

Adjoint Tomography of Surface Wave Observables from Ambient Seismic Noise

By

Kai Wang

A thesis submitted to Macquarie University
for the degree of Doctor of Philosophy
Department of Earth and Planetary Sciences
September 2018



MACQUARIE
University
SYDNEY · AUSTRALIA

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Kai Wang

Acknowledgements

During the past three-years' PhD study, I have received a lot of help from many professors, colleagues, friends, and my family. Without their encouragements, supports and accompany, I could not make all the way through the PhD program.

First and foremost, I would like to express my deepest appreciation to my supervisor Prof. Yingjie Yang, and thank for his continuous supports and extreme patience towards my research. Like every PhD student might have gone through, I was struggling with the main research topics in the early days of my PhD study. Fortunately, Yingjie guided me through this hard time with his persistence in science and insightful intuition towards seismology. I am always inspired by his innovative scientific ideas and ambitious thoughts of developing tools to advance seismic tomography, which motivates my exploration of a variety of tomographic techniques. Apart from the professional guidance, he also kindly provided me with amazing opportunities for international collaborations at two other top-ranking research groups. These invaluable experiences have significantly enriched my knowledge of seismology, and fulfilled my PhD life.

I would like to extend my sincere gratitude to Prof. Qinya Liu from University of Toronto. It is my fortune to work with her on the ambient noise adjoint tomography (ANAT) project. During my half-year visit, she passed on her knowledge and experience in spectral element method (SEM) and adjoint tomography to me unreservedly. She also devoted amounts of efforts to revising our manuscripts, and I was deeply impressed by her conscientious attitude towards science and profound knowledge in math and physics. I also want to take this opportunity to thank another collaborator of the ANAT project, Prof. Carl Tape from University of Alaska. He is a very responsible professor and always comes up with insightful suggestions for the manuscripts. I really had a great time during the visit at Toronto

and received many helps from Qi Zhao, Chuangxin Lin, Xin Song, Changyi Yu, Shan Wu, Liang Ding and other friends. I benefited considerably from the discussions with Chuangxin on SEM, and special thanks to my collegemate Qi for accommodating me during my first week in Toronto and looking after me in many aspects.

I would like to thank Prof. Shu-huei Hung from National Taiwan University for teaching me basics of finite-frequency body wave tomography and her student Yu-Hsuan Chang for sharing his experience in finite-frequency tomography and LSQR inversion scheme. Without these fundamental studies, I could not finish developing the tool for joint inversion of ambient noise surface waves and teleseismic body waves and applying it to NE China. Thanks for all the kind help from Sin-Mei Wu, Weiwei Wang, Meng-Jie Cai, Shuei-Huei You, Sin-Ying Yang and many others during my three-months' visit at Taipei.

My thanks also go to Prof. Yinhe Luo from China University of Geosciences (CUG, Wuhan), who guided me into research and trained me with basic knowledge and understanding of seismological theory in the two and half years' graduate study at CUG. The nice supervision along with the knowledge I learned from him will never be forgotten.

I would like to specially thank Chengxing Jiang, who not only helped me in many aspects of personal life but also assisted me in research with great encouragements and wonderful advice. He kindly helped review some chapters of this thesis. I also want to thank my former roommate Zhen Guo for his help in terms of both personal life and research, and I learned a lot during the collaboration with him on the joint inversion project in NE China. Thanks to all my former and current officemates, Shucheng Wu, Shilin Li, Mehdi Tork Qashqai, Farshad Salajegheh, Jun Xie, Guoliang Li, Ying Chen, Anqi Zhang, Thomas Connell, Lian Jiang, Zhongxuan Li and others. I feel enjoyable to discuss research with them, and the daily chats also made my office time delightful. Special thanks to Zhongxuan for her encouragements and accompany during the time of the thesis examination. Thanks to the great administrative team in the department, Alarnah-Jade Cullen, Wendy Southwell, Phil Dartnell, and Aida Pujol Conejero, for their supports over the past years.

I always feel lucky to be a member of the large geoscience family in our department, formed by many Post-docs, PhD students, and their family members: Siqi Zhang and his wife Wei Meng, Jinxiang Huang, Qing Xiong and his wife Jin Zhang, Yu Wang, Jianggu Lu, Chengxin and his wife Song Lu, Shucheng and his wife Qingping Hu, Guoliang, Bo Xu, Huiyuan Xu, Anqi Zhang, and some other visitors. We share experience on both research

and daily life, and take care of each other in every aspect. I never felt alone with their encouragements, understanding and supports over the past three years. The friendship we developed has been a great comfort to me and will continue to the future.

Finally, I would like to thank my parents for their unselfish and endless love to me as well as what they have scarified for the whole family. My mother spent her entire time of retirement to bring up my little niece to allow my elder brother and his wife to work outside trustingly. My sixty-year old father is still working hard to support our family and save money for the expenditure of his little son's future marriage. I am deeply touched by the fact that they never complained but have been always worried about my studies, health and every other thing. I thank my brother and his wife to take care of my parents when I am not with them, as well as my niece for being a constant source of comfort to me.

Abstract

Adjoint tomography has recently been introduced to ambient seismic noise data as a new and promising tomographic method which highlights simulation-based 3-D sensitivity kernels against the approximated analytical theory used in traditional ambient noise tomography. The basic idea of ambient noise adjoint tomography (ANAT) is to iteratively minimize the traveltimes between empirical Green's functions (EGFs) from ambient noise cross correlations and synthetic Green's functions from spectral-element simulations based on misfit gradient.

To date, most studies use vertical point-force sources to simulate the synthetics for Rayleigh wave EGFs from vertical-vertical component ambient noise cross correlation functions. In this thesis, I develop a semi-automatically iterative inversion package for ANAT based on the software SPECFEM3D Cartesian. I apply the inversion package to ANAT of 5 – 50 s Rayleigh wave EGFs in southern California and construct an improved V_{sv} model based on a community velocity model constructed by 2 – 30 s earthquake data. Furthermore, a general theoretical framework is developed to calculate sensitivity kernels for multi-component ambient noise cross correlation functions. It is demonstrated that both eastern and northern point-force sources are required to generate horizontal-component synthetics and associated sensitivity kernels based on rotation relations. Similarly, I apply this method and the inversion package to ANAT of 5 – 50 s Love wave EGFs in southern California and construct an improved V_{sh} model. The two obtained V_s models (V_{sv} and V_{sh}) also give potential insights into the complicated structures of crustal radial anisotropy in southern California. Finally, I apply the method of calculating multi-component sensitivity kernels to investigate the Rayleigh wave ellipticity kernel at regional scale. I show that the lateral variations of the 3-D ellipticity kernel in horizontal direction are non-negligible and should be considered in

future tomographic studies using ellipticity observations to image shallow structures.

Contents

Acknowledgements	v
Abstract	ix
Contents	xi
List of Figures	xv
List of Tables	xxii
List of Publications	xxiii
1 Introduction	1
1.1 Research background	1
1.2 Ambient noise adjoint tomography	3
1.3 Thesis objectives	4
1.4 Thesis structure	5
2 Theory of adjoint tomography	7
2.1 Summary	7
2.2 Introduction	8
2.3 Elastic wave Born scattering theory	10
2.3.1 Governing equations	10
2.3.2 Reciprocity and representation theorems	11
2.3.3 Perturbation theory	13
2.3.4 Lippman-Schwinger equation	14
2.3.5 Born approximation	15
2.4 3-D Fréchet kernel based on adjoint method	15
2.4.1 General Seismological Data Functional (GSDF)	16

2.4.2	GSDF adjoint source	17
2.4.3	Sensitivity kernels	18
2.5	Travel time adjoint tomography	20
2.5.1	Kernels for traveltime measurement	20
2.5.2	Event kernel for travel time misfit	21
2.5.3	Total travelttime misfit	22
2.5.4	Optimization method	23
3	Refined seismic structure of southern California by ambient noise adjoint tomography	25
3.1	Summary	25
3.2	Introduction	26
3.3	Ambient noise data in southern California	28
3.4	Background of adjoint tomography	31
3.5	Inversion Procedures	33
3.5.1	SEM mesh	33
3.5.2	Forward simulation	34
3.5.3	Preprocessing	34
3.5.4	Measurement	35
3.5.5	Adjoint simulation	36
3.5.6	Preconditioner and kernel smoothing	37
3.5.7	Updating the model and reducing the seismogram misfit	37
3.6	Results	38
3.6.1	Initial model	38
3.6.2	Updated models	40
3.6.3	Final model	42
3.7	Discussion	48
3.7.1	Misfit analysis	48
3.7.2	Model comparison	49
3.7.3	Inversion strategies	51
3.8	Conclusions	52
3.9	Appendix A: Line search and model update	54
3.10	Supplementary materials	56

4 Sensitivity kernels for multi-component ambient noise data based on adjoint methods, with application to Love-wave adjoint tomography in southern California	83
4.1 Summary	83
4.2 Introduction	84
4.3 Methodology	89
4.3.1 Sensitivity kernels for isotropic elastic models	89
4.3.2 Coordinate system rotation	90
4.3.3 Individual sensitivity kernels and adjoint source	91
4.3.4 Event kernel for ambient-noise measurements	95
4.3.5 Total misfit kernel	95
4.4 Numerical simulations	95
4.4.1 Seismogram comparison	97
4.4.2 Individual misfit kernel	97
4.4.3 Event misfit kernel	101
4.5 Love-wave ANAT in Southern California	101
4.5.1 Data processing procedures	101
4.5.2 Inversion procedures	102
4.5.3 Results	103
4.6 Discussions	107
4.7 Conclusion	109
4.8 Appendix A: Auxiliary matrix	111
5 Numerical investigation of three-dimensional Rayleigh wave ellipticity kernel based on SEM adjoint method	113
5.1 Summary	113
5.2 Introduction	114
5.3 Methodology	117
5.3.1 Amplitude kernel	117
5.3.2 Differential amplitude kernel	118
5.3.3 Rayleigh-wave ellipticity kernel	118
5.4 Numerical results	119
5.4.1 Parameters setting	120

5.4.2	Lamb's problem—a half space model	121
5.4.3	Higher-mode interferences and medium complexities	122
5.4.4	Azimuthally averaged ellipticity kernel	125
5.5	Conclusions and discussions	133

6 Conclusions and discussions 137

6.1	Conclusions	138
6.2	Future works	139

A ANAT- a script driven package for ambient noise adjoint tomography 143

A.1	Workflow	143
A.2	Package structure	144
A.2.1	Inputs	145
A.2.2	Data processing and plot	145
A.2.3	Outputs	145
A.2.4	Driven scripts	146
A.2.5	PBS scripts	146
A.3	Iterative inversion procedures	147
A.3.1	Preparing input files	147
A.3.2	Running forward and adjoint simulations	148
A.3.3	Post-processing and model update	149
A.3.4	Line search for optimal step length	149
A.4	Data processing scripts	150
A.4.1	<i>run_preprocessing.sh</i>	150
A.4.2	<i>process_data.pl</i>	151

References 153

List of Figures

3.1	(a) Map of southern California with topography, bathymetry and active faults. The solid black rectangle outlines the simulation region. The 148 stations used in this tomographic study are show as triangles, out of which 19 are selected for line search shown in red. Faults are indicated by the bold black lines. Labels 1-8 denote the major eight geological provinces with their boundaries delineated by red dash lines: 1. Coastal Ranges; 2. Great Central Valley; 3. Sierra Nevada; 4. Basin and Range; 5. Transverse Ranges; 6. Mojave Desert; 7. Peninsular Ranges; 8. Salton Trough. Geological features labeled in bold white letters as references for subsequent figures: SCR, southern Coast Range; SAF, San Andreas Fault; SJV, San Joaquin Valley; SNB, Sierra Nevada Batholith; WL: Walker Lane; WBR: Western Basin and Range; WTR, CTR, and ETR: western, central and eastern Transverse Range; ECSZ: Eastern California Shear Zone; LAB: Los Angeles Basin; ePRB and wPRB: east and west Peninsular Ranges Batholith; STB: Salton Trough Basin. (b) Locations of cross sections across San Andreas Fault: AA'-II' (red); Peninsular Ranges: (PR)aa'-cc' (dark green); Sierra Nevada: (SN)aa'-cc' (blue); Salton Trough: (ST)aa'-cc' (purple).	29
3.2	Examples of spectra of EGFs between the master station HEC and a number of other stations shown in (b)-(i). The location of the master station HEC is shown as the red star, and other stations are indicated by the blue triangles in (a).	31

3.3	Total misfit reduction over iterations. Different symbols with different colors, as shown in the legend box, are used to represent the different frequency bands used in measuring the total misfit. The solid curve connects the final total misfit reduction over the three inversion stages using different frequency bands: the first (20 – 50 s), the second (20 – 50 and 10 – 20 s), and the third (20 – 50, 10 – 20 and 5 – 10 s).	38
3.4	Histograms of traveltimes misfits obtained for model M16 (a), M17 (b), M19 (c) M21 (d). The green solid bars represent misfits for the initial model M16 , and the red bars in (b-d) are misfits for the updated models. Misfits are measured at all three period bands, 5 – 10 (left), 10 – 20 (middle), and 20 – 50 (right) s. CC_{min} in (a) are chosen as 0.80 for 5 – 10 s, 0.75 for 10 – 20 s, 0.69 for 20 – 50 s to show the misfits for the initial model. All the misfits are measured with $CC_{min} = 0.75$ in (b) and $CC_{min} = 0.80$ in (c) and (d) for comparison. Note that the measurements of 5 – 10 and 10 – 20 s bands in (a) are not used in the stage 1 inversion, and those of 5 – 10 s in (2) are not used in the stage 2 inversion.	39
3.5	Horizontal slices of shear velocity of model M16 (left panel), ln(M17/M16) (middle panel), and ln(M19/M16) (right panel) at 5, 15, 25, and 35 km depths. In the vicinity of the Moho, the slices reveal patches that represent wave speeds from the underlying mantle (or overlying crust).	41
3.6	Horizontal slices of shear velocity of model M21 (left panel), and ln(M21/M16) (right panel) at 10, 15 and 20 km depths.	45
3.7	Same as Figure 3.6 but for 25, 30, and 40 km depths	46
3.8	Cross-sections of shear velocity of model M16 and M21 across the San Andreas Faults with the profiles shown in Figure 1b.	47
3.9	Cross-sections of shear velocity of model M16 (left panel) and M21 (right panel) across the Peninsular Range Batholith (PRaa'-PRcc') with the profiles shown in Figure 1b.	49
3.10	Cross-sections of shear velocity of M16 (left panel) and M21 across the Salton Through (STaa'-STcc') as the profiles shown in Figure 1b.	50

3.11	Comparison of model M16 , M21 , Barak2015 and CVM-S4.26 along profile DD' across SAF and profile PRbb' across PRB. See more results in supplementary materials	53
3.A1	(a-e) Line search results at each iteration, showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves).	55
3.S1	(a-e) Line search results at each iteration showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves);(f) Total misfit reduction over iterations.	57
3.S2	Histograms of traveltimes misfits obtained for model M16 (a), M17 (b), M19 (c) M21 (d). The green solid bars represent misfits for the initial model, and the red bars in (b-d) are misfits for updated models through iterations. Misfits are measured at all three period bands, 5 – 50 (left), 10 – 50 (middle), and 20 – 50 (right) s. CC_{min} in (a) are specially chosen as 0.80 (5 – 50 s), 0.75 (10 – 50 s), 0.69 (20 – 50 s) to show the misfit of the initial model. All the misfits are measured with $CC_{min} = 0.75$ in (b) and $CC_{min} = 0.80$ in (c) and (d) for comparison. Note that 5 – 50 and 10 – 50 s bands of M16 , and 5 – 50 s band of M17 are not used in the inversion.	58
3.S3	Horizontal slices of shear velocity for model M21 (left panel) using broadband data, M21.narrowband (middle panel) using narrowband data, and their differences (right panel) in percentage at 5, 10, 15, and 20 km depths.	59
3.S4	Same as 3.S3 but for 25, 30, 35, 40 km depths.	60
3.S5	(a) Selected ray paths for waveform comparison of some station pairs; (b)-(d) waveforms of station pairs filtered at 5 – 50 (1st column), 10 – 50 (2nd column), and 20 – 50 (3rd column) s. The EGFs are displayed as black solid lines, and SGFs are plotted as green solid lines for M16 and red solid lines for M21 . Traveltimes misfits and uncertainties are also shown with corresponding colors.	61
3.S6	(a)-(d) Volumetric sensitivities for Vs tomographic models plotted at 10, 15, 20, 25 km depths. (e)-(h) Masked volumetric sensitivities determined by a subjective threshold value of $K = 4.0 \times 10^{-13} m^{-3}$	62

3.S7	Model resolution tests. The left panels show the locations of the fast/slow shear velocity anomalies at the depth of 15 km: (a) one with a 5 km radius put at the LA basin and (b) the other with a 10 km radius at north-eastern part region. The right panels are the corresponding Hessian kernels showing the degree of blurring in recovering the model perturbations	63
3.S8	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile AA' .	64
3.S9	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile BB' .	65
3.S10	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile CC' .	66
3.S11	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile DD' .	67
3.S12	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile EE' .	68
3.S13	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile FF' .	69
3.S14	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile GG' .	70
3.S15	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile HH' .	71
3.S16	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile II' .	72
3.S17	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile PRa-PRa' .	73
3.S18	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile PRb-PRb' .	74
3.S19	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile PRc-PRc' .	75
3.S20	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile SNa-SNa' .	76
3.S21	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile SNb1-SNb1'	77
3.S22	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile SNb2-SNb2'	78
3.S23	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile STa-STa' .	79
3.S24	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile STb-STb' .	80
3.S25	Comparison of M16 , M21 , Barak2015 and CVM-S4.26 along profile STc-SNc' .	81
4.1	Definition of azimuth of radial direction at the source (θ) and station (θ') as well as station backazimuth $\phi = \theta' + 180^\circ$ in a flat (a) and spherical earth (b); For the flat earth case, $\theta' = \theta = \phi - 180^\circ$	86

4.6	(a-c) Line search results at first (1stIT), third (3stIT), and fifth (5thIT) iteration showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves);(d) Total misfit reduction over iterations.	104
4.7	Travetime misfit histograms obtained for model M16 (a), M21 (b). The green solid bars represent misfits of the initial model, and the red bars in (b) are misfits of the final model. Misfits are measured at all three period bands, 5 – 50 (left), 10 – 50 (middle), and 20 – 50 (right) s. CC_{min} in (a) are specially chosen as 0.80 (5 – 50 s), 0.75 (10 – 50 s), 0.75 (20 – 50 s) to show the misfit of the initial model. All the misfits are measured with $CC_{min} = 0.80$ in (b) for comparison.	105
4.8	Horizontal slices of shear velocity for model M16 (left panel), M21 (middle panel), and In(M21/M16) (right panel) at 5, 15, 25, and 35 km depths.	106
4.9	Horizontal slices of V_{sv} from Rayleigh wave (left panel), V_{sh} from Love waves (middle panel), and radial anisotropy (right panel) at 5, 15, 25, and 35 km depths.	108
4.A1	Individual sensitivity kernels of $K^{ET}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (left panels), $K^{NT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (middle panels), and $K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (right panels) for station pair CI.STC-TA.U11A at the depths of 5 km (a), 15 km (b), 25 km (c) and 35 km (d).	112
5.1	Three test models used in numerical simulations. (a) a half space model (M01); (b) a 1-D layered southern California model named Socal1d (M02); (c) a 3-D southern California model named Socal3d (M03). α , β , ρ denote compression-wave, shear-wave velocity, and density. M03 is derived from the 3-D model of Tape et al. [2010].	115
5.2	Seismograms of ZZ and ZR components and their particle motions filtered at (a) 5-10 s; (b) 10-20 s; (c) 20-50 s.	120
5.3	Sensitivity amplitude kernels of ZZ (first column) and ZR component (second column) and their differences (third column) at different depths filtered at 10-20 s in a half space model.	123

5.4	(a) Same as the third column of Figure 5.3 but at different depths; (b) vertical cross sections of the fundamental mode Rayleigh wave ellipticity kernel for P-wave velocity (alpha), S-wave velocity (beta) and density (rho) filtered at 10-20 s.	124
5.5	Sensitivity amplitude kernels and their differences in a 1-D layered model.	126
5.6	Same as Figure 5.4 but in a 1-D layered Socal1d model	127
5.7	Sensitivity amplitude kernels and their differences in a 3-D heterogeneous model	128
5.8	Same as Figure 5.4 but in a 3-D heterogeneous Socal3d model	129
5.9	(a) horizontal slices of azimuthally averaged ellipticity kernel at 5, 10, 20 and 30 km depths; (b) vertical cross sections of azimuthally averaged ellipticity kernel with respect to compression-wave (alpha), density (rho), and shear-wave velocity (beta). The 3-D model is the homogeneous half space model M01.	130
5.10	Same as Figure 5.9 but for M02.	131
5.11	Same as Figure 5.9 but for M03.	132
5.12	Integrated kernels for density (black), P-wave (red), and S-wave (blue). The radius of the cylinder area for integration is shown at the top of each sub-figure.	134
5.S1	Same as Figure 5.4 but at a larger distance.	136
A.1	Workflow of ANAT package.	144
A.2	Structure of ANAT package.	147
A.3	Command line of <i>run_preprocessing.sh</i>	150
A.4	Command line of job1.sh in seis_process_M01.set1	151
A.5	Command line of job1.sh in measure_adj_M01.set1.	152

List of Tables

3.1 Values of inversion parameters used over iterations. T is the period bands for filtered data used in each iteration; dT , $d \ln A$, CC_{min} are timeshifts, amplitude differences, and cross-correlation coefficients criteria for window selection; the two $\sigma_{(h,v)}$ are the horizontal and vertical radius of the Gaussian function used to smooth the misfit gradient; α represents the optimal step length chosen at each iteration.	36
--	----

List of Publications

- **Chapter 3.** Wang K., Y. Yang, P. Basini, P. Tong, C. Tape, and Q. Liu. *Refined crustal and uppermost mantle structure of southern California by ambient noise adjoint tomography*. Geophysical Journal International, **215**(2), 844-863 (2018).

Thesis Author Contribution: Concept (50 %); Data Collection (100 %); Analysis (100 %); Writing (90 %)

- **Chapter 4.** Wang K., Q. Liu, and Y. Yang. *Sensitivity kernels for multi-component ambient noise empirical Green's functions based on adjoint methods*. (GJI, under revision)

Thesis Author Contribution: Concept (90 %); Data Collection (100 %); Analysis (100 %); Writing (80 %)

- **Chapter 5.** Wang K., Y. Yang, and Q. Liu. *Numerical investigation of Rayleigh wave ellipticity kernel based on adjoint method*. (in preparation)

Thesis Author Contribution: Concept (100 %); Data Collection (n/a); Analysis (100 %); Writing (100 %)

Publications not included in this thesis:

- Guo Z., K. Wang, Y. Yang, Y. Tang, Y. Chen, and S. Hung. *The origin and mantle dynamics of Quaternary intra-plate volcanism in Northeast China from joint inversion of surface wave and body wave*. Journal of Geophysical Research: Solid Earth, **123**(3), 2410-2425 (2018).
- Wang K., Y. Luo, Y. Yang. *Correction of phase velocity bias caused by strong directional noise sources in high-frequency ambient noise tomography: a case study in*

Karamay, China. Geophysical Journal International, **205**(2), 715-727 (2016).

- Luo Y., Y. Yang, Y. Xu, H. Xu, K. Zhao, K. Wang. *On the limitations of interstation distances in ambient noise tomography.* Geophysical Journal International, **201**(2), 652-661 (2015).
- Wang K., Y. Luo, Y. Yang. *Body waves revealed by spatial stacking on long-term cross-correlation of ambient noise.* Journal of Earth Science, **25**(6), 977-984 (2014).
- Wang K., Y. Luo, H. Li. *The nature of ambient noise over a field in western Junggar basin near Karamay, China.* Proceedings of the 6th International Conference on Environmental and Engineering Geophysics, 92-97 (2014).

1

Introduction

1.1 Research background

Ambient noise tomography (ANT) has been widely used to image the crustal and upper-mantle structures of the Earth across the globe [e.g., Bensen et al., 2007, Lin et al., 2007, Shapiro et al., 2005, Yang et al., 2007, Yao et al., 2006]. The ANT technique mainly consists of two parts, seismic interferometry and seismic tomography. Seismic interferometry is based on the assumption that the Green's function between any two points can be retrieved from their cross correlation when the seismic noise field is diffuse [Lobkis and Weaver, 2001] and noise sources are evenly distributed [Roux et al., 2005, Snieder, 2004]. The most popular application of seismic interferometry is the direct-wave interferometry [Wapenaar et al., 2010], which aims to retrieve surface wave empirical Green's functions (EGFs) and then invert for shear-wave (V_s) velocities of the subsurface. Another type of seismic interferometry named reflected-wave interferometry [Wapenaar et al., 2010], is utilized to recover the reflected body wave signals in order to image the discontinuities below the subsurface

[e.g., Draganov et al., 2009, Lin et al., 2013, Nakata et al., 2015, Poli et al., 2012a,b, Wang et al., 2014b]. This thesis mainly focuses on developing tomographic method for surface wave EGFs from ambient noise.

Traditional ANT adopts a two-step surface wave tomographic method that usually relies on ray theory. First, period-dependent Rayleigh-wave phase velocity maps are constructed by inverting inter-station dispersion curves measured from cross correlation functions; Second, these phase velocity maps are used to produce a series of local dispersion curves on geographic grids, which are further inverted for local V_s velocities based on 1-D depth sensitivity kernels. The resulting 1-D profiles from all inversion grid points are assembled to form the 3-D V_s model. The innovation of ANT is the ability of interferometry technique in retrieving EGFs from random noise instead of earthquakes. Compared with traditional earthquake-based tomography, ANT is not affected by distributions of earthquakes, and can be conveniently applied to any seismic array where seismic noise is readily recorded stations. In this way, the resolution of ANT mostly depends on station distribution and density. These advantages allow ANT to construct crustal and upper mantle structures with unprecedented resolution [e.g., Lin et al., 2009, Shapiro et al., 2005, Shen et al., 2013].

However, the accuracy of ANT is still hotly debated and attracts increasing amounts of attentions lately [e.g., Basini et al., 2013, Fichtner, 2014, Fichtner et al., 2016, Tromp et al., 2010, Wang et al., 2016, Yao and Van Der Hilst, 2009]. First, the theoretical assumption of evenly distributed noise sources in seismic interferometry is normally hard to be met strictly in the real world [e.g., Stehly et al., 2006, Wang et al., 2014a, Yang and Ritzwoller, 2008, Yao et al., 2009], which could result in potential biases on the measured phase velocities from EGFs. Second, the two-step surface wave tomography is mostly built on the idealized ray theory, which is a high-frequency approximation and ignores the finite-frequency effects that might eventually affect the resulting tomographic model. In addition, the two separated inversion strategy is likely to generate additional ambiguities in the final tomographic result.

This thesis focuses on solving the second problem by developing tomographic method to advance the traditional ANT with 3-D finite-frequency sensitivity kernels. Direct inversion of surface wave inter-station dispersions for V_s velocities without constructing phase/group velocity maps has been studied both with ray-based method [e.g., An et al., 2009, Boschi and Ekström, 2002, Fang et al., 2015, Pilz et al., 2012] and analytical 3-D finite-frequency kernels [e.g., Zhou et al., 2004, 2006]. However, these approaches solve the seismic wave

propagation in a 1-D reference model without considering lateral heterogeneities. According to Zhou et al. [2011], sensitivity kernels in 1-D reference model are close to the ones in 3-D earth models in case of small perturbations. However, large lateral heterogeneities have strong effects on the amplitude and Fresnel zone of the surface wave sensitivity kernel. Based on numerical simulations by SEM adjoint method, Zhou et al. [2011] find that a large-scale anomaly located along the great-circle path may affect the overall amplitude of the sensitivity kernels due to a change in the reference wavelength of the seismic wave; when heterogeneities are located off the reference path, surface wave sensitivity becomes stronger in slow-velocity regions where the Fresnel zone narrows, and weaker in fast-velocity regions where the associated Fresnel zone widens. As a result, a strong heterogeneous crustal structure may have significant effects on short-period surface wave sensitivity kernels. In this study, I solve the tomographic problem with 3-D Fréchet kernels constructed from full wavefield modeling in a heterogeneous Earth model, which is expected to provide more accurate solution although it takes much heavier computation cost.

1.2 Ambient noise adjoint tomography

With the developments of high-performance computing, it has now become practical to simulate seismic wave propagation in complex 3-D Earth models using numerical methods such as spectral-element method (SEM) [e.g., Komatitsch and Vilotte, 1998, Tromp et al., 2008]. Utilizing forward computation capabilities, adjoint tomography based on SEM has been introduced to earthquake seismology [Tromp et al., 2005] and successfully applied to constrain crustal structures of southern California [Tape et al., 2009, 2010], and upper mantle structures of Australia, Europe, and East Asia [Chen et al., 2015, Fichtner et al., 2009, 2010, Zhu et al., 2012, 2015], among others. The basic idea of adjoint tomography is to iteratively minimize a misfit function between observations and synthetics by using 3-D Fréchet derivatives, which are constructed from the interference between a forward wavefield and an adjoint wavefield [Liu and Tromp, 2006]. The use of full numerical solvers for both wavefields provides accurate Fréchet kernels for the current model at each iteration, and allows more data to be used successively as the model improves.

Recently, surface waves extracted from ambient noise have been used in full waveform inversion through scattering-integral method based on finite-difference modeling [e.g., Gao

and Shen, 2014, Lee et al., 2014], and adjoint method based on SEM [e.g., Chen et al., 2014, Liu et al., 2017]. Compared to the traditional ANT, these simulation-based ANT studies demonstrated several advantages, including better waveform fitting, more realistic sensitivity kernels for model updating, and capability of resolving more pronounced velocity variations. However, current ambient-noise adjoint tomography (ANAT) or full waveform ANT studies have only utilized vertical-vertical (Z-Z) component Rayleigh-wave EGFs extracted from vertical-component ambient noise to resolve isotropic structures. No ANAT studies have attempted to use signals from other components, such as Love-wave signals from T-T components. However, the additional T-T and R-R component EGFs in ANAT are complementary to the Z-Z component EGFs and allow the study of radially anisotropic structures. Moreover, adjoint method and SEM can be also used to calculate the 3-D differential amplitude kernels for Rayleigh wave ellipticity. As far as we know, 3-D kernels of Rayleigh wave ellipticity have only been studied by surface mode summation with a 1-D reference model [Muir and Tsai, 2017] and their characteristics in a 3-D heterogeneous media remain to be explored.

1.3 Thesis objectives

The goal of this PhD project is to apply adjoint tomography to surface wave observables extracted from multi-component ambient seismic noise. Specifically, this includes:

1. Developing an iterative inversion package to implement adjoint tomography to ambient noise data based on full wavefield modeling using spectral element method;
2. Applying adjoint tomography to Rayleigh wave EGFs from vertical component ambient noise data at 5 – 50 s in southern California to a) validate the developed inversion package and b) improve the community velocity model that is based on relatively shorter period earthquake data.
3. Developing a general theory of calculating sensitivity kernels for multi-component ambient noise cross correlation functions based on adjoint method, and applying it to adjoint tomography of Love wave EGFs from transverse component ambient noise data at 5 – 50 s in southern California.

4. Performing numerical simulation of 3-D Rayleigh wave ellipticity kernel based on SEM in conjunction with adjoint method to investigate the 3-D geometry of Rayleigh wave ellipticity kernel and the influence on the kernel from both higher-mode interferences and medium heterogeneities.

1.4 Thesis structure

The thesis consists of six chapters, including an introduction of ANAT in chapter 1, the theory of adjoint tomography in chapter 2, the applications of Rayleigh wave and Love-wave ANAT in southern California in chapter 3 and 4 respectively, numerical simulation of 3-D Rayleigh wave ellipticity kernel in chapter 5 and finally a conclusion with future research plans in chapter 6. A simple manual of our ANAT package is attached in Appendix A.

In chapter 2, representation theorem is applied to the perturbed equation of motion to obtain first-order Born scattering theory, which connects a perturbed displacement field with model perturbations and reference displacement field. By introducing General Seismological Data Functional (GSDF) for different types of measurements to the perturbed displacement, the Fréchet kernels of GSDF relative to model perturbations are derived from the interaction between a forward wavefield and an adjoint wavefield based on adjoint method. Using travelttime measurements as an example, I summarize the basic concepts of misfit function and sensitivity kernel used in both classic finite-frequency and adjoint tomography, with a focus on their differences in kernel construction and inversion procedures.

In chapter 3, we apply adjoint tomography to 5-50 s Rayleigh wave EGFs from ambient noise to construct an improved V_{sv} model of Southern California. EGFs from ambient noise are compared with synthetic Green's functions (SGFs) from SEM-based numerical simulations. Travelttime misfits between EGFs and SGFs are measured at three period bands, namely 5-10 s, 10-20 s, and 20-50 s. We start the inversion with only data in the 20 – 50 s band and then progressively introduce data from the other two shorter-period bands. By minimizing the travelttime differences between EGFs and SGFs through local-gradient-based optimization algorithm, we iteratively improve the 3-D tomographic model constructed based on earthquake data. We compare the final model with the initial earthquake-based model as well as two other models and then discuss the improvements and reliability of our model

compared to others. All scripts and programs used in this study are assembled into a script-driven package for implementing ANAT in appendix A.

In chapter 4, we develop a general theoretical framework to calculate the sensitivity kernel for multi-component ambient noise cross-correlation functions. Based on adjoint method, we demonstrate that a horizontal component kernel can be obtained by the summation of two rotated kernels with proper amplitude modulation factors based on rotation relations. Utilizing this method for calculating T-T component sensitivity kernel, I successfully perform Love-wave ANAT in Southern California and obtained an improved V_{sh} model. Combined with the V_{sv} model from Rayleigh wave ANAT in chapter 3, I discuss the features of resolved crustal radial anisotropy in southern California and its relation with tectonic deformation. The method we developed provides the basis for future multi-component ambient noise adjoint tomography to construct high-resolution isotropic and anisotropic models.

In chapter 5, I investigate the characteristics of Rayleigh wave ellipticity kernel using adjoint method and SEM-based simulations. I use three different background models, namely a half space model, a 1-D layered model, and a 3-D heterogeneous model to investigate the 3-D ellipticity kernels for fundamental-mode Rayleigh waves, and the influences on kernel from higher-mode interferences and medium heterogeneities. Moreover, I compare the 3-D ellipticity kernel with 1-D depth sensitivity kernel that is widely used in surface wave tomography today. Finally, I discuss the feasibility of future simulation-based ellipticity tomography.

Chapter 6 draws a conclusion of this thesis. In this chapter, I summarize the main research outcomes and discuss the possible future research directions based on current studies.

2

Theory of adjoint tomography

2.1 Summary

This chapter presents the deviation of finite-frequency sensitivity kernels based on Born scattering theory using adjoint method and the basic inversion theory for travel time adjoint tomography. Based on the equation of motion in linear elasticity, Betti's reciprocity theorem is applied to obtain the representation theorem, which gives the general form of a displacement field due to body force and surface traction in terms of Green's function of point forces and medium properties. On the other hand, by adding small perturbations to medium properties in the equation of motion, we obtained the perturbed equation of motion. Applying the representation theorem to the perturbed equation of motion, we obtain the perturbed displacement field in terms of perturbed medium properties, Green's function in reference medium, and the total displacement field, which is also called Lippman-Schwinger equation. Upon replacing the total displacement field with the one in reference medium, we obtain the Born scattering theory which connects a perturbed displacement field with perturbed medium properties,

Green's function and the unperturbed displacement in a reference medium.

By introducing General Seismological Data Functional (GSDF) for different types of measurements to the perturbed displacement field, the Fréchet kernels of GSDF relative to model perturbations are constructed from the interaction of a forward wavefield and an adjoint wavefield based on adjoint method. Taking traveltime measurements as an example, I summarize the basic concepts of misfit function and sensitivity kernel used in both classic finite-frequency and adjoint tomography, in particular emphasizing their differences in kernel construction and inversion procedures.

2.2 Introduction

Ray theory has been widely adopted in early seismic tomographic studies because of its simplicity and efficiency, and it provides us a primary knowledge of the velocity structure of the Earth [e.g., Dziewonski and Anderson, 1981, Kennett and Engdahl, 1991, Kennett et al., 1995, Masters et al., 2000, Ritzwoller et al., 2002]. It is based on the simple assumption of weak and relatively large-scale lateral heterogeneities, and is a high frequency approximation that seismic waves propagate only along the geometrical ray path between source and receiver. However, this assumption becomes less valid when the wavelength is comparable or larger than structural heterogeneities. In other words, ray theory can not explain phenomena such as wave-front healing and other finite-frequency diffraction effects. To overcome such limitations, Born scattering theory [e.g., Snieder, 2002, Wu, 1989] was developed by solving perturbed seismic wave equations for a single scattering problem where each perturbed point is only scattered once in the reference wavefield. Using Born scattering theory, finite-frequency sensitivity kernels can be calculated based on no-linear asymptotic coupling theory [e.g., Li and Romanowicz, 1995], scalar exponent approximation [e.g., Marquering et al., 1999], paraxial ray approximation [e.g., Dahlen et al., 2000], surface wave summation [e.g., Zhou et al., 2004], and normal mode summation [e.g., Capdeville, 2005, Zhao and Jordan, 2006, Zhao et al., 2000]. In general, a 3-D finite-frequency kernel has volumetric sensitivities in the vicinity of the central ray path due to constructively interference based on single scattering approximation, instead of only along the infinite ray path.

In the classic banana-doughnut theory, Dahlen et al. [2000] adopt a paraxial ray approximation to account for finite-frequency effects for body waves, and they found that a 3-D

sensitivity kernel of body wave phases is a hollow banana with zero sensitivities along the ray path, and a doughnut in a cross section perpendicular to the ray. Following this, finite-frequency theory for surface wave observables was also developed by surface wave mode summation [Zhou et al., 2004]. Different from body waves, a 3-D finite frequency kernel for minor-arc surface waves has a strong-to-weak-to-strong pattern along the geometrical ray, and the horizontal plane shows alternating positive and negative side bands with an elliptical shape across the central ray. A Major-arc and multi-orbit surface wave kernel resembles a sausage link, with pinches at the receiver and source antipodes. As finite-frequency effects become more evident at longer period, finite frequency tomography has been applied to long period body and surface waves to obtain fine-scale mantle structure at both global [e.g., Montelli et al., 2006, 2004, Zhou et al., 2006] and regional scales [e.g., Hung et al., 2011, 2010, 2004, Obrebski et al., 2011, Tang et al., 2014, Yang and Forsyth, 2006]. These finite-frequency images show its advantage towards improving resolution of tomography, which can help to reveal small scale structures such as the shape of possible mantle plumes [Montelli et al., 2006], and mantel convections [Tang et al., 2014].

Nevertheless, the classic finite-frequency theory is still based on analytical methods using 1-D reference models, which means that lateral heterogeneities are not considered. However, the interior of the Earth is strongly heterogeneous, arising debates on the accuracy of classic finite-frequency theory. For example, De Hoop and van Der Hilst [2005] argued that the zero sensitivities of the banana-doughnut kernels do not hold for realistic media. To accurately describe the wavefield propagating in a 3-D heterogeneous medium, one needs to refer to numerical methods such as finite-difference [e.g., Olsen et al., 1995], pseudo-spectral [e.g., Carcione, 1994], finite-element [e.g., Akcelik et al., 2002], and spectral element method (SEM) [e.g., Komatitsch and Tromp, 1999, Tromp et al., 2008].

Based on Born scattering theory, there are generally two ways to calculate finite-frequency kernels using numerical solver: one is the scattering integral (SI) method and the other is adjoint method. Zhao et al. [2005] was the first to use the SI method to calculate finite-frequency kernels based on finite-difference method in a 3-D reference model. The SI method was then extended by Chen et al. [2010] to derive the perturbation kernels for general seismological data functional. Tromp et al. [2005] firstly introduced the adjoint method into earthquake seismology. Afterwards, finite-frequency sensitivity kernels in a 3-D heterogeneous medium using adjoint method based on SEM were developed in both isotropic [Liu

and Tromp, 2006, 2008] and anisotropic [Sieminski et al., 2007a,b] medium.

Adjoint tomography has been applied to constrain crustal structure of southern California [Tape et al., 2009, 2010], and upper mantle structure of Australia, Europe, and East Asia [Chen et al., 2015, Fichtner et al., 2009, 2010, Zhu et al., 2012, 2015], among others. More recently, surface waves from ambient noise cross-correlation functions have also been used in full waveform inversion through either scattering-integral method based on finite-difference modeling [e.g., Lee et al., 2014], or adjoint tomography based on SEM [e.g., Chen et al., 2014, Wang et al., 2018b]. Compared to the traditional ANT, these simulation-based ANT studies demonstrate several advantages, including better waveform fitting, more realistic sensitivity kernels for 3D structural updates, and resolving more pronounced velocity variations. Some recent review articles about the advances in seismic tomography can be found in Liu and Gu [2012], Rawlinson et al. [2010], Rawlinson and Sambridge [2003].

This chapter is organized as follows: in section 2, we derive the first-order Born scattering theory in the time domain, which is the foundation of finite-frequency theory; Then, based on the Born scattering theory, we calculate finite-frequency sensitivity kernels for general seismological data functionals using the adjoint method in section 3.

2.3 Elastic wave Born scattering theory

In the following, I will derive the Born scattering theory step by step, from the equation of motion to representation theorem, from perturbation theory to Lippman-Schwinger equation. Such kind of derivation can be also found in Snieder [2002], [chapter 1.7.1 Pike and Sabatier, 2001] and [chapter 3.1 Schumacher, 2014]. However, their derivation is in the frequency domain. The purpose of this section is to reproduce the derivation in the time domain in terms of the theoretical framework established in chapter 2.3.2 by Aki and Richards [2002], with the same symbols and notations.

2.3.1 Governing equations

The Newton's second law states that the rate of change of momentum of a body is directly proportional to the force applied, that is

$$\frac{\partial}{\partial t} \iiint_v \rho \frac{\partial \mathbf{u}}{\partial t} dV = \iiint_v \mathbf{f} dV + \iint_S \mathbf{T}(\mathbf{n}) dS \quad (2.1)$$

where, \mathbf{u} is the displacement field, \mathbf{f} is external body force, and \mathbf{T} is surface traction with a unit normal \mathbf{n} to the surface S . ρ , t , V are density, time, and object volume. The surface integral of traction can be converted to volume integral of stress tensor by *Gauss's divergence theorem*:

$$\iint_S T_i dS = \iint_S \tau_{ij} n_j dS = \iiint_V \partial_j \tau_{ij} dV \quad (2.2)$$

Substituting eq. (2.2) into eq. (2.1), we have the **equation of motion**:

$$\rho \ddot{u}_i = \nabla \tau_{ij} + f_i \quad (2.3)$$

In linear elasticity, we have the **strain-displacement** equation:

$$e_{ij} = \frac{1}{2} (\partial_j u_i + \partial_i u_j) \quad (2.4)$$

and the **constitute equation**:

$$\tau_{ij} = c_{ijkl} \partial_l u_k \quad (2.5)$$

where τ_{ij} is the stress tensor, e_{ij} is the strain tensor, and c_{ijkl} is fourth-order elastic tensor. Eqs. (2.3)-(2.5) are the three governing equations of Born scattering theory.

2.3.2 Reciprocity and representation theorems

First, we reproduce the reciprocity and representation theorems in chapter 2 of Aki and Richards [2002]. Let's rewrite the equation of motion:

$$\rho \ddot{u}_i = \partial_j (c_{ijkl} \partial_l u_k) + f_i \quad (2.6)$$

Betti's reciprocity theorem

Assume that $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ is the displacement due to body forces \mathbf{f} and boundary conditions on S and $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ is another displacement due to body forces \mathbf{g} . We use the notation $\mathbf{T}(\mathbf{u}, \mathbf{n})$ for the traction due to \mathbf{u} and $\mathbf{T}(\mathbf{v}, \mathbf{n})$ for the traction due to \mathbf{v} . Then, Betti's theorem states

$$\iiint_V (\mathbf{f} - \rho \ddot{\mathbf{u}}) \cdot \mathbf{v} dV + \iint_S \mathbf{T}(\mathbf{u}, \mathbf{n}) \cdot \mathbf{v} dS = \iiint_V (\mathbf{g} - \rho \ddot{\mathbf{v}}) \cdot \mathbf{u} dV + \iint_S \mathbf{T}(\mathbf{v}, \mathbf{n}) \cdot \mathbf{u} dS \quad (2.7)$$

To proof Betti's theorem, substitute eq. (2.6) into LHS of eq. (2.7)

$$\begin{aligned} LHS &= \iiint_V (f_i - \rho \ddot{u}_i) v_i dV + \iint_S \tau_{ij} n_j v_i dS \\ &= - \iiint_V \partial_j (c_{ijkl} \partial_l u_k) v_i dV + \iiint_V \partial_j (\tau_{ij} v_i) dV \\ &= - \iiint_V \partial_j (c_{ijkl} \partial_l u_k) v_i dV + \iiint_V \partial_j (c_{ijkl} \partial_l u_k v_i) dV = \iiint_V (\partial_j v_i) c_{ijkl} \partial_l u_k dV \end{aligned} \quad (2.8)$$

where *Gauss's divergence theorem* is used in the first equality, and *three-dimension integral by part* is used in the last equality. Similarly, we have

$$RHS = \iiint_v (\partial_j u_i) c_{ijkl} \partial_l v_k dV = \iiint_v (\partial_l u_k) c_{klji} \partial_j v_i dV = LHS \quad (2.9)$$

where we exchange the index pair (i, j) with (k, l) and use the symmetry $c_{ijkl} = c_{klji}$.

For displacement fields with a quiescent past [chapter 2.3.2, Aki and Richards, 2002], we have

$$\int_{-\infty}^{+\infty} \rho \{ \ddot{\mathbf{u}}(t) \cdot \mathbf{v}(\tau - t) - \mathbf{u}(t) \cdot \ddot{\mathbf{v}}(\tau - t) \} dt = 0 \quad (2.10)$$

Substituting $\mathbf{u}(\mathbf{x}, t)$ and $\mathbf{v}(\mathbf{x}, \tau - t)$ into eq. (2.7) using eq. (2.10), we obtain another form of Betti's theorem with a quiescent past:

$$\begin{aligned} & \int_{-\infty}^{+\infty} dt \iiint_v \{ \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{g}(\mathbf{x}, \tau - t) - \mathbf{v}(\mathbf{x}, \tau - t) \cdot \mathbf{f}(\mathbf{x}, t) \} dV(\mathbf{x}) \\ &= \int_{-\infty}^{+\infty} dt \iint_S \{ \mathbf{v}(\mathbf{x}, \tau - t) \cdot \mathbf{T}(\mathbf{u}(\mathbf{x}, t), \mathbf{n}) - \mathbf{u}(\mathbf{x}, t) \cdot \mathbf{T}(\mathbf{v}(\mathbf{x}, \tau - t), \mathbf{n}) \} dS(\mathbf{x}) \end{aligned} \quad (2.11)$$

Representation theorem

If the integrated form of Betti's theorem, our equation (2.11), is used with a Green function for one of the displacement fields, then a representation for the other displacement field becomes available.

In elasticity theory, the *Green's function* gives the displacement generated by a unit impulse in a certain direction. if a unit force in the n direction at location $\mathbf{x} = \mathbf{r}$ and time $t = \tau$ is applied, the i th component of displacement at any point (\mathbf{x}, t) is denoted by $G_{in}(\mathbf{x}, t; \mathbf{r}, \tau)$ and it satisfies

$$\rho \ddot{G}_{in} = \partial_j (c_{ijkl} \partial_l G_{kn}) + \delta_{in} \delta(\mathbf{x} - \mathbf{r}) \delta(t - \tau) \quad (2.12)$$

We substitute into eq. (2.11) with the body force $g_i(\mathbf{x}, t) = \delta_{in} \delta(\mathbf{x} - \mathbf{r}) \delta(t)$, for which the corresponding solution is $v_i(\mathbf{x}, t) = G_{in}(\mathbf{x}, t; \mathbf{r}, 0)$, and find

$$\begin{aligned} u_n(\mathbf{r}, \tau) &= \int_{-\infty}^{+\infty} dt \iiint_v f_i(\mathbf{x}, t) G_{in}(\mathbf{x}, \tau - t; \mathbf{r}, 0) dV(\mathbf{x}) \\ &+ \int_{-\infty}^{+\infty} dt \iint_S \{ G_{in}(\mathbf{x}, \tau - t; \mathbf{r}, 0) c_{ijkl} \partial_l u_k(\mathbf{x}, t) n_j \\ &- u_i(\mathbf{x}, t) c_{ijkl} \partial_l [G_{kn}(\mathbf{x}, \tau - t; \mathbf{r}, 0)] n_j \} dS(\mathbf{x}) \end{aligned} \quad (2.13)$$

In above, we have used the traction-stress relation $T_i = \tau_{ij} n_j$ and constitute equation eq.

(2.5). Upon interchanging the symbols \mathbf{x} and \mathbf{r} , and t and τ , we have

$$\begin{aligned} u_n(\mathbf{x}, t) &= \int_{-\infty}^{+\infty} d\tau \iiint_V f_i(\mathbf{r}, \tau) G_{in}(\mathbf{r}, t - \tau; \mathbf{x}, 0) dV(\mathbf{r}) \\ &\quad + \int_{-\infty}^{+\infty} d\tau \iint_S \{ G_{in}(\mathbf{r}, t - \tau; \mathbf{x}, 0) c_{ijkl} \partial_l u_k(\mathbf{r}, \tau) n_j \\ &\quad - u_i(\mathbf{r}, \tau) c_{ijkl} \partial_l [G_{kn}(\mathbf{r}, t - \tau; \mathbf{x}, 0)] n_j \} dS(\mathbf{r}) \end{aligned} \quad (2.14)$$

It states that the displacement \mathbf{u} at a certain point \mathbf{x} is made up from contributions due to the force \mathbf{f} throughout V , plus contributions due to the traction $\mathbf{T}(\mathbf{u}, \mathbf{n})$ and to the displacement \mathbf{u} itself on S . However, the way in which each of these three contributions is weighted is unsatisfactory, since each involves a Green's function with source at \mathbf{x} and observation point at \mathbf{r} [Page 29, Aki and Richards, 2002].

2.3.3 Perturbation theory

Then, we introduce the perturbation theory in the time domain following Herrera and Mal [1965], Wu [1989]. The corresponding deviations in the frequency domain are almost the same, and can be found in Snieder [2002] and Schumacher [2014].

In the perturbation method, a heterogeneous medium is decomposed into a reference medium and its perturbations. In the following, the properties of the reference model will always be referred to by a superscript 0, e.g. ρ^0 , c_{ijkl}^0 , \mathbf{u}^0 , etc. and the perturbations by a prefix δ , e.g. $\delta\rho$, δc_{ijkl} , $\delta\mathbf{u}$, etc. Thus, the perturbed properties may be written as

$$\begin{aligned} \rho &= \rho^0 + \delta\rho \\ c_{ijkl} &= c_{ijkl}^0 + \delta c_{ijkl} \end{aligned} \quad (2.15)$$

Let's define the vector differential operator \mathbf{L} as

$$[\mathbf{L}(\mathbf{u}, \mathbf{m})]_i = \rho \ddot{u}_i - \partial_j (c_{ijkl} \partial_l u_k) \quad (2.16)$$

Substituting eq. (2.15) into eq. (2.16), and considering a homogeneous boundary condition $\mathbf{L}(\mathbf{u}, \mathbf{m}) = \mathbf{f} = 0$ [Box 2.4, Aki and Richards, 2002], we have

$$\rho^0 \ddot{u}_i - \partial_j c_{ijkl}^0 \partial_k u_l = - (\delta\rho \ddot{u}_i - \partial_j \delta c_{ijkl} \partial_k u_l) \quad (2.17)$$

Similarly to eq. (2.15), the total displacement can be decomposed into a reference displacement and a perturbed displacement:

$$\mathbf{u} = \mathbf{u}^0 + \delta\mathbf{u} \quad (2.18)$$

where the reference displacement \mathbf{u}^0 is such that

$$[\mathbf{L}(\mathbf{u}^0, \mathbf{m}^0)]_i = \rho^0 \ddot{u}_i^0 - \partial_j (c_{ijkl}^0 \partial_l u_k^0) = 0 \quad (2.19)$$

Subtracting eq. (2.19) from eq. (2.17) using eq. (2.18), we have the *perturbed equation of motion*

$$\rho^0 \delta \ddot{u}_i - \partial_j c_{ijkl}^0 \partial_k \delta u_l = - (\delta \rho \ddot{u}_i - \partial_j \delta c_{ijkl} \partial_k u_l) = Q_i \quad (2.20)$$

where $Q_i(\mathbf{x}, t)$ represents the equivalent body force due to the interaction of the heterogeneities with the wavefield.

2.3.4 Lippman-Schwinger equation

Applying the representation theorem eq. (2.14) to *perturbed equation of motion*, we have

$$\begin{aligned} \delta u_n(\mathbf{x}, t) &= \int_{-\infty}^{+\infty} d\tau \iiint_v Q_i(\mathbf{r}, \tau) G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}, 0) dV(\mathbf{r}) \\ &\quad + \int_{-\infty}^{+\infty} d\tau \iint_S \left\{ G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}, 0) c_{ijkl}^0 \partial_l \delta u_k(\mathbf{r}, \tau) n_j \right. \\ &\quad \left. - \delta u_i(\mathbf{r}, \tau) c_{ijkl}^0 \partial_l [G_{kn}^0(\mathbf{r}, t - \tau; \mathbf{x}, 0)] n_j \right\} dS(\mathbf{r}) \end{aligned} \quad (2.21)$$

Let's first check the volume and surface integral over \mathbf{r} . The volume integral may be expanded as

$$\begin{aligned} \iiint_v Q_i(\mathbf{r}) G_{in}^0(\mathbf{r}, \mathbf{x}) dV &= - \iiint_v \left\{ \delta \rho \ddot{u}_i(\mathbf{r}) - \partial_j [\delta c_{ijkl} \partial_l u_k(\mathbf{r})] \right\} G_{in}^0(\mathbf{r}, \mathbf{x}) dV \\ &= - \iiint_v \left\{ \delta \rho \ddot{u}_i(\mathbf{r}) G_{in}^0(\mathbf{r}, \mathbf{x}) + [\partial_j G_{in}^0(\mathbf{r}, \mathbf{x})] [\delta c_{ijkl} \partial_l u_k(\mathbf{r})] \right\} dV \\ &\quad + \iiint_v \left\{ \partial_j [G_{in}^0(\mathbf{r}, \mathbf{x}) \delta c_{ijkl} \partial_l u_k(\mathbf{r})] \right\} dV \\ &= - \iiint_v \left\{ \delta \rho \ddot{u}_i(\mathbf{r}) G_{in}^0(\mathbf{r}, \mathbf{x}) + [\partial_j G_{in}^0(\mathbf{r}, \mathbf{x})] [\delta c_{ijkl} \partial_l u_k(\mathbf{r})] \right\} dV \\ &\quad + \iint_S [G_{in}^0(\mathbf{r}, \mathbf{x}) \delta c_{ijkl} \partial_l u_k(\mathbf{r})] n_j dS \end{aligned} \quad (2.22)$$

where *three-dimension integral by part* (eq. 2.8) and *Gauss's divergence theorem* (eq. 2.2) are used in the second and last equality, respectively. This adds another surface integral to eq. (2.21), which may be combined with the first surface integral

$$\begin{aligned} &\iint_S [G_{in}^0(\mathbf{r}, \mathbf{x}) \delta c_{ijkl} \partial_l u_k(\mathbf{r})] n_j dS + \iint_S [G_{in}^0(\mathbf{r}, \mathbf{x}) c_{ijkl}^0 \partial_l \delta u_k(\mathbf{r})] n_j dS \\ &= \iint_S \left\{ G_{in}^0(\mathbf{r}, \mathbf{x}) \left[\delta c_{ijkl} \partial_l u_k(\mathbf{r}) + c_{ijkl}^0 \partial_l \delta u_k(\mathbf{r}) \right] n_j \right\} dS \end{aligned} \quad (2.23)$$

As it is assumed that the boundaries are not perturbed, the above surface integral is zero [Schumacher, 2014]. We assume that the reference medium has zero tractions at the surface:

$$c_{ijkl}^0 \partial_l [G_{kn}^0(\mathbf{r}, t - \tau; \mathbf{x}, 0)] n_j = 0 \quad (2.24)$$

Thus, the last surface integral term of eq. (2.21) vanishes. Finally, adding the time integral, eq. (2.21) becomes to

$$\delta u_n(\mathbf{x}, t) = - \int_{-\infty}^{+\infty} d\tau \iiint_v \{ \delta \rho \ddot{u}_i(\mathbf{r}, \tau) G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}) + \partial_j G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}) \delta c_{ijkl} \partial_l u_k(\mathbf{r}, \tau) \} d\mathbf{r}^3 \quad (2.25)$$

It states that the perturbed displacement $\delta \mathbf{u}$ at a certain point \mathbf{x} is made up from contributions of scattered waves $u_i(\mathbf{r}, \tau)$ and $G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x})$ at any scatter \mathbf{r} generated by the perturbations of the medium.

2.3.5 Born approximation

In the right hand side of eq. (2.25), replacing the total wave field with the unperturbed wave gives the Born-approximation scattered wave

$$\delta u_n(\mathbf{x}, t) = - \int_{-\infty}^{+\infty} d\tau \iiint_v \{ \delta \rho \ddot{u}_i^0(\mathbf{r}, \tau) G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}) + \partial_j G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}) \delta c_{ijkl} \partial_l u_k^0(\mathbf{r}, \tau) \} d\mathbf{r}^3 \quad (2.26)$$

Considering a receiver at $\mathbf{x} = \mathbf{x}_r$ and a source at $\mathbf{x} = \mathbf{x}_s$, the Born-approximation scattered wave for perturbed displacement of the source-receiver pair is

$$\begin{aligned} \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s) = & - \int_0^t \iiint_v \{ \delta \rho(\mathbf{r}) \ddot{u}_i^0(\mathbf{r}, \tau; \mathbf{x}_s) G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}_r) \\ & + \partial_j G_{in}^0(\mathbf{r}, t - \tau; \mathbf{x}_r) \delta c_{ijkl}(\mathbf{r}) \partial_k u_l^0(\mathbf{r}, \tau; \mathbf{x}_s) \} d\mathbf{r}^3 d\tau \end{aligned} \quad (2.27)$$

It states that the perturbed displacement $\delta \mathbf{u}(\mathbf{x}_r, t; \mathbf{x}_s)$ from source \mathbf{x}_s to receiver \mathbf{x}_r is made of scattered waves $u_i^0(\mathbf{r}, \tau; \mathbf{x}_s)$ from source to scatter \mathbf{r} and $G_{in}^0(\mathbf{r}, t - \tau, \mathbf{x}_r)$ from scatter to receiver in a reference medium, with contributions from the perturbations of the medium.

2.4 3-D Fréchet kernel based on adjoint method

The use of adjoint methods in seismology was advocated in the 1980s [Gauthier et al., 1986, Tarantola, 1984] and expanded in the 2000s following the advent of high-performance computing and the development of finite-frequency theory [e.g., Liu and Tromp, 2006, Luo,

2012, Tape et al., 2007, Tromp et al., 2005]. Following Tromp et al. [2005], we apply Born scattering theory to obtain the 3-D Fréchet kernel based on adjoint method.

2.4.1 General Seismological Data Functional (GSDF)

Let's consider the General Seismological Data Functional (GSDF) introduced by Gee and Jordan [1992] and extended by Chen et al. [2010]. The seismogram observed on the n th component of the r th receiver from the s th source is represented by $u_n(\mathbf{x}_r, t; \mathbf{x}_s)$. For each seismogram, we define a data functionals $\phi_n(\mathbf{x}_r, t; \mathbf{x}_s)$ that measure the misfit between observed $u_n^{obs}(\mathbf{x}_r, t; \mathbf{x}_s)$ and synthetic wavefield $u_n^{syn}(\mathbf{x}_r, t; \mathbf{x}_s)$ as:

$$\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) = D [u_n^{obs}(\mathbf{x}_r, t; \mathbf{x}_s), u_n^{syn}(\mathbf{x}_r, t; \mathbf{x}_s)] , \quad (2.28)$$

where D is the function that quantifies the misfit. As the data functional is seismogram specific, the Fréchet kernel of a data functional with respect to displacement perturbation can be expressed in terms of linear integration of space-independent kernels:

$$\delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) = \int_0^T J_n(\mathbf{x}_r, t; \mathbf{x}_s) \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s) dt \quad (2.29)$$

where $\delta\phi_n(\mathbf{x}, t)$, $\delta u_n(\mathbf{x}_r, t; \mathbf{x}_s)$ represent perturbations in the data functional and displacement wavefield, respectively. $J_n(\mathbf{x}, t)$ is the so-called *seismogram perturbation kernel*, which maps a perturbation in displacement to a perturbation in the data functional. It accounts for the effects of different types of measurement operators on the waveforms as well as any instrument filtering. Here is four widely used data functionals as shown in Chen et al. [2007a]

1. waveform differences.

$$\delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) = \int_0^T [u_n^{syn}(\mathbf{x}_r, t; \mathbf{x}_s) - u_n^{obs}(\mathbf{x}_r, t; \mathbf{x}_s)] \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s) dt \quad (2.30)$$

$$\delta J_n(\mathbf{x}_r, t; \mathbf{x}_s) = u_n^{syn}(\mathbf{x}_r, t; \mathbf{x}_s) - u_n^{obs}(\mathbf{x}_r, t; \mathbf{x}_s) \quad (2.31)$$

2. Traveltime differences Dahlen et al. [2000].

$$\delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) = \delta T_n(\mathbf{x}_r, t; \mathbf{x}_s) = \int_0^T \frac{\partial_t u_n(\mathbf{x}_r, t; \mathbf{x}_s) \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, t; \mathbf{x}_s)|^2 dt} dt \quad (2.32)$$

$$\delta J_n(\mathbf{x}_r, t; \mathbf{x}_s) = \frac{\partial_t u_n(\mathbf{x}_r, t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, t; \mathbf{x}_s)|^2 dt} \quad (2.33)$$

3. Amplitude differences Dahlen and Baig [2002].

$$\delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) = \delta \ln A_n(\mathbf{x}_r, t; \mathbf{x}_s) == \int_0^T \frac{u_n(\mathbf{x}_r, t; \mathbf{x}_s)\delta u_n(\mathbf{x}_r, t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, t; \mathbf{x}_s)|^2 dt} dt \quad (2.34)$$

$$\delta J_n(\mathbf{x}_r, t; \mathbf{x}_s) = \frac{u_n(\mathbf{x}_r, t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, t; \mathbf{x}_s)|^2 dt} \quad (2.35)$$

4. GSDF frequency-dependent measurements. For a given frequency ω , the synthetic displacement waveform ($u(\omega)$) is mapped into the observed displacement ($u^{obs}(\omega)$) waveform through

$$u(\omega) = u^{obs}(\omega)e^{-i\omega[\delta\tau_p(\omega)-i\delta\tau_q(\omega)]}, \quad (2.36)$$

where the frequency-dependent phase-delay time $\delta\tau_p(\omega)$ and amplitude-reduction time $\delta\tau_q(\omega)$ measure the waveform difference between the synthetic and observed waveforms.

$$\delta\tau_p(\omega) = - \int_0^T \Im \left[\frac{e^{-i\omega t} W_n(t)}{\omega u(\omega)} \right] \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s) dt \quad (2.37)$$

$$\delta\tau_q(\omega) = - \int_0^T \Re \left[\frac{e^{-i\omega t} W_n(t)}{\omega u(\omega)} \right] \delta u_n(\mathbf{x}_r, t; \mathbf{x}_s) dt, \quad (2.38)$$

where \Re and \Im represent the real and image part of a complex, and $W_n(t)$ is the time window.

