begin above 1000 km depth and end at various heights above the CMB in the upwelling area, are chosen for sampling multiple areas.

For modeling the plastic deformation, aggregate of a pyrolite composition with 17 % periclase (MgO), 9% $CaSiO_3$ (CaPv), and 74% bridgmanite (Pv/pPv) is considered. Following the assumption from Wenk et al. (2011) and Cottaar et al. (2014), 50% of the accrued strain contributes to plastic deformation by dislocation glide. Then a visco-plastic self-consistent (VPSC) method (Lebensohn and Tomé 1993) by solving the constitutive equations under the Eshelby inclusion formalism (Eshelby 1957) is applied for simulating the evolution of texture and aggregated elastic tensors in geodynamic models. To find the transition locations between forward and reverse phases, the pressure P and temperature T (Chandler et al. (2021) Fig. 5a) , which is obtained from interpolation using PREM and the geodynamic model respectively, are taken into account with assuming a Clapeyron slope for the Pv-pPv phase boundary (Oganov and Ono 2004; Tsuchiya et al. 2004; Hirose et al. 2006). If the transition conditions are satisfied, all Pv transforms into pPv through the martensitic transformation according to the orientation relations described by Dobson et al. (2013) (Fig. 5b in Chandler et al. (2021), the c-axis is maintained but two new variants of a-axes are formed). The crystal preferred orientation (CPO) evolution along path-lines is expressed as a series of 3-D orientation distribution (ODF) which are pole figures displayed in upper hemisphere equal area projection, using the MTEX software package (Bachmann et al. 2010).

Figs 5.1-5.2 show the CPO evolution of 001 and 010 models following the development of strain rate, as the slab moves into different flow regimes denoted by: (a) during downwelling where an average equivalent strain rate of $\sim 1.15 \times 10^{-13} s^{-1}$, both models are identical except for the pPv phase transformation and develop strong (100) texture (∼8 m.r.d). (b) the strain rate increases by $\sim 1.6\times$ as the slab turns along the CMB while entering the Pv–pPv transition (red paths in Figs 5.1-5.2). Model (001) develops (001) texture which is orthogonal to the material flow direction. While Model 010 shares an (010) distribution similar to that of Model 001, but weaker in the strength. (c) the rate of strain dramatically increases to $\sim 8 \times 10^{-13} s^{-1}$ during upwelling. For Model 001, the (001) texture keeps the continues to increase in strength as well as follow with the flow direction, reaching a maximum

of ∼10 m.r.d. Model 010 develops an initial dispersion of CPO just after the reverse phase transition and only reaches a maximum of ∼4–5 m.r.d. On the other hand, the secondary cubic phases MgO and CaPv both develop CPO (Fig. 5.1 right two columns) nearly aligned with the flow direction during downwelling, and maintained throughout the simulation.

5.3 Estimations of elastic properties and seismic anisotropy

The elastic constants for MgO (Karki et al. 2000), CaPv (Kawai and Tsuchiya 2009) and Pv and pPv (Zhang et al. 2016) are applied and averaged over their orientation distributions to obtain overall polycrystal elastic properties and seismic velocities of the 3-phase aggregate. The aggregate elastic tensor is described in a 6×6 matrix (21 independent values, C_{ij}), in Voigt notation. To determine the seismic velocities, the densities are estimated for all points within path-lines based on pressure and temperature. Since the strong radial anisotropy parameter $\xi = V_{SH}^2/V_{SV}^2$ can be found in global tomographic models in the lowermost mantle. ξ is first calculated from each aggregate elastic tensor at all points using the following equation (Browaeys and Chevrot 2004):

$$\xi = \frac{V_{SH}^2}{V_{SV}^2} = \frac{\frac{1}{8}(C_{11} + C_{22}) - \frac{1}{4}(C_{12}) + \frac{1}{2}(C_{66})}{\frac{1}{2}(C_{55} + C_{66})} \tag{5.1}$$

Next, shear wave splitting is calculated for comparison with the local seismological studies of D" anisotropy which rely on splitting measurements of horizontally propagating shear waves diffracted (Sdiff) or reflected (ScS). The Christoffel matrix γ_{kl} (Babuska and Cara 1991) describes the polarized phase velocities in any given direction(x) in an elastic tensor, and is given by:

$$\gamma_{kl} = C_{ijkl} x_i x_j / \rho \tag{5.2}$$

where ρ is the density. The eigenvalues of γ_{kl} give the squared phase velocities: V_p^2, V_{s1}^2 and V_{s2}^2. The eigenvectors are the corresponding polarization directions. The strength of splitting is described as $100 \cdot 2(V_{S1} - V_{S2})/(V_{S1} + V_{S2})$. The implementation of SWS is using the Matlab Seismic Anisotropy Toolkit (MAST, Walker and Wookey 2012).

Since the dominant Pv phase has weakly anisotropic elastic structure, the downwelling part of the slab is mostly isotropic. The forward Pv–pPv transformation occurs at the depth range of 2550-2600km and the pattern of radial anisotropy of $V_{SH} > V_{SV}$ ($\xi > 1$ in Fig 5.3) appears. In both models, ξ is steadily increasing as the slab proceeds along the CMB from a value of 1–3 % . (ξ = 1.01 - 1.03) to a maximum of ∼6 % in the deepest parts of the slab. As the slab approaches the base of the upwelling, the increasing temperature initiates the reverse pPv–Pv transformation at the depths (2595–2835 km). Model 001 presents the patterns of $V_{SH} < V_{SV}$ at the base of the upwelling, which is consistent with available

seismic observations (e.g., Long 2009; Romanowicz and Wenk 2017). However, the pattern of $V_{SH} < V_{SV}$ is not obvious in Model 010 (dashed ellipses in Fig. 5.3).

The resultant shear wave splitting results are shown in Fig 5.4. Figs 5.4(a) and (b) show the regions highlighted in red for presence of pPv, where horizontal fast axis orientation continues throughout the region of horizontal flow along the CMB in agreement with seismic observations (Panning and Romanowicz 2006; Sturgeon et al. 2019). However, significant differences occur between the two pPv models: for Model 010, the fast axis direction with relatively larger SWS strength (∼1.0–3.0 %) remains horizontal after the reverse transition and the trend continues during the upwelling; Model 001 shows a complex pattern of mixed horizontal and vertical fast axis directions after the reverse pPv to Pv transition with ∼2.0 % of SWS strength, and a tilted fast axis with 1.0–2.0 % SWS strength in the upwelling segment (Fig. 5.4) also in agreement with Cottaar and Romanowicz (2013).

5.4 Effect of partial melting

Several studies (Williams and Garnero 1996; Simmons and Grand 2002; Yuan and Romanowicz 2017) have suggested that partial melting in the lowermost mantle could explain ultra-low velocity zones (ULVZs). Here a first order approximation on the effect of the addition of melting in D" is investigated. For considering of the locations of partial melting, the melting conditions of a pyrolytic sample across the P-T conditions in lower mantle (Nomura et al. 2014) is taken into account. Our result shows that the partial melting may occur at the base of upwelling near the edges with the reverse pPv–Pv phase transition (Chandler et al. (2021), Fig. 10b). Only 1 % melting is needed to observe ∼2.5 % decrease in P-wave velocity and ∼4 % decrease in S-wave velocity compared to the same path-lines without partial melting added. Increasing the melt percent to 15 % leads to reductions in V_s and V_p of ∼18 and 7 %, respectively. The direction of fast polarization in either model does not change, but a strong increase in the strength of anisotropy (Fig. 5.5). This is consistent with seismic observations using SK(K)S waves near the edges of LLSVPs at D" (Lynner and Long 2014).

5.5 Summary

By combining not only constraints from seismology but also from mineral physics and geodynamics, this study models the effect on seismic anisotropy of forward and reverse Pv/pPv phase transitions, in a scenario of subducting slab being deformed along the CMB and then upwelling. Our result shows that 001 model has a strong but complicated anisotropy, with $V_{SH} > V_{SV}$ in regions of faster than average isotropic V_s (e.g. graveyards of slabs), a tilted fast axis of SWS near the borders of the LLSVPs, and flow-aligned as observed in hotter than average regions of upwelling. All the results of the 001 model agree with the seismic observations (Panning and Romanowicz 2006; Cottaar and Romanowicz 2013; Lynner and Long 2014; Sturgeon et al. 2019).

5.6 Figures

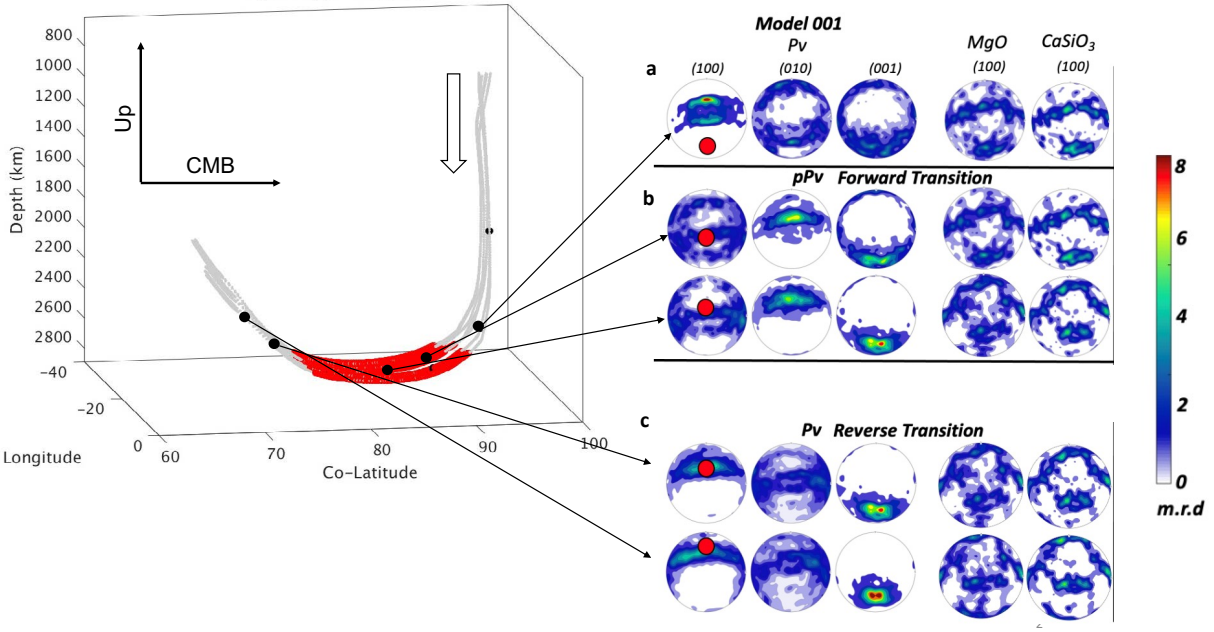

Figure 5.1: Deformation of Pv/pPv + MgO + $CaSiO_3$ for Model 001 before and after the forward and reverse phase transitions indicating at the locations of the streamlines (light grey in left panel). Red paths represent the presence of pPv. The arrow indicates the direction of slab subduction.

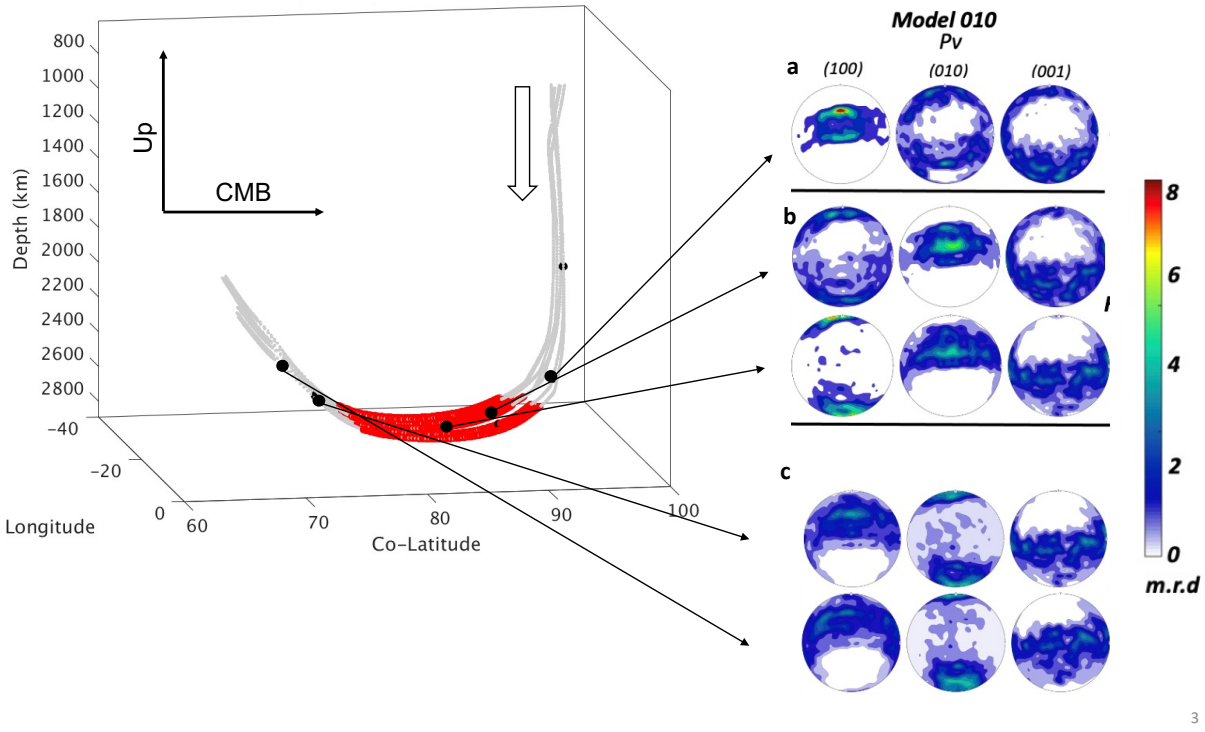

Figure 5.2: Deformation of Pv/pPv for Model 010 before and after the forward and reverse phase transitions indicating at the locations of the streamlines (light grey in left panel). Red paths represent the presence of pPv. The arrow indicates the direction of slab subduction.

