

WAVE EQUATION MIGRATION VELOCITY ANALYSIS FOR
VTI MODELS WITH GEOLOGICAL AND ROCK PHYSICS
CONSTRAINTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF GEOPHYSICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Yunyue Li
July 2014

© Copyright by Yunyue Li 2014
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Biondo L. Biondi) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Jon F. Claerbout)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Gary Mavko)

Approved for the University Committee on Graduate Studies

Abstract

Wave-equation migration velocity analysis (WEMVA) is a powerful technique for robust velocity model building when the subsurface is complex and the starting model is far from true. However, this traditional isotropic WEMVA technique cannot explain anisotropic wave phenomenon, which has significant effects when the reflectors are steeply dipping and/or the waves are traveling at large angles. Furthermore, errors in the anisotropic parameters cause similar defocusing as errors in isotropic velocity. Any defocusing caused by anisotropy may be translated to false updates in the isotropic velocity, which lead to further mispositioning and misinterpretations. However, simple extension of the isotropic WEMVA to the anisotropic medium cannot provide a unique and reliable solution due to the nonlinear and underdetermined nature of the anisotropic model building problem. Many anisotropic models with vastly different geological interpretations may explain the surface seismic data equally well.

This thesis addresses these issues by including anisotropic effects in WEMVA and by integrating other useful information from geology and rock physics to better regularize the inversion. First, I extend the isotropic WEMVA method to the anisotropic medium to evaluate and update the anisotropic Earth models. Instead of the industry standard ray-based tomographic methods, the anisotropic WEMVA technique uses wavefields as information carrier to handle both the complex subsurface and the frequency-dependent behavior of the wave propagation. I include the geological information using steering filters to regularize the gradients. Both synthetic and field 2-D examples show that the anisotropic WEMVA technique can resolve the errors in both velocity and anisotropic parameters. Consequently, the migration images are

improved with higher resolution and correct depths for the reflectors.

Second, I mitigate the ambiguities among the anisotropic parameters using the rock physics prior information. I model shale anisotropy in the seismic scale from well log measurements and the interpretation results from a previous isotropic processing workflow . By varying the key parameters to the rock physics models, I include different geological scenarios and parameter uncertainties in the rock physics modeling. As a result, multiple realizations of the anisotropic parameters are obtained and used as the rock physics constraints for the seismic data inversion. On a synthetic example, I show the correct rock physics regularization accelerates the convergence of the well constrained parameter and brings extra information for the poorly constrained parameters. Anisotropic WEMVA inversion of a 3-D field data set yields models that are consistent with seismic data, geological knowledge, as well as rock physics information. The migration image based on the inverted models shows better-resolved faulting discontinuities and better imaged salt flanks.

Preface

The electronic version of this report¹ makes the included programs and applications available to the reader. The markings **ER**, **CR**, and **NR** are promises by the author about the reproducibility of each figure result. Reproducibility is a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

ER denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g., SEG-EAGE data sets), or be available in the SEP data library². We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Windows system and the software downloadable from our website (SEP makerules, SEPScons, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library, or that the rules depend on commercial packages such as Matlab or Mathematica.

¹<http://sepwww.stanford.edu/private/docs/sep146>

²<http://sepwww.stanford.edu/public/docs/sepdata/lib/toc.html>

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. The primary reasons for the CR designation is that the processing requires 20 minutes or more.

NR denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.6 (using the Intel Fortran90 compiler) and the SEPlib-6.4.6 distribution, but the code should be portable to other architectures. Reader's suggestions are welcome. For more information on reproducing SEP's electronic documents, please visit <http://sepwww.stanford.edu/research/redoc/>.

Acknowledgments

It takes great efforts and many people's help for me to break the cocoon and to transform from a student to a researcher.

I am most grateful to Biondo Biondi, my adviser, not only for giving me the opportunity to be part of the unique research group, but also for giving me complete freedom to explore the research areas, and to try, to fail and eventually to be able to stand on my own feet. I am deeply grateful to Jon Claerbout. It was a privilege to work with him and learn from his genius mind. I thank Bob Clapp for his help with any difficulties and guidance through my own blindness to my problems. I thank Gary Mavko for being always available for discussions and for being supportive with his rock physics expertise. I thank my thesis committee member Tapan Mukerji for his well-taught lectures and his broad view of reservoir modeling. I also thank Lexing Yin for chairing my defense.

I thank my teacher, Xiaohong Chen, for revealing the interesting problems in geophysics that can be solved using mathematical tools. I thank my undergraduate thesis advisors, Bangrang Di and Jianxin Wei for their continue support and encouragement during my graduate study at Stanford.

I learned a lot from my internship experiences and many great mentors. I thank Dave Nichols for teaching me inverse theory from a different prospective and pouring me with invaluable research ideas. I thank Peng Shen for training me to validate my physical intuitions with solid mathematical derivations. I thank René-Édouard Plessix for showing me a great researcher can be successful both in academia and in

industry.

Many aspects of my research were built on previous work by Guojian Shan, Claudio Guerra, Yaxun Tang, and Gboyega Ayeni. They are my SEP mentors. I am very fortunate to have shared office with some of the finest SEPerS. Mohammad Maysami helped me greatly adapt to SEP environment and improve my English. Mandy Wong and Ohad Barak, whom after five years have practically become my families at Stanford. Many discussions, both geophysical and about general life topics, made me a maturer and wiser person. I will miss the intense discussions with Ali Almomin, Yang Zhang, and Yi Shen. I thank all the students arrived before and after me, Abdullah Al Theyab, Sjoerd de Ridder, Adam Halpert, Xukai Shen, Chris Leader, Kittinat Taweessintanon, Qiang Fu, Noha Farghal, Musa Maharramov, Jason Chang, Taylor Dahlke, Huy Le, Eileen Martin, Daniel Blatter, Guillaume Barnier, Gustavo Alves and Ettore Biondi, for being great friends and companions.

I am grateful to the geophysics staff for their excellent support. Specifically, Diane Lau ensured SEP meetings, travels, and research needs are taken care of promptly. Tara Ilich and Nancy Massarweh, the Student Services Manager, made sure all my questions about departmental requirements answered timely. Dennis Michael, Manager of CEES HPTC, provided superb technical support and HPC training during my use of the clusters.

I thank my dear friends at Stanford: Yi Wu, Jingyi Chen, Tiejuan Zhu, Jiang Bian, Yao Tong, Cheng Cheng, Qing Duan, Crystal Shi, Liyuan He and Yu Xia for spending our most beautiful youth together. I thank my long time dear friends: Hetong Xu, Hongye Zhang, Zhijian Wang, Haoran Xue, Yang Liu, Yue Sun, Pingping Dong, Wenjun Lu, Kunxuan Li, Ting Jia, and Xinglong Zhao for constantly nurturing me from my roots and reminding me that there is home to return.

Finally, I would like to thank my parents Qingguo Li and Shuzhen Wang for their unconditional support and love throughout my time at Stanford. I am utterly lucky to have found my boyfriend, Qiao Xiang, a fellow idealist and a faithful companion on the tough journey we both choose to take.

To my grandfather, Yulin Wang (1927 - 2009)
my uncle, Qingding Liu (1951 - 2011)
and my brother, Yunhui Li (1984 - 2008)

Contents

Abstract	vi
Preface	viii
Acknowledgments	xi
1 Introduction	1
2 WEMVA for VTI models	15
3 Rock physics constrained anisotropic WEMVA	53
4 3D Field Data Test	95
5 Conclusions	139
A Implicit finite differencing	141
B Rock physics modeling	145
C RTM-based WEMVA for VTI models	151

List of Tables

4.1	Computational cost for each WEMVA iteration.	113
B.1	End-member mineral elastic properties.	146

List of Figures

1.1	Left: A sketch of the wave-speed symmetry for VTI medium. Right: A sketch of the wave-speed symmetry for TTI medium. [NR]	3
1.2	Comparison of wavefronts propagated from a point source at $x = 0$, $z = 0$ in homogeneous isotropic and VTI media. Parameters used to propagate the wavefronts are: $\epsilon = 0$ and $\delta = 0$ for the solid line; $\epsilon = 0.2$ and $\delta = 0.2$ for the dashed line; and $\epsilon = 0.2$ and $\delta = -0.2$ for the dotted line. The vertical velocity of the medium is 2000m/s in all three cases. [ER]	4
1.3	Comparison between rays and wavefields. Panel (a) shows a snapshot of the rays overlaid by the Sigsbee2A velocity model, (b) shows a snapshot of the wavefields. [ER]	6
1.4	Angle domain common image gathers with (a) 5% slower velocity, (b) accurate velocity, and (c) 5% faster velocity. [CR]	7
1.5	Angle domain common image gathers with (a) 50% smaller ϵ , (b) accurate ϵ , and (c) 50% larger ϵ . [CR]	7
1.6	Angle domain common image gathers with (a) 50% smaller δ , (b) accurate δ , and (c) 50% larger δ . [CR]	8

1.7	WEMVA objective function with respect to (a) ϵ and δ , (b) vertical velocity and ϵ , and (c) vertical velocity and δ . The third parameter is accurate on each panel. Notice the different color scale in panel (a). [CR]	9
1.8	Stochastic rock physics modeling results for parameters v_v , ϵ and δ . [ER]	10
2.1	2D impulse responses for vertical velocity (left column), ϵ (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the two rows above. [ER]	25
2.2	2D impulse responses for NMO velocity (left column), η (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the top two rows. [ER]	26
2.3	Modified BP2007 model. (a): initial NMO velocity; (b): inverted NMO velocity; and (c) true NMO velocity. [CR]	32
2.4	Modified BP2007 model. (a): initial η model; (b): inverted η model; and (c) true η model. [CR]	33
2.5	Comparison of the anomaly. (a): inverted velocity anomaly; (b): true velocity anomaly; (c) inverted η anomaly and (d) true η anomaly. [CR]	34
2.6	Objective function as a function of iteration. [CR]	35
2.7	Stack image using (a): initial anisotropic model; (b): inverted anisotropic model; and (c) true anisotropic model. Flags 1 and 2 mark the depth of two representative reflectors in this section. Notice the depths of both reflectors have been corrected in the inverted image. The ellipses highlight the region with better focusing and more continuous reflectors. [CR]	36

2.8	Angle-domain common image gathers extracted at 250m interval between $x = 2000\text{m}$ and $x = 8000\text{m}$. (a): ADCIGs using initial anisotropic model, (b): ADCIGs using inverted anisotropic model. [CR]	38
2.9	Available datasets for anisotropic model building. From Bear et al. (2005). [NR]	40
2.10	Visualization of interactive modeling for anisotropy. The best anisotropic model honors sonic logs, OCS traveltimes and flattens the prestack surface seismic gathers at the same time. From Bear et al. (2005). [NR]	41
2.11	Source locations on a 2D line in (a) and corresponding receiver locations in (b). [CR]	42
2.12	Initial velocity model (a) and initial η model (b). [CR]	44
2.13	The initial stack image (Top panel) and initial angle domain common image gathers at $\text{CMP} = 7, 10, 13, 14, 14.5, 15 \text{ km}$ (Bottom row). [CR]	45
2.14	(a): Estimated dip field from the initial image on the top panel of Figure 2.13. (b): Reflectors created from a randomly distributed field. [CR]	46
2.15	Inverted velocity model (a) and inverted η model (b). [CR]	47
2.16	Updates in velocity model (a) and updates in η model (b) after inversion. [CR]	48
2.17	The final stack image (Top panel) and final angle domain common image gathers at $\text{CMP} = 7, 10, 13, 14, 14.5, 15 \text{ km}$ (Bottom row). Compared with Figure 2.13, improvements in continuity and enhancements in amplitude strength are highlighted by the oval. [CR]	49
3.1	The value of data fitting objective function extracted from (a) $v_v = 2\text{km}$ plane, (b) $\delta = 0.1$ plane, and (c) $\epsilon = 0.2$ plane. Notice the different color scale in panel (a). [CR]	57

3.2	The value of data fitting objective function extracted from (a) $v_v = 1.9\text{km}$ plane, (b) $v_v = 2\text{km}$ plane, and (c) $v_v = 2.1\text{km}$ plane. [CR] . . .	58
3.3	Comparison of the initial (left column) and the true (right column) models for BP2007 synthetic test. Top row: velocity models; middle row: ϵ models; bottom row: δ models. [CR]	62
3.4	True updates in velocity in (a), ϵ in (b) and δ in (c). Pseudo velocity logs at random locations are plotted in (d). [CR]	63
3.5	A stacked image of this section in panel (a), and the corresponding dip field in panel (b). [CR]	66
3.6	Diagonal elements of the covariance matrix of the VTI model. (a) C_{vv} : map of variance for velocity, (b) $C_{\epsilon\epsilon}$: map of variance for ϵ , and (c) $C_{\delta\delta}$: map of variance for δ . [ER]	67
3.7	Off-diagonal elements of the covariance matrix of the VTI model. (a) $C_{v\epsilon}$: map of covariance between velocity and ϵ , (b) $C_{v\delta}$: map of covariance between velocity and δ , and (c) $C_{\epsilon\delta}$: map of covariance between ϵ and δ . [ER]	68
3.8	Comparison of the interpolation results. (a): Inverted velocity using $\mathbf{C}_1^{1/2}$. (b): Inverted velocity using $\mathbf{C}_2^{1/2}$. (c): Inverted ϵ using $\mathbf{C}_2^{1/2}$. (d): Inverted δ using $\mathbf{C}_2^{1/2}$. [CR]	69
3.9	Comparison of the initial stacked image (a) and the true stacked image (b). Notice the depth shift between the images and the unfocused reflectors around $z = 3\text{km}$ and $x = 4\text{km}$. [CR]	71
3.10	Comparison of the initial ADCIGs (a) and the true ADCIGs (b). Notice the upward moveout throughout the section in the initial gathers. [CR]	72

3.11	Velocity updates after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [CR]	77
3.12	Updates in ϵ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. [CR]	78
3.13	Updates in δ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. [CR]	79
3.14	Velocity updates after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [CR]	80
3.15	Updates in ϵ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. [CR]	81
3.16	Updates in δ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. [CR]	82
3.17	Inverted vertical velocity model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true vertical velocity model for comparison. [CR]	83
3.18	Inverted ϵ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ model for comparison. [CR]	84
3.19	Inverted δ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ model for comparison. [CR]	85

3.20	Final stacked images with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Panel (d) shows the true stacked image. [CR]	88
3.21	Final ADCIGs with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Notice the improved angle coverage in the ellipse on the left and the improved flatness on the right. [CR]	90
3.22	Comparison of the initial objective function map (a), with the final objective function map with (b) no preconditioning, (c) diagonal preconditioning, and (d) full preconditioning. [CR]	91
3.23	Differences between the objective function maps. (a) Difference between Figure 3.22(d) and 3.22(b). (b) Difference between Figure 3.22(d) and 3.22(c). [CR]	92
3.24	Objective function value of the data fitting goal in (a) and the size of the gradient in (b) as a function of the iteration. Blue line: no preconditioning; red line: diagonal preconditioning; magenta line: full preconditioning. [CR]	92
4.1	(a): Input of the rock physics modeling from well log measurements. From left to right: gamma ray measurements, deduced shale contents from gamma ray and porosity measurements. (b): One instance of the rock physics modeling experiment. From left to right, panels are ϵ profile from the lamination model and inclusion model; δ profile from the lamination model and inclusion model; shale content at the well location. [ER]	100
4.2	Interpreted shale content (a) and modeled porosity (b). Average ϵ model (c) and average δ model (d). Both models are obtained from the stochastic rock physics modeling experiment. [ER]	102

4.3	Stochastic rock physics modeling results for (a) vertical velocity; (b) ϵ ; and (c) δ . The solid line in each panel denotes the numerical histogram for each model variable. The dashed line denotes the approximated multivariate Gaussian distribution based on the numerical realizations. [ER]	103
4.4	Diagonal elements of the covariance matrix. (a): Variance of vertical velocity; (b): Variance of ϵ ; and (c): Variance of δ . [ER]	104
4.5	Off-diagonal elements of the covariance matrix. (a): Covariance between vertical velocity and ϵ ; (b): Covariance between vertical velocity and δ ; and (c): Covariance between ϵ and δ . [ER]	105
4.6	Area covered by the E-Dragon Phase II acquisition. Ocean bottom cables were laid out on the sea bed under shallow water with an average depth of 120 m. [NR]	107
4.7	Cross-spread geometry of the GoM dataset. Panel (a): Source locations; Panel (b): Receiver locations. The source lines and the receiver lines are perpendicular to each other. I define the direction along the receiver lines the in-line direction and the perpendicular direction cross-line direction. [ER]	108
4.8	Acquisition pattern for one particular receiver. Source locations are denoted by "o" and the receiver location is denoted as the solid circle. For any given receiver, the maximum in-line offset is about 6 km and the maximum cross-line offset is about 3 km. [ER]	108
4.9	One receiver gather shows the dense cross-line sampling and very coarse in-line sampling. The slow arrivals after 3.5 s are the low frequency surface wave energies. [ER]	109
4.10	Receiver gather in Figure 4.9 after bandpass, muting and time weighting. Surface waves have been attenuated. Reflection events become more continuous especially in the later time. [ER]	110

4.11	(a): Isotropic velocity model provided by ray-based tomography. (b): Vertical velocity model, (c): initial ϵ model and (d): initial δ model from the rock physics modeling. [CR]	114
4.12	(a): Isotropic migration image using the isotropic velocity in Figure 4.11(a). (b): Anisotropic migration image using the models in Figure 4.11(b) to 4.11(d). [CR]	116
4.13	Same as Figure 4.12 but at a different location. [CR]	117
4.14	First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ using all shots. [CR]	118
4.15	First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ using one-third of all the shots. [CR]	119
4.16	First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ after rock physics preconditioning. [CR]	121
4.17	(a): Normalized objective function curve. (b): Normalized gradient magnitude curve. Both curves show fast convergence in the first few iterations. [CR]	123
4.18	Total updates after ten anisotropic WEMVA iterations in (a): vertical velocity model, (b): ϵ model and (c): δ model after ten anisotropic WEMVA iterations. [CR]	125
4.19	(a): Vertical velocity model, (b): ϵ model and (c): δ model after ten anisotropic WEMVA iterations. [CR]	126
4.20	Migration image using the initial anisotropic model. Labels 2 and 3 highlights areas with broken reflectors in the in-line and cross-line direction, respectively. [CR]	127

4.21	Migration image using the updated anisotropic model after ten iterations. Comparing to Figure 4.20, label 1 points out a distinct fault near the salt body which is much better imaged with the updated model. Dipping Reflectors in both circles 2 and 3 are more continuous. [CR]	128
4.22	Migration image using the initial anisotropic model. Sediments truncation and salt flank are ambiguous in circle 2, 3, and 4. [CR]	129
4.23	Migration image using the updated anisotropic model. Compared to Figure 4.22, sediments truncation and salt flank are much better imaged. On the depth slice, a set of faults (circle 1) caused by the salt intrusion are better focused. [CR]	130
4.24	Migration image using the initial anisotropic model. [CR]	131
4.25	Migration image using the updated anisotropic model. Compare to Figure 4.24, flag 1 points out the much higher resolution dipping reflector around the salt. The reflections in circle 2 are imaged deeper with better continuity. Two normal faults on the cross-line section are distinctly imaged due to the improved focusing. Circle 3 highlights one of the two faults. [CR]	132
4.26	ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 5$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. Although the shallow curvatures are roughly unchanged, the updated ADCIGs show significantly improved angle coverage at medium depths (between 2 km and 4 km). [CR]	133

4.27	ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 13.5$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. The downward curvatures in the shallow gathers (above 2 km) are flattened in the updated ADCIGs. Angle coverages of the ADCIGs at medium depths (between 2 km and 4 km) are improved, especially for the horizon at roughly 3.5 km depth. In the deep section, the updated ADCIGs show wider angle coverage and higher resolution. [CR]	134
4.28	ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 10$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. Updated ADCIGs show strong improvements below 4 km (more significantly on the three left panels) due to the cumulative effects of anisotropy along the wavepaths. [CR]	135
A.1	(a) Table for α and (b) table for β at discrete ϵ and δ values. [ER]	142
A.2	Derivatives of the coefficients with respect to ϵ and δ ; (a) table of $\frac{\partial\alpha}{\partial\epsilon}$; (b) table of $\frac{\partial\beta}{\partial\epsilon}$; (c) table of $\frac{\partial\alpha}{\partial\delta}$, and (d) table of $\frac{\partial\beta}{\partial\delta}$. [ER]	144
C.1	Smooth velocity model (a) and reflectivity model (b) used to generate the synthetic Born data. [CR]	163
C.2	The ϵ model (a) and δ model (b) used to generate the synthetic Born data. [CR]	164
C.3	(a) The $\delta\epsilon$ model to test the objective function. (b) Objective function vs. ϵ perturbation. The angle-domain objective function C.45 has a minimum at the correct epsilon model, and it has a semi-quadratic shape with respect to the model perturbation. [CR]	165
C.4	(a) True ϵ model and (b) true δ model used to generate the synthetic data. [CR]	166

C.5	(a) Initial ϵ model and (b) initial angle-domain common-image gathers. Gathers are taken at every 10th common image point from $x = 4$ km to $x = 8$ km. [CR]	166
C.6	(a) Inverted ϵ model and (b) final angle-domain common-image gathers using the inverted ϵ model in (a). Compared with Figure C.5(b), panel (b) shows more even energy across different angles. Gathers are taken at every 10th common image point from $x = 4$ km to $x = 8$ km. [CR]	167
C.7	Initial velocity model (a) and the angle gathers (b) obtained using the initial velocity model. The initial ϵ model is shown in Figure C.5(a). Model error causes significant curvatures in the angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [CR]	169
C.8	Inverted velocity model (a) in m/s and ϵ model (b) after 40 iterations. Angle gathers (c) obtained by the inverted model. Angle gathers are extracted at the same CIP as those in Figure C.7(b). Improved velocity and ϵ flattens the corresponding angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [CR]	170

Chapter 1

Introduction

Seismic anisotropy in complex earth subsurface has become increasingly important in seismic imaging due to the increasing offset and azimuth in modern seismic data. To account for anisotropic wave propagation, sophisticated wave extrapolation and imaging techniques have been developed for subsurface areas that are as complex as orthorhombic media. However, building an accurate and reliable anisotropic earth model remains a challenge to the seismic industry due to its nonlinear and underdetermined nature. This thesis presents an automated, wave-equation-based anisotropic model building methodology that integrates surface seismic data, geological information, and prior rock physics knowledge of the subsurface.

EARTH MODEL BUILDING

Subsurface models, including both the background wave-speed model and the structural image, are essential for interpretation, drilling and reservoir modeling. In this thesis, I aim to build a reliable wave speed model that enables the creation of an accurate structural image of the subsurface.

Isotropy or anisotropy

In the seismic-exploration industry, seismic wave speed models are widely needed to explain the travel time of the recorded data. Seismic waves travel in heterogeneous media at different speeds in different directions. This directional dependence of the seismic wave speed is usually referred to as *seismic anisotropy*. Although seismic anisotropy was reported in exploration seismology in the 1930s (McCollum and Snell, 1932; Postma, 1955; Helbig, 1956), it did not play a significant role in seismic imaging until the 1980s. However, as seismic exploration develops to more complex geology and deeper targets, the necessity of including seismic anisotropy has been recognized. Modern seismic acquisition is performed with long offsets (sometimes with more than 10 km) and full azimuthal coverage. Consequently, the data collected recently are highly sensitive to anisotropy and cannot be explained by simple isotropic models.

To account for the kinematic effects of anisotropy, various wave extrapolation schemes have been developed for subsurface models that are as complex as orthorhombic media (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008). However, the most common anisotropic model in seismic imaging and model building is still the transverse isotropic (TI) model. The TI model is the simplest anisotropy model, and it is a reasonable approximation for many sedimentary environments in the subsurface.

Figure 1.1 shows two different types of TI models: the vertical transverse isotropy (VTI) and the tilted transverse isotropy (TTI). For P-wave, the wave speeds in the transverse plane are the same, but the wave speed in the perpendicular direction is different. When the symmetry axis is vertical, the vertical velocity v_v , horizontal velocity v_h , and normal moveout (NMO) velocity v_n , are all needed to define the P-wave wavefront in a homogeneous VTI medium (Tsvankin and Thomsen, 1994; Alkhalifah and Tsvankin, 1995). When the symmetric axis is tilted, the dip angle θ and the azimuth angle α , are also needed to determine the symmetry axis. In my thesis study, I focus on the VTI model because it is the most needed anisotropic model for seismic imaging, and it is sufficient to characterize most of the shale-rich

sedimentary basins in the Gulf of Mexico.

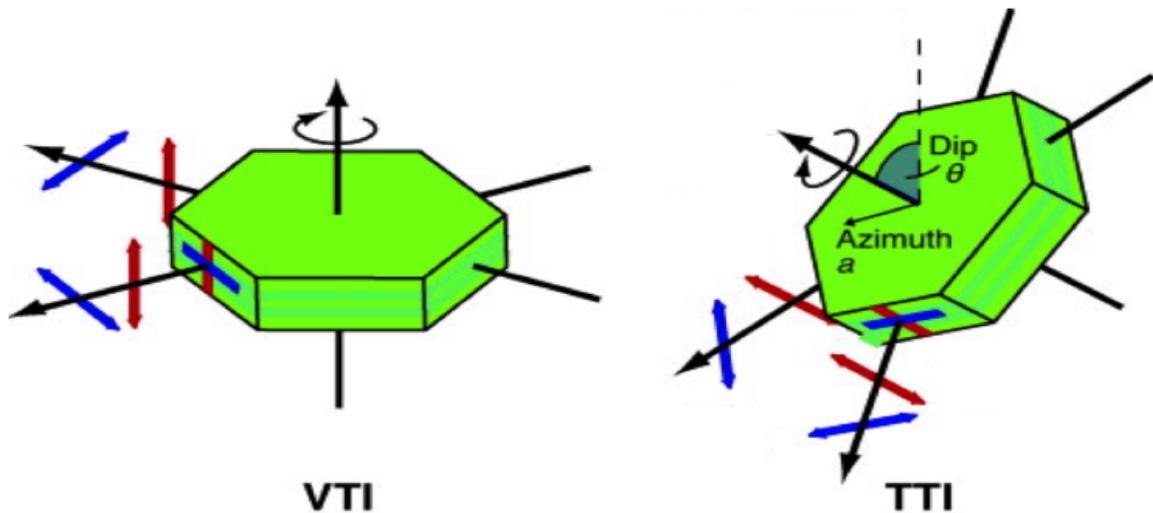


Figure 1.1: Left: A sketch of the wave-speed symmetry for VTI medium. Right: A sketch of the wave-speed symmetry for TTI medium. [NR] chap1/. timedia

Thomsen (1986) introduced dimensionless parameters that relate the three velocities in a TI medium as follows:

$$v_h^2 = v_v^2(1 + 2\epsilon), \quad (1.1)$$

$$v_n^2 = v_v^2(1 + 2\delta). \quad (1.2)$$

Parameters ϵ and δ define the degree of anisotropy of a medium. The medium is isotropic when $\epsilon = \delta = 0$.