2.4.2 GSDF adjoint source

Substituting eq. (2.27) into eq. (2.29), we have

$$\begin{aligned} \delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) &= \int_0^T J_n(\mathbf{x}_r, t; \mathbf{x}_s) \left[- \int_0^t \iiint_v \{ \delta\rho(\mathbf{r}) \ddot{u}_i^0(\mathbf{r}, \tau; \mathbf{x}_s) G_{in}^0(\mathbf{r}, t-\tau; \mathbf{x}_r) \right. \\ &\quad \left. + \partial_j G_{in}^0(\mathbf{r}, t-\tau; \mathbf{x}_r) \delta c_{ijkl}(\mathbf{r}) \partial_k u_l^0(\mathbf{r}, \tau; \mathbf{x}_s) \} d\mathbf{r}^3 d\tau \right] dt \end{aligned} \quad (2.39)$$

Using the following transforming for *double integrals*:

$$\int_o^T \int_0^t d\tau dt = \int_0^T \int_\tau^T dt d\tau, \quad (2.40)$$

we have

$$\begin{aligned} \delta\phi_n(\mathbf{x}_r, t; \mathbf{x}_s) &= - \iiint_v \int_0^T \int_\tau^T \{ \delta\rho(\mathbf{r}) J_n(\mathbf{x}_r, t; \mathbf{x}_s) \ddot{u}_i^0(\mathbf{r}, \tau; \mathbf{x}_s) G_{in}^0(\mathbf{r}, t-\tau; \mathbf{x}_r) \\ &\quad + \partial_j G_{in}^0(\mathbf{r}, t-\tau; \mathbf{x}_r) \delta c_{ijkl}(\mathbf{r}) J_n(\mathbf{x}_r, t; \mathbf{x}_s) \partial_k u_l^0(\mathbf{r}, \tau; \mathbf{x}_s) \} d\mathbf{r}^3 dt d\tau. \end{aligned} \quad (2.41)$$

Upon reversing time by making the substitution $t \rightarrow T-t$ and $\int_\tau^T dt = \int_0^{T-\tau} dt$ we obtain

$$\begin{aligned} \delta\phi_n(\mathbf{x}_r, T-t; \mathbf{x}_s) &= - \iiint_v \int_0^T \int_0^{T-\tau} \{ \delta\rho(\mathbf{r}) J_n(\mathbf{x}_r, T-t; \mathbf{x}_s) \ddot{u}_i^0(\mathbf{r}, \tau; \mathbf{x}_s) G_{in}^0(\mathbf{r}, T-t-\tau; \mathbf{x}_r) \\ &\quad + \partial_j G_{in}^0(\mathbf{r}, T-t-\tau; \mathbf{x}_r) \delta c_{ijkl}(\mathbf{r}) J_n(\mathbf{x}_r, T-t; \mathbf{x}_s) \partial_k u_l^0(\mathbf{r}, \tau; \mathbf{x}_s) \} d\mathbf{r}^3 dt d\tau \end{aligned} \quad (2.42)$$

Now, we define the *GSDF adjoint field* as

$$u_n^\dagger(\mathbf{r}, T - \tau; \mathbf{x}_r) = \int_0^{T-\tau} J_n(\mathbf{x}_r, T-t; \mathbf{x}_s) G_{in}^0(\mathbf{r}, T-t-\tau; \mathbf{x}_r) dt \quad (2.43)$$

Upon reversing time by making the substitution $\tau \rightarrow T - \tau$,

$$\begin{aligned} u_n^\dagger(\mathbf{r}, \tau; \mathbf{x}_r) &= \int_0^\tau f_n^\dagger(\mathbf{x}_r, t; \mathbf{x}_s) G_{in}^0(\mathbf{r}, \tau-t; \mathbf{x}_r) dt \\ &= \int_0^\tau f_n^\dagger(\mathbf{x}_r, \tau-t; \mathbf{x}_s) G_{in}^0(\mathbf{r}, t; \mathbf{x}_r) dt \end{aligned} \quad (2.44)$$

The corresponding *GSDF adjoint source* is:

$$f_n^\dagger(\mathbf{x}, t; \mathbf{x}_s) = J_n(\mathbf{x}_r, T-t; \mathbf{x}_s) \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.45)$$

For the four widely used data functionals, the corresponding adjoint sources are:

1. Waveform adjoint source.

$$f_n^\dagger(\mathbf{x}, t; \mathbf{x}_s) = [u_n^{syn}(\mathbf{x}_r, t; \mathbf{x}_s) - u_n^{obs}(\mathbf{x}_r, t; \mathbf{x}_s)] \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.46)$$

2. Traveltime adjoint source.

$$f_n^\dagger(\mathbf{x}, t; \mathbf{x}_s) = \frac{\partial_t u_n(\mathbf{x}_r, T-t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, T-t; \mathbf{x}_s)|^2 dt} \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.47)$$

3. Amplitude adjoint source.

$$f_n^\dagger(\mathbf{x}, t; \mathbf{x}_s) = \frac{u_n(\mathbf{x}_r, T-t; \mathbf{x}_s)}{\int_{t1}^{t2} |u_n(\mathbf{x}_r, T-t; \mathbf{x}_s)|^2 dt} \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.48)$$

4. GSDF frequency-dependent adjoint sources.

$$f_{n,p}^\dagger(\mathbf{x}, t; \mathbf{x}_s) = -\Re \left[\frac{e^{-i\omega t} W_n(t)}{\omega u(\omega)} \right] \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.49)$$

$$f_{n,q}^\dagger(\mathbf{x}, t; \mathbf{x}_s) = -\Im \left[\frac{e^{-i\omega t} W_n(t)}{\omega u(\omega)} \right] \delta(\mathbf{x} - \mathbf{x}_r) \quad (2.50)$$

2.4.3 Sensitivity kernels

The misfit perturbation can be linearized relative to model perturbations through 3-D Fréchet kernels as

$$\delta\phi = \int_V [K_\rho(\mathbf{x}) \delta \ln \rho(\mathbf{x}) + K_{cijkl}(\mathbf{x}) \delta \ln c_{ijkl}(\mathbf{x})] d^3\mathbf{x}, \quad (2.51)$$

where $K_\rho, K_{c_{ijkl}}$ are the Fréchet kernels (also known as *sensitivity kernels*) defined for fractional density and elastic tensor. Upon substituting eq. (2.43) into eq. (2.41), and make the substitution $\tau \rightarrow t$ and $\mathbf{r} \rightarrow \mathbf{x}$, we have

$$K_\rho = -\rho(\mathbf{x}) \int_0^T [\mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)] dt \quad (2.52)$$

$$K_{c_{ijkl}} = -\rho(\mathbf{x}) \int_0^T [\epsilon_{ij}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r) c_{ijkl}(\mathbf{x}) \partial_t^2 \epsilon_{kl}(\mathbf{x}, t; \mathbf{x}_s)] dt, \quad (2.53)$$

where $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)$ is the forward wavefield emanated from the source in the reference model, and $\mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)$ is the adjoint wavefield generated by setting time-reversed data residual signals at receivers as simultaneous fictitious adjoint sources. ϵ_{kl} and ϵ_{ij}^\dagger are the strain and adjoint strain tensors, respectively.

For isotropic models, we can choose density ρ , shear modulus μ , and bulk modulus κ as model parameters. By using $c_{ijkl} = (\kappa - 2\mu/3)\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$, we can rewrite the variation of ϕ as

$$\delta\phi = \int_V [K_\rho(\mathbf{x})\delta \ln \rho(\mathbf{x}) + K_\mu(\mathbf{x})\delta \ln \mu(\mathbf{x}) + K_\kappa(\mathbf{x})\delta \ln \kappa(\mathbf{x})] d^3\mathbf{x}, \quad (2.54)$$

where K_μ, K_κ are the sensitivity kernels defined for fractional shear and bulk moduli perturbations. More specifically,

$$K_\mu = -2\mu(\mathbf{x}) \int_0^T [\mathbf{D}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r) : \mathbf{D}(\mathbf{x}, t; \mathbf{x}_s)] dt \quad (2.55)$$

$$K_\kappa = -\kappa(\mathbf{x}) \int [\nabla \cdot \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)] [\nabla \cdot \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)] dt, \quad (2.56)$$

where \mathbf{D} and \mathbf{D}^\dagger denote the strain deviators for the forward and adjoint wavefield, respectively. Alternatively, we may express the above kernels in terms of model variations in density, shear wave speed, and compressional wave speed:

$$K_{\rho'} = K_\rho + K_\mu + K_\kappa, K_\alpha = 2 \left(K_\mu - \frac{4}{3} \frac{\mu}{\kappa} K_\kappa \right), K_\beta = 2 \left(\frac{K_\kappa + \frac{4}{3}\mu}{\kappa} \right) K_\kappa, \quad (2.57)$$

such that

$$\delta\phi = \int_V [K_{\rho'}(\mathbf{x})\delta \ln \rho(\mathbf{x}) + K_\alpha(\mathbf{x})\delta \ln \alpha(\mathbf{x}) + K_\beta(\mathbf{x})\delta \ln \beta(\mathbf{x})] d^3\mathbf{x}. \quad (2.58)$$

Note that these formal kernel expressions in eq. (2.52)-(2.57) are true for different types of misfit functions or measurements ϕ . The choice of misfit function only affects the adjoint sources (i.e., data residual injected at receivers) and hence the adjoint wavefields.

2.5 Travel time adjoint tomography

Taking traveltime measurements as an example, now I summarize the basic concepts of misfit function and sensitivity kernel used in both classic finite-frequency and adjoint tomography, in particular emphasizing their differences in kernel construction and inversion procedures.

2.5.1 Kernels for traveltime measurement

In the following derivations, we use $m(\mathbf{x})$ to abstractly represent the properties of the medium. To illustrate our point, we will take travelttime misfit functions as an example, and as stated in section 2.4.3, all other choices of misfit functions only affect adjoint sources, and do not change the formal kernel expressions in section 2.4.3.

In the classic finite-frequency theory [Dahlen et al., 2000, Hung et al., 2004, Zhao et al., 2000], the travelttime perturbation is linearly related to model perturbations through a Fréchet kernel (also known as *banana-doughnut kernel* for body-wave travelttime) that illuminates mostly the volume in the vicinity of the source-receiver ray path as

$$\delta T_i = \int_V K_i(\mathbf{x}) \delta \ln m(\mathbf{x}) d^3\mathbf{x}, \quad (2.59)$$

where $K_i(\mathbf{x})$ is the individual sensitivity kernel for the i th travelttime measurement and $\delta \ln m(\mathbf{x})$ is the fractional model perturbation from the reference (or current) model $m(\mathbf{x})$. In practice, the model perturbation is often discretized based on a set of basis functions $B_k(\mathbf{x})$, $k = 1, \dots, M$, as

$$\delta \ln m(\mathbf{x}) = \sum_{k=1}^M B_k(\mathbf{x}) \delta m_k, \quad (2.60)$$

where δm_k is the model perturbation coefficient related to the k th basis function B_k . Substituting eq. (2.60) into (2.59), we obtain the linear equation

$$\delta T_i = \sum_{k=1}^M G_{ik} \delta m_k, \quad (2.61)$$

where

$$G_{ik} = \int_V K_i(\mathbf{x}) B_k(\mathbf{x}) d^3\mathbf{x} \quad (2.62)$$

is the discretized kernel for the i th measurement (i.e., specific to a source-receiver path). As shown in section 2.4.3, the construction of kernel $K_i(\mathbf{x})$ depends on the reference (or current) model $m(\mathbf{x})$.

Assuming small perturbations, one can assemble eq. (2.61) for traveltime measurements from all source-receiver pairs (from multiple events) into the matrix form

$$\Delta\mathbf{T} = \mathbf{G}\Delta\mathbf{m}, \quad (2.63)$$

where the row vectors of the *design matrix* \mathbf{G} are given by individual sensitivity kernels $K_i(\mathbf{x})$. In classical ray-based or finite-frequency-based tomography, as individual sensitivity kernels $K_i(\mathbf{x})$ are easily computed, the *design matrix* \mathbf{G} in eq. (2.63) can be formed and model perturbations $\Delta\mathbf{m}$ can be solved by least-square inversion methods with regularization [e.g., Hung et al., 2004, Menke, 1984, Paige and Saunders, 1982].

However, when sensitivity kernels are computed based on full numerical simulations [Liu and Tromp, 2006, 2008], it is computationally prohibitive to calculate each individual kernel $K_i(\mathbf{x})$ and construct the full \mathbf{G} matrix. Instead, sensitivity kernel for the misfit of all measurements from a single event (also known as *event kernel*) is computed and used as the basic building blocks of adjoint tomography [e.g., Tape et al., 2007].

2.5.2 Event kernel for travel time misfit

We denote the total travelttime misfit of an event e as the sum of travelttime misfit for all measurements ($i = 1, \dots, N_e$) made for this event,

$$\phi^e(m) = \sum_{i=1}^{N_e} \phi_i^e(m) = \frac{1}{2} \sum_{i=1}^{N_e} \left[T_i^{e,obs} - T_i^e(m) \right]^2 = \frac{1}{2} \sum_{i=1}^{N_e} [\Delta T_i^e(m)]^2 \quad (2.64)$$

where $T_i^{e,obs}$ is the i th observed travelttime for this event, and $T_i^e(m)$ is the corresponding predicted travelttime calculated for model m . Using eq. (2.59) and (2.60), the variation of eq. (2.64) is

$$\delta\phi^e = \sum_{i=1}^{N_e} \delta\phi_i^e = \int_V \widehat{K}^e(\mathbf{x}) \delta \ln m d^3\mathbf{x} = \sum_{k=1}^M \widehat{G}_k^e \delta m_k, \quad (2.65)$$

where

$$\widehat{K}^e(\mathbf{x}) = - \sum_{i=1}^{N_e} \Delta T_i^e K_i^e(\mathbf{x}), \quad (2.66)$$

$$\widehat{G}_k^e = \int_V \widehat{K}^e(\mathbf{x}) B_k(\mathbf{x}) d^3\mathbf{x}. \quad (2.67)$$

The *event misfit kernel* $\widehat{K}^e(\mathbf{x})$ is therefore a sum of individual banana-doughnut kernels $K_i^e(\mathbf{x})$ weighted by travelttime anomalies ΔT_i^e . \widehat{G}^e gives the discretized event kernel vector. It should be noted that the summation of individual misfit kernels is implicitly implemented

by introducing the *combined traveltime adjoint field*, which is generated by injecting time-reversed adjoint sources at all receivers as simultaneous, fictitious sources [Tromp et al., 2005].

If we have a total E events, each with associated event kernels as in eq. (2.65), then we have in matrix form,

$$\Delta\phi = \widehat{\mathbf{G}}\Delta\mathbf{m}, \quad (2.68)$$

where $\Delta\phi$ is the event misfit vector of size $E \times 1$, $\widehat{\mathbf{G}}$ is a $E \times M$ matrix with discretized event kernels as row vectors, and $\Delta\mathbf{m}$ is the $M \times 1$ model column vector. This linear system is much reduced from the full least-square problem of eq. (2.63). Often in adjoint tomography it is only numerically feasible to construct event kernels and only eq. (2.68) is available for model update. For most inversions, the number of model parameters M is much larger than the number of events E , and eq. (2.68) is an under-determined system that cannot provide a unique model update. However, it can be used to successfully update the model through iterative methods.

2.5.3 Total traveltime misfit

We now consider the total traveltime misfit as the sum of misfit functions for a total E event

$$\phi = \sum_{e=1}^E \phi^e = \frac{1}{2} \sum_{e=1}^E \sum_{i=1}^{N_e} [\Delta T_i^e(m)]^2 \quad (2.69)$$

where N_e indicates the number of measurements made for the e th event. We can take its variation as

$$\delta\phi = \sum_{i=1}^E \delta\phi^e = \int_V \widehat{K}(\mathbf{x}) \delta \ln m \, d^3\mathbf{x} = \sum_{k=1}^M \widehat{G}_k \delta m_k, \quad (2.70)$$

where the total misfit kernel $\widehat{K}(\mathbf{x})$ is related to event misfit kernels and individual banana-doughnut kernels as

$$\widehat{K}(\mathbf{x}) = \sum_{e=1}^E \widehat{K}^e(\mathbf{x}) = - \sum_{e=1}^E \sum_{i=1}^{N_e} \Delta T_i^e K_i^e(\mathbf{x}) \quad (2.71)$$

with similar relations for the discretized kernels

$$\widehat{G}_k = \sum_{e=1}^E \widehat{G}_k^e = - \sum_{e=1}^E \sum_{i=1}^{N_e} \Delta T_i^e G_{ik}^e. \quad (2.72)$$

Here \widehat{G}_k is the gradient of the misfit function with respect to discretized model parameters m_k .

We rewrite eq. (2.70) in matrix form:

$$\Delta\phi = \widehat{\mathbf{G}}\Delta\mathbf{m}, \quad (2.73)$$

where $\Delta\phi$ is a 1×1 , $\widehat{\mathbf{G}}$ is a $1 \times M$ matrix, and $\Delta\mathbf{m}$ is the $M \times 1$ model column vector. Instead of solving this odd equation as a linear system, it is more general to treat it as an optimization problem and minimize the total misfit function using the gradient or Hessian-based algorithm.

2.5.4 Optimization method

Considering the quadratic Taylor expansion of misfit function [Tarantola, 2005] :

$$\phi(\mathbf{m} + \delta\mathbf{m}) \approx \phi(\mathbf{m}) + \mathbf{g}^T \delta\mathbf{m} + \frac{1}{2} \delta\mathbf{m}^T \mathbf{H} \delta\mathbf{m} + O(\delta\mathbf{m}^3), \quad (2.74)$$

where the misfit gradient \mathbf{g} and Hessian \mathbf{H} are expressed as

$$\mathbf{g} = \left. \frac{\partial \phi}{\partial \mathbf{m}} \right|_{\mathbf{m}}, \quad \mathbf{H} = \left. \frac{\partial^2 \phi}{\partial \mathbf{m} \partial \mathbf{m}} \right|_{\mathbf{m}} \quad (2.75)$$

The total misfit can be reduced by either updating the model based on the negative gradient direction

$$\delta\mathbf{m} \sim -\mathbf{g} \quad (2.76)$$

or Hessian-based algorithm

$$\delta\mathbf{m} = -\mathbf{H}^{-1} \mathbf{g}, \quad (2.77)$$

where the gradient is "preconditioned" by the generalized inverse of the Hessian. We can examine the *Hessian* matrix of the misfit function by taking the second-order derivative of eq. (2.69) as

$$H_{jk} = \frac{\partial^2 \phi}{\partial m_j \partial m_k} = \sum_{e=1}^E \sum_{i=1}^{N_e} \left[\frac{\partial T_i^e}{\partial m_j} \frac{\partial T_i^e}{\partial m_k} - \Delta T_i^e \frac{\partial^2 T_i^e}{\partial m_j \partial m_k} \right], \quad (2.78)$$

where in practice, the second term on the right-hand side is often deemed small compared to the first term by assuming small data residuals ΔT_i^e . By using eq. (2.61), the resulting approximated Hessian is formally given by

$$\tilde{H}_{jk} = \sum_{e=1}^E \sum_{i=1}^{N_e} \frac{\partial T_i^e}{\partial m_j} \frac{\partial T_i^e}{\partial m_k} = \sum_{e=1}^E \sum_{i=1}^{N_e} G_{ij}^e G_{ik}^e \quad \text{or} \quad \mathbf{H} = \mathbf{G}^T \mathbf{G}, \quad (2.79)$$

where G_i are the discretized kernels for individual measurements. As discussed in section 2.5.2, we do not have access to individual kernels in adjoint tomography, therefore, we can

not explicitly form the Hessian matrix, or the second-order derivative of the misfit function. Therefore the model needs to be updated successively only based on local gradients as the summation of event misfit kernels generated in eq. (2.72). However, a variety of preconditioners can be used to approximate the Hessian and applied to the gradient vector to improve convergence.

In a different flavor of full waveform inversion [e.g., Virieux and Operto, 2009], known as scattering integral method [Chen et al., 2007b, Zhao et al., 2005], it is possible to calculate all individual kernels using full numerical simulations. In that case, the whole strain and velocity wavefield of $3N_s$ (N_s being the total number of stations) numerical simulations need to be saved spatially and temporally, which may pose a daunting storage challenge when N_s is excessive or model domain is extensive. By comparison, adjoint tomography strikes a balance between storage demand and computational cost by only calculating and saving event misfit kernels, instead of strain Green's tensors at spatial-temporal grid points.

3

Refined seismic structure of southern California by ambient noise adjoint tomography

3.1 Summary

We construct an improved shear-wave velocity (V_s) model of southern Californian crust and uppermost mantle by adjoint tomography of Rayleigh wave empirical Green's functions (EGFs) at 5-50 s periods from ambient noise cross correlations. Our initial model is the isotropic V_s model **M16** from Tape et al. [2009], which was generated by three-component seismograms at 2-30 s periods from local earthquake data. Synthetic Green's functions (SGFs) from **M16** show good agreement with the EGFs at 5-10 s and 10-20 s period bands, but they have an average 2.1 s time advance at 20-50 s. By minimizing the traveltimes differences between the EGFs and SGFs using a gradient-based algorithm, we successively refine

the Vs model, and the total misfit is reduced from 1.75 to 0.41 after five iterations.

Relative to **M16**, our new Vs model reveals: (1) a lower crust with the mean Vs about 6% slower; (2) a faster Vs speed throughout the crust in the regions of the Los Angeles Basin and Central Transverse Range; (3) higher Vs in the lower crust beneath the westernmost Peninsular Range Batholith (PRB); an enhanced high-velocity zone in the middle crust beneath Salton Trough Basin. Our updated model also reveals refined lateral velocity gradients across PRB, Sierra Nevada Batholith (SNB), and San Andreas Fault, which helps define the west-east compositional boundary between PRB and SNB.

Our study applies adjoint tomography to ambient noise data in southern California, and shows the improvement of lateral coverage and depth sensitivity from using ambient noise instead of only earthquakes. The numerical spectral-element solver used in adjoint tomography provides accurate structural sensitivity kernels, and hence generates more robust images than those by traditional ambient noise tomography based on approximate methods.

3.2 Introduction

Regional- and global-scale tomographic models are typically constructed using measurements of P- or S-wave traveltimes [e.g., Hung et al., 2004, Masters et al., 2000] or surface wave dispersions [e.g., Ritzwoller et al., 2002, Yang and Forsyth, 2006, Zhou et al., 2006] extracted from teleseismic events [Liu and Gu, 2012]. However, there are limitations from using these teleseismic observations. Teleseismic body-wave tomography has limited resolution at lithospheric depths since ray paths from teleseismic body waves are nearly vertical at shallow depths and thus hardly cross each other. Surface waves from teleseismic events often lack high-frequency information as a result of intrinsic attenuation and scattering. Furthermore, local earthquake tomography is only feasible in seismically active regions with dense seismic networks, and source parameters may need to be estimated simultaneously [Zhang and Thurber, 2003]. All earthquake-based tomography methods are affected by uneven distributions of earthquakes.

Ambient noise tomography (ANT) has become an innovative technique in seismic imaging because it can retrieve empirical Green's functions (EGFs) between a pair of receivers from random ambient noise [e.g., Lobkis and Weaver, 2001, Roux et al., 2005]. Using this method, Shapiro et al. [2005] successfully performed the first ambient noise surface wave

tomography of southern California using Rayleigh waves extracted from cross correlations of ambient noise. Since then, ANT has been widely employed to image crustal and upper mantle structures from regional [e.g., Lin et al., 2008, Yang et al., 2008, Yao et al., 2006] to continental [e.g., Bensen et al., 2009, Yang et al., 2007] and global [e.g., Haned et al., 2015, Nishida et al., 2009] scales. Compared with traditional earthquake-based tomography, ANT is not affected by distributions of earthquakes, and can be applied to any seismic array where seismic noises are readily recorded.

Benefiting from the developments of high-performance computing, it has become practical to simulate seismic wave propagation in complex 3-D Earth models using numerical methods such as spectral-element method (SEM) [e.g., Komatitsch and Vilotte, 1998, Tromp et al., 2008]. Utilizing forward computation capabilities, adjoint tomography based on SEM has been introduced to earthquake seismology [Tromp et al., 2005] and successfully applied to constrain crustal structures of southern California [Tape et al., 2009, 2010], and upper mantle structures of Australia, Europe, and East Asia [Chen et al., 2015, Fichtner et al., 2009, 2010, Zhu et al., 2012, 2015], among others. The idea of adjoint tomography is to iteratively minimize a misfit function between observations and synthetics by using 3-D Fréchet derivatives, which are constructed from the interference between a forward wavefield and an adjoint wavefield [Liu and Tromp, 2006]. The use of full numerical solvers for both wavefields provides accurate Fréchet kernels for the current model at each iteration, and allows more data to be used successively as the model improves. The resulting images typically provide better data fits than those generated by classical tomography based on ray theory or finite-frequency theory using asymptotic or semi-analytical methods for 1-D reference models [Liu and Gu, 2012].

In recent years, surface waves extracted from ambient noise have been used in full waveform inversion through scattering-integral method based on finite-difference modeling [e.g., Gao and Shen, 2014, Lee et al., 2014], and adjoint method based on SEM [e.g., Chen et al., 2014, Liu et al., 2017]. For example, Gao and Shen [2014] used 7-200 s broadband ambient noise to constrain the upper mantle structures of the Cascadia subduction zone. Chen et al. [2014] applied adjoint tomography to short-period (10-40 s) ambient noise recorded at 25 stations to refine the crustal model in southeastern Tibet.

Southern California is a tectonically active region with a 200-300 km wide zone of transpressive deformation spanning the Pacific-North America plate boundary [Hauksson,

2000]. The complex tectonic history has left its marks in the form of several provinces such as the Transverse Ranges, the Peninsular Ranges, the Mojave Desert, and the Salton Trough. (Figure 3.1a). Owing to abundant seismicity and one of the earliest regional permanent broadband seismic networks, traditional seismic tomographic inversions have been performed based on P and S arrivals [e.g., Hauksson, 2000, Schmandt and Humphreys, 2010], and surface waves from earthquake data [e.g., Yang and Forsyth, 2006] and ambient noise [e.g., Shapiro et al., 2005, Zygone et al., 2015]. Tape et al. [2009, 2010] generated a high-resolution 3-D crustal model **M16** by applying adjoint tomography to a dataset of waveforms recorded at 203 stations from 143 local earthquakes. Lee et al. [2014] applied a combination of scattering-integral and adjoint methods based on finite-difference modeling to both earthquake data and ambient noise at 5 – 50 s periods. In this study, we apply adjoint tomography to 5 – 50 s Rayleigh waves extracted from ambient noise to further improve the **M16** model, especially structures in the lower crust. We use SEM to simulate wave propagation through **M16** and successively update models.

In this chapter, we first describe the construction of EGFs from ambient noise in section 3.3. Then, we summarize the basic theory of adjoint tomography in section 3.4, and describe the inversion procedures in section 3.5. In section 3.6, we investigate the characteristics of the initial model and present updated models from three stages of the inversion with increasing data from different period bands. Finally, we show the robustness of our inversion results by misfit evolution analysis and comparison with two other published models.

3.3 Ambient noise data in southern California

We first build a database of vertical-vertical component cross-correlations of ambient noise in southern California (Figure 3.1a). We download vertical components of continuous long-period (LHZ) seismic data recorded at 148 stations between January 2006 and December 2012 from southern California Earthquake Data Center (SCEDC), and process them following the standard noise data processing procedures outlined in Bensen et al. [2007] to obtain cross-correlation functions at a period band of 5 – 50 s. We first cut the raw continuous data into a series of daily segments, then remove the trend, mean and instrument responses and bandpass filter them at period band of 5 – 50 s. Afterwards, we normalize the pre-processed

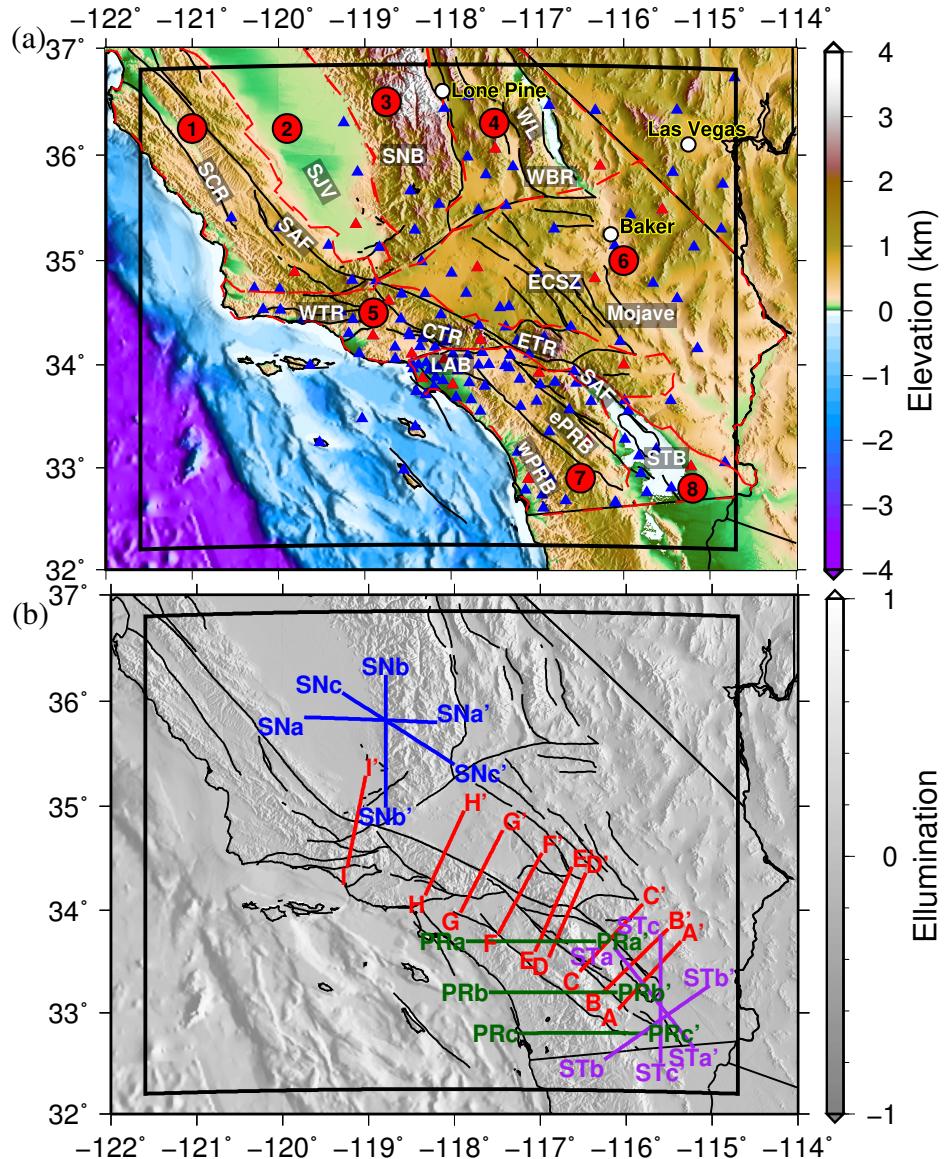


FIGURE 3.1: (a) Map of southern California with topography, bathymetry and active faults. The solid black rectangle outlines the simulation region. The 148 stations used in this tomographic study are shown as triangles, out of which 19 are selected for line search shown in red. Faults are indicated by the bold black lines. Labels 1-8 denote the major eight geological provinces with their boundaries delineated by red dash lines: 1. Coastal Ranges; 2. Great Central Valley; 3. Sierra Nevada; 4. Basin and Range; 5. Transverse Ranges; 6. Mojave Desert; 7. Peninsular Ranges; 8. Salton Trough. Geological features labeled in bold white letters as references for subsequent figures: SCR, southern Coast Range; SAF, San Andreas Fault; SJV, San Joaquin Valley; SNB, Sierra Nevada Batholith; WL: Walker Lane; WBR: Western Basin and Range; WTR, CTR, and ETR: western, central and eastern Transverse Range; ECSZ: Eastern California Shear Zone; LAB: Los Angeles Basin; ePRB and wPRB: east and west Peninsular Ranges Batholith; STB: Salton Trough Basin. (b) Locations of cross sections across San Andreas Fault: AA'-II' (red); Peninsular Ranges: (PR)aa'-cc' (dark green); Sierra Nevada: (SN)aa'-cc' (blue); Salton Trough: (ST)aa'-cc' (purple).

seismograms in the time domain and whiten their spectra in spectral domain in order to suppress the effects of earthquake signals and instrumental irregularities on cross-correlations. Finally, we compute the daily cross-correlations for each possible station pair and stack them linearly to produce the final cross-correlations. The positive and negative legs of each stacked cross-correlation are stacked further to obtain the symmetric cross-correlations. The resulting cross-correlations are dominated by Rayleigh wave signals as we only process vertical data.

In total, we obtain 10,862 cross-correlations with high quality Rayleigh waves between these stations. The symmetric cross-correlation function $C_{a,b}(t)$ between a station pair, a and b , is related to its empirical Green's function $G_{a,b}(t)$ as

$$-\frac{\partial C_{a,b}(t)}{\partial t} = G_{a,b}(t). \quad (3.1)$$

One potential concern of ANT is the influence of the uneven distribution of noise on the phases of surface wave signals emerging from cross-correlation functions [Basini et al., 2013, Wang et al., 2014a]. ANT is based on the assumption that ambient noise is diffuse and evenly distributed [Lobkis and Weaver, 2001]. However, a number of studies have shown that the distributions of ambient noise are often directional and uneven [e.g., Stehly et al., 2006, Yang and Ritzwoller, 2008]. Phase velocity biases stemming from unevenly distributed noise sources have been investigated by plane-wave modeling [Tsai, 2009] and can be mostly corrected by an iterative procedure [Wang et al., 2016, Yao and Van Der Hilst, 2009].

In the context of full waveform modeling, Tromp et al. [2010] showed that how it is theoretically possible to account for the non-uniform character of the noise sources distribution, and more recently Fichtner [2014] showed that both unevenly distributed noise sources and seismic data processing can affect cross-correlations of ambient noise and their sensitivities to structures, suggesting that carrying out full waveform ambient noise tomography without taking source heterogeneities and data processing schemes into account could introduce tomographic artifacts. On the other hand, Fichtner [2014] also suggested that, if the phase traveltimes of surface waves extracted from ambient noise are measured in narrow bands, these traveltimes are only weakly affected by source heterogeneities and data processing schemes.

In order to choose proper period bands in measuring phase traveltimes of EGFs, we examine the dominant energy in the spectra of the noise cross-correlations. As seen in Figure

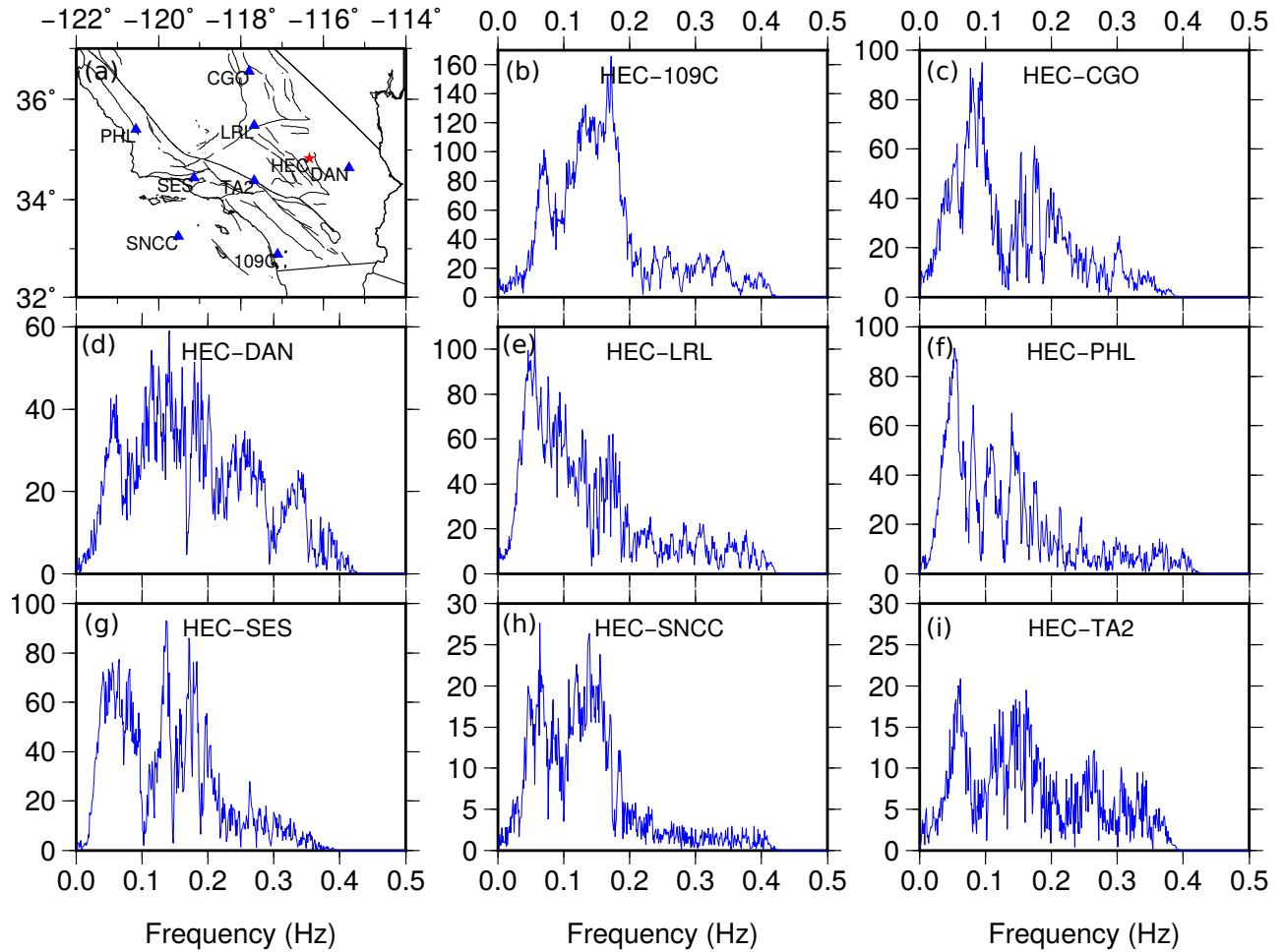


FIGURE 3.2: Examples of spectra of EGFs between the master station HEC and a number of other stations shown in (b)-(i). The location of the master station HEC is shown as the red star, and other stations are indicated by the blue triangles in (a).

3.2(b)-(i), our noise cross-correlations exhibit two peaks in the spectra between 10–20 s and 5–10 s, which are referred to as primary and secondary microseism bands [Stehly et al., 2006]. Therefore, in our inversion, we filter waveforms of EGFs in three narrow period bands, namely 5–10 s, 10–20 s, and 20–50 s. The selection of period bands will be further discussed in detail in the pre-processing steps of the inversion procedures (section 3.5).

3.4 Background of adjoint tomography

The use of adjoint methods in seismology was advocated in the 1980s [Gauthier et al., 1986, Tarantola, 1984] and expanded in the 2000s following the advent of high-performance computing and the development of finite-frequency theory [e.g., Dahlen et al., 2000, Hung et al., 2000, Liu and Tromp, 2006, Tromp et al., 2005, Zhao et al., 2000, 2005]. Here, we outline

the formal expressions for kernel calculations based on adjoint method, the basic concepts of event misfit kernel and gradient used in adjoint tomography.

If we assume ϕ to be a scalar function that quantifies the misfit of measurements between observations and synthetic predictions based on a reference (or current) model $m(\mathbf{x})$, then the misfit perturbation can be linearized relative to the reference model as

$$\delta\phi = \int_V K(\mathbf{x}) \delta \ln m(\mathbf{x}) d^3\mathbf{x}, \quad (3.2)$$

where $K(\mathbf{x})$ is the Fréchet kernels (also known as *sensitivity kernels*) defined for fractional model parameters, such as density, shear and bulk moduli perturbations. Based on adjoint method [Liu and Tromp, 2006, Tromp et al., 2005], these kernel can be expressed in a general form:

$$K(\mathbf{x}) = \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s) \otimes \mathbf{u}^\dagger(\mathbf{x}, T - t; \mathbf{x}_r), \quad (3.3)$$

where $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)$ is the forward wavefield emitted from the source, $\mathbf{u}^\dagger(\mathbf{x}, T - t; \mathbf{x}_r)$ is the adjoint wavefield generated by the time-reversed adjoint sources injected at receivers. The \otimes operator represents general interaction for different types of kernel which can be found in Tromp et al. [2005].

Adjoint tomography strikes a balance between storage demand and computational cost by only calculating and saving event misfit kernels [Tape et al., 2007], instead of individual kernel for each source-receiver pair. Assuming we have a total E events, each with associated event kernels, then we have in matrix form,

$$\Delta\phi = \mathbf{G}\Delta\mathbf{m}, \quad (3.4)$$

where $\Delta\phi$ is the event misfit vector of size $E \times 1$, \mathbf{G} is a $E \times M$ matrix with discretized event kernels as row vectors, and $\Delta\mathbf{m}$ is the $M \times 1$ model column vector. For most inversions, M is much larger than E , and eq. (3.4) is an under-determined system that cannot provide a unique model update. There are numerous options to derive the model update using event kernels [Modrak and Tromp, 2016]. We choose an update that is linearly proportional to the sum of event kernels determined as follows.

Considering the quadratic Taylor expansion of misfit function [Tarantola, 2005] :

$$\phi(\mathbf{m} + \delta\mathbf{m}) \approx \phi(\mathbf{m}) + \mathbf{g}^T \delta\mathbf{m} + \frac{1}{2} \delta\mathbf{m}^T \mathbf{H} \delta\mathbf{m} + O(\delta\mathbf{m}^3), \quad (3.5)$$

where the misfit gradient \mathbf{g} and Hessian \mathbf{H} are expressed as

$$\mathbf{g} = \left. \frac{\partial \phi}{\partial \mathbf{m}} \right|_{\mathbf{m}}, \mathbf{H} = \left. \frac{\partial^2 \phi}{\partial \mathbf{m} \partial \mathbf{m}} \right|_{\mathbf{m}} \quad (3.6)$$

The total misfit can be reduced by either updating the model based on the negative gradient direction

$$\delta\mathbf{m} \sim -\mathbf{g} \quad (3.7)$$

or Hessian-based algorithm

$$\delta\mathbf{m} = -\mathbf{H}^{-1}\mathbf{g} \quad (3.8)$$

where, the gradient is "preconditioned" by the generalized inverse of the Hessian. In adjoint tomography, we do not have access to individual kernels nor the Hessian matrix. However, a variety of preconditioners can be used to approximate the Hessian and applied to the gradient vector to improve convergence (see also section 3.5.6).

3.5 Inversion Procedures

In this section, we outline the adjoint tomography procedures used to construct a new Vs model of southern California by minimizing the phase traveltimes differences between the EGFs of station pairs and the synthetic Green's functions (SGFs) computed based on numerical simulations. As illustrated in section 3.3, the EGFs are derived from negative time derivatives of cross-correlations of ambient noise using the empirical relation eq. (3.1). We simulate the set of corresponding SGFs, measure the phase traveltimes misfits between EGFs and SGFs, and then successively update our model based on misfit gradient. All the forward and adjoint simulations in this study are performed using the open-source spectral-element package SPECFEM3D_Cartesian.

3.5.1 SEM mesh

The mesh in this study is almost identical to that used in Tape et al. [2009] except that we use the new meshing tool in the latest version of SPECFEM3D_Cartesian. It has 336 elements in longitude (639 km), 288 elements in latitude (503 km), and 11 layers in depth. The 11 layers from the bottom to the surface (1-11) are divided into four subregions: from the bottom to the Moho (3 layers), the Moho to the basin basement(5 layers), the basin basement to the sediment (2 layers), the sediment to the surface (1 layer). The depth of sediment layer is set to be one-fifth of that of basin basement. Two mesh doubling occurs right beneath the sediment (10th layer) and basin basement (8th layer). The mesh yields a horizontal grid

spacing of about 2 km and an increasing vertical spacing from ~ 1 km at the top to ~ 12 km at the bottom, which gives sufficient simulation accuracy at 5 s and longer periods. As our new mesh is slightly different from that of Tape [2009], the density, compression and shear-wave velocities of our new mesh are assigned from the nearest neighboring points in **M16** as the initial model.

3.5.2 Forward simulation

In the forward simulation for Rayleigh-wave-type SGFs, a single vertical point force $\mathbf{f} = (0, 0, f\hat{\mathbf{z}})$ is placed at the surface of a master station while all other stations are treated as generic receivers. A Gaussian function is used as the source time function of the point force:

$$g(t) = \frac{1}{\sqrt{\pi\tau}} e^{-(\frac{t}{\tau})^2}, \quad (3.9)$$

where τ is the half-duration of the source, and we set it to be 1.0 s since bandpass filters are applied in the processing later. The wavefield generated by this point-force vector source is recorded at all the other stations, and the vertical-component seismograms can be regarded as SGFs between the master station and each individual receiver. With the SEM mesh discussed in section 3.5.1, we choose a time step of 0.01 s, smaller than the suggested maximum time step of 0.012 s based on Courant condition. The SEM calculations are performed in parallel and distributed to $14 \times 12 = 168$ CPU cores (Intel Xeon E5540) at 2.53GHz. Each forward simulation typically takes about 25 mins to generate seismograms of 240 s duration.

3.5.3 Preprocessing

A number of preprocessing steps are performed to construct the needed adjoint sources based on frequency-dependent traveltimes measurements between observed (EGFs) and synthetic (SGFs) seismograms.

First, the EGFs and SGFs are cut from 20 s before to 240 s after the initiation time, and filtered to multiple period bands. A 20-second window is added before the initiation time to better measure near-field Rayleigh waves. EGFs are bandpass filtered at different frequency bands and their amplitudes are normalized to the maximum absolute values of corresponding SGFs. The selection of multiple frequency bands is a multi-scale strategy employed to reduce the non-linearity in the inversion [Akcelik et al., 2002, Bunks et al., 1995]. To this end, typically two different workflows can be used: 1) gradually introduce short-period data after

the model has improved based on long-period data, as in some continental-scale inversions [Zhu et al., 2015]; 2) alternatively, filter waveforms through multiple over-lapped period bands, as commonly used in finite-frequency body wave tomography [Hung et al., 2004] and other regional-scale inversions [Chen et al., 2014]. Here, we adopt the first option by first inverting long-period 20 – 50 s waveforms and gradually add 10 – 20 s and 5 – 10 s data.

Second, a window selection procedure is performed to isolate Rayleigh waves from the filtered EGF and SGF pairs. Chen et al. [2014] used the automatic windowing algorithm FLEXWIN [Maggi et al., 2009] to isolate impulsive energy packets in both data and synthetics for misfit measurements. However, as EGFs from ambient noise are dominated by surface waves, it is sufficient to select windows based on the estimated arrivals of surface waves for measurements. For each pair of EGF and SGF, we define a time window around the predicted Rayleigh-wave group-velocity arrival time as $[\frac{D}{U_{max}} - \frac{T}{2}, \frac{D}{U_{min}} + \frac{T}{2}]$, where D represents the inter-station distance, T denotes the maximum period of data used in the current inversion, U_{min} and U_{max} are the typical minimum and maximum group velocities at these period bands. We choose U_{min} and U_{max} from dispersion analysis, allowing the entire surface-wave package to be included in the misfit measurements.

3.5.4 Measurement

We measure the frequency-dependent phase travelttime misfit over a single time window for a particular event e based on the multi-taper technique as [e.g., Tape et al., 2010]

$$\phi_i^e(\mathbf{m}) = \int_{-\infty}^{+\infty} \frac{h_i(\omega)}{H_i} \left[\frac{\Delta T_i(\omega, \mathbf{m})}{\sigma_i} \right]^2 d\omega, \quad (3.10)$$

where \mathbf{m} is a model vector, $\Delta T_i(\omega, \mathbf{m}) = T_i^{obs}(\omega) - T_i(\omega, \mathbf{m})$ represents the frequency-dependent travelttime difference between the SGFs and EGFs over this time window i of a particular station with its uncertainty σ_i , $h_i(\omega)$ is a frequency-domain window to which we associate the normalization constant $H_i = \int_{-\infty}^{+\infty} h_i(\omega) d\omega$. The detailed expressions of the misfit and corresponding adjoint source can be found in Appendix C of Tape [2009].

A set of quality control parameters are applied to only choose those windows with good fits between data and synthetics. This includes minimum/maximum values of timeshift (dT), amplitude difference ($d \ln A$), and minimum cross-correlation coefficient (CC_{min}). In Table 3.1, we list the values of these parameters used at each iteration. As we normalize the EGFs to the corresponding SGFs, $d \ln A$ is set to [-1,1]. We use tighter quality control parameters

Summary of misfit measurements and gradient smoothness parameters

	1st iteration	2nd-3rd iteration		4th-5th iteration		
$T(sec)$	[20, 50]	[10, 20]	[20, 50]	[5, 10]	[10, 20]	[20, 50]
$dT(sec)$	[-4.5, 4.5]	[-3.5, 3.5]	[-4.5, 4.5]	[-2.5, 2.5]	[-3.5, 3.5]	[-4.5, 4.5]
$d \ln A$	[-1.0, 1.0]	[-1.0, 1.0]	[-1.0, 1.0]	[-1.0, 1.0]	[-1.0, 1.0]	[-1.0, 1.0]
CC_{min}	0.69	0.75	0.75	0.80	0.80	0.80
$\sigma_{(h,v)}(km)$	(20.0, 10.0)	(15.0, 7.0)	(15.0, 7.0)	(10.0, 5.0)	(10.0, 5.0)	(10.0, 5.0)

Optimal step length

	1st iteration	2nd iteration	3rd iteration	4th iteration	5th iteration
α	0.12	0.05	0.03	0.02	0.02

TABLE 3.1: Values of inversion parameters used over iterations. T is the period bands for filtered data used in each iteration; dT , $d \ln A$, CC_{min} are timeshifts, amplitude differences, and cross-correlation coefficients criteria for window selection; the two $\sigma_{(h,v)}$ are the horizontal and vertical radius of the Gaussian function used to smooth the misfit gradient; α represents the optimal step length chosen at each iteration.

as the inverted model improves iteratively, including smaller dT for shorter period bands and larger CC_{min} .

The total misfit for all measurement windows is then expressed as the sum of misfit for individual windows in eq. (3.10):

$$\phi(\mathbf{m}) = \frac{1}{E} \frac{1}{N_e} \sum_{e=1}^E \sum_{i=1}^{N_e} \phi_i^e(\mathbf{m}), \quad (3.11)$$

where E and N_e are the number of events and measurements for e th event. If multiple period bands are used, we simply combine misfits or adjoint-sources of different period bands with the same weight.

3.5.5 Adjoint simulation

For a master station, once adjoint sources at all receivers are computed, we perform the adjoint simulation by putting time-reversed adjoint sources at the receivers simultaneously to generate the adjoint wavefield. At the same time, we recover the forward wavefield from the last time step of displacement and velocity wavefield saved in the forward simulation. The multiplication of adjoint wavefield and backward reconstructed forward wavefield generate

the event kernel for this master station.

These simulations are performed on the same 168 processors and it takes about 1.5 h to complete both the forward and adjoint simulation for a single master station. In our study, we ignore the effect of attenuation, because we do not expect its minor influence on our target periods (≥ 5 s) to be worth the computational cost.

3.5.6 Preconditioner and kernel smoothing

As discussed in section 3.4, adjoint tomography only provides the first-order derivative of total misfit function (i.e., gradient vector), not its second-order derivative (i.e., the Hessian). Nevertheless, a variety of preconditioners can be used to approximate the Hessian and applied to the gradient vector to improve convergence. Luo [2012] tested four different types of preconditioners and recommended a preconditioner based on the second term of the diagonal Hessian for its faster and more steady convergence in both data and model space, which is expressed as

$$P(x) = \int \partial_t^2 \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_s) \cdot \partial_t^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s) dt \quad (3.12)$$

In SPECFEM3D_Cartesian, this preconditioner is calculated by the multiplication of the adjoint accelerations and forward acceleration.

In this study, we use a 3-D Gaussian function to smooth the preconditioned misfit gradient. Generally, the smoothing radius should be at least one element size (to eliminate sensitivity singularity at sources and receivers) and close to the wavelength of resolvable structures by the current period band. See Table 3.1 for the choices made in this study.

3.5.7 Updating the model and reducing the seismogram misfit

We update the model using a line search within a gradient-based optimization method (see Appendix 3.9). In Figure 3.3, we plot the total misfit reduction for the individual period band used in the various stages of the inversion. After the first iteration, we introduce measurements made at the 10 – 20 s band. It produces a slight increase in the misfit for the 20 – 50 s band which decreases in the last two iterations. Furthermore, the value of the averaged total misfit for the two bands continues to decrease during the second and third iterations. After introducing the third band of 5 – 10 s, we see that the overall misfit continues to decrease, albeit slower than previous iterations. After the fifth iteration, the averaged total misfit of

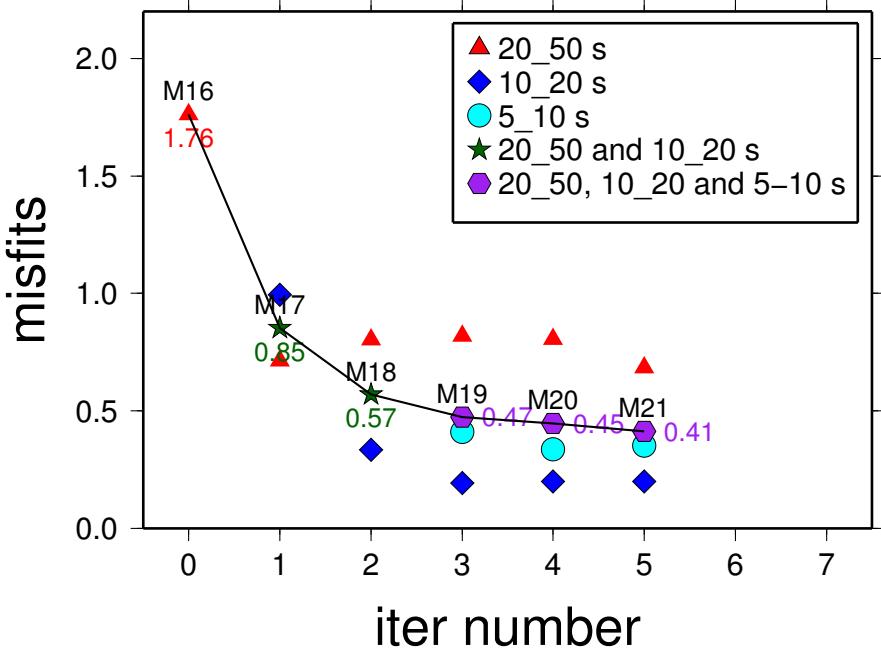


FIGURE 3.3: Total misfit reduction over iterations. Different symbols with different colors, as shown in the legend box, are used to represent the different frequency bands used in measuring the total misfit. The solid curve connects the final total misfit reduction over the three inversion stages using different frequency bands: the first ($20 - 50$ s), the second ($20 - 50$ and $10 - 20$ s), and the third ($20 - 50$, $10 - 20$ and $5 - 10$ s).

these three bands as well as the single period bands over one iteration is less than 2% , at which point we terminate our inversions.

3.6 Results

Here, we present and discuss the final shear wave speed model of southern California obtained after three stages of adjoint tomographic inversion from **M16** to **M17** (stage 1), **M17** to **M19** (stage 2), and **M19** to **M21** (stage 3).

3.6.1 Initial model

We start our inversion using the model **M16** obtained by Tape et al. [2010] as described in section 3.2. Following the procedures described in section 3.5, we first perform a set of forward simulations using each individual station as the master station. As suggested by the energy distribution in the spectra of EGFs (Figure 3.2), we filter waveforms at three different period bands, i.e. $5 - 10$ s, $10 - 20$ s and $20 - 50$ s, and measure frequency-dependent traveltime misfits of SGF-EGF pairs. In Figure 3.4(a), we present histograms of timeshift

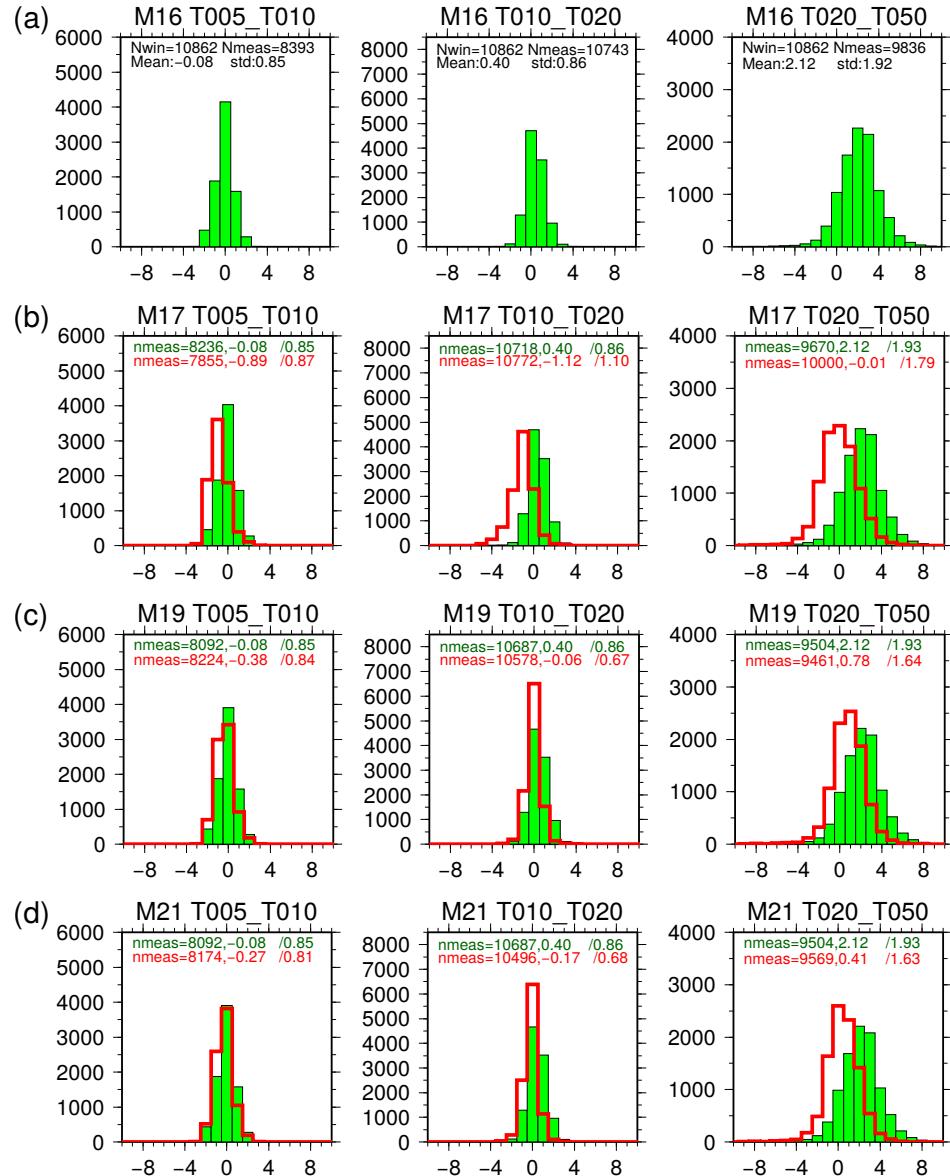


FIGURE 3.4: Histograms of traveltime misfits obtained for model **M16** (a), **M17** (b), **M19** (c) **M21** (d). The green solid bars represent misfits for the initial model **M16**, and the red bars in (b-d) are misfits for the updated models. Misfits are measured at all three period bands, 5 – 10 (left), 10 – 20 (middle), and 20 – 50 (right) s. CC_{min} in (a) are chosen as 0.80 for 5 – 10 s, 0.75 for 10 – 20 s , 0.69 for 20 – 50 s to show the misfits for the initial model. All the misfits are measured with $CC_{min} = 0.75$ in (b) and $CC_{min} = 0.80$ in (c) and (d) for comparison. Note that the measurements of 5 – 10 and 10 – 20 s bands in (a) are not used in the stage 1 inversion , and those of 5 – 10 s in (2) are not used in the stage 2 inversion.

measurements made at these three bands in selected windows based on criteria outlined in section 3.5.3. For the initial model **M16**, in the 5–10 s band (Figure 3.4a, left), the histogram is symmetric around zero and most timeshifts are within the range of [−2, +2] s, indicating that **M16**, built solely on earthquake data, also produces good fits between SGFs and our EGFs. However, there is still room for improvements, e.g. the number of windows that pass quality control criteria in misfit measurements (i.e., *good measurements*, thereafter) is small (8393 out of 10862). Similar characteristics are also observed in the 10–20 s band (Figure 3.4a, middle), with a slightly wider distribution of timeshift values in [−2, +3] s, a larger number of good measurements (10743 out of 10862), and a small skew towards positive timeshift. However, in the 20–50 s band (Figure 3.4a, right), the distribution is clearly asymmetric and biased towards positive timeshift, which indicates that parts of the model may be significantly slower than **M16**. It is expected that our inversion will focus first on reducing the misfit in this 20–50 s band.