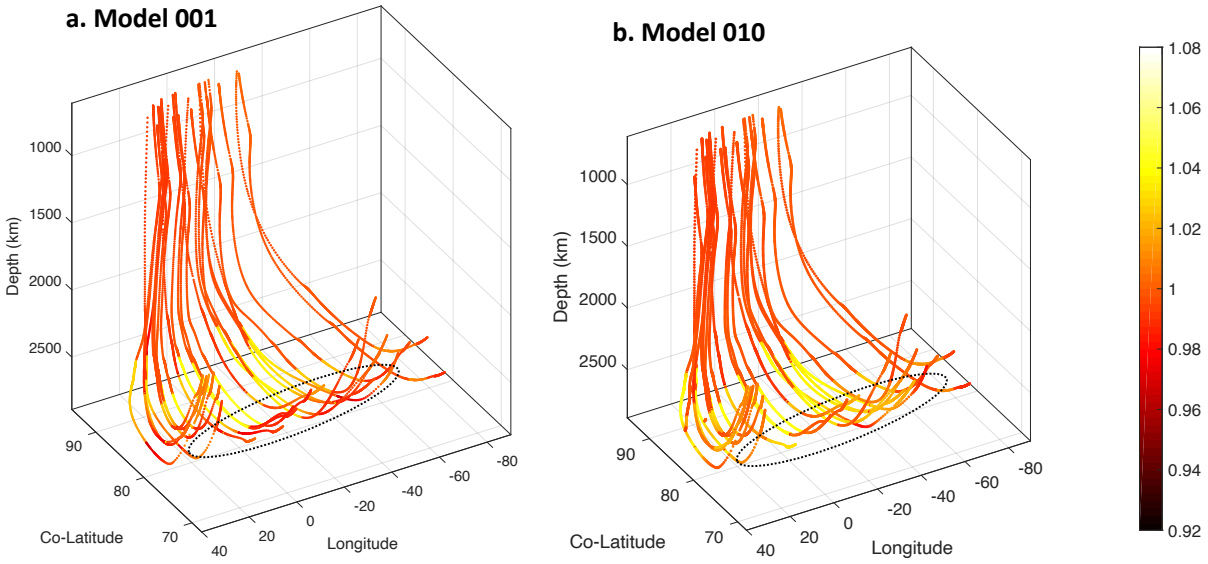

Figure 5.3: Spatial distribution of radial anisotropy ($\xi = \frac{V_{SH}^2}{V_{SV}^2}$) computed for (a) Model 001 and (b) Model 010 along 25 streamlines. Color scale represents the strength of ξ. Dashed ellipses indicate the different ξ patterns of two models along the slab upwelling.

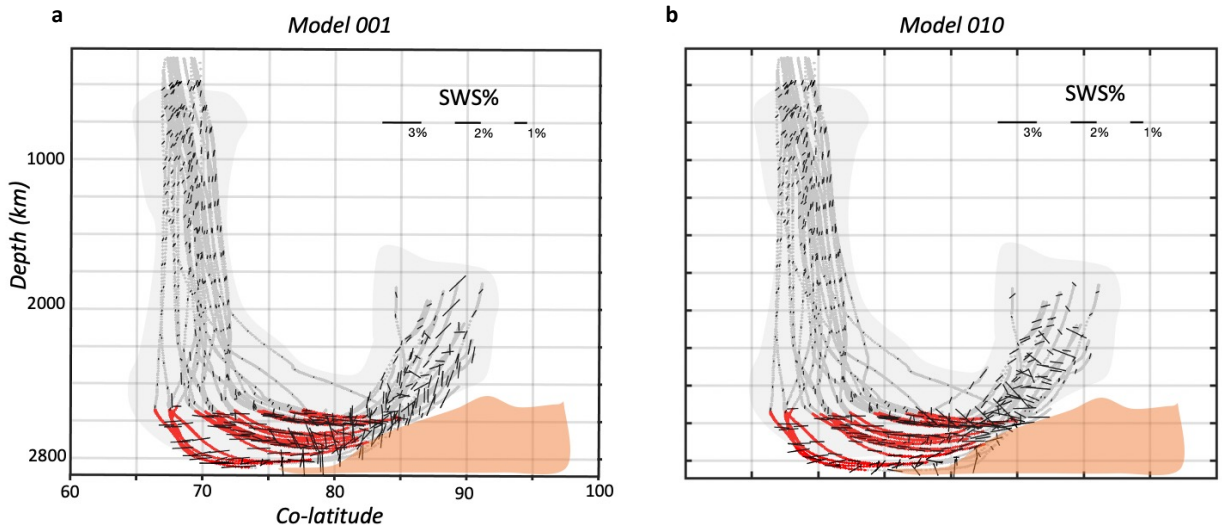

Figure 5.4: Shear wave splitting strength and fast axis direction for waves propagating in lowermost mantle for (a) Model 001 and (b) Model 010. Red paths highlight areas where the pPv phase is present. Orange areas indicate areas of increased temperature at the base of the upwelling. Black ticks indicate the fast polarization direction scaled by %. Incident seismic wave direction is into the page.

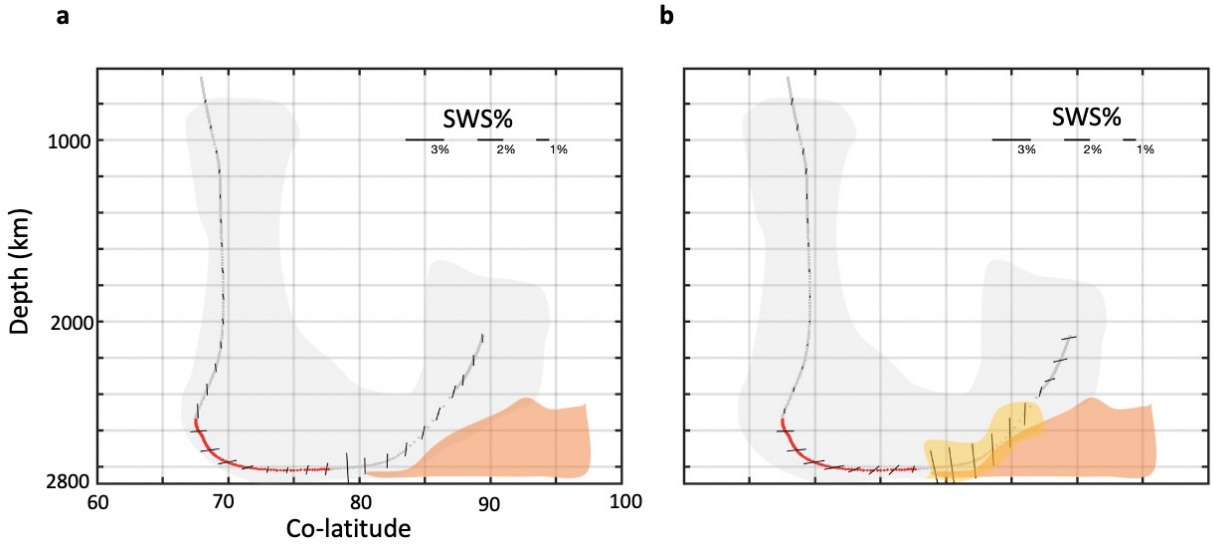

Figure 5.5: Shear wave splitting strength and fast axis direction for waves propagating in lowermost mantle (a), without partial melting and (b), with 1% partial melt for Model (001). Black bars indicate the fast axis direction scaled by %. Red paths represent the presence of pPv. The area of the slab where partial melt is present is shown in yellow, while orange represents increased temperature at the base of the upwelling. All incident wave propagation directions point into the page.

Chapter 6

Temporal variation of near-surface seismic velocity and anisotropy in Hokkaido

---

This chapter presents the results of collaborative research between Berkeley Seismology Lab (Li-Wei Chen, Taka'aki Taira) and Earthquake Research Institute in the University of Tokyo (Hiroe Miyake).

6.1 Motivation

Temporal variations in the elastic properties of crustal structure provide a unique constraint on its behaviors, such as sudden changes due to static and dynamic stress imparted by local and distant earthquakes, seasonality from the responses to hydrological phenomena, healing process of fractures, which helps to better understand the underlying mechanisms. Noise-based seismic interferometry has become a very well-developed and efficient seismological took for time-lapse monitoring the temporal variations of elastic properties, but mostly for shear velocity structure. Here an earthquake-based method (Chen et al. 2017) is applied through interferometry of earthquake coda waves by using borehole-surface station pairs. The technique provides comprehensive observations not only of S-wave velocity (Vs) but, additionally, of P-wave (Vp) and S-wave azimuthal anisotropy between two sensors. The temporal variations of elastic properties is monitored at the KiK-net station IBUH03 in southern Hokkaido for more than 15 years, which experienced the 2003 Mw 8.3 Tokachi-Oki (PGA= $\sim 350 cm/s^2$) and the 2018 Mw 6.7 Hokkaido Eastern Iburi (PGA= $\sim 500 cm/s^2$) earthquakes. Note that the station is located only 8km away from the epicenter of the 2018 event (Fig. 6.1), which provides an unique opportunity to observe the near-surface changes of elastic properties during earthquake and also the healing process after. In what follows, the prominent temporal variations over the post-seismic period and seasonality in the observations are discussed.

6.2 Data and method

Taking advantage of the strong-motion seismograph network KiK-net which comprises pairs of triggering mode seismometers installed together in a borehole, all available waveform data ($>$ 800 events) at IBUH03 (the inter-station distance is 150 m) in a frequency band of 1-15 Hz from 2003 to 2018 is collected. Next, seismic interferometry for earthquake coda waves is performed by following the procedure of Chen et al. (2017) for constructing the cross-correlation functions (CCFs) (Fig. 6.2). V_p was extracted by computing the CCF of vertical component (C_{ZZ}). To evaluate the V_S azimuthal anisotropy, four combinations of CCFs between station-pairs: $C_{NN}, C_{NE}, C_{EE}, C_{EN}$ are computed. The CCF in any directions is given by

$$C_{\theta\theta} = C_{NN}cos\theta cos\theta + C_{EN}sin\theta cos\theta + C_{NE}sin\theta cos\theta + C_{EE}sin\theta sin\theta \tag{6.1}$$

where θ is the azimuthal angle measured clockwise from the north. By following relationship $V_s(\theta) = V_{iso} + V_{ani}cos2(\theta - \phi)$ (e.g. Alford 1986), the elastic properties: the isotropic V_s (V_{iso}), anisotropy coefficient (V_{ani}), and the fast polarization direction (ϕ) are inverted based on least-squared fitting. Vs_{fast} and Vs_{slow} are evaluated as V_s in the fast and slow polarization directions. Then the strength of anisotropy can be defined as $(Vs_{fast} - Vs_{slow})/Vs_{fast}$.

6.2.1 Stability of stacked CCFs

Fig. 6.3 shows the resultant CCFs stacked from different numbers of single CCF for 3 components. The results show that at least 6-12 single CCFs are needed to be stacked to have a stable CCF. On the other hand, unlike noise-based seismic interferometry, the earthquake-based method depends on the occurrence of earthquakes and cannot evenly sample the temporal variation of the elastic parameters. To better observe them in different situations (i.e. post-seismic change or seasonality), two stacking methods to stabilize the CCFs are proposed.

1. Event bin stack: a certain number of CCFs are stacked from the aftershocks of large events which are recorded at a given station and choose the occurrence time of the middle event as representative time. The same weighting is applied to every single CCF while stacking. This method is applied after a large event due to abundant aftershocks which provide dense monitoring during post-seismic periods.

2. Gaussian-weighted stack: by defining a certain length of time window (L) where the central event is in the middle (T_0), the Gaussian-weighted CCFs are stacked from the events which have occurred in this time window. The weighting scheme is defined by a Gaussian distribution where L represents the whole length of the distribution. (i.e. the largest weighting for the central CCF is 1.0. the weighting for CCF occurrence time of $T_0 + 1\sigma$ and $T_0 - 1\sigma$ is about 0.63.) For the interest of monitoring seasonal variations, this method is applied due to uneven sampling by earthquakes in time and lack of earthquakes.

Since CCFs before and after the large earthquake could be very different due to the co-seismic velocity reduction, the CCFs are prevented from stacking across large events for two methods.

6.3 Temporal variation of seismic velocity and V_S azimuthal anisotropy

Fig. 6.4 presents the temporal variations of V_{Siso} (a), V_P (b), V_P/V_S (c), strength of anisotropy (d) and fast direction (e), after applying a 180-day Gaussian-weighted stack to the CCFs at IBUH03. The site with low seismic velocities and high V_P/V_S ratio is located in the lowlands with unconsolidated fluid-saturated sediments, according to the well log data report. Also, the observed anisotropy here can be categorized as SAA (Stress-Aligned Anisotropy) which is related to the opening of fluid-filled fractures or cracks subject to lateral compression or extension (e.g. Crampin 1987, 1994; Chen et al. 2017).