Figure 1.2 shows the comparison of wavefronts propagated from a point source in both isotropic and VTI media. The travelttime from the source to the wavefront is 1 sec. The solid line shows the (semi-)circular wavefront in an isotropic medium; the dashed line shows the (semi-)elliptical wavefront in a VTI medium with $\epsilon = \delta = 0.2$; and the dotted line shows the wavefront in a VTI medium with $\epsilon = 0.2$ and $\delta = -0.2$. The positive ϵ leads to the advanced wavefront in the horizontal direction compared with the isotropic wavefront. The positive or negative δ determines the faster or

slower velocity in the near-vertical direction compared with the vertical velocity.

The differences between both the VTI wavefronts and the isotropic wavefront are small when the propagation direction is within 20° around the vertical. The differences between the $\epsilon = \delta = 0.2$ wavefront and the isotropic wavefront become more apparent as the propagation direction deviates from vertical. Tradeoff effects can be seen from the dotted wavefront when $\epsilon = 0.2$ and $\delta = -0.2$. The positive ϵ effect has been compensated by the negative δ effect. Because the vertical velocity is the same, the differences between the dotted wavefront and the isotropic wavefront are rather small when the propagation direction is within 60° around the vertical. The differences only become significant when the propagation angle is greater than 60° . Therefore, presence of dipping reflectors and/or long offset acquisition will magnify the anisotropic effects.

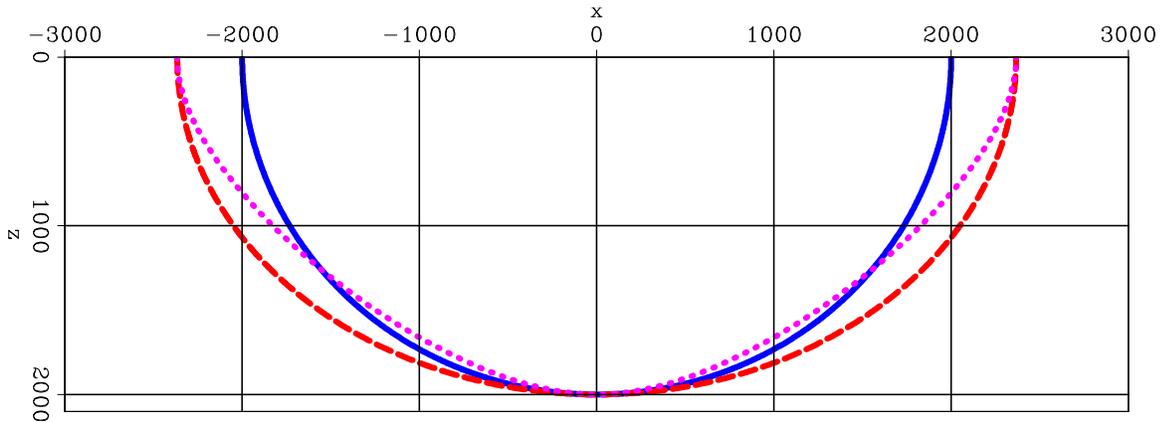


Figure 1.2: Comparison of wavefronts propagated from a point source at $x = 0$, $z = 0$ in homogeneous isotropic and VTI media. Parameters used to propagate the wavefronts are: $\epsilon = 0$ and $\delta = 0$ for the solid line; $\epsilon = 0.2$ and $\delta = 0.2$ for the dashed line; and $\epsilon = 0.2$ and $\delta = -0.2$ for the dotted line. The vertical velocity of the medium is 2000m/s in all three cases. [ER] [chap1/. vtiwavefronts](#)

Alkhalifah and Tsvankin (1995) introduced the anellipticity parameter η which directly relates v_h with v_n as follows:

$$v_h^2 = v_n^2(1 + 2\eta), \quad (1.3)$$

with

$$\eta = \frac{\epsilon - \delta}{1 + 2\delta}. \quad (1.4)$$

The parameter η describes the anellipticity of the wavefront. When $\eta = 0$, the wavefront is an ellipsoid, and the medium is elliptical anisotropic. Alkhalifah and Tsvankin (1995) showed that v_n and η are the only two parameters that determine the long-spread (nonhyperbolic) P-wave moveout assuming horizontal reflectors.

Now we have six parameters to describe the VTI media, but only three are independent. In this thesis, I refer to any three-parameter combination which fully describes a VTI medium as a *VTI model*, and its components as the *VTI parameters*. I refer to the dimensionless parameters ϵ , δ , and η as the *anisotropic parameters*. Based on the available information, only a subset of the three VTI parameters can be reliably resolved.

Ray-based or wave-equation-based tomography

The state of the art in anisotropic model building is represented by ray-based tomography (Woodward et al., 2008). Ray-based methods are efficient and adequate in relatively simple geological settings with mild velocity contrasts. However, they often fail to describe the complexity of the wave propagation due to its infinite-frequency approximation. Figure 1.3 shows the comparison between a snapshot of the traced rays and a snapshot of the wavefields. The ray approximation over simplifies the wavefields above the salt body and fails to model the complex transmitted wavefronts through the salt body.

Furthermore, the sensitivity kernels for velocity and anisotropy are frequency dependent (Woodward, 1992). To capture the frequency-dependence of the model sensitivity, we should back project the model updates along the band-limited wavepaths rather than the infinite-bandwidth rays.

To overcome the limitations of ray-based tomography, many wave-equation-based tomography methods (known as wave-equation migration velocity analysis (WEMVA))

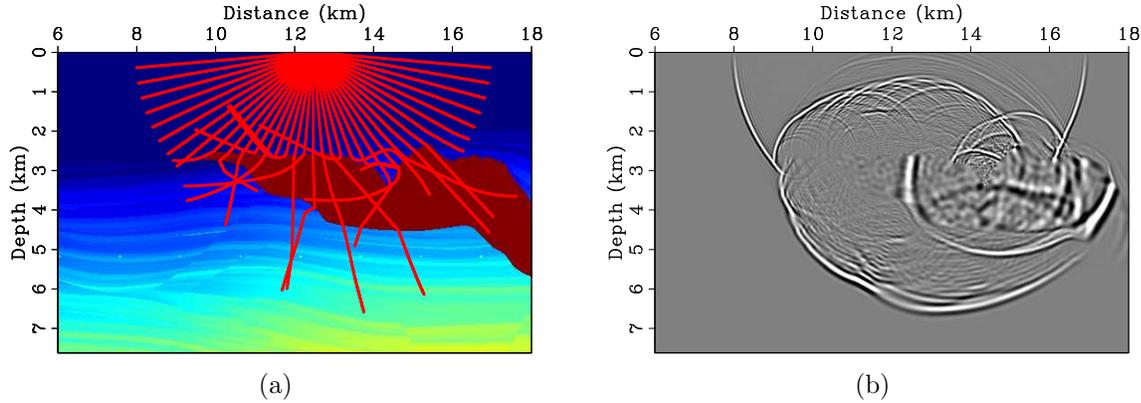


Figure 1.3: Comparison between rays and wavefields. Panel (a) shows a snapshot of the rays overlaid by the Sigsbee2A velocity model, (b) shows a snapshot of the wavefields. [ER] `chap1/. sigsb2a-rays-vmod,sigsb2a-wave-snapshot`

(Biondi and Sava, 1999; Sava and Biondi, 2004a,b; Shen and Symes, 2008) were proposed to use the wavefields as information carriers. These methods work in the prestack image space and utilize the incoherency with reflection angle or data offset in the image gathers to update velocity. Figure 1.4 shows an example of the different moveouts in the angle domain common image gathers (ADCIGs) when different velocity models are used in imaging. The ADCIGs are flat (Panel (b)) when the migration velocity is accurate. When the migration velocity is lower or higher than the true velocity, the ADCIGs show upward or downward residual moveouts, respectively (Panel (a) and (c)). The WEMVA algorithms then translate the residual moveouts into model updates and improve the background velocity model.

We can apply the same criteria to evaluate the accuracy of the anisotropic parameters ϵ and δ . Figures 1.5 and 1.6 show the ADCIGs when different anisotropic parameters are used for imaging (while the other VTI parameters are accurate). We observe similar moveout patterns for the anisotropic parameters as for velocity. However, the absolute moveouts (depth differences from zero angle) with respect to 50% perturbation in the anisotropic parameters are much smaller than that with respect to 5% perturbation in the vertical velocity (Figure 1.4). This suggests the WEMVA objectives have much lower sensitivities for anisotropic parameters than for velocity.

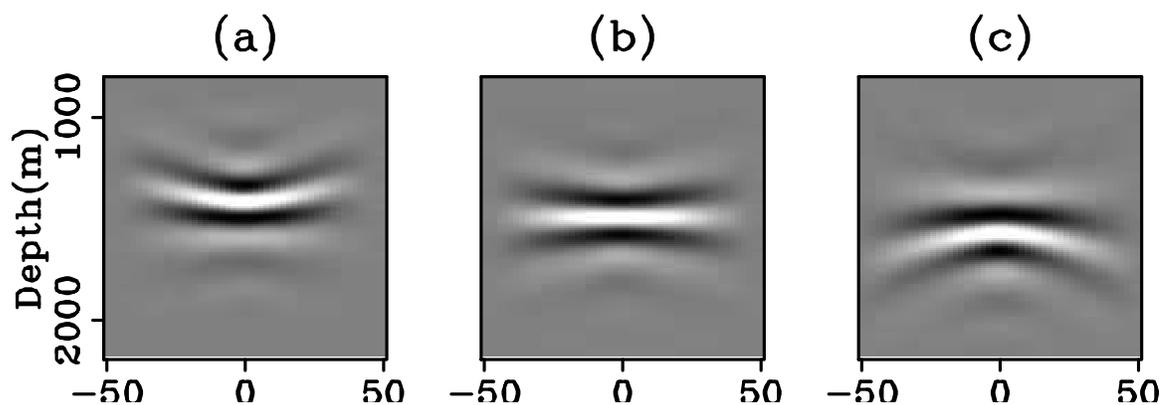


Figure 1.4: Angle domain common image gathers with (a) 5% slower velocity, (b) accurate velocity, and (c) 5% faster velocity. [CR] [chap1/. aimg-vel](#)

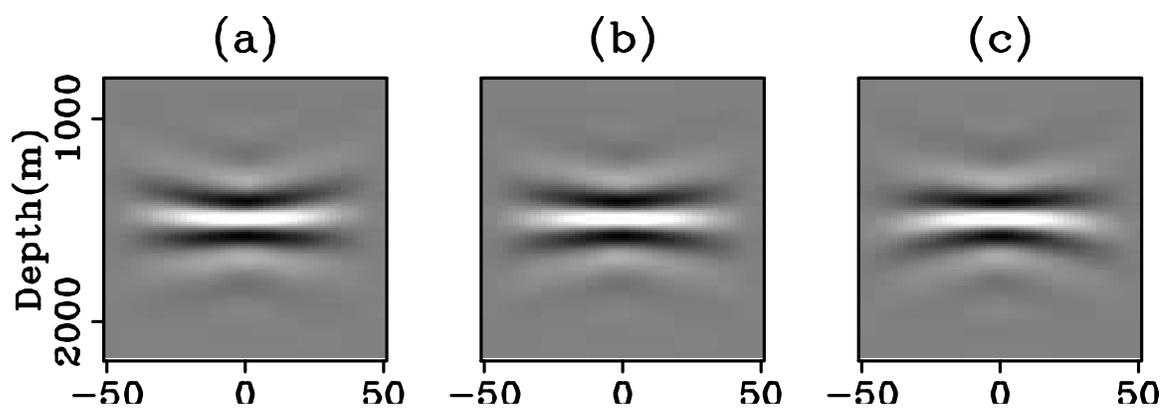


Figure 1.5: Angle domain common image gathers with (a) 50% smaller ϵ , (b) accurate ϵ , and (c) 50% larger ϵ . [CR] [chap1/. aimg-eps](#)

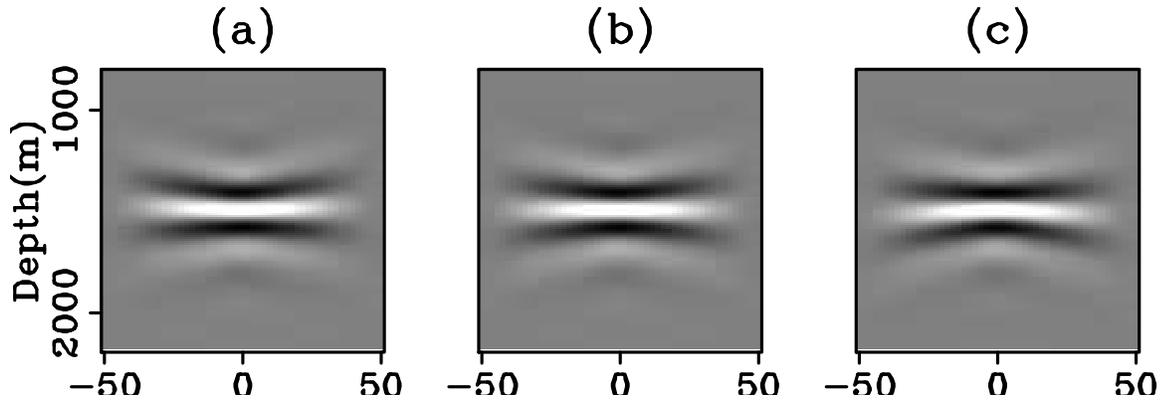


Figure 1.6: Angle domain common image gathers with (a) 50% smaller δ , (b) accurate δ , and (c) 50% larger δ . [CR] `chap1/.aimg-del`

Data integration

Earth model building is an underdetermined and hence challenging inverse problem, especially at the exploration stage. In general, the input information can be obtained from geological information, surface seismic data, and rock physics modeling from well logs. During early exploration, the geological information often comes from plate or regional tectonics to identify the potential target area. Surface seismic data are acquired to image the structure of the subsurface after establishing the area of interest. Based on the preliminary interpretation results, a few exploration wells will be drilled to verify the earth model. The initial model building results are often far from the true subsurface. Therefore, corrections of the earth model by integrating all available information is necessary.

Surface seismic data have the best compromise between accuracy and coverage among the three types of information. Geological knowledge covers large regions without accurate positions, whereas well logs provide accurate, high resolution information only at sparse locations. Therefore, most of the current practices of information integration occur after seismic imaging and structural interpretation. First, seismic images are stretched vertically according to the well markers. Then, borehole core analysis is propagated from the well location to the rest of the region based on

the seismic images and the underlying geological assumptions. However, this conventional workflow does not include a feedback loop to verify if the modified seismic images honor the original seismic data. Therefore, inconsistencies may be introduced by the sequential evaluations of the data.

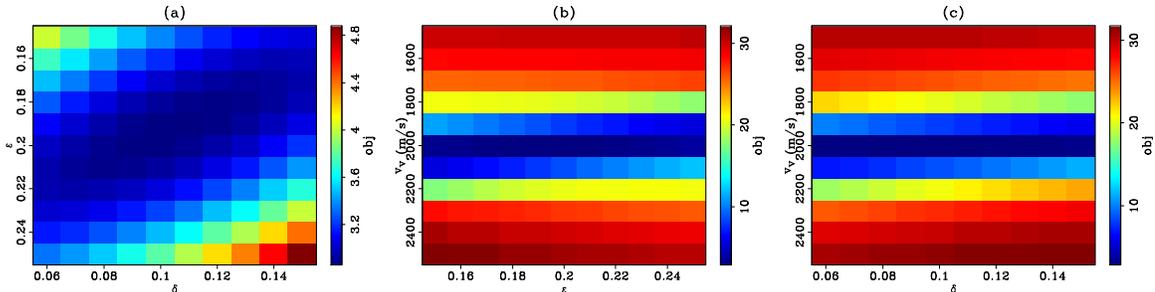


Figure 1.7: WEMVA objective function with respect to (a) ϵ and δ , (b) vertical velocity and ϵ , and (c) vertical velocity and δ . The third parameter is accurate on each panel. Notice the different color scale in panel (a). [CR] chap1/. objfun1

Simultaneous integration of different data types becomes essential when the earth model becomes more complex. With the increasing number of unknowns at each subsurface location, the model building problem based on a single type of data becomes more non-linear and underdetermined.

For example, Figure 1.7 plots the WEMVA objective function based on the surface seismic data with respect to a range of the model parameters. This objective function evaluates the coherence in the ADCIGs when different model parameters are used to migrate the surface seismic data (such as in Figures 1.4, 1.5, and 1.6). The WEMVA objective function can resolve the vertical velocity accurately regardless of the accuracy of the anisotropic parameters. However, it cannot correctly resolve the anisotropic parameters when the vertical velocity is inaccurate. Even when the vertical velocity is accurate, the nearly flat objective function in the ϵ - δ space (Figure 1.7(a)) suggests that the WEMVA objective function cannot sufficiently constrain the anisotropic parameters.

Moreover, severe tradeoffs among the VTI parameters are shown in the objective function plots. A strong tradeoff can be seen in Figure 1.7(a). The WEMVA

objective function cannot resolve ϵ or δ independently as long as the summation of the two remains the same (unless events propagating at more than 60° angles are recorded), as has been observed from the wavefronts in Figure 1.2. Tradeoffs between the anisotropic parameters and the vertical velocity can also be seen from Figures 1.7(b) and 1.7(c), where positive vertical velocity errors are compensated by negative errors in the anisotropic parameters and vice versa. Nevertheless, these tradeoff effects are much less obvious, further demonstrating the low sensitivity of seismic data to the anisotropic parameters.

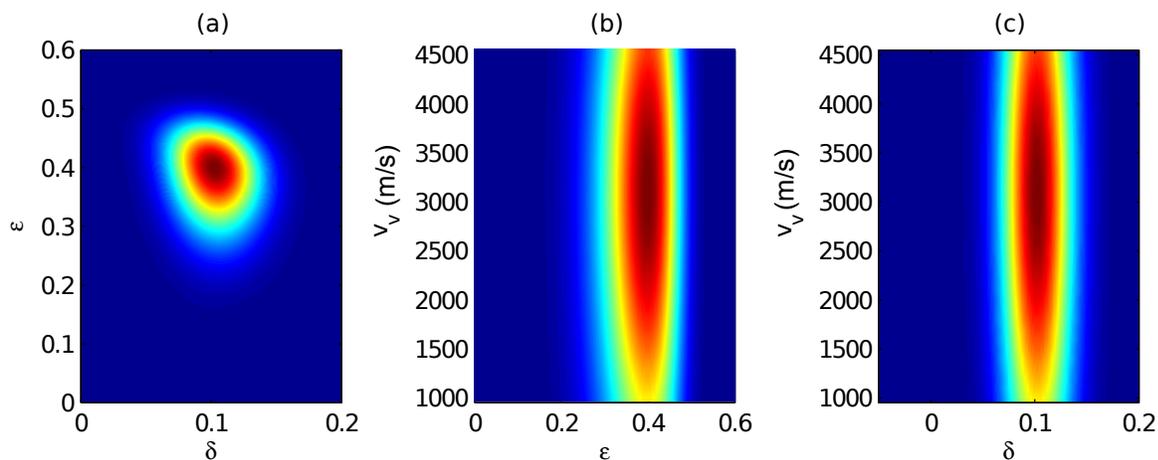


Figure 1.8: Stochastic rock physics modeling results for parameters v_v , ϵ and δ . [ER] chap1/. rockphy

On the other hand, our prior knowledge suggests that the VTI parameters are correlated with each other. Locally, the lithology determines the correlation/covariance among the VTI parameters. These correlations can be measured by stochastic rock physics modeling using the well logs as input (Bachrach et al., 2011). Figure 1.8 shows the rock physics objective function based on a compacting shale model. Compared to the WEMVA objective function (Figure 1.7), the rock physics objective function has much looser control of velocity, suggesting similar lithology can result in a wide range of seismic velocities. In contrast, the rock physics objective function defines a much tighter possible range for anisotropic parameters ϵ and δ . These lithological-based distributions can be summarized as a cross-parameter covariance.

Spatially, the correlations of the parameters are defined by the geological model in the sedimentary basin. In a compaction-based model, the wave speeds vary smoothly parallel to the sea bottom (e.g., Slotnick (1936); Kaufman (1953)). In a lithology-based model, the wave speeds follow the structural dip (Clapp, 2000; Woodward et al., 2008). The anisotropic parameters, being the ratio between the wave speeds, are often considered spatially smoother than the wave speeds themselves. These spatial correlations of each parameter can be summarized as the spatial covariance based on the underlying geological models.

Surface seismic data, rock physics modeling from the well logs, and the geological models each define the VTI parameters from different perspectives and provide complementary information for each other. Therefore, the focus of this thesis is to define effective methods which intelligently combine all information to produce a coherent and consistent model is the focus of this thesis.

OVERVIEW OF THESIS

The goal of this thesis is to integrate different sources of information about the subsurface at different scales in order to produce a reliable anisotropic subsurface model that is consistent with all the available information.

WEMVA for VTI models

In Chapter 2, I extend the concept of wave-equation migration velocity analysis (WEMVA) to anisotropic media. I define the anisotropic WEMVA objective function in the image space by evaluating the flatness of the angle domain common image gathers and the focusing (stacking power) of the stacked image. I derive the gradients of the WEMVA objective function with respect to the VTI parameters and reparameterize the inversion using NMO velocity v_n and the anellipticity parameter η due to the lack of complementary information. I introduce a regularization term to include the geological information into the inversion.

Using both a synthetic and a field data example, I show how anisotropic WEMVA can identify localized anomalies in both NMO velocity and the η model. On the synthetic data, I show that anisotropic WEMVA improves the flatness in the angle domain common image gathers and corrects for the positioning error due to the velocity and η anomalies. On the field data, I show that anisotropic WEMVA improves the resolution of the stacked image and enhances the definition of a major fault in the 2D section.

Anisotropic WEMVA with rock physics constraints

In Chapter 3, I build the framework to regularize the wave-equation tomography that utilizes the available geological and rock physics information. I use the geological information to approximate the spatial covariance for each individual parameter and the rock physics information to approximate the multi-parameter covariance at a particular location. From a simple missing data problem, I demonstrate the extra information the cross-covariance (cross-correlation) among the VTI parameters contributes to the inversion.

On a synthetic example, I compare and contrast the inversion behaviors of the anisotropic WEMVA when three different rock physics covariance matrices are applied. I show that the additional rock physics information will not change the solution of the inversion problem for the well-constrained parameter. Nevertheless, additional rock physics information improves the convergence and the resolution of the well-constrained parameter during the early iterations. By comparing and contrasting the inversion results, I show that the rock physics information is essential to resolve the less-constrained parameters. I show that, the anisotropic WEMVA successfully improves the focusing and the positioning of the migration image in all three inversion tests.

Field data tests

In Chapter 4, I apply the anisotropic WEMVA with geological and rock physics constraints to a 3D dataset from the Gulf of Mexico. I construct the rock physics covariance at every grid point in the model space based on the well logs and the previously interpreted seismic results. Based on the rock physics modeling results, I show that the general trends of velocity-anisotropy correlation are positive in the shallow region, but become negative when the mineral transition of the shales become the dominant factor for both velocity and anisotropy.

I invert the seismic data acquired by an ocean-bottom-system using the proposed anisotropic WEMVA methodology. The inversion results improve both focusing and resolution of the migrated images.

Appendices

In Appendix A, I describe the optimization-based implicit finite differencing scheme I use to extrapolate the wavefield. The optimized coefficients and their derivatives are further used when differentiating the wave-equation with respect to the anisotropic parameters.

In Appendix B, I provide the details for the rock physics models I use to model the compaction-induced shale anisotropy and thin-layering-induced anisotropy in Chapter 4. I show how the previous interpretation results can be used to construct the rock physics covariance matrix in space.

In Appendix C, I provide a synthetic example of the anisotropic WEMVA based on reverse time migration (RTM). RTM handles the large-angle propagation better than one-way wave-equation migration (WEM). However, I keep the one-way WEM implementation in my thesis study to maintain a low computational cost and thus to be able to test the method on 3D dataset, given the computational resource available for this project.

ASSUMPTIONS AND LIMITATIONS

Model assumption is one of the weaknesses of the method presented in this thesis. Although the VTI model can sufficiently describe many geological scenarios, it may not adequately explain formations with steep dips or layered media with one or multiple sets of fractures. More complex models (TTI or orthorhombic model) with more parameters are needed to fully characterize the P-wave propagation. In those cases, the concept of constraining the model building process using the rock physics covariance is even more valuable because the inversion will be more underdetermined with stronger ambiguities among the parameters.

The second assumption comes from the rock physics models. Although I capture the uncertainties in the rock physics modeling through a stochastic modeling process, the additional information I introduce into the inversion is as accurate as the rock physics modeling. The geological and lithological scenarios are limited by the ranges of the rock physics models. However, when in doubt, we can reduce the strength of the rock physics constraints by changing the value of the trade-off parameter for the regularization.

Finally, the resolution and accuracy of the inversion presented in this thesis is limited by the fact that we ignore the presence of shear waves. Although the signal-to-noise ratio of shear wave is usually lower than that of pressure wave, shear waves are more sensitive to the anisotropic properties of the subsurface. Therefore, jointly inverting pressure and shear waves with proper constraints may produce better solutions to the anisotropic model building problem.

Chapter 2

WEMVA for VTI models

Anisotropic models are needed for wave simulation and inversion where a complex geological environment exists. In this chapter, I extend the theory of wave equation migration velocity analysis (WEMVA) to vertical transverse isotropic (VTI) media. The accuracy of the VTI model is mapped to the focusing of the prestack image in the subsurface offset domain, which is then evaluated by a combined differential semblance optimization objective function and a stacking power maximization objective function. Because of the lack of constraints, I assume δ can be accurately obtained from other sources of information, and invert for the NMO slowness and the anellipticity parameter η .

To regularize the multi-parameter inversion, I build a regularization framework to adapt the geological information to guide the updates in both NMO slowness and η . This regularization step is crucial to stabilize the inversion and to produce geologically meaningful results. I test the proposed approach on both a synthetic dataset and a 2-D Gulf of Mexico dataset, in both cases started with fairly good initial VTI models. The inversion results reveal shallow anomalies collocated in NMO velocity and η and improve both the continuity and the resolution of the final stacked images.

INTRODUCTION

When compared with isotropic models, anisotropic models are recognized as more realistic representations of the subsurface in complex geological environment (McCollum and Snell, 1932; Postma, 1955; Helbig, 1956). Many authors (Shan, 2009; Fletcher et al., 2009; Zhang and Zhang, 2009; Fei and Liner, 2008) have developed migration and processing schemes for vertical transverse isotropic (VTI) and tilted transverse isotropic (TTI) media; however, the challenge of building a reliable anisotropic Earth model remains a bottleneck for the exploration workflow.

In my thesis study, I am interested in retrieving the kinematics of the pressure waves in acoustic media. Surface seismic data inversion for anisotropy is ill-posed and highly underdetermined due to the rapidly increasing size of the model space with the increasing complexity of the subsurface (Woodward et al., 2008; Bakulin et al., 2010b,a). To reduce the size of the model space, Sarkar and Tsvankin (2003, 2004) and Takanashi and Tsvankin (2012) parametrize their model space using factorized blocks or layers, where the anisotropic parameters ϵ and δ are constant and the velocity variations are simple linear or quadratic functions of the space coordinates. These factorization methods dramatically decrease the number of unknowns to be solved by the inversion, and hence stabilize the inversion. However, these models tend to over simplify the Earth where heterogeneities may occur at all scales.