3.6.2 Updated models

We start our inversion with measurements only from the long-period band of 20–50 s, and progressively add measurements from the 10–20 s and 5–10 s bands as the misfit reduction over iterations from existing band(s) becomes less evident. Alternative multi-scale strategy using broad bands (e.g., 5–50 s, 10–50 s, and 20–50 s) is also possible. We demonstrate that, in our case, this strategy does not alter the main features of the adjoint inversion (see Figures S1–S4).

In Figure 3.5, we display the absolute shear wave speed of **M16** (left column) and cumulative variations in percentage from **M16** for model **M17** (central column) obtained after the first stage inversion and model **M19** (right column) obtained after the second stage inversion. The models are masked based on amplitudes of volumetric coverage kernels shown in the supporting document (Figure S6). After the first iteration using solely 20–50 s band measurements, we obtain the model **M17**. As expected from the positive bias in the traveltimes misfit distribution (Figure 3.4a, right), V_s is broadly slowed down over the whole crust and uppermost mantle with variations at places reaching as much as $\sim -6\%$ in the lowest crust and $\sim -12\%$ in the uppermost mantle (35 km). The total misfit for this band also drops significantly (Figure 3.3). In the next iteration, misfits seem to change very little in the line search (Figure 3.A1b). It suggests that we already fit most measurements at this band in

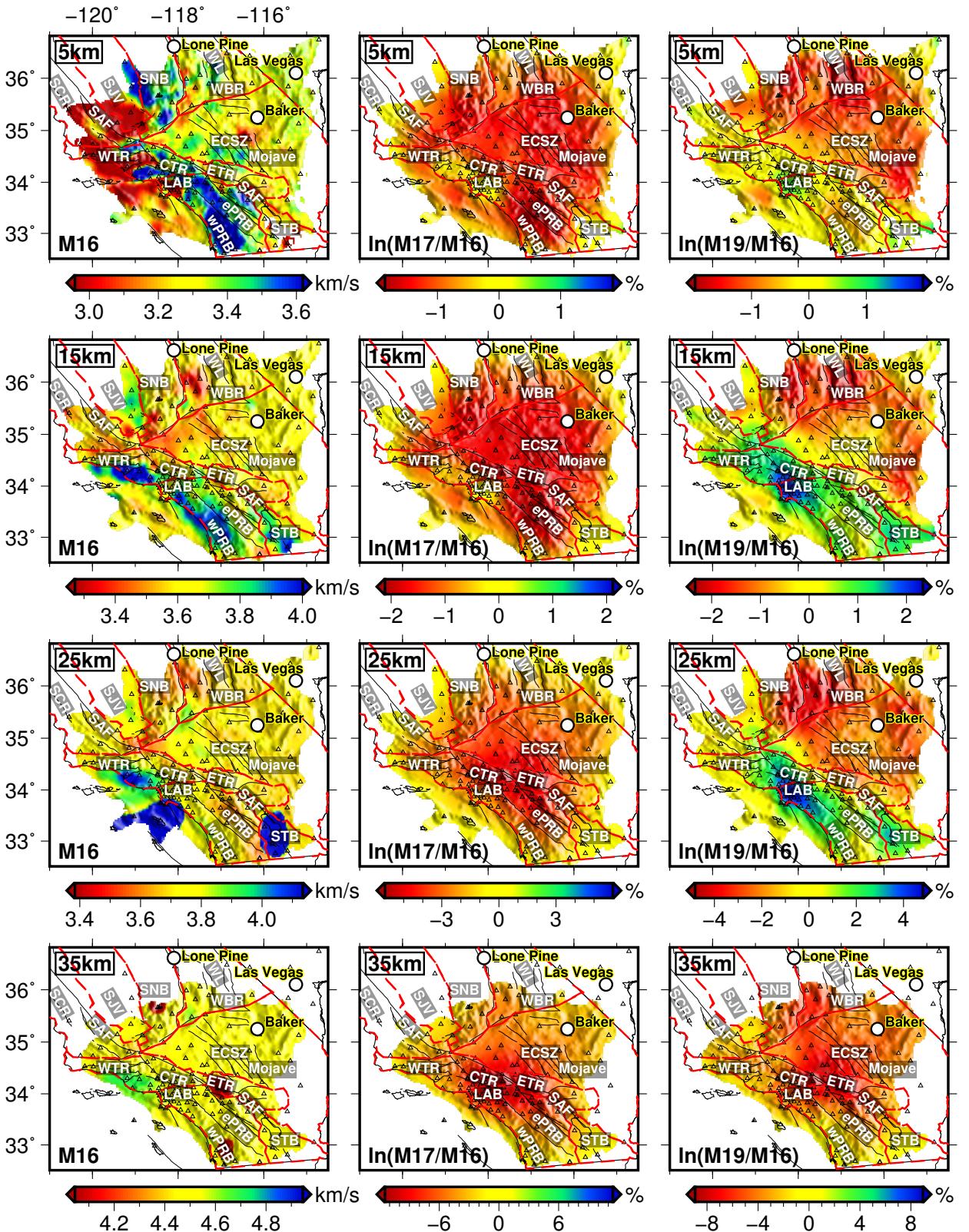


FIGURE 3.5: Horizontal slices of shear velocity of model **M16** (left panel), **In(M17/M16)** (middle panel), and **In(M19/M16)** (right panel) at 5, 15, 25, and 35 km depths. In the vicinity of the Moho, the slices reveal patches that represent wave speeds from the underlying mantle (or overlying crust).

one iteration, prompting the inclusion of 10 – 20 s band data. We then perform two more iterations using joint 20 – 50 and 10 – 20 s band data, and obtain **M19**. With the additional data, **M19** shows an increase of V_s in the Central Transverse Range (CTR), Los Angeles Basin (LAB, with maximum $\sim +4\%$ perturbation), Peninsular Ranges Botholith (PRB), and Salton Trough Basin (STB) while further slowing down V_s in parts of the Mojave Desert, western edge of Basin and Ranges (WBR), and southern Sierra Nevada Botholith (SNB), particularly at the lower crustal depth. The emergence of these features coincides with the fact that traveltimes misfit distribution for the 10 – 20 s period band becomes fairly symmetric from Figure 3.4a to Figure 3.4c.

3.6.3 Final model

We then perform another two iterations by including the measurements from the band of 5 – 10 s to obtain the final model **M21**. As stated in section 3.3, EGFs in the secondary microseism band of 5 – 10 s have strong energy that can constrain the shallowest part of the model. As shown in Figures 3.6, the inclusion of the 5 – 10 s measurements helps refine small-scale V_s structures in the crust, with perturbations ranging between $[-4.5, +4.5]\%$ from the initial model.

In the following, we compare our final model **M21** with the initial model **M16** from Tape et al. [2010] by investigating a series of horizontal slices (Figures 3.6-3.7) and vertical cross-sections (Figures 3.8-3.10), and discuss their differences introduced by the EGFs from ambient noise data.

Horizontal slices

In the uppermost crust (≤ 5 km), we do not see significant improvements. This is not unexpected, considering the few number of measurements at short-period band (Figure 3.4(d), 5 – 10 s) in comparison to the large number of shorter-period band (2 – 30 s) measurements used to generate the initial model **M16** [Tape et al., 2010]. In the middle crust ($10 - 15$ km), the model update $\ln(\mathbf{M21}/\mathbf{M16})$ reveals three distinct regions with shear velocity reductions, including the southern SNB, the Walker Lane (WL) region, and the eastern Mojave Desert. These areas are not well resolved in the previous earthquake-data based adjoint inversion [Tape et al., 2010] due to the limited distribution of earthquakes. Small increases in shear velocity are observed for the LAB, CTR, and STB area. In the lower crust (20 – 30 km),

high velocity anomalies persist with changes up to $\sim +3\%$ in the CTR, STB, and the maximum reaching $+4\%$ in the LAB. We also observe higher Vs beneath the westernmost PRB. Another distinct feature is that the crust in the SNB, WL and the eastern Mojave and WBR is significantly slowed down by $\sim 6\%$. The whole uppermost mantle is also significantly slowed down by up to 12%.

Vertical cross-sections

We display several vertical profiles through SAF, PRB, and STB (see Figure 3.1b for locations). These profiles are plotted in a way similar to those shown in Barak et al. [2015] (with some changes in width) for the purpose of comparison.

As shown in Figure 3.8, both the initial model **M16** and our final model **M21** show shear velocity contrasts (~ 3.8 km/s in the west and ~ 3.5 km/s in the east) across SAF in the crust. At the southern part of SAF (profiles AA' to CC') through ePRB, ETR and STB, **M16** only shows this lateral velocity contrast across SAF down to the middle crust, while, in our model, it extends to the lower crust. Both models show that the high velocities in the west dip gradually to the northeast. For the central part of SAF (profiles DD' to FF') through ETR, our model exhibits a high velocity layer in the middle crust that dips slightly to the northeast, but this feature is not seen in the initial model. Towards the northwest (profiles GG', and HH') from ETR to SCR, our model reveals enhanced velocity contrasts that are not exactly collocated with the SAF.

In the PRB region (Figure 3.9), a west-east velocity contrast is observed, dividing the seismic structure of PRB into western (wPRB) and eastern (ePRB) parts. Shear velocities reach ~ 3.9 km/s in the middle and lower crust beneath wPRB, while remaining low (≤ 3.7 km/s) throughout the crust beneath ePRB. In the lower crust, high velocities are only observed at the westernmost part of wPRB, and they are significantly enhanced in our model relative to the initial model. The high velocity zone beneath wPRB dips slightly to the northeast in the northeastern part, while the southern part shows a vertical-to-steep dipping angle.

From Figure 3.10, it is observed that Vs in the STB region is enhanced in the middle and lower crust . At the southeastern edge of STB, a shallow thin sheet-like high-velocity structure is seen in the middle crust (10 – 20 km), about 30 km wide and surrounded by low shear velocities. Similar high velocities observed separately at the northwestern part in

M16 appear to be connected to the southeastern part ones in our new model (STaa' in Figure 3.10).

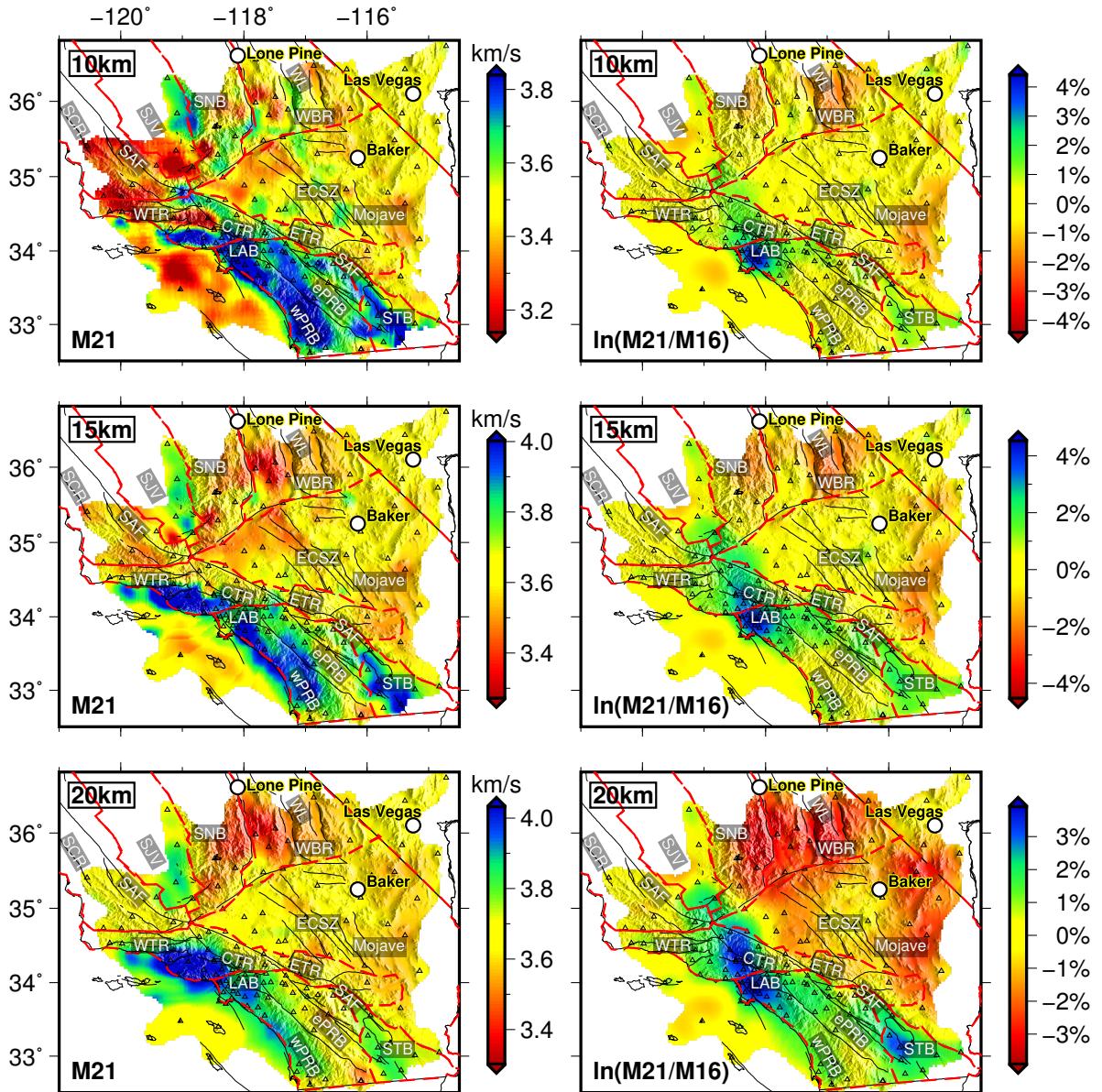


FIGURE 3.6: Horizontal slices of shear velocity of model **M21** (left panel), and **ln(M21/M16)** (right panel) at 10, 15 and 20 km depths.

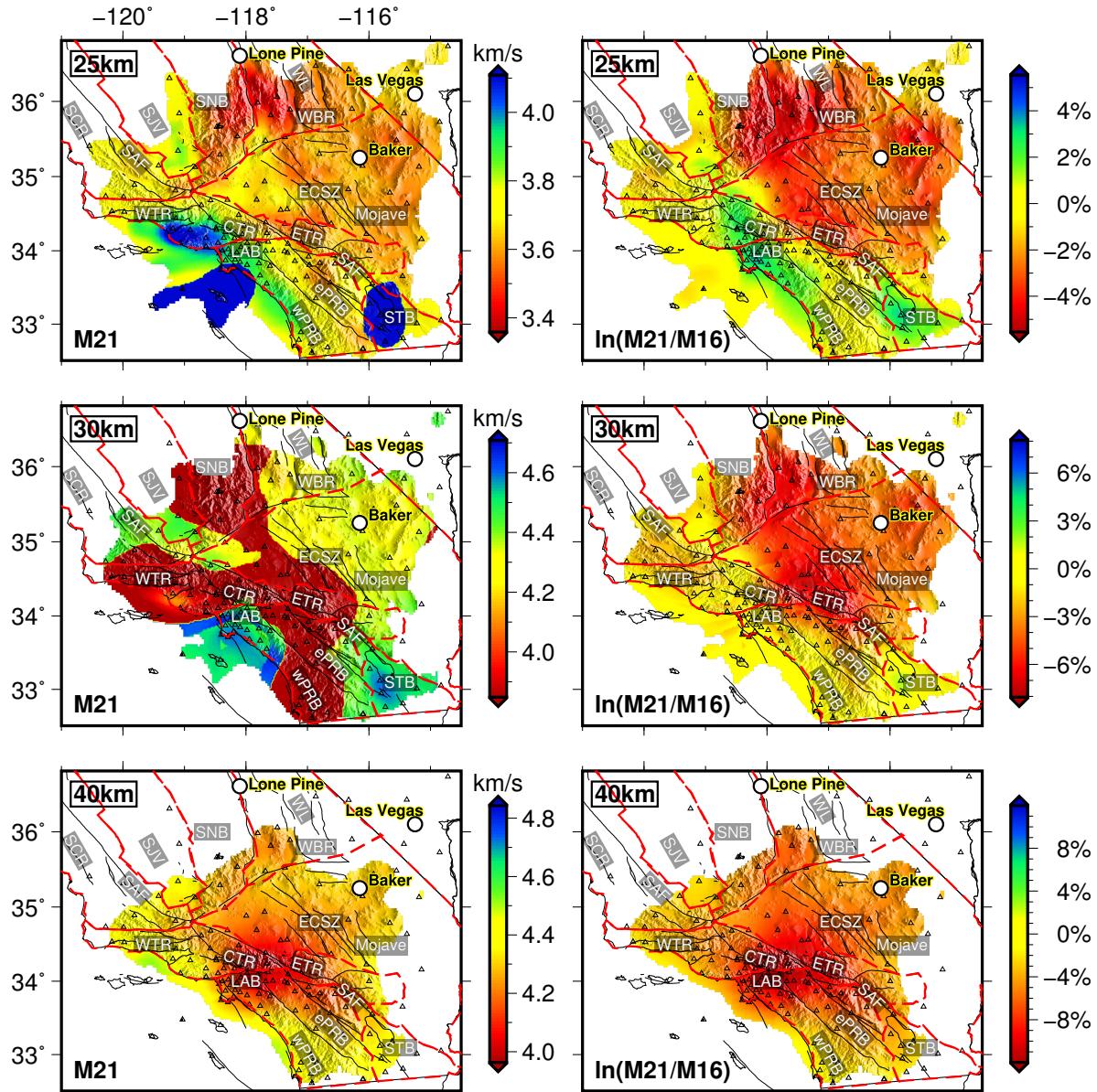


FIGURE 3.7: Same as Figure 3.6 but for 25, 30, and 40 km depths

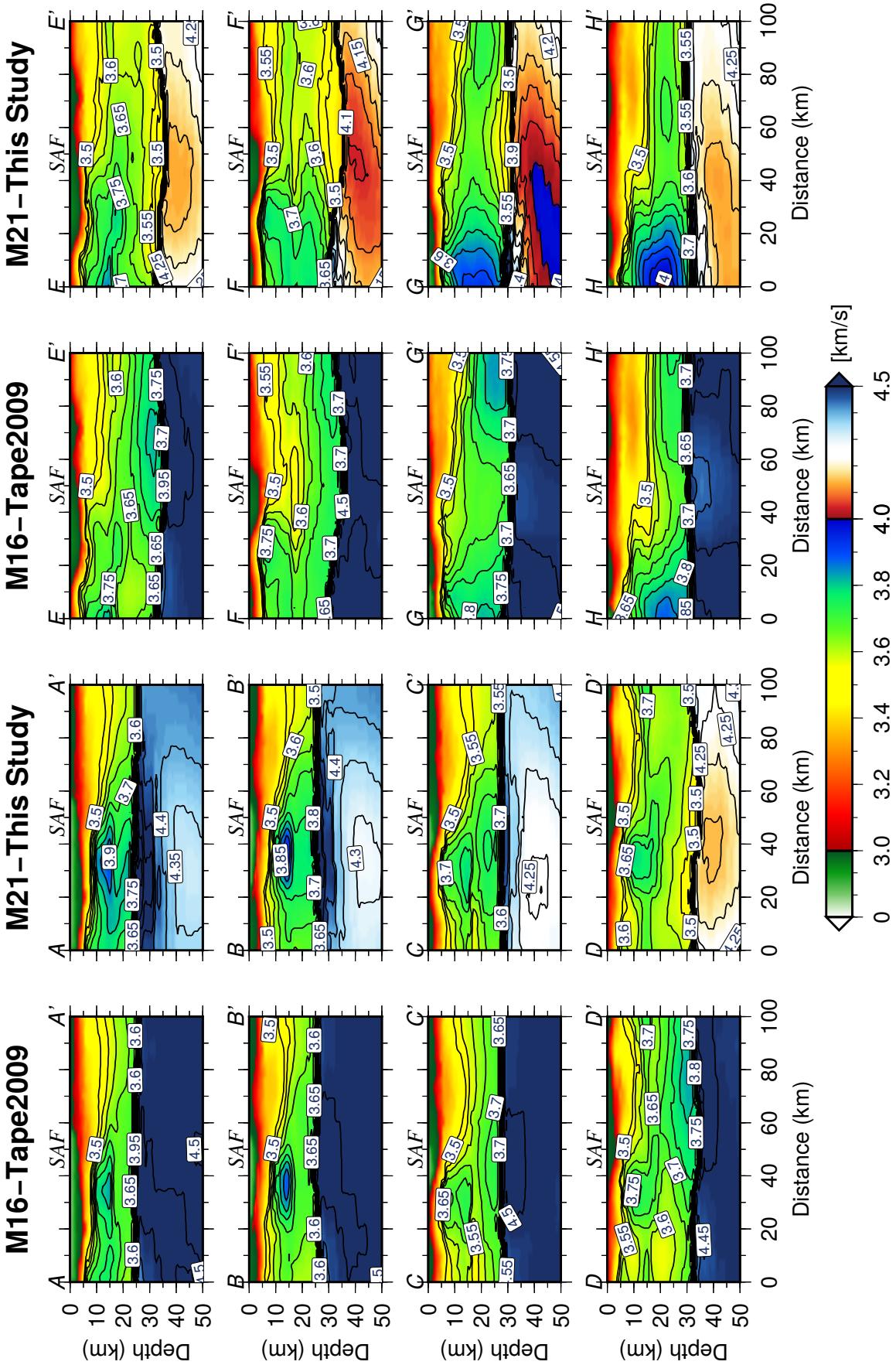


FIGURE 3.8: Cross-sections of shear velocity of model M16 and M21 across the San Andreas Faults with the profiles shown in Figure 1b.

3.7 Discussion

We discuss the robustness of our inversion results from three different aspects: (i) traveltime misfit evolution through histograms comparison; (ii) comparison of our 3-D model with three other models; (iii) inversion strategies towards future adjoint tomography. Additional information on waveform fitting, model uncertainties and resolution analysis are also provided in the supporting document (Figures S5-S7).

3.7.1 Misfit analysis

We first check the number of good measurements (N_{meas}) obtained from the initial model and updated models to see whether the number is improved by our inversion procedure. As stated in section 3.5, misfit measurements are strongly influenced by the dT , $d \ln A$, and CC_{min} criteria used in window selections. The dT and $d \ln A$ of the three bands are fixed in our inversion, while the CC_{min} is improved from 0.69 (**M16**) to 0.75 (**M17** to **M18**), and to 0.80 (**M19** to **M21**). To make reliable comparisons, we additionally measure cross-correlation travelttime misfits of the three bands for **M16** using $CC_{min} = 0.75$ (Figure 3.4(b)) when comparing the number between the stage 1 and 2 inversions and using $CC_{min} = 0.80$ (Figure 3.4(c)-(d)) between stage 1 and 3 inversions. For the 20 – 50 s band, the N_{meas} improves from 9670 to 10000 (Figure 3.4(b), right) and final 9504 to 9569 (Figure 3.4(d), right) in the three inversion stages, respectively. The same phenomena are also observed with the 10 – 20 s band where N_{meas} improves from the initial 10718 to 10772 (Figure 3.4(b), middle). However, it slightly decreases during the last inversion stage. In the third inversion stage for the additional 5 – 10 s band, the N_{meas} improves from the initial 8092 to final 8174 (Figure 3.4(d), left).

Then, we investigate the changes of travelttime misfits for the three updated models (**M17**, **M19**, **M21**) relative to those misfits for the initial model **M16**. The most significant improvements are certainly obtained in the long-period band of 20–50 s, for which travelttime misfits are significantly reduced with the mean and standard deviation (STD) values changing from the initial 2.12 ± 1.93 s to 0.41 ± 1.63 s (Figure 3.4(d), right). For the 10 – 20 s band, the travelttime misfit reductions are relatively small, with the mean and STD values reduced from 0.40 ± 0.86 s to -0.17 ± 0.68 s (Figure 3.4(d), middle). The one for the 5 – 10 s band shows almost no change in the timeshift distribution. Furthermore, we notice that the number of

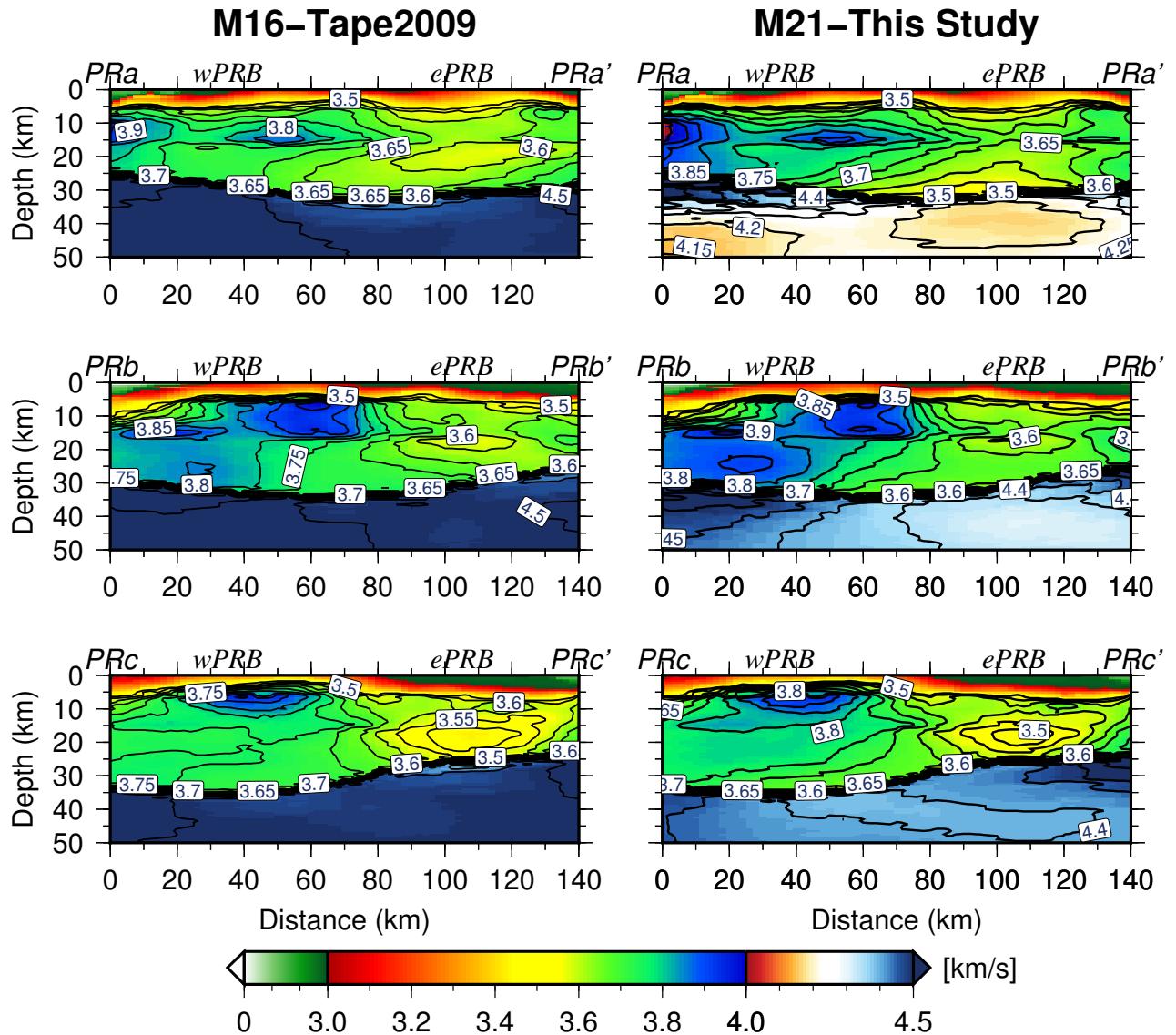


FIGURE 3.9: Cross-sections of shear velocity of model **M16** (left panel) and **M21** (right panel) across the Peninsular Range Batholith (PRaa'-PRcc') with the profiles shown in Figure 1b.

good measurements with $|\Delta T| \leq 1.0$ s increases in all three bands.

3.7.2 Model comparison

We now compare the initial model **M16** and our final model **M21** with two other models. The first is a model named CVM-S4.26 from Lee et al. [2014], which uses both earthquake data and ambient noise at 5 – 50 s, finite-difference forward modeling, and a combination of scattering-integral and adjoint methods for the tomographic inversion. The second model is Barak2015, constructed by traditional ambient noise tomography [Barak et al., 2015]. It is built by first obtaining group velocity maps at 4 – 40 s periods from 2-D surface wave tomography, and then averaging the median result of 13 inversions with different starting

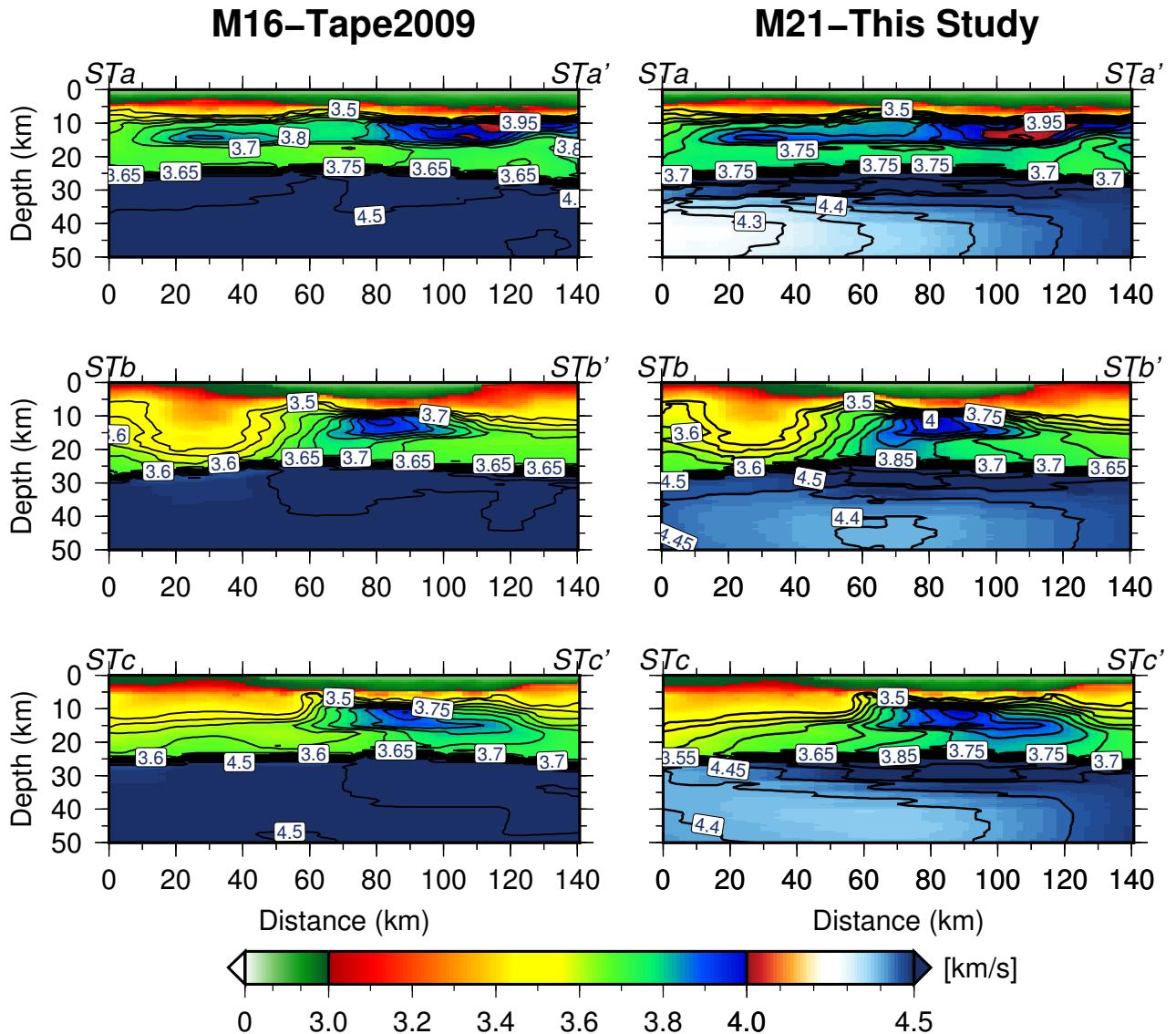


FIGURE 3.10: Cross-sections of shear velocity of **M16** (left panel) and **M21** across the Salton Through (STAa'-STcc') as the profiles shown in Figure 1b.

models that are partly based on **M16**. We compare our final model with the other models by examining a series of vertical cross-sections in the supporting document (Figures S8-S25).

In general, the CVM-S4.2.6 model contains more small-scale features than the other three models and has much higher velocities, perhaps due to the larger number of data used and many iterations performed. Based on the similar initial model and data set, our final model shows velocity variations that are generally in agreement with those of Barak2015: (1) lower-crust and upper-mantle wave speeds are significantly lower by global standards; (2) high-shear velocities are observed beneath the westernmost PRB and southeastern STB in the lower crust (Figures S9-S26).

However, our final model differs from the Barak2015 model in several aspects as

shown in Figure 3.11. At the central part of SAF (such as profiles DD'), our model exhibits a high velocity zone in the middle crust that dips to the northeast slightly, while the Barak2015 model shows that it dips into the lower crust. Moreover, the high velocity zone in the lower crust of wPRB in our model is not as evident as that shown in Barak2015 model and is only observed in the westernmost part (profile PRbb'). Our model shows that the southern boundary of wPRB and ePRB dips steeply towards the northeast into the lower crust, while it only dips slightly towards the northeast in the middle crust (PRbb') in Barak2015. In addition, our model shows increased velocities beneath LAB and CTR (profiles GG'-HH', Figures S15-S16), which are not identified in the Barak2015 model.

3.7.3 Inversion strategies

The differences among these four models (**M16**, **M21**, CVM-S4.26, Barak2015) are expected, given that different data and tomographic inversion techniques were used. Given the major efforts involved in simulation-based inversions (as in the first three models above), here we offer some perspectives for future efforts in seismic tomography.

The five iterations we conduct using ambient noise data in this study can be regarded as an extension to the earthquake-based adjoint tomography by Tape [2009]. Using the same underlying technique as used to build **M16**—spectral-element and adjoint methods—we demonstrate that surface waves from ambient noise data are very helpful for resolving lower crustal and upper mantle structures.

The CVM-S4.26 model of Lee et al. [2014] began its inversion with ambient noise data, then added shorter-period earthquake data after several iterations. This sequence has the advantage that earthquake sources can be introduced into the inversion after the velocity structure has been improved.

If all techniques and data are available, we would advocate the following steps sequentially: (1) classical tomography using ambient noise, (2) adjoint-based ambient noise tomography, (3) earthquake-based adjoint tomography. To our knowledge, this sequence has not yet been performed by the same research group at any scale.

Southern California provides the challenge and opportunity for using additional data sets to improve the seismic velocity models. For example, significant improvements to basin models and the Moho model were made following the completion of **M16** [CVM-H 15.1, Shaw et al., 2015]. These detailed improvements are not necessarily compatible with the

larger scale structures sampled by earthquake and ambient noise data, and therefore further tomographic iterations are warranted.

Looking to the future, the prevalence of adjoint tomography will likely increase with the improvement of workflows [Bozdağ et al., 2016, Krischer et al., 2015, Modrak and Tromp, 2016] and the wider availability of computational resources for these problems.

3.8 Conclusions

In this work, we apply the adjoint tomography [Tromp et al., 2005] to EGFs constructed from ambient noise data in southern California in order to iteratively improve the 3-D tomographic model **M16** constructed based on earthquake data [Tape et al., 2010]. EGFs from ambient noise are compared with SGFs from numerical simulation base on SEM. Traveltime misfits between EGFs and SGFs are measured at three period bands, 5–10 s, 10–20 s, and 20–50 s. The EGFs from ambient noise show a good agreement with the SGFs from **M16** at 5 – 10 s and 10–20 s, while exhibiting an average 2.12 s time advance at 20–50 s. By minimizing the traveltimes differences between EGFs and SGFs through local-gradient-based optimization algorithm, we obtain a final model **M21** with better data fits compared with the initial model **M16**.

We start our inversion with only data in the 20 – 50 s band and then progressively introduce the other two shorter-period bands. After five iterations, the averaged total misfit combining the three bands is reduced from 1.75 to 0.41. The final model shows features generally in agreement with the initial model, especially at the shallowest depth (≤ 5 km). However, the new Vs model reveals several new features in the middle and lower crust, including: (1) the mean speed of lower crust is slowed down by about 6%; (2) higher Vs anomalies (up to +4%) are observed in the LAB and CTR throughout the crust; (3) higher Vs anomalies are seen in the lower crust beneath the westernmost PRB; (4) an enhanced shallow high velocity zone in the middle crust is observed beneath STB. Our model also shows refined lateral velocity gradients across PRB, SNB, SAF, which provides constraints on the west-east compositional boundary of PRB, SNB, as well as the dip angle and the depth extent of SAF.

As ambient noise cross-correlations are available between any station pairs, the new

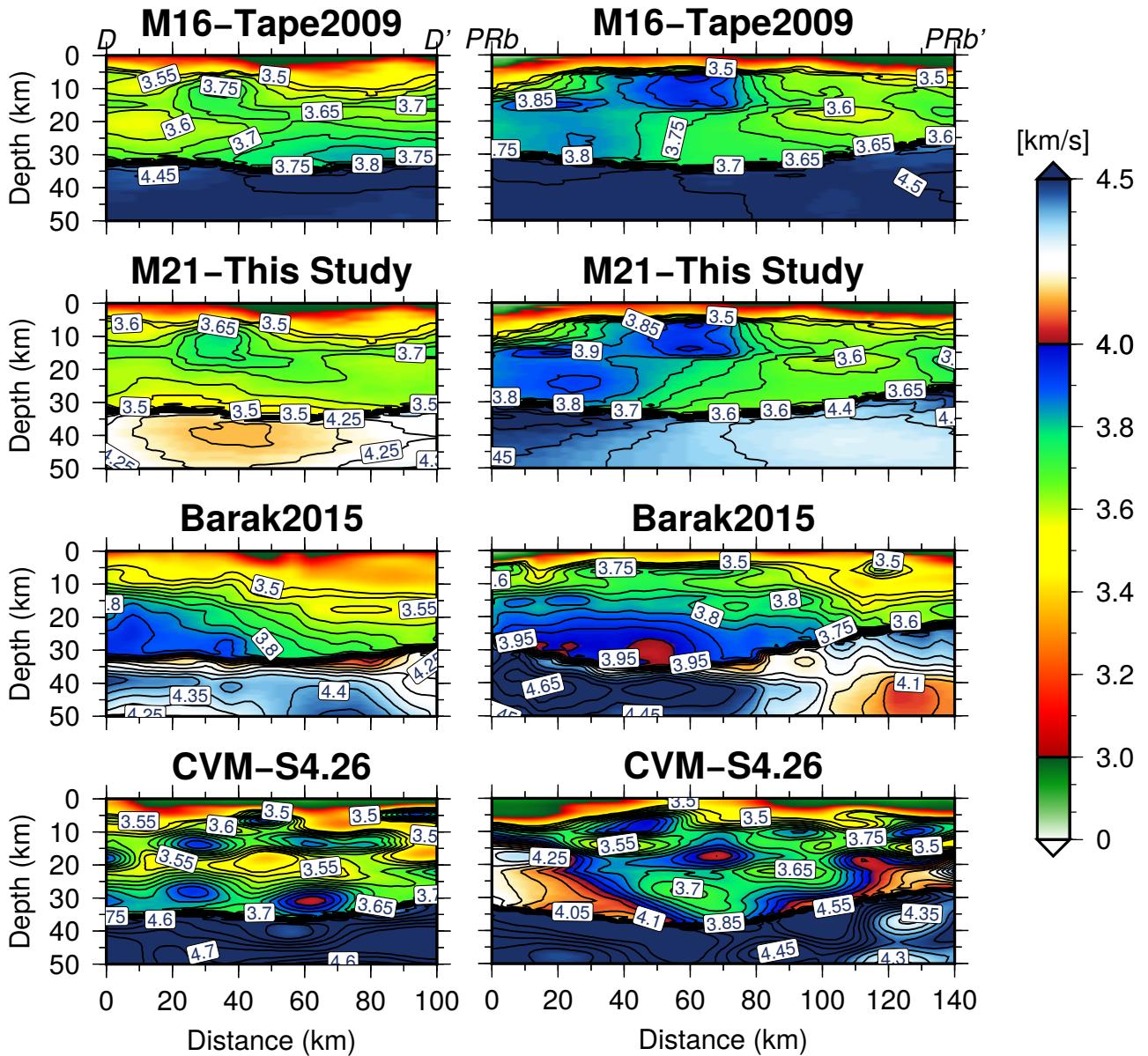


FIGURE 3.11: Comparison of model **M16**, **M21**, Barak2015 and CVM-S4.26 along profile DD' across SAF and profile PRbb' across PRB. See more results in supplementary materials

tomographic model provides a better resolution relative to the initial one in areas not well-covered by event-station paths. Also owing to the longer-period features of ambient noise than the seismic data from local earthquakes, the lower-crust is better illuminated. The total misfit reduction, and a series of point-source resolution tests (Figure S7) all support the robustness of our new tomographic model.

3.9 Appendix A: Line search and model update

We update the model $\delta m = (\delta\alpha, \delta\beta, \delta\rho)$ based on a search direction \mathbf{d} defined as the negative preconditioned and smoothed misfit gradient

$$\Delta \mathbf{m} = \alpha \mathbf{d} = \alpha \left(-\mathbf{P}^{-1} \mathbf{g} \right), \quad (3.13)$$

where α represents the step length, \mathbf{g} denotes the misfit gradient, and \mathbf{P} is the preconditioner. As surface waves from ambient noise data are not very sensitive to density variations, we scale the perturbation of density to the perturbation of shear wave velocity by a factor of 0.33 [Montagner and Anderson, 1989]. A few trial models can be built for a series of α values, and their corresponding total misfits are calculated through forward simulations, from which the optimal step length can be determined. However, this involves N_s number of forward simulations for each α value and therefore is computationally very expensive. Instead, a representative subset of 19 master stations that evenly cover the surface region (red triangles in Figure 3.1) are selected to do the line search. We have tested some other subsets of stations and found that including a number of stations in the LAB area helps better represent the total misfit.

Figures 3.A1(a-e) shows the line search results at each iteration, with the corresponding optimal step lengths listed in Table 3.1. We add 10 – 20 s and 5 – 10 s bands at the second (2ndIT) and fourth (4thIT) iteration when the averaged total misfit at longer-period bands changes little. When multiple frequency bands are included, we usually choose the optimal step length from the minimum of the averaged total misfit curve (solid lines in Figure 3.A1a-e). This sometimes results in slight increase of misfit for the band(s) that have already been reduced to local minimum in previous iterations. For example, at the second iteration, the total misfit for 20 – 50 s band increases slightly when we introduce 10 – 20 s band (Figure 3.A1b). At the third iteration, we choose a more conservative step length of 0.03 instead of the optimal one at 0.05 from averaged total misfit curve to ensure the total misfit for 10 – 20 s drops.

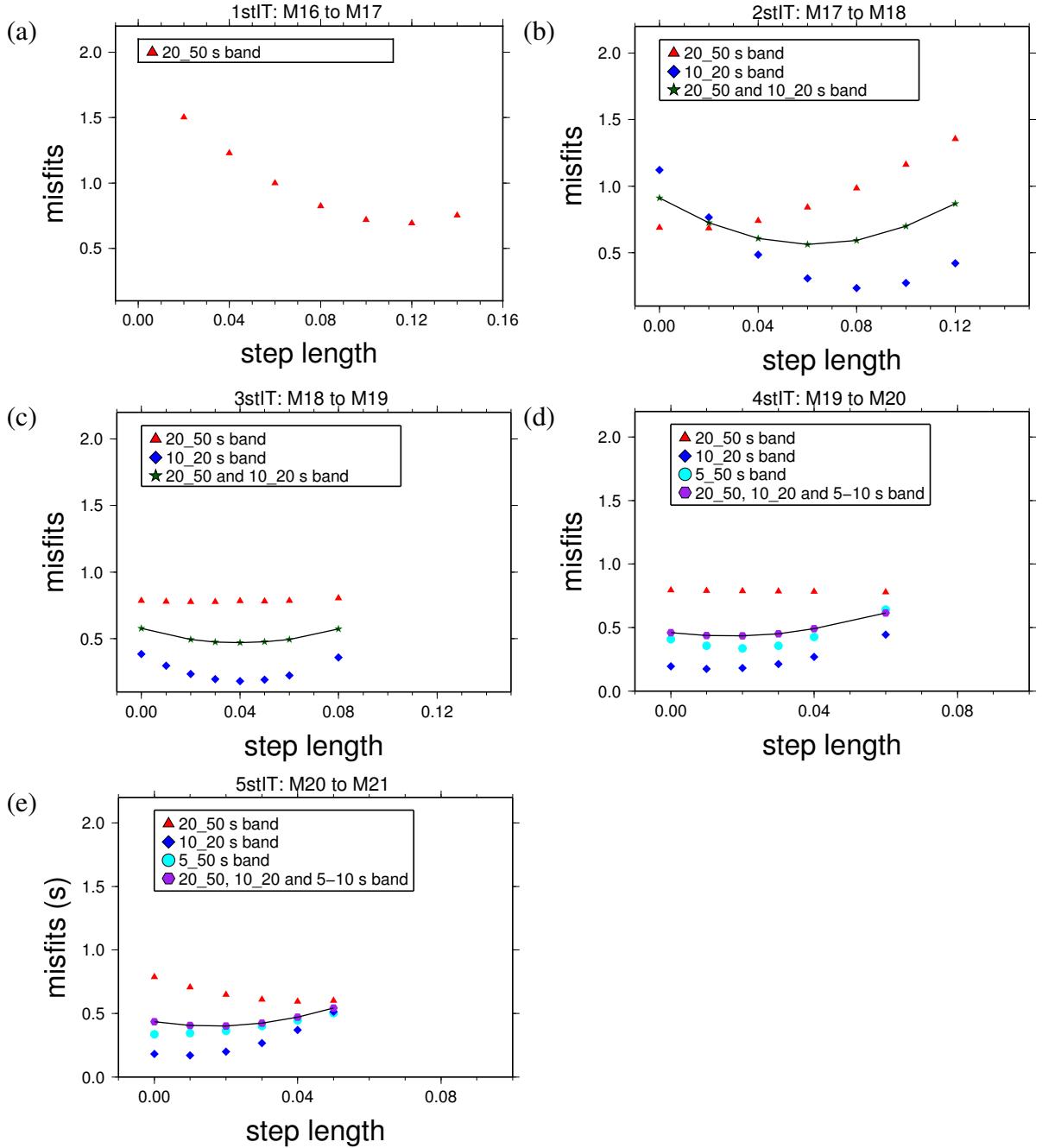


FIGURE 3.A1: (a-e) Line search results at each iteration, showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves).

3.10 Supplementary materials

This supporting document contains: (1) S1-S4: inversion results using broadband data filtered at 5 – 50, 10 – 50, 20 – 50 s bands. Figure S1 shows the line search results, and total misfit reduction using broadband data and the corresponding misfit evolutions are shown in Figure S2. Figures S3-S4 show the model comparison between using narrow-band and broadband data. (2) S5-S7: additional information on waveform fits (S5), volumetric sensitivity kernels (S6) and point source resolution tests (S7); (3) S8-S25: comparisons of model **M16**, **M21**, Barak2015 and CVM-S4 .26 at nine profiles along SAF (AA'-II'), three profiles at PRB (PRBaa'-cc'), three profiles across ST (STaa-'STcc') and three profiles across SNB (SNaa'-SNcc').

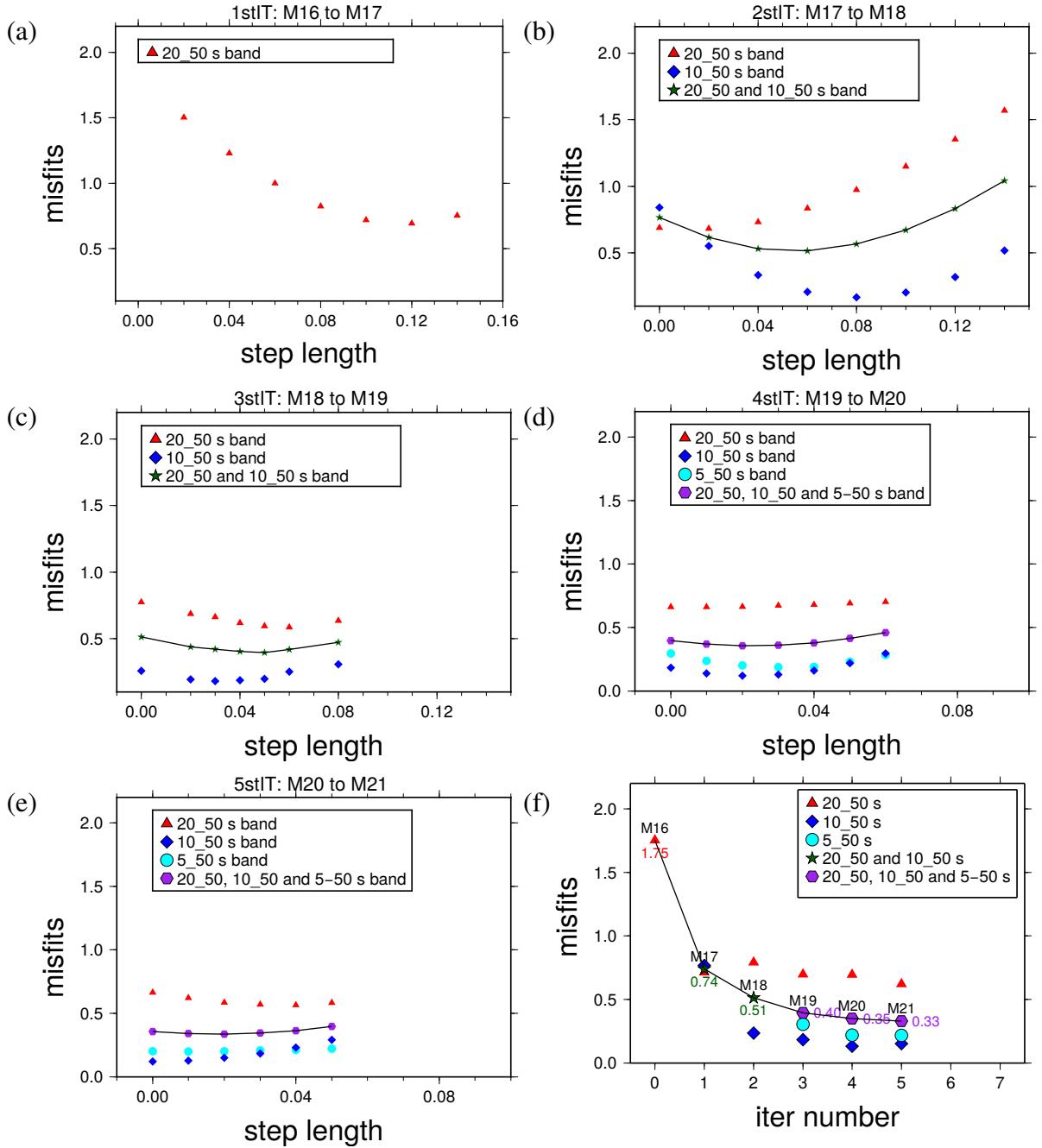


FIGURE 3.S1: (a-e) Line search results at each iteration showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves);(f) Total misfit reduction over iterations.

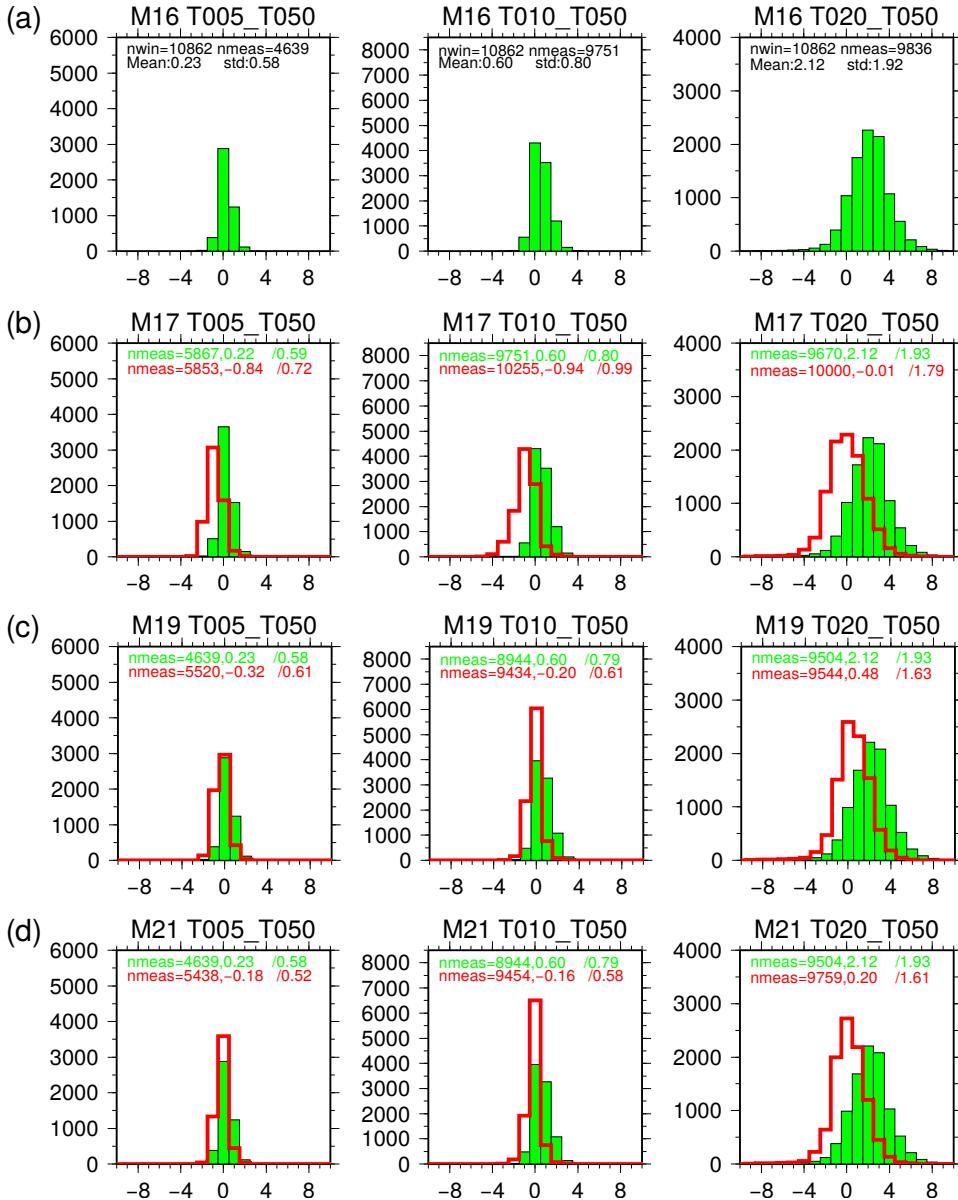


FIGURE 3.S2: Histograms of traveltimes misfits obtained for model **M16** (a), **M17** (b), **M19** (c) and **M21** (d). The green solid bars represent misfits for the initial model, and the red bars in (b-d) are misfits for updated models through iterations. Misfits are measured at all three period bands, 5 – 50 (left), 10 – 50 (middle), and 20 – 50 (right) s. CC_{min} in (a) are specially chosen as 0.80 (5 – 50 s), 0.75 (10 – 50 s), 0.69 (20 – 50 s) to show the misfit of the initial model. All the misfits are measured with $CC_{min} = 0.75$ in (b) and $CC_{min} = 0.80$ in (c) and (d) for comparison. Note that 5 – 50 and 10 – 50 s bands of **M16**, and 5 – 50 s band of **M17** are not used in the inversion.

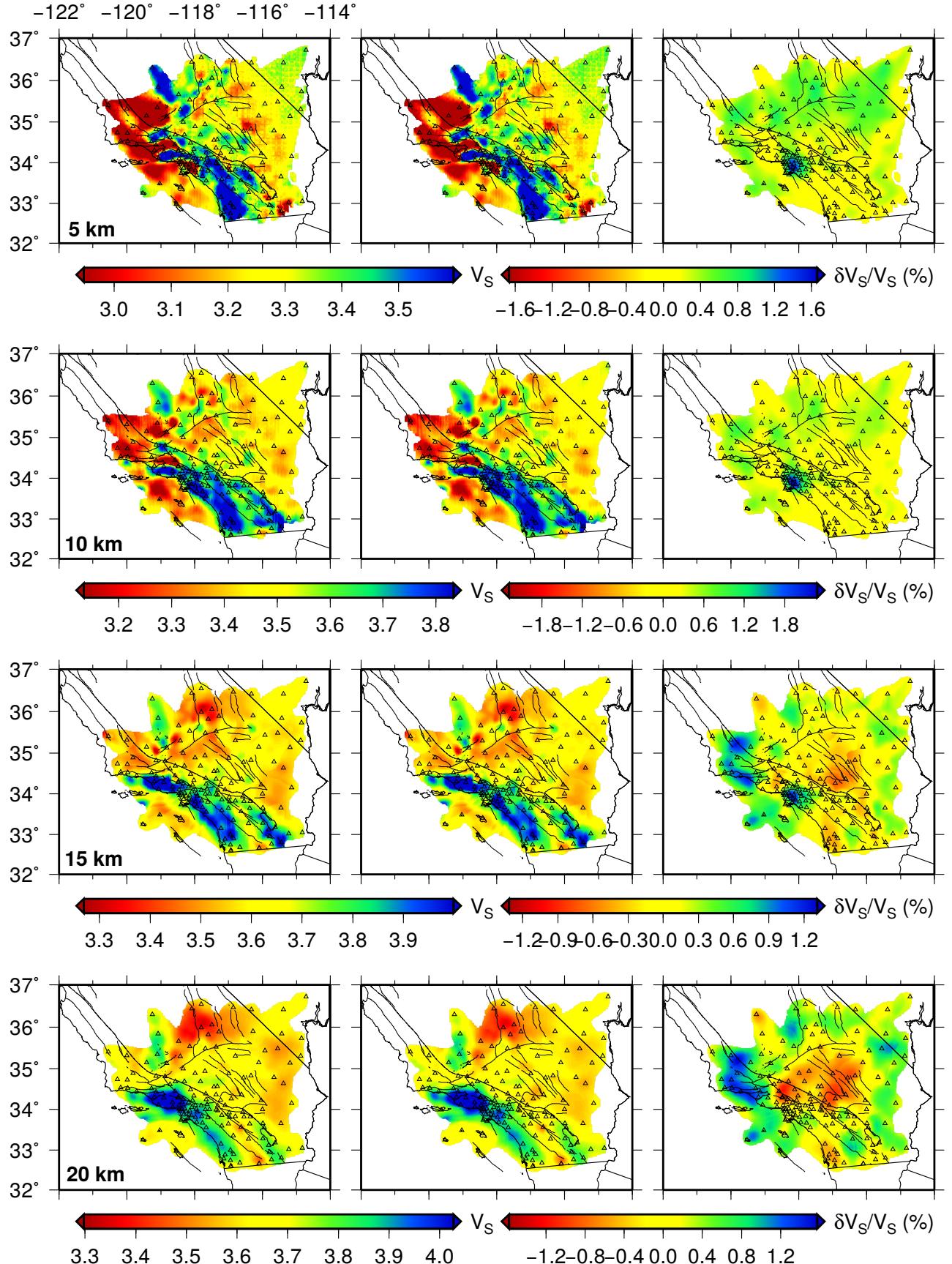


FIGURE 3.S3: Horizontal slices of shear velocity for model **M21** (left panel) using broadband data, **M21.narrowband** (middle panel) using narrowband data, and their differences (right panel) in percentage at 5, 10, 15, and 20 km depths.