After the 2003 Mw 8.3 Tokachi-Oki (the 2003 event) and the 2018 Mw 6.7 Hokkaido Eastern Iburi earthquakes (the 2018 event), there are sudden reductions of V_{Siso} of 5-7%

and 8-10% respectively, sudden increase in V_P/V_S ratios, and perturbations in the fast directions of anisotropy (Fig. 6.4). No clear change in P-wave velocity or anisotropy strength is observed. However, the temporal variation resolution can be improved since a relatively long time window (i.e. 180-day) is used here for implementing the Gaussian-weighted stack, which could smooth out short-term variations. The results of event bin stacking will be shown in next section. On the other hand, a decade-long V_{Siso} recovery trend with seasonal variations can be found during the period between the 2003 and 2018 events. To evaluate the V_{Siso} changes, the fitting scheme (black curve in Fig. 6.4a) is implemented by considering seasonal velocity variations, co-seismic velocity reduction, and post-seismic velocity recovery (Taira et al. 2018).

$$f(x) = A + Bt + [Cexp(-(T - T_{eq})/D)]H(T - T_{eq}) + Esin(\omega t) \\ +Fcos(\omega t) + Jsin(2\omega t) + Kcos(2\omega t) \tag{6.2}$$

where A is the constant offset, B is the linear trend of change, C is the co-seismic change during a large earthquake (i.e. 2003 or 2018 event) and D is the characteristic recovery time. E, F, J, and K are annual and semiannual variations, and ω in the sine and cosine terms were fixed to be $2\pi yr^{-1}$. $H(T)$ is the Heaviside step function and T_{eq} is the occurrence time of an large earthquake; T is an elapsed time(yr).

Taking advantage of the abundant aftershocks after the two events, a 5-event bin stack is applied to improve the temporal resolution, while a 21-event bin stack before the events are used. The left panel in Fig. 6.5 shows the 4-year (1 year before and 3 years after the event) temporal variations around the time of the 2003 event; the right panel presents the monitoring for the 2018 event for 3+ years (3 years before and 2 months after). Note that there are only 2-months of observations after the 2018 event. Notable drops in V_{Siso} (5-7% for the 2003 event and 8-10% for the 2018 event) and V_P (2-2.5% for the 2003 event and 1.5-2% for the 2018 event) can be found after the events. V_{Siso} has a prominent reduction and only $\sim$50% recovery of V_{Siso} in the first two months after the events. But V_P fully recovers within only a few days after the events. A sudden increase of V_P/V_S ratio can be seen in the first two months after both 2003 and 2018 events, and gradually decreasing after a slow year-long recovery trend. The strength of anisotropy rapidly increases within days after the events but quickly returns to background values (i.e. back to the seasonal fluctuations for the 2003 event) after that. The fast direction changes vary for the two events. For the 2003 event, the fast direction first rotates by 30° clockwise and then rotates back by -60°; for the 2018 event, the fast direction shows a direct rotation of -40°.

6.4 Discussion

In this study, an earthquake-based interferometry is used to monitor temporal changes of V_P, V_{Siso}, and V_S azimuthal anisotropy in the top 150 m below the surface associated with the occurrence of the two large events. Trends of Co-seismic velocity reduction, post-seismic

recovery and a long-term decade-long velocity recovery with seasonal velocity variations can be found. The detail will be discussed in the following sections.

6.4.1 Near-surface changes of elastic property after large events

After the two large events, large velocity reductions followed by short-term (in 10-15 days) rapid recovery and subsequent long-term (in 3-5 years) slow recovery can be detected. Since this observation is likely categorized as the SAA (Stress-Aligned Anisotropy) type, the opening and closure of cracks can be sensitive to external stress changes and are often considered as a mechanism of seismic velocity changes associated with earthquakes. The V_{Siso} and V_P velocity reductions could be due to the dynamic stress changes associated with the large event which lead to temporary opening or connection of fluid-saturated fractures or cracks. Also, these observations of increasing strength of anisotropy, V_P/V_S ratios and the rotations of fast directions could also support this stress-induced velocity change interpretation. For the 2003 and 2018 events, the changes in V_{Siso} (5-7% and 8-10%, respectiely) are larger than in V_P (2-2.5% and 1.5-2%) which further suggests that fluids are involved in the process (e.g. Schaff and Beroza 2004). During the co-seismic period, the short-term rapid recovery of V_{Siso}, V_P, V_P/V_S ratio and anisotropy can be explained as due to closure of the temporarily opened cracks. Note that only $\sim$50% recovery of V_{Siso} occurs in this first 15 days. As for the subsequent long-term exponential recovery of V_{Siso} and V_P/V_S ratio in the post-seismic period, poro-elastic contraction caused by post-seismic deformation could be the cause. The long-term recovery process can last for decades. However, other possible explanations (i.e. the change of non-linear soil response, hydrological perturbations, etc..) are currently under investigation.

6.4.2 Influence of equivalent water thickness, precipitation, and temperature

Seasonal V_{Siso} variations with annual and semiannual trends can be well described by fitting the equation 6.2. The amplitude of the signal is relatively small but still observable (Fig. 6.4a). There are underlying mechanisms of the seasonal velocity variations, i.e. fluid-related changes due to precipitation or groundwater level changes (e.g. Sens-Schönfelder and Wegler 2006; Gassenmeier et al. 2014), thermo-elastic strain variations (e.g. Meier et al. 2010), or thermally induced stress (e.g. Gassenmeier et al. 2016). So the corresponding gravity (water equivalent thickness) and meteorological data (precipitation and temperature) are examined. Due to the absence of groundwater observations in the vicinity of the seismic station, gravity data obtained by the GRACE satellite mission (e.g. Swenson and Wahr 2002; Tapley et al. 2004; Wahr et al. 2004) can be used to infer changes in mass balance due to hydrological variations. Water equivalent thickness (provided by the Center for Space Research of the University of Texas, Austin) is then taken into account but the changes could be caused by any water-related mass changes (i.e. groundwater, soil moisture, surface water, snow, and ice). Another type of meteorological data such as precipitation and temperature

data recorded by a station 8-km away from IBUH03 are also collected for comparison, to understand their relation to temporal variations in seismic parameters. Fig. 6.6b shows the short-term variation of V_{Siso} is negative correlated with water equivalent thickness; On the other hand, the long-term variation of V_{Siso} has a decade-long correlation trend with water equivalent thickness. The temperature shows a positive correlation trend (Fig. 6.6a) and there is no clear seasonality for precipitation (Fig. 6.6c).

To observe the seasonal variations by month, monthly stacking the CCFs is applied through a 3-month moving window during 2008-2018 (i.e. February CCF are stacked by taking the CCFs from January to March over the 10-year period). The same averaging process is also applied to water equivalent thickness, precipitation and temperature. Fig. 6.7 shows the result of monthly stacking results comparing the monthly average gravity and meteorological data. Maximum peaks of V_P, V_P/V_S ratio, and strength of anisotropy can be found in the winter-spring seasons, and correlate with the main peak of water equivalent thickness. In the summer season, a maximum peak of V_{Siso} is found as well as secondary maximum peaks of V_P, V_{Siso}, V_P/V_S ratio, and strength of anisotropy, all of which can be correlated with corresponding peaks of precipitation and temperature. V_P/V_S ratio, and strength of anisotropy increase in the winter-spring time. This could be explained by the rise of water-related mass, which could increase pore pressure and fluid content which, in turn would increase the fluid filled crack density. In the summer time, local precipitation, which could also be related to temperature or thermal effects, would dominates the variations but with a weaker strength (Fig. 6.7c, the secondary maximum peak in V_P/V_S ratio).

6.4.3 Modeling crack behavior by Hudson-Crampin anisotropic effective model

A Hudson-Crampin anisotropic effective model (e.g. Crampin 1984) is implemented to investigate the seasonal poro-elastic effect on fluid filled cracks. In order to find the best-fitting aspect ratio of cracks and crack density (defined as the number of cracks per unit volume) for the monthly observations, a grid-search approach is applied with a dimensionless L1 misfit function to constrain the V_P/V_S ratio and anisotropy strength (Fig. 6.8 shows the misfit distribution). The medium of the effective model has, on average, the V_P and V_{Siso} of station IBUH03 and inclusions only contain water. The schematics of an anisotropic effective model can be found in Fig. 6.9. In this modeling, the fixed elastic properties used are listed below:

1. medium: $V_P = 1030.29\ m/sec, V_S = 217.67\ m/sec, \rho = 2200\ kg/m^3$
2. penny-shape inclusion (crack): $V_P = 1475.00\ m/sec, V_S = 0.00\ m/sec, \rho = 1000\ kg/m^3$

Although the results are computed by considering an over-simplified effective model, the behaviors of crack density and the aspect ratio can provide insights for understanding the underlying mechanisms. Fig. 6.10 shows the decreasing aspect ratio and increasing crack density in the winter-spring time which indicates temporary opening of prolonged cracks

and increases in crack density. This can be related to the increasing pore pressure and fluid content due to the increase of water-related mass. In the summer time, the minor effect of decreasing aspect ratio and increasing crack density can be related to the rise of local precipitation which causes a slight increase of pore pressure.

6.4.4 The capability for detecting subtle variations

To ensure the capability of detecting subtle temporal variations by this method, a synthetic test is applied for the scenarios of insufficient sampling rate (100 sps) in seismic data with different noise levels. A set of wavelets considering the travel times with a low-level of strength of anisotropy is computed with a sufficient sampling rate (10,000 sps).The parameters for computation are listed below:

1. V_P= 1027.40 m/sec
2. V_{Siso}= 217.85 m/sec
3. Strength of anisotropy = 0.57 %
4. Fast direction = 34.37 °
5. downsample to 100 sps and add different levels (in Fig. 6.11 showing the cases with noise free, 0.5σ, 1σ, 2σ of the RMS amplitude) of white noise

Table 6.1 shows that the difference between the input parameters and measurements is very small even for the case with 2σ noise added. The experiments further confirm the confidence in the ability to detect tiny changes even with a relatively low sampling rate, and an anisotropy strength as small as only 0.57%.

6.5 Summary

Observations of temporal variations of near-surface seismic velocity and V_S azimuthal anisotropy are presented in this study over a 15-year period in Southern Hokkaido. The seismic waveforms (∼800 events) are analyzed at station IBUH03 in a frequency band of 1-15 Hz in the KiK-net station IBUH03, following the method of Chen et al. (2017). To ensure the reliability of observations, the temporal robustness of the reconstructed cross-correlation functions (CCFs) in each time segment is evaluated. The preliminary results show sudden reductions of V_{Siso} and V_P after the two major earthquakes considered, and these were subsequently recovered with short-term and long-term recovery rates. Changes in the fast S-wave polarization direction and strength of anisotropy were also detected during the coseismic periods. The results indicate that dynamic stress changes caused by the large events enhance the opening of fluid filled cracks. This results in reductions of seismic velocities and changes in V_S azimuthal anisotropy. Furthermore, a clear seasonal trend of elastic parameters can

be observable, which implies that the changes in elastic properties might also be affected by seasonal hydrological responses. Comparing the results with temporal variations in equivalent water thickness, precipitation, and temperature, a Hudson-Crampin anisotropic effective model is applied to understand the seasonal crack behavior. The modeling result indicates that crack density and the aspect ratio could play an important role in this observations: the winter-spring time variations could be due to the increase of water-related mass; the variations in summer are less strong and dominated by the local precipitation.

6.6 Table

	Noise Free	0.5X std	1.0X std	2.0X std
Vp (m/sec)	1.438	4.212	12.329	26.712
Vs (m/sec)	0.305	0.349	0.196	0.218
Vp/Vs ratio	0.0001	0.0118	0.05659	0.11790
Aniso Strength(%)	0.0007	0.0200	0.0798	0.1397
Fast Direction (°)	0.04	2.03	0.52	0.95

Table 6.1: Differences between the input parameters and the measurements of different noise levels.

6.7 Figures

Figure 6.1: The geographical map of the study area. IBUH03 station is shown as a red triangle. The 2018 Mw 6.7 Hokkaido Eastern Iburi earthquake was just several kilometers away. Figure courtesy of Google Earth.

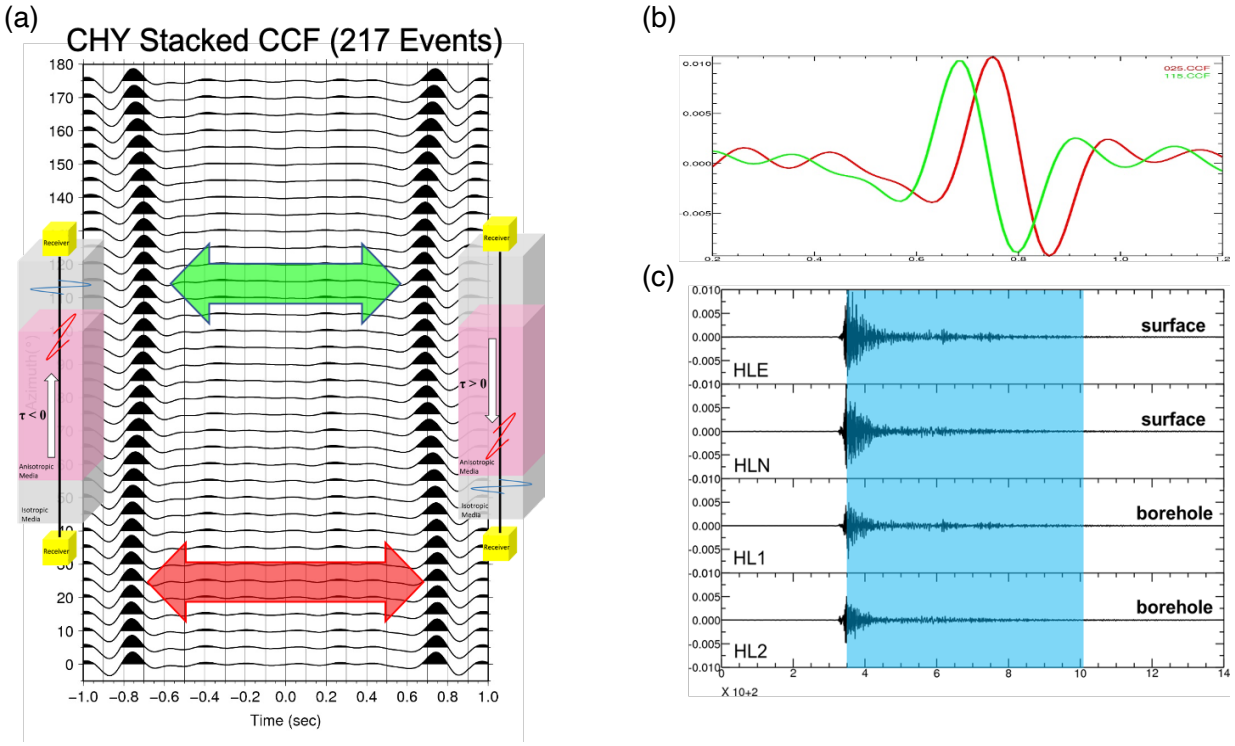

Figure 6.2: An example of earthquake-based seismic interferometry by using coda waves. The panel (a) shows the CCFs sorted by the azimuthal angles and the schematic sketch denotes the direction of traveling waves. The panel (b) shows the comparison of CCF between the fast and slow directions.) The panel (c) shows the earthquake coda waves used for cross-correlating.