In principle, grid-based tomographic methods handle both anisotropy and heterogeneity at the same time. Most of the existing grid-based anisotropic model-building schemes are based on ray-based tomography by measuring the non-hyperbolic moveout to flatten the common image gathers (CIG). The residual moveout is then back-projected along the ray paths to update the anisotropic models (Zhou et al., 2003, 2004; Yuan et al., 2006; Cai et al., 2009; Woodward et al., 2008). However, ray-based methods are prone to inaccurate results when multi-pathing exists in areas of complex overburden. Furthermore, due to the ill-posed and underdetermined nature of the anisotropic tomography, small amounts of noise in the residual moveout on a CIG gather can lead to significant variations in the inverted model. Therefore, I propose

a wave-equation-based tomography method to build the anisotropic model.

In this chapter, I first generalize the methodology of image-space WEMVA (Biondi and Sava, 1999; Sava and Biondi, 2004a,b) from an isotropic medium to an anisotropic medium. I use grid-based representations of velocity, ϵ and δ to parametrize the VTI medium. I show that theoretically the gradient of the tomographic objective functional for an isotropic medium can be modified to describe an anisotropic medium by simply adding corresponding terms for the additional parameters. Due to the lack of constraints, I reparameterize the subsurface using NMO velocity, η and δ , but keep δ fixed during the iterative WEMVA updates. Moreover, I precondition the model using geological information. Finally, I demonstrate on a synthetic dataset and a field dataset that the proposed method can identify anomalies in the velocity and the anisotropic model, and hence produce better migrated images with better spatial continuity and higher resolution.

WAVE EQUATION MIGRATION VELOCITY ANALYSIS FOR ANISOTROPY

In this section, I extend the methodology of WEMVA to VTI media. I provide a concise derivation of the WEMVA gradients using the Lagrangian augmented functional. The derivation using perturbation theory can be found in Appendix A. The interpretations of the adjoint-state equations suggest the same implementation as perturbation theory suggests.

Governing equations

Assuming that the shear wave velocity is much smaller than the pressure wave velocity, one approximation of the up-going and down-going one-way acoustic wave equations for VTI media is shown as follows (Shan, 2009):

$$\left(\frac{\partial}{\partial z} \mp i\Lambda \right) P = 0, \quad (2.1)$$

where $P = P(x, y, z, \omega)$ is the wavefield in the space-frequency domain and Λ describes the dispersion relationship in terms of P-wave vertical velocity v_v and Thomsen parameters ϵ and δ (Thomsen, 1986):

$$\Lambda = \frac{\omega}{v_v} \sqrt{\frac{\frac{\omega^2}{v_v^2} - (1 + 2\epsilon)|\mathbf{k}|^2}{\frac{\omega^2}{v_v^2} - 2(\epsilon - \delta)|\mathbf{k}|^2}}, \quad (2.2)$$

where $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector.

When implementing the wave-equation, I adapt the optimized implicit finite difference scheme (Shan, 2009) to approximate the dispersion relation in Equation 2.2 using a rational function:

$$\Lambda = \frac{\omega}{v_v} \left(1 - \sum_{i=1}^n \frac{\alpha_i |\mathbf{k}|^2}{\frac{\omega^2}{v_v^2} - \beta_i |\mathbf{k}|^2} \right), \quad (2.3)$$

where the coefficients α_i and β_i are the optimized coefficients for particular ϵ and δ values. Prior to wave-equation simulation, I build tables of coefficients for the whole range of possible ϵ and δ pairs. During the propagation, I look up the optimized implicit finite difference coefficients according to the local ϵ and δ values. Details about the optimized implicit finite differencing scheme is described in Appendix A.

In shot-profile migration, both source wavefields \mathbf{D} and receiver wavefields \mathbf{U} are downward continued using the following one-way wave equations and boundary conditions:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\Lambda \right) D(\mathbf{x}, \mathbf{x}_s) = 0 \\ D(x, y, z = 0, \mathbf{x}_s) = f_s \delta(\mathbf{x} - \mathbf{x}_s) \end{cases}, \quad (2.4)$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\Lambda \right) U(\mathbf{x}, \mathbf{x}_s) = 0 \\ U(x, y, z = 0, \mathbf{x}_s) = d_{\mathbf{x}_s} \delta(\mathbf{x} - \mathbf{x}_r) \end{cases}, \quad (2.5)$$

where f_s is the source wavelet, $d_{\mathbf{x}_s}$ is the shot record by a shot at $\mathbf{x} = \mathbf{x}_s$. To

differentiate from the anisotropic parameter δ , I use $\boldsymbol{\delta}$ to denote the Dirac delta. Equations 2.4 and 2.5 can be summarized in matrix forms as follows:

$$\mathbf{L}\mathbf{D} = \mathbf{f} \quad (2.6)$$

and

$$\mathbf{L}^*\mathbf{U} = \mathbf{d}, \quad (2.7)$$

where $\mathbf{L} = \frac{\partial}{\partial z} - i\Lambda$ and $*$ denotes the adjoint.

I use an extended imaging condition (Sava and Formel, 2006) to compute the image cube with subsurface offsets:

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{D})^*(\mathbf{S}_{-\mathbf{h}}\mathbf{U}), \quad (2.8)$$

where $\mathbf{S}_{+\mathbf{h}}$ is a shifting operator which shifts the wavefield $+\mathbf{h}$ in the \mathbf{x} direction. Notice that $(\mathbf{S}_{+\mathbf{h}})^* = \mathbf{S}_{-\mathbf{h}}$. Equations 2.6, 2.7 and 2.8 are the state equations, and \mathbf{D} , \mathbf{U} and $I_{\mathbf{h}}$ are the state variables.

WEMVA gradients for VTI parameters

To evaluate the accuracy of the subsurface model, I use a combined differential semblance optimization (DSO) (Symes and Carazzone, 1991; Shen, 2004) and stacking power objective function (Soubaras and Gratacos, 2007):

$$J = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \langle I_{\mathbf{h}}\boldsymbol{\delta}(\mathbf{h}), I_{\mathbf{h}}\boldsymbol{\delta}(\mathbf{h}) \rangle. \quad (2.9)$$

where \mathbf{h} is the subsurface offset. The first term in the objective function J is the DSO term, and the second term denotes the stacking power term. The trade-off parameter α controls the relative strength between both terms. This composite objective function shares the merit of both objectives: the global convergence of DSO and the high resolution of stacking power maximization.

To derive the gradient of the WEMVA objective function with respect to the VTI parameters, I follow the adjoint-state recipe provided by Plessix (2006). First, I form the Lagrangian augmented functional:

$$\begin{aligned}
\mathcal{L}(\mathbf{D}, \mathbf{U}, I_{\mathbf{h}}; \lambda, \mu, \gamma_{\mathbf{h}}; v_v, \epsilon, \delta) = & \\
& \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h} I_{\mathbf{h}}, \mathbf{h} I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \langle I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}), I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}) \rangle \\
& + \langle \lambda, \mathbf{f} - \mathbf{L}(v_v, \epsilon, \delta) \mathbf{D} \rangle \\
& + \langle \mu, \mathbf{d} - \mathbf{L}^*(v_v, \epsilon, \delta) \mathbf{U} \rangle \\
& + \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}}, (\mathbf{S}_{+\mathbf{h}} \mathbf{D})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) - I_{\mathbf{h}} \rangle. \tag{2.10}
\end{aligned}$$

Then the adjoint-state equations are obtained by taking the derivative of \mathcal{L} with respect to the state variables \mathbf{D} , \mathbf{U} and $I_{\mathbf{h}}$:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{D}} = -\mathbf{L}^*(v_v, \epsilon, \delta) \lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^* (\mathbf{S}_{-\mathbf{h}} \mathbf{U}) \gamma_{\mathbf{h}} = \mathbf{0}, \tag{2.11}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{U}} = -\mathbf{L}(v_v, \epsilon, \delta) \mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^* (\mathbf{S}_{+\mathbf{h}} \mathbf{D}) \gamma_{\mathbf{h}} = \mathbf{0}, \tag{2.12}$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} - \alpha I_{\mathbf{h}} \boldsymbol{\delta}(\mathbf{h}) = 0, \forall \mathbf{h}. \tag{2.13}$$

Equation 2.11, 2.12, and 2.13 are the adjoint-state equations. Parameters λ , μ and $\gamma_{\mathbf{h}}$ are the adjoint-state variables, and can be calculated from the adjoint-state equations.

The physical interpretation of the adjoint-state equations offers better understanding of the physical process and provides insights for implementation. Clearly, the solution to Equation 2.13, $\gamma_{\mathbf{h}}$, is the perturbed (residual) image at a certain subsurface offset. Equations 2.11 and 2.12 define the perturbed source and receiver wavefields, respectively. Notice the perturbed source wavefield λ at location \mathbf{x} depends on the image at $(\mathbf{x} - \mathbf{h}, \mathbf{h})$ and the background receiver wavefield \mathbf{U} at $\mathbf{x} - 2\mathbf{h}$. The same rule applies to the perturbed receiver wavefield μ .

With the solutions to the equations above, I can now derive the gradients of the

objective function 2.9 by taking the derivative of the augmented functional \mathcal{L} with respect to the model variables v_v , ϵ and δ as follows:

$$\nabla_{v_v} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \mathbf{v}_v} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \mathbf{v}_v} \mathbf{U} \right\rangle, \quad (2.14)$$

$$\nabla_{\epsilon} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \epsilon} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \epsilon} \mathbf{U} \right\rangle, \quad (2.15)$$

$$\nabla_{\delta} J = \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \delta} \mathbf{D} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \delta} \mathbf{U} \right\rangle. \quad (2.16)$$

Operators $\frac{\partial \mathbf{L}}{\partial v_v}$, $\frac{\partial \mathbf{L}}{\partial \epsilon}$, and $\frac{\partial \mathbf{L}}{\partial \delta}$ and their adjoint operators are computed based on a polynomial approximation of the dispersion relation in Equation 2.2. Details are described in Appendix A.

Note that if the forward modeling operator \mathbf{L} is isotropic, Equation 2.14 reduces to the gradient of the isotropic WEMVA objective function with respect to the isotropic velocity. Therefore, the extension of WEMVA to the anisotropic medium is simply adding related terms to the additional anisotropic parameters.

Figure 2.1 shows the gradients of the objective function when an impulsive image $I_{\mathbf{h}} = \delta(x, y, z - 1500, \mathbf{h})$ is evaluated in a homogeneous VTI medium. The background VTI parameters are: $v_v = 2\text{km/s}$, $\epsilon = 0.1$ and $\delta = 0.05$. These impulse responses are often referred as ‘‘sensitivity kernels’’ (eg. Marquering et al. (1998, 1999); Rickett (2000)). The top row displays the kernels for v_v , ϵ , and δ when a single pair of source and receiver is placed at zero offset. The zero-offset kernel for vertical velocity in panel (a) nicely covers the wavepaths from the source to the image point and from the image point to the receiver. Due to the lack of sensitivity of the vertical rays to the anisotropic parameter ϵ , the zero-offset kernel for ϵ in panel (d) is saturated with the near field artifacts. In panel (g), the zero-offset kernel for δ shows similar pattern as for v_v , except for the reduced sensitivity with depth.

The middle row in Figure 2.1 displays the kernels when a single source-receiver pair is placed at 4000m offset. All three kernels cover the wavepaths from the source to the image point and from the image point to the receiver. Reading from the scale bar in panel (b), a lower sensitivity of the slanted rays to the vertical velocity is observed compared with that of the vertical rays. To the contrary, higher sensitivities for ϵ and δ are observed from panel (e) and panel (h).

To show the relative strength of the sensitivity kernels, the summation of the kernels in the first two rows is plotted in the bottom row in Figure 2.1. It is obvious and natural that the vertical rays have higher sensitivities for vertical velocity. Meaningful sensitivity for ϵ only comes from the slanted traveling waves. Similarly, sensitivity for δ mainly comes from the slanted traveling waves.

Parameterization

Many authors (Tsvankin and Thomsen, 1994; Alkhalifah and Tsvankin, 1995) have shown that surface recorded P-wave traveltimes can be characterized by the NMO velocity, v_n , and the anellipticity parameter η . Based on the definition in Equation 1.2 and 1.4, the one-way wave-equation can be formulated in terms of v_n , η and δ by a change of variables:

$$\left(\frac{1}{\sqrt{1+2\delta}} \frac{\partial}{\partial z} \mp i\Lambda' \right) P = 0 \quad (2.17)$$

where

$$\Lambda' = \frac{\omega}{v_n} \sqrt{1 - \frac{|\mathbf{k}|^2}{\frac{\omega^2}{s_n^2} - 2\eta|\mathbf{k}|^2}}. \quad (2.18)$$

Notice that in this formulation (Equation 2.17), Thomsen parameter δ and the derivative in depth $\frac{\partial}{\partial z}$ are coupled with each other. This is a theoretical proof of the well-observed δ and depth ambiguity (Plessix and Rynja, 2010). To constrain the δ parameter, well log or well marker information is needed to add the depth dimension into the inversion.

Considering the surface seismic data only, I assume δ can be correctly obtained

from other sources of information (such as check shots and well logs) and keep it fixed through the WEMVA iterations. When $\eta = 0$, in the so-called elliptical anisotropic medium, the dispersion relation 2.18 reduces to the isotropic dispersion relation with a vertically stretched NMO slowness.

Instead of re-deriving the gradients using the new formulation of the dispersion relation 2.17, I derive the gradients of the objective function 2.9 with respect to NMO velocity v_n and η using chain rule and the previously derived gradients (Equation 2.14 to 2.16):

$$\begin{aligned}\nabla_{v_n} J &= \left(\frac{\partial J}{\partial v_n} \right)^T = \left(\frac{\partial J}{\partial v_v} \frac{\partial v_v}{\partial v_n} \right)^T + \left(\frac{\partial J}{\partial \delta} \frac{\partial \delta}{\partial v_n} \right)^T \\ &= \left(\frac{\partial v_v}{\partial v_n} \right)^T \nabla_{v_v} J + \left(\frac{\partial \delta}{\partial v_n} \right)^T \nabla_{\delta} J,\end{aligned}\tag{2.19}$$

$$\begin{aligned}\nabla_{\eta} J &= \left(\frac{\partial J}{\partial \eta} \right)^T = \left(\frac{\partial J}{\partial \epsilon} \frac{\partial \epsilon}{\partial \eta} \right)^T + \left(\frac{\partial J}{\partial \delta} \frac{\partial \delta}{\partial \eta} \right)^T \\ &= \left(\frac{\partial \epsilon}{\partial \eta} \right)^T \nabla_{\epsilon} J + \left(\frac{\partial \delta}{\partial \eta} \right)^T \nabla_{\delta} J,\end{aligned}\tag{2.20}$$

where

$$\frac{\partial v_v}{\partial v_n} = \frac{1}{\sqrt{1+2\delta}}; \quad \frac{\partial \delta}{\partial v_n} = \frac{\sqrt{1+2\delta}}{v_v};\tag{2.21}$$

$$\frac{\partial \epsilon}{\partial \eta} = 1+2\delta; \quad \frac{\partial \delta}{\partial \eta} = -\frac{(1+2\delta)^2}{1+2\epsilon}.\tag{2.22}$$

Figure 2.2 shows the same impulse responses as in Figure 2.1, except for a different parametrization: NMO velocity, η and δ . As shown by the first column, the sensitivity for the NMO velocity is very similar to that for the vertical velocity. Compared with Figure 2.1(b), the sensitivity for NMO velocity to the slanted rays is slightly higher; however, the increase in sensitivity is small (less than 2%) due to the small influence of δ on NMO velocity. Impulse response of η has a negative sign, reflecting the fact

that the surface reflection P-wave data have higher sensitivity for δ than for ϵ . The interferences between the ϵ gradients and δ gradients distribute the sensitivity of η more evenly across different propagation angles.

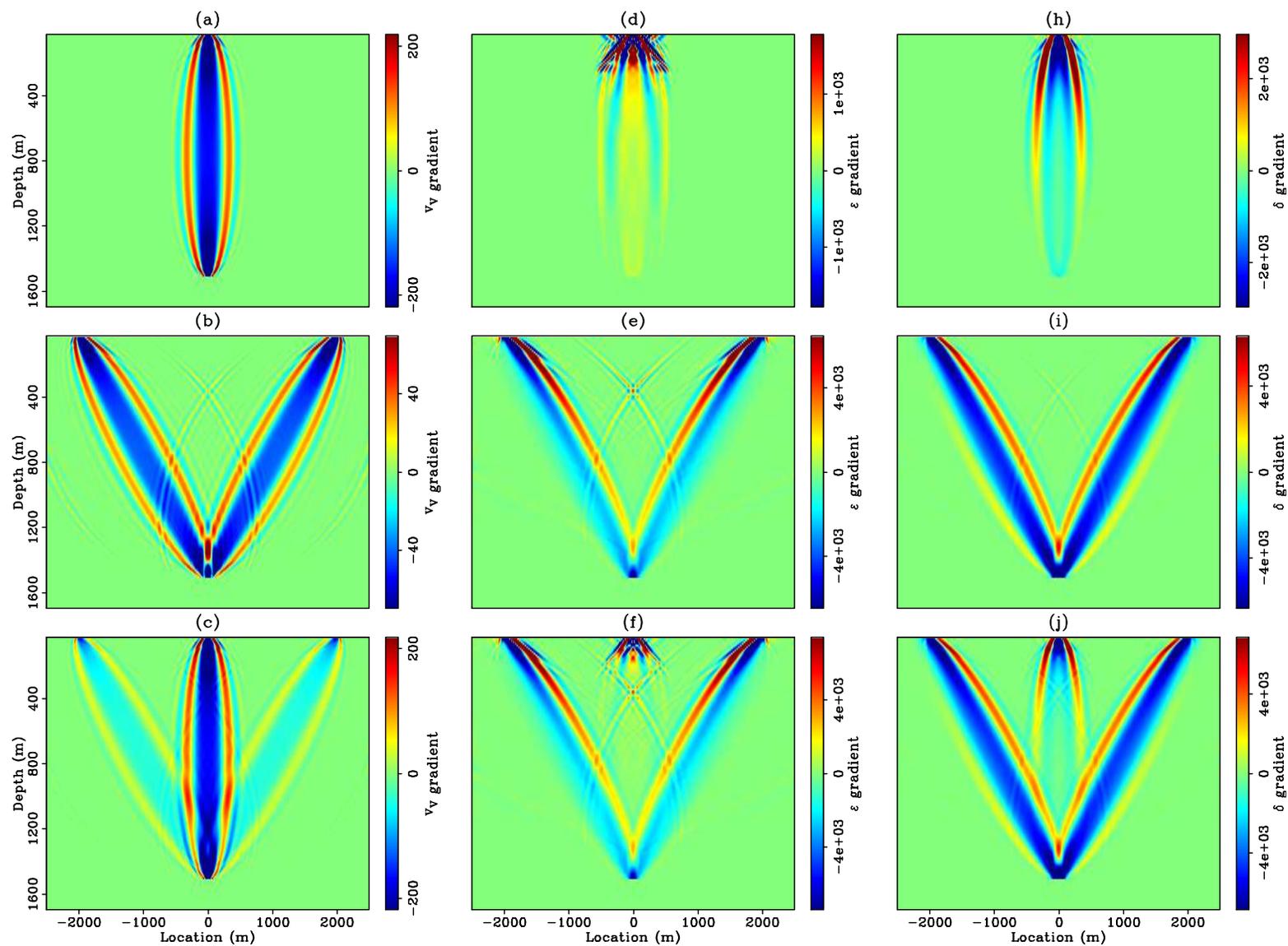


Figure 2.1: 2D impulse responses for vertical velocity (left column), ϵ (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the two rows above. [ER] chap2/. 2dkernelchap2

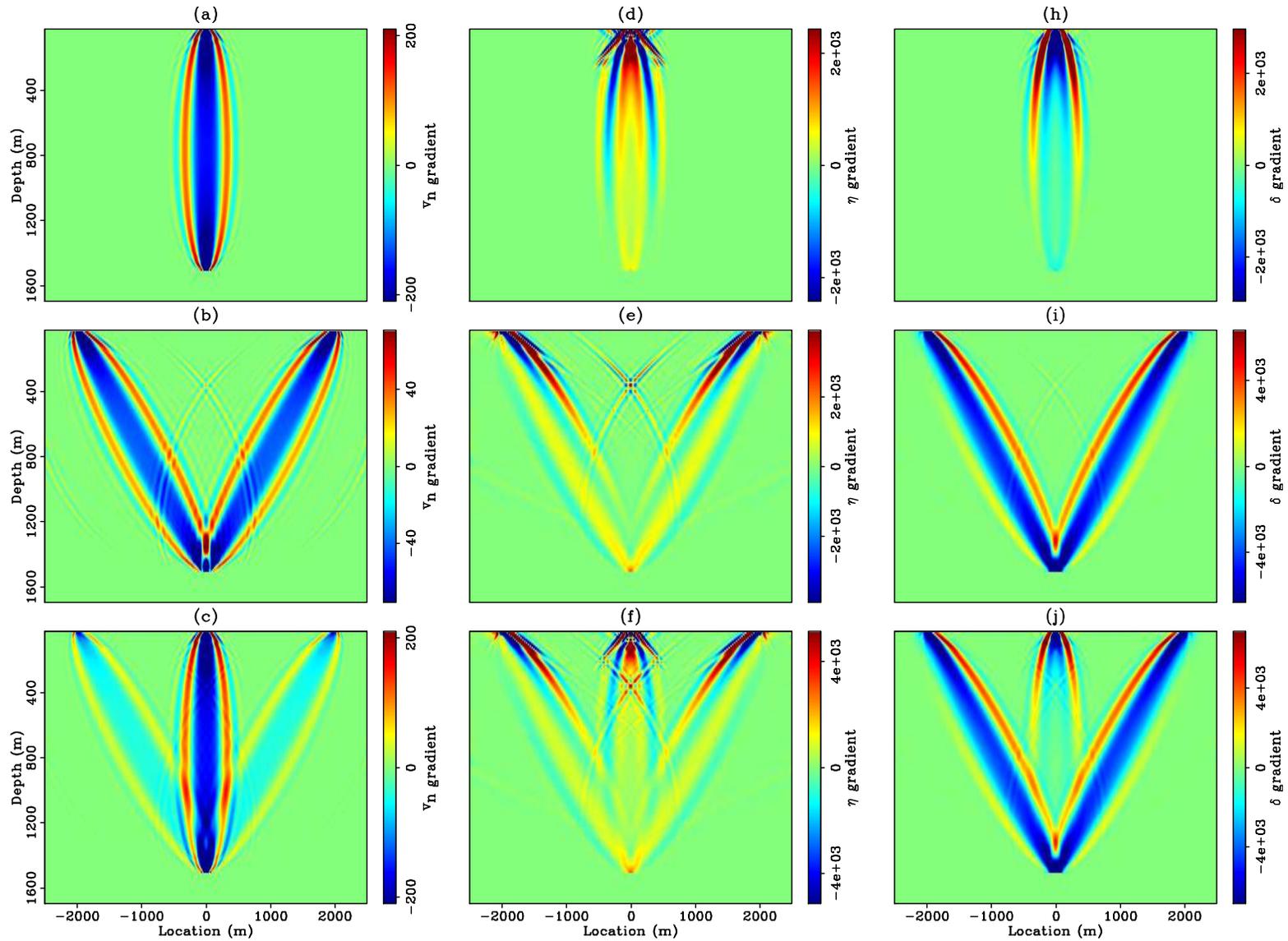


Figure 2.2: 2D impulse responses for NMO velocity (left column), η (middle column), and δ (right column). Top row: zero offset impulse responses. Middle row: impulse responses when source-receiver offset is 4 km. Bottom row: summation of the top two rows. [ER] chap2/. 2dkernv

Preconditioning the WEMVA gradient

Subsurface model building is a highly underdetermined and nonlinear inverse problem. Therefore, prior knowledge of the subsurface is needed to stabilize the inversion and to define a plausible subsurface model. Assuming Gaussian distribution, Tarantola (1984) characterizes the prior information using the mean $\mathbf{m}_{\text{prior}}$ and the covariance \mathbf{C}_M of the model and includes it as a regularization term. The resulting regularized objective function is as follows:

$$J(\mathbf{m}) = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \left\| \sum_{\theta} \mathbf{I}(\mathbf{x}, \theta) \right\|^2 + \beta \frac{1}{2} (\mathbf{m} - \mathbf{m}_{\text{prior}})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{\text{prior}}). \quad (2.23)$$

where the first two terms define the “data fitting” objective, and the third “model regularization” term. The scalar β is to balance the relative strength of the objectives.

A classic regularizer is a roughening operator (Claerbout, 2009), so that the roughness of the model is penalized. However, this regularization formulation suffers from low convergence due to the small length of the roughening operator (Clapp, 2000). To speed up the convergence, a regularization problem is often reformulated to a preconditioning problem by introducing a preconditioning variable (Polak, 1997). After preconditioning, the short roughening operator is replaced by a long smoothing operator with. The convergence is hence improved both by the longer spread of the information, and by the improved condition number of the matrix we are to invert.

If we define the preconditioned variable $\mathbf{n} = \mathbf{C}_M^{-\frac{1}{2}} (\mathbf{m} - \mathbf{m}_{\text{prior}})$, the regularized objective function 2.23 is reformulated as follows:

$$J(\mathbf{n}) = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle - \alpha \frac{1}{2} \left\| \sum_{\theta} \mathbf{I}(\mathbf{x}, \theta) \right\|^2 + \beta \frac{1}{2} \|\mathbf{n}\|^2. \quad (2.24)$$

Setting $\beta = 0$, the preconditioned objective function 2.24 is essentially the same as the original data fitting objective function 2.9, except for the different model parameters. Hence, the gradient of the preconditioned WEMVA objective function (2.24) with

respect to this preconditioned variable \mathbf{n} can be obtained by the chain rule:

$$\begin{aligned}\nabla_{\mathbf{n}}J &= \left(\frac{\partial\mathbf{m}}{\partial\mathbf{n}}\right)^*\nabla_{\mathbf{m}}J \\ &= \mathbf{C}_M^{\frac{1}{2}*}\nabla_{\mathbf{m}}J,\end{aligned}\tag{2.25}$$

where $\nabla_{\mathbf{m}}J = [\nabla_{s_n}J \ \nabla_{\eta}J \ \nabla_{\delta}J]^T$ is the gradient of the objective function 2.9 with respect to the original model parameters.

In the single-parameter (velocity) tomography, the geological dip information can be included into the inversion using the steering filters to spread the velocity updates along the structure (Clapp, 2000). The generalization of the same idea to the multi-parameter anisotropic tomography is fairly straightforward: similar steering filters can also be applied to smooth the anisotropic parameters (Woodward et al., 2008).