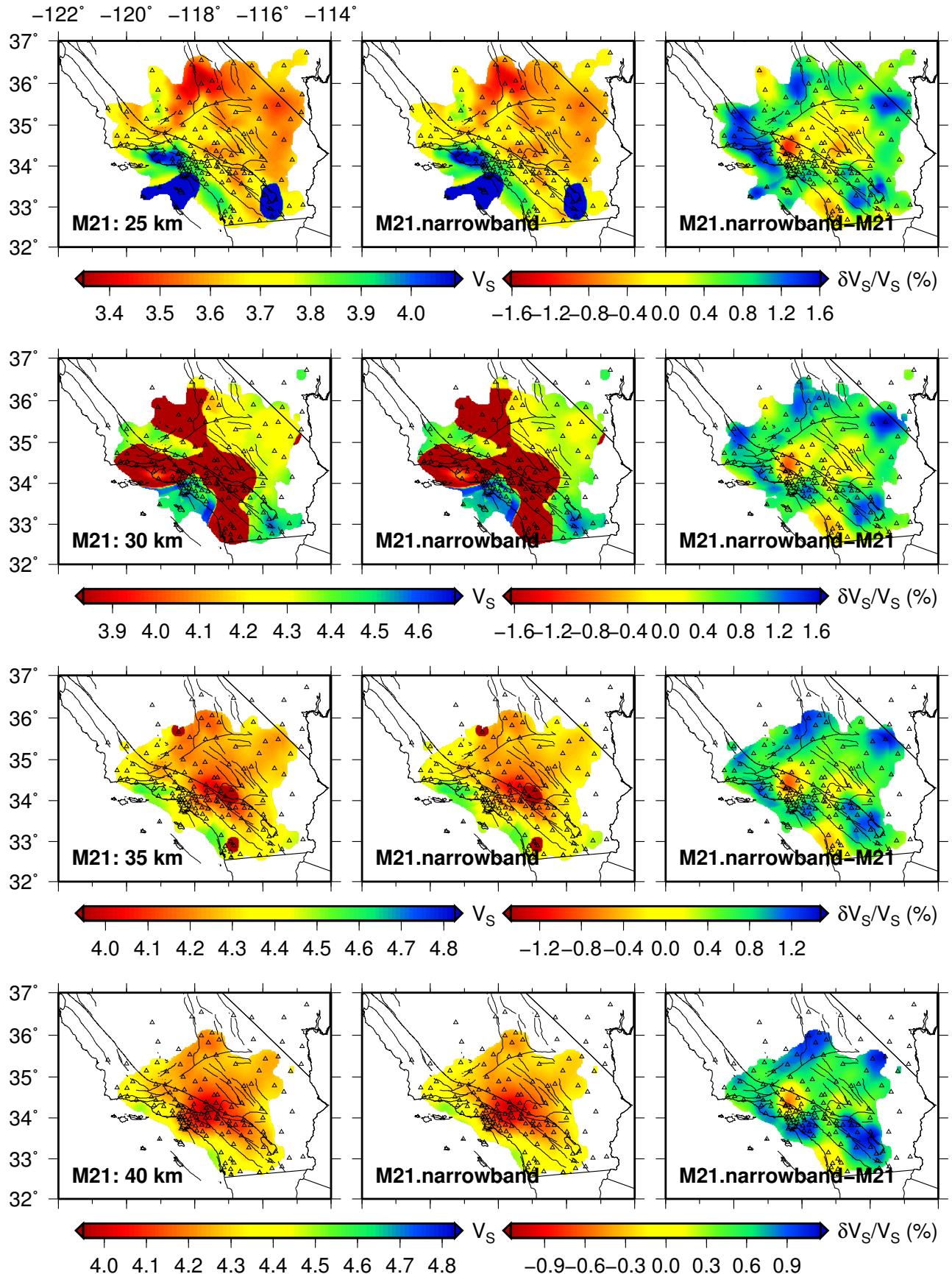


FIGURE 3.S4: Same as 3.S3 but for 25, 30, 35, 40 km depths.

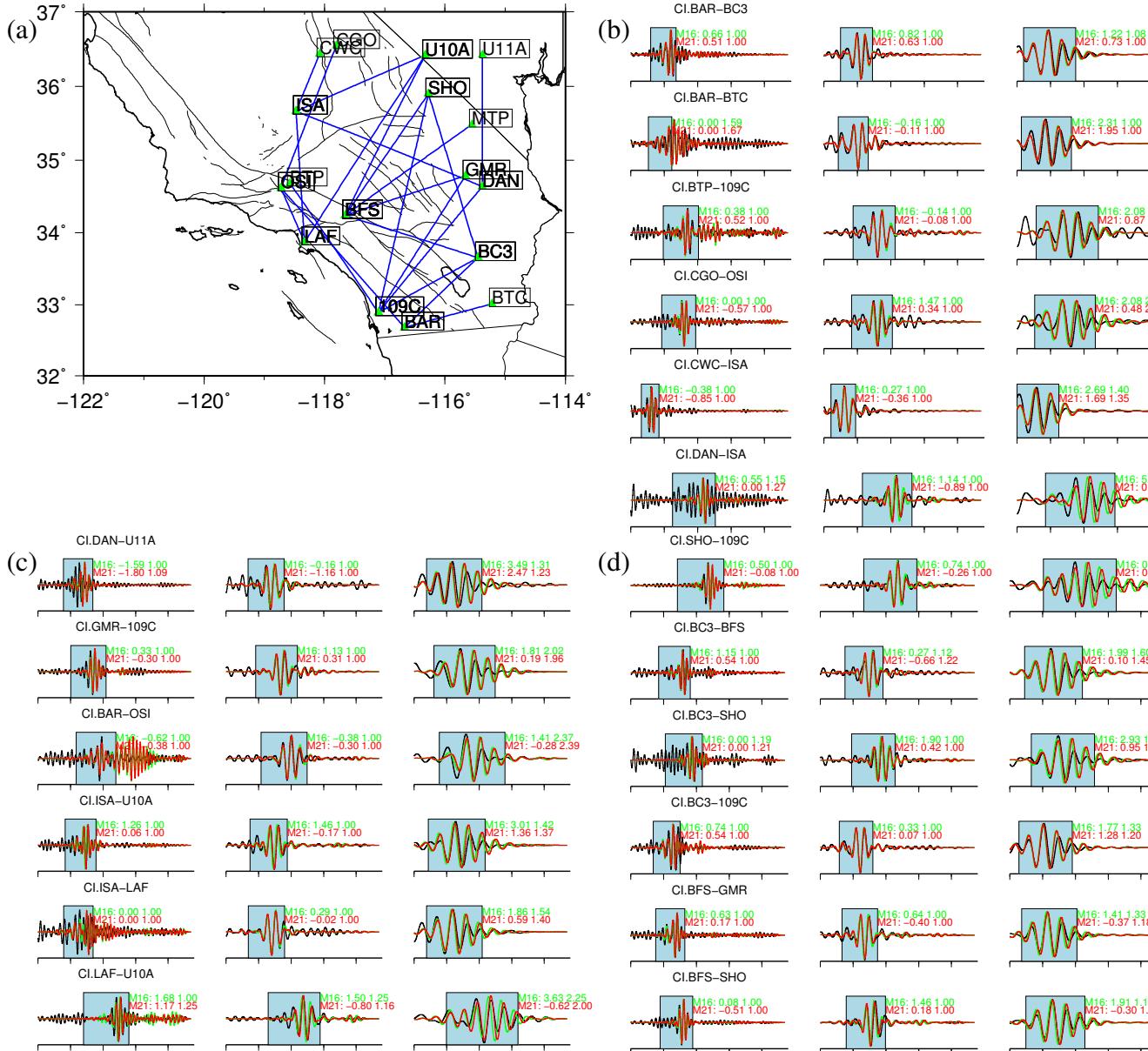


FIGURE 3.S5: (a) Selected ray paths for waveform comparison of some station pairs; (b)-(d) waveforms of station pairs filtered at 5 – 50 (1st column), 10 – 50 (2nd column), and 20 – 50 (3rd column) s. The EGFs are displayed as black solid lines, and SGFs are plotted as green solid lines for **M16** and red solid lines for **M21**. Traveltime misfits and uncertainties are also shown with corresponding colors.

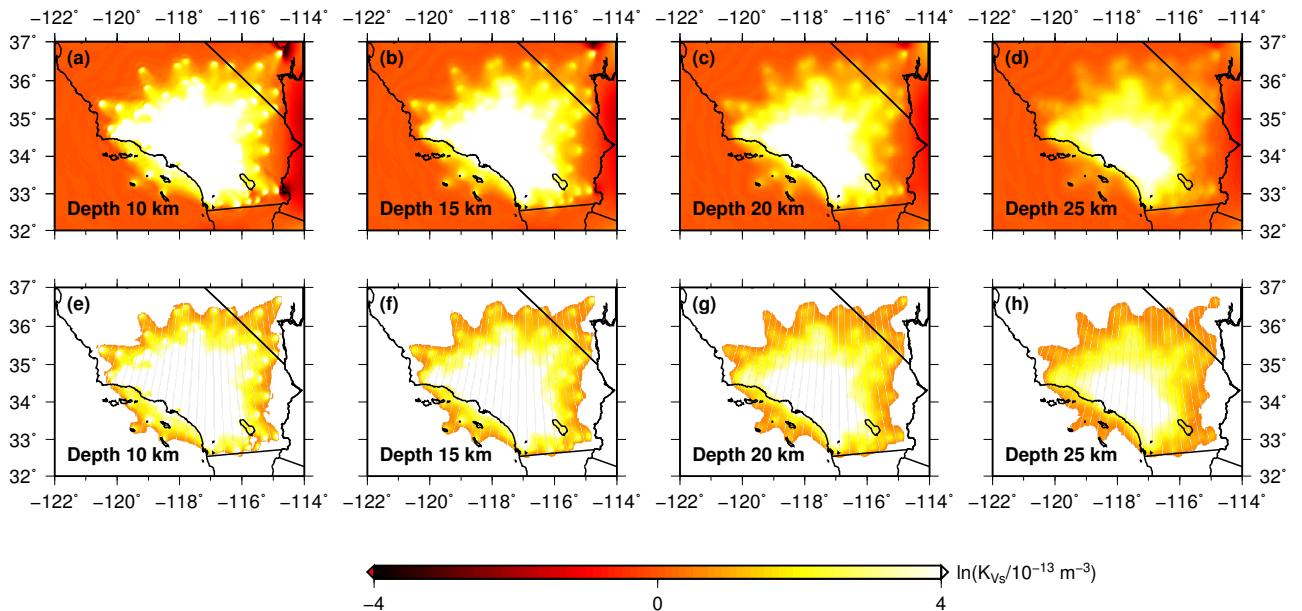


FIGURE 3.S6: (a)-(d) Volumetric sensitivities for Vs tomographic models plotted at 10, 15, 20, 25 km depths. (e)-(h) Masked volumetric sensitivities determined by a subjective threshold value of $K = 4.0 \times 10^{-13} m^{-3}$.

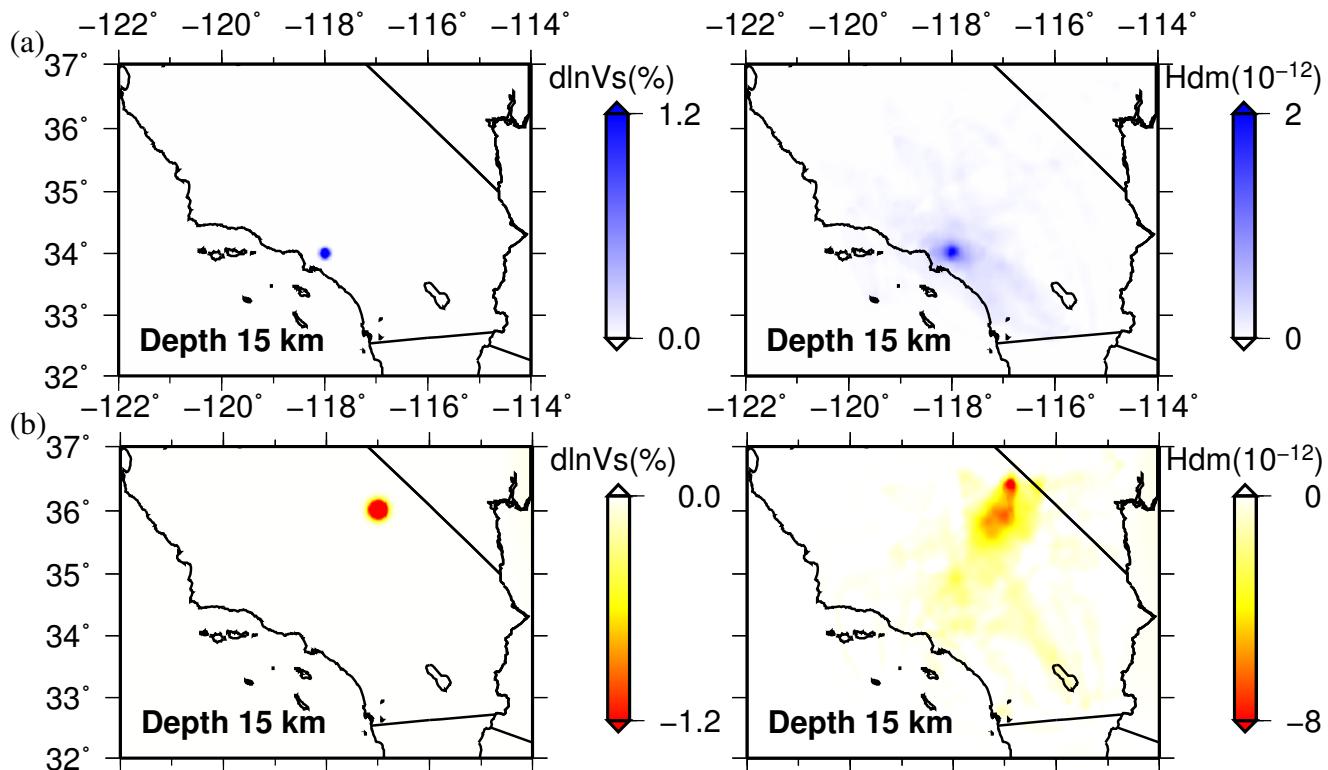


FIGURE 3.S7: Model resolution tests. The left panels show the locations of the fast/slow shear velocity anomalies at the depth of 15 km: (a) one with a 5 km radius put at the LA basin and (b) the other with a 10 km radius at north-eastern part region. The right panels are the corresponding Hessian kernels showing the degree of blurring in recovering the model perturbations

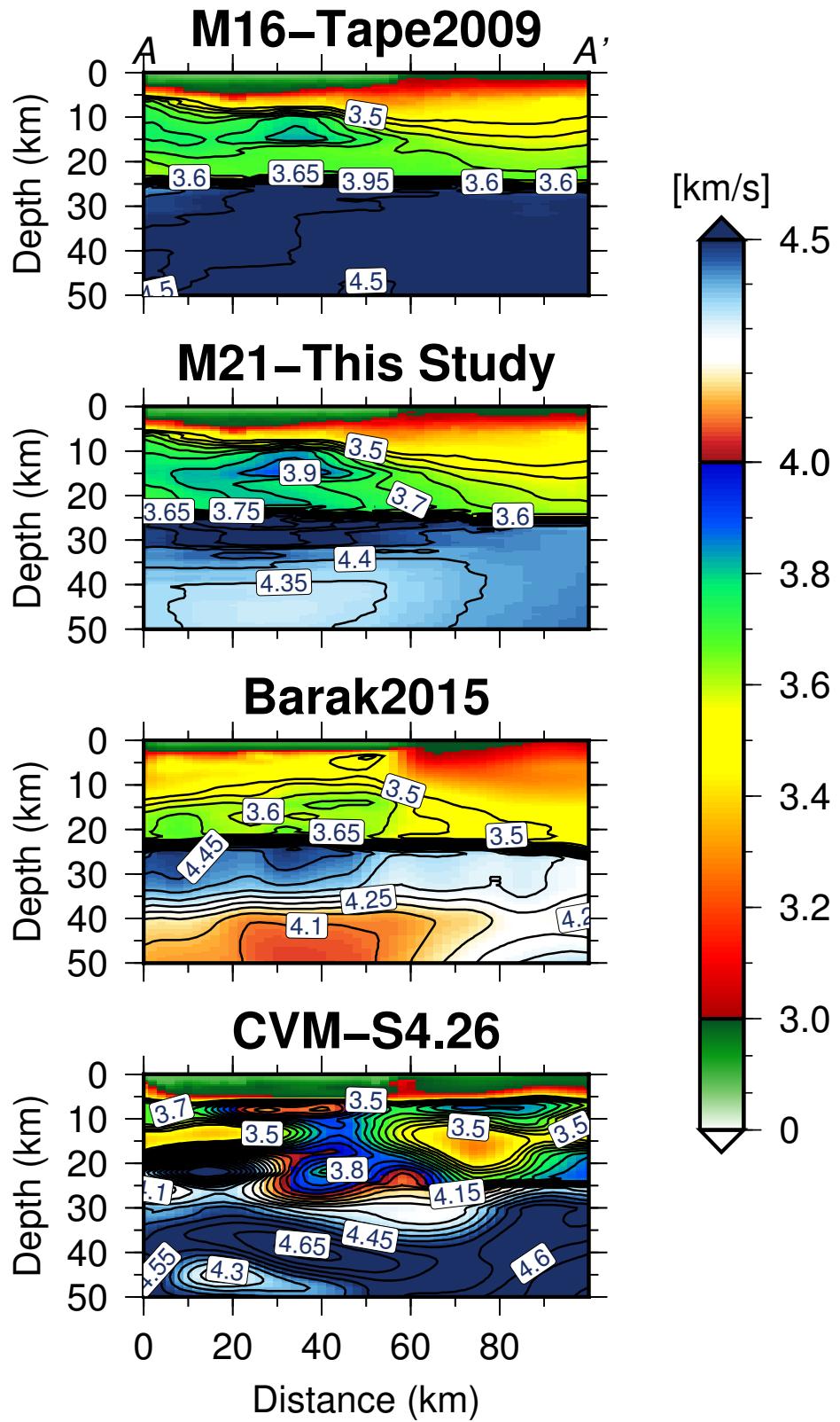


FIGURE 3.S8: Comparison of **M16**, **M21**, Barak2015 and CVM-S4.26 along profile AA'

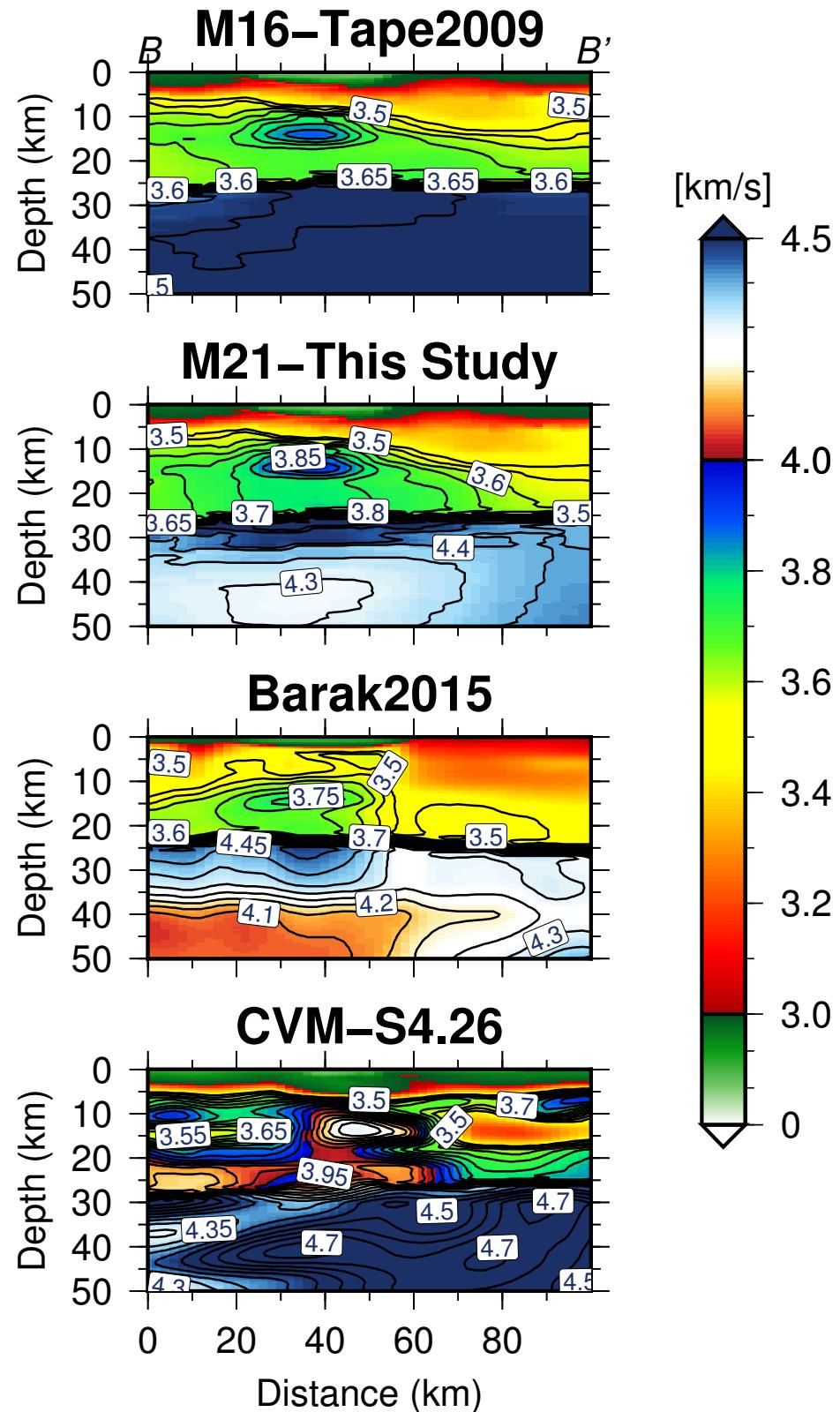


FIGURE 3.S9: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile BB'

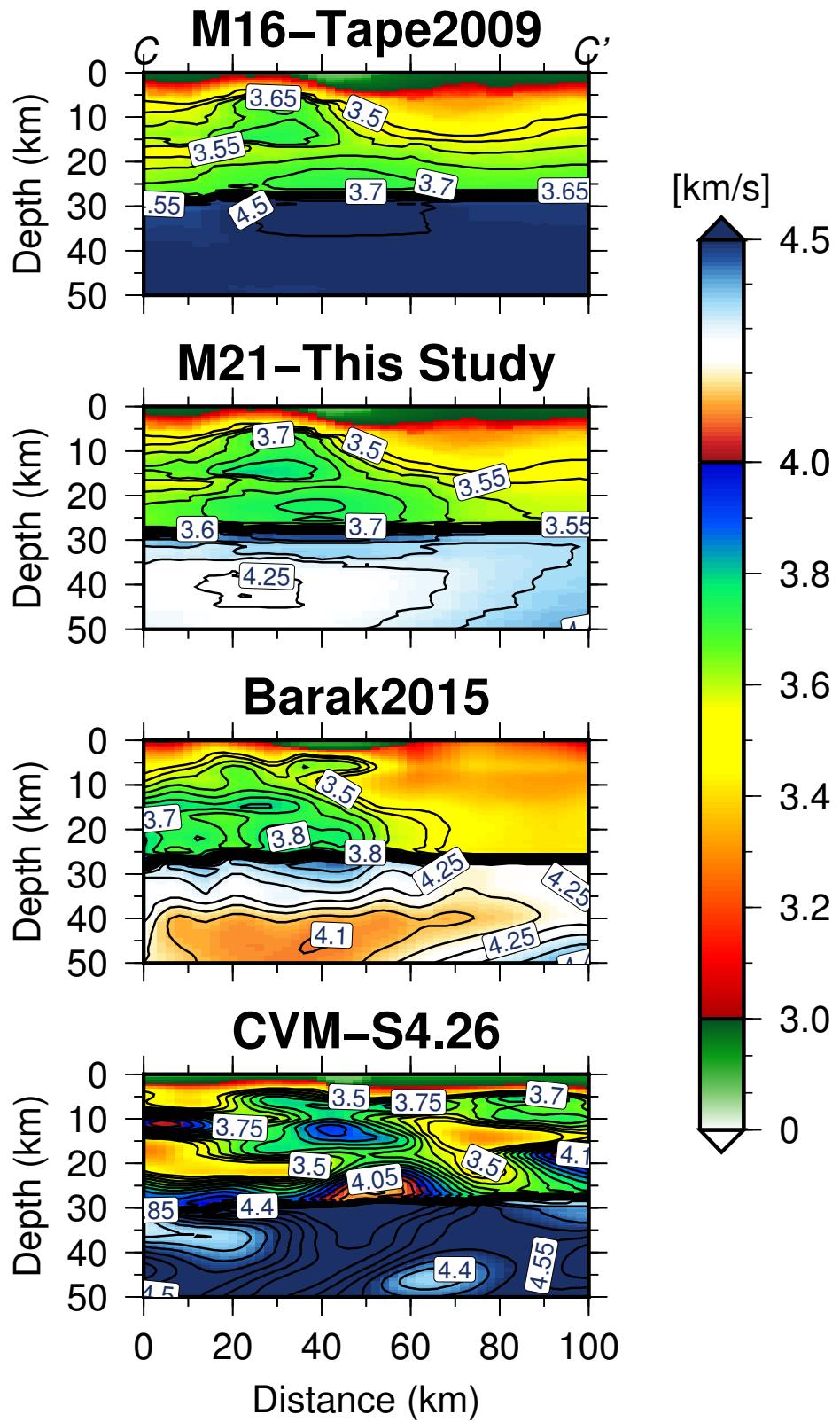


FIGURE 3.S10: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile CC'

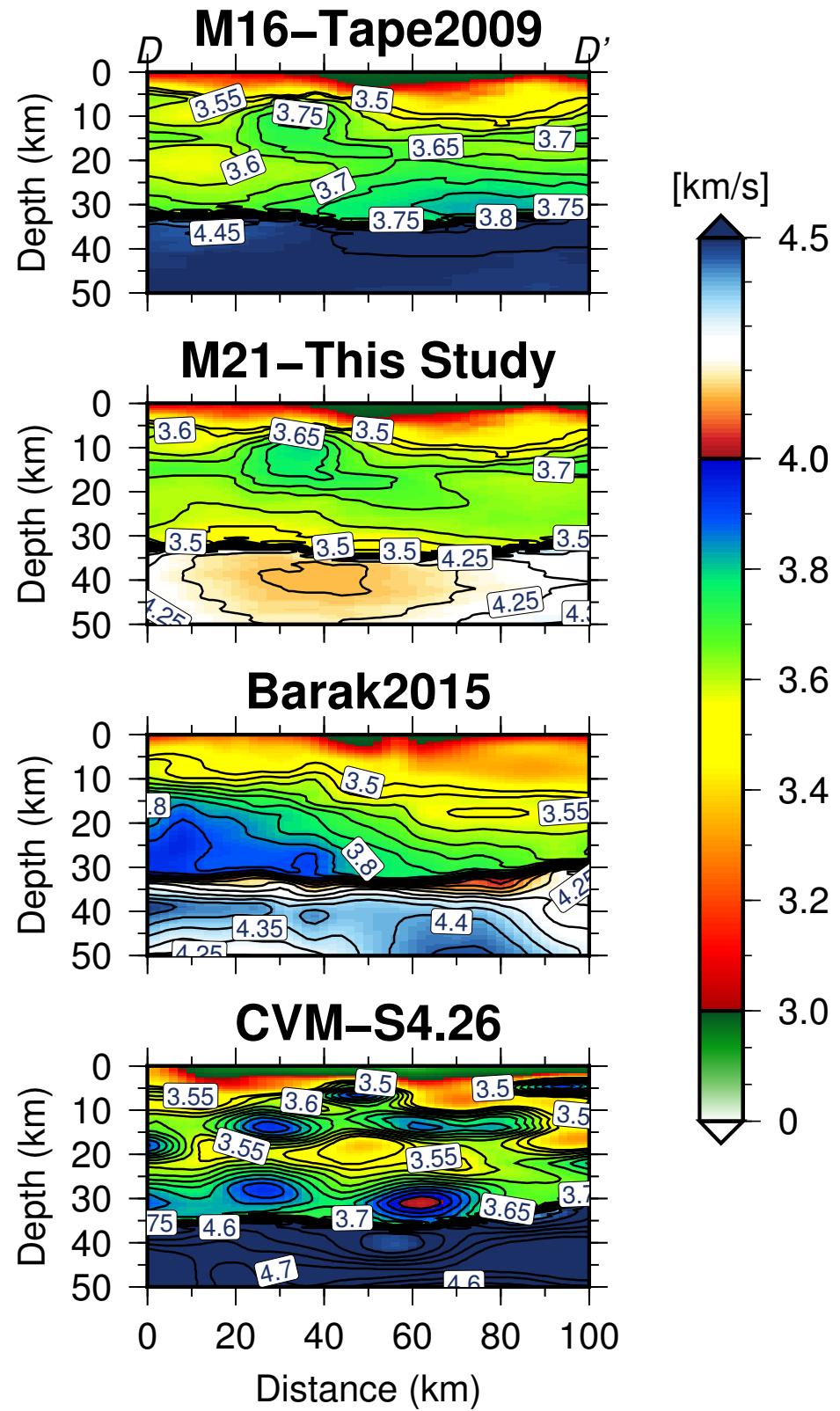


FIGURE 3.S11: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile DD'

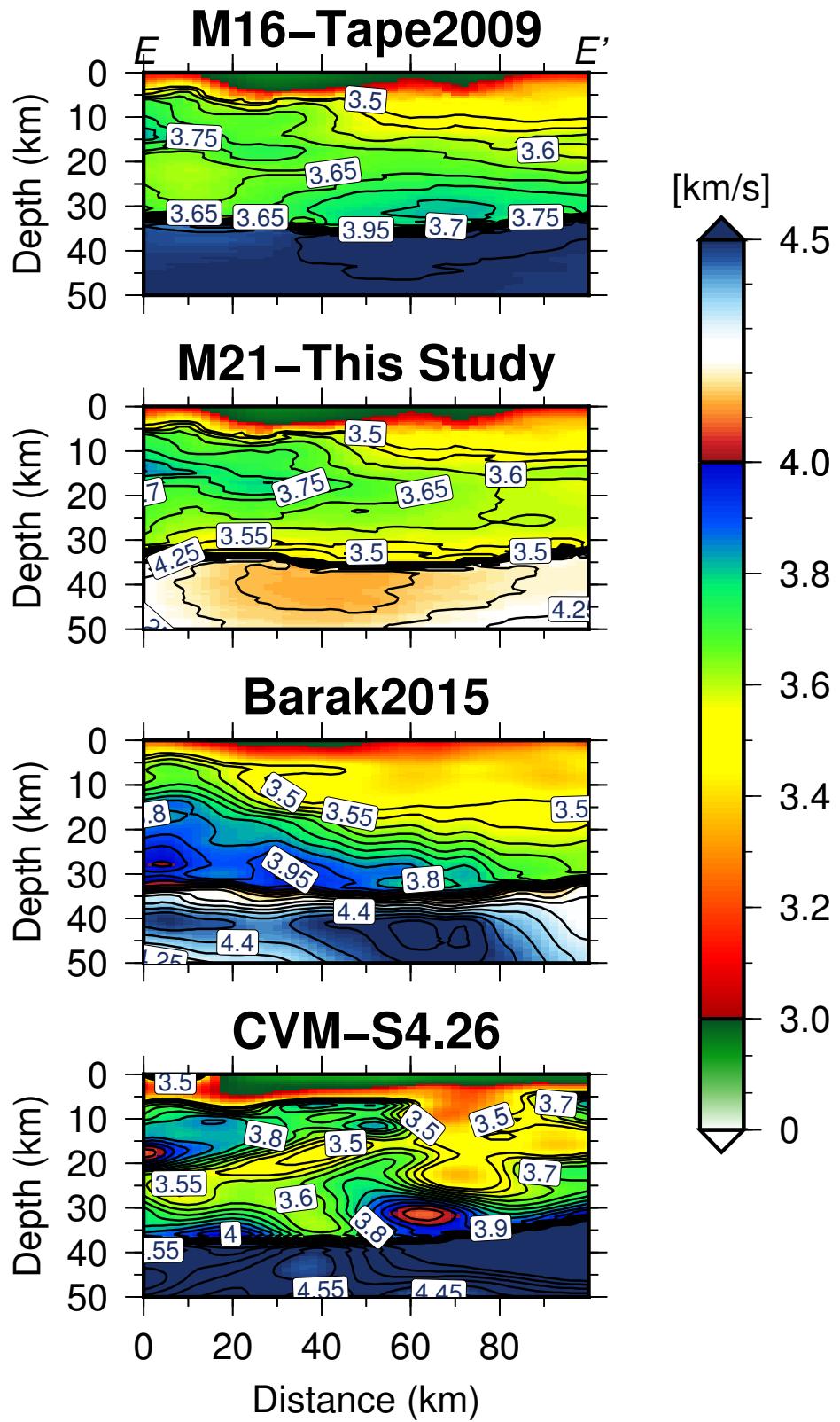


FIGURE 3.S12: Comparison of **M16**, **M21**, Barak2015 and CVM-S4.26 along profile EE'

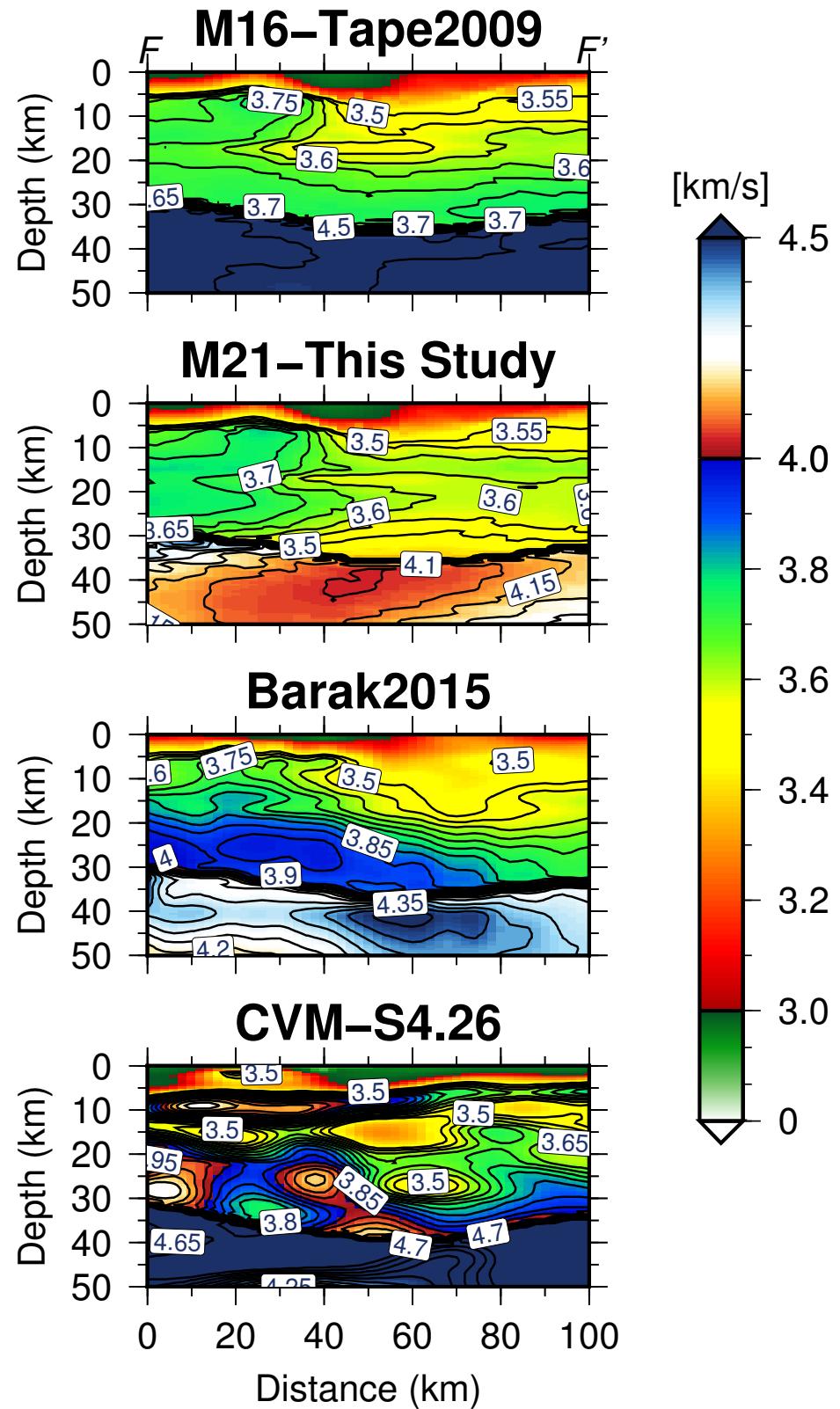


FIGURE 3.S13: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile FF'

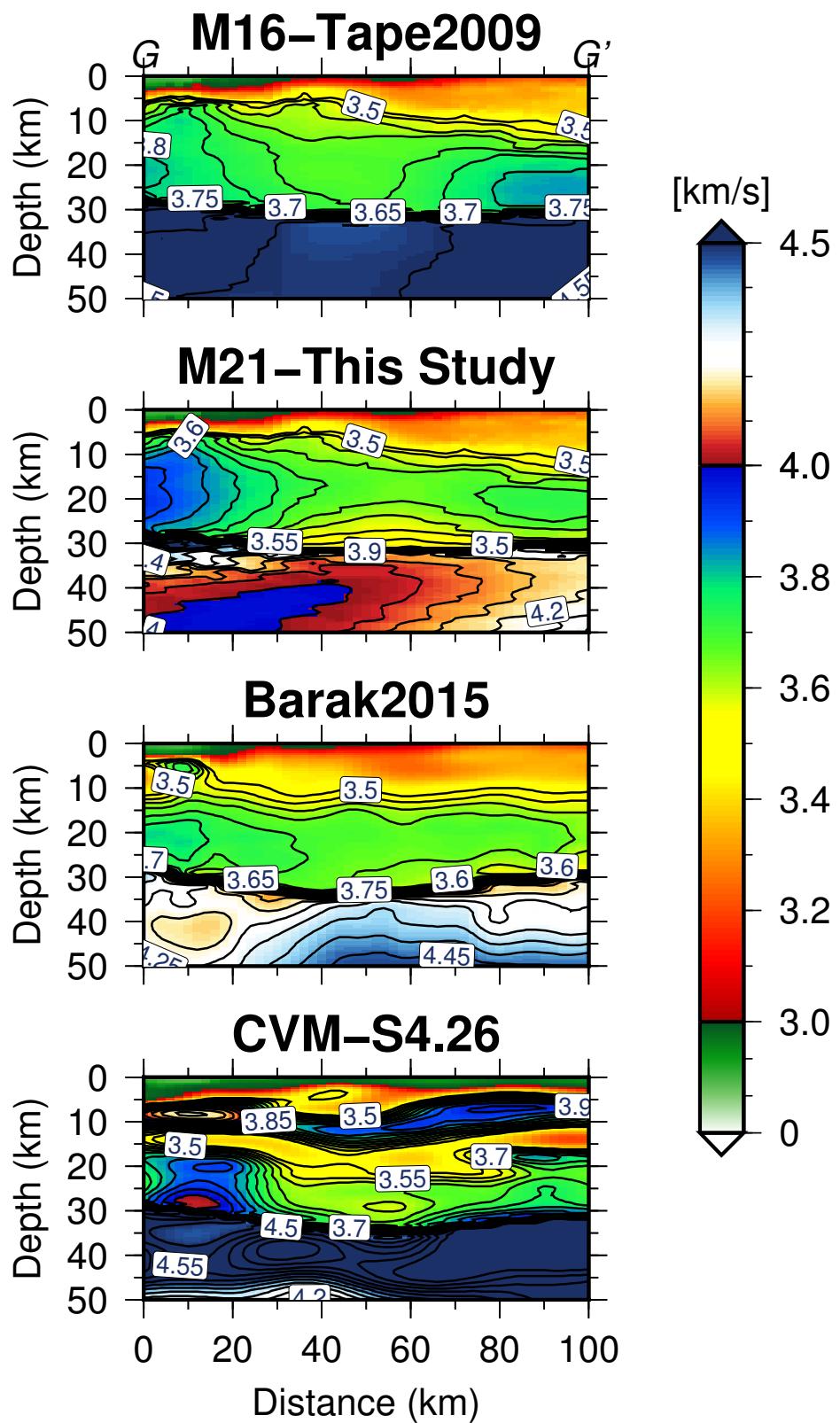


FIGURE 3.S14: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile GG'

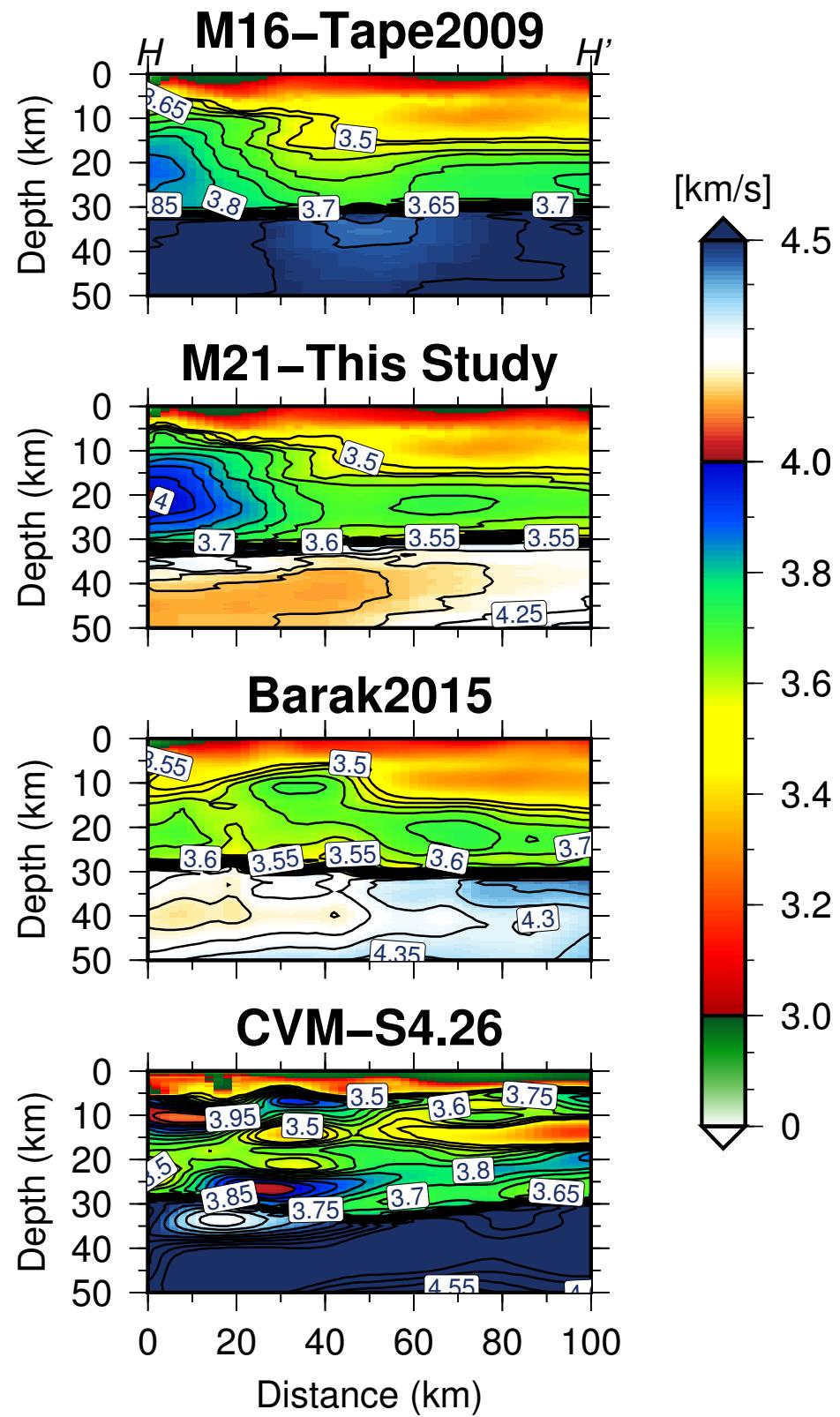


FIGURE 3.S15: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile HH'

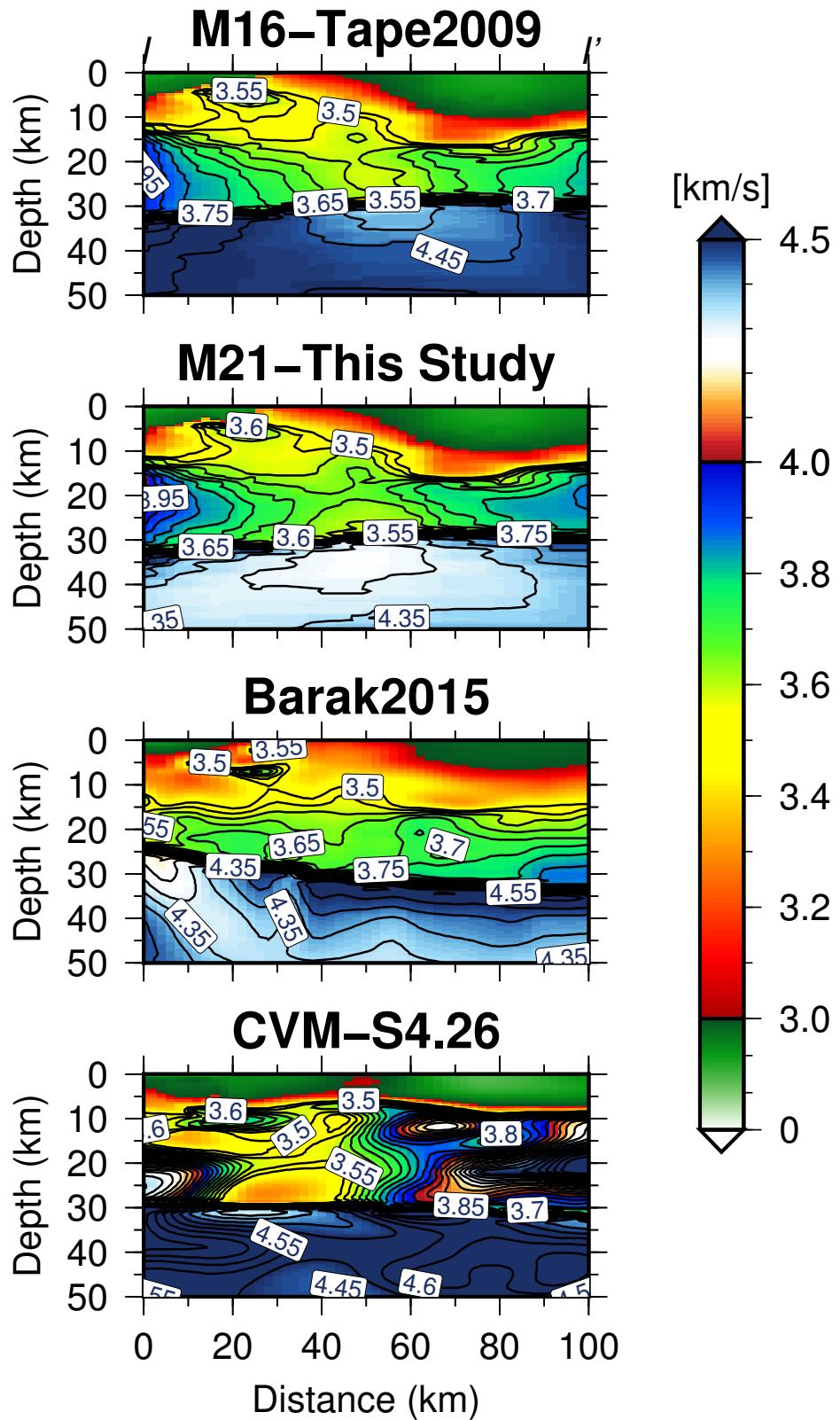


FIGURE 3.S16: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile II'

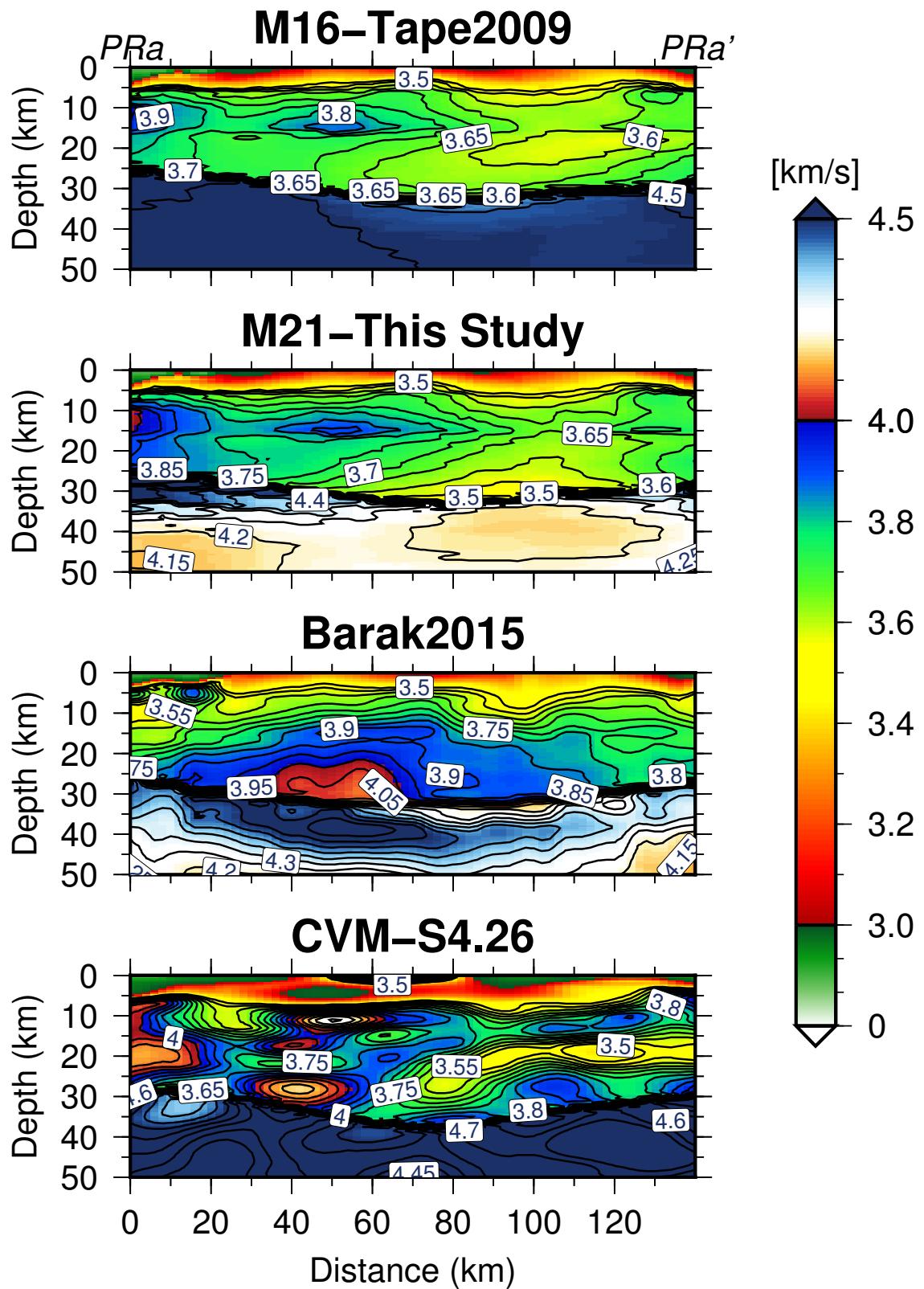


FIGURE 3.S17: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile PRa-PRa'

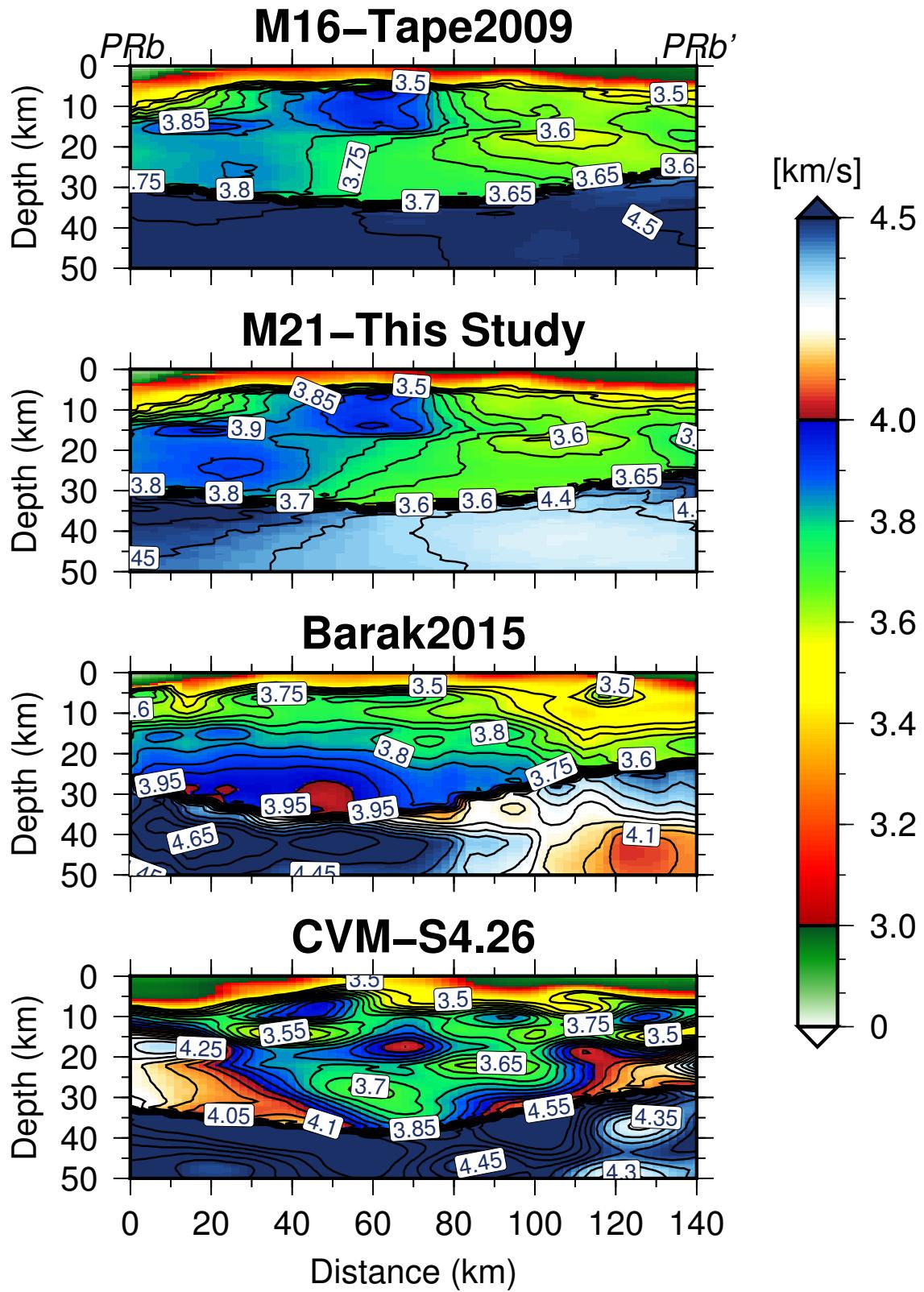


FIGURE 3.S18: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile PRb-PRb'

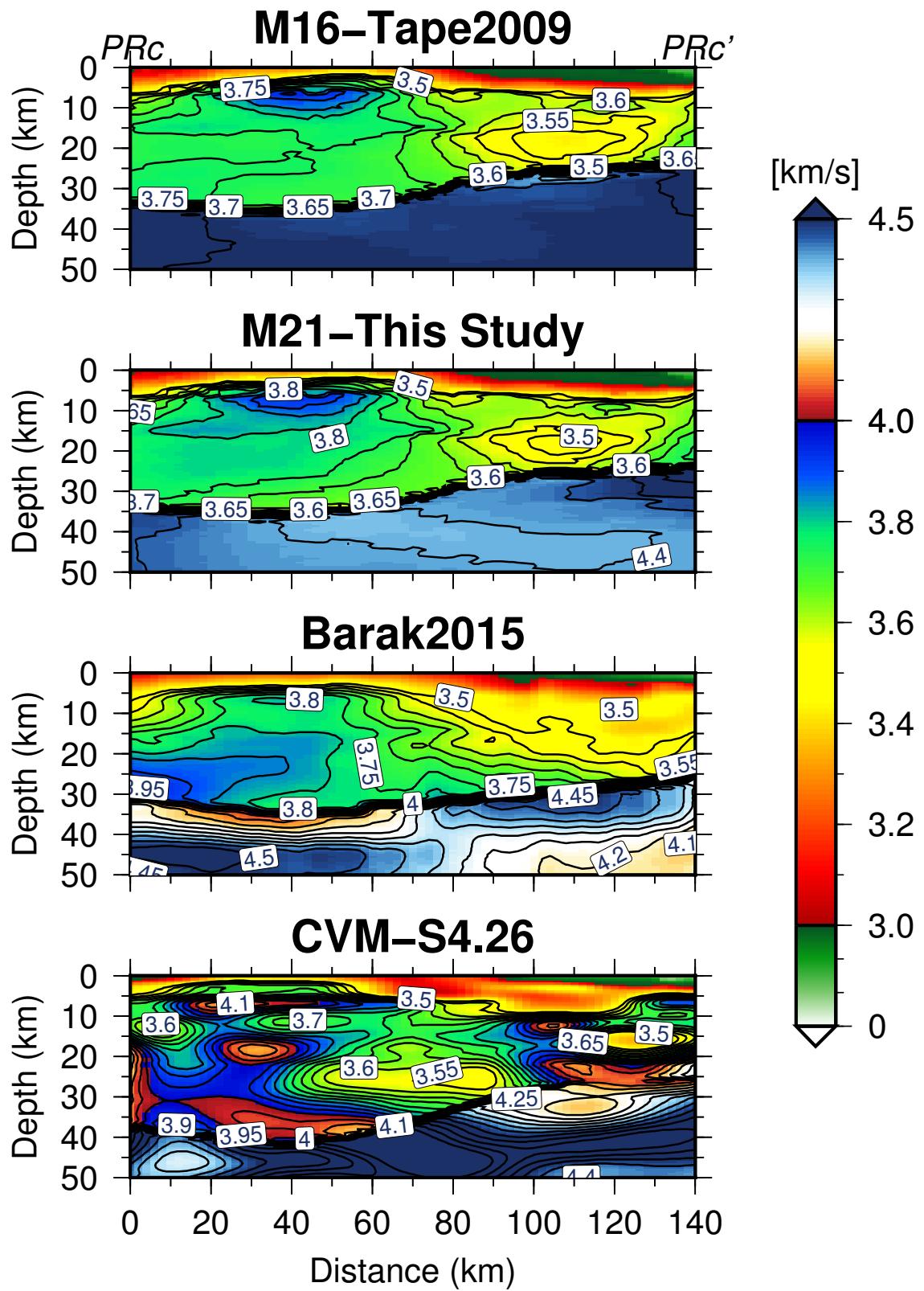


FIGURE 3.S19: Comparison of **M16**, **M21**, Barak2015 and CVM-S4.26 along profile $PRc-PRc'$

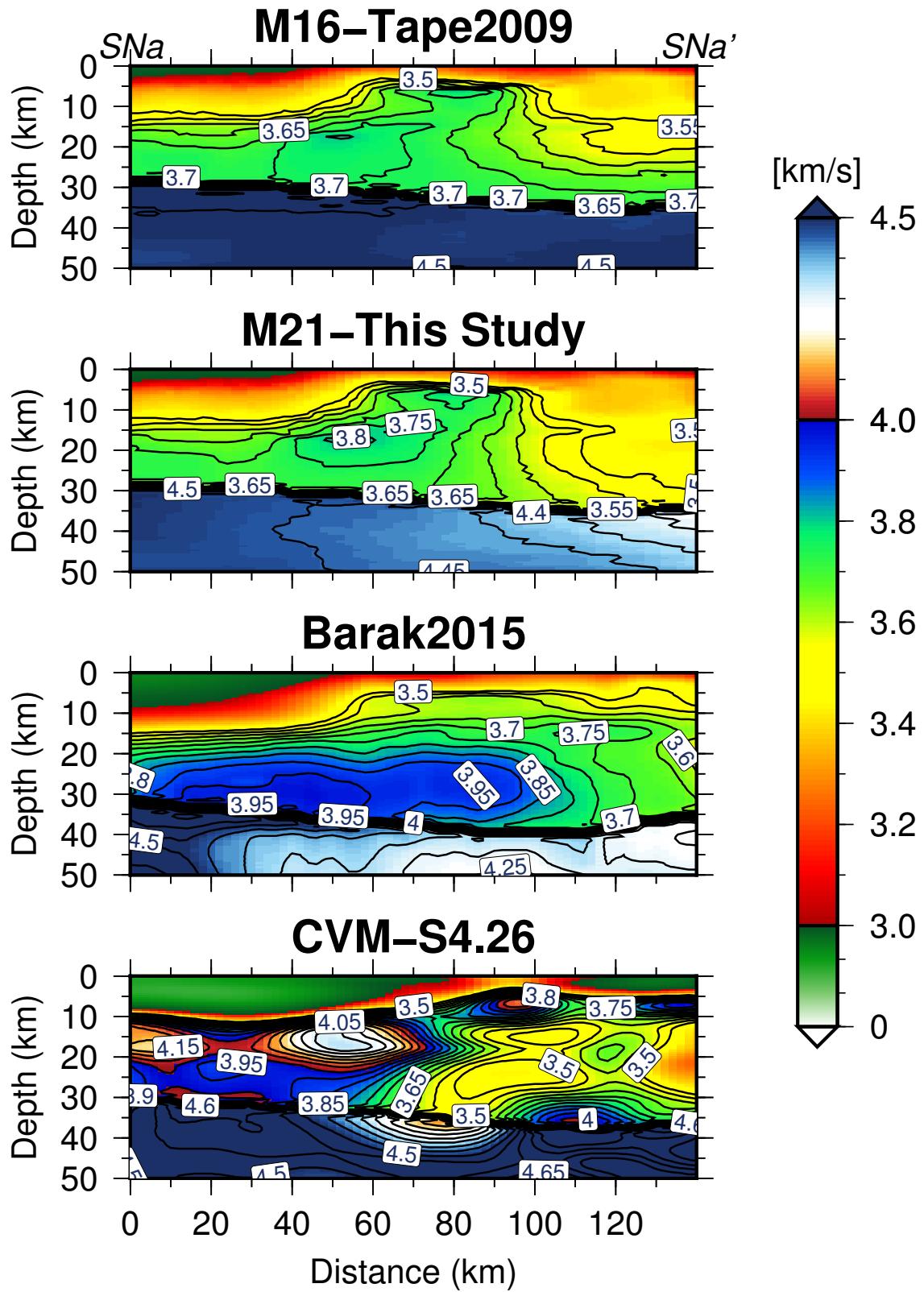


FIGURE 3.S20: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile SNa-SNa'