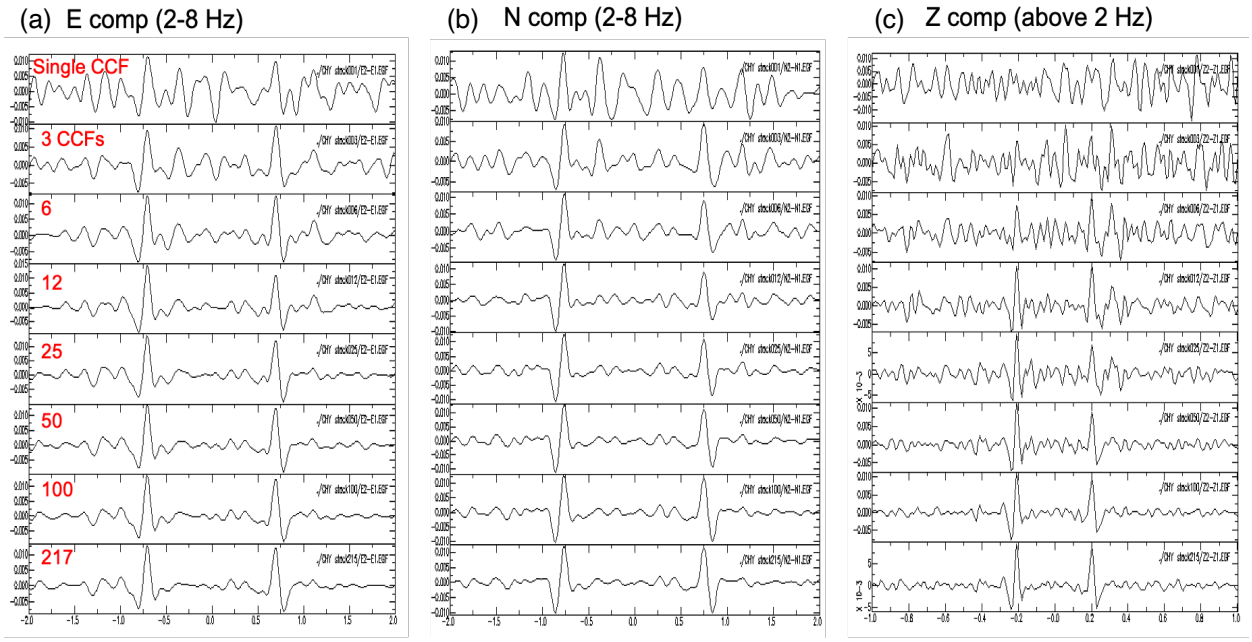

Figure 6.3: Different numbers (in red color) of single CCFs are stacked in 3 individual components for evaluating the robustness of the stacked cross-correlation functions (CCFs).

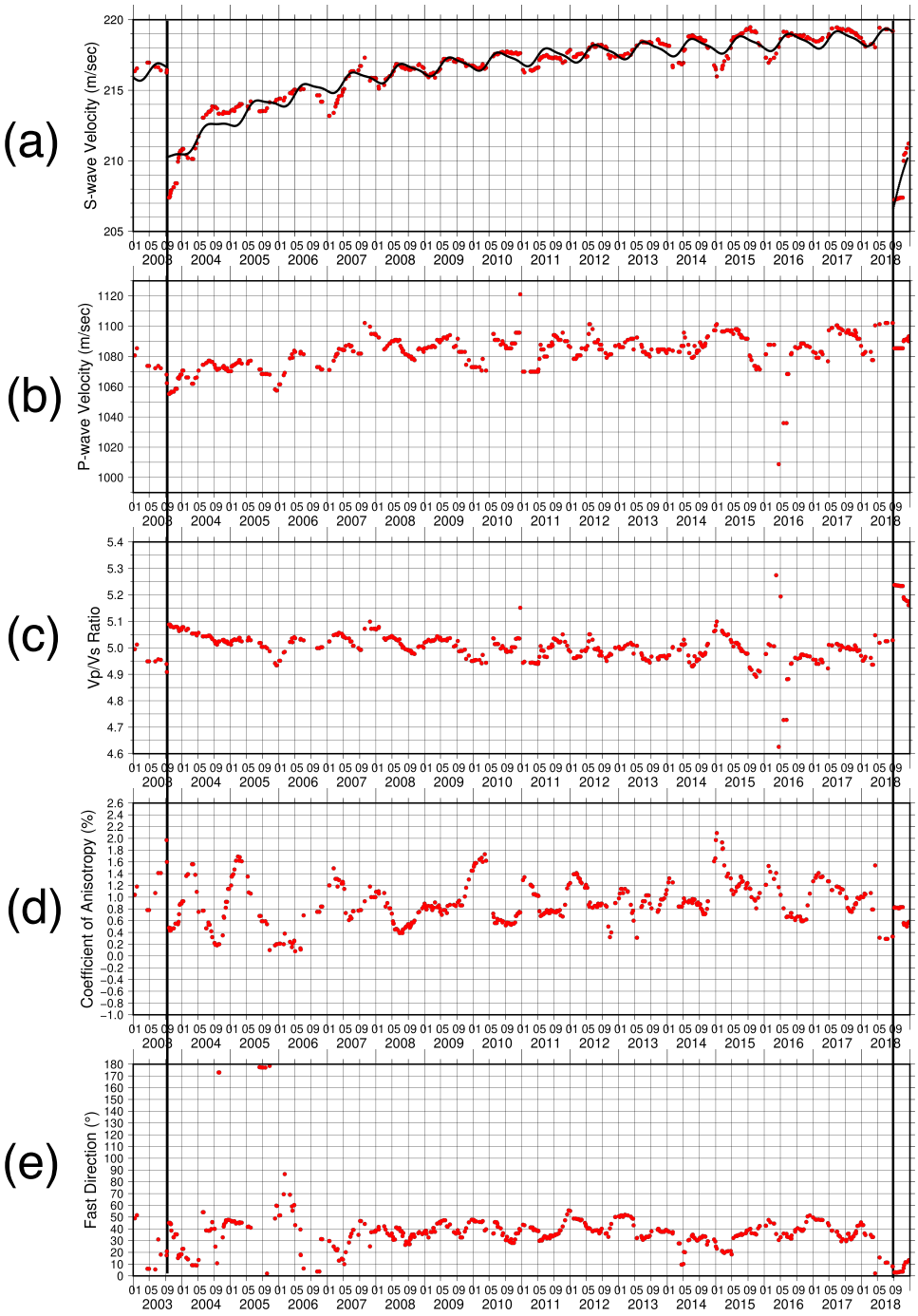

Figure 6.4: Temporal variation of V_P (a), V_{Siso} (b), V_P/V_S ratio (c) and V_S azimuthal anisotropy (d)(e) from 2003 to 2019 at IBUH03. The black curve represents the fitting result of the long-term V_{Siso} changes described in equation 6.2.

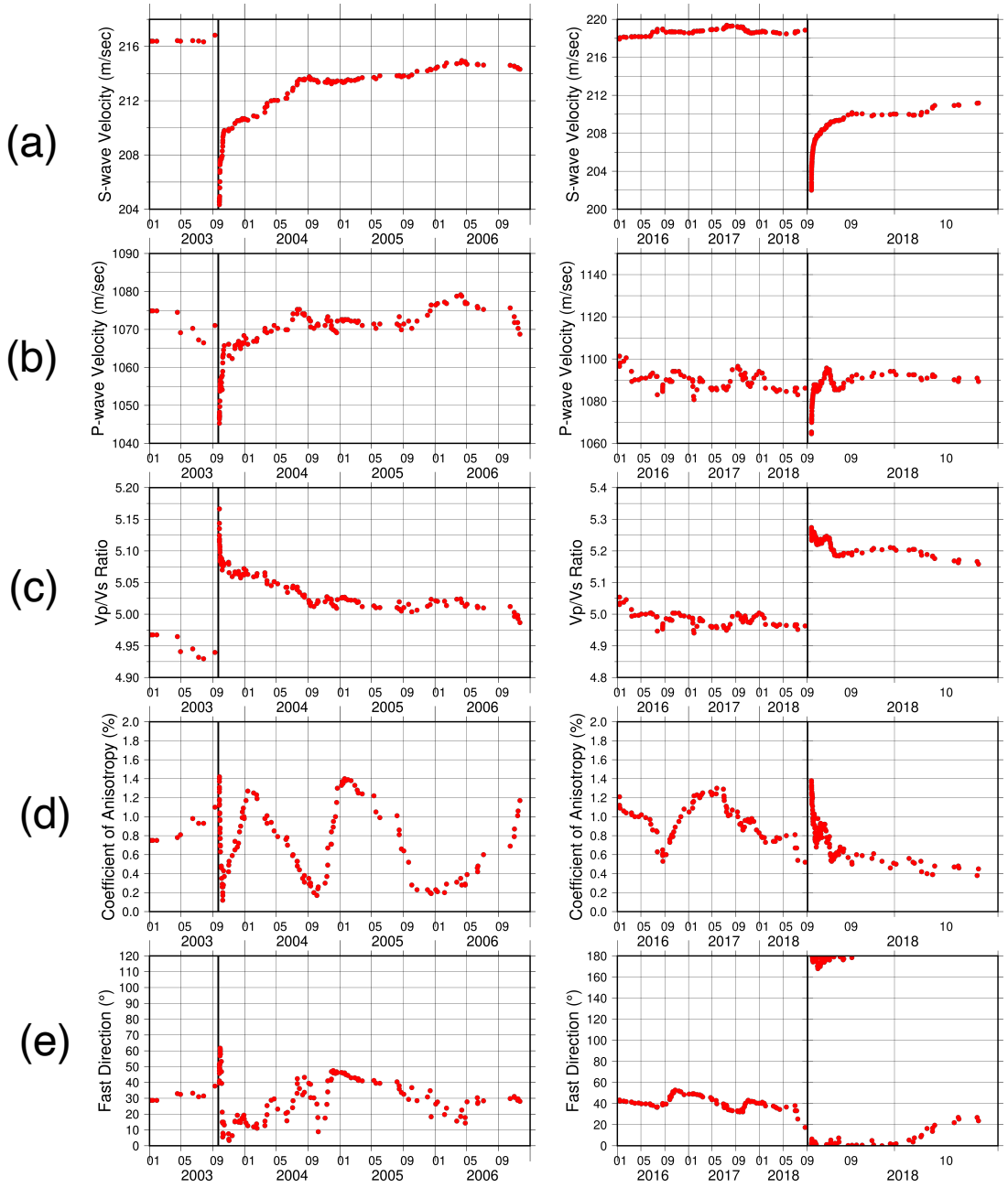

Figure 6.5: Temporal variations of V_P (a), V_{Siso} (b), V_P/V_S ratio (c) and V_S azimuthal anisotropy (d)(e) at IBUH03 after the 2003 and 2018 events. Solid lines represent the occurrence time of earthquakes. Left panel: 3 years after the 2003 Mw 8.3 Tokachi-Oki earthquake. Right panel: 2 months after the 2018 Mw 6.7 Hokkaido Eastern Iburi earthquake.