The additional character about the multi-parameter anisotropic tomography is that the VTI parameters are correlated with each other. The correlations are mainly determined by the local lithology. Therefore, a cross-parameter covariance is also needed to constrain the anisotropic tomography. In many exploration areas, the lithological information is not available. In this situation, the only information we know is the relative ratio between velocity and the anisotropic parameters. Assuming proportional errors in each parameter, we have a rough estimate of the diagonal elements of the cross-parameter covariance matrix.

In this chapter, I assume there are two component of the covariance matrix: spatial covariance and local cross-parameter covariance. I assume they can be separated independently from each other (Li et al., 2011). A set of precondition operators are used: smoothing filters to approximate square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the model covariance matrix is approximated by the following formulation:

$$\mathbf{C}_M^{\frac{1}{2}} = \mathbf{\Sigma}\mathbf{S}.\tag{2.26}$$

In Equation 2.26, the smoothing operator \mathbf{S} is a band-limited diagonal matrix:

$$\mathbf{S} = \begin{vmatrix} \mathbf{S}_v & 0 & 0 \\ 0 & \mathbf{S}_\eta & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad (2.27)$$

with potentially different smoothing operators for NMO velocity and η , according to the geological information in the study area. The standard deviation matrix Σ is the square-root of the covariance matrix:

$$\Sigma = \begin{vmatrix} \mathbf{C}_{vv} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{C}_{\eta\eta} \mathbf{I} & 0 \\ 0 & 0 & 0 \end{vmatrix}^{\frac{1}{2}}. \quad (2.28)$$

The diagonal elements \mathbf{C}_{vv} and $\mathbf{C}_{\eta\eta}$ denote the variance of the NMO velocity and η , respectively. They are used to balance the relative strength of the velocity and η updates. Notice that both the bottom rows of \mathbf{S} and Σ contain all zeros, which remove updates in the δ model.

To explain the preconditioning scheme, I analyze the preconditioning effect assuming a non-linear steepest decent inversion framework. The initial preconditioned model \mathbf{n}_0 is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \langle \mathbf{m}_0 - \Sigma \mathbf{S} \mathbf{n}_0, \mathbf{m}_0 - \Sigma \mathbf{S} \mathbf{n}_0 \rangle. \quad (2.29)$$

For the i_{th} iteration, the preconditioned variable is obtained by

$$\mathbf{n}_{i+1} = \mathbf{n}_i + \lambda_i \nabla_{\mathbf{n}} J, \quad (2.30)$$

and hence the original model variable is

$$\begin{aligned}
\mathbf{m}_{i+1} &= \mathbf{S}\Sigma\mathbf{n}_{i+1} \\
&= \mathbf{S}\Sigma\mathbf{n}_i + \alpha_i\mathbf{S}\Sigma\nabla_{\mathbf{n}}J \\
&= \mathbf{m}_i + \alpha_i\mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\nabla_{\mathbf{m}}J.
\end{aligned} \tag{2.31}$$

Equation C.44 suggests that preconditioning a non-linear inversion is equivalent to filtering the gradients so that the resulting updates have the desired spectrum. Therefore, instead of explicitly reformulating the preconditioned inversion, we can make use of the original non-linear conjugate gradient algorithm implementation with minimal changes. The preconditioning step is highlighted in red in Algorithm 1.

Algorithm 1 Optimization algorithm

```

initialize the model:  $\mathbf{m}_0$ 
compute the migrated image:  $I_0$ 
compute the gradient:  $\mathbf{g}_0$ 
precondition the gradient:  $\mathbf{g}_s^0 = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_0$ 
initialize the search direction:  $\mathbf{p}_0 = -\mathbf{g}_s^0$ 
for  $k = 1 \dots N_k$  do
  perform a line search: optimize  $\lambda$ ,  $\operatorname{argmin}_{\lambda} J(\mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1})$ 
  update the velocity model:  $\mathbf{m}_k = \mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1}$ 
  compute the migrated image:  $I_k$ 
  compute the gradient:  $\mathbf{g}_k$ 
  precondition the gradient:  $\mathbf{g}_s^k = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_k$ 
  find the search direction:  $\mathbf{p}_k = -\mathbf{g}_k + \frac{(\mathbf{g}_s^k)^T(\mathbf{g}_s^k - \mathbf{g}_s^{k-1})}{(\mathbf{g}_s^{k-1})^T\mathbf{g}_s^{k-1}}$ 
end for

```

SYNTHETIC DATA TEST

I test the proposed anisotropic WEMVA method on a modified BP2007 model.¹ Instead of using the original TTI synthetic data, I model my synthetic data assuming

¹The dataset was created by Hemang Shah and is provided at <http://software.seg.org/> courtesy of BP Exploration Operation Company Limited.

VTI. I also add a high velocity and high anisotropy anomaly in the shallow region of the model. This simulates the geological setting that local diagenetic recrystallization of clay minerals significantly increases both the velocity and the anisotropy.

The initial and the true velocity models are shown in Figure 2.3(a) and 2.3(c). The initial and the true η models are shown in Figure 2.4(a) and 2.4(c). I assume parameter δ is accurately obtained from other studies and use the true δ model in the inversion without updating it. I model the synthetic data using one-way VTI born modeling and perform anisotropic WEMVA for 60 iterations. I construct the preconditioner using the same triangular smoother for \mathbf{S}_v and \mathbf{S}_η , as well as a stationary matrix $\mathbf{\Sigma}$ with zero off-diagonal components to scale the gradients.

The inverted NMO velocity and η model are shown in Figure 2.3(b) and 2.4(b), respectively. Compared with the initial models, the inverted models are closer to the true models with improved definition of the shallow geology. I compare the true anomaly and the inverted anomaly in Figure 2.5. The anisotropic WEMVA inversion successfully identified the location of the anomaly in both NMO velocity and the η model. I increase the trade-off parameter α from 0.5 to 0.8 after 30 iterations, because the stacking power objective function can provide higher resolution gradients once the models are close to the true ones. However, the resolution of the η model is still limited by the illumination angles. Nevertheless, the cumulative contribution of this lower resolution anomaly in η explains the kinematic error in the P-wave propagation.

Figure 2.6 shows the data fitting objective function value (Equation 2.9) as a function of the iteration number. To make the objective function value comparable, I keep $\alpha = 0.5$ for this plot even though α is increased to 0.8 after 30 iterations. The objective function curve flattens out between iteration 20 and 30, suggesting the inversion has converged at the current resolution. After increasing the trade-off parameter α , the objective function further decreases until the inversion converges around 60 iterations.

Figure 2.7 compares the initial migrated image in panel (a) with the inverted

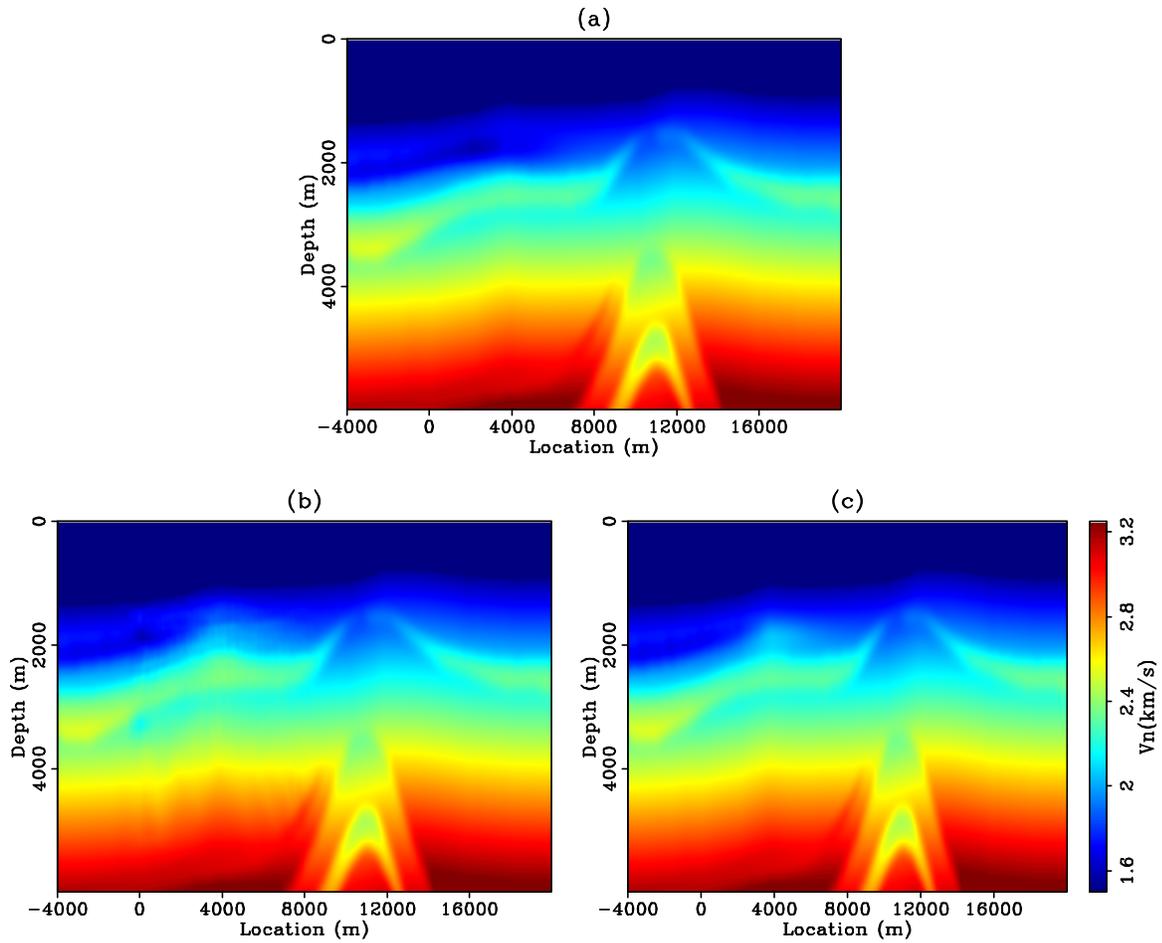


Figure 2.3: Modified BP2007 model. (a): initial NMO velocity; (b): inverted NMO velocity; and (c) true NMO velocity. [CR] chap2/. bp1-vp

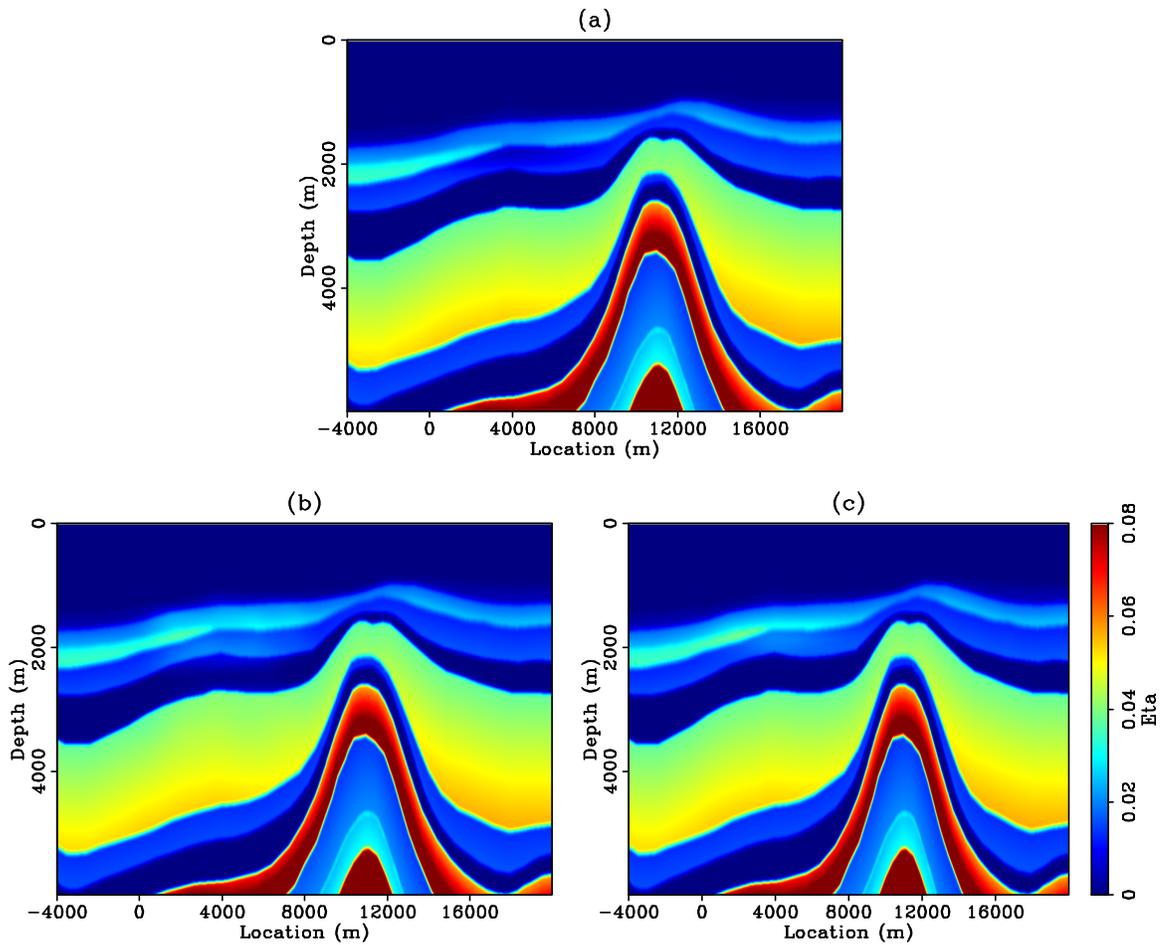


Figure 2.4: Modified BP2007 model. (a): initial η model; (b): inverted η model; and (c) true η model. [CR] chap2/. bp1-eta

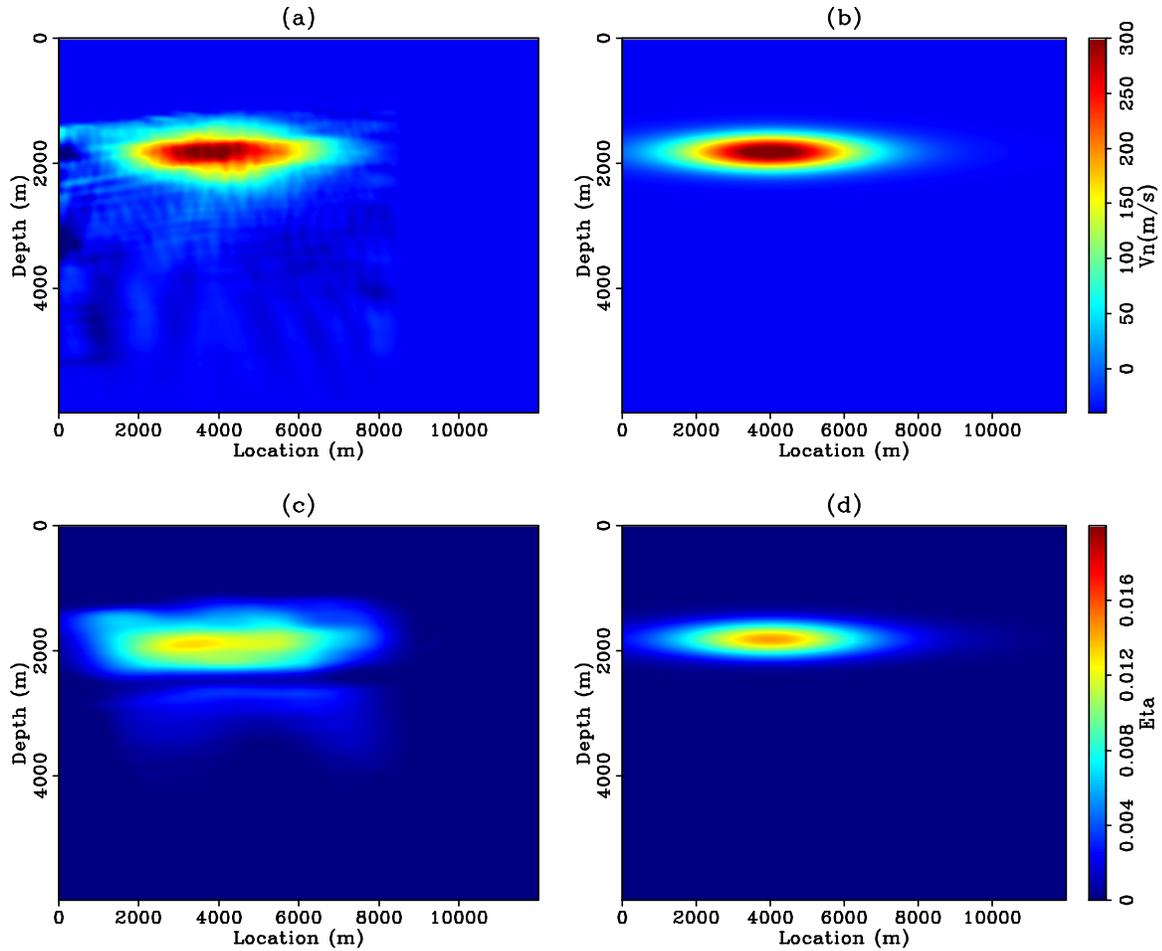
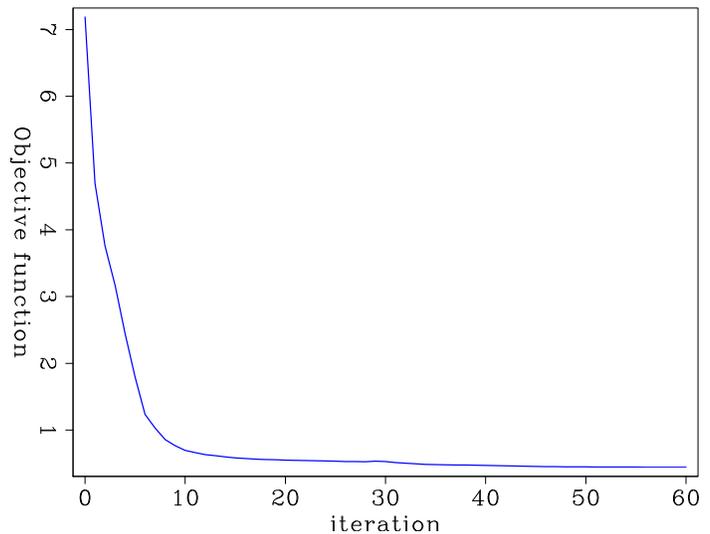


Figure 2.5: Comparison of the anomaly. (a): inverted velocity anomaly; (b): true velocity anomaly; (c) inverted η anomaly and (d) true η anomaly. [CR]

chap2/. bp1-updt

Figure 2.6: Objective function as a function of iteration. [CR]

chap2/. bp1-obj



migrated image in panel (b) and the image obtained with true models in panel (c). The improved velocity and η model yield a better image with correct depth, better focusing, and different interpretations of the subsurface. Compared with the initial image, the reflectors in the inverted image have been pushed down to the correct depths. Flag 1 and 2 in Figure 2.7 point out two representative reflectors whose positions have been corrected by the inversion. The vertical shifts of the deeper reflectors at 5 km depth are larger than 200m. Highlighted by the ellipses, the image shows more continuous reflectors and better focusing after the anisotropic WEMVA inversion. Moreover, flag 3 marks a region where an anticline trap could have been misinterpreted with the leaking point at 4 km in the initial image. It is no longer misleading in the inverted image.

Figure 2.8 compares the initial ADCIGs in panel (a) with the inverted ADCIGs in panel (b). The ADCIGs are extracted every 250m from $x = 2000\text{m}$ to $x = 8000\text{m}$. The initial gathers show strong upward curvatures in the shallow region, indicating undermigration. Interestingly, although both the velocity and η perturbations are positive, we observe downward curvatures in the deeper sections below 4km. This is because the wavepaths bend significantly due to the velocity and η anomaly so that the effective average velocity along the initial wavepaths are higher than that along

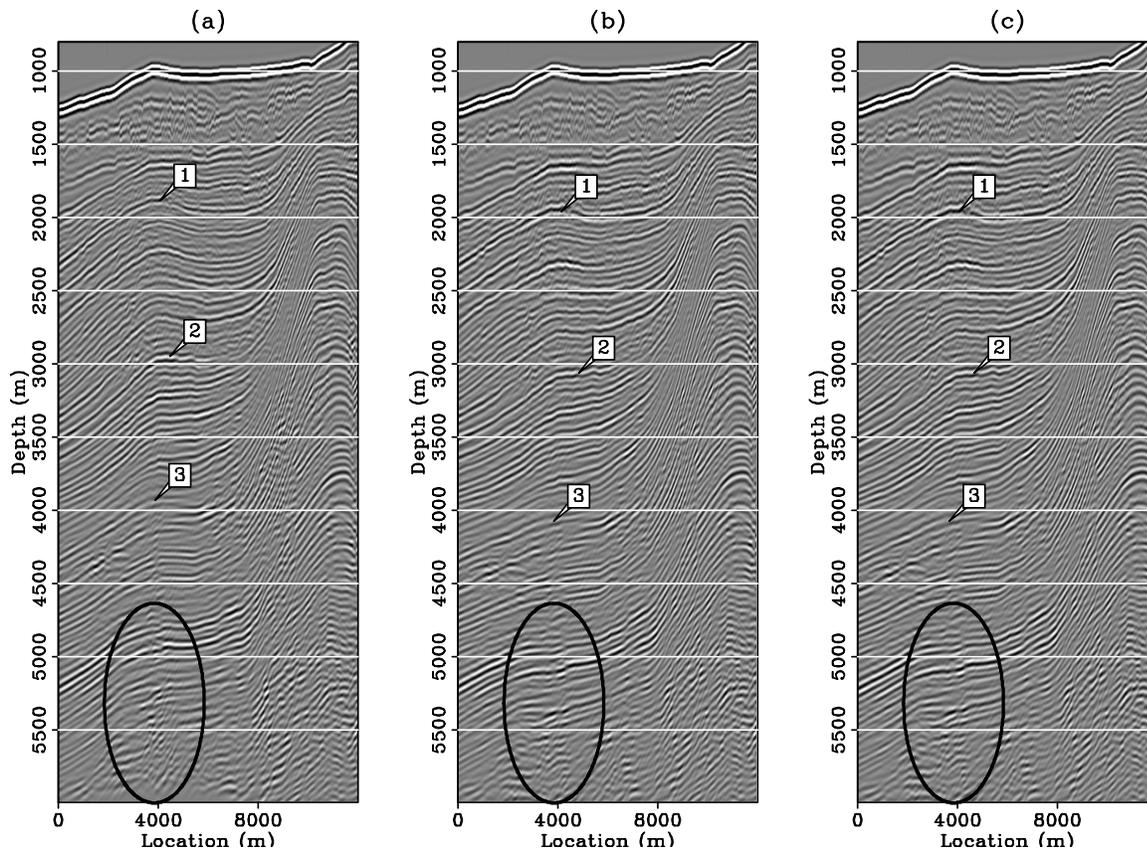


Figure 2.7: Stack image using (a): initial anisotropic model; (b): inverted anisotropic model; and (c) true anisotropic model. Flags 1 and 2 mark the depth of two representative reflectors in this section. Notice the depths of both reflectors have been corrected in the inverted image. The ellipses highlight the region with better focusing and more continuous reflectors. [CR] chap2/. bp1-imag

the correct wavepaths.

After 60 iterations of the anisotropic WEMVA inversion, the events in the ADCIGs are almost completely flattened. Notice both the upward and the downward curving residual moveouts are corrected, confirming that the inverted velocity and η model explains the kinematics of the data.

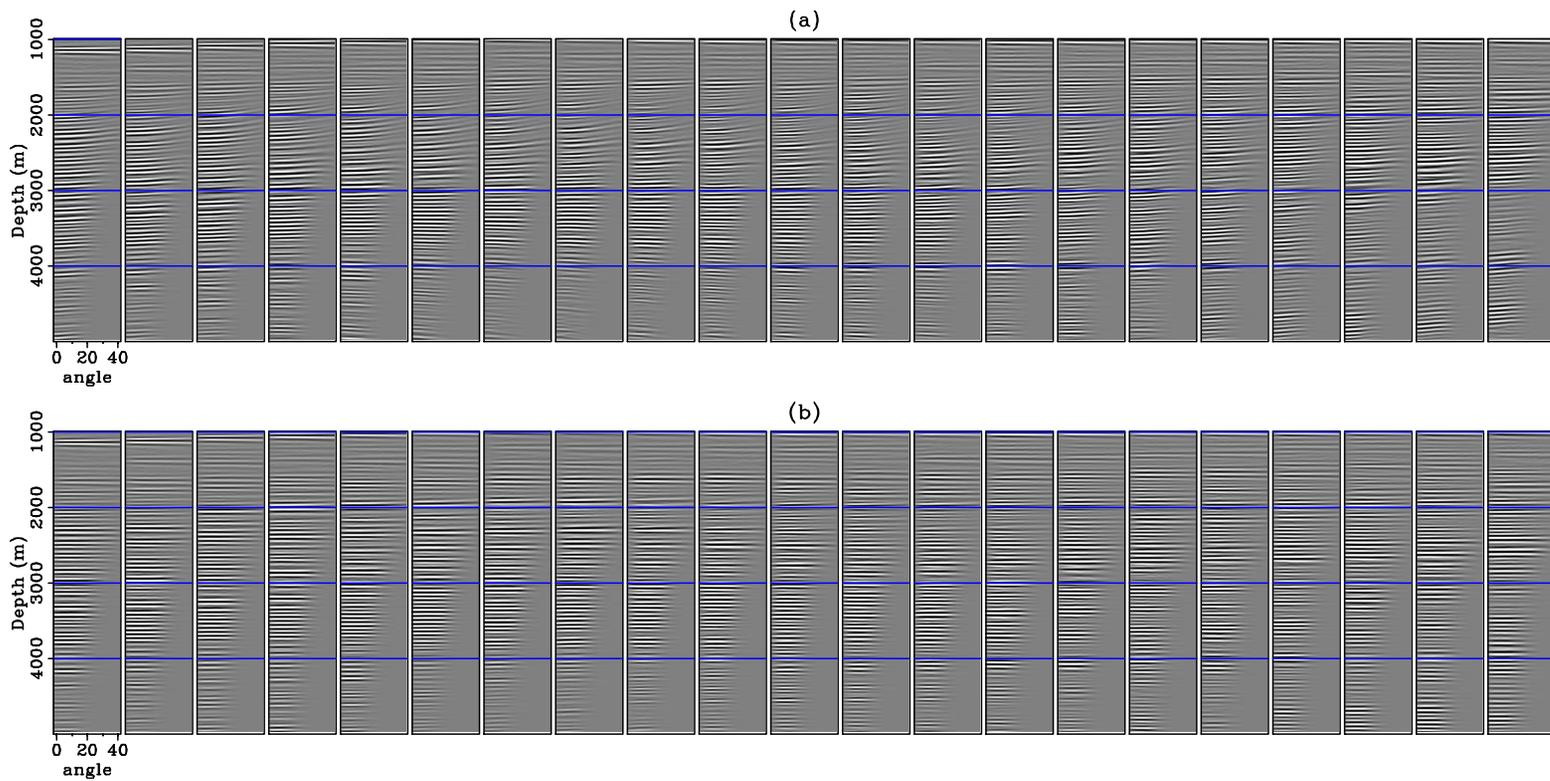


Figure 2.8: Angle-domain common image gathers extracted at 250m interval between $x = 2000\text{m}$ and $x = 8000\text{m}$. (a): ADCIGs using initial anisotropic model, (b): ADCIGs using inverted anisotropic model. [CR]

chap2/. bp1-aimg

FIELD DATA TEST

Anisotropic models are better defined in well-developed and production areas, where various well log, check shots and VSP data are available. The Gulf of Mexico dataset used in this field data test is from a producing area in an offshore sedimentary basin close to a salt body. The current VTI model was built using an interactive visualization method (Bear et al., 2005) that integrates surface seismic, sonic logs, vertical check shots, and offset check shots. Figure 2.9 illustrates the available datasets in this area.