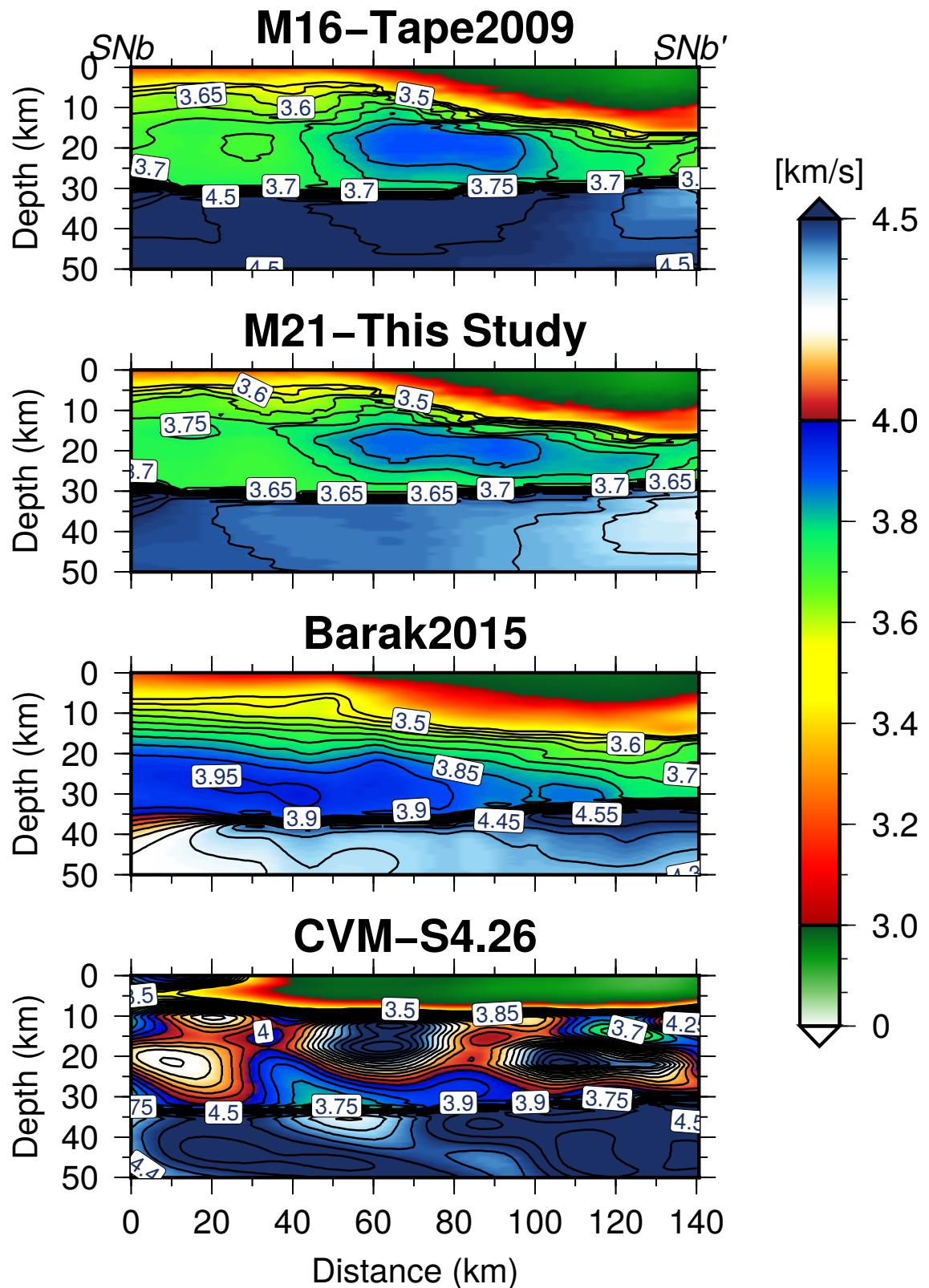


FIGURE 3.S21: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile SNb1-SNb1'

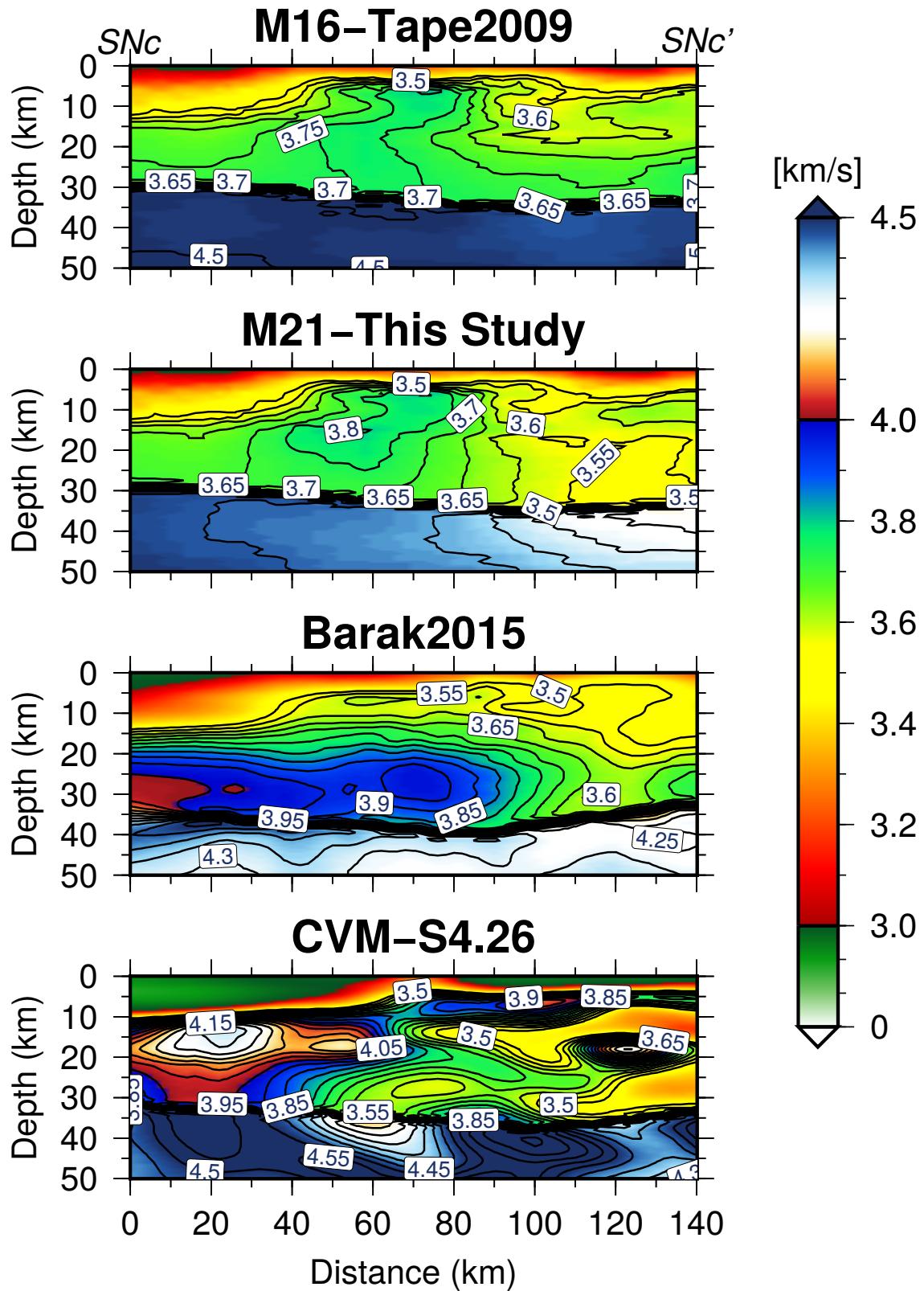


FIGURE 3.S22: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile $SNb2-SNb2'$

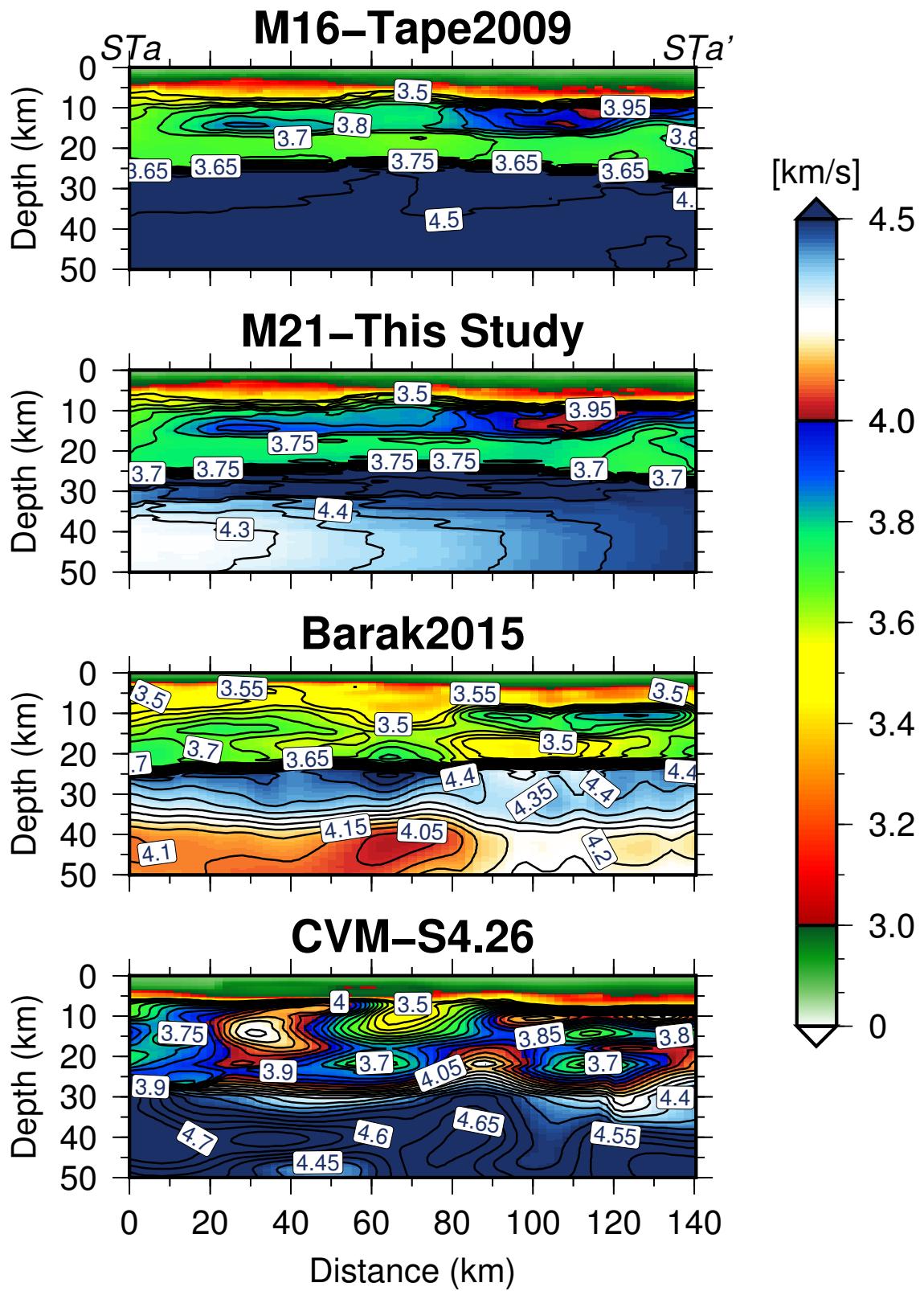


FIGURE 3.S23: Comparison of **M16**, **M21**, Barak2015 and CVM-S4.26 along profile STA-STA'

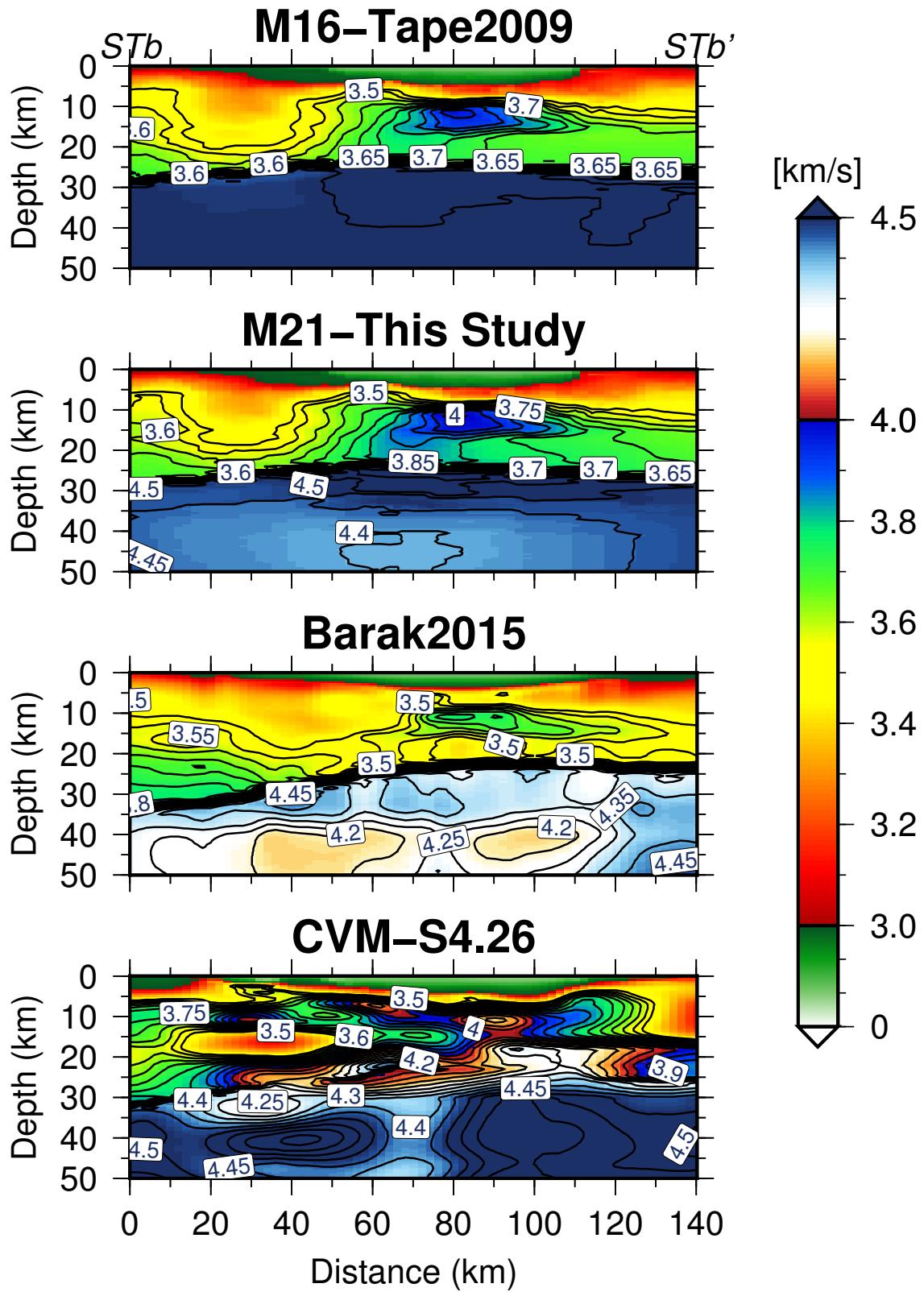
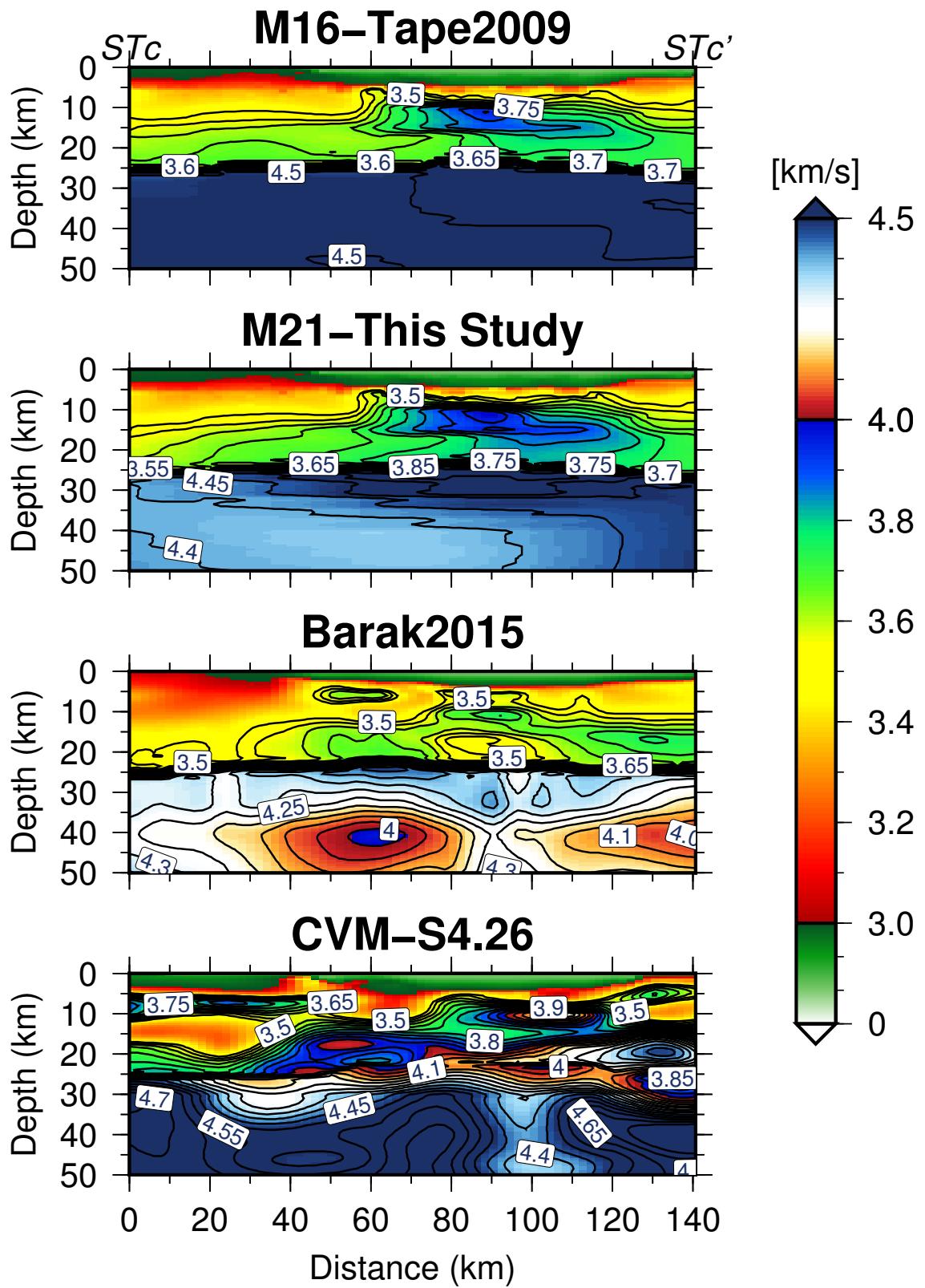


FIGURE 3.S24: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile STb-STb'

FIGURE 3.S25: Comparison of M16, M21, Barak2015 and CVM-S4.26 along profile ST_c-ST_{c'}

4

Sensitivity kernels for multi-component ambient noise data based on adjoint methods, with application to Love-wave adjoint tomography in southern California

4.1 Summary

Empirical Green's functions (EGFs) from ambient noise have been introduced to adjoint tomography based on spectral-element method in recent studies. However, only vertical-vertical components are used to infer V_{sv} velocity structure. We propose an efficient method to calculate multi-component (e.g., transverse-transverse (T-T), and radial-radial (R-R)) sensitivity kernels for EGFs from ambient noise through 3-D numerical simulations of seismic wave propagation. Based on adjoint method, we demonstrate that a T-T component kernel

can be obtained by the summation of two kernels, each of which is constructed from the interaction of a forward field due to a delta-function point force in the east or north direction, and an adjoint field by injecting T-T component adjoint sources at each individual receiver in the east and north direction with proper amplitude modulation factors based on rotation relations. Therefore, two forward simulations and two adjoint simulations are required to compute the sensitivity kernel for T-T component measurements of an event. The sensitivity kernel for R-R component measurements can be constructed similarly to those of T-T component kernels, but with R-R component adjoint source and different amplitude modulation factors. The rotation relations for seismograms, individual and event misfit kernels are validated by three groups of numerical simulations. Our results show the kernels constructed based on rotation relations are almost identical to the reference ones with less than 1% error, sufficient for adjoint tomographic inversion.

Utilizing this method for calculating T-T component sensitivity kernels, we successfully perform Love-wave ambient noise adjoint tomography (ANAT) in Southern California and obtained a new V_{sh} model. Compared with V_{sv} model from Rayleigh-wave ANAT, we observed positive ($V_{sh} > V_{sv}$) radial anisotropy in the whole uppermost mantle ($\sim +7\%$) and in the north-eastern part of the crust ($\sim +4\%$). In the middle and lower crust (15 – 25 km), the whole Transverse Range shows negative ($V_{sh} < V_{sv}$) radial anisotropy ($\sim -6\%$). The efficient computation method we developed provides us the basis for future multi-component ambient noise adjoint tomography for high-resolution isotropic and anisotropic structures.

4.2 Introduction

In the past decade, ambient noise tomography (ANT) has become a well established technique in seismic imaging because of its ability in retrieving Empirical Green's Functions (EGFs) from continuous random noise field [Lobkis and Weaver, 2001, Roux et al., 2005, Wapenaar et al., 2010]. Benefiting from increasing dense seismic arrays across the globe, ANT has tremendously advanced surface wave studies on passive seismic imaging of isotropic crustal and upper mantle shear-wave structure at regional [e.g., Lin et al., 2008, Shapiro et al., 2005, Yang et al., 2008, Yao et al., 2006], continental [e.g., Bensen et al., 2009, Yang et al., 2007] and global [e.g., Haned et al., 2015, Nishida et al., 2009] scales. Meanwhile, ANT has also been used to infer radial [e.g., Bensen et al., 2009, Luo et al., 2013, Moschetti et al.,

2010a,b, Xie et al., 2013] and azimuthal anisotropic structure [e.g., Lin et al., 2011, Yao et al., 2010]. Several studies also demonstrated it is possible to retrieve the body-wave part of Green's function from ambient noise [Lin et al., 2013, Nishida, 2013, Poli et al., 2012a, Wang et al., 2014b, Zhan et al., 2010]. While surface wave tomography is still the major application, a number of researchers managed to retrieve anelastic structure from coherency of ambient noise field, and apply it to ground motion prediction with great success [e.g., Denolle et al., 2013, 2014, Prieto and Beroza, 2008, Prieto et al., 2011, 2009]. Compared with traditional earthquake-based tomography, ANT is not restricted by the distributions of earthquakes, and the resolution depends mainly on station density. Furthermore, the broadband nature (5 – 300 s) of surface waves extracted from ambient noise makes it complementary to surface waves from teleseismic and local earthquake data.

To date, most ANT studies focus on retrieving Rayleigh-wave EGFs from vertical-vertical (Z-Z) component cross-correlation functions (CCFs), which are subsequently used to invert for a 3-D velocity model. This type of model is essentially a vertical polarized shear-wave velocity (V_{sv}) model and the inversion procedure usually consists of two steps. Firstly, period-dependent Rayleigh-wave phase velocity maps are constructed by inverting inter-station dispersion curves measured from CCFs; Secondly, local dispersion curves are extracted from these phase velocity maps, and then are used to invert for local shear velocities based on 1-D depth sensitivity kernels. The resulting 1-D profiles are assembled to form the 3-D V_{sv} model. The same procedure can be applied to Love-wave EGFs from transverse-transverse (T-T) component CCFs to obtain period-dependent Love-wave phase velocity maps, which are subsequently inverted for 3-D horizontally polarized shear-wave velocity (V_{sh}) model. The study of radial anisotropy also typically involves two steps. First, inversion with isotropic parameterization (V_s) is performed to fit local Rayleigh and Love wave dispersion curves simultaneously. Second, the points that can not be fitted well with isotropic model defines the area with so-called *Rayleigh-Love discrepancy*. Thus, radial anisotropy has to be introduced for these points in order to produce a better data fitting. In general, radial anisotropy can be obtained either by inverting for V_{sv} from Rayleigh-wave phase maps and for V_{sh} from Love-wave phase maps separately [e.g., Bensen et al., 2009, Luo et al., 2013], or more accurately by inverting for V_{sv} and V_{sh} simultaneously [e.g., Moschetti et al., 2010a,b, Xie et al., 2013].

Parallel to the development of ANT, adjoint tomography, a variant of Full Waveform Inversion (FWI) [e.g., Virieux and Operto, 2009], was formulated for earthquake seismological

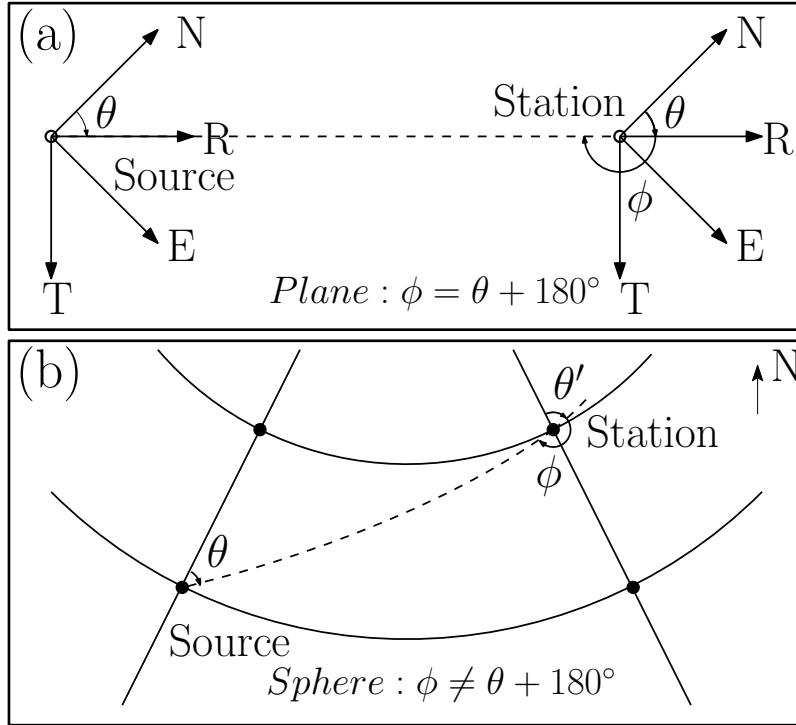


FIGURE 4.1: Definition of azimuth of radial direction at the source (θ) and station (θ') as well as station backazimuth $\phi = \theta' + 180^\circ$ in a flat (a) and spherical earth (b); For the flat earth case, $\theta' = \theta - 180^\circ$.

problems [e.g., Liu and Gu, 2012, Liu and Tromp, 2006, 2008, Tromp et al., 2005]. In particular, more accurate sensitive kernels can be computed based on 3-D full numerical solvers such as the spectral element method (SEM) [e.g., Komatitsch and Tromp, 1999, Tromp et al., 2008], an improvement from the ray theory based or finite-frequency sensitivity kernels used in classical tomography [Liu and Tromp, 2006, 2008]. For a decade or so, adjoint tomography based on earthquake data has been successfully applied to constrain crustal structures in southern California [Tape et al., 2009], and upper mantle structures in Australia , Europe, and East Asia [Chen et al., 2015, Fichtner et al., 2009, 2010, Zhu et al., 2012, 2015], among others.

However, whether the Fréchet kernels based on adjoint technique can be directly applied to ambient noise cross-correlation measurements is still under hot debate. From theoretical perspective, Tromp et al. [2010] developed a method by using ensemble-averaged cross correlations to obtain ensemble sensitivity kernels to account for the non-uniform character of noise sources distribution. Fichtner [2014] showed that both unevenly distributed noise sources and seismic data processing can affect cross-correlations of ambient noise

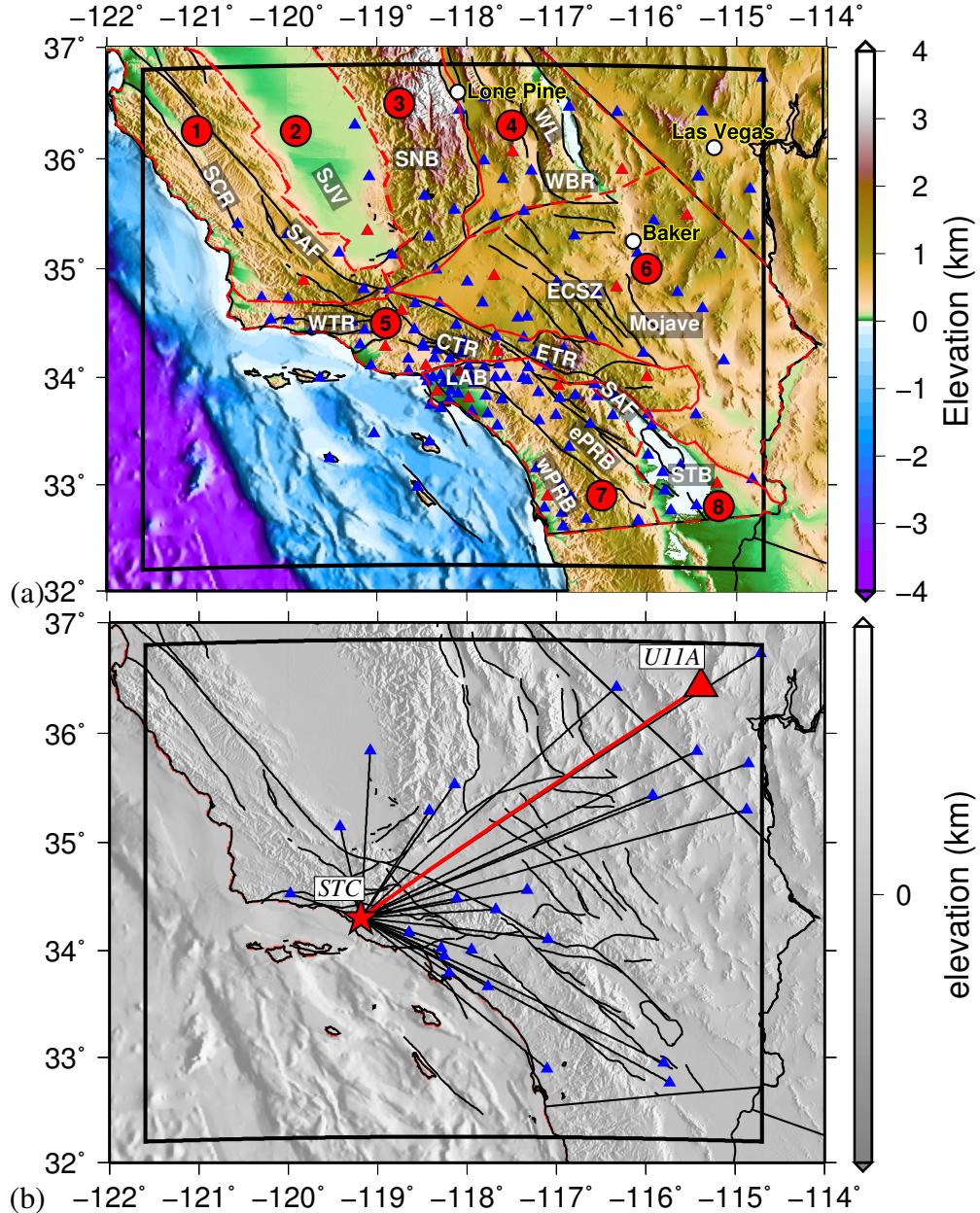


FIGURE 4.2: (a) Map of southern California with topography, bathymetry and active faults. The solid black outline denotes the simulation region. The 148 stations used in this tomographic study are shown as triangles out of which 19 selected for line search are shown as red ones. Faults are shown in bold black lines. Labels 1-8 denote eight geomorphic provinces: 1. Coast Ranges; 2. Great Central Valley; 3. Sierra Nevada; 4. Basin and Range; 5. Transverse Ranges; 6. Mojave Desert; 7. Peninsular Ranges; 8. Colorado Desert. Geological features labelled for reference in subsequent figures are drawn in bold white in transparent black boxes: SCR, southern Coast Range; SJV, San Joaquin Valley SAF, San Andreas Fault; LAB: Los Angeles Basin; ePRB and wPRB: east and west Peninsular Ranges Batholith; SNB, Sierra Nevada Batholith; STB: Salton Trough Basin. (b) Stations in southern California used for kernel calculations. The red star is the master station CI.STC and the blue triangles are 26 other regional stations with TA.U11A shown in red color. The red solid line connecting CI.STC and TA.U11A shows the selected station pair for calculating synthetics and individual kernels, with azimuth of the radial direction at the source $\theta = 54.68119^\circ$, and at the station $\theta' = 56.8734^\circ$, as well as station back-azimuth $\phi = 236.8734^\circ$.

and their sensitivities to structures, suggesting that carrying out full waveform ANT without taking source heterogeneities and data processing schemes into account could introduce tomographic artifacts. More recently, Fichtner et al. [2016] proposed a general theory to derive the Fréchet kernels for seismic interferometry using noise correlation by transforming the actual wavefield sources and wave propagation physics into effective sources and wave propagations.

Nevertheless, surface waves that dominate noise cross-correlation functions have been used for full waveform inversion (FWI) through either scattering-integral method based on finite-difference modeling [e.g., Gao and Shen, 2014, Lee et al., 2014], or adjoint tomography based on SEM [e.g., Chen et al., 2014, Wang et al., 2018b]. For example, Gao and Shen [2014] used 7–200 s surface waves from broadband ambient noise to constrain the upper-mantle structure of the Cascadia subduction zone. Chen et al. [2014] conducted adjoint tomography based on short-period (10–40 s) waveforms from ambient noise recorded at 25 stations in southeastern Tibet, and helps refine the previous crustal model. Particularly, Lee et al. [2014] applied full 3-D tomography using a combination of scattering-integral and adjoint method based on finite-difference method to both earthquake and ambient noise data at 5 – 50 s period range. Wang et al. [2018b] applied adjoint tomography to 5 – 50 s Rayleigh wave empirical Green’s Functions constructed from ambient noise in southern California, and the results helped improve the 3-D earthquake-based tomographic model **M16** from [Tape et al., 2009]. Compared to the traditional ANT, these studies demonstrated several advantages of full waveform ANT, including better waveform fitting, more realistic 3-D sensitivity kernels, and results displaying more pronounced velocity variations.

However, current ambient-noise adjoint tomography (ANAT) or full-waveform ANT studies only utilizes vertical-vertical (Z-Z) component Rayleigh-wave EGFs extracted from vertical-component ambient noise to resolve isotropic structures. No ANAT studies have attempted to use signals from other components, such as Love-wave signals from T-T components. This may be partially due to the significant computational cost needed to construct T-T or R-R component kernels based on numerical simulations. However, the additional T-T and R-R component EGFs in ANAT are complementary to the Z-Z component EGFs and allow studies of radial anisotropy. This calls for more efficient computation methods of constructing event kernels [Tape et al., 2007] for horizontal-component EGFs in ANAT, which is the primary goal of this study.

For Rayleigh-wave type ANAT using only Z-Z component EGFs, vertical point forces are implemented as virtual sources at receivers to generate the adjoint wave fields, which are then time-convolved with the forward wave fields to generate the event kernels. In comparison, for Love-wave type ANAT, a transverse-component point-force is required to be injected at the source for the forward field as well as at each receiver for the adjoint field. However, since each source-receiver pair has different azimuths, it becomes impossible to construct similar event kernel for all receivers with only one forward and adjoint simulation. In order to solve this problem, we propose to perform two forward simulations (with point force injected in the N or E direction at the source) and two adjoint simulations. The resulting two kernels from time-convolution of these forward and adjoint fields can be added to obtain the T-T kernel.

This chapter is organized as follows: (1) in section 4.3, we derive the rotation formulas for constructing horizontal-component sensitivity kernel base on SEM and adjoint method; (2) Then, we conduct several numerical simulations to verify our method in section 4.4; (3) Finally, we apply the new method in Love-wave ambient noise adjoint tomography of Southern California in section 4.5.

4.3 Methodology

The theory of adjoint tomography has been well documented in literatures [e.g., Fichtner, 2011, Luo, 2012, Tape et al., 2007, Tromp et al., 2008, 2005]. In this section, we only outline the formal expressions for kernel calculations based on adjoint method. Using rotation relations, we show how sensitivity kernel for multi-component EGFs can be obtained through two forward and adjoint simulations with proper amplitude modulation factors.

4.3.1 Sensitivity kernels for isotropic elastic models

If we assume ϕ to be a scalar function that quantifies measurements or misfit values between the observed data and synthetic predictions based on a reference (or current) model $m(\mathbf{x})$, then ϕ can be linearized relative to the reference model. For example, for an isotropic model, we can represent the model using parameters of density, shear and bulk modulus (ρ, μ, κ),

and rewrite the variation of ϕ as

$$\delta\phi = \int_V [K_\rho(\mathbf{x})\delta \ln \rho(\mathbf{x}) + K_\mu(\mathbf{x})\delta \ln \mu(\mathbf{x}) + K_\kappa(\mathbf{x})\delta \ln \kappa(\mathbf{x})] d^3\mathbf{x}, \quad (4.1)$$

where K_ρ , K_μ and K_κ are the Fréchet kernels for fractional density, shear and bulk moduli perturbations. Based on single-scattering (Born) approximation [Dahlen et al., 2000], and formulated by the adjoint method [Liu and Tromp, 2006, Tromp et al., 2005], these Fréchet derivatives can be computed by

$$K_\rho(\mathbf{x}) = -\rho(\mathbf{x}) \int_0^T [\partial_t^2 \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s) \cdot \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)] dt, \quad (4.2)$$

$$K_\mu(\mathbf{x}) = -2\mu(\mathbf{x}) \int_0^T [\mathbf{D}(\mathbf{x}, t; \mathbf{x}_s) : \mathbf{D}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)] dt, \quad (4.3)$$

$$K_\kappa(\mathbf{x}) = -\kappa(\mathbf{x}) \int [\nabla \cdot \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)] [\nabla \cdot \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)] dt, \quad (4.4)$$

where $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)$ is the forward wavefield emitted from the source, $\mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)$ is the adjoint wavefield generated by the time-reversed adjoint source injected at the receiver, and their strain deviators are denoted by \mathbf{D} and \mathbf{D}^\dagger , respectively. These kernels are known as *individual sensitivity kernels* for one source-receiver pair. If data residual signal for all receivers are injected at the same time as adjoint sources to generate the combined adjoint field $\mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)$, then eq. (4.2)-(4.4) give the so-called *event kernels* for density, shear modulus and bulk modulus, which are the basic building blocks for adjoint tomographic inversions. Note that these formal kernel expressions in eq. (4.2)-(4.4) hold for different types of misfit functions or measurements ϕ . The selection of different misfit functions only influences the adjoint source (i.e., data residual injected at receivers) and hence the adjoint wavefield. To reduce symbol clutter, we write the expression for a generic kernel as

$$K(\mathbf{x}) = \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s) \otimes \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r) \quad (4.5)$$

where the interpretation of the \otimes operator for a specific type of kernel can be made based on eq. (4.2)-(4.4).

4.3.2 Coordinate system rotation

As rotations of seismograms are involved in constructing multi-component adjoint sources, we define an orthonormal matrix A that represents the rotation from one coordinate system

$\{\hat{x}_j\}$ with unit vectors in east, north and vertical directions ($\{\hat{x}_e, \hat{x}_n, \hat{x}_z\}$) to another coordinate system $\{\hat{x}^I\}$ with unit vectors in transverse, radial and vertical directions ($\{\hat{x}^T, \hat{x}^R, \hat{x}^Z\}$), such that

$$\hat{x}^I = A_j^I \hat{x}_j, \quad \text{and} \quad A_j^I(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.6)$$

where Einstein summation is applied to repeated lowercase index, I and j represent the row and column numbers of matrix A , and θ is the azimuth of the radial vector \hat{x}^R clockwise from \hat{x}_n , as shown in Figure 4.1(a). For the spherical earth (Figure 4.1b), the azimuth of the radial vector at the source (θ) may be different from the azimuth of the radial vector at the receiver (θ'), but for flat earth geometry, $\theta = \theta'$. Also, for a fixed source at \mathbf{x}_s , θ and θ' vary as a function of the receiver position \mathbf{x}_r , i.e., they are generally different for different receivers.

4.3.3 Individual sensitivity kernels and adjoint source

Let us first consider the sensitivity kernel for the measurements made on a single station pair. The source-receiver plane is defined as a horizontal plane containing the source-receiver path with a surface azimuth of θ (Figure 4.1a) at the source and surface back azimuth ϕ at the receiver. For the master station at \mathbf{x}_s , we define the forward field due to a Delta-function point force in the \hat{x}^I direction as $\mathbf{u}^I(\mathbf{x}, t; \mathbf{x}_s)$. It can be related to forward wavefield generated by a Delta-function point force placed in the \hat{x}_j direction, $\mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s)$, as

$$\mathbf{u}^I(\mathbf{x}, t; \mathbf{x}_s) = A_j^I(\theta) \mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s). \quad (4.7)$$

Note that here j and I indices indicate the source direction instead of the component of the wavefield. For an EGF extracted from cross-correlation between the I -th component of the displacement at the master station, and J -th component of the displacement at the receiver station \mathbf{x}_r , referred to as $d^{IJ}(\mathbf{x}_r, t; \mathbf{x}_s)$, the corresponding synthetic Greens function can be extracted from J -th component of the forward field $\mathbf{u}^I(\mathbf{x}_r, t; \mathbf{x}_s)$ at receivers.

In a general coordinate system (section 4.3.2), the source force direction \hat{x}^I (e.g., $I = T, R$), may vary based on the azimuth of the receiver θ relative to the source, and therefore, the

synthetics $\mathbf{u}^I(\mathbf{x}_r, t; \mathbf{x}_s)$, for all receivers cannot be computed in one simple forward simulation. However, using the rotation relation eq. (4.6) and (4.7), its \hat{x}^J component becomes

$$\begin{aligned}\hat{\mathbf{x}}^J(\theta') \cdot \mathbf{u}^I(\mathbf{x}_r, t; \mathbf{x}_s) &= [A_k^J(\theta')\hat{x}_k] \cdot [A_j^I(\theta)\mathbf{u}_j(\mathbf{x}_r, t; \mathbf{x}_s)] \\ &= \hat{x}_k \cdot \mathbf{u}_j(\mathbf{x}_r, t; \mathbf{x}_s) [A_j^I(\theta)A_k^J(\theta')].\end{aligned}\quad (4.8)$$

Here, $\mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s)$, the wavefield due to a source in $\hat{\mathbf{x}}_j$ direction, can be obtained through one forward simulation for all receivers. Then the synthetics corresponding to d^{IJ} can be constructed through the weighted summations of \mathbf{u}_j ($j = e, n, z$). For specific I and J , we can define an auxiliary matrix C^{IJ} for a specific source-receiver pair as

$$C_{jk}^{IJ}(\theta, \theta') = A_j^I(\theta)A_k^J(\theta'), \quad (4.9)$$

where j and k are the row and column numbers of C^{IJ} . For the measurements made for Z-Z, T-T, and R-R component EGFs, the auxiliary matrices are given by

$$\begin{aligned}C^{ZZ} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C^{TT} = \begin{bmatrix} \cos \theta \cos \theta' & -\cos \theta \sin \theta' & 0 \\ -\sin \theta \cos \theta' & \sin \theta \sin \theta' & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ C^{RR} &= \begin{bmatrix} \sin \theta \sin \theta' & \sin \theta \cos \theta' & 0 \\ \cos \theta \sin \theta' & \cos \theta \cos \theta' & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned}\quad (4.10)$$

Hence the corresponding synthetics seismograms for these three types measurements become

$$\hat{x}^Z \cdot \mathbf{u}^Z = \hat{x}_z \cdot \mathbf{u}_z. \quad (4.11)$$

$$\begin{aligned}\hat{x}^T \cdot \mathbf{u}^T &= [(\cos \theta \cos \theta')\hat{x}_e - (\cos \theta \sin \theta')\hat{x}_n] \cdot \mathbf{u}_e \\ &\quad + [-(\sin \theta \cos \theta')\hat{x}_e + (\sin \theta \sin \theta')\hat{x}_n] \cdot \mathbf{u}_n\end{aligned}\quad (4.12)$$

$$\begin{aligned}\hat{x}^R \cdot \mathbf{u}^R &= [(\sin \theta \sin \theta')\hat{x}_e + (\sin \theta \cos \theta')\hat{x}_n] \cdot \mathbf{u}_e \\ &\quad + [(\cos \theta \sin \theta')\hat{x}_e + (\cos \theta \cos \theta')\hat{x}_n] \cdot \mathbf{u}_n.\end{aligned}\quad (4.13)$$

Eq. (4.12-4.13) state that synthetics for horizontal component EGFs are the weighted sum of the forward fields with point sources placed in the east and north directions with proper sine/cosine rotation factors as weights.

The synthetic seismogram, \hat{x}^J component of $\mathbf{u}^I(\mathbf{x}_r, t; \mathbf{x}_s)$ at receiver \mathbf{x}_r for a master station at \mathbf{x}_s , together with the EGF, $d^{IJ}(\mathbf{x}_r, t; \mathbf{x}_s)$, can then be used to construct the associated

adjoint source $f^{IJ}(\mathbf{x}_r, t; \mathbf{x}_s)$ [Tromp et al., 2005]. This adjoint source when injected in the \hat{x}^J direction of the receiver results in the adjoint field

$$\begin{aligned} (\mathbf{u}^\dagger)^J(\mathbf{x}, t; \mathbf{x}_r) &= \mathbf{u}^J(\mathbf{x}, t; \mathbf{x}_r) * f^{IJ}(\mathbf{x}_r, t; \mathbf{x}_s) \\ &= A_k^J(\theta') \mathbf{u}_k(\mathbf{x}, t; \mathbf{x}_r) * f^{IJ}(\mathbf{x}_r, t; \mathbf{x}_s), \end{aligned} \quad (4.14)$$

where $*$ represents time convolution and no summation is applied to repeated uppercase indices. Therefore, the generic sensitivity kernel for the measurement between this source-receiver pair becomes

$$\begin{aligned} K^{IJ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= \mathbf{u}^I(\mathbf{x}, t; \mathbf{x}_s) \otimes (\mathbf{u}^\dagger)^J(\mathbf{x}, T - t; \mathbf{x}_r) \\ &= \mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s) \otimes \left\{ [A_j^I(\theta) A_k^J(\theta') f^{IJ}(T - t)] * \mathbf{u}_k(\mathbf{x}, T - t; \mathbf{x}_r) \right\} \\ &= \mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s) \otimes \left\{ [C_{jk}^{IJ}(\theta, \theta') f^{IJ}(T - t)] * \mathbf{u}_k(\mathbf{x}, T - t; \mathbf{x}_r) \right\}, \end{aligned} \quad (4.15)$$

where we drop the $(\mathbf{x}_r, \mathbf{x}_s)$ -dependency of the adjoint source f^{IJ} for more compact expressions. Using the auxiliary matrix shown in eq. (4.10), we can write out detailed expressions for sensitivity kernels of different ambient noise measurements.

Z-Z component sensitivity kernel

Let $I = J = Z$ in eq. (4.15), then we have

$$K^{ZZ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) = \mathbf{u}_z(\mathbf{x}, t; \mathbf{x}_s) \otimes \left\{ f^{ZZ} * \mathbf{u}_z(\mathbf{x}, T - t; \mathbf{x}_r) \right\} \quad (4.16)$$

which suggests that the individual kernel for Z-Z component EGFs can be constructed from the interaction of the forward field from a Delta-function vertical point force at \mathbf{x}_s and the adjoint field due to an adjoint source f^{ZZ} injected at the receiver in the vertical direction. The adjoint source f^{ZZ} is constructed from measurement made between the Z-Z component EGFs and the vertical synthetic seismogram of the aforementioned forward field at the receiver. In the above expressions, explicit dependence of f^{TT} on $(\mathbf{x}_r, T - t; \mathbf{x}_s)$ is omitted to reduce symbol clutter.

T-T component sensitivity kernel

Let $I = J = T$ in eq. (4.15), then we have

$$\begin{aligned}
 K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= K^{ET}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) + K^{NT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) \\
 &= \mathbf{u}_e(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\cos \theta \cos \theta' f^{TT}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\
 &\quad - [\cos \theta \sin \theta' f^{TT}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\} \\
 &\quad + \mathbf{u}_n(\mathbf{x}, t; \mathbf{x}_s) \otimes \{-[\sin \theta \cos \theta' f^{TT}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\
 &\quad + [\sin \theta \sin \theta' f^{TT}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\},
 \end{aligned} \tag{4.17}$$

which suggests that the sensitivity kernel for measurements on T-T component EGFs can be constructed by the summation of two kernels. The first kernel is from the interaction of a forward field due to a delta-function point force in the east direction, $\mathbf{u}_e(\mathbf{x}, t; \mathbf{x}_s)$, and an adjoint field by injecting f^{TT} in the east and north direction at the individual receiver, with amplitude modulated by $\cos \theta \cos \theta'$ and $-\cos \theta \sin \theta'$, respectively. The other kernel can be constructed similarly, except that the forward field is due to a point force in the north direction, and the amplitude modulation factors for the adjoint source are $-\sin \theta \cos \theta'$ and $\sin \theta \sin \theta'$, respectively. Therefore, two forward simulations and two adjoint simulations are required to compute the sensitivity kernel for a measurement made on the T-T component EGF for a source-receiver station pair.

R-R component sensitivity kernel

Let $I = J = R$ in eq. (4.15), then we have

$$\begin{aligned}
 K^{RR}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= K^{ER}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) + K^{NR}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) \\
 &= \mathbf{u}_e(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\sin \theta \sin \theta' f^{RR}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\
 &\quad + [\sin \theta \cos \theta' f^{RR}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\} \\
 &\quad + \mathbf{u}_n(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\cos \theta \sin \theta' f^{RR}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\
 &\quad + [\cos \theta \cos \theta' f^{RR}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\}.
 \end{aligned} \tag{4.18}$$

Hence the sensitivity kernel for R-R component EGFs is also the sum of two kernels, similar to that for T-T component kernel in eq. (4.17), except that the adjoint source f^{RR} and the amplitude modulation factors are different. In practice, once the T-T component kernel is computed, no more forward simulations are needed, and only two additional adjoint simulations are necessary for the R-R component kernel.

4.3.4 Event kernel for ambient-noise measurements

Now we consider the event kernel associated with a selected master station at \mathbf{x}_s and a number of receivers \mathbf{x}_r , $r = 1, \dots, N_s$, which is effectively a summation of individual kernels in eq. (4.15), and can be written as

$$\begin{aligned} K^{IJ}(\mathbf{x}; \mathbf{x}_s) &= \sum_{r=1}^{N_s} \mathbf{u}^I(\mathbf{x}, t; \mathbf{x}_s) \otimes (\mathbf{u}^\dagger)^J(\mathbf{x}, T-t; \mathbf{x}_r) \\ &= \mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s) \otimes \sum_{r=1}^{N_s} \left\{ [C_{jk}^{IJ}(\theta, \theta') f^{IJ}] * \mathbf{u}_k(\mathbf{x}, T-t; \mathbf{x}_r) \right\}, \end{aligned} \quad (4.19)$$

which suggests that computing the event kernel for a selected source station and a number of receivers also requires two forward and two adjoint simulations as in eqs. (4.17-4.18), except that the modulated adjoint sources for individual receivers are injected at the same time to compute the adjoint fields. We will distinguish the kernels constructed by this efficient method as ‘rotated’ kernels.

4.3.5 Total misfit kernel

For a number of master stations \mathbf{x}_s , $s = 1, \dots, S$, the total misfit kernel is the summation of all event kernels as

$$K^{IJ}(\mathbf{x}) = \sum_{s=1}^S K^{IJ}(\mathbf{x}; \mathbf{x}_s) = \sum_{s=1}^S \sum_{r=1}^{N_s} K^{IJ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) \quad (4.20)$$

4.4 Numerical simulations

In this section, we conduct three groups of numerical simulations to validate : (1) The rotation relations between horizontal-component synthetic seismograms ($\hat{x}^R \cdot \mathbf{u}^R$, $\hat{x}^R \cdot \mathbf{u}^T$, $\hat{x}^T \cdot \mathbf{u}^R$ and $\hat{x}^T \cdot \mathbf{u}^T$) and the north and east components of the two forward fields ($\hat{x}_e \cdot \mathbf{u}_e$, $\hat{x}_n \cdot \mathbf{u}_e$, $\hat{x}_e \cdot \mathbf{u}_n$, $\hat{x}_n \cdot \mathbf{u}_n$) as in eq. (4.8); (2) individual kernel, $K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$, computed based on the

algorithm given by eq. (4.17); (3) event kernel, $K^{TT}(\mathbf{x}; \mathbf{x}_s)$, based on eq. (4.19). Although we only show the results for T-T component kernels, the methods also apply to R-R component kernels.

All the numerical simulations are performed for the 3-D Southern Californian Vs model, **M16**, constructed by [Tape et al., 2009] based on SEM simulations [Komatitsch and Tromp, 1999]. The parameters for hexhedral meshing and SEM solver are the same as those used in our recent Rayleigh-wave ANAT study [Wang et al., 2018b]. We use irregular finite-element mesh which honors surface topography and bathmetry, the lower part of the sedimentary basin underneath Los Angeles [Süss and Shaw, 2003], and the shape of Moho [Zhu and Kanamori, 2000]. The mesh yields a horizontal grid spacing of about 2 km and an increasing vertical spacing from ~ 1 km at the top to ~ 12 km at the bottom, which gives sufficient simulation accuracy down to 2 second. A Gaussian function is used as the source time function of the point force, and we set its half-duration to be 1.0 s since bandpass filters are applied in the processing later. All measurements are made at the period range of 20 – 50 s, and so as the individual and event kernels.

For illustration purpose, we choose *CI.STC* as the master station (red star in Figure 4.2b), and conduct a number of forward and adjoint simulations with point force sources in north, east, radial and transverse directions, respectively. We choose a receiver station *TA.U11A* with azimuth $\theta = 54.68119^\circ$ and back azimuth $\phi = 236.8734^\circ$ (hence $\theta' = 56.8734^\circ$) for the comparison of seismograms and individual misfit kernels. For the validation of event misfit kernel, a total of 26 receivers are used (blue rectangles in Figure 4.2b) relative to the master station *CI.STC*.

To simplify the illustration, we rewrite the J-th component of $\mathbf{u}^I(\mathbf{x}_r, t; \mathbf{x}_s)$ as $\hat{x}^J \cdot \mathbf{u}^I = u^{IJ}$. For each group of simulations, We first perform forward simulations by placing a point force source in the transverse or radial direction and produce the reference seismograms ($\underline{u}^{RR}, \underline{u}^{RT}, \underline{u}^{TR}, \underline{u}^{TT}$), individual kernel ($\underline{K}^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$). The summation of all the reference individual kernels between the master station and 26 generic receivers forms the reference event kernel ($\underline{K}^{TT}(\mathbf{x}; \mathbf{x}_s)$). Then, forward simulations are performed for north and east-component point forces and the wavefields are converted using relations in eq. (4.8) to the rotated seismograms($u^{RR}, u^{RT}, u^{TR}, u^{TT}$). Rotated individual kernel ($K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$) and event misfit kernel ($K^{TT}(\mathbf{x}; \mathbf{x}_s)$) can be constructed similarly based on these two forward simulations. Finally, we compare the difference between the reference seismograms/kernels with those

based on rotation relations.

4.4.1 Seismogram comparison

For the station pair CI.STC-TA.U11A (Figure 4.2b), we first generate three-component seismograms from four single point force sources polarized in different directions: (1) radial: $\mathbf{u}^R = (u^{RE}, u^{RN}, u^{RZ})$; (2) transverse: $\mathbf{u}^T = (u^{TE}, u^{TN}, u^{TZ})$; (3) east: $\mathbf{u}^E = (u^{EE}, u^{EN}, u^{EZ})$; (4) north: $\mathbf{u}^N = (u^{NE}, u^{NN}, u^{NZ})$. The corresponding horizontal components are shown in Figure 4.3(a)-(c). Using eq. (4.6), we obtain the reference seismograms (\underline{u}^{RR} , \underline{u}^{RT}), (\underline{u}^{TR} , \underline{u}^{TT}) (black solid lines in Figure 4.3(d)) from (u^{RE}, u^{RN}) , and (u^{TE}, u^{TN}) , respectively. Then, we rotate the four north-east components $(u^{EE}, u^{EN}, u^{NE}, u^{NN})$ to obtain the four rotated seismograms (\underline{u}^{RR} , \underline{u}^{RT} , \underline{u}^{TR} , \underline{u}^{TT}) (red dashed lines in 4.3(d)) based on eq. (4.8). As we can see, the reference and rotated seismograms are the same within numerical errors. Note that we can use $\theta' = \theta$ in a plane for rotation, and the results are almost the same as reference ones in such a small region ($\sim 600 \text{ km} \times 600 \text{ km}$).

4.4.2 Individual misfit kernel

Using the reference seismogram \underline{u}^{TT} , we construct the transverse-component adjoint source f^{TT} to generate the adjoint field $(\underline{\mathbf{u}}^\dagger)^T$. The interaction of the forward field $\underline{\mathbf{u}}^T$ and the corresponding adjoint field $(\underline{\mathbf{u}}^\dagger)^T$ results in the reference individual misfit kernel $\underline{K}^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$. On the other hand, the rotated seismogram \underline{u}^{TT} can also be used to compute the adjoint source f^{TT} . With the proper amplitude modulation factors, this adjoint source is injected in the east and north direction at the receiver to construct the adjoint field, which can then interact with the two forward field \mathbf{u}^E and \mathbf{u}^N to generate two individual kernels ($K^{ET}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ and $K^{NT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$) in two separate adjoint simulations (Figure A1) as indicated by eq. (4.17). The summation of these two kernels gives the final individual sensitivity kernel $K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ for measurements on T-T component EGF (Figure 4.4). Compared against the reference individual kernel $\underline{K}^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (Figure 4.4b), the rotated kernel is very similar to the reference kernel, with less than 1% difference (Figure 4.4c). This small error is probably a result of numerical errors in mesh discretization and wave simulations in SEM. Nevertheless, the error is negligible and the rotated kernel is accurate enough for practical tomographic applications.

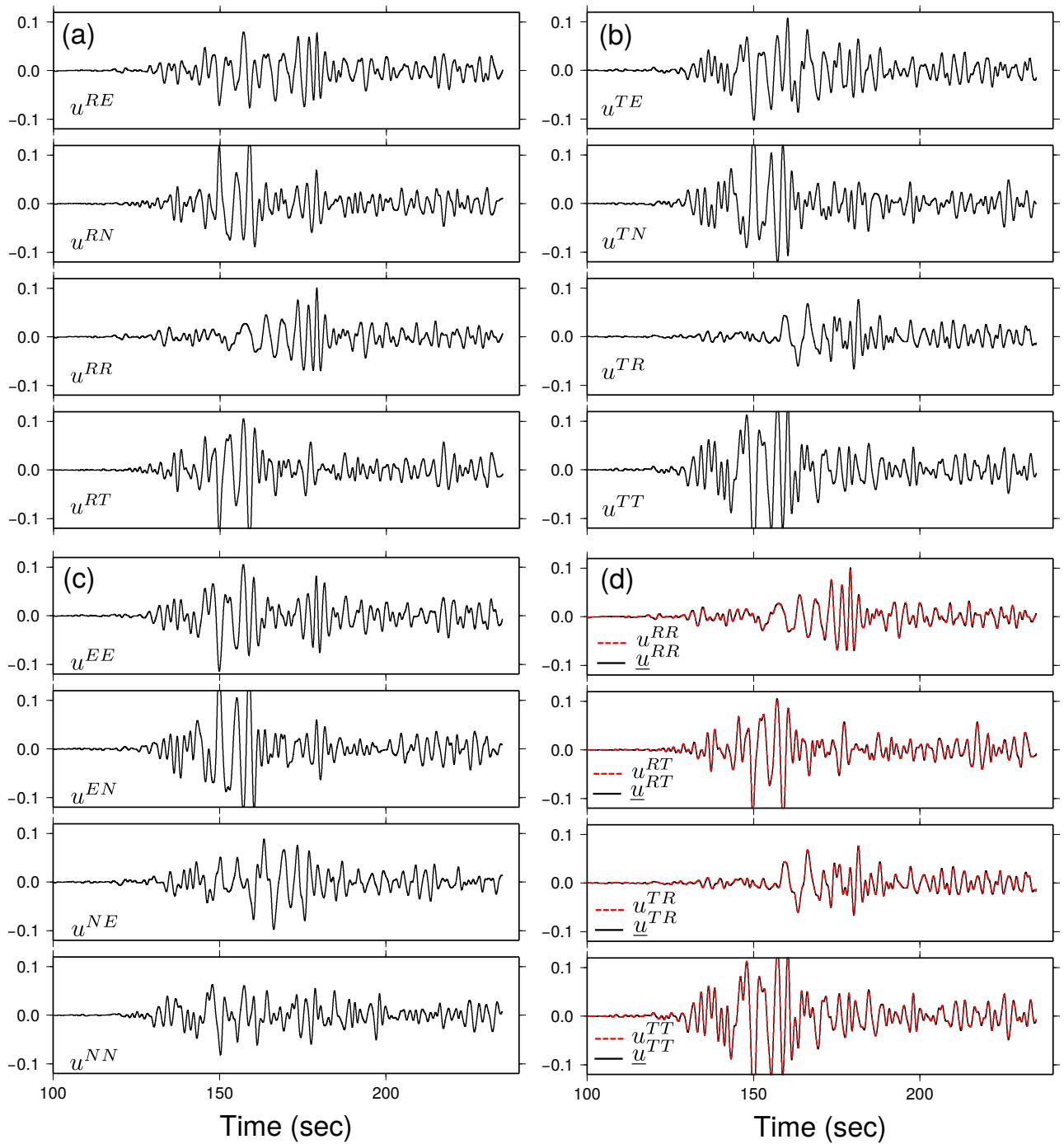


FIGURE 4.3: Synthetic seismograms for station pair CI.STC-TA.U11A from a (a)radial; (b) transverse; (c)north and east-component point source.(d) the black solid lines are reference components from a radial or transverse point sources shown in (a) and (b); the red dashed lines are those rotated from the u^{EE} , u^{EN} , u^{NE} , and u^{NN} components using eq. (4.8).