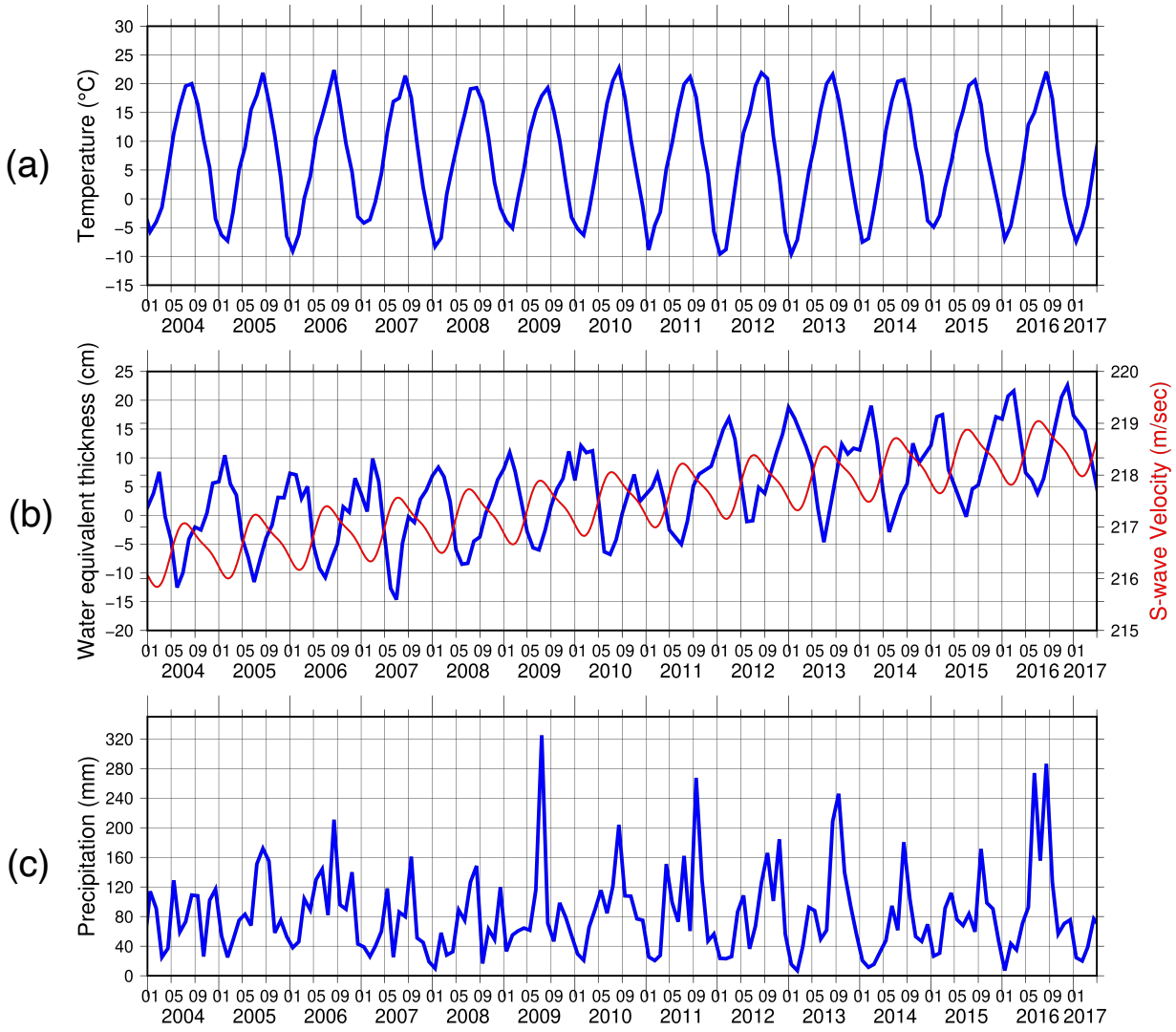

Figure 6.6: Temporal variation of equivalent water thickness (a), precipitation (b), and temperature (c). Comparison with the fitting result of S-wave (red curve) without considering coseismic velocity reduction and postseismic recovery.

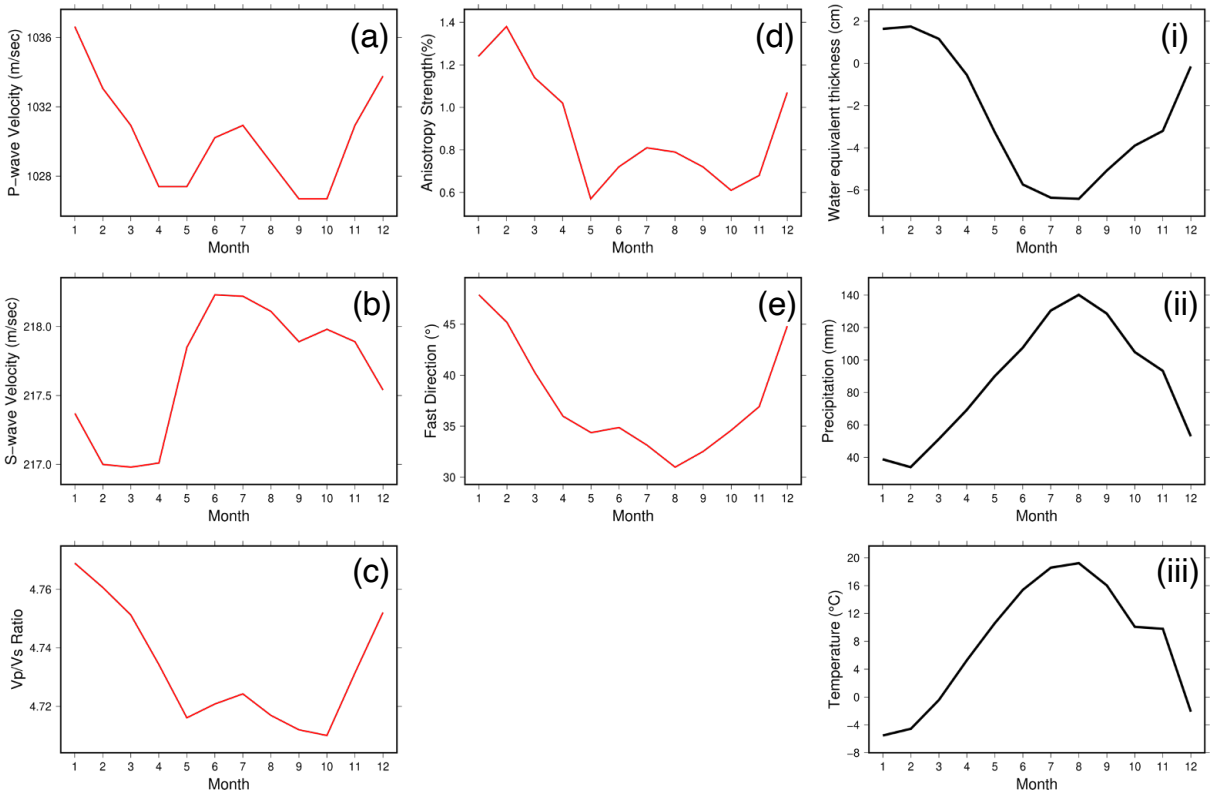

Figure 6.7: Monthly stacking results of V_P (a), V_{Siso} (b), V_P/V_S ratio (c) and V_S azimuthal anisotropy (d)(e) during the period 2008-2018. Monthly averaging results of water equivalent thickness (i), precipitation (ii), and temperature (iii) are shown on the right panel.

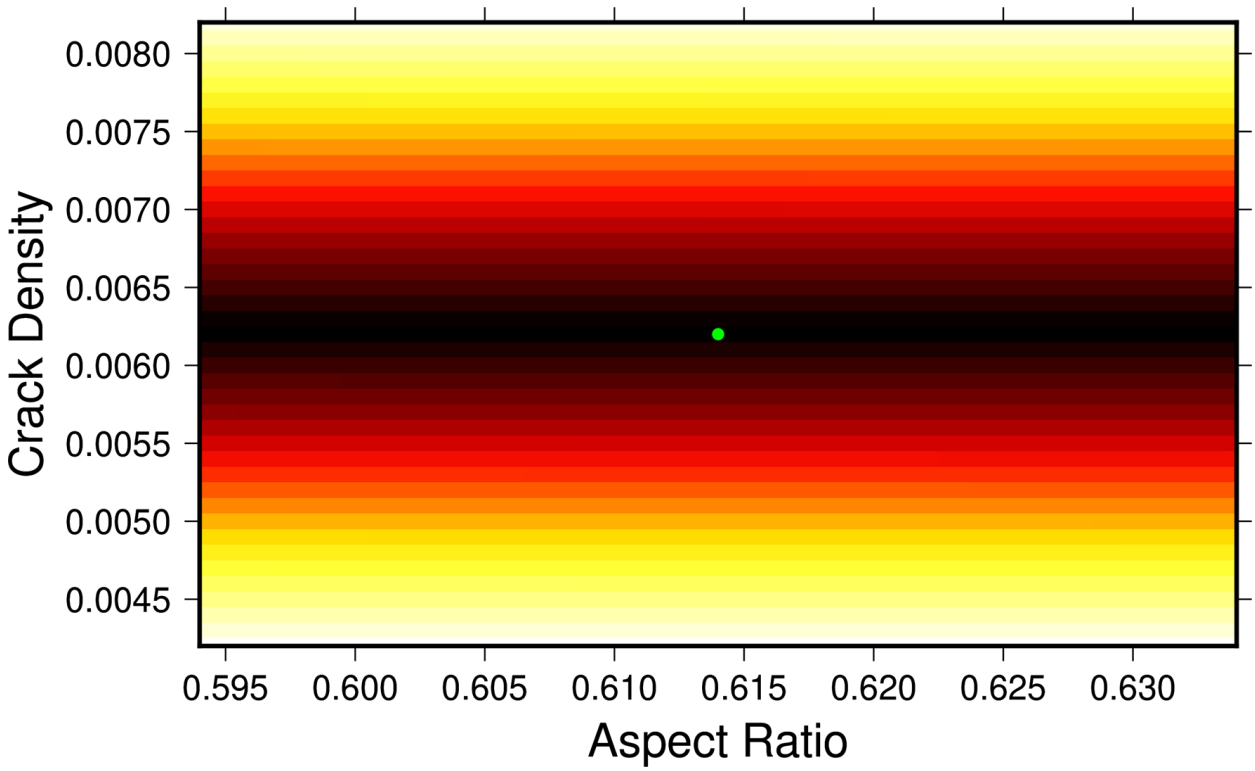

Figure 6.8: Distribution of the L1 misfit for searching the results of May . Green dot indicates the local minimum of the searching range. Note that the misfit function for the range is not sensitive to the aspect ratio, but a local minimum still can be found.

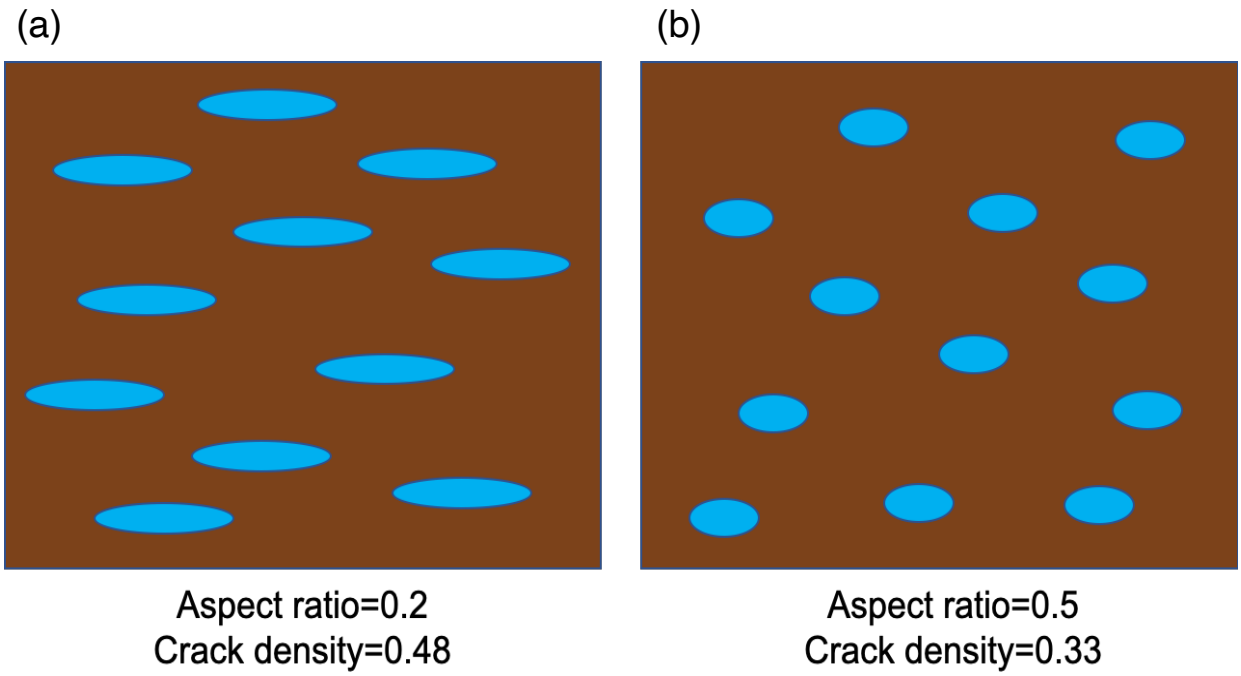

Figure 6.9: Schematic illumination of an anisotropic effective model. The left panel shows a case of stronger seismic anisotropic strength and a higher Vp/Vs ratio.