From the bottom row in Figure 2.10, we can see that the PSDM gathers are almost flat, inverted velocities have a reasonable match with the sonic logs, and the modeled traveltimes agrees with the offset check-shot measurements. Therefore, I assume the current anisotropic model for δ is accurate enough and keep it fixed through the WEMVA iterations.

However, according to the color code in Figure 2.9, although most of the traveltimes are fitted very well for near- to mid-offset check shots (indicated by rays in white), the travel times modeled between long-offset shots and the downhole receivers are still greater than the measured travel times (indicated by rays in green). Therefore, a fully automated anisotropic model building method utilizing all types of data would be highly valuable to improve the current model.

I extract a source line from the Gulf of Mexico 3D dataset away from the salt body. The source and receiver positions are plotted in Figure 2.11. Ignoring the cross line offsets, sources are grouped every 100m, and receivers are grouped every 50m. The maximum offset is about 4 km. Using the initial models in Figure 2.12, I obtain the initial migrated image which is shown on the top panel in Figure 2.13.

Many small-scale faults are visible in this area, as shown in the initial migrated image. Migration artifacts at $x = 13$ km and $z = 2500$ m are caused by the sharp discontinuity in the velocity defined by a major fault running from $x = 14$ km on the top to the bottom of the section. The initial angle gathers are shown in the bottom row in Figure 2.13. Because of the streamer geometry, the subsurface reflectors are

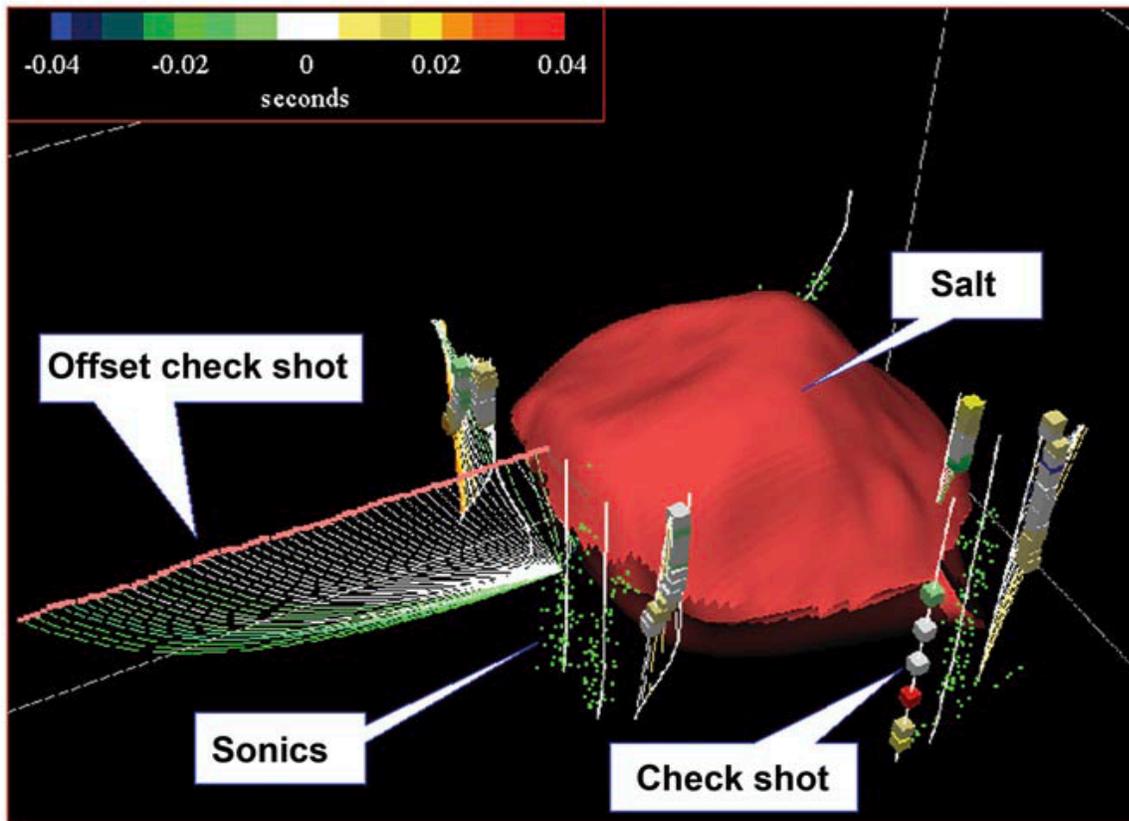


Figure 2.9: Available datasets for anisotropic model building. From Bear et al. (2005).

[NR] chap2/. alldata

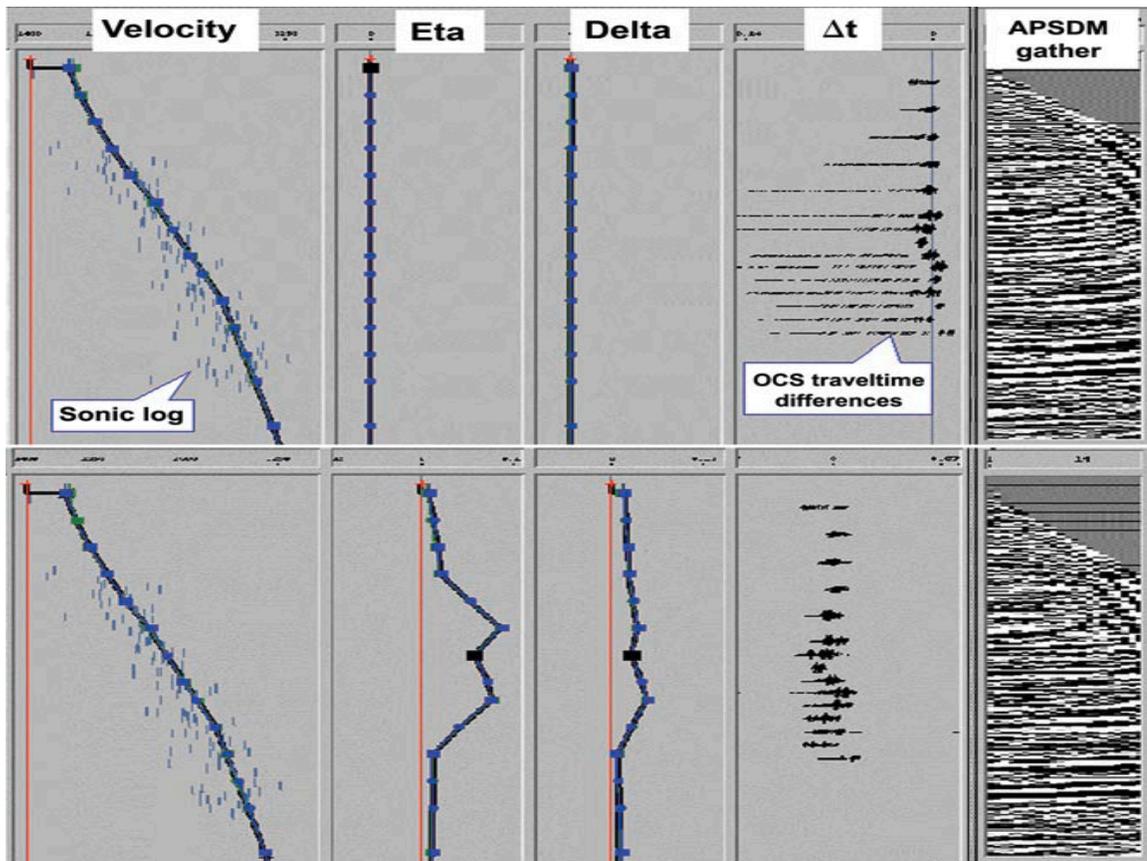


Figure 2.10: Visualization of interactive modeling for anisotropy. The best anisotropic model honors sonic logs, OCS traveltimes and flattens the prestack surface seismic gathers at the same time. From Bear et al. (2005). [NR] chap2/. interact

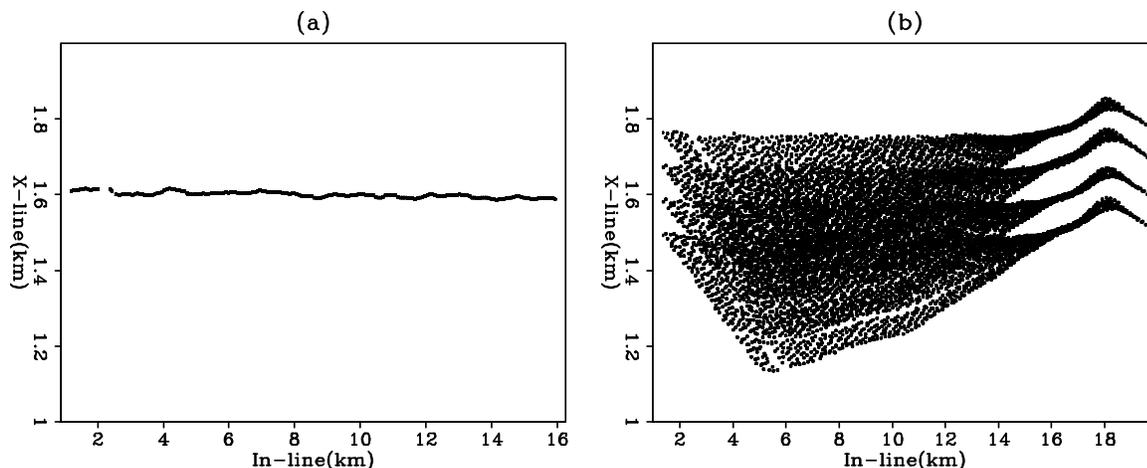


Figure 2.11: Source locations on a 2D line in (a) and corresponding receiver locations in (b). [CR] chap2/. exxongeo

only illuminated from positive angles. Although the gathers are fairly flat, small upward residual moveouts remain in the angle domain common image gathers. This agrees well with the negative residual travel time misfit in the OCS survey. Therefore, there is a chance to improve the model and the image by flattening the gathers.

To constrain the tomography, I estimate the dip field (Figure 2.14(a)) from the initial migrated image. The section shows mostly positive dips. The small negative dips correspond to the large reverse fault to the right of this section. Based on the dip field, non-stationary steering filters are constructed at each grid point. To demonstrate the steered smoothing effect, I apply the steering filter on a section filled with random numbers. The resulting field is shown in Figure 2.14(b). Notice the “synthetic reflectors” created by the steering filters follow the structural dip.

During inversion, the non-stationary steering filters are used to smooth the gradients for both NMO velocity and η . Due to the lack of rock physics information, a stationary matrix Σ with zero off-diagonal components is applied, simply to scale the relative amplitude of the slowness and η gradient. I use a large weight $\alpha = 0.8$ for the stacking power objective function, because the current VTI model is already well defined.

The inverted anisotropic model after 8 iterations is shown in Figure 2.15. Updates between the initial and the inverted VTI models are shown in Figure 2.16(a) and 2.16(b). The inversion successfully resolves a localized shallow anomaly between 13 km and 15 km at around 800 m depth. Notice that the updates in velocity are less than 10% of the background model, indicating the high quality of the initial model. The updates in η are around 25%. These positive updates in both velocity and η agree well with the negative travel time misfits in the previous OCS modeling results (Bear et al., 2005).

Comparing the initial migrated image (top panel in Figure 2.13) with the final migrated image (top panel in Figure 2.17), improved continuity and signal strength can be seen in the area highlighted by the oval. The faulting in this area is also better defined in the final image. If I correlate the inverted models with the final migrated image, I might change the interpretation of the major fault in the initial model: the emerging location on the surface might be shifted from 14.2 km to around 16 km. However, this change of interpretation can only be confirmed with other areal geological information.

The final ADCIGs produced using the inverted models are shown in the bottom row in Figure 2.17. To better illustrate the effects of the model updates, the ADCIGs are sampled more densely between $\text{CMP} = 13\text{km}$ and 16km and sparsely outside of this range. In general, we can see improved flatness for almost all the reflectors. However, DSO and stacking power objective functions are global optimization schemes. They improve the overall flatness and focusing in each iteration but cannot guarantee flatter events at each image point. For example, the ADCIGs at $\text{CMP} = 15\text{km}$ between the depth of 500m and 900m are slightly over compensated to flatten the gathers in the deeper region. More iterations should eventually flatten all the gathers in principle; however, we do not pursue higher resolution results with more iterations in order to avoid overfitting the noise in the image space.

We can verify the effect of the improvements in velocity and η on the angle gathers at different image points. Specifically, for the shallower events above 1km, most improvements happen at large angles over 35° . Therefore, I interpret the improvements

for the shallow events primarily as the contribution of the improved η model. For the deeper events at the same CMP location, both the depth and the flatness of the angle gather have been changed by inversion. The upward-curving events in the angle domain from the initial migration has been flattened by the improvements in both velocity and η model.

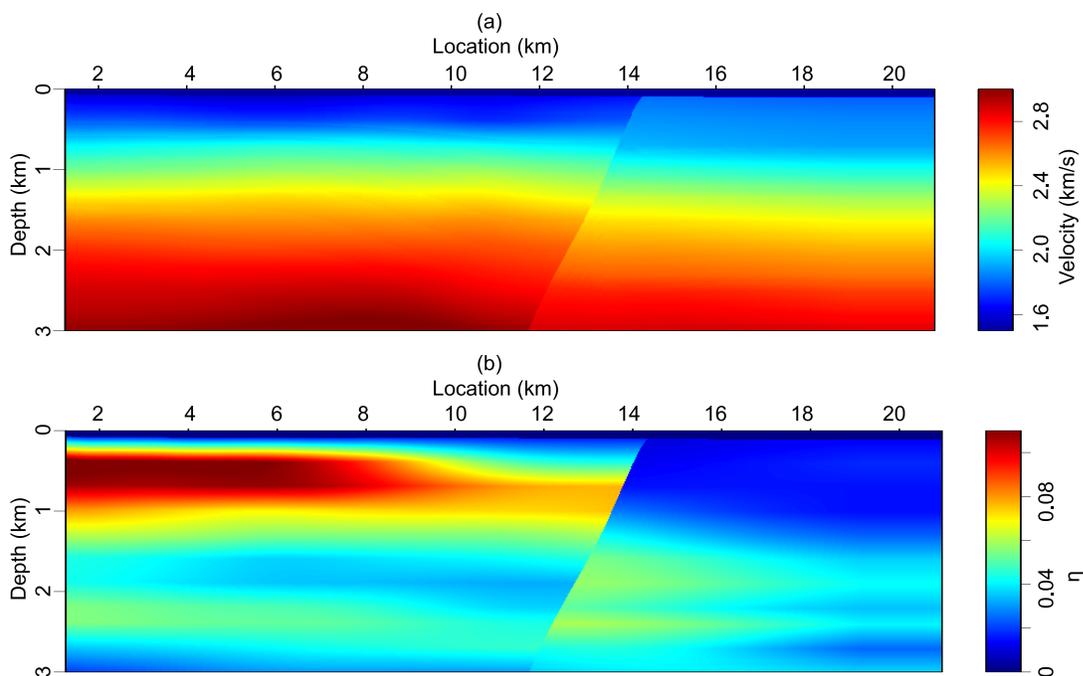


Figure 2.12: Initial velocity model (a) and initial η model (b). [CR]
 chap2/. initmodel

DISCUSSIONS AND CONCLUSIONS

I have presented a methodology of wave equation migration velocity analysis for anisotropic media. By utilizing the Lagrangian adjoint-state method, I demonstrated

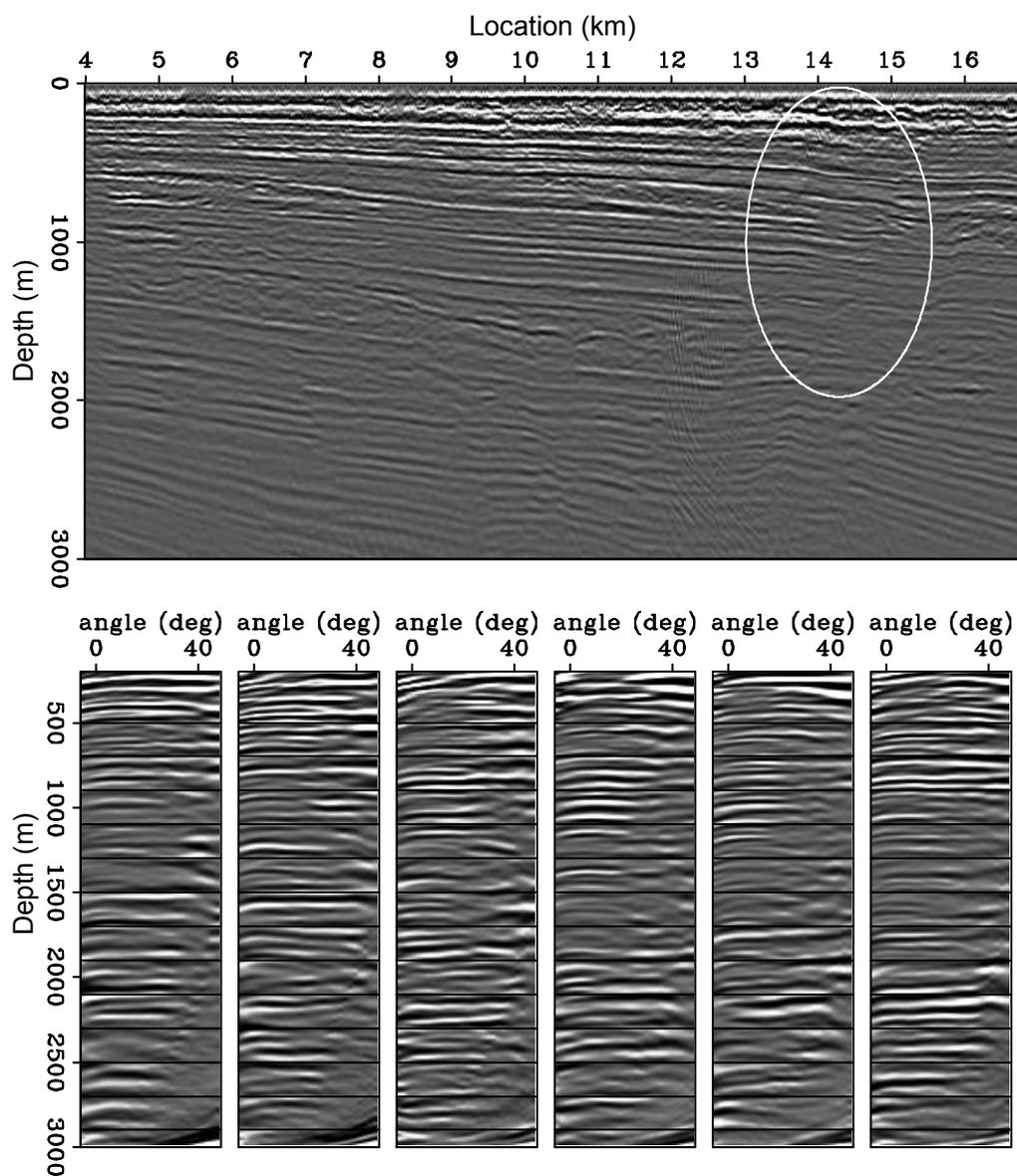


Figure 2.13: The initial stack image (Top panel) and initial angle domain common image gathers at $\text{CMP} = 7, 10, 13, 14, 14.5, 15$ km (Bottom row). [CR] chap2/. image-init-an

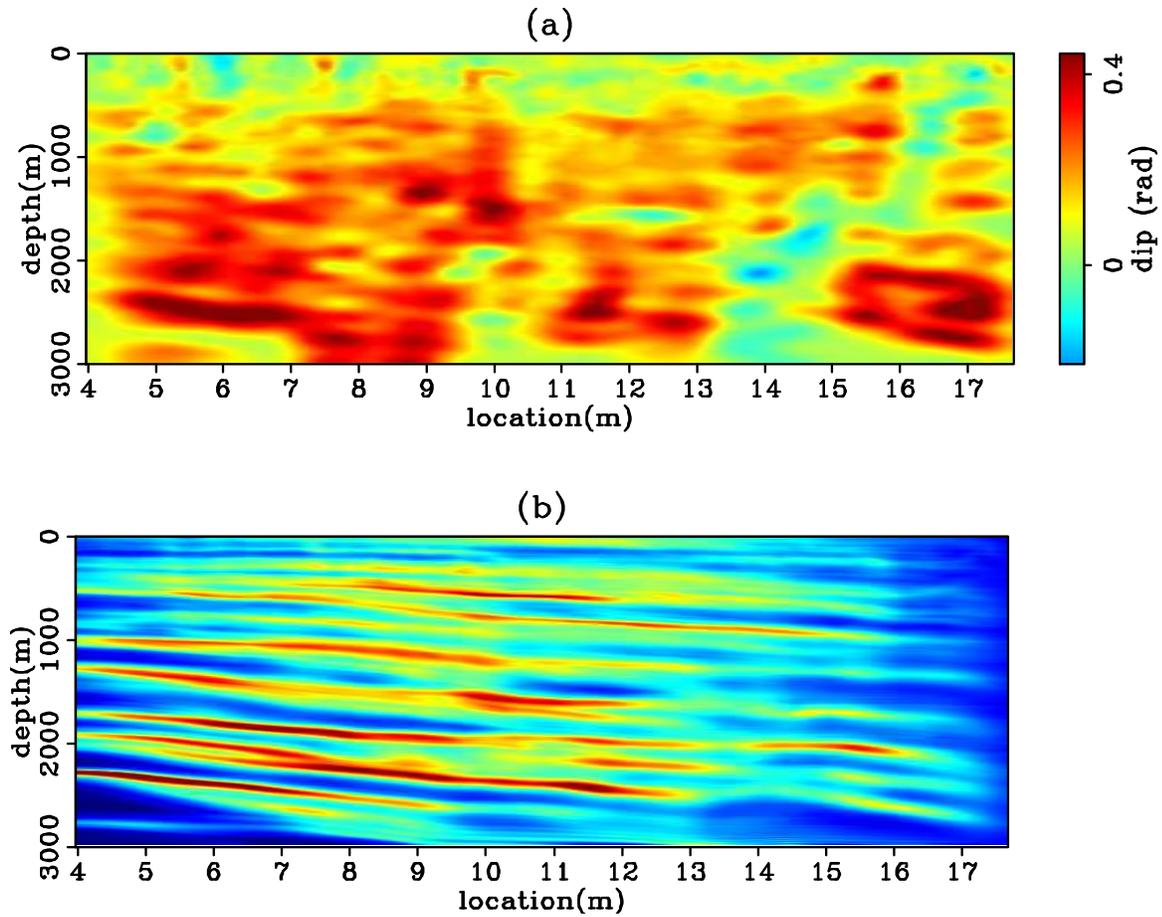


Figure 2.14: (a): Estimated dip field from the initial image on the top panel of Figure 2.13. (b): Reflectors created from a randomly distributed field. [CR]

chap2/. showsteer

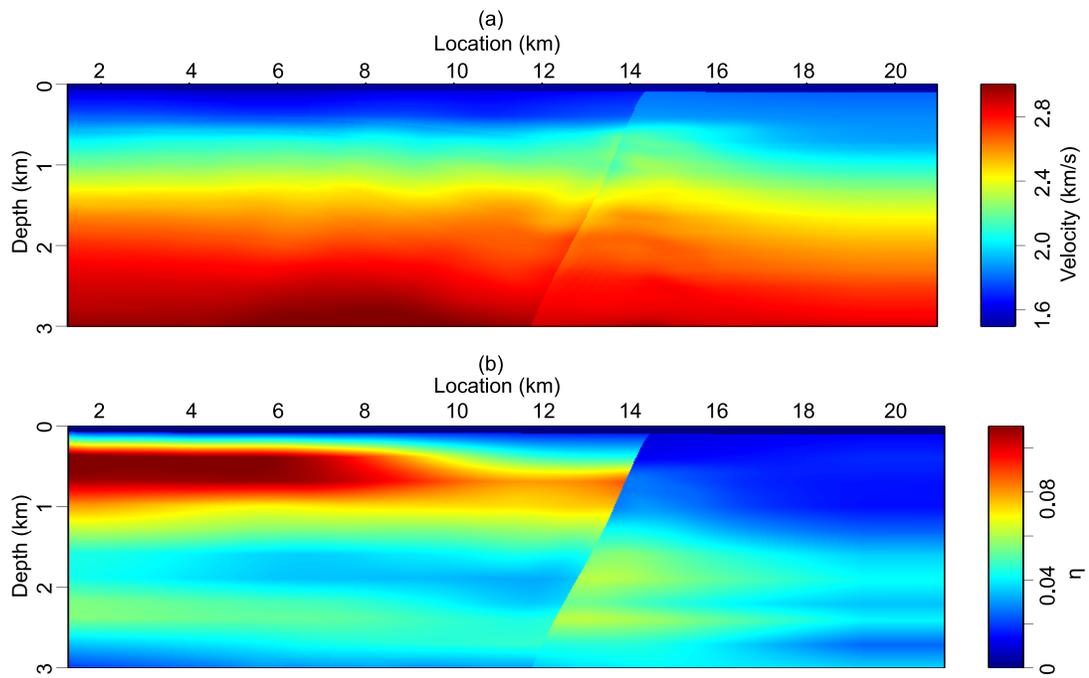


Figure 2.15: Inverted velocity model (a) and inverted η model (b). [CR]
chap2/. finalmodel