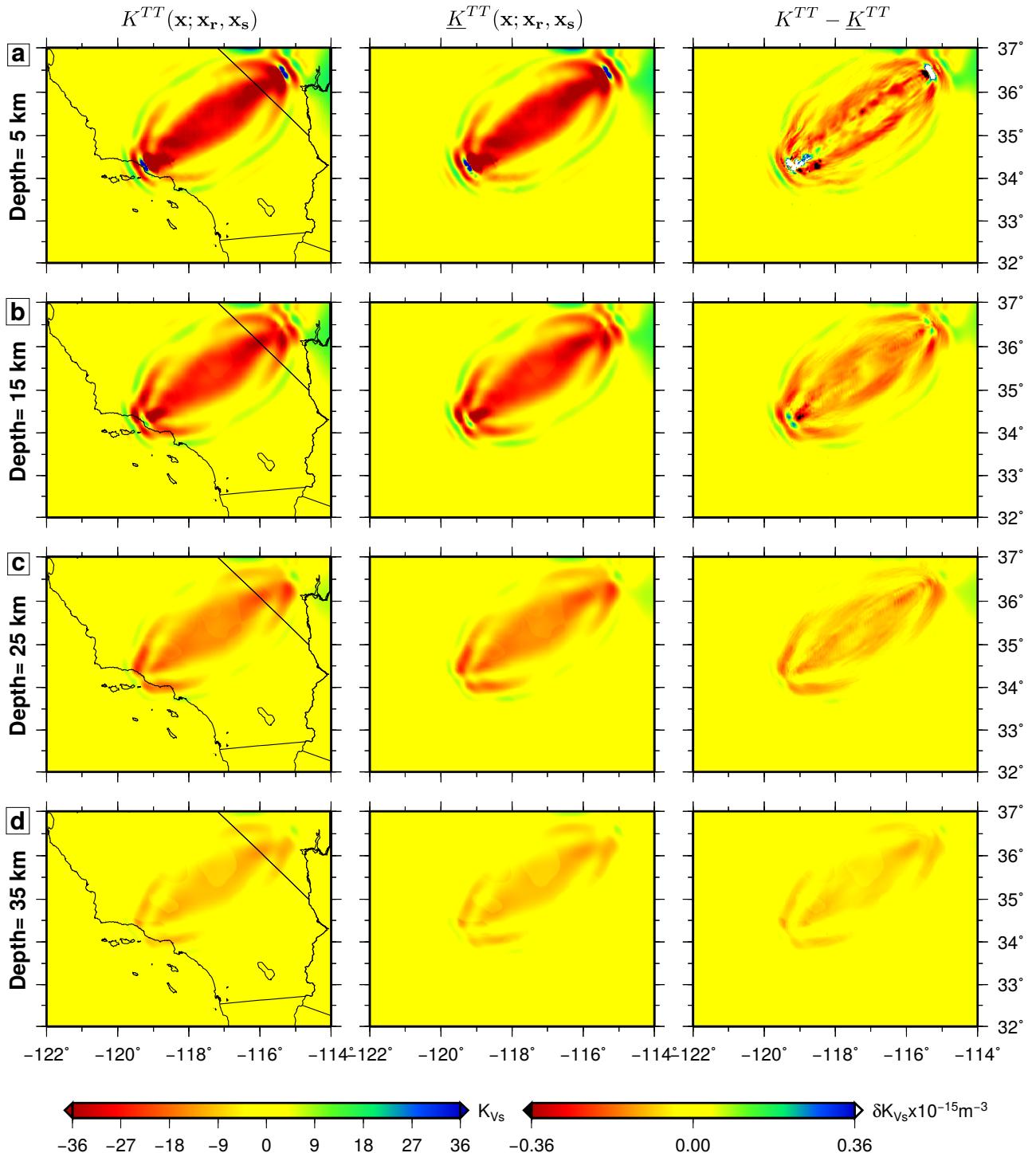


FIGURE 4.4: Rotated (left panels) and reference (middle panels) individual misfit kernel of station pair CI.STC-TA.U11A and their differences (right panels). Note the color range of differences are set to be 1 percent of the kernel.

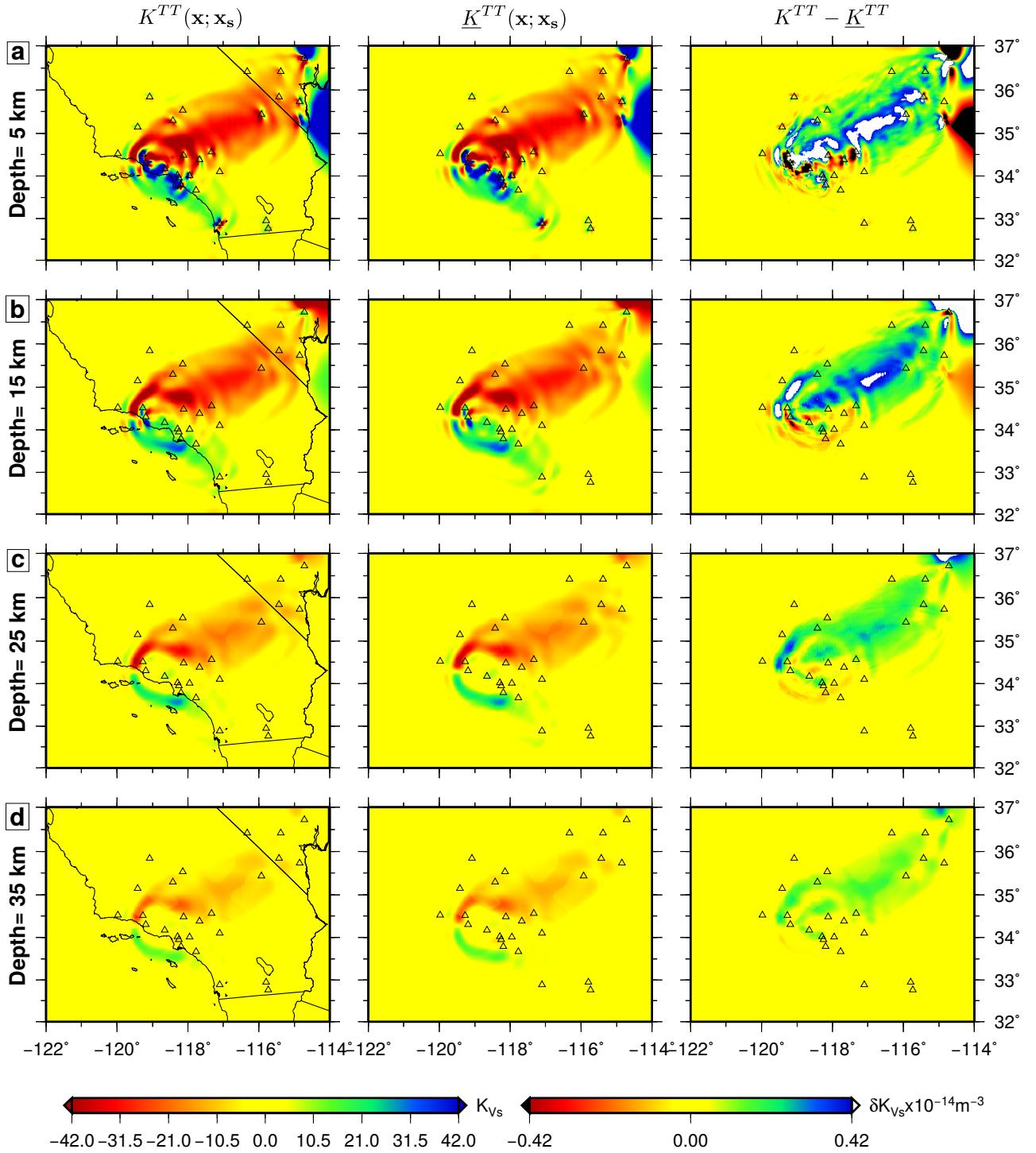


FIGURE 4.5: Rotated (left panels) and reference (middle panels) event misfit kernel of the master station CI.STC with 26 generic receivers (triangles) and their differences (right panels).

4.4.3 Event misfit kernel

For an event misfit kernel of T-T component measurements at receivers, the construction of a reference and rotated event kernel requires different number of numerical simulations. The reference T-T component event misfit kernel $\underline{K}^{TT}(\mathbf{x}, \mathbf{x}_s)$ is a summation of 26 reference individual misfit kernels, each of which needs one forward and one adjoint simulation. In comparison, the rotated T-T component event misfit kernel $K^{TT}(\mathbf{x}, \mathbf{x}_s)$ is obtained by summation of two event kernels, each of which is calculated by the interaction of a forward field from point-source force and an adjoint field from a properly constructed adjoint source injected at all receivers simultaneously. Figure 4.5 shows the rotated T-T event kernel $K^{TT}(\mathbf{x}, \mathbf{x}_s)$ (left panels), the reference event kernel $\underline{K}^{TT}(\mathbf{x}, \mathbf{x}_s)$ (middle panels), and their differences (right panels) at different depths. Similar to Figure 4.4, the errors are less than 1% and negligible in adjoint tomography.

In summary, we have demonstrated in this section that the rotated seismograms, individual and event misfit kernels using eq. (4.8), (4.17), and (4.19) are sufficient enough to approximate the corresponding reference ones.

4.5 Love-wave ANAT in Southern California

Based on the proposed method for calculating T-T component event kernel, we perform Love-wave ambient noise adjoint tomography in Southern California to obtain a V_{sh} model. In this section, we firstly introduce the data processing procedures to retrieve Love wave signals from ambient noise, and then briefly explain the inversion procedures. Lastly, we present the misfit revolution and our final V_{sh} model and derived radial anisotropy.

4.5.1 Data processing procedures

We first build a database of T-T component ambient-noise cross-correlation functions obtained from 148 stations in southern California (Figure 4.2a) with ambient noise data recorded from January 2006 to December 2012. First, north and east components of continuous long-period (LHN, LHE) seismic data at these stations are downloaded from southern California Earthquake Data Center (SCEDC, <http://service.scedc.caltech.edu/>) and processed using the methodology outlined in Bensen et al. [2007]. For each component (E and N), we first

cut them into a series of daily segments, then remove the trend, mean and instrument responses from the daily seismograms and bandpass filter them at the period band of 5 – 50 second. Afterwards, time domain normalization is performed to the pre-processed E- and N- component seismograms together (see Lin et al. [2008]), in order to suppress the effects of earthquake signals and instrumental irregularities on cross-correlation functions. Spectral whitening is further applied to balance the energy within our interested period range. At last, daily cross-correlation functions for each possible station pair are computed and stacked linearly to produce the final cross-correlation functions. The positive and negative legs of each stacked cross-correlation function are stacked further to obtain the symmetric cross-correlation functions.

In total, we obtain 10,869 T-T component cross-correlation functions with high quality Love wave signals, which are converted to the EGFs by taking its negative time derivatives as done in Wang et al. [2018b]. In this study, we assume the noise distribution is homogeneous, thus not considering the potential effects on EGFs from non-uniform noise sources [Wang et al., 2014a, 2016].

4.5.2 Inversion procedures

The inversion procedures applied here are similar to that of Rayleigh-wave ANAT, which is fully described in Wang et al. [2018b]. We summarize it briefly here with a focus on the Love wave data processing. (1) Forward simulation. For each master station, we conduct two forward simulations with point force sources in north and east direction, separately. The resulting four component seismograms at receivers are rotated to obtain T-T component Synthetic Green’s Functions (SGFs) using eq. (4.12). (2) Preprocessing. We measure the frequency-dependent traveltime misfit as well as adjoint source for each EGF-SGF pair using a multitaper technique at three frequency bands, namely 5 – 50 s, 10 – 50 s, and 20 – 50 s. The adjoint sources of three frequency bands are added up to form the final adjoint source. (3) Adjoint simulation. The T-T component adjoint sources are rotated to east and north directions, and used to generate adjoint fields, which are subsequently interacted with the forward fields by two adjoint simulations to obtain the two rotated event kernels, respectively. (4) Postprocessing. All rotated event kernels are summed, preconditioned and smoothed to obtain the final misfit gradient. (5) Line search. The optimal step length for model update is determined based on the total misfit curve of 19 stations (red triangles in Figure 4.2(b))

from a number of forward simulations based on trial models with different step lengths. (6) Model update. If the total misfit does not change too much, iteration stops; otherwise, go to step (1) with the new model using the optimize step length.

Following the above procedures, we iteratively update the initial model **M16** using a multi-scale strategy to improve data fitting. In the Rayleigh-wave type ANAT, we start our inversion with measurements made only from long period band of 20–50 s, and progressively add measurements at 10–50 and 5–50 period bands when the misfit reduction over iterations from existing band(s) becomes less evident. This strategy is effective and essential to avoid being trapped in local minima when the initial model has a relatively large misfit (average 2.1 s for 20 – 50 s band). However, for Love waves in this study, the misfits of our initial model are quite small (see Figure 4.7a). Thus, we decide to include the measurements at 10 – 50 s and 20 – 50 s period bands simultaneously in the first two iterations (**M16** to **M18**), and then add the 5 – 50 s band in the subsequent iterations. Figure 4.6(a)-(c) show some examples of the linear search results and Figure 4.6(d) displays the evolution of the total misfit reduction over iterations. The total misfit is calculated from all available measurements between station pairs in the three frequency bands (black lines in Figure 4.6). As can be seen from Fig 4.6, the total misfit decreases dramatically in the first three iterations before it slows down. After the fifth iteration, the total misfit almost has no change, thus we decide to stop the iteration.

4.5.3 Results

We obtain **M21** from the 5th iteration and take it as our final V_{sh} model. Combining the V_{sv} model from Rayleigh-wave ANAT, we further constructed a radial anisotropic model of southern California. In this subsection, we show the improvement of our model by comparing the cross-correlation timeshift histograms between **M16** and **M21**. Then, we describe some interesting features shown both in the V_{sh} and radial anisotropic velocity model.

Misfit

As shown in Figure 4.7(a), most time shifts of the three frequency bands for the initial model **M16** are distributed within the range of $[-2, +2]$ s, indicating that **M16** built solely on earthquake data, produces good fits between SGFs and EGFs. However, the asymmetric distribution of the misfit histograms as well as the relative large standard deviation (STD) indicates that there is still room for improvement. For example, at 5–50 and 10–50 s period bands, the

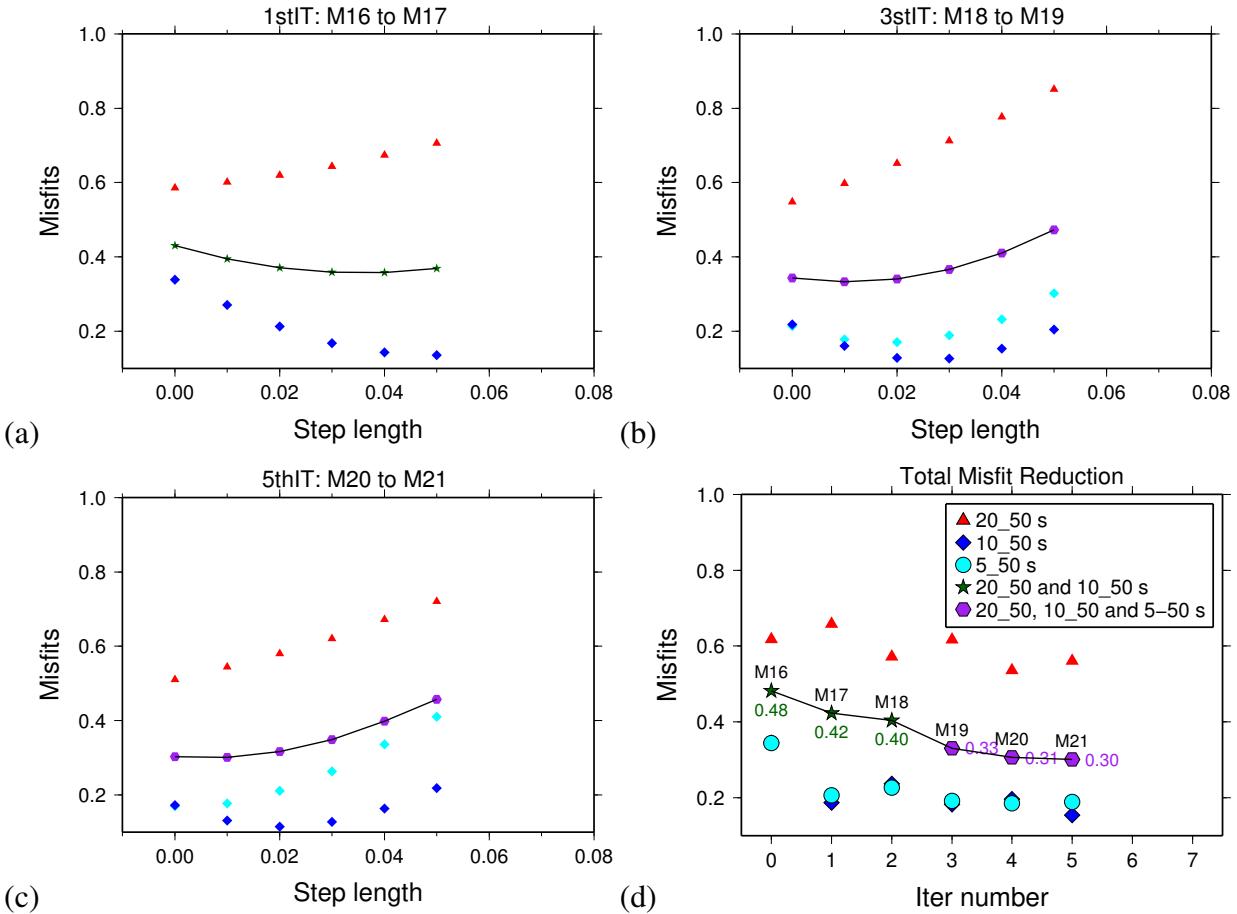


FIGURE 4.6: (a-c) Line search results at first (1stIT), third (3stIT), and fifth (5thIT) iteration showing the variation of misfit function values as a function of step length for various period bands and the total misfit (also connected by solid curves);(d) Total misfit reduction over iterations.

timeshift distributions are slightly biased towards the negative axis, which means the shallow structure of the model is relatively faster than **M16**. The situation is opposite for the 20 – 50 s band, at which the distribution is biased towards positive axis, suggesting the velocities are relatively slower than **M16** at deeper depths.

To check the improvements of our inversion results, we make a histogram comparison between **M16** and our final model **M21** as shown in Figure 4.7(b). The most significant improvements are obtained in the 10 – 50 and 5 – 50 second bands, for which the timeshifts are drastically reduced, with the mean and STD values changing from -0.45 ± 0.66 s to 0.04 ± 0.51 s for 5 – 50 s band (Figure 4.7(b), left), and from -0.48 ± 0.75 s to -0.17 ± 0.60 s for 10 – 50 s band (Figure 4.7(b), middle). In the long period band of 20 – 50 s, the timeshift reductions are relatively small, with the mean and STD values reduced from 0.59 ± 1.66 s to 0.48 ± 1.64 s.

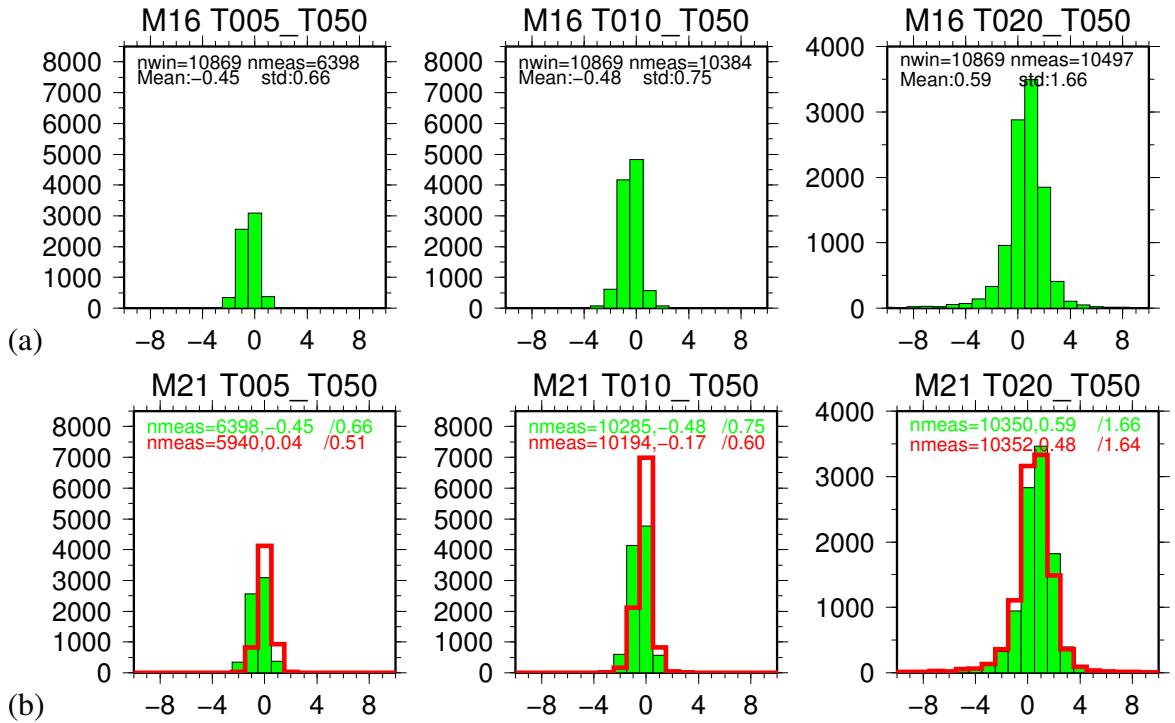


FIGURE 4.7: Travetime misfit histograms obtained for model **M16** (a), **M21** (b). The green solid bars represent misfits of the initial model, and the red bars in (b) are misfits of the final model. Misfits are measured at all three period bands, 5 – 50 (left), 10 – 50 (middle), and 20 – 50 (right) s. CC_{min} in (a) are specially chosen as 0.80 (5 – 50 s), 0.75 (10 – 50 s), 0.75 (20 – 50 s) to show the misfit of the initial model. All the misfits are measured with $CC_{min} = 0.80$ in (b) for comparison.

V_{sh} model and Radial anisotropy

Figure 4.8 displays the absolute shear wave speed of **M16** (left column), **M21** (central column) and the differences between the two models in percentage (right column). The models are masked based on amplitudes of volumetric coverage kernels as shown in Rayleigh-wave type ANAT. As inferred from the negative biases in the timeshift distributions for 5 – 50 and 10 – 50 s bands (Figure 4.7a), the Vs velocities in our model are generally enhanced at most area in the upper crust (5 – 15 km), including the southernmost edge of San Joaquin Valley (SJV) and Sierra Nevada Batholith (SNB), most of the Mojave Desert, and the Salton Trough Basin (STB). The maximum variations reach as much as $\sim +6\%$ at the southern area of SJV and SNB at 5 km depth. Absolute velocities at the Western Transverse Range (WTR), the Eastern Transverse Range (ETR) and the eastern Peninsular Ranges Batholith (ePRB) are slowed down by $\sim 4\%$ at the middle crust (15 km). In the lower crust and uppermost mantle (≥ 25 km), Vs generally becomes slower, which is particularly evident at CTR, ETR and Los Angeles Basin (LAB) where the variations reach as much as -4% . This is reasonable considering the positive mean of the misfit for **M16** at 20 – 50 s period band.

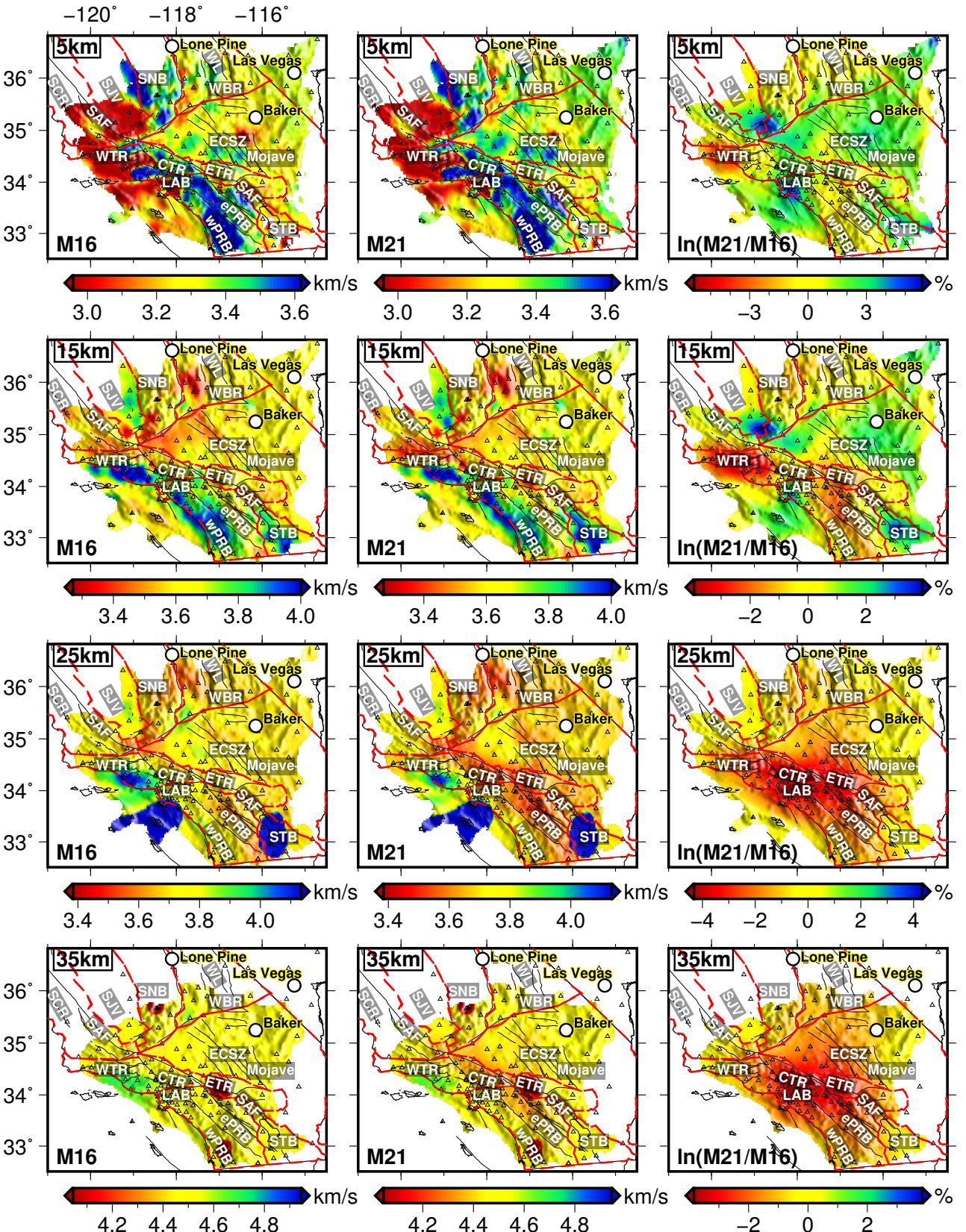


FIGURE 4.8: Horizontal slices of shear velocity for model M16 (left panel), M21 (middle panel), and $\ln(M21/M16)$ (right panel) at 5, 15, 25, and 35 km depths.

The discrepancies between V_{sv} from Rayleigh waves and V_{sh} from Love waves give rise to radial anisotropy. Figure 4.9 shows V_{sv} model from Rayleigh-wave ANAT, V_{sh} model from this study and radial anisotropy defined as $2 * (V_{sh} - V_{sv})/(V_{sv} + V_{sh})$ at different depths. In the uppermost crust (5 km), we observed three distinct positive ($V_{sh} > V_{sv}$) radial anisotropic zones ($\sim +4\%$) concentrated at the southern edge of SJV and SNB, southwestern coast of LAB, and eastern Mojave. In the middle and lower crust (15 – 25 km), the whole Transverse Range shows negative ($V_{sh} < V_{sv}$) radial anisotropy with an amplitude of $\sim -6\%$; whereas the rest region in the northern-east exhibits $\sim +4\%$ positive radial anisotropy, including the Western Basin Range (WBR) and eastern Mojave. In the uppermost mantle, most areas show positive radial anisotropy, and this is especially evident at the Eastern California Shear Zone (ECSZ), where the amplitude reaches as much as $\sim +7\%$.

4.6 Discussions

We show that only two forward and two adjoint simulations are needed to calculate either the individual or event kernels for T-T or R-R component EGF measurements. The derivation can be easily extended to measurements on EGFs across components (Appendix A), although these type of EGFs are used much less in practice. In terms of the SEM simulations, each T-T or R-R type of event kernel costs twice amount of simulation time as a Z-Z type of kernel, although the forward simulations for east and north component point sources only need to be run once if both T-T and R-R kernels are to be constructed.

In our derivation, we simply use Green's functions from point-force sources as the synthetics for EGFs, ignoring the influence of heterogeneous noise sources [e.g., Wang et al., 2016, Yang and Ritzwoller, 2008]. However, such noise source effects can be accounted for by extending our method with ensemble forward and adjoint wavefield [Tromp et al., 2010] or transforming the actual wavefield sources and wave propagation physics into effective sources and wavefield [Fichtner et al., 2016].

In this study, we made a simple assumption that Rayleigh waves are only sensitive to V_{sv} structure while Love waves are only sensitive to V_{sh} structure, and we perform adjoint tomography of the two type of waves separately based on 3-D Fréchet kernel from isotropic model. A more accurate and elegant way is to use a radial anisotropic initial model during the adjoint tomography, and fit the waveforms of Rayleigh and Love waves simultaneously.

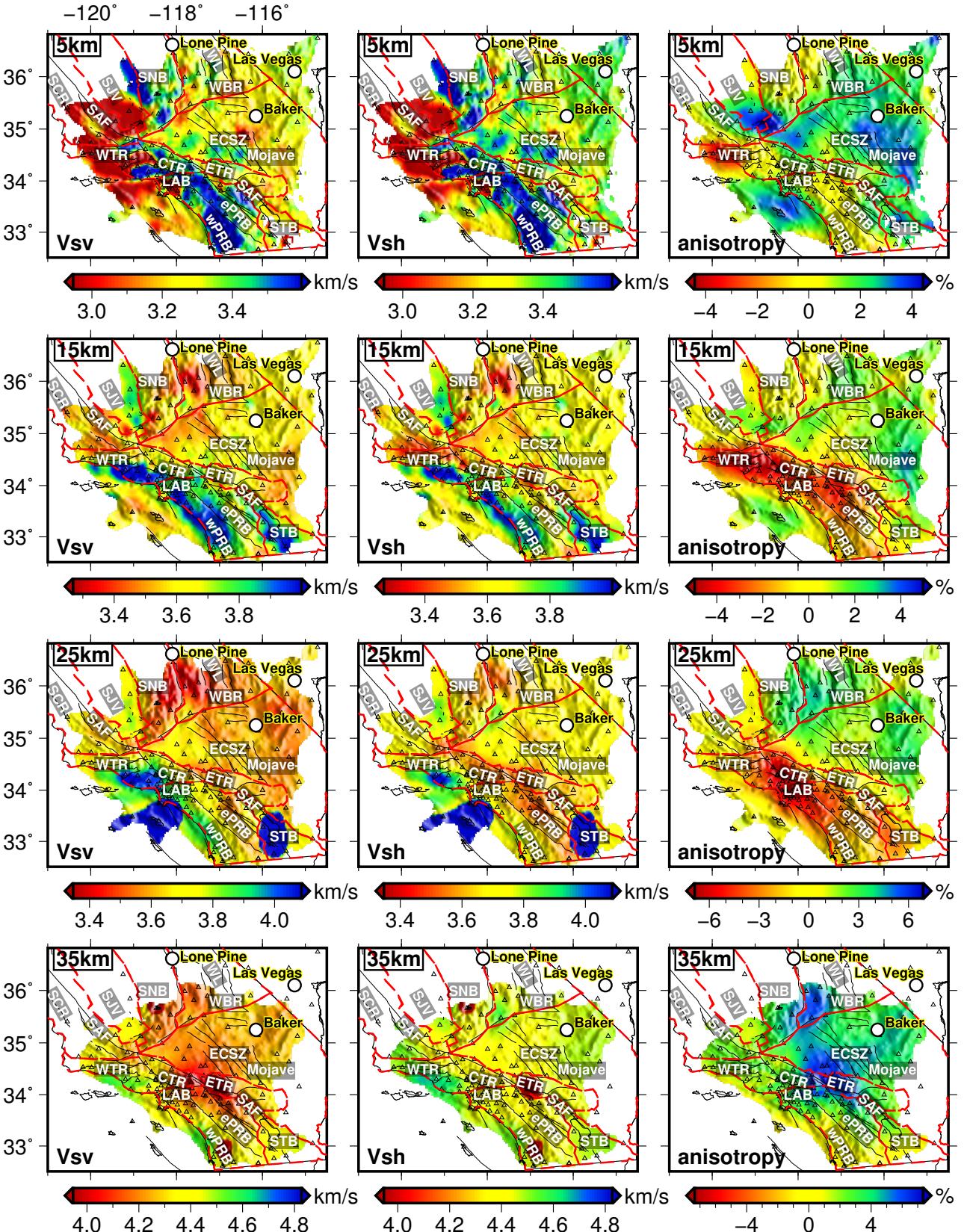


FIGURE 4.9: Horizontal slices of V_{sv} from Rayleigh wave (left panel), V_{sh} from Love waves (middle panel), and radial anisotropy (right panel) at 5, 15, 25, and 35 km depths.

Nevertheless, the radial anisotropies in the crust and uppermost of Southern California we observe are highly consistent with previous studies, and show more interesting details in the crust. Compared to Moschetti et al. [2010a], we observed similar features of positive radial anisotropy in the entire uppermost mantle ($\sim +7\%$) and the north-eastern part of the crust ($\sim +4\%$). Seismic anisotropy in the mantle is mostly attributed to the lattice-preferred orientation of mantle anisotropic minerals (e.g. olivine) and is used to infer characteristics about the mantle flow field [Moschetti et al., 2010a]. Moschetti et al. [2010a] interpret the observed crustal radial anisotropy as resulting from the lattice-preferred orientation of seismically anisotropic crustal minerals (e.g., micas, amphiboles) induced by the finite strains accompanying extension. They suggest that the deep-crustal response to extension in the western United States is widespread and relatively uniform. Moreover, our results reveal $\sim 6\%$ negative radial anisotropies along the whole WTR in the middle-lower crust, which were not observed by Moschetti et al. [2010a]. A more detailed interpretation along with geologic features will be explored in following studies.

4.7 Conclusion

In this study, we propose an efficient method for calculating multi-component (e.g., T-T, R-R) sensitivity kernels for EGFs from ambient noise based on adjoint method. We demonstrate that a T-T component kernel can be obtained by the summation of two kernels, each of which is constructed from the interaction of a forward field due to a delta-function point force in the east or north direction, and an adjoint field by injecting T-T component adjoint source at individual receiver in the east and north direction with proper amplitude modulation factors. Therefore, two forward simulations and two adjoint simulations are required to compute the sensitivity kernel for T-T component measurements. The sensitivity kernel for R-R component measurement is similar to those of T-T component kernels, but with R-R component adjoint source and different amplitude modulation factors. To validate our method, we conduct three groups of numerical simulations to validate the rotation relations for seismograms, individual and event misfit kernels. Our results show the rotated kernels are almost identical to the reference ones with the error less than 1%, which is sufficiently accurate for application.

Utilizing the method for calculating T-T component sensitivity kernel, we have successfully performed Love-wave ANAT in Southern California and obtained a V_{sh} model after five iterations. Compared with V_{sv} model from Rayleigh-wave ANAT, we observe positive radial anisotropies in the whole uppermost mantle ($\sim 7\%$) and the north-eastern part of the crust ($\sim 4\%$). Moreover, our results reveal $\sim 6\%$ negative radial anisotropies along the whole WTR in the middle-lower crust, which were not observed by Moschetti et al. [2010a].

With increasing dense arrays across the globe providing continuous ambient noise field, the developed method can be readily applied to multi-component ambient noise adjoint tomography for high resolution anisotropic studies, especially for resolving complex crustal anisotropic structures.

4.8 Appendix A: Auxiliary matrix

The auxiliary matrix related to the other multi-component EGFs are given as

$$C^{RT}(\theta, \theta') = \begin{bmatrix} \sin \theta \cos \theta' & -\sin \theta \sin \theta' & 0 \\ \cos \theta \cos \theta' & -\cos \theta \sin \theta' & 0 \\ 0 & 0 & 0 \end{bmatrix}, C^{TR}(\theta, \theta') = \begin{bmatrix} \cos \theta \sin \theta' & \cos \theta \cos \theta' & 0 \\ -\sin \theta \sin \theta' & -\sin \theta \cos \theta' & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$C^{ZT}(\theta, \theta') = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \cos \theta' & -\sin \theta' & 0 \end{bmatrix}, C^{ZR}(\theta, \theta') = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sin \theta' & \cos \theta' & 0 \end{bmatrix}$$

$$C^{TZ}(\theta, \theta') = \begin{bmatrix} 0 & 0 & \cos \theta \\ 0 & 0 & -\sin \theta \\ 0 & 0 & 0 \end{bmatrix}, C^{RZ}(\theta, \theta') = \begin{bmatrix} 0 & 0 & \sin \theta \\ 0 & 0 & \cos \theta \\ 0 & 0 & 0 \end{bmatrix}$$

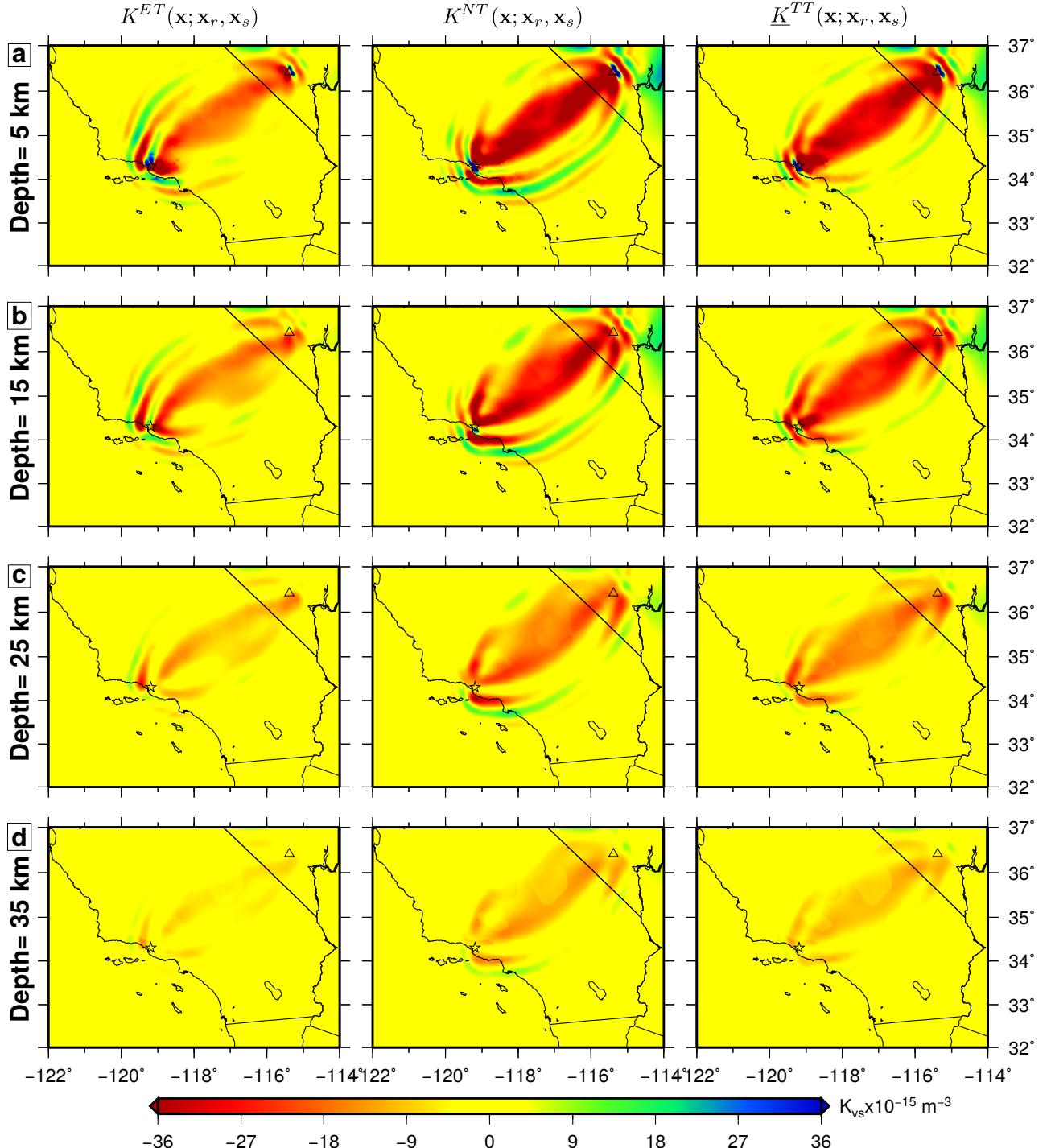


FIGURE 4.A1: Individual sensitivity kernels of $K^{ET}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (left panels), $K^{NT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (middle panels), and $K^{TT}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s)$ (right panels) for station pair CI.STC-TA.U11A at the depths of 5 km (a), 15 km (b), 25 km (c) and 35 km (d).

5

Numerical investigation of three-dimensional Rayleigh wave ellipticity kernel based on SEM adjoint method

5.1 Summary

Rayleigh wave ellipticity is exploited to constrain the shallow structure beneath a single station by assuming the lateral variations of its sensitivities in the horizontal direction can be neglected. However, earlier studies have demonstrated that the horizontal variations of its sensitivities in a heterogeneous medium can be significant. Thus, in order to achieve more accurate and higher resolution results using Rayleigh wave ellipticity, one needs to consider the lateral variations in the inversion for structures. In this study, we investigate the characteristics of 3-D Rayleigh wave ellipticity kernel using adjoint method in conjunction with

spectral-element based full wavefield simulations. Our results show that Rayleigh wave ellipticity kernels in a half-space or 1-D layered medium have sensitivities only around the receiver, and higher-mode interferences broaden the influence zone by extending it towards the source. However, in a 3-D heterogeneous medium, Rayleigh wave ellipticity has sensitivities over the whole elliptical region around the source-receiver path due to intrinsic scattering. In addition, Rayleigh wave ellipticity kernel has large sensitivities to S-wave velocity, relatively small but non-negligible sensitivities to density, and little sensitivities to P-wave velocity.

The 3-D ellipticity kernel has a complex shape in a lateral homogeneous medium. The horizontal slices show alternating positive and negative sidebands with an elliptical shape truncated at the receiver. The vertical transects along the source-receiver path show a positive-negative striped sensitivity tail beneath the receiver. By summing individual ellipticity kernels from different azimuths, an azimuthally averaged ellipticity kernel can be constructed, which has a quasi circular sensitivity zone with a pattern of alternating positive-negative sensitivities centered at the station. The circular influence zone is broadened by higher mode interferences in a 1-D layered model and becomes asymmetric and significantly disturbed due to strong scattering in a 3-D heterogeneous medium.

The variations of 3-D ellipticity kernels in both vertical and horizontal directions suggest that it is inaccurate to simply replace the 3-D kernel by using a classic 1-D depth kernel with sensitivities just beneath the receiver. 3-D sensitivity kernels based on numerical simulations are expected to be used in future ellipticity tomographic studies.

5.2 Introduction

Most surface wave studies extract phase or group velocity dispersion curves from Rayleigh or Love waves, and then invert them for depth-dependent shear wave velocities. Surface wave dispersion measurements can be extracted from both earthquakes [e.g., Ritzwoller et al., 2002, Yang and Forsyth, 2006, Zhou et al., 2006] and ambient noise data [e.g., Bensen et al., 2007, Jiang et al., 2014, Lin et al., 2007, Moschetti et al., 2010a, Shapiro et al., 2005, Shen et al., 2013, Yang et al., 2007, Yao et al., 2006], which has greatly improved our capability of imaging the crustal and upper mantle structures of the Earth. However, in regional- and global- scale studies, only surface waves at period longer than ~ 5 s can be obtained, which provide weak constraints on the shallow crustal structure. On the other hand, Rayleigh wave

ellipticity, typically defined as the ratio of horizontal and vertical component amplitude (H/V ratio), has been proved by theoretical studies [e.g., Boore and Toksöz, 1969, Malischewsky and Scherbaum, 2004, Tsuboi and Saito, 1983] to be more sensitive to shallow structures than phase/group velocity dispersion data. It helps better constrain near-surface velocities, which is important in studying surface geology, characterizing ground motions, etc [Lin et al., 2012].

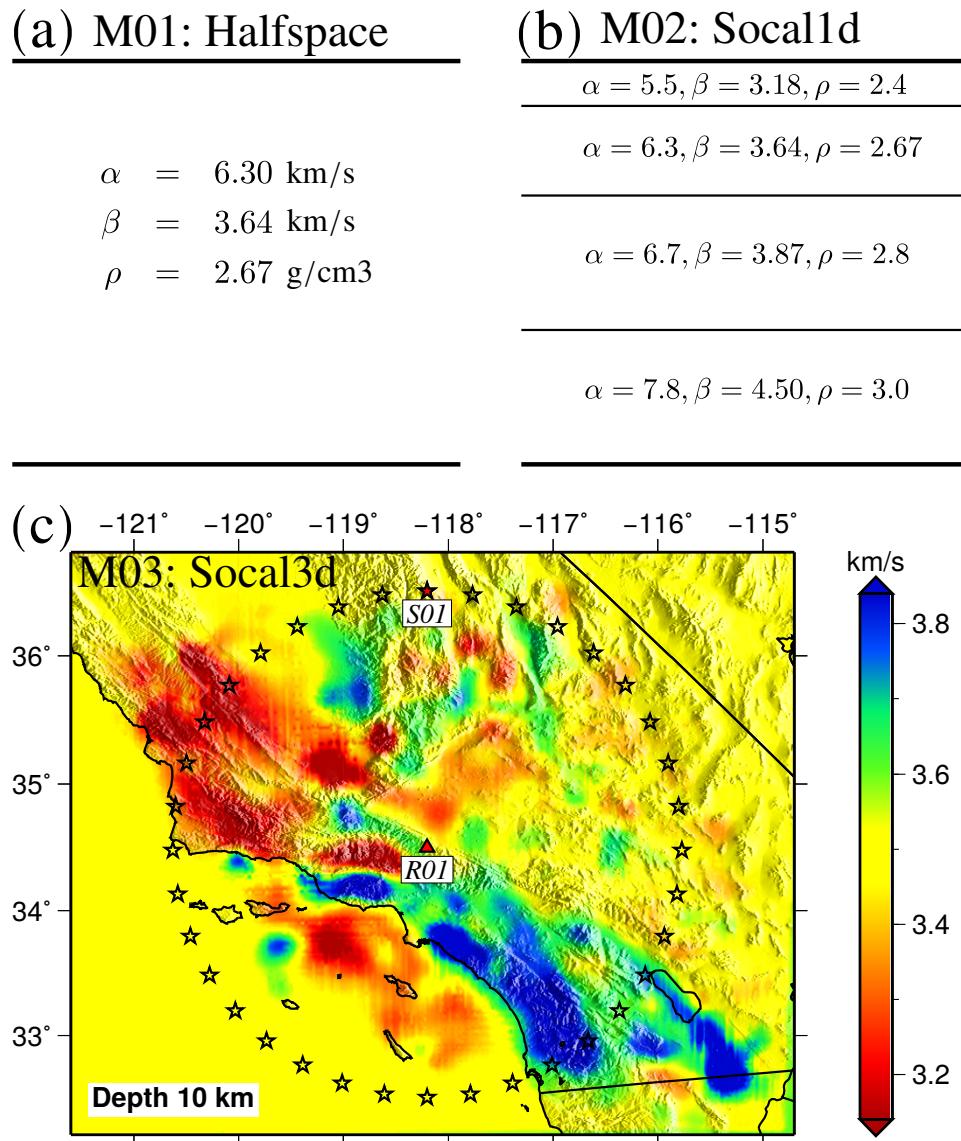


FIGURE 5.1: Three test models used in numerical simulations. (a) a half space model (M01); (b) a 1-D layered southern California model named Socal1d (M02); (c) a 3-D southern California model named Socal3d (M03). α, β, ρ denote compression-wave, shear-wave velocity, and density. M03 is derived from the 3-D model of Tape et al. [2010].

Studies of Rayleigh wave ellipticity can be dated back to the 1960s. Boore and Toksöz [1969] found that the ellipticity is more sensitive to near-surface structures beneath a receiver

than phase velocities, thus providing additional constraints on shallow structures than dispersion data. Rayleigh wave ellipticity did not attract as much attentions as phase/group dispersions in early days [Sexton et al., 1977], mainly due to the low-quality of multi-component seismograms and the lack of reliable techniques in measuring ellipticity data accurately. Recently, with the developments of high-quality three-component seismometers and advances in modern data processing techniques, Rayleigh wave ellipticity can be extracted very well and has been applied to constrain local structures beneath a single station at both short-period (5 – 50 s) and long-period (to \sim 150 s) bands [e.g., Ferreira and Woodhouse, 2007, Tanimoto and Alvizuri, 2006, Tanimoto and Rivera, 2008]. Especially, in ambient noise tomography, Rayleigh wave ellipticity has been used to infer density and shear-wave velocities at shallow depths [e.g., Li et al., 2016, Lin et al., 2012, 2014, Muir and Tsai, 2017].

Traditional methods using Rayleigh-wave ellipticity to constrain shallow structure is based on 1-D depth sensitivity kernels for model parameters (density, P-wave and S-wave velocity). These 1-D sensitivity kernels are constructed either by numerical differential methods [e.g., Boore and Toksöz, 1969] or by variational principle [e.g., Tanimoto and Tsuboi, 2009, Tsuboi and Saito, 1983], based on ray theory assumption for surface waves propagating over a 1-D layered medium. However, ray theory becomes less valid when the wavelength of surface wave is smaller than the scale of structural heterogeneities as shown by classic finite-frequency theory [e.g., Dahlen et al., 2000, Zhou et al., 2004]. To solve this, Maupin [2017] investigated the 3-D sensitivity kernels of Rayleigh wave ellipticity based on surface wave mode summation with near-field terms included [Liu and Zhou, 2016, Zhou et al., 2004]. However, the method of calculating a 3-D sensitivity kernel of Rayleigh wave ellipticity proposed by Maupin [2017] is only valid for a lateral homogeneous medium. For a highly heterogeneous medium, a method of calculating 3-D sensitivity kernels considering heterogeneities is greatly desired.

With the development of high-performance computation, it has now become practical to derive more accurate Fréchet kernels based on numerical simulations such as spectral element method (SEM) [Komatitsch and Tromp, 1999]. Finite-frequency kernels have been derived by scattering integral method based on finite-difference [Chen et al., 2007a, Zhao et al., 2005] and adjoint method based on SEM [Liu and Tromp, 2006, 2008, Tromp et al., 2005]. Adjoint tomography has been successfully applied to both earthquakes [e.g., Chen et al., 2015, Tape et al., 2009] and ambient noise data [e.g., Chen et al., 2014, Wang et al.,

2018b] to construct more refined structure than traditional methods.

In this study, we demonstrate a method to numerically construct the 3-D kernel of Rayleigh wave ellipticity by adopting SEM-based adjoint method. The key questions to be addressed here includes: (1) what is the shape of 3-D Rayleigh wave ellipticity kernel ? (2) how higher-mode interferences and heterogeneous structures affect the 3-D Rayleigh wave ellipticity kernel? In addition, we will compare 3-D ellipticity kernels with 1-D depth sensitivity kernels. We will firstly introduce the adjoint method to calculate Rayleigh wave ellipticity kernel in section 5.3. Then, we present our numerical simulations in section 5.4 followed by discussions in section 5.5.

5.3 Methodology

In this section, we will first introduce the basic idea of constructing individual amplitude kernel and differential amplitude kernel based on adjoint method [Tromp et al., 2005]. Then, based on the theoretical framework of multi-component sensitivity kernel of Wang et al. [2018a], we derive the two variants of sensitivity kernels for Rayleigh wave ellipticity.

5.3.1 Amplitude kernel

If we assume $\delta \ln A$ to be a scalar function that quantifies the logarithmic amplitude residual between observations and synthetic predictions based on a reference (or current) model $m(\mathbf{x})$, then it can be linearized relative to the reference model as:

$$\delta \ln A = \int_V K_A(\mathbf{x}) \delta \ln m(\mathbf{x}) d^3\mathbf{x}, \quad (5.1)$$

where $K_A(\mathbf{x})$ is the Fréchet amplitude kernels defined for fractional model parameters, such as density, shear and bulk moduli perturbations. Based on adjoint method [Liu and Tromp, 2006, Tromp et al., 2005], these kernels can be expressed in a general form:

$$K_A(\mathbf{x}) = \mathbf{u}(\mathbf{x}, t; \mathbf{x}_s) \otimes \mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r), \quad (5.2)$$

where $\mathbf{u}(\mathbf{x}, t; \mathbf{x}_s)$ is the forward wavefield emitted from the source, $\mathbf{u}^\dagger(\mathbf{x}, T-t; \mathbf{x}_r)$ is the adjoint wavefield generated by the time-reversed amplitude adjoint sources injected at receivers. The \otimes operator represents general interaction for different types of kernel which can be found in Tromp et al. [2005].

5.3.2 Differential amplitude kernel

For differential amplitude measurements, the residual of the logarithm of the amplitude ratio, relative to the background unperturbed earth model, is expressed as

$$\delta \ln(A'/A) = \int_V K_{A'/A}(\mathbf{x}) \delta \ln m(\mathbf{x}) d^3\mathbf{x}, \quad (5.3)$$

where we use prime (A') and unprimed (A) subscripts of amplitude to represent measurements of different phases or components. Since $\ln(A'/A) = \ln A' - \ln A$, we have

$$K_{A'/A} = K_{A'} - K_A \quad (5.4)$$

Eq. (5.4) states that the differential amplitude kernel is simply the difference of the individual amplitude kernels from different phases or components.

5.3.3 Rayleigh-wave ellipticity kernel

Following [Tanimoto and Rivera, 2008], we define the Rayleigh wave ellipticity (E) as the ratio of vertical to radial amplitude:

$$E = \frac{A_{ZZ}}{A_{ZR}} = \frac{A_{RZ}}{A_{RR}} \quad (5.5)$$

In the above four multi-component amplitudes, the first letter of the subscript represents the direction of point force at the source and the second letter denotes the component at the receiver. According to eq. (5.3) and (5.4), the receiver-side ellipticity kernel has two variants:

$$K_E^Z = K_A^{ZZ} - K_A^{ZR} \quad (5.6)$$

$$K_E^R = K_A^{RZ} - K_A^{RR}, \quad (5.7)$$

where K_E^Z, K_E^R are a vertical and radial point force source generated ellipticity kernel, respectively. Based on the general expressions of multi-component sensitivity kernels derived by Wang et al. [2018a], the sensitivity kernel of the J -th component from a source in I -th direction is constructed by the interaction of forward and adjoint fields excited by point sources in geographic directions (north, east, vertical) with corresponding auxiliary matrix.

$$\begin{aligned} K^{IJ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= \mathbf{u}^I(\mathbf{x}, t; \mathbf{x}_s) \otimes (\mathbf{u}^J)^{\dagger}(\mathbf{x}, T-t; \mathbf{x}_r) \\ &= \mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s) \otimes \left\{ [C_{jk}^{IJ}(\theta, \theta') f^{IJ}(\mathbf{x}_r, T-t; \mathbf{x}_s)] * \mathbf{u}_k(\mathbf{x}, T-t; \mathbf{x}_r) \right\}, \end{aligned} \quad (5.8)$$

where $\mathbf{u}_j(\mathbf{x}, t; \mathbf{x}_s)$ is the forward field generated by a Delta-function point force source placed in the x_j direction and $\mathbf{u}_k(\mathbf{x}, t; \mathbf{x}_r)$ is the adjoint field generated by a Delta-function point force source placed in the x_k direction. $C_{jk}^{IJ}(\theta, \theta')$ is the auxiliary matrix of amplitude modulation factors in terms of azimuth at source (θ) and receiver (θ'), $f^{IJ}(\mathbf{x}_r, T - t; \mathbf{x}_s)$ is the adjoint source and $*$ represents time convolution.

Given the auxiliary matrix in the appendix of Wang et al. [2018a], we have the following four individual sensitivity kernels:

$$K^{ZZ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) = \mathbf{u}_z(\mathbf{x}, t; \mathbf{x}_s) \otimes \{f^{ZZ}(\mathbf{x}_r, T - t; \mathbf{x}_s) * \mathbf{u}_z(\mathbf{x}, T - t; \mathbf{x}_r)\} \quad (5.9)$$

$$\begin{aligned} K^{ZR}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= \mathbf{u}_z(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\sin \theta' f^{ZR}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\ &\quad + [\cos \theta' f^{ZR}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\} \end{aligned} \quad (5.10)$$

$$\begin{aligned} K^{RZ}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= \mathbf{u}_e(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\sin \theta f^{RZ}] * \mathbf{u}_z(\mathbf{x}, T - t; \mathbf{x}_r) \\ &\quad + \mathbf{u}_n(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\cos \theta f^{RZ}] * \mathbf{u}_z(\mathbf{x}, T - t; \mathbf{x}_r)\} \end{aligned} \quad (5.11)$$

$$\begin{aligned} K^{RR}(\mathbf{x}; \mathbf{x}_r, \mathbf{x}_s) &= \mathbf{u}_e(\mathbf{x}, t; \mathbf{x}_s) \otimes \{[\sin \theta \sin \theta' f^{RR}] * \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\ &\quad + [\sin \theta \cos \theta' f^{RR}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r) \\ &\quad + \mathbf{u}_n(\mathbf{x}, t) \otimes \{[\cos \theta \sin \theta' f^{RR}] * (\mathbf{x}_r, t)\} \mathbf{u}_e(\mathbf{x}, T - t; \mathbf{x}_r) \\ &\quad + [\cos \theta \cos \theta' f^{RR}] * \mathbf{u}_n(\mathbf{x}, T - t; \mathbf{x}_r)\} \end{aligned} \quad (5.12)$$

These kernel expressions work for different types of measurements, such as traveltime, amplitude, etc. Substituting eqs. (5.9)-(5.12) into (5.6-5.7), we can obtain the differential amplitude kernel for Rayleigh wave ellipticity.

5.4 Numerical results

In this section, we present the main results of our numerical simulations. We begin with the classic Lamb's problem, in which a vertical point force is inserted at the surface of a half-space homogeneous model to generate only fundamental mode Rayleigh waves. Special attentions are paid on the frequency dependent particle motion and spatial shape of the 3-D ellipticity kernel for fundamental mode Rayleigh waves. Then, the same simulations are conducted for 1-D layered and 3-D heterogeneous models separately to investigate the influences of higher-mode interferences and medium complexities on Rayleigh wave ellipticity kernel. Last, we investigate the characteristics of azimuthally averaged ellipticity kernel.

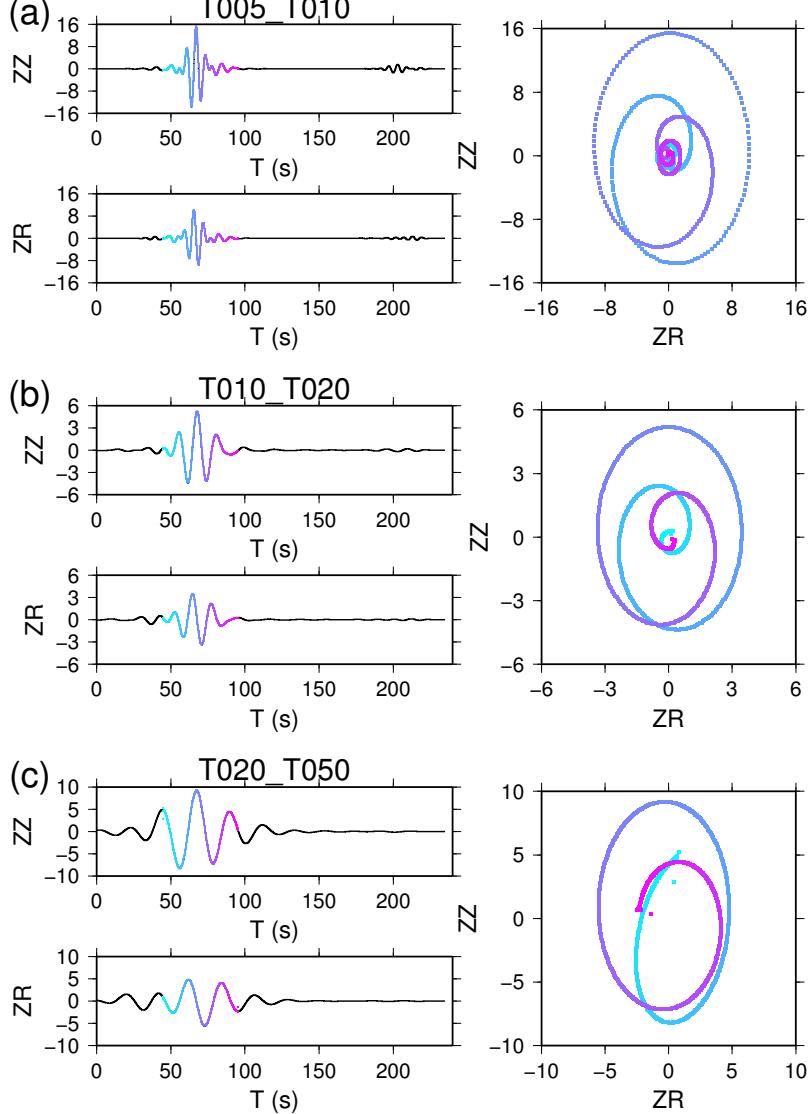


FIGURE 5.2: Seismograms of ZZ and ZR components and their particle motions filtered at (a) 5-10 s; (b) 10-20 s; (c) 20-50 s.