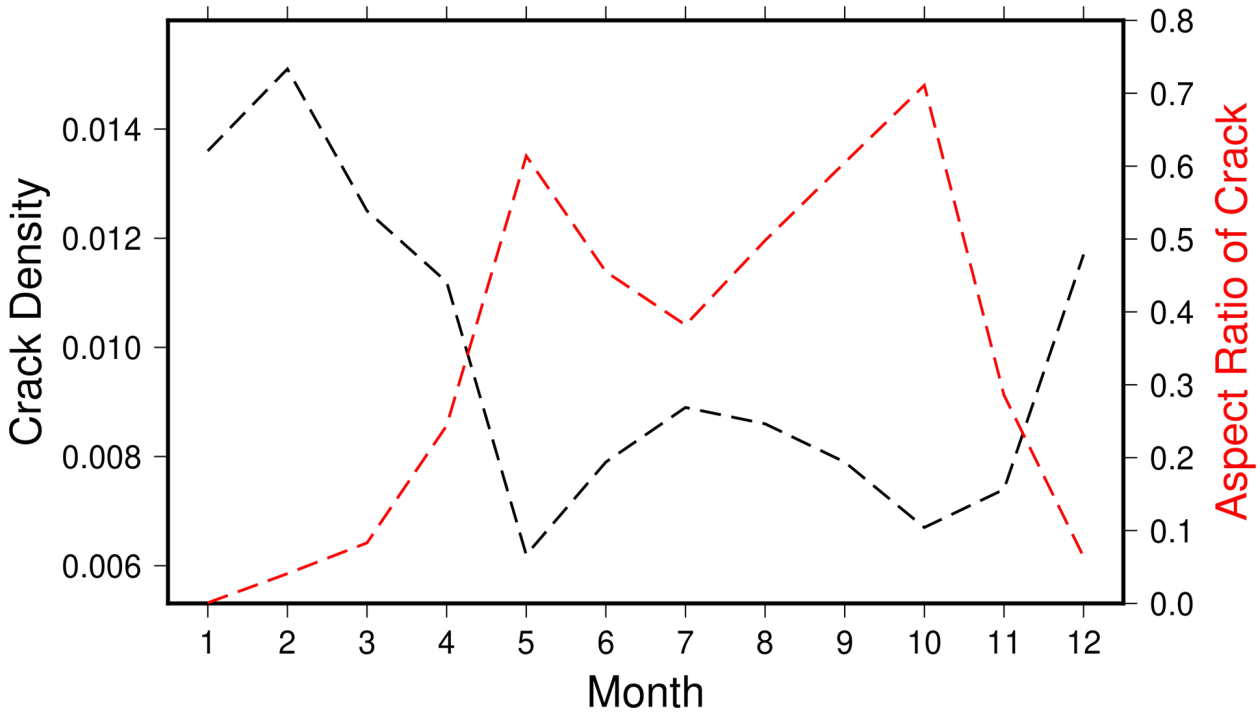

Figure 6.10: Modeled crack density and aspect ratio of each month by implementing the Hudson-Crampin anisotropic effective model.

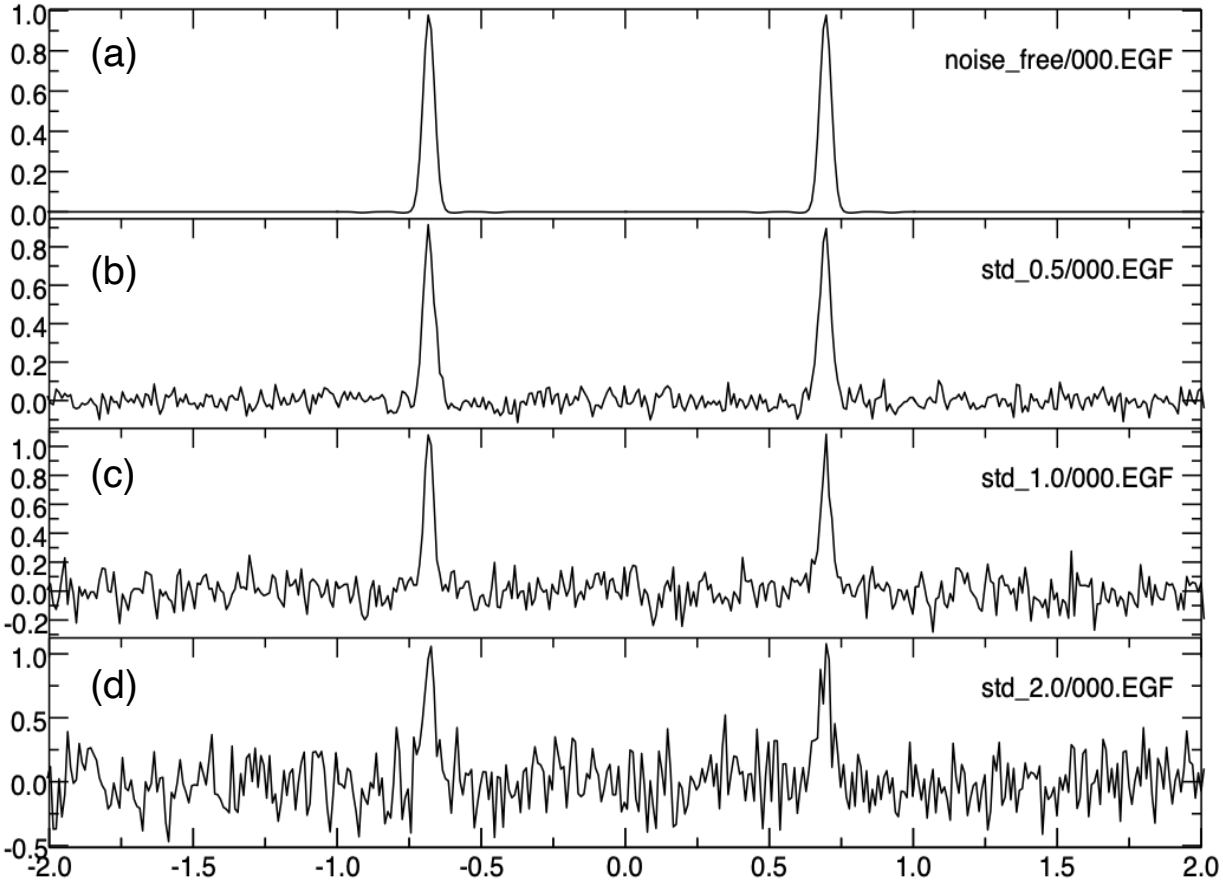

Figure 6.11: Synthetic test for capability of detecting tiny temporal variation by earthquake-based interferometry. Four synthetic waveforms in 100 samplings per second with different noise levels added.

Chapter 7

Conclusions and future perspectives

7.1 Conclusions

The source stacking method has been proposed to effectively reduce the numerical cost in wavefield computations for global tomography by at least an order of magnitude. Multiple earthquake sources are simultaneously triggered in one forward simulation only and the resulting summed wavefield can be compared with the corresponding observed summed wavefield at each station. In chapter 2, starting with proof-of-concept synthetic tests through a series of experiments, a toy 3D radially anisotropic model 'BLOB' which contains shear wave and radial anisotropy anomalies at different scales is considered. Based on the resultant 10,000 s synthetic waveform data, with and without realistic noise added, source stacked waveforms and station-pair cross-correlations of the summed waveforms are inverted to investigate if this second approach could overcome the issue like dominated large-amplitude fundamental surface waves. The use of cross-correlations brings out lower-amplitude overtone and body wave information on deeper mantle structure to the inversion and also makes it possible to implement path-weighting, an important step in seismic tomography.

When applying the Spectral Element Method (SEM), the effects of crustal and uppermost mantle on the waveforms must be carefully accounted for. However, the presence of thin slow layers in the crust largely slows down already heavy SEM computations. An alternative approach is to use the concept of homogenization to construct a smooth crustal model by global dispersion data. To address this, in chapter 3, global crust and uppermost mantle models are constructed by inverting fundamental mode Rayleigh and Love wave dispersion data and then combined the shallow part of the dispersion-based models (Moho-80 km depth) with the deeper part of SEMUCB-WM1 in order to assess how well such models can fit body and teleseismic waveform data. For that purpose, comparison between predicted and observed waveforms collected from 5 globally distributed events recorded at 200 very broadband stations are shown. The dispersion models show patterns of lag-time distribution that are correlated with tectonic provinces which indicates that fundamental mode dispersion-based models do not represent global shear velocity structure adequately at depths greater than 80 km. This is contrary to the assumption that fundamental mode dispersion-based tomographic inversion has the capability to retrieve shear wave velocity structure accurately down to 200 km depth. A workflow for global full-waveform inversion (FWI) is proposed, in which the top 60-80 km depth is first constrained by dispersion, using an approach based on the concept of effective medium, and the FWI is subsequently performed only at larger depths. This can significantly save the computation time involved in FWI.

In chapter 4, the first application of the methodology described in chapters 2 and 3 is applied to long period synthetic data computed in a realistic 3D global radially anisotropic model. Then the capability of retrieving both Vs and ξ perturbations is evaluated in the upper mantle at somewhat longer wavelengths than present in the target model. To understand the contribution of smaller amplitude overtones and body waves in CCF to the inversion, applying windowing on the cross-correlation function of the summed waveforms between pairs of stations are tested. The results show similar Vs retrieving capabilities of the

higher, fundamental, and mixed windowed CCFs. But the ξ results show that the different segments in CCFs individually provide varying resolving power. The issue of missing data could be reduced by replacing them by synthetics computed using NACT in the previous iteration 3D model. Also, grouping the data can help to minimize the number of missing waveforms in each group.

In chapter 5, a multidisciplinary collaborative research is presented and the seismological input is contributed to those provided by mineral physics and geodynamics, on anisotropy in the deepest 200-300 kilometers of the Earth's mantle. This study modeled the effect on seismic anisotropy of the forward and reverse Pv/pPv phase transition, in a scenario of a subducting slab reaching the CMB and the associated deformation included, spreading horizontally along the CMB and then turning into an upwelling. The results show that the 001 model, which is dominated by 001 slip plane in pPv, has a strong but complicated anisotropy, with $V_{SH} > V_{SV}$ in regions of faster than average isotropic V_s (e.g. graveyards of slabs), a tilted fast axis of SWS near the borders of the LLSVPs, and flow-aligned as observed in hotter than average regions of upwelling. All the results of the 001 model agree with the seismic observations (Panning and Romanowicz 2006; Cottaar and Romanowicz 2013; Lynner and Long 2014; Sturgeon et al. 2019) and better than the results of the 010 model.

Finally, in chapter 6, the results of temporal monitoring of near-surface seismic velocity and V_s azimuthal anisotropy over an interval of 15 years in Southern Hokkaido, by using earthquake-based interferometry are shown. The preliminary results show that station (IBUH03) has sudden reductions of V_{Siso} and V_P after the 2003 Mw 8.3 Tokachi-Oki and the 2018 Mw 6.7 Hokkaido Eastern Iburi earthquakes. and then recovers with different short-term and long-term recovery rates; changes in the fast S-wave polarization direction and strength of anisotropy are also observable during the coseismic periods. The results indicate that dynamic stress changes during the events induce the opening of fluid filled cracks, which is also responsible for the reductions of seismic velocities and the changes of V_s azimuthal anisotropy. Furthermore, seasonality is notable for the observations which implies the changes of elastic properties might also be affected by the hydrological response in the near surface. The results are compared with equivalent water thickness, precipitation, and temperature to understand the possible mechanism of the temporal variations. A Hudson-Crampin anisotropic effective model is implemented for modeling the seasonal crack behaviors. However, the other possible explanations are currently under investigation (i.e. fluid-related changes due to precipitation or groundwater level changes, thermo-elastic strain variations, or thermally induced stress). In the near future, applying the same analysis to more stations will help understand the underlying mechanisms.

7.2 Future perspectives

The source stacking method has been tested, and its resolving capability demonstrated for V_S and ξ. The next step is applying it to the real data. In chapter 4, a new global dataset that

includes 355 events and 340 very broadband recording stations for the real data is introduced. Future steps include improving quality control of the real data, since the conventional phase picking algorithm based on the comparison of single observed records with corresponding 3D synthetics cannot be applied directly. Another step in the continuation of this work is to deal with missing data issue by replacing subsets of NACT-computed synthetics and also grouping the data. The tests on weighting and combining the contributions from higher, fundamental, and mixed windowed CCFs, or even summed waveforms, are also necessary.

On the other hand, top portion of crustal model can be improved progressively as better independent constraints are available globally or regionally. For example, introducing shorter period dispersion data from ambient noise measurements or combing the results with receiver functions by utilizing teleseismic waveforms, can be taken into consideration. Further improvement to adopt shorter period dispersion data will require denser parametrization in the crust. Increasing the number of GLL points used to represent depth variations in the crust is necessary for waveform modeling at periods shorter than 30s.

Finally, in order to increase the resolution of the inversion, to attain wavelengths of interest for research on earth's deep structure, a finer model parameterization for the mantle will need to be implemented, and higher frequency synthetic and real waveforms considered. This means the current inversion code, as previously mentioned in chapters 2 and 3, needs to be modified to allow the construction of larger Hessian matrices. The technique has been developed, in particular for whole mantle inversions, using the UPC++ abstraction for the assembly of large matrices in the conventional global mantle FWI (e.g. French et al. 2015, IEEE).

Bibliography

Aki, K. (1957). "Space and time spectra of stationary stochastic waves, with special reference to microtremors". In: *Bulletin of the Earthquake Research Institute* 35, pp. 415–456.

Alford, R. (1986). "Shear data in the presence of azimuthal anisotropy: Dilley, Texas". In: *SEG Technical Program Expanded Abstracts 1986*. Society of Exploration Geophysicists, pp. 476–479.

Babuska, V. and M. Cara (1991). *Seismic anisotropy in the Earth*. Vol. 10. Springer Science & Business Media.

Bachmann, F., R. Hielscher, and H. Schaeben (2010). "Texture analysis with MTEX–free and open source software toolbox". In: *Solid State Phenomena*. Vol. 160. Trans Tech Publ, pp. 63–68.

Backus, G. E. (1962). "Long-wave elastic anisotropy produced by horizontal layering". In: *Journal of Geophysical Research* 67.11, pp. 4427–4440.

Becker, T. W. and L. Boschi (2002). "A comparison of tomographic and geodynamic mantle models". In: *Geochemistry, Geophysics, Geosystems* 3.1, p. 1003.

Bozdağ, E., D. Peter, M. Lefebvre, D. Komatitsch, J. Tromp, J. Hill, N. Podhorszki, and D. Pugmire (2016). "Global adjoint tomography: first-generation model". In: *Geophysical Journal International* 207.3, pp. 1739–1766.

Browaeys, J. T. and S. Chevrot (2004). "Decomposition of the elastic tensor and geophysical applications". In: *Geophysical Journal International* 159.2, pp. 667–678.