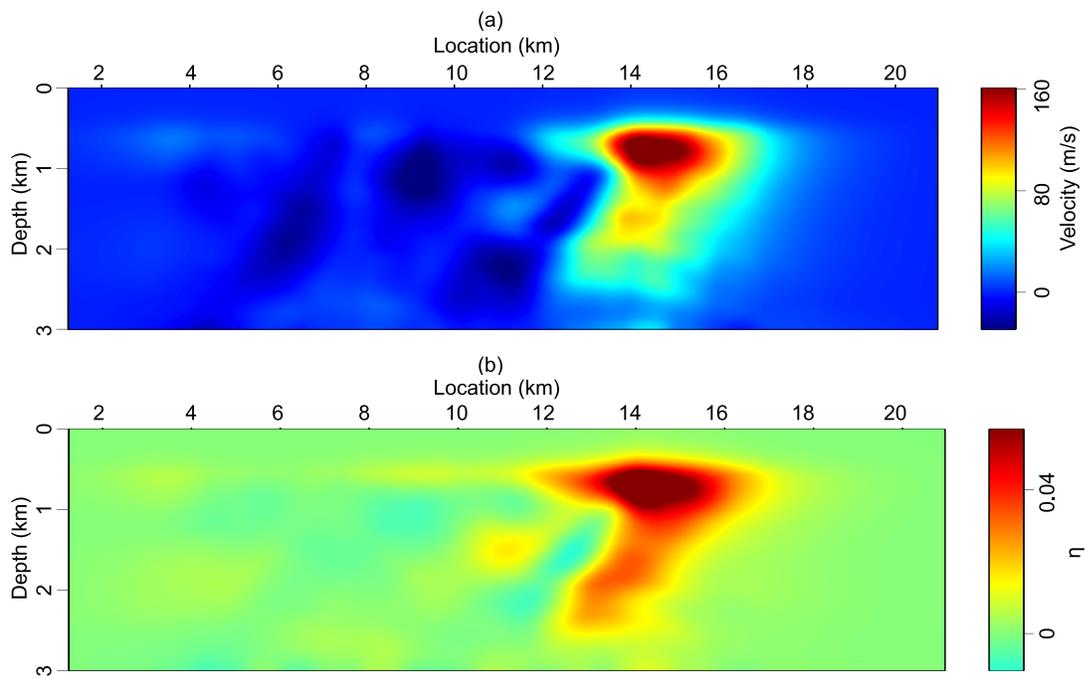


Figure 2.16: Updates in velocity model (a) and updates in η model (b) after inversion.
[CR] chap2/. updates

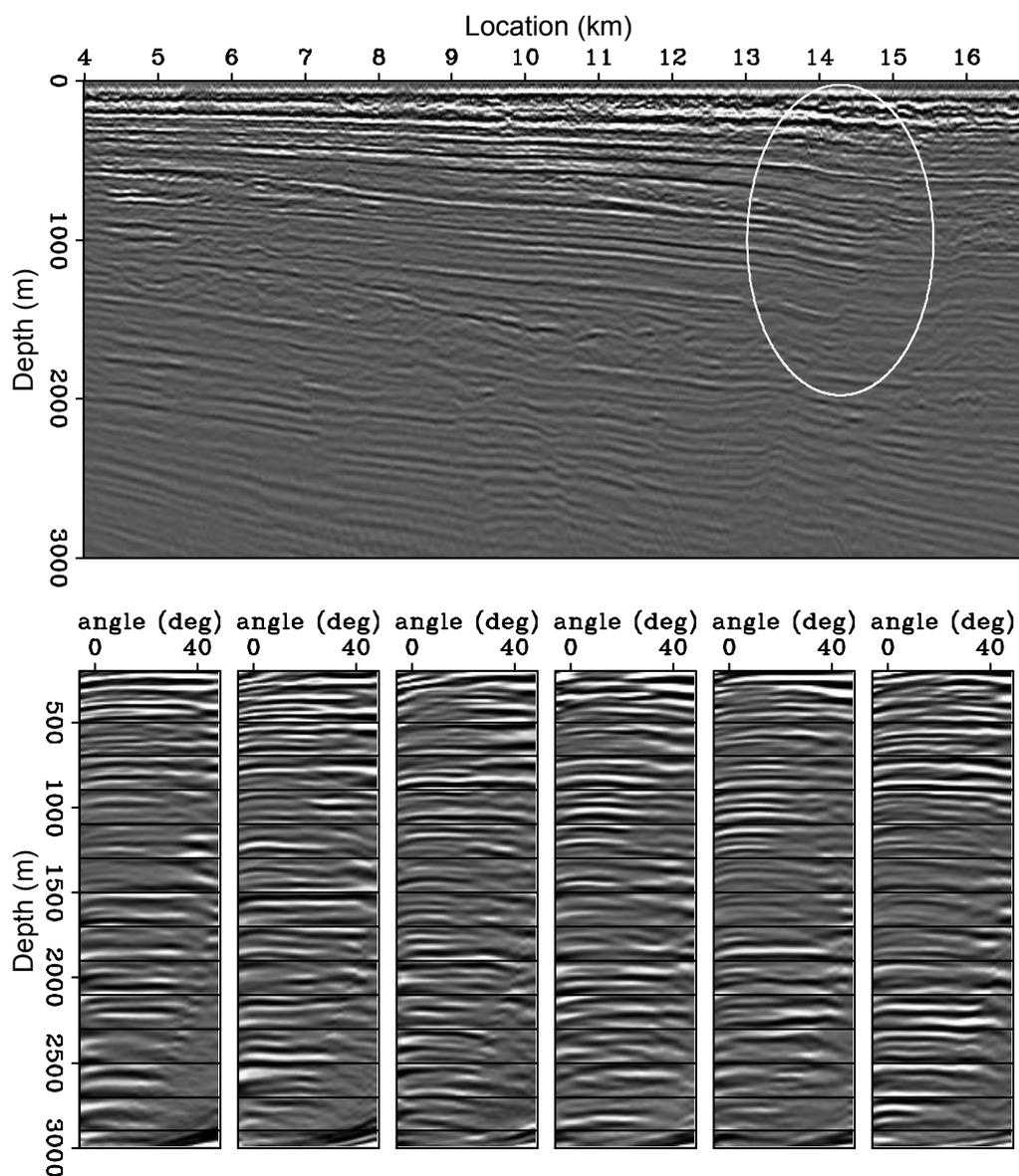


Figure 2.17: The final stack image (Top panel) and final angle domain common image gathers at $\text{CMP} = 7, 10, 13, 14, 14.5, 15$ km (Bottom row). Compared with Figure 2.13, improvements in continuity and enhancements in amplitude strength are highlighted by the oval. [CR] chap2/. image-fnal-an

that our anisotropic WEMVA methodology is a straightforward extension of the well-established isotropic WEMVA methodology: simply adding an additional term for the additional parameter η . Our method is fully automated and picking-free. The grid-based parameterization for the velocity and anisotropic parameters provides the flexibility to reveal perturbations with different scales.

To evaluate the accuracy of the inverted anisotropic models, I combine the DSO objective function and the stacking power objective function, benefiting from the global convergence of the former and the high resolution from the latter. In practice, I turned up the weights on the stacking power objective function as I got closer to the true solution. Nevertheless, the resolution for the anellipticity parameter η is lower than that for NMO velocity due to its weaker influence on the data.

The preconditioning scheme of our proposed method is crucial to obtain geologically meaningful results for both velocity and anisotropic parameters. This scheme can also easily incorporate human interpretation and other prior knowledge. When the data are fitted equally well, I subjectively choose to trust a smoother anisotropic model rather than a highly heterogeneous isotropic model. In addition, I assume the anellipticity parameter η , being the ratio between the directional velocities, is smoother in space than the velocities. These smoothness assumptions mitigate the ambiguities between anisotropy and heterogeneity.

When the lithological and well log information is not available, I can only estimate the diagonal elements of the cross-covariance matrix, reflecting the autocorrelation of velocity and η perturbation separately. As a result, it takes many iterations and higher weights on the stacking power term in the objective function to obtain high resolution models and sort out the ambiguities between the parameters where the data permit (namely, in the shallow region). However, had I roughly known the lithological information in a certain region, I would be able to build a cross-covariance matrix which reflects the cross-correlation between velocity and η . Consequently, I would be able to speed up the convergence and better resolve the ambiguities (Li et al., 2011). I will explain the framework in Chapter 3.

Another way to mitigate the ambiguity between the parameters is to include more data. The proposed anisotropic WEMVA methodology works in the image space with reflection data, which are most sensitive to the NMO velocity. On the other hand, most of the current full waveform inversion practices work in the data space with long-offset refractions and diving waves (e.g. Gholami et al. (2013)). These transmission components of the seismic data are most sensitive to the horizontal velocity. Jointly inverting for both reflection and transmission data should provide better constraints on the anisotropic properties of the subsurface.

ACKNOWLEDGEMENT

The authors thank BP for providing the synthetic model and ExxonMobil for providing the field data.

Chapter 3

Rock physics constrained anisotropic WEMVA

This chapter presents a regularization scheme utilizing available rock physics data to better constrain the anisotropic wave-equation migration velocity analysis (WEMVA) and to better resolve the ambiguity among the anisotropic parameters. This additional constraint is necessary because velocity model building is a highly underdetermined and nonlinear problem. Using surface reflection data, the sensitivity of the image-space objective function to the anisotropic parameters ϵ and δ is much lower than the sensitivity to velocity. As will be shown in a simple 2D synthetic example, the values of the image-space objective function at a range of model realizations are almost the same. These models are called “equiprobable” models. As a result, many VTI models with vastly different geological interpretations may explain the surface seismic data equally well.

In addition to the spatial covariance to constrain the spatial correlation of each VTI parameter individually introduced in Chapter 2, a cross-parameter covariance at each subsurface location is included in the inversion in this chapter. The cross-parameter covariance in an ideal case is constructed based on the local lithological information at every grid point in the subsurface. There are two significant effects

that the regularization with cross-parameter covariance brings to the updates for the VTI parameters. First, instead of spreading the updates evenly along the wavepath, the regularization term allows more updates in the regions where the models are highly uncertain. Using a preconditioning scheme, the resulting update direction then follows a priori defined spatial distribution since the first iteration. The same distribution could be obtained by many iterations. Second, the regularization term brings extra information of parameter updates from the correlation of the other parameters. These improvements help the inversion converge faster and yield VTI models that are more consistent with the underlying geological and lithological assumptions. In this chapter, I will present the theory for the regularized anisotropic WEMVA and demonstrate the improvements on a synthetic dataset.

INTRODUCTION

Anisotropic model building tries to resolve more than one parameter at each model location. This number could be three for a vertical transverse isotropic (VTI) media, and can increase to five for a tilted transverse isotropic (TTI) media. Traditional surface seismic tomography may be able to produce an accurate isotropic earth model efficiently for a large area when the acquisition is dense and the earth is well-illuminated by rays at a wide range of angles. However, surface seismic data inversion becomes ill-posed and highly underdetermined due to the rapidly increasing model space with the increasing complexity of the subsurface.

One important disadvantage of surface seismic tomography is its lack of depth information. During tomography, neither the low wavenumber velocity nor the high wavenumber reflectivity are known. This issue is more severe when we consider anisotropy. Several localized tomography around the wells are studied to add the depth dimension into the inversion (Bear et al., 2005; Bakulin et al., 2010b,a). Joint inversion of surface seismic data and borehole data (check-shots and walkaway VSPs) in these studies shows great potential to yield better defined earth models. Due to the ambiguity between the parameters, it is difficult to resolve a reliable, unique

anisotropic model in 3D even the borehole aided localized tomography (Bakulin et al., 2009).

In the anisotropic case, the uncertainties are further increased because the sensitivities of the kinematics of the surface seismic data to the anisotropic parameters are much lower than to velocity. Large offsets and a wide range of illumination angles are required to constrain the anisotropic parameters. Consequently, the recoverable depth range for the anisotropic parameters is much shallower than when a simple isotropic velocity is estimated. Even in the shallow region, where the seismic waves travel with wide angles and large offsets, the kinematic effects of the velocity can still overwhelm the inversion.

To help with the inversion for anisotropy, we need to use our prior knowledge of the subsurface. In addition to the two-point (spatial) covariance I introduced in Chapter 2, a single point (local) cross-parameter covariance can be used to fully describe the subsurface. One way to estimate the cross-parameter covariance is from rock physics studies (Hornby et al., 1995; Sayers, 2004, 2010; Bachrach, 2010b). Many authors (Dræge et al., 2006; Bandyopadhyay, 2009; Bachrach, 2010a) have built depth trends for seismic purposes. In particular, Bachrach (2010a) developed both deterministic and stochastic modeling schemes based on the rock physics effective-media models for compacting shale and sandy shale. The stochastic modeling enables us to explore the range of possible anisotropic parameters based on the rock physics modeling parameters. Further corroborated by core measurements, the parameters needed by the rock physics model are limited to a certain range, which greatly reduces the possible range of the VTI parameters. These rock physics modeling results can be used to construct the initial VTI models and the covariance relationships among the VTI parameters. Li et al. (2011) and Yang et al. (2012) have demonstrated that the rock physics prior models are helpful in constraining ray-based tomography.

In this chapter, I first analyze the sensitivities of the WEMVA objective function with respect to different VTI parameters on a homogeneous VTI model. Using an interpolation example, I demonstrate the additional information that the cross-parameter covariance brings to the inversion. I then test three different regularization

schemes when a synthetic dataset is inverted using anisotropic WEMVA. The inversion results show that when the accurate full covariance matrix is applied, the convergence of the WEMVA inversion as well as the lithological definition of the inverted models are improved.

WEMVA AND ROCK-PHYSICS REGULARIZATION

As introduced in Chapter 2, anisotropic wave-equation migration velocity analysis (WEMVA) aims at building an anisotropic earth model that minimizes the residual image from the surface seismic data. This optimization problem is highly non-linear and underdetermined. Therefore, we commonly add a model regularization term to the anisotropic WEMVA objective function 2.9 defined in the image space to constrain the null space and stabilize the inversion. The resulting objective function is as follows:

$$S(\mathbf{m}) = \frac{1}{2} \|\mathbf{D}_\theta \mathbf{I}(\mathbf{x}, \theta)\| - \alpha \frac{1}{2} \left\| \sum_\theta \mathbf{I}(\mathbf{x}, \theta) \right\| + \beta \frac{1}{2} (\mathbf{m} - \mathbf{m}_{\text{prior}})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{\text{prior}}), \quad (3.1)$$

where the first two terms define the “data fitting” objective, and the third defines the “model regularization” objective. The first term is to minimize the differential semblance in the data fitting objective, and the second term is to maximize the stacking power. Model \mathbf{m} is the VTI subsurface model, $\mathbf{I}(\mathbf{x}, \theta)$ is the migration image in the angle domain with θ the aperture angle and \mathbf{D}_θ a derivative operator along the angle axis. In the model regularization objective, $\mathbf{m}_{\text{prior}}$ and \mathbf{C}_M define a Gaussian distribution of a prior model that is ideally independent of the seismic data. This regularization will bring more information into the optimization. Parameters α and β balance the relative weights among different objectives.

The data fitting objective relates the incoherence in the angle domain common image gathers to the inaccuracy in the subsurface models. To test the data objectives, I model a simple synthetic dataset using a homogeneous VTI model ($v_v = 2\text{km/s}$, $\epsilon = 0.2$, and $\delta = 0.1$) with one flat reflector. The maximum offset is 6km, and the depth

of the reflector is 1.5km. The data are then migrated using all possible combinations of v_v , ϵ , and δ when v_v varies in $[1.5, 2.5]$ km, $\epsilon \in [0.1, 0.3]$, and $\delta \in [0, 0.2]$. Based on the migrated images in the angle domain, the data fitting objective at each model point in this subspace is then computed according to the first two terms in Equation 3.1.

Figure 3.1 shows the data fitting objective function assuming one of the three VTI parameters is accurate. Panels (a), (b), and (c) are extracted from the $v_v = 2$ km plane, $\delta = 0.1$ plane, and $\epsilon = 0.2$ plane, respectively. When the third parameter is accurate, the data fitting objective function is convex and comes to a minimum at the correct solution for the other two parameters. However, the resolution of the objective function with respect to each parameter is dramatically different. A much higher resolution for velocity than for ϵ and δ can be seen from Figures 3.1(b) and (c). The resolution of ϵ and δ is at least an order of magnitude lower than that of velocity (notice the smaller range of values on the colorbar in Figure 3.1(a)).

Moreover, a significant trade-off among the VTI parameters can be observed from Figure 3.1. The better resolved direction on panel (a) is the η direction, which is roughly parallel to $\epsilon - \delta$ when δ is small. An obvious trade-off between v_v and δ is also observed (Figure 3.1(c)). Due to the limited source-receiver offset, the objective function has the lowest resolution for ϵ .

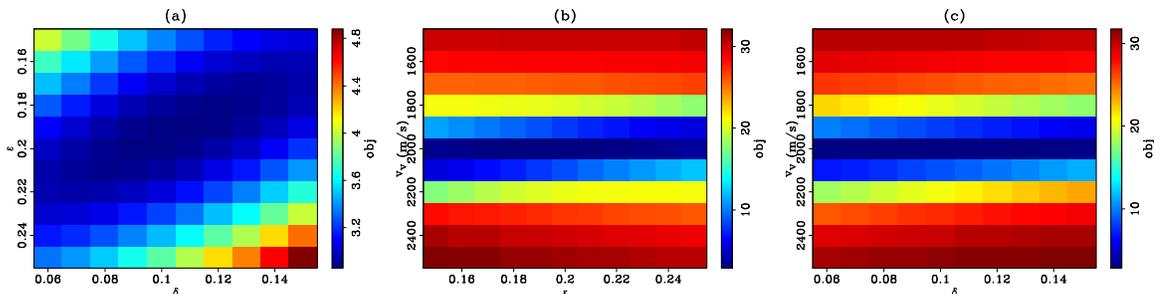


Figure 3.1: The value of data fitting objective function extracted from (a) $v_v = 2$ km plane, (b) $\delta = 0.1$ plane, and (c) $\epsilon = 0.2$ plane. Notice the different color scale in panel (a). [CR] chap3/. objfun2

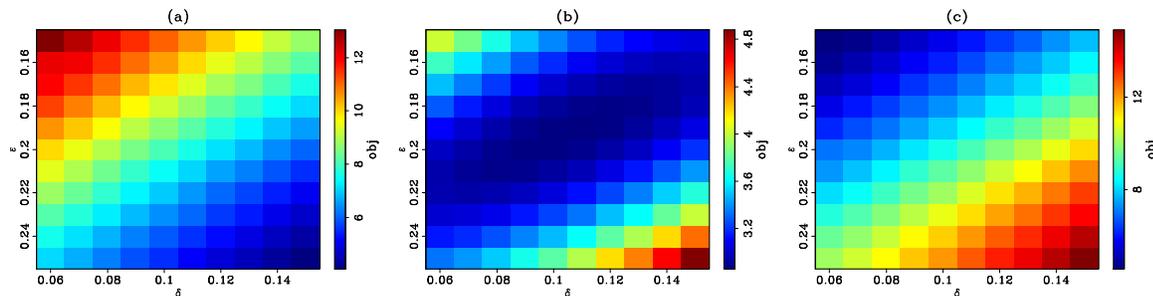


Figure 3.2: The value of data fitting objective function extracted from (a) $v_v = 1.9\text{km}$ plane, (b) $v_v = 2\text{km}$ plane, and (c) $v_v = 2.1\text{km}$ plane. [CR] chap3/. objfun3

For pressure waves, velocity has a dominant effect on the kinematics due to its first-order influence. As a result, the objective function is biased towards the velocity error despite the error in the other VTI parameters. Figure 3.2 shows the topography of the data fitting objective function extracted from the (a) $v_v = 1.9\text{km}$ plane, (b) $v_v = 2\text{km}$ plane, and (c) $v_v = 2.1\text{km}$ plane, respectively. When the velocity is inaccurate, the objective function loses its convexity in the $\epsilon - \delta$ plane. The objective function always dips towards higher ϵ and higher δ when velocity is slower, and towards lower ϵ and lower δ when velocity is faster. In these cases, the gradient of the objective function would not be able to guide the inversion to the correct solution for the anisotropic parameters if the errors in ϵ and δ are in the opposite direction than the velocity error.

In real cases, the “data fitting” objective function has worse behavior because the structure of the earth subsurface is highly complex with heterogeneities at all scales and densely distributed dipping reflectors. These complexities weaken the convexity of the objective function and create local minima.

To better constrain the model and to mitigate severe ambiguities among the VTI parameters, we include additional information in the inversion. I utilize the regularization term in Equation 3.1 assuming a multivariate Gaussian distribution for the VTI model parameters. To speed up the convergence, I use the same preconditioning scheme introduced in Chapter 2 instead of the original regularization scheme.

WEMVA with the rock physics regularization

Assuming a Gaussian distribution, Tarantola (1984) characterizes the prior information using the mean and the covariance of the model and includes it as a regularization term. I separate the covariance into two parts as follows: a spatial covariance between the same parameter at different locations, and a cross parameter covariance between different parameters at the same location.

The spatial covariance is mainly defined by the structure of the area. As presented in Chapter 2, the spatial covariance can be estimated using a set of steering filters. The cross-parameter covariance can be inferred from the lithological information at the model location. In Chapter 2, I introduce the cross-parameter matrix with diagonal terms only due to the lack of the lithological information. This diagonal preconditioning balances the relative scales among the VTI parameters, but it ignores the correlation among them.

However, when a rough estimate of the lithological environment is known, we can build a more complete cross-parameter covariance using rock physics modeling. As an example, the Gulf of Mexico is populated with large shale deposits. Therefore, rock physics principles can be used to estimate the range of anisotropic parameters considering a compacting shale model (Bachrach et al., 2011). As demonstrated in Appendix B, the rock physics stochastic modeling for shale anisotropy shows that in the shallow sediments, a high velocity zone is generally collocated with sand layers where the anisotropy is low. On the contrary, diagenesis processes in the more compacted deeper sediments alter the clay mineral from smectite to illite, which increases both velocity and anisotropy simultaneously. Valuable prior knowledge about the subsurface can be included in the inversion via both the diagonal and the off-diagonal terms in the cross-parameter covariance matrix.

I assume the spatial covariance and local cross-parameter covariance components are independent from each other (Li et al., 2011). I use the same preconditioning scheme as in Chapter 2 to speed up the convergence. The preconditioning scheme is as follows: steering filters to approximate the square-root of the spatial covariance, and

a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the preconditioning variable \mathbf{n} is related to the original model \mathbf{m} as follows:

$$\mathbf{m} = \mathbf{\Sigma}\mathbf{S}\mathbf{n}. \quad (3.2)$$

In Equation C.37, the smoothing operator \mathbf{S} is a band-limited diagonal matrix:

$$\mathbf{S} = \begin{vmatrix} \mathbf{S}_v & 0 & 0 \\ 0 & \mathbf{S}_\epsilon & 0 \\ 0 & 0 & \mathbf{S}_\delta \end{vmatrix}, \quad (3.3)$$

with potentially different smoothing operators for velocity, ϵ , and δ , according to the geological information in the study area. The standard deviation matrix $\mathbf{\Sigma}$ is the square-root of the covariance matrix:

$$\mathbf{\Sigma} = \begin{vmatrix} C_{vv} \mathbf{I} & C_{v\epsilon} \mathbf{I} & C_{v\delta} \mathbf{I} \\ C_{\epsilon v} \mathbf{I} & C_{\epsilon\epsilon} \mathbf{I} & C_{\epsilon\delta} \mathbf{I} \\ C_{\delta v} \mathbf{I} & C_{\delta\epsilon} \mathbf{I} & C_{\delta\delta} \mathbf{I} \end{vmatrix}^{1/2}. \quad (3.4)$$

The diagonal elements C_{vv} , $C_{\epsilon\epsilon}$, and $C_{\delta\delta}$ denote the variance of the velocity, ϵ , and δ , respectively. The off-diagonal elements $C_{v\epsilon}$, $C_{v\delta}$, and $C_{\epsilon\delta}$ denote the cross-variance between the velocity and ϵ , between velocity and δ , and between ϵ and δ , respectively. These elements can be obtained by rock-physics modeling and/or lab measurements (Bachrach et al., 2011; Li et al., 2011). The covariance matrix $\mathbf{\Sigma}$ is symmetric. As a result, there are only six independent components in the covariance matrix. In an ideal case, matrix $\mathbf{\Sigma}$ should be estimated at each subsurface location to reflect the local lithological information.

Interpolation test with the preconditioning operators

To demonstrate the extra information the preconditioning operators introduce to the inversion, I test these preconditioners on a classic missing data problem (Claerbout, 2009). Consider the synthetic example in Figure 3.3, where the left column shows the current best spatial estimation of the VTI parameters (initial model) and the right column shows the true VTI parameters (true model). To build the initial velocity model, I assume the water bottom topography is known and the water velocity is 1.5km/s. The basement interface is assumed to be flat and the velocity of the basement rock is 3.5km/s. Then, I linearly interpolate the velocity between the water bottom and the basement interface. I follow a similar approach to build the initial ϵ and δ model. The resulting initial VTI model shows smooth, linearly increasing trends for velocity, ϵ , and δ .

The true velocity, ϵ , and δ updates are plotted in Figures 3.4(a), (b) and (c), respectively. The true velocity updates show two anticline structures up shallow and a major anticline structure down deep. The updates in the anisotropic parameters contain more layering information than the velocity update. These are the solutions to the model building problem.

Suppose that, among the three unknown parameters, only velocity has been measured at random well locations in this section as shown in Figure 3.4(d). The goal of the inversion is to interpolate the well information onto a regular 2D grid and to provide reasonable estimations for the unmeasured anisotropic parameters.