Throughout this section, we use a vertical point force as the source to generate the ellipticity kernel at a receiver as shown in eq. (5.6). The results should be the same if a radial point force is used as the source following eq. (5.7).

5.4.1 Parameters setting

We conduct three groups of numerical simulations based on different models, namely M01, M02 and M03 as shown in Figure 5.1. M01 is a half space model with constant model parameters, M02 is a 1-D layered Southern Californian model and M03 is a 3-D earthquake-based tomographic model from [Tape et al., 2009, 2010]. All numerical simulations in this

study are performed in a domain for Southern California (Figure 5.1c) based on SEM [Komatitsch and Tromp, 1999] as we did in Wang et al. [2018b]. The simulation domain has a horizontal scale of 639 km in longitude and 503 km in latitude and extends to the depth of 60 km. To save computation cost, we use irregular finite-element mesh with increasing vertical spacing from about 2 km at the surface to about 12 km at the bottom [Wang et al., 2018a,b]. For a single kernel, a vertical point source is inserted at the surface of a source WK.S01 ($-118.2^\circ, 36.5^\circ$) and three-component seismograms are collected at a station WK.R01 ($-118.2^\circ, 34.5^\circ$). A circle of sources with a radius of 2° relative to WK.R01 are chosen to generate azimuthally averaged ellipticity kernel.

5.4.2 Lamb's problem— a half space model

Figure 5.2 shows the frequency dependent seismograms of ZZ and ZR and their particle motions at station WK.R01 from the source WK.S01 in the half space model M01. We can see that the particle motions at the free surface are an ellipse with the amplitude of ZZ components larger than ZR components. Upon applying a Hilbert transform to the radial components, we find they have a $\pi/2$ phase advance relative to the vertical components. Another feature identified here is that the particle is moving (from blue to magenta in Figure 5.2) in a way referred to as "retrograde particle motion". By assessing the arrival times of the envelope peak of ZZ components at various frequency bands, we can see that the group velocity does not change with period which means the fundamental mode Rayleigh wave in a homogeneous half space is non-dispersive .

The 3-D amplitude sensitivity kernels (10 – 20 s) of ZZ and ZR component and their differences are displayed in Figure 5.3. The ZZ component amplitude kernel exhibits a strong-weak-strong pattern along the source-receiver path and alternating positive and negative sidebands with elliptical shapes and symmetric energy across the ray path, which is the same as that shown by Zhou et al. [2004]. Different from the ZZ kernel, the ZR component amplitude kernel is not as smooth as that of ZZ component and there are some oscillations in the sidebands especially at the receiver side. The differences of the two amplitude kernels

can be explained by surface wave mode summation as indicated by Maupin [2017]:

$$\begin{aligned}\delta s_z(\mathbf{x}) &= \sum_m U^m(z) \delta \Phi_m(x, y) \\ \delta s_x(\mathbf{x}) &= \sum_m \{1/k_m[-V^m(z)\partial/\partial x + W^m(z)\partial/\partial x]\} \delta \Phi_m(x, y) \\ \delta s_y(\mathbf{x}) &= \sum_m \{1/k_m[-W^m(z)\partial/\partial x - V^m(z)\partial/\partial x]\} \delta \Phi_m(x, y),\end{aligned}\quad (5.13)$$

where $\delta s_{x,y,z}(\mathbf{x})$ are perturbed displacements, $\Phi_m(x, y)$ is the perturbation of m mode potential field, and k_m is the wavenumber of m mode. U^m , V^m are the vertical and horizontal component eigenfunctions of m mode Rayleigh waves respectively, and W^m is the eigenfunction of m mode Love waves. In a half space model, the vertical component only depends on the perturbed potential of fundamental mode Rayleigh waves, while the radial component has additional horizontal derivatives which enhance the singularity of the potential close to the station, and contains additional Love wave contributions [Maupin, 2017].

More horizontal slices of the differential amplitude kernel and vertical cross-sections along the ray path for different model parameters are displayed in Figure 5.4. In general, the horizontal cross-sections show alternating positive and negative sidebands with an elliptical shape truncated at the receiver. The vertical plane along the source-receiver path displays a positive-negative striped tail beneath the receiver. Another feature is that the ellipticity kernel is more sensitive to S-wave velocity than to density and P-wave velocity. However, the amplitude of ellipticity density kernel is non-negligible compared with S-wave kernel, which is different from phase velocity data that has much larger sensitivities with respect to S-wave velocity than density and P-wave velocity. This is why Rayleigh wave ellipticity is also used to constrain shallow density structure [Lin et al., 2012].

5.4.3 Higher-mode interferences and medium complexities

To investigate the influence of Rayleigh wave overtones on ellipticity kernels, we perform another group of numerical simulations for 1-D layered model M02, and the results are shown in Figures 5.5-5.6. Compared with the ellipticity kernels in a half space model, two main differences are identified: (1) the magnitude of the individual amplitude kernel of ZZ or ZR component increases as a result a superposition of higher-modes on fundamental mode; (2) the energy of sidebands across the ray path is still symmetric, however some oscillations emerge at ZZ component kernel and the vibrations become more evident at ZR component

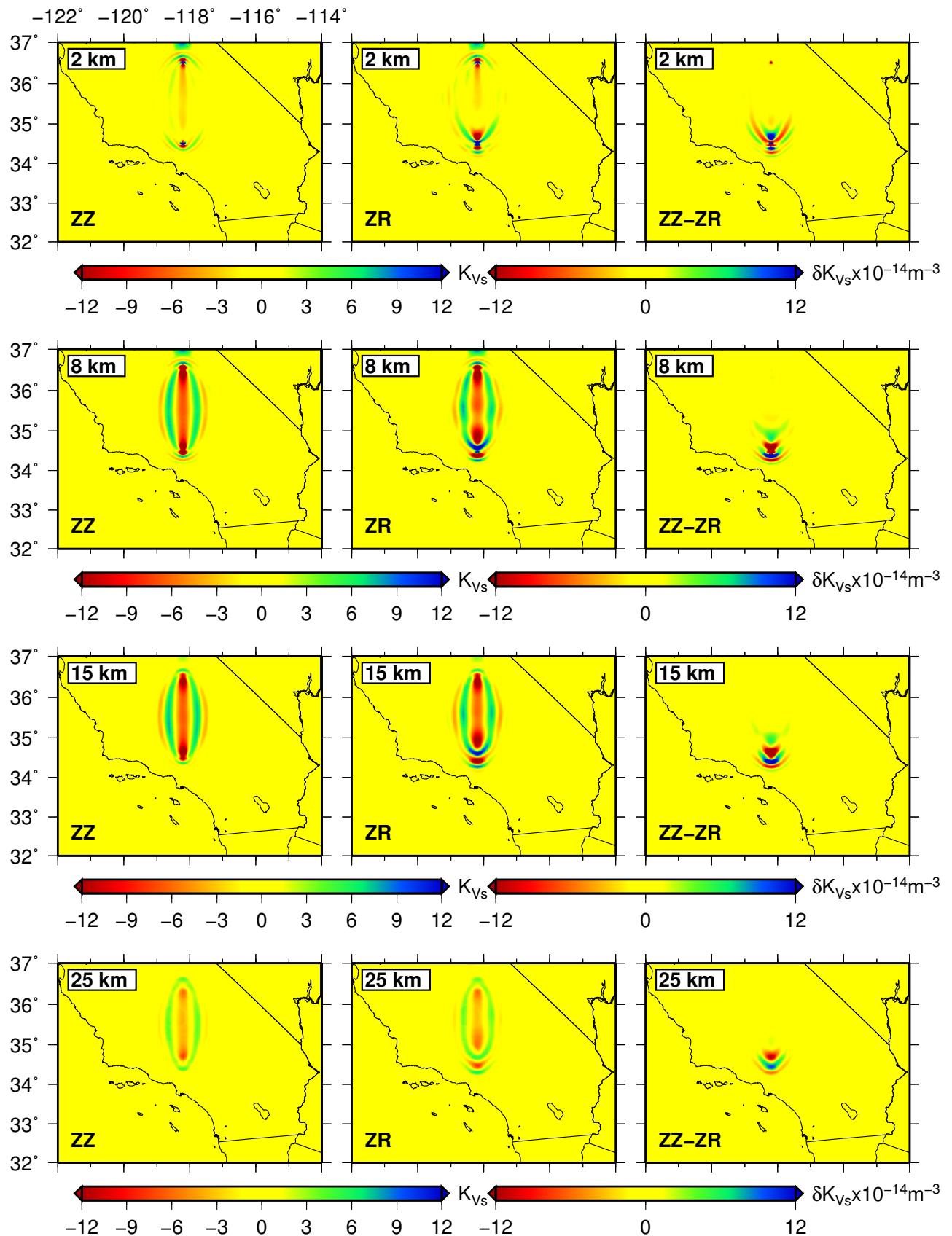


FIGURE 5.3: Sensitivity amplitude kernels of ZZ (first column) and ZR component (second column) and their differences (third column) at different depths filtered at 10-20 s in a half space model.

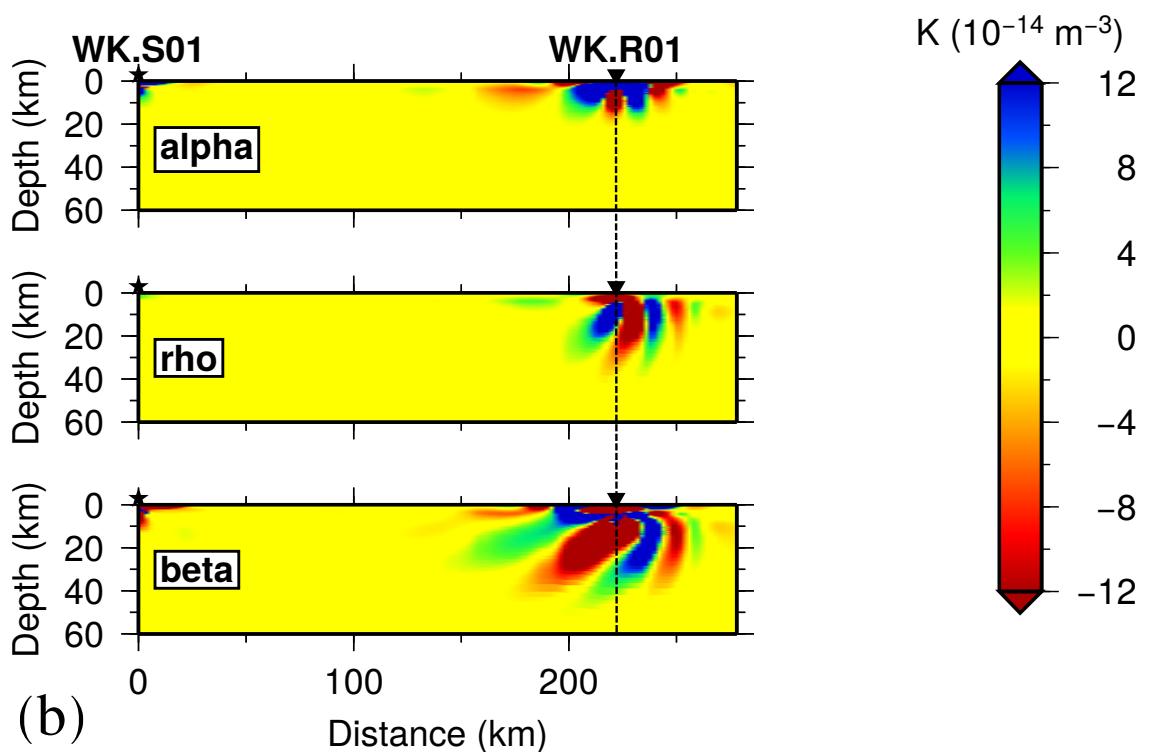
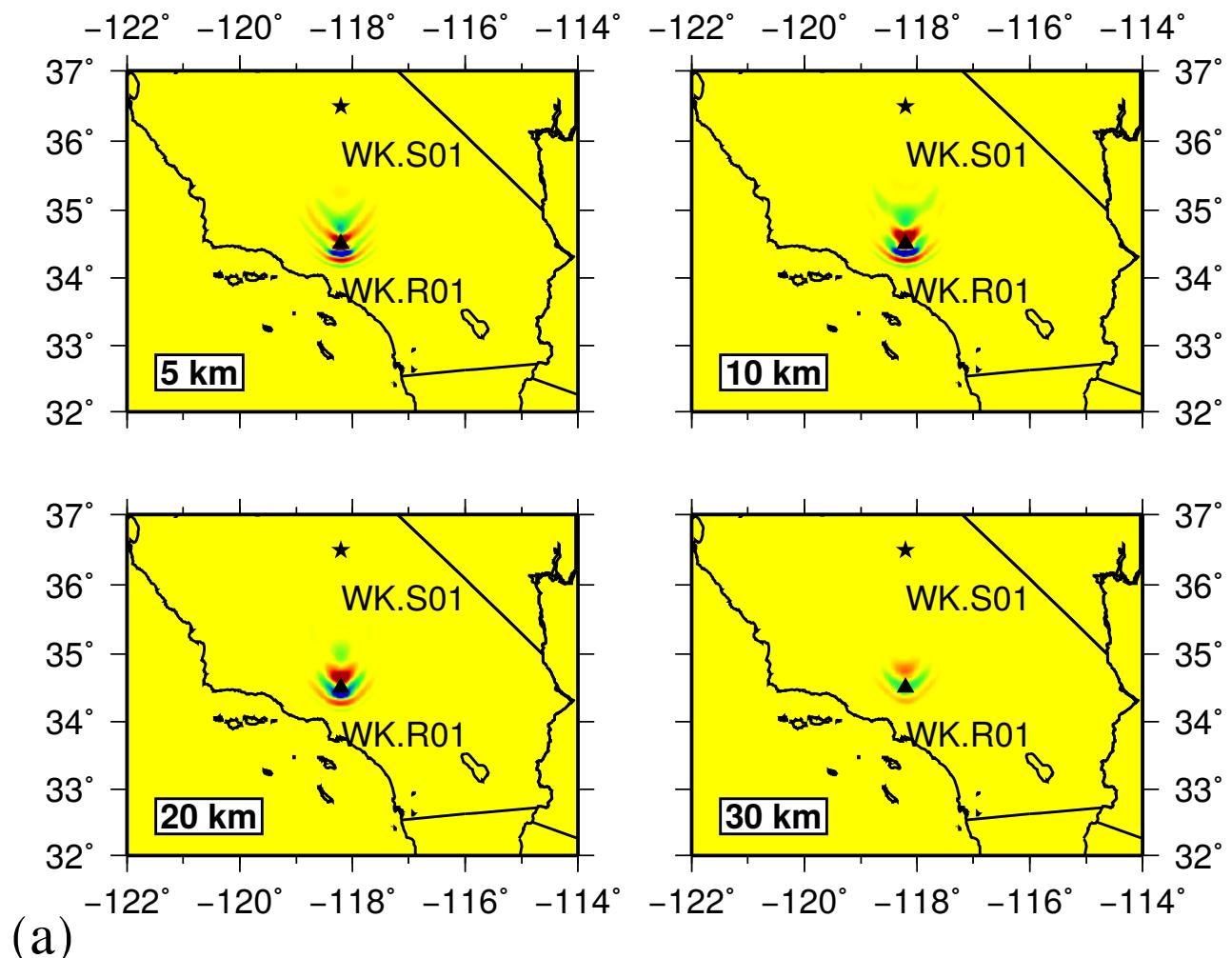


FIGURE 5.4: (a) Same as the third column of Figure 5.3 but at different depths; (b) vertical cross sections of the fundamental mode Rayleigh wave ellipticity kernel for P-wave velocity (α), S-wave velocity (β) and density (ρ) filtered at 10–20 s.

kernel; (3) the ellipticity sensitivities are mainly concentrated at areas close to the receiver but with larger influence zone towards the source.

Figures. 5.7-5.8 exhibit the ellipticity kernels in a 3-D heterogeneous medium. As expected, the vibrations in ellipticity kernel introduced by higher-mode interferences now are strengthened by strong scattering in heterogeneous medium. Both the magnitude of the individual amplitude kernel and the oscillations of the sidebands are stronger and the across-ray symmetry is destroyed by medium heterogeneities. In Figure S1, we also show that at larger distance, the Rayleigh wave ellipticity kernel might be distributed not only around the receiver but also within the whole elliptical zone.

5.4.4 Azimuthally averaged ellipticity kernel

The complicated pattern of an individual Rayleigh wave ellipticity kernel around the receiver makes it unstable to measure ellipticity from a single earthquake. In reality, averaging over different azimuths is often implemented to obtain reliable results [e.g., Li et al., 2016, Tanimoto and Rivera, 2008].

Following Maupin [2017], we obtain a azimuthally averaged ellipticity kernel by summing 36 individual kernels with 10° bins in azimuth (Figure 5.1c). Horizontal slices and vertical cross-sections of the kernel for the three models M01, M02 and M03 are displayed in Figures. 5.9-5.11. In a half space homogeneous model (M01), the averaged ellipticity kernel has a quasi-circular sensitivity kernel with a pattern of alternating positive-negative sensitivities centered at the station. The circular influence zone of the ellipticity kernel is broadened by higher-mode interferences in the 1-D layered model (M02). In 3-D heterogeneous medium (M03), the influence zone of the ellipticity kernel becomes asymmetric and significantly disturbed due to strong scattering.

To compare with 1-D depth kernels obtained by numerical methods or analytical methods, we calculate the 1-D integrated kernels for the 3-D model over a cylinder area around the receiver with different radius as shown in Figure 5.12. With increasing radius, the pattern of the integrated kernel becomes stable. For shear and compression waves, the integrated sensitivity kernels are very similar with that from 1-D depth kernels [e.g., Tanimoto and Alvizuri, 2006]. The 10 – 20 s Rayleigh wave ellipticity is most sensitive to shear-wave velocity with mainly negative sensitivities at depths of 5-20 km. The P-wave ellipticity sensitivity kernel has large sensitivities at the surface and decreases dramatically with increasing depths, and

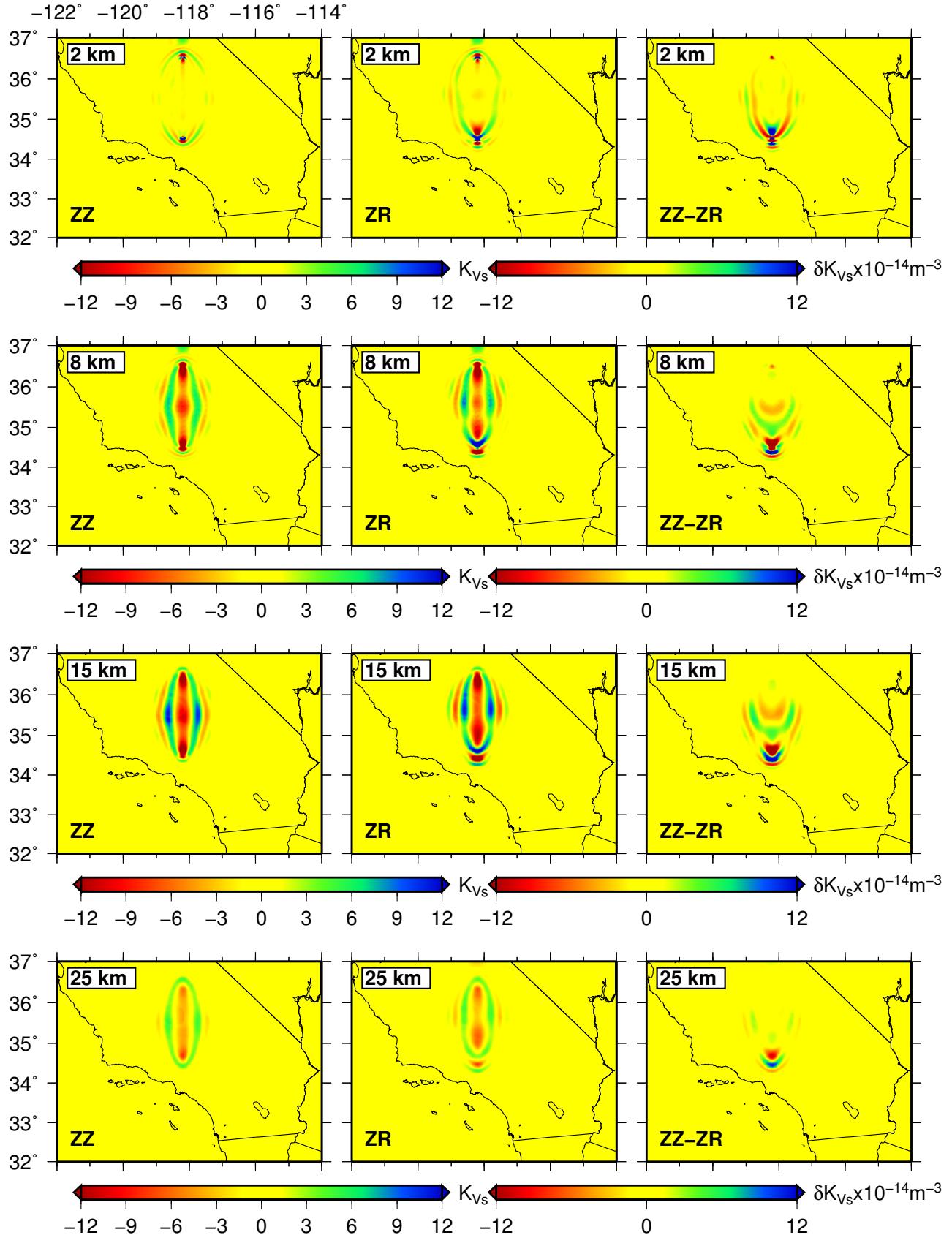


FIGURE 5.5: Sensitivity amplitude kernels and their differences in a 1-D layered model.

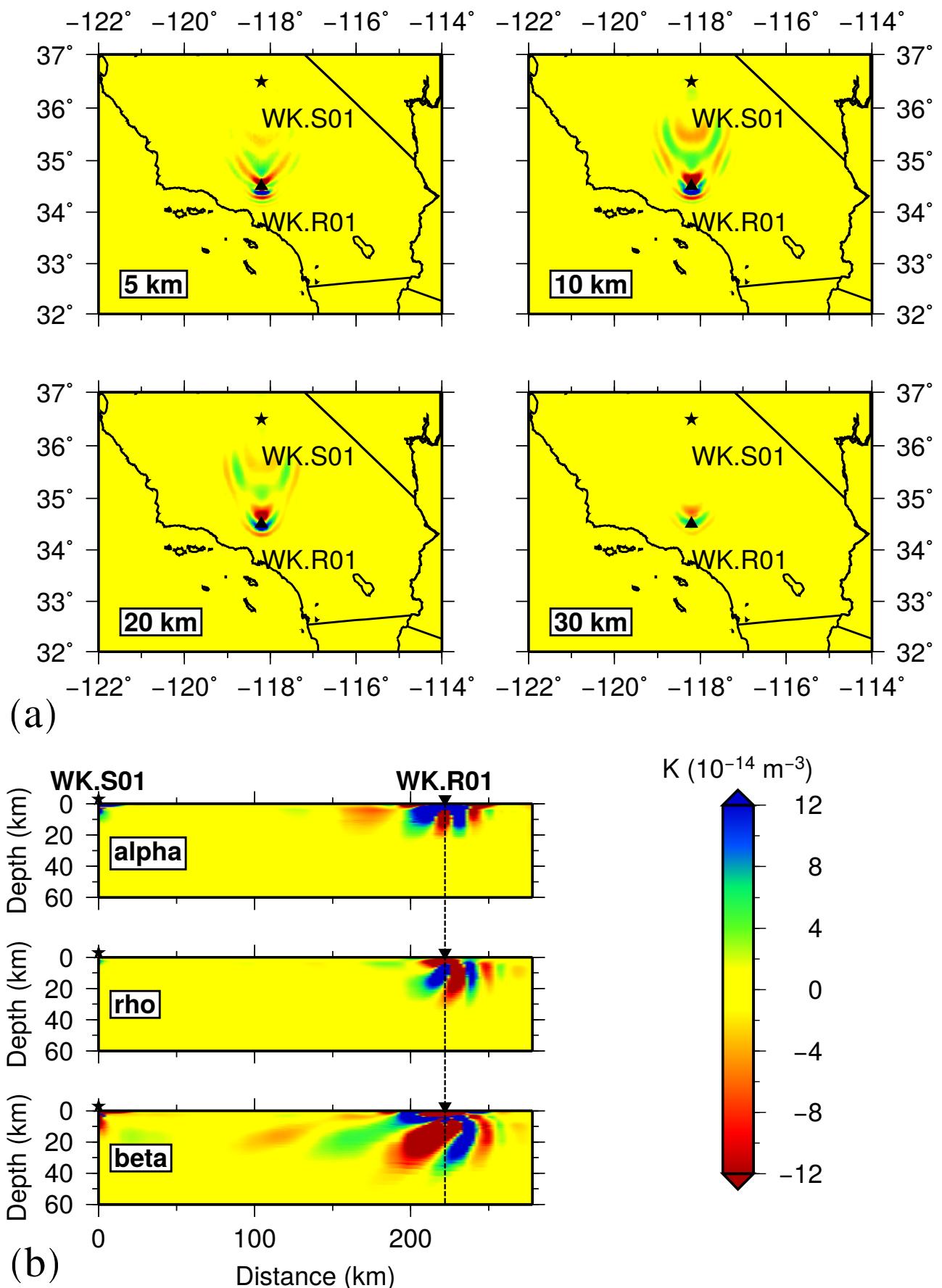


FIGURE 5.6: Same as Figure 5.4 but in a 1-D layered Socal1d model

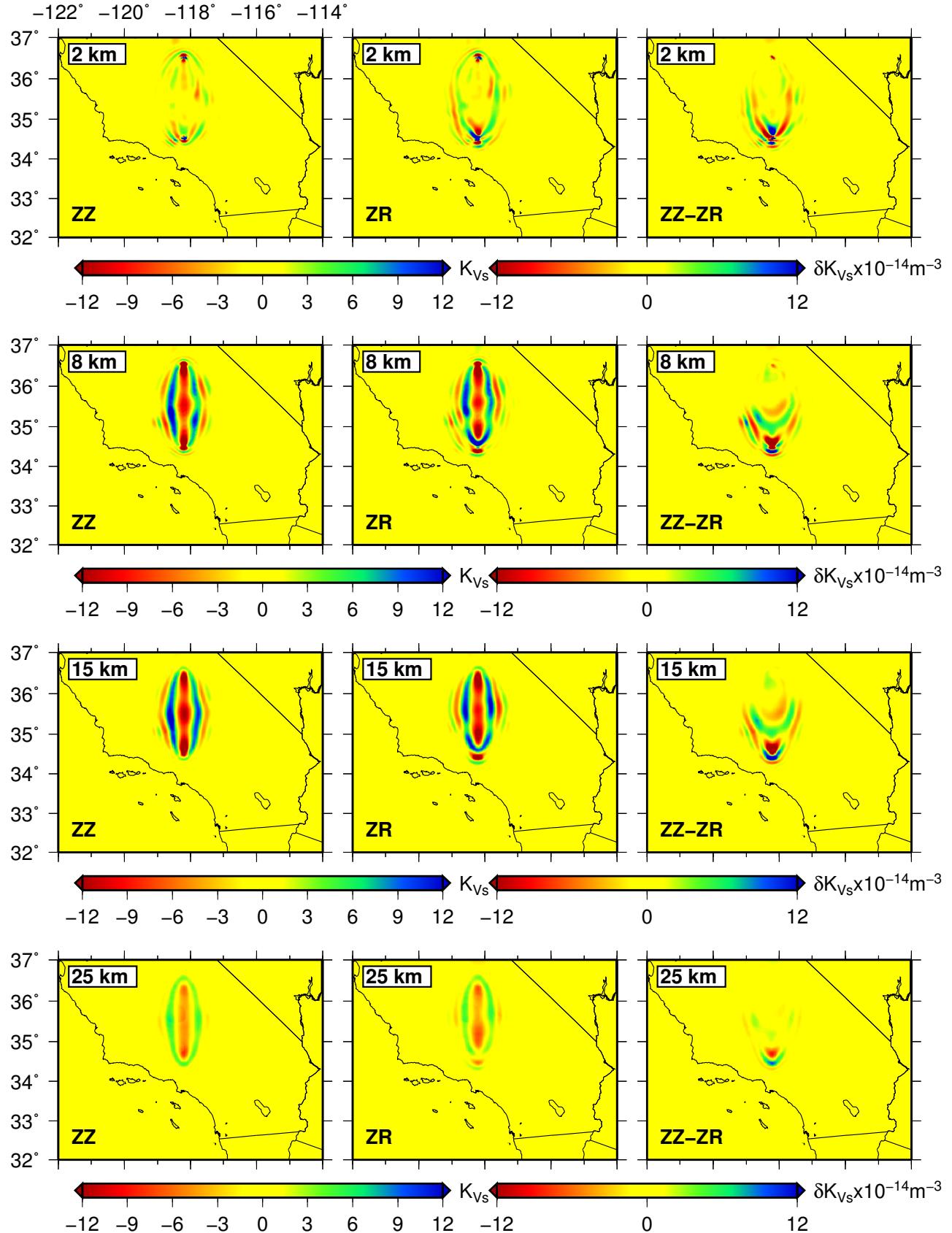


FIGURE 5.7: Sensitivity amplitude kernels and their differences in a 3-D heterogeneous model

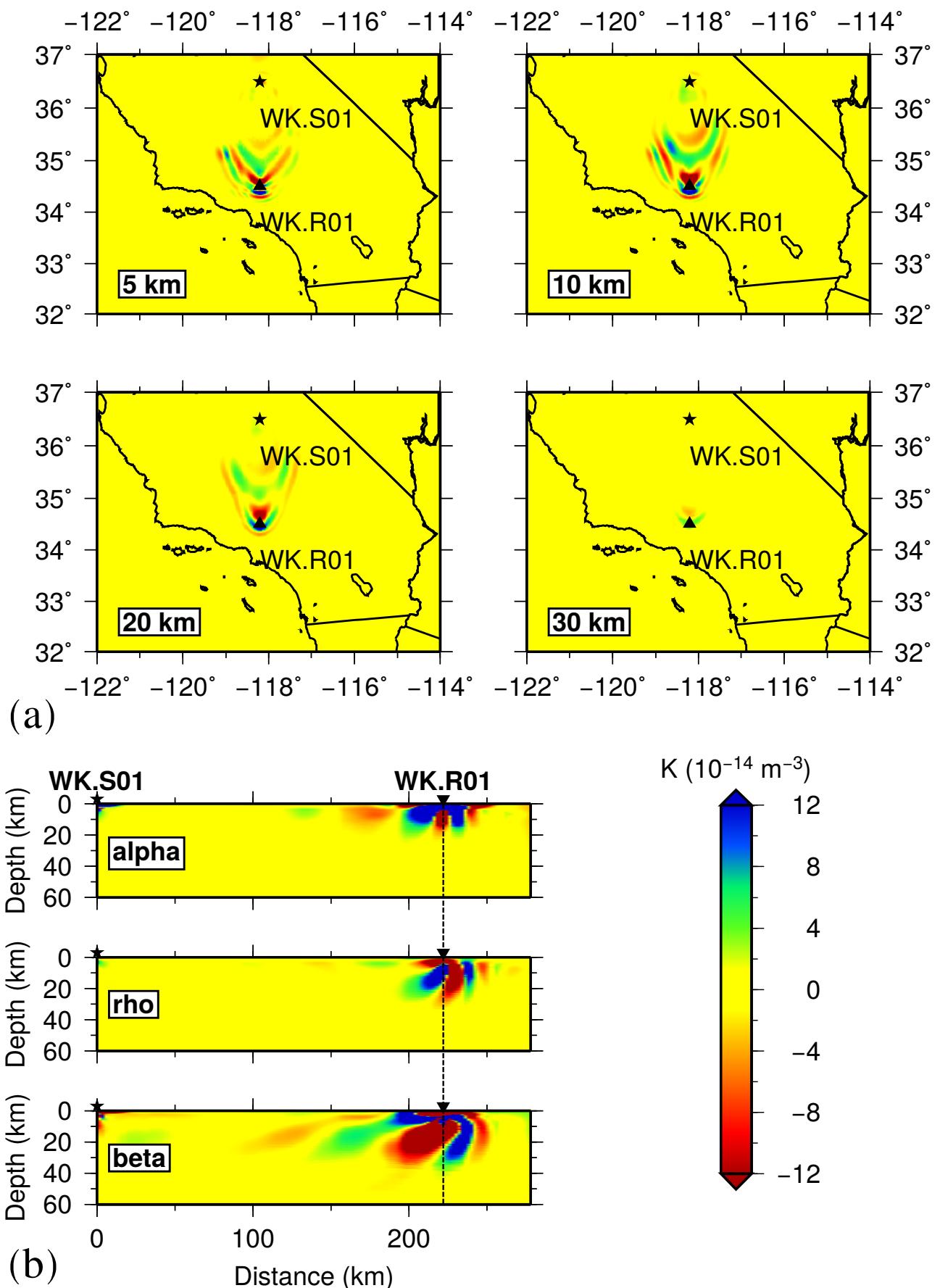


FIGURE 5.8: Same as Figure 5.4 but in a 3-D heterogeneous Socal3d model

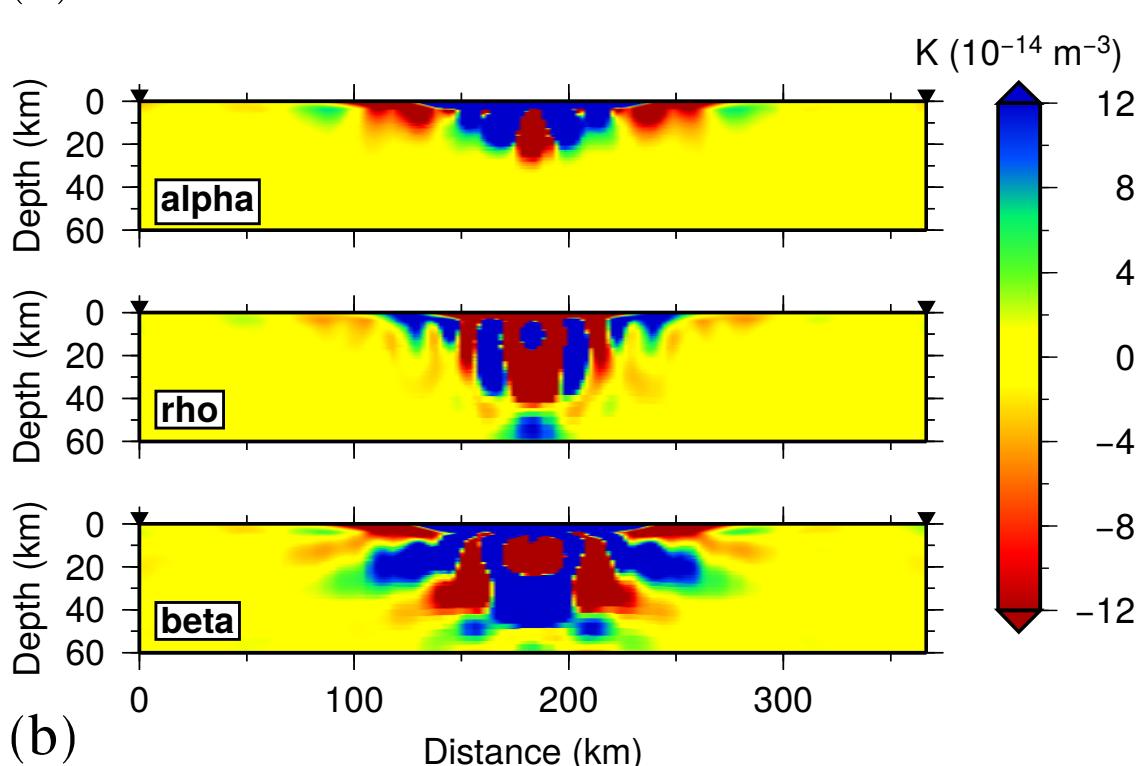
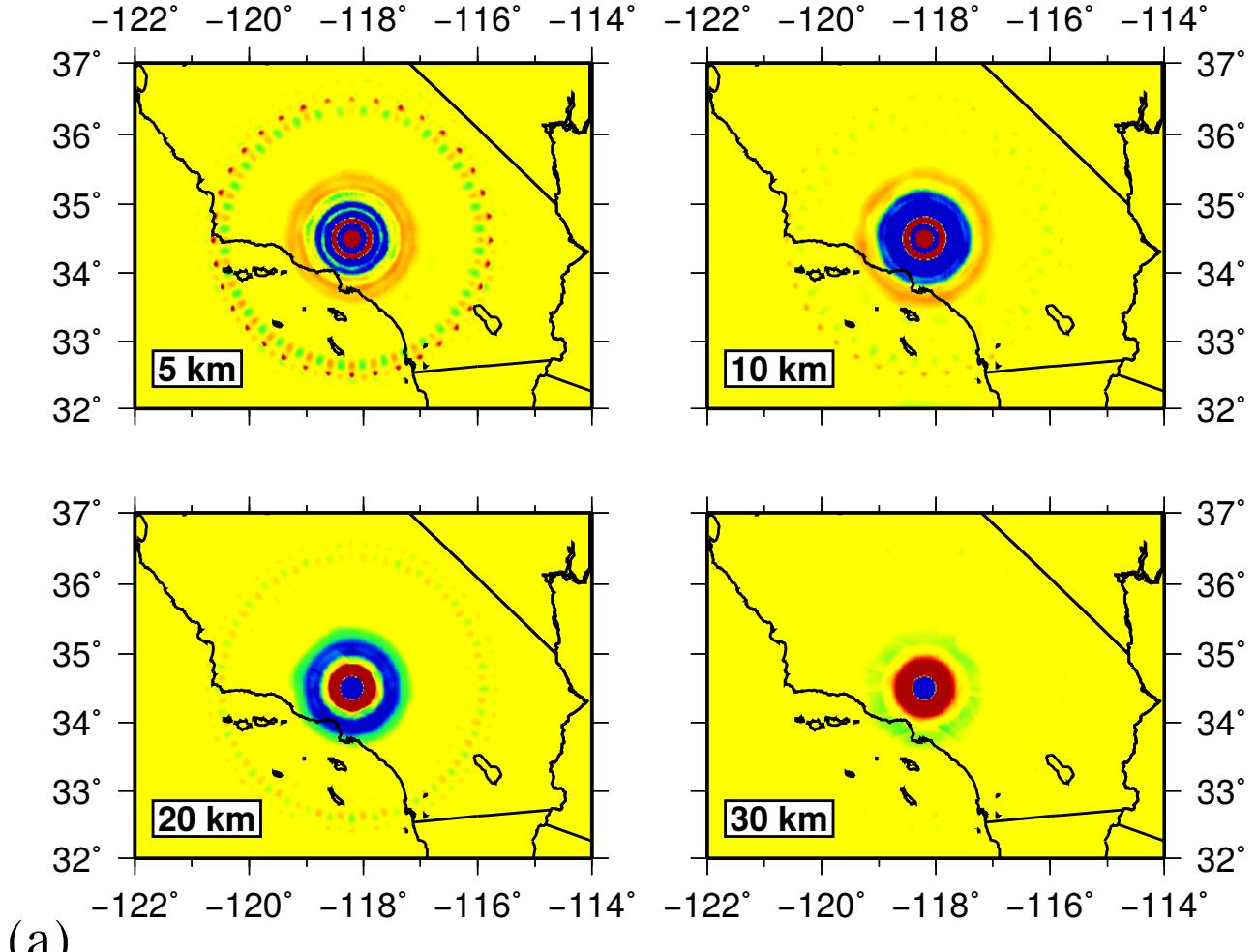


FIGURE 5.9: (a) horizontal slices of azimuthally averaged ellipticity kernel at 5, 10, 20 and 30 km depths; (b) vertical cross sections of azimuthally averaged ellipticity kernel with respect to compression-wave (alpha), density (rho), and shear-wave velocity (beta). The 3-D model is the homogeneous half space model M01.

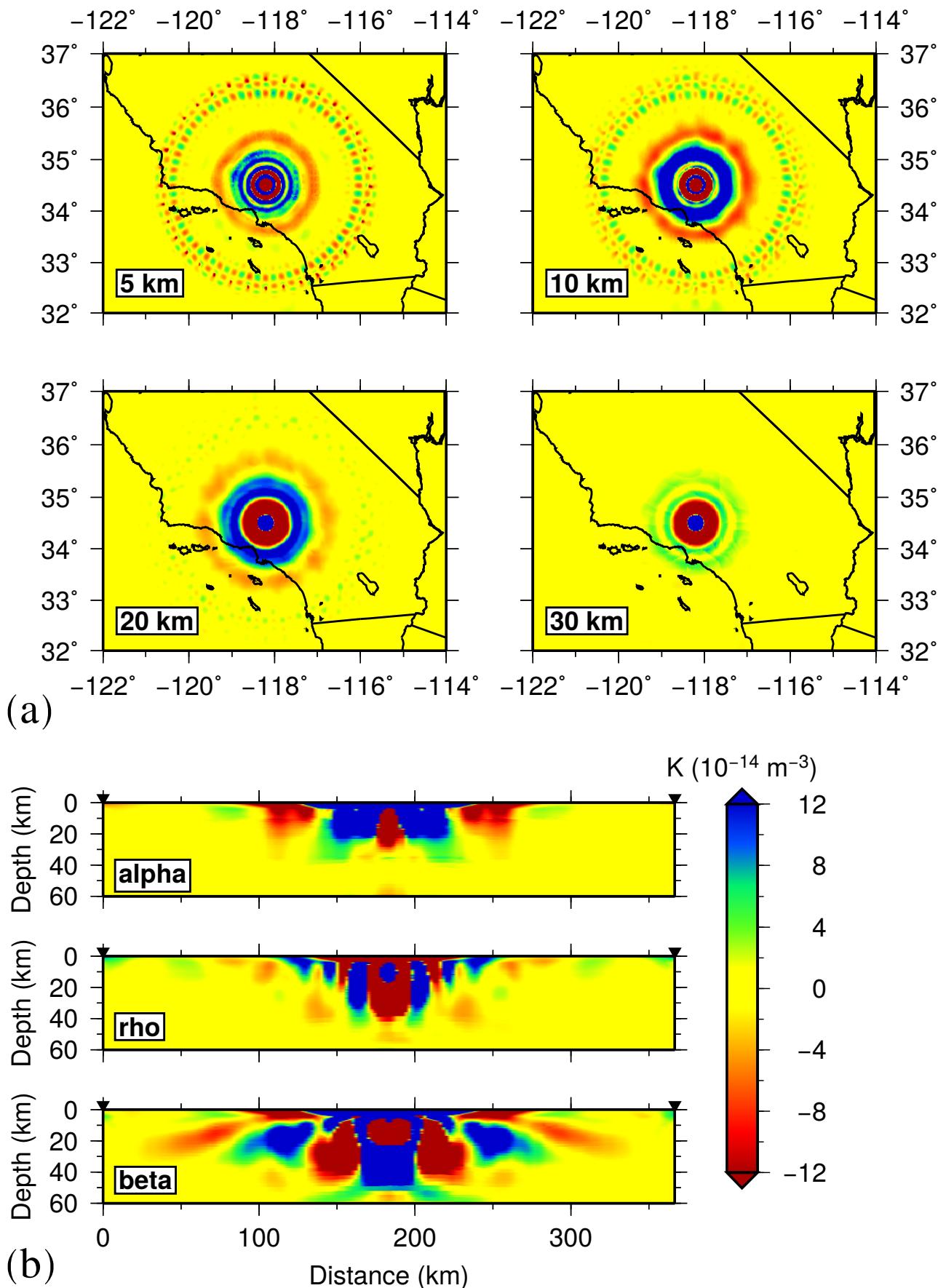


FIGURE 5.10: Same as Figure 5.9 but for M02.

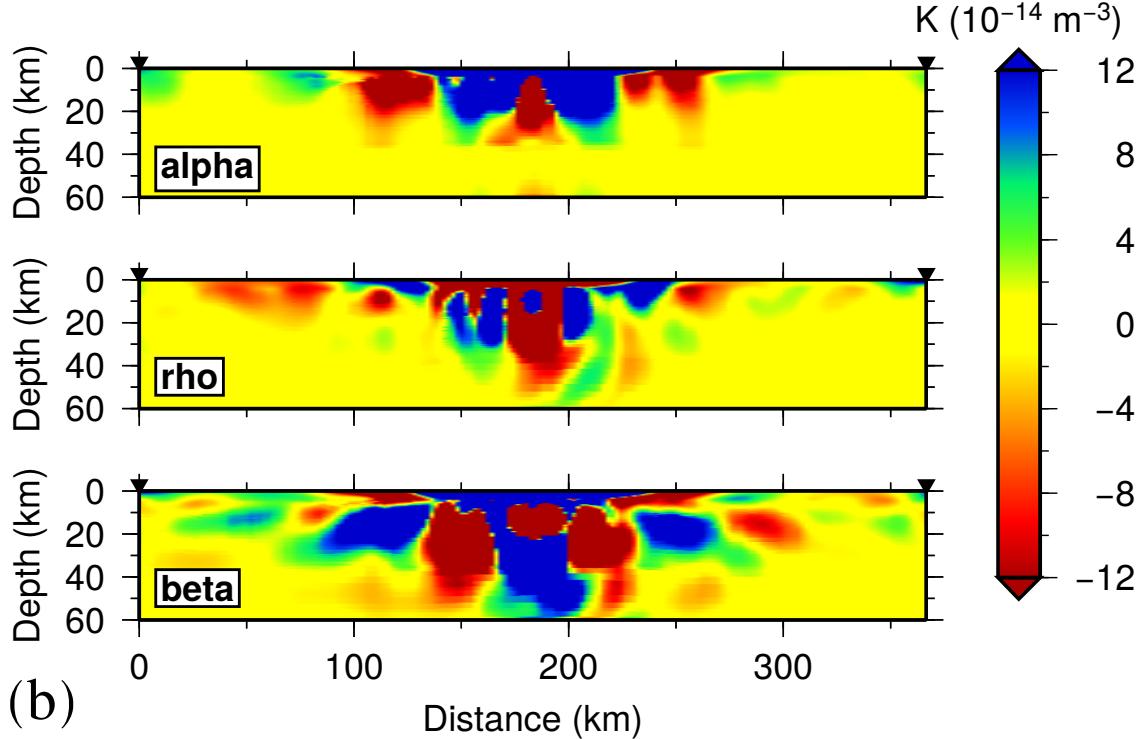
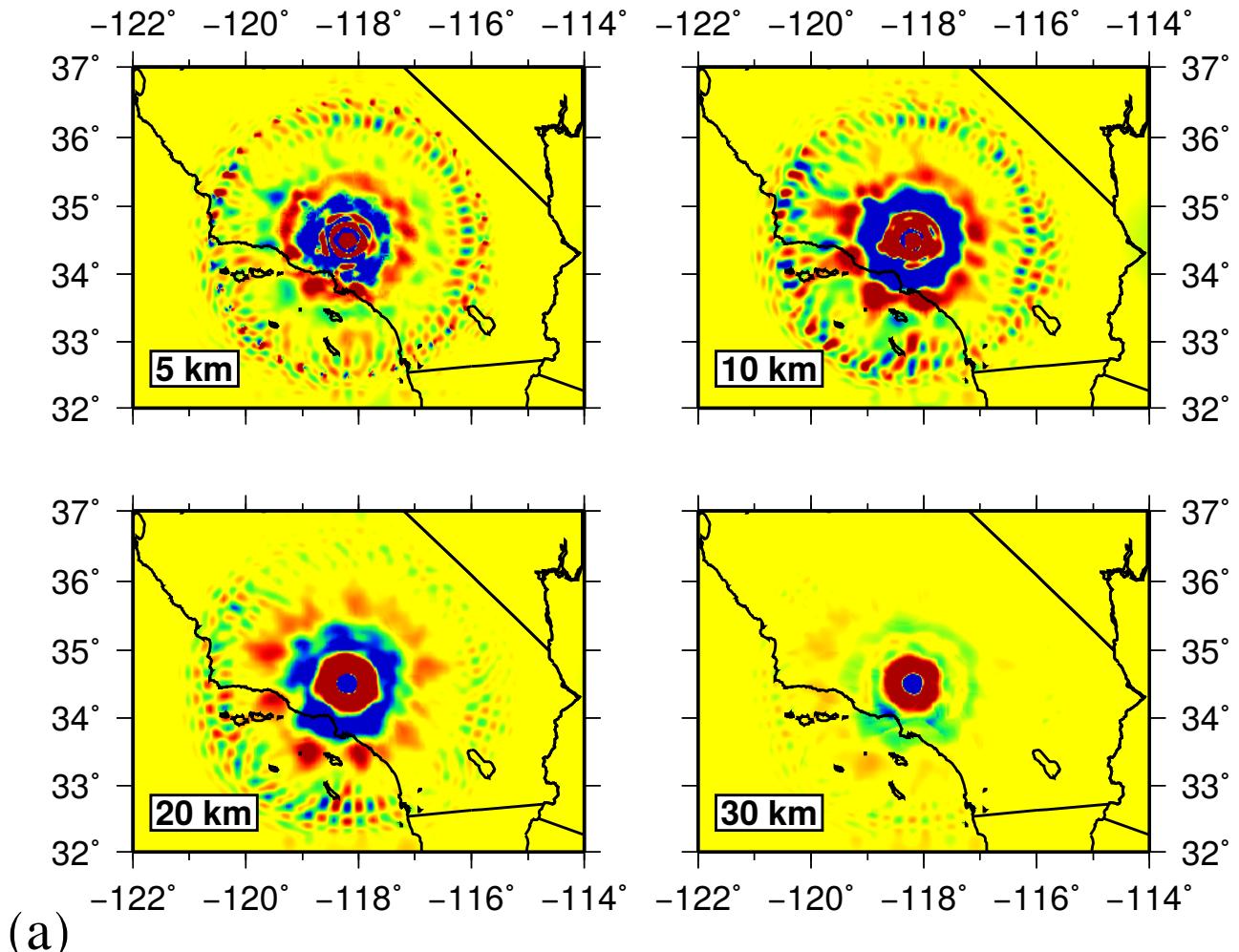


FIGURE 5.11: Same as Figure 5.9 but for M03.

disappears at about 10 km. The integrated ellipticity kernel for density is almost around zero, which is different from the negative sensitivities from 1-D depth kernels. This is because that the positive and negative sidebands of the 3-D density ellipticity kernel is almost symmetric, which cancel each other when integrated.

5.5 Conclusions and discussions

In this study, we investigate the 3-D Rayleigh wave ellipticity kernel using adjoint method based on numerical simulations. We demonstrate that the ellipticity kernel is actually the differential amplitude kernel between vertical and radial components. Our results show that Rayleigh wave ellipticity kernels in a half space and 1-D layered medium show dominant sensitivities around the receiver, and higher-mode interferences broaden the influence zone by extending it towards the source. Benefiting from numerical solver, ellipticity kernels in a heterogeneous medium are investigated based on adjoint method for the first time. We find that medium heterogeneities have an intrinsic scattering effect that makes the ellipticity kernel distributed within the whole elliptical zone around the source-receiver path, rather than only focused at the receiver in a lateral homogeneous medium.

The 3-D ellipticity kernel has a complex shape in a laterally homogeneous medium. Horizontal slices show alternating positive and negative sidebands with an elliptical shape truncated at the receiver. The vertical transects along the source-receiver path show a positive-negative striped sensitivity tail beneath the receiver. By summing individual ellipticity kernels with different azimuths, an azimuthally averaged ellipticity kernel is constructed, which shows a quasi circular sensitivity zone with a pattern of alternating positive-negative sensitivities centered at the station. The circular influence zone is broadened by higher-mode interferences in a 1-D layered model and becomes asymmetric and significantly perturbed due to strong scattering in a 3-D heterogeneous medium.

The variations of 3-D ellipticity kernel in vertical and horizontal directions suggest that it is inaccurate to simply replace the 3-D kernel by using a classic 1-D depth kernel with sensitivities just beneath the receiver. Yano et al. [2009] observed that the inversion of Rayleigh wave ellipticity using 1-D kernel is unstable and may lead to spurious low-velocity zones. Obviously, the differences between 1-D depth kernels and 3-D realistic kernels have some effects on the inversion for Earth's structures. Thus, we recommend to use 3-D ellipticity

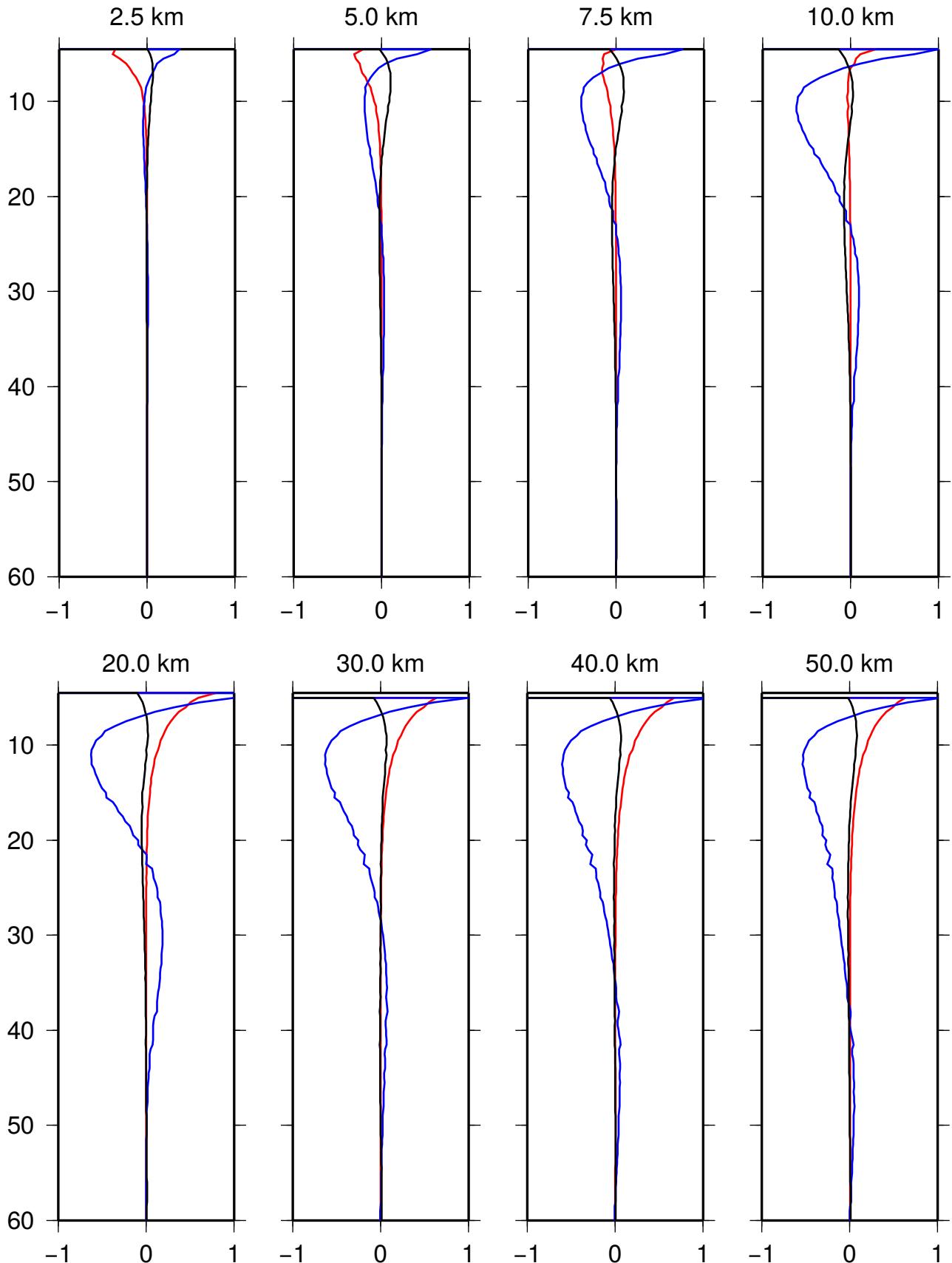


FIGURE 5.12: Integrated kernels for density (black), P-wave (red), and S-wave (blue). The radius of the cylinder area for integration is shown at the top of each sub-figure.

kernels in future tomographic studies whenever computation resources are allowed.

The question is whether there is a straight way to conduct ambient noise adjoint tomography based on Rayleigh wave ellipticity alone or in combination with other types of data? In traveltimes adjoint tomography, it is efficient to construct a so-called event kernel [Tape et al., 2007] by the interaction of a forward field emitted by a source and an combined adjoint field generated by inserting adjoint sources at receivers simultaneously. However, the situation is obviously opposite for a receiver-side Rayleigh wave ellipticity kernel. A ‘receiver kernel’ calls for a combined forward field emitted from various sources simultaneously and two adjoint fields with adjoint sources in radial and vertical directions. But, in this case, we can not identify the seismograms of each earthquake from the combined forward simulation, which is expected to be chaos because of contributions from different sources at the same time. In other words, it is impossible to measure the individual Rayleigh wave ellipticity as well as the azimuthally averaged ellipticity from a combined forward field. The problem of ellipticity adjoint tomography is that we can not separate misfit measurements and kernel construction like what is done in travel time adjoint tomography.

An alternative way is to calculate the individual 3-D Rayleigh wave ellipticity kernel for each event-station pair by using the scattering integral (SI) method [Chen et al., 2015, Zhao et al., 2005]. In the SI method, the whole strain and velocity wavefield of $3N_s$ (N_s is the number of stations) simulations with sources placed at individual receiver need to be saved spatially and temporally. Thus, each individual kernel can be constructed by the interaction of source- and receiver-side wavefield and strain field. The azimuthally averaged kernel is subsequently obtained by summing individual kernels for different source-receiver paths over one central station. However, the drawback of the SI method is that it may pose a daunting storage challenge when N_s is enormous or the model domain is extensive.

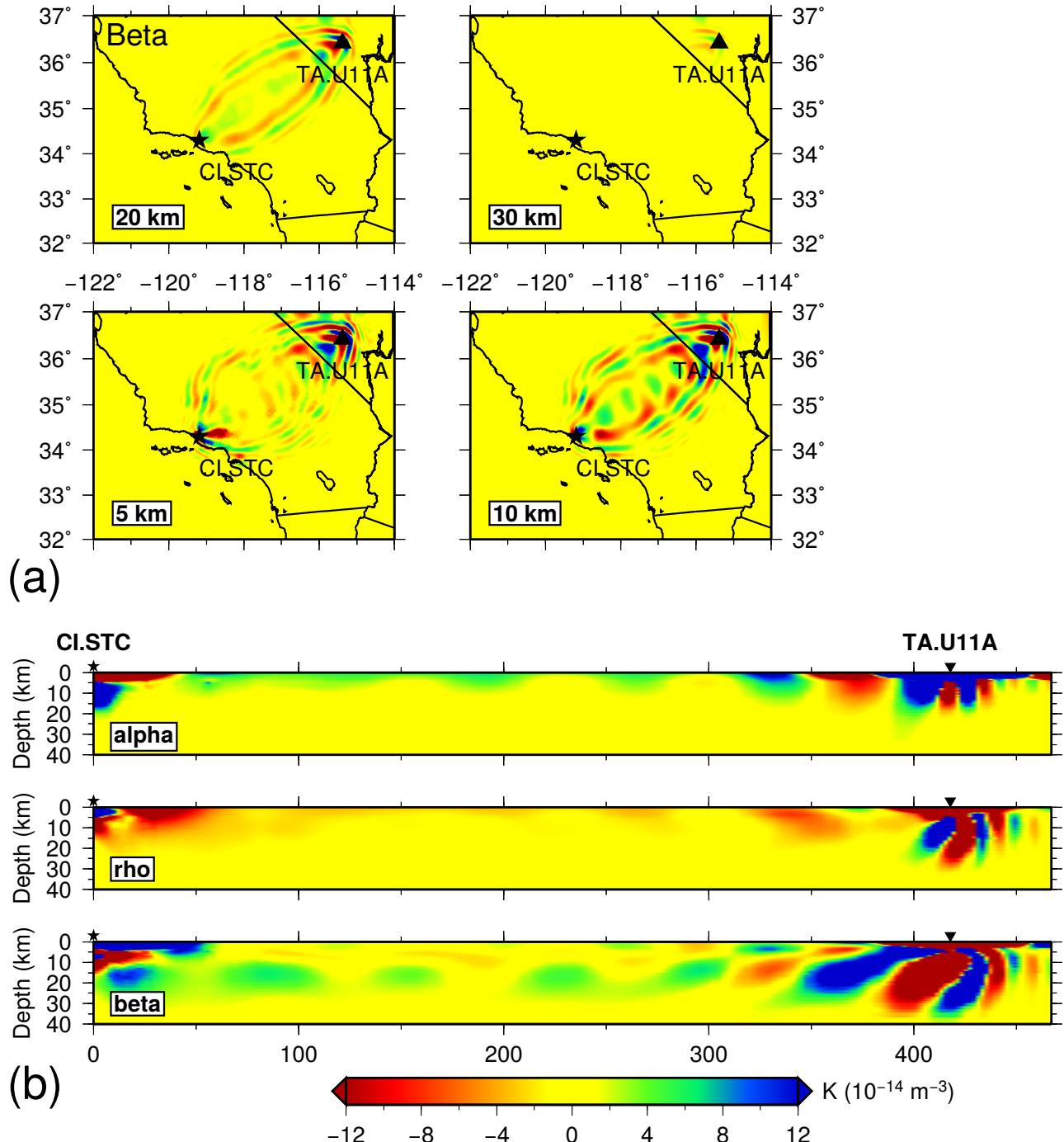


FIGURE 5.S1: Same as Figure 5.4 but at a larger distance.

6

Conclusions and discussions

The primary goal of this thesis is to develop an iterative inversion package for multi-component ambient noise adjoint tomography. This objective has been achieved by developing a general theoretical framework to calculate sensitivity kernels for multi-component ambient noise cross correlation functions and a associated semi-automatically script-driven ambient noise adjoint tomography package (Appendix A). Based on the theory and inversion package, I successfully apply adjoint tomography to Rayleigh wave and Love wave empirical Green's functions (EGFs) from three-component ambient noise at 5 – 50 s in southern California to improve an earthquake-based tomographic model. In addition, 3-D Rayleigh-wave ellipticity kernels are investigated based on spectral-element method (SEM) and adjoint method, which is fundamental for future simulation-based ellipticity tomography. The main conclusions from this thesis are summarized below, followed by a short discussion on some potential research directions in the future.