Bullen, K. (1949). "Compressibility-pressure hypothesis and the Earth's interior". In: *Geophysical Journal International* 5, pp. 335–368.

Capdeville, Y., E. Chaljub, and J. P. Montagner (2003). "Coupling the spectral element method with a modal solution for elastic wave propagation in global earth models". In: *Geophysical Journal International* 152.1, pp. 34–67.

Capdeville, Y., Y. Gung, and B. Romanowicz (2005). "Towards global earth tomography using the spectral element method: a technique based on source stacking". In: *Geophysical Journal International* 162.2, pp. 541–554.

Capdeville, Y. and J.-J. Marigo (2007). "Second order homogenization of the elastic wave equation for non-periodic layered media". In: *Geophysical Journal International* 170.2, pp. 823–838.

Chandler, B. C., L.-W. Chen, M. Li, B. Romanowicz, and H.-R. Wenk (2021). "Seismic anisotropy, dominant slip systems and phase transitions in the lowermost mantle". In: *Geophysical Journal International* 227.3, pp. 1665–1681.

Chen, L.-W., Y.-N. Chen, Y. Gung, J.-C. Lee, and W.-T. Liang (2017). "Strong near-surface seismic anisotropy of Taiwan revealed by coda interferometry". In: *Earth and Planetary Science Letters* 475, pp. 224–230.

Cottaar, S., M. Li, A. K. McNamara, B. Romanowicz, and H.-R. Wenk (2014). "Synthetic seismic anisotropy models within a slab impinging on the core–mantle boundary". In: *Geophysical Journal International* 199.1, pp. 164–177.

Cottaar, S. and B. Romanowicz (2013). "Observations of changing anisotropy across the southern margin of the African LLSVP". In: *Geophysical Journal International* 195.2, pp. 1184–1195.

Crampin, S. (1984). "Effective anisotropic elastic constants for wave propagation through cracked solids". In: *Geophysical Journal International* 76.1, pp. 135–145.

— (1987). "Geological and industrial implications of extensive-dilatancy anisotropy". In: *Nature* 328.6130, pp. 491–496.

— (1994). "The fracture criticality of crustal rocks". In: *Geophysical Journal International* 118.2, pp. 428–438.

Dobson, D. P., N. Miyajima, F. Nestola, M. Alvaro, N. Casati, C. Liebske, I. G. Wood, and A. M. Walker (2013). "Strong inheritance of texture between perovskite and post-perovskite in the D" layer". In: *Nature Geoscience* 6.7, pp. 575–578.

Dziewonski, A. M. and D. L. Anderson (1981). "Preliminary reference Earth model". In: *Physics of the Earth and Planetary Interiors* 25.4, pp. 297–356.

Dziewonski, A. M., B. H. Hager, and R. J. O'Connell (1977). "Large-scale heterogeneities in the lower mantle". In: *Journal of Geophysical Research* 82.2, pp. 239–255.

Dziewonski, A., T.-A. Chou, and J. H. Woodhouse (1981). "Determination of earthquake source parameters from waveform data for studies of global and regional seismicity". In: *Journal of Geophysical Research: Solid Earth* 86.B4, pp. 2825–2852.

Ekström, G., M. Nettles, and A. Dziewonski (2012). "The global CMT project 2004 – 2010 : Centroid-moment tensors". In: *Physics of the Earth and Planetary Interiors* 201, pp. 1–9.

Eshelby, J. D. (1957). "The determination of the elastic field of an ellipsoidal inclusion, and related problems". In: *Proceedings of the royal society of London. Series A. Mathematical and physical sciences* 241.1226, pp. 376–396.

Ferreira, A., J. Woodhouse, K. Visser, and J. Trampert (2010). "On the robustness of global radially anisotropic surface wave tomography". In: *Journal of Geophysical Research: Solid Earth* 115.B4, B04313.

Fichtner, A. and J. Trampert (2011). "Hessian kernels of seismic data functionals based upon adjoint techniques". In: *Geophysical Journal International* 185.2, pp. 775–798.

Fichtner, A., J. Trampert, P. Cupillard, E. Saygin, T. Taymaz, Y. Capdeville, and A. Villasenor (2013). "Multiscale full waveform inversion". In: *Geophysical Journal International* 194.1, pp. 534–556.

French, S., V. Lekic, and B. Romanowicz (2013). "Waveform tomography reveals channeled flow at the base of the oceanic asthenosphere". In: *Science* 342.6155, pp. 227–230.

French, S., Y. Zheng, B. Romanowicz, and K. Yelick (2015). "Parallel Hessian assembly for seismic waveform inversion using global updates". In: *2015 IEEE International Parallel and Distributed Processing Symposium*. IEEE, pp. 753–762.

French, S. W. and B. Romanowicz (2015). "Broad plumes rooted at the base of the Earth's mantle beneath major hotspots". In: *Nature* 525.7567, pp. 95–99.

French, S. and B. Romanowicz (2014). "Whole-mantle radially anisotropic shear velocity structure from spectral-element waveform tomography". In: *Geophysical Journal International* 199.3, pp. 1303–1327.

Gassenmeier, M., C. Sens-Schönfelder, M. Delatre, and M. Korn (2014). "Monitoring of environmental influences on seismic velocity at the geological storage site for CO_2 in Ketzin (Germany) with ambient seismic noise". In: *Geophysical Journal International* 200.1, pp. 524–533.

Gassenmeier, M., C. Sens-Schönfelder, T. Eulenfeld, M. Bartsch, P. Victor, F. Tilmann, and M. Korn (2016). "Field observations of seismic velocity changes caused by shaking-induced damage and healing due to mesoscopic nonlinearity". In: *Geophysical Journal International* 204.3, pp. 1490–1502.

Gerstoft, P., P. M. Shearer, N. Harmon, and J. Zhang (2008). "Global P, PP, and PKP wave microseisms observed from distant storms". In: *Geophysical Research Letters* 35.23, p. L23306.

Goryaeva, A. M., P. Carrez, and P. Cordier (2016). "Low viscosity and high attenuation in $MgSiO_3$ post-perovskite inferred from atomic-scale calculations". In: *Scientific Reports* 6.1, pp. 1–10.

Haned, A., E. Stutzmann, M. Schimmel, S. Kiselev, A. Davaille, and A. Yelles-Chaouche (2016). "Global tomography using seismic hum". In: *Geophysical Journal International* 204.2, pp. 1222–1236.

Hirose, K., R. Sinmyo, N. Sata, and Y. Ohishi (2006). "Determination of post-perovskite phase transition boundary in $MgSiO_3$ using Au and MgO pressure standards". In: *Geophysical Research Letters* 33.1.

Karki, B., R. M. Wentzcovitch, S. De Gironcoli, and S. Baroni (2000). "High-pressure lattice dynamics and thermoelasticity of MgO". In: *Physical Review B* 61.13, pp. 8793–8800.

Kawai, K. and T. Tsuchiya (2009). "Temperature profile in the lowermost mantle from seismological and mineral physics joint modeling". In: *Proceedings of the National Academy of Sciences* 106.52, pp. 22119–22123.

Komatitsch, D. and J. Tromp (1999). "Introduction to the spectral element method for three-dimensional seismic wave propagation". In: *Geophysical Journal International* 139.3, pp. 806–822.

— (2002). "Spectral-element simulations of global seismic wave propagation—I. Validation". In: *Geophysical Journal International* 149.2, pp. 390–412.

Komatitsch, D. and J.-P. Vilotte (1998). "The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures". In: *Bulletin of the Seismological Society of America* 88.2, pp. 368–392.

Krebs, J. R., J. E. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M.-D. Lacasse (2009). "Fast full-wavefield seismic inversion using encoded sources". In: *Geophysics* 74.6, WCC177–WCC188.

Kustowski, B., G. Ekström, and A. Dziewoński (2008). "The shear-wave velocity structure in the upper mantle beneath Eurasia". In: *Geophysical Journal International* 174.3, pp. 978–992.

Lebensohn, R. A. and C. Tomé (1993). "A self-consistent anisotropic approach for the simulation of plastic deformation and texture development of polycrystals: application to zirconium alloys". In: *Acta Metallurgica et Materialia* 41.9, pp. 2611–2624.

Lei, W., Y. Ruan, E. Bozdağ, D. Peter, M. Lefebvre, D. Komatitsch, J. Tromp, J. Hill, N. Podhorszki, and D. Pugmire (2020). "Global adjoint tomography—model GLAD-M25". In: *Geophysical Journal International* 223.1, pp. 1–21.

Lekić, V., M. Panning, and B. Romanowicz (2010). "A simple method for improving crustal corrections in waveform tomography". In: *Geophysical Journal International* 182.1, pp. 265–278.

Lekić, V. and B. Romanowicz (2011). "Inferring upper-mantle structure by full waveform tomography with the spectral element method". In: *Geophysical Journal International* 185.2, pp. 799–831.

Li, X.-D. and B. Romanowicz (1995). "Comparison of global waveform inversions with and without considering cross-branch modal coupling". In: *Geophysical Journal International* 121.3, pp. 695–709.

— (1996). "Global mantle shear velocity model developed using nonlinear asymptotic coupling theory". In: *Journal of Geophysical Research: Solid Earth* 101.B10, pp. 22245–22272.

Li, X.-D. and T. Tanimoto (1993). "Waveforms of long-period body waves in a slightly aspherical Earth model". In: *Geophysical Journal International* 112.1, pp. 92–102.

Lin, F.-C., V. C. Tsai, B. Schmandt, Z. Duputel, and Z. Zhan (2013). "Extracting seismic core phases with array interferometry". In: *Geophysical Research Letters* 40.6, pp. 1049–1053.

Long, M. D. (2009). "Complex anisotropy in D" beneath the eastern Pacific from SKS–SKKS splitting discrepancies". In: *Earth and Planetary Science Letters* 283.1-4, pp. 181–189.

Lynner, C. and M. D. Long (2014). "Lowermost mantle anisotropy and deformation along the boundary of the African LLSVP". In: *Geophysical Research Letters* 41.10, pp. 3447–3454.

Maggi, A., C. Tape, M. Chen, D. Chao, and J. Tromp (2009). "An automated time-window selection algorithm for seismic tomography". In: *Geophysical Journal International* 178.1, pp. 257–281.

McNamara, A. K., P. E. van Keken, and S.-I. Karato (2003). "Development of finite strain in the convecting lower mantle and its implications for seismic anisotropy". In: *Journal of Geophysical Research: Solid Earth* 108.B5, ETG 3-1 –ETG 3-14.

McNamara, A. K., P. E. Van Keken, and S.-I. Karato (2002). "Development of anisotropic structure in the Earth's lower mantle by solid-state convection". In: *Nature* 416.6878, pp. 310–314.

Mégnin, C. and B. Romanowicz (2000). "The three-dimensional shear velocity structure of the mantle from the inversion of body, surface and higher-mode waveforms". In: *Geophysical Journal International* 143.3, pp. 709–728.

Meier, U., A. Curtis, and J. Trampert (2007). "Fully nonlinear inversion of fundamental mode surface waves for a global crustal model". In: *Geophysical Research Letters* 34.16.

Meier, U., N. M. Shapiro, and F. Brenguier (2010). "Detecting seasonal variations in seismic velocities within Los Angeles basin from correlations of ambient seismic noise". In: *Geophysical Journal International* 181.2, pp. 985–996.

Merkel, S., A. K. McNamara, A. Kubo, S. Speziale, L. Miyagi, Y. Meng, T. S. Duffy, and H.-R. Wenk (2007). "Deformation of (Mg, Fe) SiO_3 post-perovskite and D" anisotropy". In: *Science* 316.5832, pp. 1729–1732.

Miyagi, L., W. Kanitpanyacharoen, P. Kaercher, K. K. Lee, and H.-R. Wenk (2010). "Slip systems in $MgSiO_3$ post-perovskite: implications for D" anisotropy". In: *Science* 329.5999, pp. 1639–1641.

Mochizuki, E. (1986). "Free oscillations and surface waves of an aspherical Earth". In: *Geophysical Research Letters* 13.13, pp. 1478–1481.

Montagner, J.-P. and D. L. Anderson (1989). "Constrained reference mantle model". In: *Physics of the Earth and Planetary Interiors* 58.2-3, pp. 205–227.

Nataf, H.-C., I. Nakanishi, and D. L. Anderson (1986). "Measurements of mantle wave velocities and inversion for lateral heterogeneities and anisotropy: 3. Inversion". In: *Journal of Geophysical Research: Solid Earth* 91.B7, pp. 7261–7307.

Nishida, K. (2013). "Earth's background free oscillations". In: *Annual Review of Earth and Planetary Sciences* 41, pp. 719–740.

Nolet, G. and F. Dahlen (2000). "Wave front healing and the evolution of seismic delay times". In: *Journal of Geophysical Research: Solid Earth* 105.B8, pp. 19043–19054.

Nomura, R., K. Hirose, K. Uesugi, Y. Ohishi, A. Tsuchiyama, A. Miyake, and Y. Ueno (2014). "Low core-mantle boundary temperature inferred from the solidus of pyrolite". In: *Science* 343.6170, pp. 522–525.

Nowacki, A., J. Wookey, and J.-M. Kendall (2011). "New advances in using seismic anisotropy, mineral physics and geodynamics to understand deformation in the lowermost mantle". In: *Journal of Geodynamics* 52.3-4, pp. 205–228.

Oganov, A. R. and S. Ono (2004). "Theoretical and experimental evidence for a post-perovskite phase of MgSiO 3 in Earth's D" layer". In: *Nature* 430.6998, pp. 445–448.

Panning, M. and B. Romanowicz (2006). "A three-dimensional radially anisotropic model of shear velocity in the whole mantle". In: *Geophysical Journal International* 167.1, pp. 361–379.