The interpolation problem can be formulated as a data fitting objective

$$\mathbf{WB} d \approx \mathbf{W}m; \tag{3.5}$$

to match the interpolated model m with the binned data $\mathbf{B} d$ at the well location defined by \mathbf{W} . Additionally, the distribution of the model must follow a user-defined

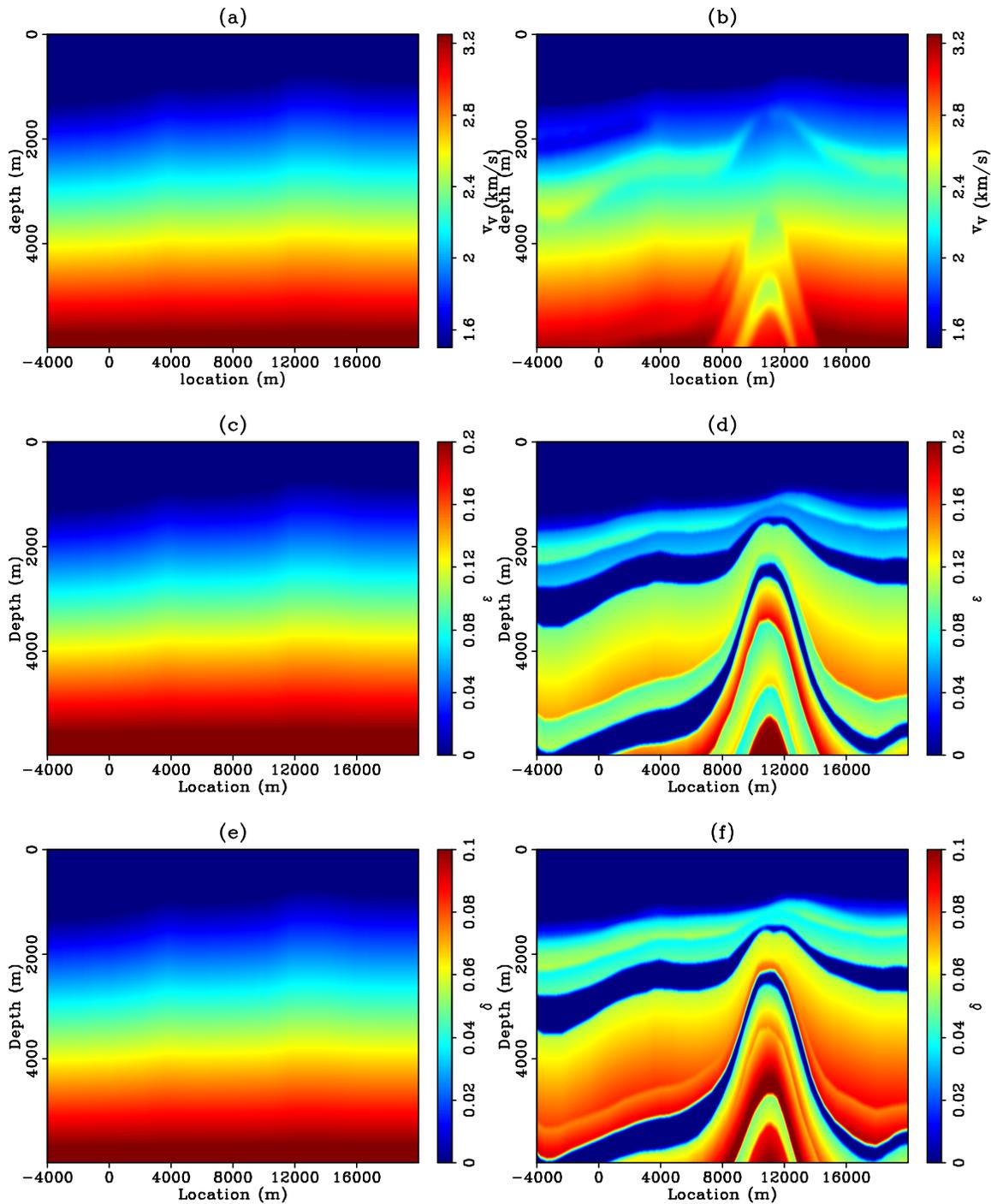


Figure 3.3: Comparison of the initial (left column) and the true (right column) models for BP2007 synthetic test. Top row: velocity models; middle row: ϵ models; bottom row: δ models. [CR] chap3/. init-cvn-model

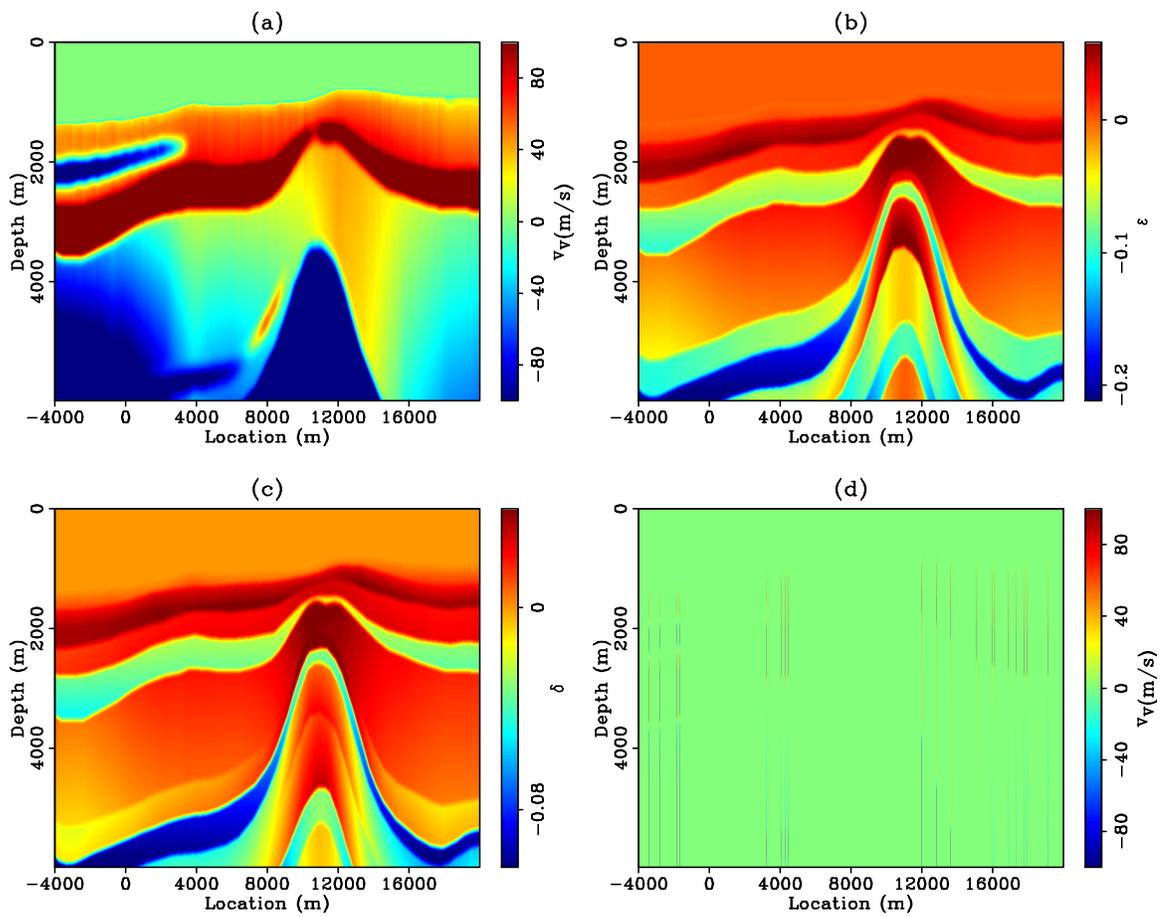


Figure 3.4: True updates in velocity in (a), ϵ in (b) and δ in (c). Pseudo velocity logs at random locations are plotted in (d). [CR] chap3/. cvn-mismdl

prior covariance \mathbf{C} :

$$0 \approx \mathbf{C}^{-1/2}m. \quad (3.6)$$

To speed up the inversion, a preconditioning strategy is used by introducing a preconditioning variable

$$n = \mathbf{C}^{-1/2}m. \quad (3.7)$$

Now, the system of equations I solve is as follows:

$$\mathbf{WB} d \approx \mathbf{WC}^{1/2}n; \quad (3.8)$$

$$0 \approx n. \quad (3.9)$$

Different choices can be made for the preconditioner $\mathbf{C}^{1/2}$. As discussed in the previous chapter, the structure dip filters can be good estimates of the spatial covariance of the velocity and anisotropic model individually. Figure 3.5 shows a stacked image of the studied area (Panel (a)) and the dip field estimated from it (Panel (b)). Steering filters \mathbf{S}_v (Clapp, 2000) are estimated at each point on this section according to the dip field. I use the same steering filters for ϵ and δ , that is $\mathbf{S}_\epsilon = \mathbf{S}_\delta = \mathbf{S}_v$.

In addition, correlations among the VTI parameters at a single location are obvious. Point-by-point, I calculate the variance and the cross-variance using the true updates in Figure 3.4, and I plot the six independent components of the covariance matrix in Figure 3.6 and 3.7. The diagonal components in Figure 3.6 describe the spatial distribution of the variance between the initial and true model for each VTI parameter. The off-diagonal components in Figure 3.7 describe the cross-covariance between two given VTI parameters.

To demonstrate the value of the cross-parameter covariance, I compare the inversion results of the following two preconditioners:

$$\mathbf{C}_1^{1/2} = \mathbf{S}, \quad (3.10)$$

and

$$\mathbf{C}_2^{1/2} = \mathbf{S}\Sigma. \quad (3.11)$$

The interpolation results of both tests after six iterations are plotted in Figure 3.8.

The interpolation result for velocity using $\mathbf{C}_1^{1/2}$ is shown in Figure 3.8(a). I do not plot the interpolation results for ϵ and δ because both of them are zero due to the lack of measurements. The inversion results are better to the right side of the section where the velocity logs are denser. The shallow sand layer has been almost perfectly recovered. On the left side, the inversion interpolates between and extrapolates from the well logs along the structural dip. The result hints the shallow sand layer and the deep anticline structure; however, the blank spaces between the logs are too large to be filled.

The inverted velocity using $\mathbf{C}_2^{1/2}$ (Figure 3.8(b)) is very similar to that using $\mathbf{C}_1^{1/2}$. However, extra information added by the off-diagonal terms in the covariance matrix shows correct updates in the inverted results for ϵ and δ . On the right side, the inversion almost perfectly recovers the ϵ and δ updates along with the proper layering in both the shallow and the deep sections. The reconstruction of the deep layering is mostly attributed to the accurate correlation between velocity and the anisotropic parameters. To the left, the correct ϵ and δ updates are obtained wherever the velocity has been correctly resolved. In the regions where velocity is missing, the off-diagonal terms do not falsely “create” updates for ϵ and δ .

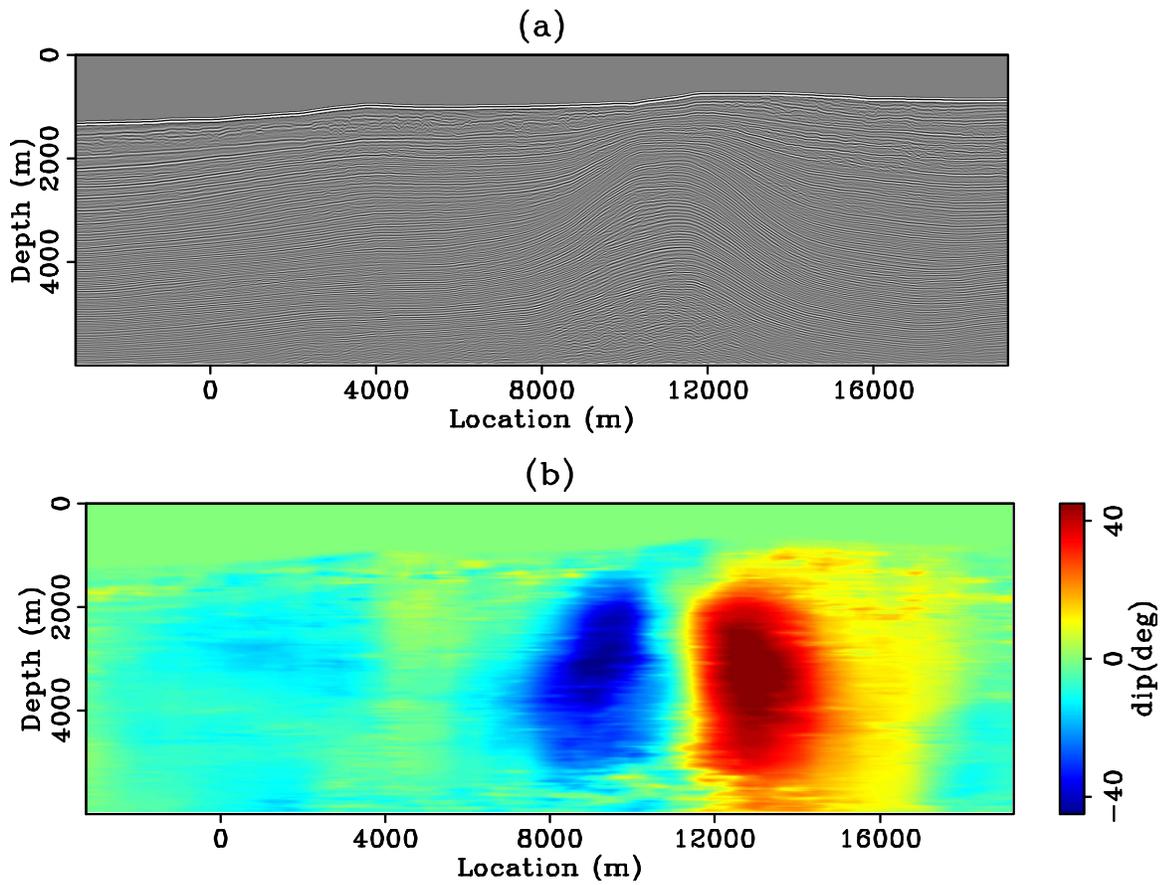


Figure 3.5: A stacked image of this section in panel (a), and the corresponding dip field in panel (b). [CR] chap3/. cvn-dip

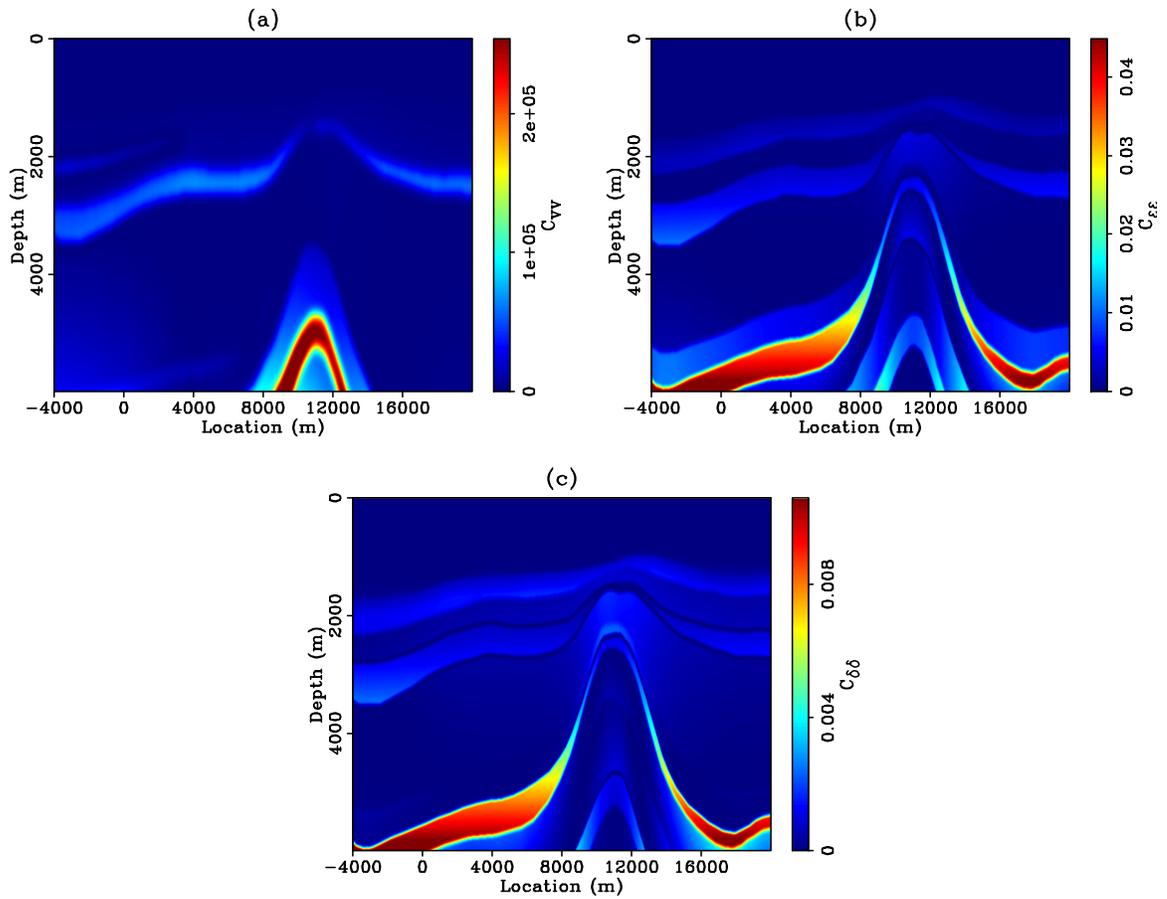


Figure 3.6: Diagonal elements of the covariance matrix of the VTI model. (a) C_{vv} : map of variance for velocity, (b) $C_{\epsilon\epsilon}$: map of variance for ϵ , and (c) $C_{\delta\delta}$: map of variance for δ . [ER] `chap3/. cvn-stdv-diag`

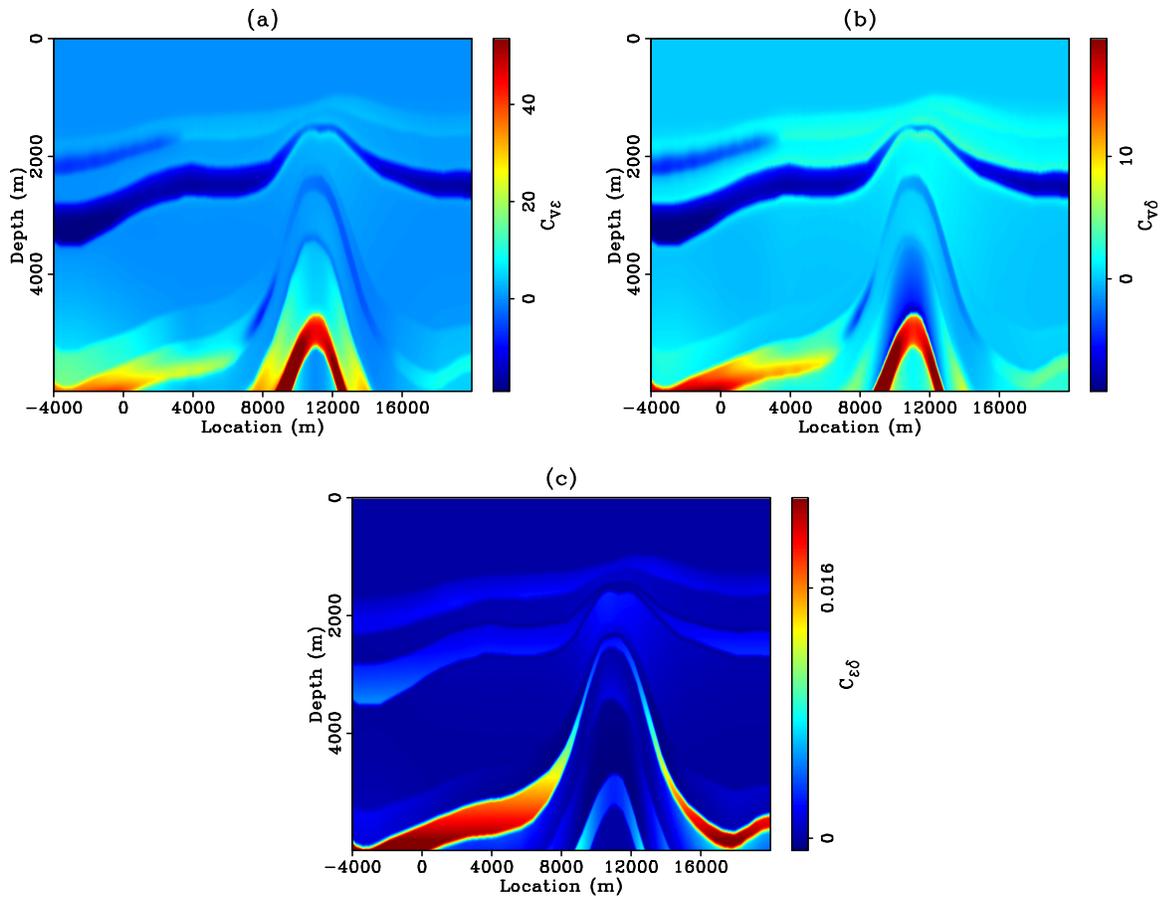


Figure 3.7: Off-diagonal elements of the covariance matrix of the VTI model. (a) $C_{v\epsilon}$: map of covariance between velocity and ϵ , (b) $C_{v\delta}$: map of covariance between velocity and δ , and (c) $C_{\epsilon\delta}$: map of covariance between ϵ and δ . [ER]

chap3/. cvn-stdv-offdiag

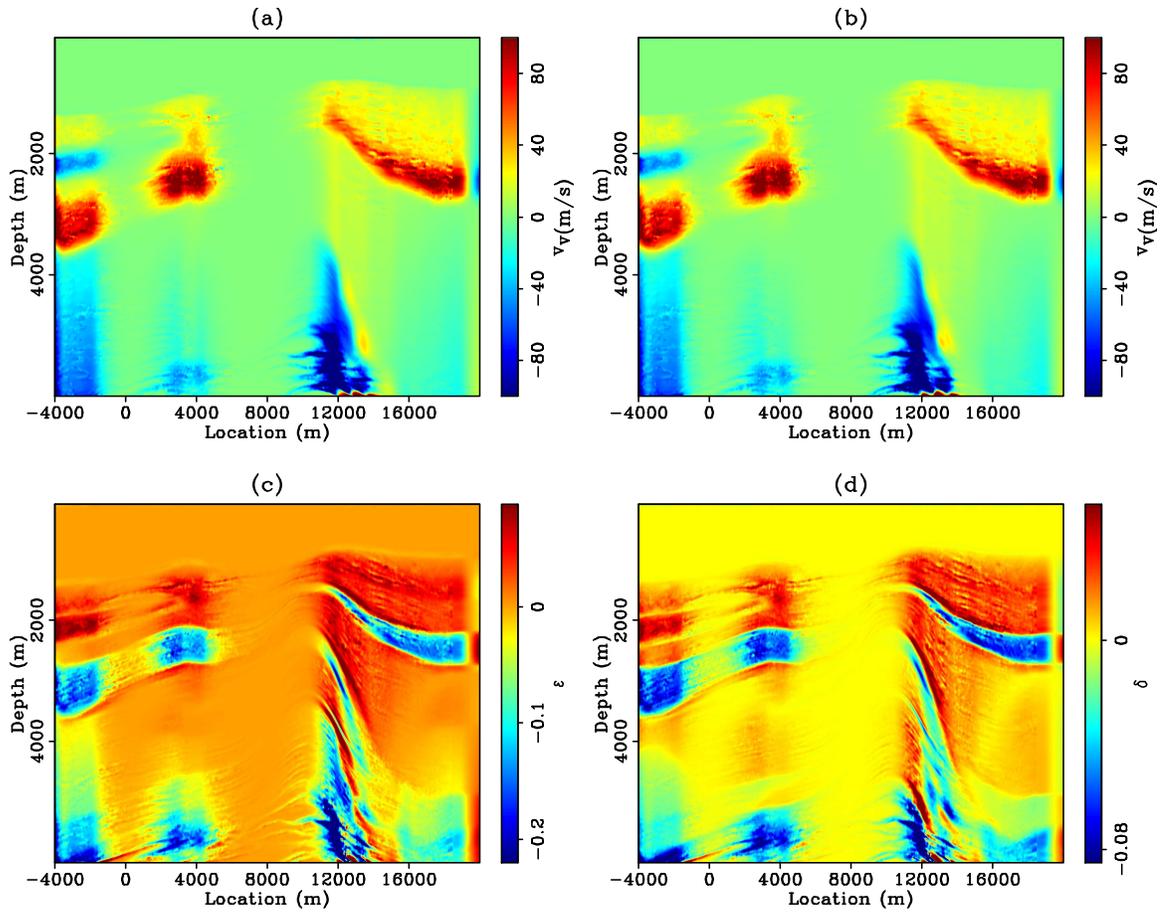


Figure 3.8: Comparison of the interpolation results. (a): Inverted velocity using $C_1^{1/2}$. (b): Inverted velocity using $C_2^{1/2}$. (c): Inverted ϵ using $C_2^{1/2}$. (d): Inverted δ using $C_2^{1/2}$. [CR] chap3/. cvn-interp

WEMVA SYNTHETIC EXAMPLE WITH ROCK PHYSICS CONSTRAINTS

In this section, I will test the preconditioning scheme using different estimations of the covariance matrix on the modified BP2007 VTI dataset. The synthetic data are modeled with a streamer geometry where the maximum offset is 6km. Shot spacing is 100m, and receiver spacing is 12.5m. A total of 100 shots have been modeled and inverted using the preconditioned anisotropic WEMVA technique.

The initial and the true models are shown in Figure 3.3. Compared with the true models, the initial model captures the overall increasing trend of the VTI parameters. However, the layering in the models is missing. The following two important lithological layers in the shallow region (above 4km) are missing from the initial models: the first shale layer with low velocity and high anisotropy, and the third sand layer with high velocity and low anisotropy. The sand-shale inter-layering in the deep region (below 4km) is obvious along the anticline in the true model. In contrast with the shallow region, the high velocity layers in the deep section correlate with the shale layers and with high anisotropy. Because the deeper structures below 4km will be difficult to resolve due to the limited acquisition, I will focus on the discussion of the shallow sediments.

The initial migration image using the initial models is shown in Figure 3.9. There are large vertical shifts between the initial stacked image and the true stacked image. The reduced amplitudes in the deeper region on the initial stacked image indicate a weaker focusing effect compared with the true stacked image. Figure 3.10 shows the comparison of the initial ADCIGs in (a) with the true ADCIGs in (b). The upward moveouts in the initial ADCIGs indicate that the average migration velocity is lower than the true average velocity along the wavepaths. Although the true anisotropy updates in the sand layer are negative by my construction, the overwhelming effects of the slower velocity on the kinematics of the acoustic waves are predominant. In the iterative WEMVA process, these upward moveouts in the ADCIGs will translate into positive updates in both velocity and anisotropic parameters ϵ and δ .