6.1 Conclusions

In chapter 3, we construct an improved Vs model of southern California, and the new Vs model reveals several new features in the middle and lower crust compared to the earthquake-based model, including: (1) the mean speed of lower crust is slowed down by about 6%; (2) higher Vs anomalies (up to +4%) are observed in the LAB and CTR throughout the crust; (3) higher Vs anomalies are seen in the lower crust beneath the westernmost PRB; (4) an enhanced shallow high velocity zone in the middle crust is observed beneath STB. Our model also shows refined lateral velocity gradients across PRB, SNB, SAF, which provides constraints on the west-east compositional boundary of PRB, SNB. As ambient noise cross-correlations can be obtained between any station pairs, the new tomographic model provides a better resolution relative to the initial one in areas not well-covered by event-station paths. Also owing to the longer-period features of ambient noise than the seismic data from local earthquakes, the lower-crust is better illuminated.

In chapter 4, we propose an efficient method to calculate multi-component (e.g., transverse-transverse (T-T), and radial-radial (R-R)) sensitivity kernels for EGFs from ambient noise based on 3-D numerical simulations of seismic wave propagation. Based on adjoint method, we demonstrate that a T-T component kernel can be obtained by the summation of two kernels, each of which is constructed from the interaction of a forward field due to a delta-function point force in the east or north direction, and an adjoint field by injecting T-T component adjoint sources at each individual receiver in the east and north direction with proper amplitude modulation factors based on rotation relations. Therefore, two forward simulations and two adjoint simulations are required to compute the sensitivity kernel for T-T component measurements of an event. The sensitivity kernel for R-R component measurements can be constructed similarly to those of T-T component kernels, but with R-R component adjoint source and different amplitude modulation factors. Our results show the kernels constructed based on rotation relations are almost identical to the reference ones with less than 1% error, sufficient for adjoint tomographic inversion. Utilizing this method for calculating T-T component sensitivity kernels, I successfully perform Love-wave ambient noise adjoint tomography (ANAT) in Southern California and obtain a new V_{sh} model. Compared with V_{sv} model from Rayleigh-wave ANAT, I observe positive ($V_{sh} > V_{sv}$) radial anisotropy in the whole uppermost mantle ($\sim +7\%$) and the north-eastern part of the crust ($\sim +4\%$). In the middle and lower crust (15 – 25 km), the whole Transverse Range shows

negative ($V_{sh} < V_{sv}$) radial anisotropy ($\sim -6\%$). The efficient computation method we developed provides us the basis for future multi-component ambient noise adjoint tomography for high-resolution isotropic and anisotropic structures.

In chapter 5, I show that Rayleigh wave ellipticity kernels in a half-space or a 1-D layered medium have sensitivities only around the receiver, and higher-mode interferences broaden the influence zone by extending it towards the source. However, in a 3-D heterogeneous medium, Rayleigh wave ellipticity has sensitivities over the whole elliptical region around the source-receiver path due to intrinsic scattering. In addition, Rayleigh wave ellipticity kernel has large sensitivities to S-wave velocity, relatively small but non-negligible sensitivities to density, and little sensitivities to P-wave velocity. The 3-D ellipticity kernel has a complex shape in a lateral homogeneous medium. The horizontal cross sections show alternating positive and negative sidebands with an elliptical shape truncated at the receiver. The vertical transects along the source-receiver path show a positive-negative striped sensitivity tail beneath the receiver. By summing individual ellipticity kernels from different azimuths, an azimuthally averaged ellipticity kernel is obtained, which has a quasi circular sensitivity zone with a pattern of alternating positive-negative sensitivities centered at the station. The circular influence zone is broadened by higher-mode interferences in a 1-D layered model and becomes asymmetric and significantly disturbed due to strong scattering in a 3-D heterogeneous medium.

6.2 Future works

Based on our current applications of ANAT in southern California, we summarize here some interesting research directions that can be explored further in the future.

1. Multi-component ANAT in southern California. In this thesis, we make the simple assumption that Rayleigh waves are only sensitive to V_{sv} structure while Love waves are only sensitive to V_{sh} structure, and we perform adjoint tomography of the two types of waves separately based on 3-D Fréchet kernel constructed from an isotropic model. A more accurate and elegant way is to use a radially anisotropic initial model for the adjoint tomography, and fit the waveforms of Rayleigh and Love waves simultaneously during the inversion. Adjoint tomography has been used to study continental scale anisotropic structure using earthquake data based on a global version of SPECFEM3D

with anisotropic sensitivity kernels [Chen and Tromp, 2007, Sieminski et al., 2007a, Zhu et al., 2015]. However, such tomographic scheme has not been applied to ambient noise data to image regional-scale structure. A natural extension of current studies is to conduct a multi-component ambient noise adjoint tomography in southern California using a transversely isotropic initial model based on the V_{sv} and V_{sh} model from this thesis.

2. Inversion strategy in ambient noise adjoint tomography. In this thesis, we use an earthquake-based tomographic model as our starting model, which already fits well with the observed earthquake data. In many areas where no adequate initial model is available, we have to use traditional ANT based models or simple 1-D layered models as starting models. In that case, it is still not clear how much the adjoint tomography can improve the results of traditional ANT and whether it is worthy of the expensive computation cost? Although Chen et al. [2014] have shown that both 1-D model or ANT based adjoint tomography can improve the results of traditional ANT. However, there is no other high-resolution models can quantify the improvements. A potential approach to solve this problem is to use a 3-D checker board model to simulate the surface waves. Then, we apply both adjoint tomography and traditional ANT to the synthetics and evaluate the performance by comparing their tomographic results.
3. Adjoint tomography of both earthquake and ambient noise noise. Throughout this thesis, we focus on applying adjoint tomography to ambient noise alone, although the initial model used is an earthquake-based tomographic model. We seek to use the long-period surface waves from ambient noise to refine seismic structure of the middle-lower crust, and in the meantime to keep the main features of shallow crust from short-period earthquake data. Although it will take a large amount of computation by fitting earthquake and ambient noise together, it should definitely be presumed in future studies. If all data and techniques are available, we would recommend a three-stage inversion: (1) isotropic adjoint tomography of three-component earthquake and ambient noise data; (2) attenuation adjoint tomography of three-component earthquake data; (3) anisotropic adjoint tomography of three-component earthquake and ambient noise data.
4. Adjoint tomography of Rayleigh wave ellipticity. The geometrical variations of 3-D

ellipticity kernel in vertical and horizontal directions suggest that it is inaccurate to simply replace the 3-D kernel by a classic 1-D depth kernel with sensitivities just beneath the receiver. 3-D sensitivity kernels based on numerical simulations are expected to be applied in future ellipticity tomographic studies.

5. Full waveform inversion of ambient noise based on Scattering Integral (SI) and SEM. Adjoint tomography strikes a balance between storage demand and computational cost by only calculating and saving event misfit kernels, instead of strain Green's tensors at spatial-temporal grid points as in the SI method. Adjoint tomography is more computationally efficient than SI only when the number of stations is larger than that of events. However, the number of stations and virtual sources are the same in ambient noise tomography. In this case, the SI method could be more efficient and suitable for ambient noise, especially when the number of station is small such that no heavy storage is required. Currently, SI-based full waveform inversion of ambient noise is often based on finite-difference numerical modeling [e.g., Gao and Shen, 2014, Lee et al., 2014]. Schumacher [2014] developed a modularized iterative full waveform inversion package named ASKI [Schumacher and Friederich, 2016], which uses the SI method in conjunction with SEM. However, applications of the ASKI are only restricted to waveform inversion due to the lack of travetime sensitivity kernels. It is feasible to advance the ASKI package with travetime SI-SEM based full waveform inversion in the future.
6. Ambient noise adjoint tomography with directional sources. In this thesis, we simply use Green's functions from point-force sources as the synthetics for EGFs, ignoring the influence of heterogeneous noise sources [e.g., Wang et al., 2016, Yang and Ritzwoller, 2008]. However, such noise source effects can be accounted for by extending our method with ensemble forward and adjoint wavefield [Tromp et al., 2010] or transforming the actual wavefield sources and wave propagation physics into effective sources and wavefield [Fichtner et al., 2016].

A

ANAT- a script driven package for ambient noise adjoint tomography

This appendix serves as a simple guide for conducting ambient noise adjoint tomography using our ANAT package. We first introduce the workflow of the three-stage iterative inversion in section A.1, and the structure of the package in section A.2. Then, a step-by-step tutorial of running iterative inversion is given in section A.3. Finally, we give the details for some scripts and programs in section A.4 as references.

A.1 Workflow

The basic idea of ANAT is to iteratively minimize the traveltime misfit between empirical Green's functions (EGFs) from ambient noise and synthetic Green's functions (SGFs) from spectral-element simulations based on misfit gradient. According to the job submission schedule, we divide the inversion into three stages as shown in Fig. A.1. In stage one, we

conduct the forward and adjoint simulation in one PBS job (including misfit measurement) to obtain the event kernel of each virtual source (or master station). After all the simulations are finished, event kernels are summed, preconditioned, and smoothed in stage two to obtain the final misfit gradient. In stage three, we use line search to determine the optimal step length for model updating.

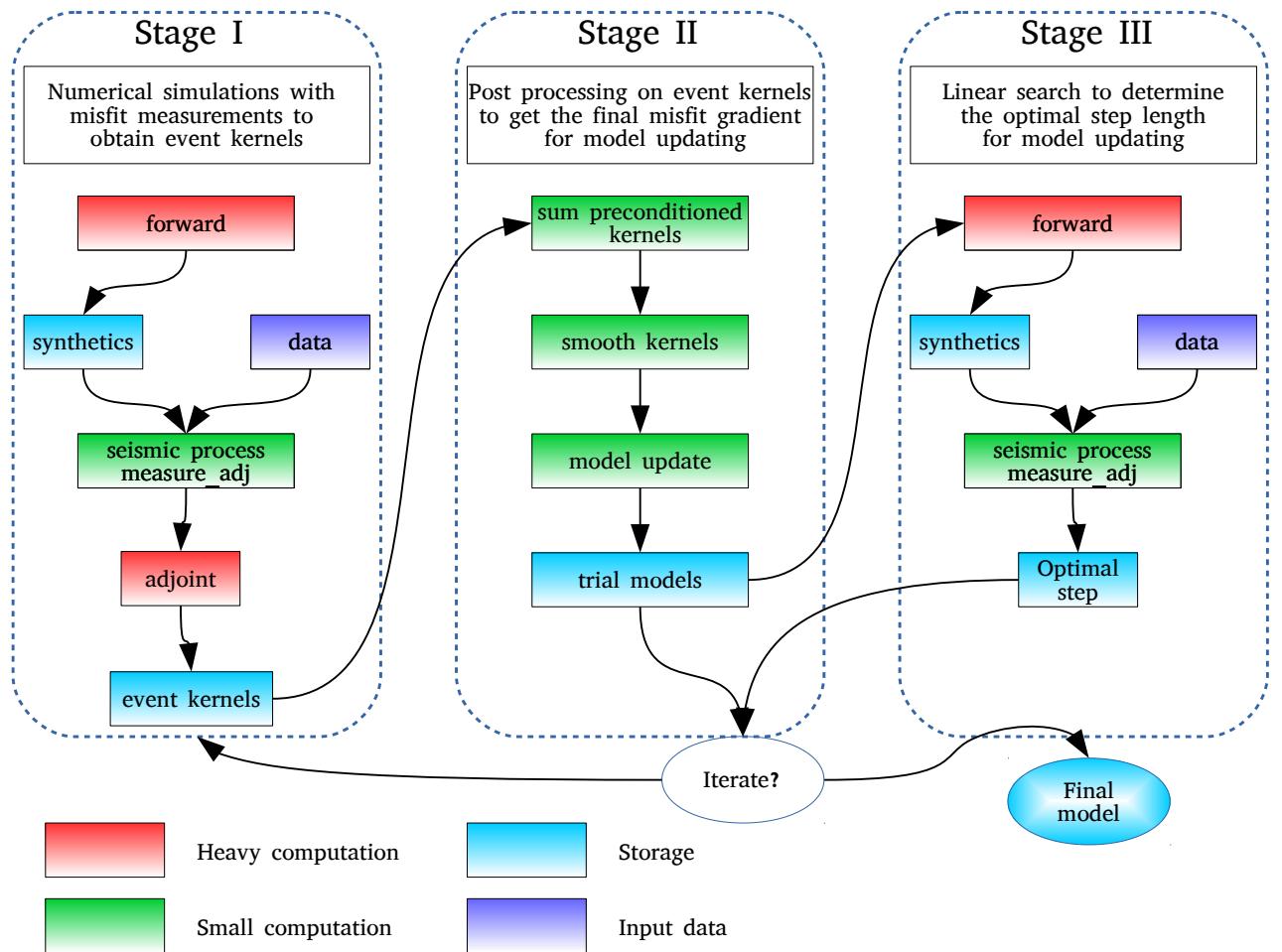


FIGURE A.1: Workflow of ANAT package.

A.2 Package structure

ANAT package consists of some workflow-control bash scripts and programs in several folder.

A.2.1 Inputs

- **data** is a database of ambient noise symmetric cross correlation functions sorted by virtual sources. All the data are in SAC format and named as data/net1.evt1/net1.sta1.comp.sac, where net1 and evt1 are the network and name for virtual source 1, sta1 is for station 1, comp is the component. For example, data/CI.STC/CI.BAK.HXZ.sac is the EGF between the master station CI.STC and general station CI.BAK.
- **src_rec** stores all source and station information needed for the inversion. There are three files prepared for the SPECFEM3D Cartesian, namely CMTSOLUTION_evt1, FORCESOLUTION_evt1, and STATIONS_evt1. The format of these files should be the same as used in SPECFEM3D Cartesian. The other two files sources_set1.dat and sources_ls.dat are name lists of event (virtual sources), which are used for making directories for simulations by event name. We can divide the events into different sets, such as set1, set2, ..., etc.
- **specfem3d** contains three inputs for the software SPECFEM3D Cartesian, **DATA**, **OUTPUT_FILES**, and **bin**. These three inputs are essential for launching numerical simulations, and will be linked to all forward simulations directories.

A.2.2 Data processing and plot

- **ADJOINT_TOMOGRAPHY_TOOLS** is a bunch of seismic tools, among which we use **flexwin** for picking window and **measure_adj** for misfit measuring.
- **seismic_process** contains some C shell and Perl scripts for data preprocessing, such as converting ASCII seismogram to SAC files, bandpass filtering, etc.
- **plots** contains some example scripts for plotting misfit, model, kernel, waveform. The corresponding directories are **misfit_plot**, **model_plot**, **kernel_plot**, **seismo_plot**.

A.2.3 Outputs

- **output** stores all the misfits files for a specific model. For example, all misfits for model M01 are stored in ./output/misfits/M01.

- **optimize** is where postprocessing procedure applied. All event kernels are processed in the directories **sum_kernels_M01**, and model updating files are stored in **SD_M01**. All inputs and outputs in these two directories are listed in Fig. A.2.
- **solver** is where numerical simulations are performed and stored. For event1 in the source list set1, the forward simulation root directory is solver/M01.set1/event1. For line search at step length of 0.01, the corresponding directory is solver/M01.slen0.01/event1. Inside each forward simulation root directory, the three sub-directories of **specfem3d** are linked as the inputs for SPECFEM3D Cartesian. The script *change_simulation_type.pl* and directory **SEM** is generated for adjoint simulation.

A.2.4 Driven scripts

- *run_iteration.bash* is the main scripts for running forward and adjoint simulations. This scripts will prepare all inputs in the forward simulation directories located at **solver** and call the other two scripts *run_preprocessing.3band.bash* and *pbs_mesh_fwd_measure_adj.bash*.
- *run_preprocessing.3band.bash* is to make all inputs and directories for preprocessing in **seis_process_M01.set1**, and for misfit measurement in **measure_adj_M00.set1**.
- *run_postprocessing.bash* is to prepare all inputs and directories for postprocessing needed by the script *pbs_postprocessing.bash*.
- *run_line_search.bash* is the same as *run_iteration.bash* but for a linear search.

A.2.5 PBS scripts

- *pbs_mesh_fwd_measure_adj.bash* is the PBS script to do meshing, forward simulations, misfit measuring, and adjoint simulations.
- *pbs_postprocessing.bash* is the PBS script to do postprocessing procedures include, kernel summation, precondition, and smoothing.

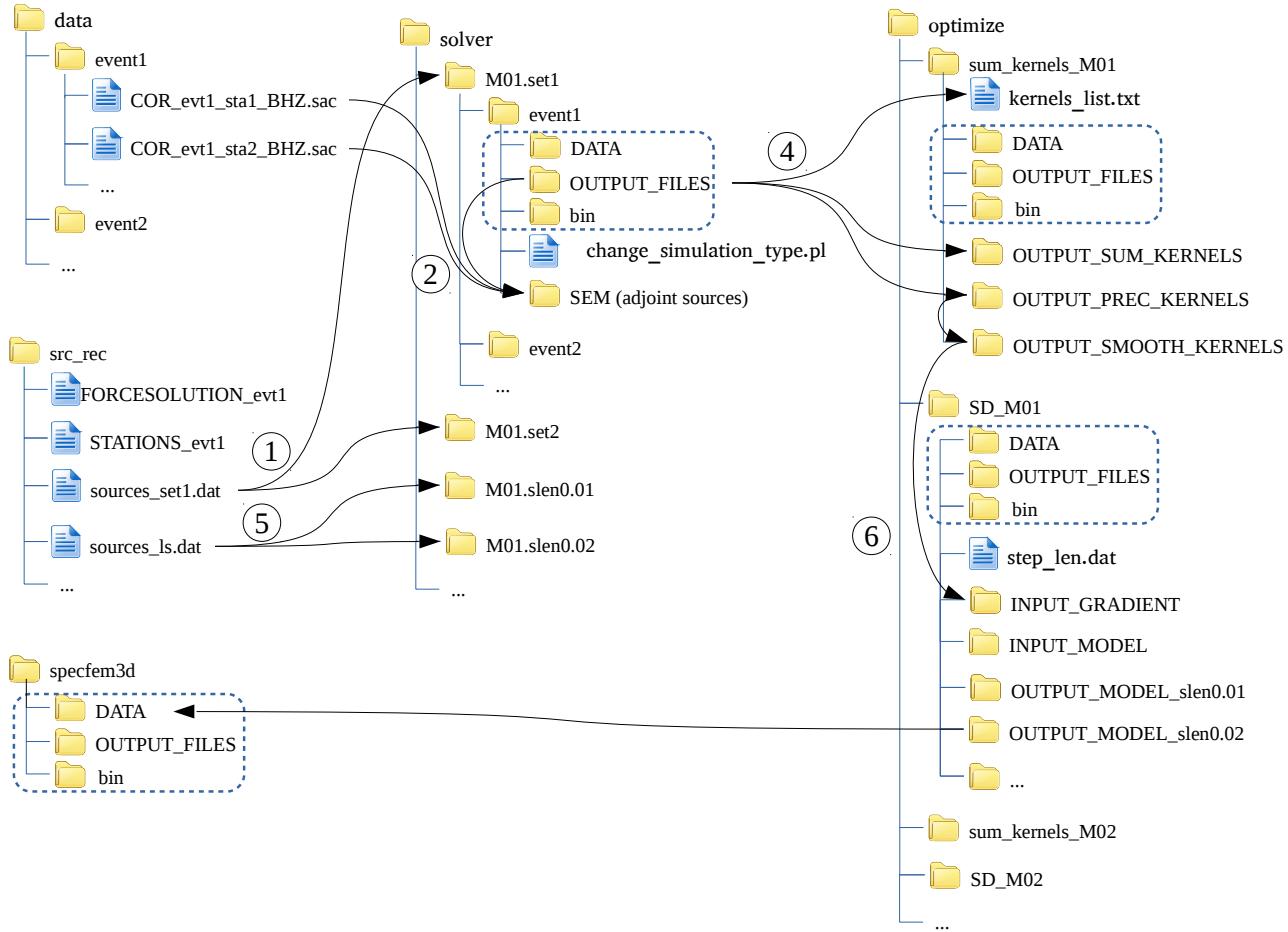


FIGURE A.2: Structure of ANAT package.

A.3 Iterative inversion procedures

There are four basic steps to do the iterative inversion using our ANAT package.

A.3.1 Preparing input files

Before running any iteration, the users should write their own scripts to make **data** and **src_rec** in the format described above. The corresponding example scripts are *generate_data.bash* and *mk_forcesolution.bash*.

The inputs files in **specfem3d** should be tested to make sure that they are ready for running any simulations. This means that the user should set up the meshing files at *DATA/fem3D_files*, initial xyz model at *specfem3d/DATA/tomo_files* or gll model in a user defined folder, and *DATA/Par_file*. Set APPROXIMATE_HESS_KL = .true. and USE_RHO_SCALING = .true. in *Par_file*. For testing if the mesh and initial model is correctly set up, set

`SAVE_MESH_FILE = .true.` in *Par_file* and plot the velocity model *.bin files in *OUTPUT_FILES/DATABASES_MPI*, and set it back to `.false.` after the test. The user should also check *output_mesher.txt* and *output_meshfem3D.txt* for the minimum period resolved, suggested time step, and other useful information regarding the meshing and solver parameters.

A.3.2 Running forward and adjoint simulations

The first step of the iterative inversion is to run the script *run_iteration.bash*. We should first set the model name (such as `mod=M01`), and step length (such as `step=0.02`) for reading updated model from previous iterations. Basically, this script can be divided into three parts, including:

1. Making forward simulation directory. The script will make all forward simulation directories by looping over event sets defined by a variable `ipart` according to the file `sources_set$ipart.dat`. Then, it will link the all input files in **specfem3d**, corresponding STATIONS and FORCESOLUTION files in **src_rec** to each forward directory. The script also have an option to change the model for the first iteration and other iterations.
2. Running *run_preprocessing.sh*. We adopt a multi-scale strategy in misfit measurement, thus there are different version of these scripts, such as *run_preprocessing.1band.sh*, *run_preprocessing.2band.sh*, *run_preprocessing.3band.sh*. The author should choose how many and what frequency bands needed in current iteration depending on the data, then make changes to these pre-processing scripts. There will be two directories created, namely **seis_process_M01.set1** and **measure_adj_M01.set1**. The former is to do some pre-processing procedures on data and synthetics, and the latter is to measure the misfits between each data and synthetic pair and calculate the corresponding adjoint sources.
3. Submitting *pbs_mesh_fwd_measure_adj.sh*. As long as all forward and data processing directories are created correctly, this script is ready to be submit after setting the proper walltime. After all the forward and adjoint simulations are successfully finished, the event kernels and Hessians are generated at `solver/M01.set1/event1/OUTPUT_FILES/DATABASES_MPI/*kernel.bin`.

A.3.3 Post-processing and model update

Then, we use the following two steps to do post-processing on event kernels to obtain the total misfit gradient and update the model.

1. Run *run_postprocessing.sh*. We should first set the current model name (such as mod=M01), and step length (such as step=0.02) for reading the current model from previous iterations. This script is also divided into two parts, which is controlled by setting is_sumkern=.true. and is_update=.true. The resulting two directories are **sum_kernels_M01** and **SD_M01**, where "SD" refers to Steep Descent method. In the model update part, the user should set proper step lengths for line search in the text stage, and also the input model for updating. If it is the first iteration, the input model is the initial model, otherwise, it is the updated model from previous iteration.
2. Submit *pbs_postprocessing.sh*. This script has three parts: is_sumkern, is_smooth, and is_update. The three part can submit in one time by setting them to true, or submit one by one sequentially. The most time consuming part is the kernel smoothness, which might take a few hours. The summed, preconditioned, and smoothed gradient are stored OUTPUT_SUM KERNELS, OUTPUT_PREC_KERNELS, and OUTPUT_SMOOTH_KERNELS. And the updated models are stored at OUTPUT_MODEL_slen0.01, OUTPUT_MODEL_slen0.02, ..., etc.

A.3.4 Line search for optimal step length

After the updated models are generated, we use them as trial models to do a line search in order to obtain the optimal step length. The script responsible for this is *run_line_search.bash*, which is changed from the script *run_iteration.bash*. The script will make all forward simulation directories by looping over step lengths, and each of them has a selected number of events from the file sources_ls.dat. The data processing script is still *run_preprocessing.sh*, but the PBS script is *pbs_mesh_fwd_measure.sh* without adjoint simulation. The option is_ls is turned to true, which will delete the forward output directories because we only need the misfit measurements.

As long as all the forward simulations for different trial models are finished, the corresponding traveltimes misfits are collected in the directory **outputs/M01**. Then, we use the script *plt_line_search.3band.bash* to plot the total misfit curve over step length.

A.4 Data processing scripts

A.4.1 *run_preprocessing.sh*

This long script can be divided into two parts: pre_processing and measure_adj. In each part, the script will write all commands into an extra job*i*.sh for paralleling.

```
#!/bin/bash

mod=M01.set1
srfile=sources_set1.dat

R0=2.5 # minimum group velocity ### need to change this
R1=4.5 # maximum group velocity ### need to change this
#=====
is_prepoc=true
is_meas_adj=true

is_bd=false # if banana-doughnought or multi-taper kernel
i_plot=false # if plot misfit measurements
#-----
preprocdir=seis_process_{$mod}
measdir=measure_adj_{$mod}
kerndir=sum_kernels
#=====
if $is_prepoc;then
i=1
cat src_rec/$srfile |
while read evtfile; do # loop over all virtual evts
  eid=`echo $evtfile |awk '{printf "%s.%s",$2,$1}'`#
  fwd_dir=solver/$mod/$eid
  #####
  echo "seis_process commands" >$preprocdir/job${i}.sh
  #####
  let i=i+1
done # end loop over evt
fi
#=====
if $is_meas_adj;then
i=1
cat src_rec/$srfile |
while read evtfile; do # loop over all virtual evts
  eid=`echo $evtfile |awk '{printf "%s.%s",$2,$1}'`#
  fwd_dir=solver/$mod/$eid
  #####
  echo "meas_adj commands" >$measdir/job${i}.sh
  #####
  let i=i+1
done # end of loop over evts
fi
#=====
```

FIGURE A.3: Command line of *run_preprocessing.sh*.

In pre_processing, data and synthetics are re-sampled to 1Hz, cut to preferred length and band-passed filtered before window selection in the next step as suggested by FLEXWIN manual. Filtered data are normalized by the maximum of its absolute amplitude and then multiply the maximum of aboulute amplitude of synthetics.

The pre-processed data and synthetics are saved as:

```
data/CI.ADO/CI.BAK.BXZ.sac.T010_T020
solver/m00/CI.ADO/CI.BAK.BXZ.fwd.semd.sac.T010_T020
```

```
cat /dev/null >CI.ADO/process.log
perl process_syn.pl -m CI.ADO/CMTSOLUTION -a CI.ADO/STATIONS -s 100.0 -l -20/235 -t 5/10 -x T005_T010 ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd >>CI.ADO/process.log
synmin=`saclst depmin f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T005_T010 |awk '{print $2}'` 
synmax=`saclst depmax f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T005_T010 |awk '{print $2}'` 
norm=`echo $synmin $synmax |awk '{if($1*$1>$2*$2) {print sqrt($1*$1);} else {print sqrt($2*$2)}}` 

perl process_data.pl -m CI.ADO/CMTSOLUTION -s 100.0 -l -20/235 -t 5/10 -v -n -A $norm -x T005_T010 CI.ADO/DATA_NORM/CI.AL.P.HXZ.sac >>CI.ADO/process.log
perl process_syn.pl -m CI.ADO/CMTSOLUTION -a CI.ADO/STATIONS -s 100.0 -l -20/235 -t 10/20 -x T010_T020 ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd >>CI.ADO/process.log
synmin=`saclst depmin f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T010_T020 |awk '{print $2}'` 
synmax=`saclst depmax f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T010_T020 |awk '{print $2}'` 
norm=`echo $synmin $synmax |awk '{if($1*$1>$2*$2) {print sqrt($1*$1);} else {print sqrt($2*$2)}}` 

perl process_data.pl -m CI.ADO/CMTSOLUTION -s 100.0 -l -20/235 -t 10/20 -v -n -A $norm -x T010_T020 CI.ADO/DATA_NORM/CI.AL.P.HXZ.sac >>CI.ADO/process.log
perl process_syn.pl -m CI.ADO/CMTSOLUTION -a CI.ADO/STATIONS -s 100.0 -l -20/235 -t 20/50 -x T020_T050 ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd >>CI.ADO/process.log
synmin=`saclst depmin f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T020_T050 |awk '{print $2}'` 
synmax=`saclst depmax f ..../solver/M01.set1/CI.ADO/OUTPUT_FILES/CI.AL.P.HXZ.fwd.semd.sac.T020_T050 |awk '{print $2}'` 
norm=`echo $synmin $synmax |awk '{if($1*$1>$2*$2) {print sqrt($1*$1);} else {print sqrt($2*$2)}}` 

perl process_data.pl -m CI.ADO/CMTSOLUTION -s 100.0 -l -20/235 -t 20/50 -v -n -A $norm -x T020_T050 CI.ADO/DATA_NORM/CI.AL.P.HXZ.sac >>CI.ADO/process.log
```

FIGURE A.4: Command line of job1.sh in seis_process_M01.set1

In measure_adj, we adapt several Perl scripts from Tape et al. [2009] to do misfit measurement and make adjoint sources.

run_measure_adj.pl– main Perl script for running the program MEAS_ADJ.

combine_3_adj_src.pl– combine adjoint sources from three different frequency band

A.4.2 *process_data.pl*

I modyfied "process_data.pl" by adding -v option to get EGFs from time domain derivation of cross-correlation functions, -n option to normalize the amplitude and -A option to make the ammplitude of EGFs comparable to the their corresponding SGFs from SEM.

```

cd CI.ADO
cat /dev/null >run.log
cp MEASUREMENT.WINDOWS.T005_T010 MEASUREMENT.WINDOWS
./run_measure_adj.pl M01.set1 -20/235 0 0 0 -2.0/0.01/24000 7 HX 5/10 0/0/0/1 -2.5/2.5/-1.0/1.0/0
.80 1/1.0/0.5 1/0.02/2.5/2.0/2.5/3.5/1.5 >>run.log
mv ADJOINT_SOURCES ADJOINT_SOURCES_T005_T010
mv window_chi ../../output/misfits/M01.set1_T005_T010_CI.ADO_window_chi
cp MEASUREMENT.WINDOWS.T010_T020 MEASUREMENT.WINDOWS
./run_measure_adj.pl M01.set1 -20/235 0 0 0 -2.0/0.01/24000 7 HX 10/20 0/0/0/1 -3.5/3.5/-1.0/1.0/
0.80 1/1.0/0.5 1/0.02/2.5/2.0/2.5/3.5/1.5 >>run.log
mv ADJOINT_SOURCES ADJOINT_SOURCES_T010_T020
mv window_chi ../../output/misfits/M01.set1_T010_T020_CI.ADO_window_chi
cp MEASUREMENT.WINDOWS.T020_T050 MEASUREMENT.WINDOWS
./run_measure_adj.pl M01.set1 -20/235 0 0 0 -2.0/0.01/24000 7 HX 20/50 0/0/0/1 -4.5/4.5/-1.0/1.0/
0.80 1/1.0/0.5 1/0.02/2.5/2.0/2.5/3.5/1.5 >>run.log
mv ADJOINT_SOURCES ADJOINT_SOURCES_T020_T050
mv window_chi ../../output/misfits/M01.set1_T020_T050_CI.ADO_window_chi
./combine_3_adj_src.pl HX ADJOINT_SOURCES_T005_T010 ADJOINT_SOURCES_T010_T020 ADJOINT_SOURCES_T020_T050 ADJOINT_SOURCES iker07 iker07 iker07 >>run.log
cd ../..
$cdir=`pwd`/
cd solver/M01.set1/CI.ADO
rm -rf SEM
mkdir -p SEM

cd SEM
for meas_adj in `ls $cdir/measure_adj_M01.set1/CI.ADO/ADJOINT_SOURCES/*.adj`;do
    adj=`echo $meas_adj |awk -F$cdir/measure_adj_M01.set1/CI.ADO/ADJOINT_SOURCES/ '{print $2}'` 
    stnm=`echo $adj |awk -F. '{print $1}'` 
    net=`echo $adj |awk -F. '{print $2}'` 
    ch=`echo $adj |awk -F. '{print $3}'` 
    adj_new=$net.$stnm.$ch.adj
    mv $meas_adj $adj_new
    cat $cdir/measure_adj_M01.set1/CI.ADO/ADJOINT_SOURCES/STATIONS_ADJOINT |sed -n '2,$p' >../DATA/STATIONS_ADJOINT
done
cd .. # done of make adj
cd $cdir

```

FIGURE A.5: Command line of job1.sh in measure_adj_M01.set1.

References

- Akcelik, V., Biros, G., and Ghattas, O. (2002). Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation. In *Supercomputing, ACM/IEEE 2002 Conference*, pages 41–41. IEEE.
- Aki, K. and Richards, P. G. (2002). *Quantitative seismology*.
- An, M., Feng, M., and Zhao, Y. (2009). Destruction of lithosphere within the north China craton inferred from surface wave tomography. *Geochemistry, Geophysics, Geosystems*, 10(8).
- Barak, S., Klemperer, S. L., and Lawrence, J. F. (2015). San Andreas Fault dip, Peninsular Ranges mafic lower crust and partial melt in the Salton Trough, Southern California, from ambient-noise tomography. *Geochemistry, Geophysics, Geosystems*, 16(11):3946–3972.
- Basini, P., Nissen-Meyer, T., Boschi, L., Casarotti, E., Verbeke, J., Schenk, O., and Giardini, D. (2013). The influence of nonuniform ambient noise on crustal tomography in Europe. *Geochemistry, Geophysics, Geosystems*, 14(5):1471–1492.
- Bensen, G., Ritzwoller, M., Barmin, M., Levshin, A., Lin, F., Moschetti, M., Shapiro, N., and Yang, Y. (2007). Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements. *Geophysical Journal International*, 169(3):1239–1260.
- Bensen, G., Ritzwoller, M., and Yang, Y. (2009). A 3-D shear velocity model of the crust and uppermost mantle beneath the United States from ambient seismic noise. *Geophysical Journal International*, 177(3):1177–1196.

- Boore, D. M. and Toksöz, M. N. (1969). Rayleigh wave particle motion and crustal structure. *Bulletin of the Seismological Society of America*, 59(1):331–346.
- Boschi, L. and Ekström, G. (2002). New images of the Earth’s upper mantle from measurements of surface wave phase velocity anomalies. *Journal of Geophysical Research: Solid Earth*, 107(B4).
- Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2016). Global adjoint tomography: first-generation model. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 207(3):1739–1766.
- Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, 60(5):1457–1473.
- Capdeville, Y. (2005). An efficient Born normal mode method to compute sensitivity kernels and synthetic seismograms in the Earth. *Geophysical Journal International*, 163(2):639–646.
- Carcione, J. M. (1994). The wave equation in generalized coordinates. *Geophysics*, 59(12):1911–1919.
- Chen, M., Huang, H., Yao, H., Hilst, R., and Niu, F. (2014). Low wave speed zones in the crust beneath SE Tibet revealed by ambient noise adjoint tomography. *Geophysical Research Letters*, 41(2):334–340.
- Chen, M., Niu, F., Liu, Q., Tromp, J., and Zheng, X. (2015). Multiparameter adjoint tomography of the crust and upper mantle beneath East Asia: 1. Model construction and comparisons. *Journal of Geophysical Research: Solid Earth*, 120(3):1762–1786.
- Chen, M. and Tromp, J. (2007). Theoretical and numerical investigations of global and regional seismic wave propagation in weakly anisotropic earth models. *Geophysical Journal International*, 168(3):1130–1152.
- Chen, P., Jordan, T. H., and Lee, E.-J. (2010). Perturbation kernels for generalized seismological data functionals (GSDF). *Geophysical Journal International*, 183(2):869–883.

- Chen, P., Jordan, T. H., and Zhao, L. (2007a). Full three-dimensional tomography: a comparison between the scattering-integral and adjoint-wavefield methods. *Geophysical Journal International*, 170(1):175–181.
- Chen, P., Zhao, L., and Jordan, T. H. (2007b). Full 3D tomography for the crustal structure of the Los Angeles region. *Bulletin of the Seismological Society of America*, 97(4):1094–1120.
- Dahlen, F. and Baig, A. M. (2002). Fréchet kernels for body-wave amplitudes. *Geophysical Journal International*, 150(2):440–466.
- Dahlen, F., Hung, S.-H., and Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes-I. Theory. *Geophysical Journal International*, 141(1):157–174.
- De Hoop, M. V. and van Der Hilst, R. D. (2005). On sensitivity kernels for 'wave-equation' transmission tomography. *Geophysical Journal International*, 160(2):621–633.
- Denolle, M., Dunham, E., Prieto, G., and Beroza, G. (2013). Ground motion prediction of realistic earthquake sources using the ambient seismic field. *Journal of Geophysical Research: Solid Earth*, 118(5):2102–2118.
- Denolle, M., Dunham, E., Prieto, G., and Beroza, G. (2014). Strong ground motion prediction using virtual earthquakes. *Science*, 343(6169):399–403.
- Draganov, D., Campman, X., Thorbecke, J., Verdel, A., and Wapenaar, K. (2009). Reflection images from ambient seismic noise. *Geophysics*, 74(5):A63–A67.
- Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the earth and planetary interiors*, 25(4):297–356.
- Fang, H., Yao, H., Zhang, H., Huang, Y.-C., and van der Hilst, R. D. (2015). Direct inversion of surface wave dispersion for three-dimensional shallow crustal structure based on ray tracing: Methodology and application. *Geophysical Journal International*, 201(3):1251–1263.
- Ferreira, A. M. and Woodhouse, J. H. (2007). Observations of long period Rayleigh wave ellipticity. *Geophysical Journal International*, 169(1):161–169.

- Fichtner, A. (2011). Full seismic waveform modelling and inversion: Springer Science & Business Media.
- Fichtner, A. (2014). Source and processing effects on noise correlations. *Geophysical Journal International*, 197(3):1527–1531.
- Fichtner, A., Kennett, B. L., Igel, H., and Bunge, H.-P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, 179(3):1703–1725.
- Fichtner, A., Kennett, B. L., Igel, H., and Bunge, H.-P. (2010). Full waveform tomography for radially anisotropic structure: new insights into present and past states of the Australasian upper mantle. *Earth and Planetary Science Letters*, 290(3):270–280.
- Fichtner, A., Stehly, L., Ermert, L., and Boehm, C. (2016). Generalised interferometry-I. Theory for inter-station correlations. *Geophysical Journal International*, page ggw420.
- Gao, H. and Shen, Y. (2014). Upper mantle structure of the Cascades from full-wave ambient noise tomography: Evidence for 3D mantle upwelling in the back-arc. *Earth and Planetary Science Letters*, 390:222–233.
- Gauthier, O., Virieux, J., and Tarantola, A. (1986). Two-dimensional nonlinear inversion of seismic waveforms: Numerical results. *Geophysics*, 51(7):1387–1403.
- Gee, L. S. and Jordan, T. H. (1992). Generalized seismological data functionals. *Geophysical Journal International*, 111(2):363–390.
- Haned, A., Stutzmann, E., Schimmel, M., Kiselev, S., Davaille, A., and Yelles-Chaouche, A. (2015). Global tomography using seismic hum. *Geophysical Journal International*, 204(2):1222–1236.
- Hauksson, E. (2000). Crustal structure and seismicity distribution adjacent to the Pacific and North America plate boundary in southern California. *Journal of Geophysical Research: Solid Earth*, 105(B6):13875–13903.
- Herrera, I. and Mal, A. (1965). A perturbation method for elastic wave propagation: 2. Small inhomogeneities. *Journal of Geophysical Research*, 70(4):871–883.

- Hung, S.-H., Chen, W.-P., and Chiao, L.-Y. (2011). A data-adaptive, multiscale approach of finite-frequency, traveltime tomography with special reference to P and S wave data from central Tibet. *Journal of Geophysical Research: Solid Earth*, 116(B6).
- Hung, S.-H., Chen, W.-P., Chiao, L.-Y., and Tseng, T.-L. (2010). First multi-scale, finite-frequency tomography illuminates 3-D anatomy of the Tibetan plateau. *Geophysical Research Letters*, 37(6).
- Hung, S.-H., Dahlen, F., and Nolet, G. (2000). Fréchet kernels for finite-frequency traveltimes-II. Examples. *Geophysical Journal International*, 141(1):175–203.
- Hung, S.-H., Shen, Y., and Chiao, L.-Y. (2004). Imaging seismic velocity structure beneath the Iceland hot spot: A finite frequency approach. *Journal of Geophysical Research: Solid Earth*, 109(B8).
- Jiang, C., Yang, Y., and Zheng, Y. (2014). Penetration of mid-crustal low velocity zone across the Kunlun Fault in the NE Tibetan Plateau revealed by ambient noise tomography. *Earth and Planetary Science Letters*, 406:81–92.
- Kennett, B. and Engdahl, E. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International*, 105(2):429–465.
- Kennett, B., Engdahl, E., and Buland, R. (1995). Constraints on seismic velocities in the Earth from traveltimes. *Geophysical Journal International*, 122(1):108–124.
- Komatitsch, D. and Tromp, J. (1999). Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical journal international*, 139(3):806–822.
- Komatitsch, D. and Vilotte, J.-P. (1998). The spectral element method: An efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the seismological society of America*, 88(2):368–392.
- Krischer, L., Fichtner, A., Zukauskaite, S., and Igel, H. (2015). Large-Scale Seismic Inversion Framework. *Seismological Research Letters*.
- Lee, E.-J., Chen, P., Jordan, T. H., Maechling, P. B., Denolle, M. A., and Beroza, G. C. (2014). Full-3-D tomography for crustal structure in southern California based on the

- scattering-integral and the adjoint-wavefield methods. *Journal of Geophysical Research: Solid Earth*, 119(8):6421–6451.
- Li, G., Chen, H., Niu, F., Guo, Z., Yang, Y., and Xie, J. (2016). Measurement of Rayleigh wave ellipticity and its application to the joint inversion of high-resolution S wave velocity structure beneath northeast China. *Journal of Geophysical Research: Solid Earth*, 121(2):864–880.
- Li, X.-D. and Romanowicz, B. (1995). Comparison of global waveform inversions with and without considering cross-branch modal coupling. *Geophysical Journal International*, 121(3):695–709.
- Lin, F.-C., Moschetti, M. P., and Ritzwoller, M. H. (2008). Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps. *Geophysical Journal International*, 173(1):281–298.
- Lin, F.-C., Ritzwoller, M. H., and Snieder, R. (2009). Eikonal tomography: surface wave tomography by phase front tracking across a regional broad-band seismic array. *Geophysical Journal International*, 177(3):1091–1110.
- Lin, F.-C., Ritzwoller, M. H., Townend, J., Bannister, S., and Savage, M. K. (2007). Ambient noise Rayleigh wave tomography of New Zealand. *Geophysical Journal International*, 170(2):649–666.
- Lin, F.-C., Ritzwoller, M. H., Yang, Y., Moschetti, M. P., and Fouch, M. J. (2011). Complex and variable crustal and uppermost mantle seismic anisotropy in the western United States. *Nature Geoscience*, 4(1):55.
- Lin, F.-C., Schmandt, B., and Tsai, V. C. (2012). Joint inversion of Rayleigh wave phase velocity and ellipticity using USArray: Constraining velocity and density structure in the upper crust. *Geophysical Research Letters*, 39(12).
- Lin, F.-C., Tsai, V. C., and Schmandt, B. (2014). 3-D crustal structure of the western United States: application of Rayleigh-wave ellipticity extracted from noise cross-correlations. *Geophysical Journal International*, 198(2):656–670.
- Lin, F.-C., Tsai, V. C., Schmandt, B., Duputel, Z., and Zhan, Z. (2013). Extracting seismic core phases with array interferometry. *Geophysical Research Letters*, 40(6):1049–1053.

- Liu, K. and Zhou, Y. (2016). Travelling-wave Green tensor and near-field Rayleigh-wave sensitivity. *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, 205(1):134–145.
- Liu, Q. and Gu, Y. (2012). Seismic imaging: From classical to adjoint tomography. *Tectonophysics*, 566:31–66.
- Liu, Q. and Tromp, J. (2006). Finite-frequency kernels based on adjoint methods. *Bulletin of the Seismological Society of America*, 96(6):2383–2397.
- Liu, Q. and Tromp, J. (2008). Finite-frequency sensitivity kernels for global seismic wave propagation based upon adjoint methods. *Geophysical Journal International*, 174(1):265–286.
- Liu, Y., Niu, F., Chen, M., and Yang, W. (2017). 3-D crustal and uppermost mantle structure beneath NE China revealed by ambient noise adjoint tomography. *Earth and Planetary Science Letters*, 461:20–29.
- Lobkis, O. I. and Weaver, R. L. (2001). On the emergence of the Green’s function in the correlations of a diffuse field. *The Journal of the Acoustical Society of America*, 110(6):3011–3017.
- Luo, Y. (2012). Seismic imaging and inversion based on spectral-element and adjoint methods.
- Luo, Y., Xu, Y., and Yang, Y. (2013). Crustal radial anisotropy beneath the Dabie orogenic belt from ambient noise tomography. *Geophysical Journal International*, 195(2):1149–1164.
- Maggi, A., Tape, C., Chen, M., Chao, D., and Tromp, J. (2009). An automated time-window selection algorithm for seismic tomography. *Geophysical Journal International*, 178(1):257–281.
- Malischewsky, P. G. and Scherbaum, F. (2004). Love’s formula and H/V-ratio (ellipticity) of Rayleigh waves. *Wave motion*, 40(1):57–67.

- Marquering, H., Dahlen, F., and Nolet, G. (1999). Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox. *Geophysical Journal International*, 137(3):805–815.
- Masters, G., Laske, G., Bolton, H., and Dziewonski, A. (2000). The relative behavior of shear velocity, bulk sound speed, and compressional velocity in the mantle: implications for chemical and thermal structure. *Earth's deep interior: mineral physics and tomography from the atomic to the global scale*, pages 63–87.
- Maupin, V. (2017). 3-D sensitivity kernels of the Rayleigh wave ellipticity. *Geophysical Journal International*, 211(1):107–119.
- Menke, W. (1984). Geophysical Data Analysis: Discrete Inverse Theory Academic. *New York*.
- Modrak, R. and Tromp, J. (2016). Seismic waveform inversion best practices: regional, global and exploration test cases. *Geophysical Journal International*, 206(3):1864–1889.
- Montagner, J.-P. and Anderson, D. L. (1989). Petrological constraints on seismic anisotropy. *Physics of the earth and planetary interiors*, 54(1-2):82–105.
- Montelli, R., Nolet, G., Dahlen, F., and Masters, G. (2006). A catalogue of deep mantle plumes: New results from finite-frequency tomography. *Geochemistry, Geophysics, Geosystems*, 7(11).
- Montelli, R., Nolet, G., Dahlen, F., Masters, G., Engdahl, E. R., and Hung, S.-H. (2004). Finite-frequency tomography reveals a variety of plumes in the mantle. *Science*, 303(5656):338–343.
- Moschetti, M., Ritzwoller, M., Lin, F., and Yang, Y. (2010a). Seismic evidence for widespread western-US deep-crustal deformation caused by extension. *Nature*, 464(7290):885.
- Moschetti, M., Ritzwoller, M., Lin, F.-C., and Yang, Y. (2010b). Crustal shear wave velocity structure of the western United States inferred from ambient seismic noise and earthquake data. *Journal of Geophysical Research: Solid Earth*, 115(B10).

- Muir, J. B. and Tsai, V. C. (2017). Rayleigh-Wave H/V via Noise Cross Correlation in Southern California. *Bulletin of the Seismological Society of America*, 107(5):2021–2027.
- Nakata, N., Chang, J. P., Lawrence, J. F., and Boué, P. (2015). Body wave extraction and tomography at Long Beach, California, with ambient-noise interferometry. *Journal of Geophysical Research: Solid Earth*, 120(2):1159–1173.
- Nishida, K. (2013). Global propagation of body waves revealed by cross-correlation analysis of seismic hum. *Geophysical Research Letters*, 40(9):1691–1696.
- Nishida, K., Montagner, J.-P., and Kawakatsu, H. (2009). Global surface wave tomography using seismic hum. *Science*, 326(5949):112–112.
- Obrebski, M., Allen, R. M., Pollitz, F., and Hung, S.-H. (2011). Lithosphere–asthenosphere interaction beneath the western United States from the joint inversion of body-wave traveltimes and surface-wave phase velocities. *Geophysical Journal International*, 185(2):1003–1021.
- Olsen, K. B., Pechmann, J. C., and Schuster, G. T. (1995). Simulation of 3D elastic wave propagation in the Salt Lake Basin. *Bulletin of the Seismological Society of America*, 85(6):1688–1710.
- Paige, C. C. and Saunders, M. A. (1982). An algorithm for sparse linear equations and sparse least quarts. *ACM Trans. Math. Softw*, 8.
- Pike, E. R. and Sabatier, P. C. (2001). *Scattering, Two-Volume Set: Scattering and Inverse Scattering in Pure and Applied Science*. Elsevier.
- Pilz, M., Parolai, S., Picozzi, M., and Bindi, D. (2012). Three-dimensional shear wave velocity imaging by ambient seismic noise tomography. *Geophysical Journal International*, 189(1):501–512.
- Poli, P., Campillo, M., Pedersen, H., Group, L. W., et al. (2012a). Body-wave imaging of earth's mantle discontinuities from ambient seismic noise. *Science*, 338(6110):1063–1065.

- Poli, P., Pedersen, H., and Campillo, M. (2012b). Emergence of body waves from cross-correlation of short period seismic noise. *Geophysical Journal International*, 188(2):549–558.
- Prieto, G. A. and Beroza, G. C. (2008). Earthquake ground motion prediction using the ambient seismic field. *Geophysical Research Letters*, 35(14).
- Prieto, G. A., Denolle, M., Lawrence, J. F., and Beroza, G. C. (2011). On amplitude information carried by the ambient seismic field. *Comptes Rendus Geoscience*, 343(8):600–614.
- Prieto, G. A., Lawrence, J. F., and Beroza, G. C. (2009). Anelastic Earth structure from the coherency of the ambient seismic field. *Journal of Geophysical Research: Solid Earth*, 114(B7).
- Rawlinson, N., Pozgay, S., and Fishwick, S. (2010). Seismic tomography: a window into deep earth. *Physics of the Earth and Planetary Interiors*, 178(3-4):101–135.
- Rawlinson, N. and Sambridge, M. (2003). Seismic traveltimes tomography of the crust and lithosphere. *Advances in Geophysics*, 46:81–199.
- Ritzwoller, M. H., Shapiro, N. M., Barmin, M. P., and Levshin, A. L. (2002). Global surface wave diffraction tomography. *Journal of Geophysical Research: Solid Earth*, 107(B12).
- Roux, P., Sabra, K. G., Kuperman, W. A., and Roux, A. (2005). Ambient noise cross correlation in free space: Theoretical approach. *The Journal of the Acoustical Society of America*, 117(1):79–84.
- Schmandt, B. and Humphreys, E. (2010). Seismic heterogeneity and small-scale convection in the southern California upper mantle. *Geochemistry, Geophysics, Geosystems*, 11(5).
- Schumacher, F. (2014). Modularized iterative full seismic waveform inversion for 3D-heterogeneous media based on waveform sensitivity kernels.
- Schumacher, F. and Friederich, W. (2016). ASKI: A modular toolbox for scattering-integral-based seismic full waveform inversion and sensitivity analysis utilizing external forward codes. *SoftwareX*, 5:252–259.
- Sexton, J. L., Rudman, A., and Mead, J. (1977). Ellipticity of Rayleigh waves recorded in the Midwest. *Bulletin of the Seismological Society of America*, 67(2):369–382.

- Shapiro, N. M., Campillo, M., Stehly, L., and Ritzwoller, M. H. (2005). High-resolution surface-wave tomography from ambient seismic noise. *Science*, 307(5715):1615–1618.
- Shaw, J. H., Plesch, A., Tape, C., Suess, M. P., Jordan, T. H., Ely, G., Hauksson, E., Tromp, J., Tanimoto, T., Graves, R., et al. (2015). Unified structural representation of the southern California crust and upper mantle. *Earth and Planetary Science Letters*, 415:1–15.
- Shen, W., Ritzwoller, M. H., and Schulte-Pelkum, V. (2013). A 3-D model of the crust and uppermost mantle beneath the central and western US by joint inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research: Solid Earth*, 118(1):262–276.
- Sieminski, A., Liu, Q., Trampert, J., and Tromp, J. (2007a). Finite-frequency sensitivity of body waves to anisotropy based upon adjoint methods. *Geophysical Journal International*, 171(1):368–389.
- Sieminski, A., Liu, Q., Trampert, J., and Tromp, J. (2007b). Finite-frequency sensitivity of surface waves to anisotropy based upon adjoint methods. *Geophysical Journal International*, 168(3):1153–1174.
- Snieder, R. (2002). General theory of elastic wave scattering. In *Scattering*, pages 528–542. Elsevier.
- Snieder, R. (2004). Extracting the Green’s function from the correlation of coda waves: A derivation based on stationary phase. *Physical Review E*, 69(4):046610.
- Stehly, L., Campillo, M., and Shapiro, N. (2006). A study of the seismic noise from its long-range correlation properties. *Journal of Geophysical Research: Solid Earth*, 111(B10).
- Süss, M. P. and Shaw, J. H. (2003). P wave seismic velocity structure derived from sonic logs and industry reflection data in the Los Angeles basin, California. *Journal of Geophysical Research: Solid Earth*, 108(B3).
- Tang, Y., Obayashi, M., Niu, F., Grand, S. P., Chen, Y. J., Kawakatsu, H., Tanaka, S., Ning, J., and Ni, J. F. (2014). Changbaishan volcanism in northeast China linked to subduction-induced mantle upwelling. *Nature Geoscience*, 7(6):470.

- Tanmoto, T. and Alvizuri, C. (2006). Inversion of the HZ ratio of microseisms for S-wave velocity in the crust. *Geophysical Journal International*, 165(1):323–335.
- Tanmoto, T. and Rivera, L. (2008). The ZH ratio method for long-period seismic data: sensitivity kernels and observational techniques. *Geophysical Journal International*, 172(1):187–198.
- Tanmoto, T. and Tsuboi, S. (2009). Variational principle for Rayleigh wave ellipticity. *Geophysical Journal International*, 179(3):1658–1668.
- Tape, C. (2009). *Seismic tomography of southern California using adjoint methods*. PhD thesis, California Institute of Technology.
- Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2009). Adjoint tomography of the southern California crust. *Science*, 325(5943):988–992.
- Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2010). Seismic tomography of the southern California crust based on spectral-element and adjoint methods. *Geophysical Journal International*, 180(1):433–462.
- Tape, C., Liu, Q., and Tromp, J. (2007). Finite-frequency tomography using adjoint methods—Methodology and examples using membrane surface waves. *Geophysical Journal International*, 168(3):1105–1129.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*. SIAM.
- Tromp, J., Komattisch, D., and Liu, Q. (2008). Spectral-element and adjoint methods in seismology. *Communications in Computational Physics*, 3(1):1–32.
- Tromp, J., Luo, Y., Hanasoge, S., and Peter, D. (2010). Noise cross-correlation sensitivity kernels. *Geophysical Journal International*, 183(2):791–819.
- Tromp, J., Tape, C., and Liu, Q. (2005). Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International*, 160(1):195–216.

- Tsai, V. C. (2009). On establishing the accuracy of noise tomography travel-time measurements in a realistic medium. *Geophysical Journal International*, 178(3):1555–1564.
- Tsuboi, S. and Saito, M. (1983). Partial derivatives of Rayleigh wave particle motion. *Journal of Physics of the Earth*, 31(2):103–113.
- Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.
- Wang, K., Liu, Q., and Yang, Y. (2018a). Sensitivity kernels for multi-component ambient noise empirical Green's functions based on adjoint method. *Geophysical Journal International (under revision)*.
- Wang, K., Luo, Y., and Li, H. (2014a). The nature of ambient noise over a field in western Junggar basin near Karamay, China. In *The 6th International Conference on Environmental and Engineering Geophysics*, pages 92–97.
- Wang, K., Luo, Y., and Yang, Y. (2016). Correction of phase velocity bias caused by strong directional noise sources in high-frequency ambient noise tomography: a case study in Karamay, China. *Geophysical Journal International*, 205(2):715–727.
- Wang, K., Luo, Y., Zhao, K., and Zhang, L. (2014b). Body waves revealed by spatial stacking on long-term cross-correlation of ambient noise. *Journal of Earth Science*, 25(6):977.
- Wang, K., Yang, Y., Basini, P., Tong, P., Tape, C., and Liu, Q. (2018b). Refined crustal and uppermost mantle structure of southern California by ambient noise adjoint tomography. *Geophysical Journal International*, 215(2):844–863.
- Wapenaar, K., Draganov, D., Snieder, R., Campman, X., and Verdel, A. (2010). Tutorial on seismic interferometry: Part 1-Basic principles and applications. *Geophysics*, 75(5):75A195–75A209.
- Wu, R.-S. (1989). The perturbation method in elastic wave scattering. *pure and applied geophysics*, 131(4):605–637.
- Xie, J., Ritzwoller, M. H., Shen, W., Yang, Y., Zheng, Y., and Zhou, L. (2013). Crustal radial anisotropy across eastern Tibet and the western Yangtze craton. *Journal of Geophysical Research: Solid Earth*, 118(8):4226–4252.

- Yang, Y. and Forsyth, D. W. (2006). Rayleigh wave phase velocities, small-scale convection, and azimuthal anisotropy beneath southern California. *Journal of Geophysical Research: Solid Earth*, 111(B7).
- Yang, Y. and Ritzwoller, M. H. (2008). Characteristics of ambient seismic noise as a source for surface wave tomography. *Geochemistry, Geophysics, Geosystems*, 9(2).
- Yang, Y., Ritzwoller, M. H., Levshin, A. L., and Shapiro, N. M. (2007). Ambient noise Rayleigh wave tomography across Europe. *Geophysical Journal International*, 168(1):259–274.
- Yang, Y., Ritzwoller, M. H., Lin, F.-C., Moschetti, M., and Shapiro, N. M. (2008). Structure of the crust and uppermost mantle beneath the western United States revealed by ambient noise and earthquake tomography. *Journal of Geophysical Research: Solid Earth*, 113(B12).
- Yano, T., Tanimoto, T., and Rivera, L. (2009). The ZH ratio method for long-period seismic data: inversion for S-wave velocity structure. *Geophysical Journal International*, 179(1):413–424.
- Yao, H., Campman, X., Maarten, V., and van der Hilst, R. D. (2009). Estimation of surface wave Green's functions from correlation of direct waves, coda waves, and ambient noise in SE Tibet. *Physics of the Earth and Planetary Interiors*, 177(1):1–11.
- Yao, H. and Van Der Hilst, R. D. (2009). Analysis of ambient noise energy distribution and phase velocity bias in ambient noise tomography, with application to SE Tibet. *Geophysical Journal International*, 179(2):1113–1132.
- Yao, H., van Der Hilst, R. D., and Maarten, V. (2006). Surface-wave array tomography in SE Tibet from ambient seismic noise and two-station analysis. Phase velocity maps. *Geophysical Journal International*, 166(2):732–744.
- Yao, H., Van Der Hilst, R. D., and Montagner, J.-P. (2010). Heterogeneity and anisotropy of the lithosphere of SE Tibet from surface wave array tomography. *Journal of Geophysical Research: Solid Earth*, 115(B12).

- Zhan, Z., Ni, S., Helmberger, D. V., and Clayton, R. W. (2010). Retrieval of Moho-reflected shear wave arrivals from ambient seismic noise. *Geophysical Journal International*, 182(1):408–420.
- Zhang, H. and Thurber, C. H. (2003). Double-difference tomography: The method and its application to the Hayward fault, California. *Bulletin of the Seismological Society of America*, 93(5):1875–1889.
- Zhao, L. and Jordan, T. H. (2006). Structural sensitivities of finite-frequency seismic waves: a full-wave approach. *Geophysical Journal International*, 165(3):981–990.
- Zhao, L., Jordan, T. H., and Chapman, C. H. (2000). Three-dimensional Fréchet differential kernels for seismic delay times. *Geophysical Journal International*, 141(3):558–576.
- Zhao, L., Jordan, T. H., Olsen, K. B., and Chen, P. (2005). Fréchet kernels for imaging regional earth structure based on three-dimensional reference models. *Bulletin of the Seismological Society of America*, 95(6):2066–2080.
- Zhou, Y., Dahlen, F., and Nolet, G. (2004). Three-dimensional sensitivity kernels for surface wave observables. *Geophysical Journal International*, 158(1):142–168.
- Zhou, Y., Liu, Q., and Tromp, J. (2011). Surface wave sensitivity: mode summation versus adjoint SEM. *Geophysical Journal International*, 187(3):1560–1576.
- Zhou, Y., Nolet, G., Dahlen, F., and Laske, G. (2006). Global upper-mantle structure from finite-frequency surface-wave tomography. *Journal of Geophysical Research: Solid Earth*, 111(B4).
- Zhu, H., Bozdağ, E., Peter, D., and Tromp, J. (2012). Structure of the European upper mantle revealed by adjoint tomography. *Nature Geoscience*, 5(7):493–498.
- Zhu, H., Bozdağ, E., and Tromp, J. (2015). Seismic structure of the European upper mantle based on adjoint tomography. *Geophysical Journal International*, 201(1):18–52.
- Zhu, L. and Kanamori, H. (2000). Moho depth variation in southern California from teleseismic receiver functions. *Journal of Geophysical Research: Solid Earth*, 105(B2):2969–2980.

- Zigone, D., Ben-Zion, Y., Campillo, M., and Roux, P. (2015). Seismic tomography of the Southern California plate boundary region from noise-based Rayleigh and Love waves. *Pure and Applied Geophysics*, 172(5):1007–1032.