Park, J. (1987). "Asymptotic coupled-mode expressions for multiplet amplitude anomalies and frequency shifts on an aspherical earth". In: *Geophysical Journal International* 90.1, pp. 129–169.

Pasyanos, M. E. and A. A. Nyblade (2007). "A top to bottom lithospheric study of Africa and Arabia". In: *Tectonophysics* 444.1-4, pp. 27–44.

Peterson, J. R. (1993). "Observations and modeling of seismic background noise". In: 93-322.

Pham, T.-S., H. Tkalčić, M. Sambridge, and B. L. Kennett (2018). "Earth's correlation wavefield: Late coda correlation". In: *Geophysical Research Letters* 45.7, pp. 3035–3042.

Rickers, F., A. Fichtner, and J. Trampert (2013). "The Iceland–Jan Mayen plume system and its impact on mantle dynamics in the North Atlantic region: evidence from full-waveform inversion". In: *Earth and Planetary Science Letters* 367, pp. 39–51.

Romanowicz, B. (1987). "Multiplet-multiplet coupling due to lateral heterogeneity: asymptotic effects on the amplitude and frequency of the Earth's normal modes". In: *Geophysical Journal International* 90.1, pp. 75–100.

Romanowicz, B., L.-W. Chen, and S. W. French (2020). "Accelerating full waveform inversion via source stacking and cross-correlations". In: *Geophysical Journal International* 220.1, pp. 308–322.

Romanowicz, B. and H.-R. Wenk (2017). "Anisotropy in the deep Earth". In: *Physics of the Earth and Planetary Interiors* 269, pp. 58–90.

Romanowicz, B. A., M. P. Panning, Y. Gung, and Y. Capdeville (2008). "On the computation of long period seismograms in a 3-D earth using normal mode based approximations". In: *Geophysical Journal International* 175.2, pp. 520–536.

Schaff, D. P. and G. C. Beroza (2004). "Coseismic and postseismic velocity changes measured by repeating earthquakes". In: *Journal of Geophysical Research: Solid Earth* 109.B10.

Schuster, G. T., X. Wang, Y. Huang, W. Dai, and C. Boonyasiriwat (2011). "Theory of multisource crosstalk reduction by phase-encoded statics". In: *Geophysical Journal International* 184.3, pp. 1289–1303.

Sens-Schönfelder, C. and U. Wegler (2006). "Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia". In: *Geophysical Research Letters* 33.21, p. L21302.

Shapiro, N. M. and M. Campillo (2004). "Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise". In: *Geophysical Research Letters* 31.7.

Shapiro, N. and M. Ritzwoller (2002). "Monte-Carlo inversion for a global shear-velocity model of the crust and upper mantle". In: *Geophysical Journal International* 151.1, pp. 88–105.

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

Simmons, N. and S. Grand (2002). "Partial melting in the deepest mantle". In: *Geophysical Research Letters* 29.11, pp. 47-1 - 47–4.

Stehly, L., M. Campillo, and N. Shapiro (2006). "A study of the seismic noise from its long-range correlation properties". In: *Journal of Geophysical Research: Solid Earth* 111.B10, B10306.

Steinberger, B. (2000). "Plumes in a convecting mantle: Models and observations for individual hotspots". In: *Journal of Geophysical Research: Solid Earth* 105.B5, pp. 11127–11152.

Sturgeon, W., A. M. Ferreira, M. Faccenda, S.-J. Chang, and L. Schardong (2019). "On the origin of radial anisotropy near subducted slabs in the midmantle". In: *Geochemistry, Geophysics, Geosystems* 20.11, pp. 5105–5125.

Suda, N., K. Nawa, and Y. Fukao (1998). "Earth's background free oscillations". In: *Science* 279.5359, pp. 2089–2091.

Swenson, S. and J. Wahr (2002). "Methods for inferring regional surface-mass anomalies from Gravity Recovery and Climate Experiment (GRACE) measurements of time-variable gravity". In: *Journal of Geophysical Research: Solid Earth* 107.B9, ETG 3-1 –ETG 3-13.

Taira, T., A. Nayak, F. Brenguier, and M. Manga (2018). "Monitoring reservoir response to earthquakes and fluid extraction, Salton Sea geothermal field, California". In: *Science Advances* 4.1, e1701536.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004). "GRACE measurements of mass variability in the Earth system". In: *Science* 305.5683, pp. 503–505.

Tarantola, A. (1984). "Inversion of seismic reflection data in the acoustic approximation". In: *Geophysics* 49.8, pp. 1259–1266.

Tsai, V. C. (2009). "On establishing the accuracy of noise tomography travel-time measurements in a realistic medium". In: *Geophysical Journal International* 178.3, pp. 1555–1564.

Tsuchiya, T., J. Tsuchiya, K. Umemoto, and R. M. Wentzcovitch (2004). "Elasticity of post-perovskite $MgSiO_3$". In: *Geophysical Research Letters* 31.14.

Wahr, J., S. Swenson, V. Zlotnicki, and I. Velicogna (2004). "Time-variable gravity from GRACE: First results". In: *Geophysical Research Letters* 31.11.

Walker, A., A. Forte, J. Wookey, A. Nowacki, and J.-M. Kendall (2011). "Elastic anisotropy of D" predicted from global models of mantle flow". In: *Geochemistry, Geophysics, Geosystems* 12.10.

Wang, Z. and F. Dahlen (1995). "Spherical-spline parameterization of three-dimensional Earth models". In: *Geophysical Research Letters* 22.22, pp. 3099–3102.

Wenk, H.-R., S. Cottaar, C. N. Tomé, A. McNamara, and B. Romanowicz (2011). "Deformation in the lowermost mantle: From polycrystal plasticity to seismic anisotropy". In: *Earth and Planetary Science Letters* 306.1-2, pp. 33–45.

Williams, Q. and E. J. Garnero (1996). "Seismic evidence for partial melt at the base of Earth's mantle". In: *Science* 273.5281, pp. 1528–1530.

Woodhouse, J. (1998). "The calculation of eigenfrequencies and eigenfunctions of the free oscillations of the Earth and the Sun". In: *Seismological Algorithms*, pp. 321–370.

Woodhouse, J. H. and A. M. Dziewonski (1984). "Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms". In: *Journal of Geophysical Research: Solid Earth* 89.B7, pp. 5953–5986.

Wu, X., J.-F. Lin, P. Kaercher, Z. Mao, J. Liu, H.-R. Wenk, and V. B. Prakapenka (2017). "Seismic anisotropy of the D" layer induced by (001) deformation of post-perovskite". In: *Nature Communications* 8.1, pp. 1–6.

Yuan, K. and B. Romanowicz (July 2017). "Seismic evidence for partial melting at the root of major hot spot plumes". en. In: *Science* 357.6349, pp. 393–397.

Zhang, Q., W. Mao, H. Zhou, H. Zhang, and Y. Chen (2018). "Hybrid-domain simultaneous-source full waveform inversion without crosstalk noise". In: *Geophysical Journal International* 215.3, pp. 1659–1681.

Zhang, S., S. Cottaar, T. Liu, S. Stackhouse, and B. Militzer (2016). "High-pressure, temperature elasticity of Fe-and Al-bearing $MgSiO_3$: Implications for the Earth's lower mantle". In: *Earth and Planetary Science Letters* 434, pp. 264–273.

Zhong, S. (2006). "Constraints on thermochemical convection of the mantle from plume heat flux, plume excess temperature, and upper mantle temperature". In: *Journal of Geophysical Research: Solid Earth* 111.B4, B04409.

Zhu, H., E. Bozdağ, and J. Tromp (2015). "Seismic structure of the European upper mantle based on adjoint tomography". In: *Geophysical Journal International* 201.1, pp. 18–52.

Chapter A

Appendix

A.1 Misfit function based on cross-correlations

A.1.1 Discrete misfit function

Let E be the set of seismic events to be summed. Further, let R be the common subset of receivers spanning event set E. For all receiver pairs $(p, q) \in R \times R$, we wish to minimize the discrete, weighted cross-correlogram difference misfit function

$$\Phi_E(\mathbf{m}) = \sum_{(p,q)\in R\times R} \Phi^{(p,q)}(\mathbf{m}) \tag{A.1}$$

where $\mathbf{m}$ is the earth model and the inner, trace-pair misfit $\Phi^{(p,q)}(\mathbf{m})$ is given by:

$$\Phi^{(p,q)}(\mathbf{m}) = \frac{1}{2} \sum_{i=-N+1}^{N-1} \mathbf{W}_{i,i}^{(p,q)} \left[\Delta\mathcal{C}_i^{(p,q)}\right]^2 \tag{A.2}$$

in which the difference of discrete cross-correlograms $\Delta\mathcal{C}_i^{(p,q)}$ (i.e. the residual) is given by

$$\Delta\mathcal{C}_i^{(p,q)} = \mathcal{C}_i(\mathbf{d}^{(p)}, \mathbf{d}^{(q)}) - \mathcal{C}_i(\mathbf{G}^{(p)}(\mathbf{m}), \mathbf{G}^{(q)}(\mathbf{m}))$$

where $\mathbf{d}$ and $\mathbf{G}(\mathbf{m})$ are the summed data and synthetics for E (each a time series of length N), $\mathcal{C}_i(\cdot,\cdot)$ represents the discrete cross-correlation operator evaluated at time point t_i, and $\mathbf{W}$ is a diagonal weighting matrix reflecting both path uniqueness (for station pair (p, q)) and time-dependent windowing.

A.1.2 Minimization

In order to solve this minimization problem in the context of the generalized least-squares formalism, we require expressions for both the gradient of eq. A.1 and the Jacobian of the

discrete cross-correlogram, both with respect to the model $\mathbf{m}$ (the former also implicitly depends upon the latter). After some manipulation of eq. A.2, we find that the misfit gradient for a single station pair is given by

$$\frac{\partial}{\partial \mathbf{m}_k}\Phi^{(p,q)}(\mathbf{m}) = -\sum_{i=-N+1}^{N-1} \mathbf{W}_{i,i}^{(p,q)} \Delta\mathcal{C}_i^{(p,q)} \mathbf{G}_{i,k}^{(p,q)}$$

where the Jacobian of the discrete station-pair cross-correlogram is give by

$$\mathbf{G}_{i,k}^{(p,q)} = \frac{\partial}{\partial \mathbf{m}_k}\left\{\mathcal{C}_i(\mathbf{G}^{(p)}(\mathbf{m}), \mathbf{G}^{(q)}(\mathbf{m}))\right\}$$

For our purposes, it suffices to develop an expression for the latter.

A.1.3 Cross-correlogram Jacobian

Defining the above discrete cross-correlation of two real time series as

$$\mathcal{C}_i(s^{(p)}, s^{(q)}) = \sum_j s_j^{(p)} s_{i+j}^{(q)}$$

it follows that

$$\begin{aligned}
\mathbf{G}_{i,k}^{(p,q)} &= \frac{\partial}{\partial \mathbf{m}_k}\sum_j s_j^{(p)} s_{i+j}^{(q)} \\
&= \sum_j \frac{\partial s_j^{(p)}}{\partial \mathbf{m}_k} s_{i+j}^{(q)} + s_j^{(p)} \frac{\partial s_{i+j}^{(q)}}{\partial \mathbf{m}_k} \\
&= \sum_j \mathbf{G}_{j,k}^{(p)} s_{i+j}^{(q)} + s_j^{(p)} \mathbf{G}_{i+j,k}^{(q)}
\end{aligned}$$

where $\mathbf{G}^{(p)}$ and $\mathbf{G}^{(q)}$ are the traditional summed-wavefield Jacobians for event set E, evaluated at receivers p and q, respectively:

$$\mathbf{G}_{i,k}^{(p)} = \frac{\partial}{\partial \mathbf{m}_k}\mathbf{G}_i^{(p)}(\mathbf{m}) \quad \text{and} \quad \mathbf{G}_{i,k}^{(q)} = \frac{\partial}{\partial \mathbf{m}_k}\mathbf{G}_i^{(q)}(\mathbf{m})$$

Finally, it can be seen that

$$\mathbf{G}^{(p,q)} = \mathbf{S}_L^{(q)}\mathbf{G}^{(p)} + \mathbf{S}_R^{(p)}\mathbf{G}^{(q)}$$

where $\mathbf{S}_L$ and $\mathbf{S}_R$ are matrices containing left and right shifted synthetic time series, of the form

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & s_0 \\ 0 & 0 & \cdots & s_0 & s_1 \\ & & \ddots & & \\ 0 & s_0 & \cdots & s_{N-3} & s_{N-2} \\ s_0 & s_1 & \cdots & s_{N-2} & s_{N-1} \\ s_1 & s_2 & \cdots & s_{N-1} & 0 \\ & & \ddots & & \\ s_{N-2} & s_{N-1} & \cdots & 0 & 0 \\ s_{N-1} & 0 & \cdots & 0 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} s_{N-1} & 0 & \cdots & 0 & 0 \\ s_{N-2} & s_{N-1} & \cdots & 0 & 0 \\ & & \ddots & & \\ s_1 & s_2 & \cdots & s_{N-1} & 0 \\ s_0 & s_1 & \cdots & s_{N-2} & s_{N-1} \\ 0 & s_0 & \cdots & s_{N-3} & s_{N-2} \\ & & \ddots & & \\ 0 & 0 & \cdots & s_0 & s_1 \\ 0 & 0 & \cdots & 0 & s_0 \end{pmatrix}$$

respectively. Thus, $\mathbf{G}^{(p,q)}$ may be computed as a post-processing step, combining the individual station Jacobians $\mathbf{G}^{(p)}$ and $\mathbf{G}^{(q)}$ with the corresponding synthetic traces through $\mathbf{S}_R^{(p)}$ and $\mathbf{S}_L^{(q)}$.