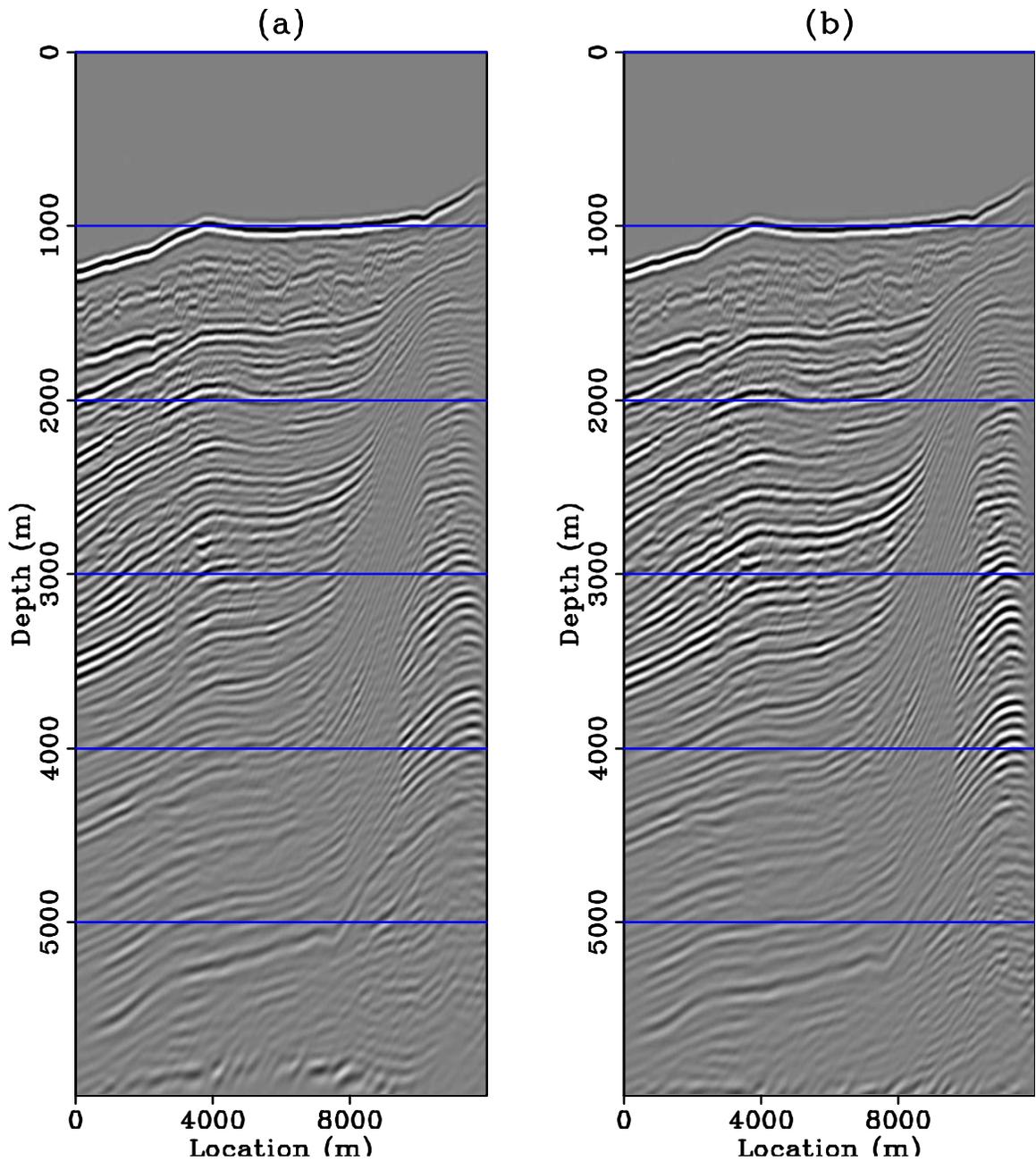


Figure 3.9: Comparison of the initial stacked image (a) and the true stacked image (b). Notice the depth shift between the images and the unfocused reflectors around $z = 3\text{km}$ and $x = 4\text{km}$. [CR] [chap3/. cvn-init-bimg](#)

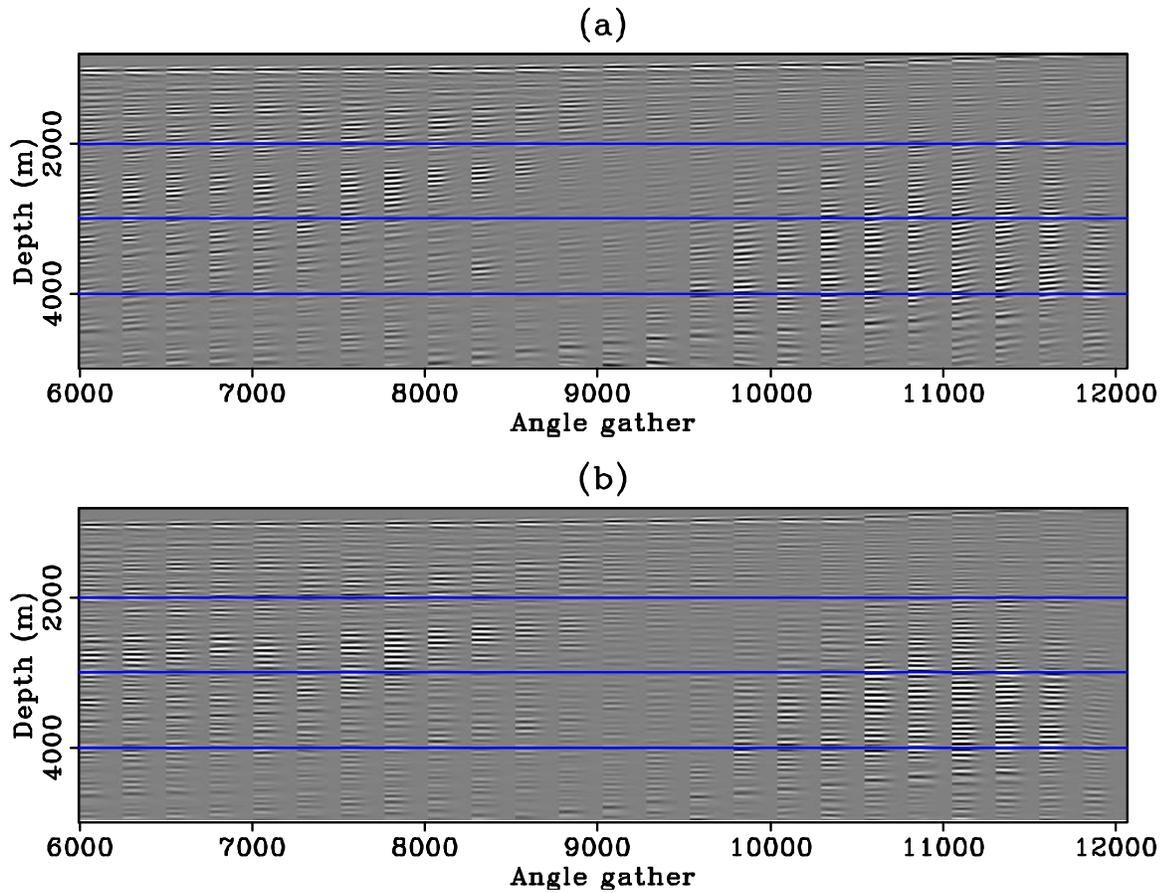


Figure 3.10: Comparison of the initial ADCIGs (a) and the true ADCIGs (b). Notice the upward moveout throughout the section in the initial gathers. [CR] `chap3/.cvn-init-aimg`

The six independent components of the covariance matrix at each point are plotted in Figures 3.6 and 3.7. Notice the negative correlation between velocity and anisotropy in the shallow region, and the positive correlation in the deep region (Figure 3.7(a) and 3.7(b)). Figure 3.7(c) shows that the positive correlation trend between ϵ and δ is largely invariant of the depth in this section.

To test the effect of different preconditioning schemes, I perform anisotropic WEMVA with three different preconditioning matrices:

$$\Sigma_1 = \begin{vmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{vmatrix}^{1/2}, \quad (3.12)$$

$$\Sigma_2 = \begin{vmatrix} C_{vv} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_{\epsilon\epsilon} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & C_{\delta\delta} \mathbf{I} \end{vmatrix}^{1/2}, \quad (3.13)$$

$$\Sigma_3 = \begin{vmatrix} C_{vv} \mathbf{I} & C_{v\epsilon} \mathbf{I} & C_{v\delta} \mathbf{I} \\ C_{\epsilon v} \mathbf{I} & C_{\epsilon\epsilon} \mathbf{I} & C_{\epsilon\delta} \mathbf{I} \\ C_{\delta v} \mathbf{I} & C_{\delta\epsilon} \mathbf{I} & C_{\delta\delta} \mathbf{I} \end{vmatrix}^{1/2}. \quad (3.14)$$

The preconditioning matrix Σ_1 indicates that no preconditioning is applied; Σ_2 indicates that a diagonal preconditioning is applied; and Σ_3 indicates that the full preconditioning approach is applied. The spatial covariance matrix for the three tests are the same.

The initial preconditioning model \mathbf{n}_0 is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \langle \mathbf{m}_0 - \Sigma \mathbf{B} \mathbf{n}_0, \mathbf{m}_0 - \Sigma \mathbf{B} \mathbf{n}_0 \rangle. \quad (3.15)$$

The gradient of the WEMVA objective function (3.1) with respect to this preconditioning variable \mathbf{n} is

$$\begin{aligned}\nabla_{\mathbf{n}}J &= \left(\frac{\partial\mathbf{m}}{\partial\mathbf{n}}\right)^*\nabla_{\mathbf{m}}J \\ &= \mathbf{B}^*\Sigma^*\nabla_{\mathbf{m}}J,\end{aligned}\tag{3.16}$$

where $\nabla_{\mathbf{m}}J = [\nabla_v J \ \nabla_\epsilon J \ \nabla_\delta J]^T$. I use a nonlinear conjugate gradient algorithm to minimize the objective function.

Figure 3.11 shows the velocity updates after the first iteration. In panel (a), I display the raw update from the moveouts on the ADCIGs. The overall update direction for velocity is positive. The vertical resolution of the updates is very low after the first iteration: a bulk positive velocity shift is indicated by the raw gradient. In panel (b), the spatial distribution of the velocity update has been modified by the diagonal elements of the covariance matrix. Because of the prior knowledge of a sand layer with high velocity perturbation, the preconditioning concentrates the strong velocity update within that layer. Compared with the true velocity update in panel (d), the velocity update with the diagonal preconditioning already achieves high vertical resolution from the first iteration. In panel (c), I utilize all the elements of the covariance matrix to precondition the gradient. The resulting update is very similar to the update with diagonal preconditioning. This similarity suggests that the off-diagonal elements of the covariance matrix have limited influence on velocity.

Figure 3.12 shows the updates in ϵ after the first iteration. In panel (a), the raw update in ϵ appears lower resolution than the raw updates for velocity. All the updates are in the positive direction, thus compensating for the upward moveouts in the initial ADCIGs. However, the updates in ϵ are very small due to its low influence to the kinematics of the waves. Although the diagonal preconditioning has changed the spatial distribution of the update with better definition of the layers, the positive values in the ϵ update in the sand layer is still in the opposite direction of the true update (Figure 3.12b). In panel (c), when the off-diagonal elements of the covariance matrix are included to account for the cross-variance between velocity

and ϵ , the preconditioning provides ϵ updates in accordance with the lithology as follows: positive ϵ updates in the shale layer and negative ϵ updates in the sand layer. Considering the significant difference between panel (b) and panel (c), I conclude that the lithological information mainly comes from the off-diagonal elements of the covariance matrix. Similar analysis and conclusions can also be made for anisotropic parameter δ (Figure 3.13).

Figure 3.14 shows the updates in velocity after 20 iterations. Panel (a) shows the velocity update when no preconditioning is applied. The spatial resolution for velocity gradually improves with iterations. However, the resolution remains significantly lower than the true update at iteration 20. More iterations could further improve the resolution, with a higher weighting on the stacking power term. However, to make a fair comparison, I stop all the tests at 20 iterations. When implementing diagonal preconditioning (panel (b)) and full preconditioning (panel (c)), the overall structure and resolution of the velocity updates remain stable with increasing iterations. This shows an early convergence. Nevertheless, the later iterations improve the definition of the thin sand layer with high velocity at $z = 1.5\text{km}$ and $x = 10\text{km}$, and the anticline structure with low velocity around $z = 4\text{km}$ and $x = 11.5\text{km}$. The final inversion results for velocity are nearly identical between the diagonal and the full preconditioning schemes. As will be shown by the final stacked image, the effective imaging velocity from the three different preconditioning schemes are very similar.

Figure 3.15 shows the updates in ϵ after 20 iterations. When no preconditioning is applied (panel (a)), the resolution of the ϵ is lower than the resolution of velocity as predicted by the topography of the objective function in the previous section. Compared with Figure 3.14(a), although the updates in ϵ in Figure 3.15(a) are mostly in phase with the updates in velocity, the positive updates on the flanks and the top of the anticline are in the correct direction despite the negative velocity updates at the same locations. When preconditioned with diagonal covariance (panel (b)), the inversion has a better resolution for the shale layer near the water bottom. However, the updates in the sand layer are in the opposite direction of the true ϵ updates. The inverted updates for ϵ are very close to the true updates when the off-diagonal

components of the covariance matrix are also included in the preconditioning (panel (c)). The inversion almost perfectly recovers both the shale and the sand layer due to the prior knowledge of the cross-correlation between the velocity and ϵ perturbation within each layer. Similar analysis and conclusions can also be made for the inverted updates in δ (Figure 3.16). Notice that on the flanks and the top of the anticline, the inversion successfully resolves the ϵ perturbation no matter which preconditioning scheme is applied. This success is in debt to the rich angular coverage in these regions, especially the wide angle imaging rays with higher sensitivity to the anisotropic parameters.

Figure 3.17 shows the inverted vertical velocity models after twenty iterations. When no preconditioning is applied (Panel (a)), the inversion only recovers low wavenumber components of the velocity model. When diagonal or full preconditioning is applied (Panels (b) and (c)), the layering with variable depths in the shallow region is very well recovered. In the deeper section, only the strong low velocity anomaly at the top of the anticline has been retrieved due to the limited angle illumination. The lack of inversion success in the deeper region suggests that powerful as the proposed preconditioning scheme is, it cannot “create” information where the seismic data has little information.

Figure 3.18 shows the inverted ϵ models after twenty iterations. Inversions without the cross-parameter correlation (Panel (a) and (b)) have barely moved the solution from the initial models. The inversion can resolve a high resolution and rock physics plausible ϵ model only when the full preconditioning scheme is applied.

Figure 3.19 shows the inverted δ models after twenty iterations. Similar to the inversion results for ϵ , the information contributed by the off-diagonal terms in the full covariance matrix provides much better spatial and rock physical constraints. Notice that in the shallow shale layers, the stronger anisotropic thin shale layers are very well resolved on Panel (c). This high resolution result is attributed to both the high resolution preconditioning and the high resolution stacking power objective function.

It is important to note the remaining ambiguities among the VTI parameters

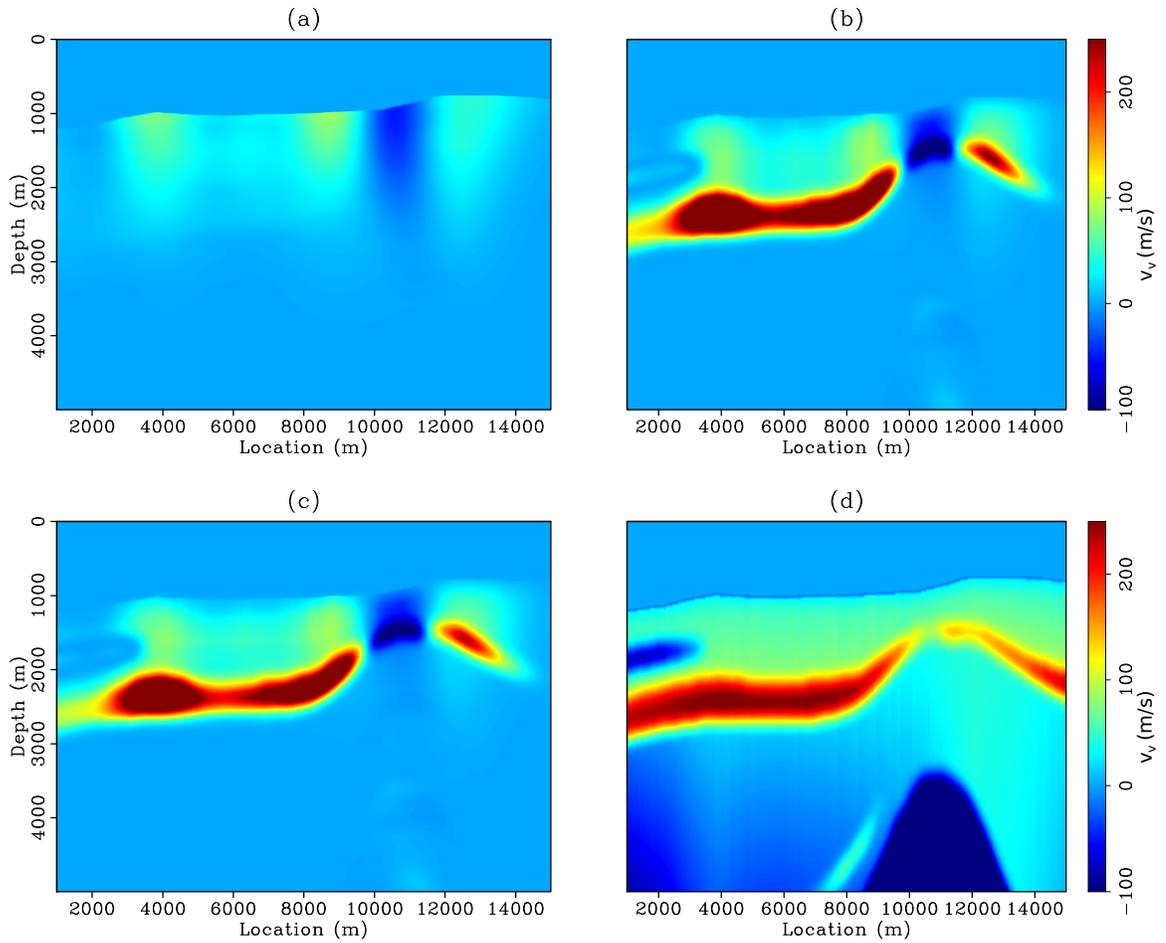


Figure 3.11: Velocity updates after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [CR] [chap3/. cvn-vupdt-2](#)

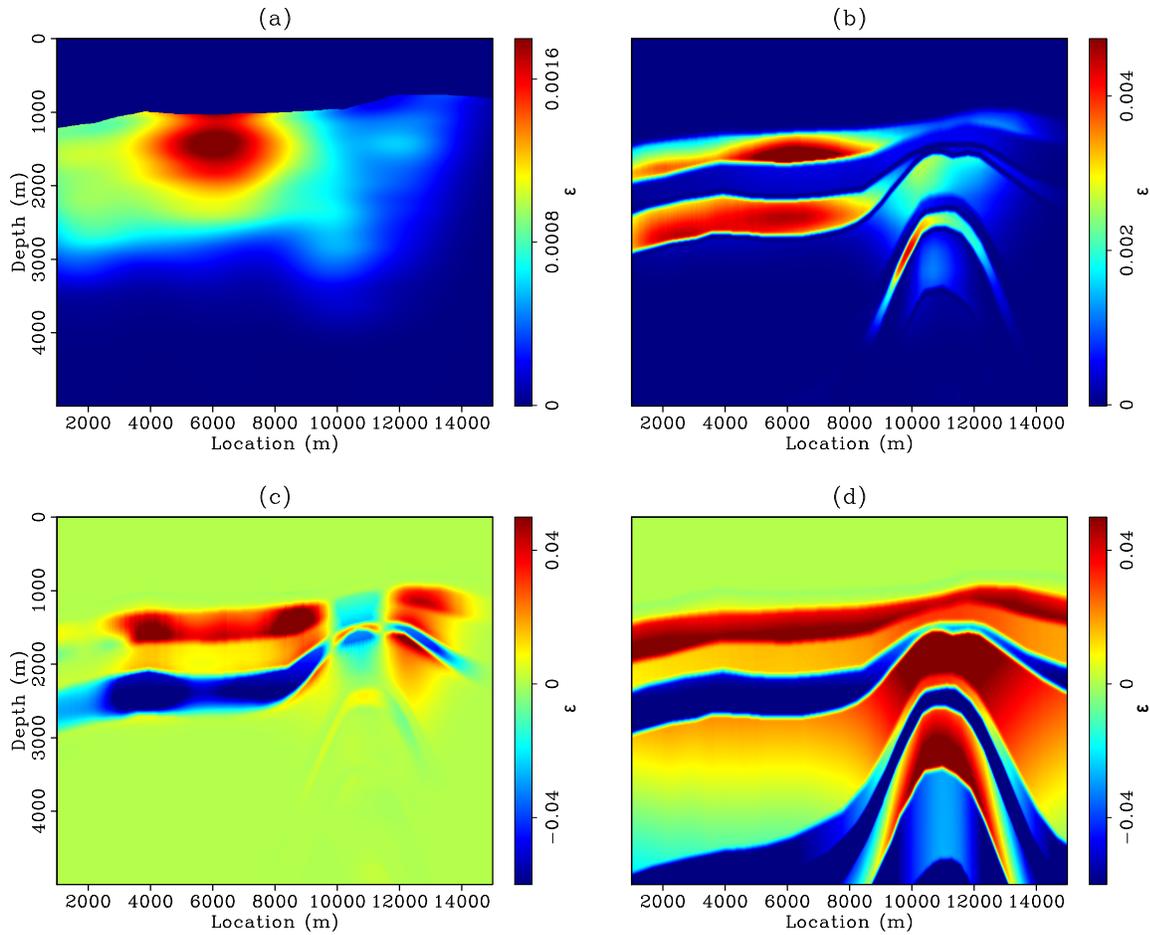


Figure 3.12: Updates in ϵ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. [CR] chap3/. cvn-eupdt-2

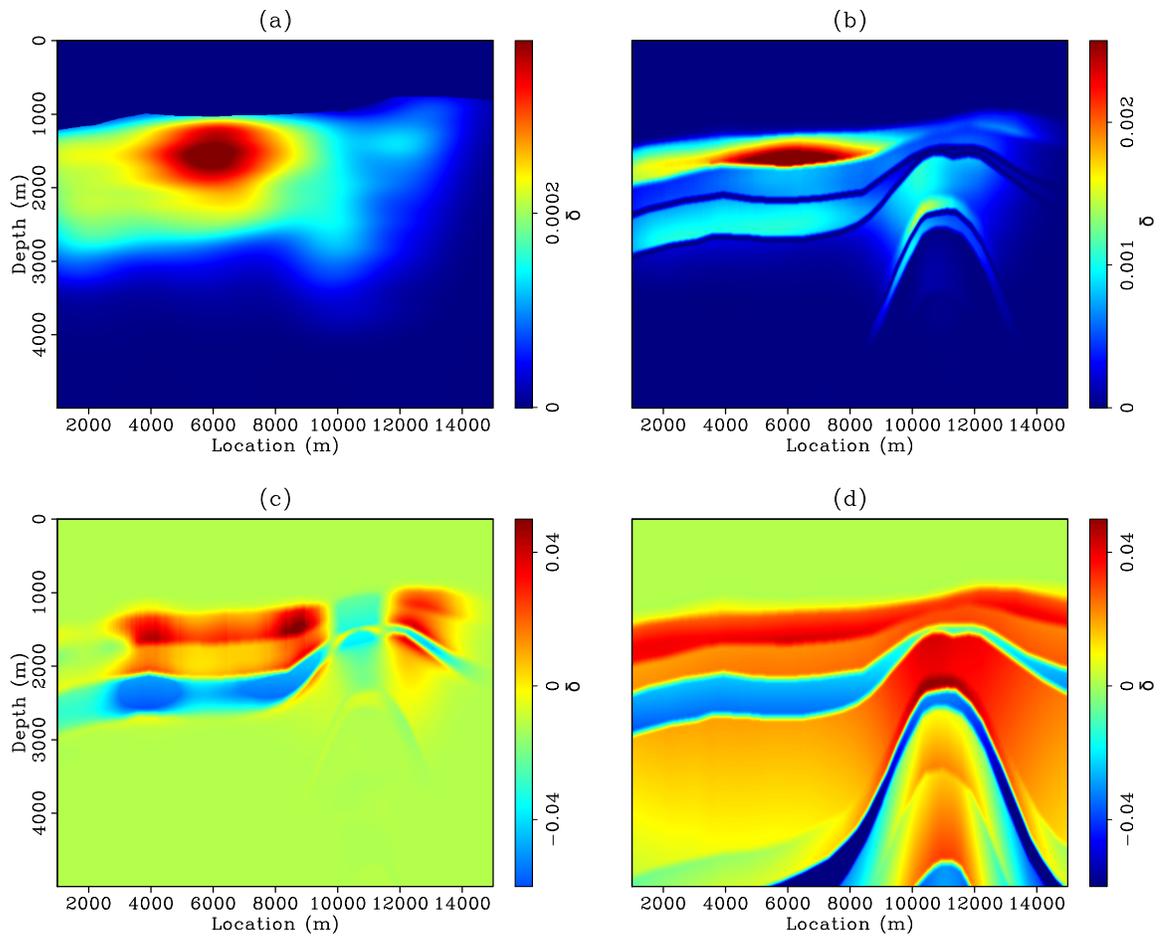


Figure 3.13: Updates in δ after the first iteration with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. [CR] chap3/. cvn-dupdt-2

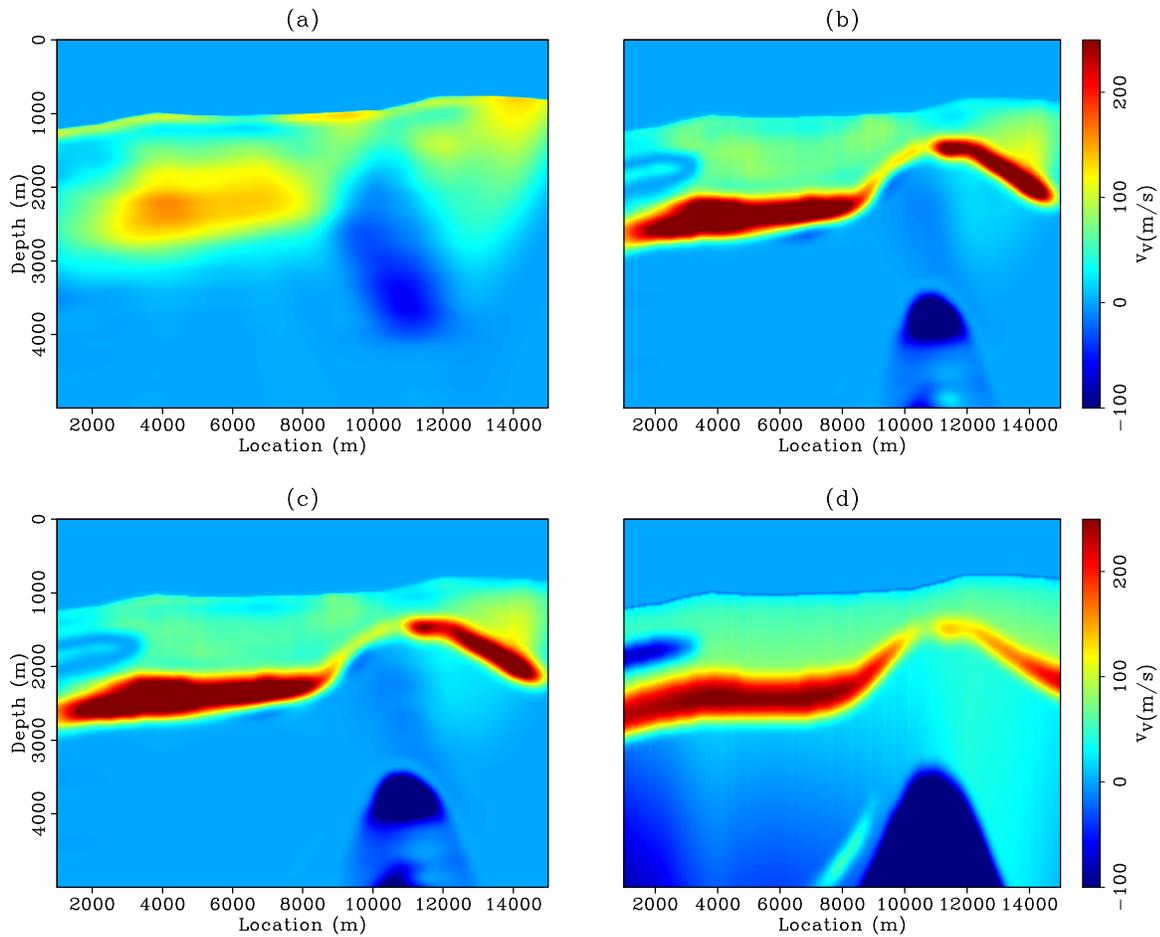


Figure 3.14: Velocity updates after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true velocity updates for comparison. [CR] chap3/. cvn-vupdt-21

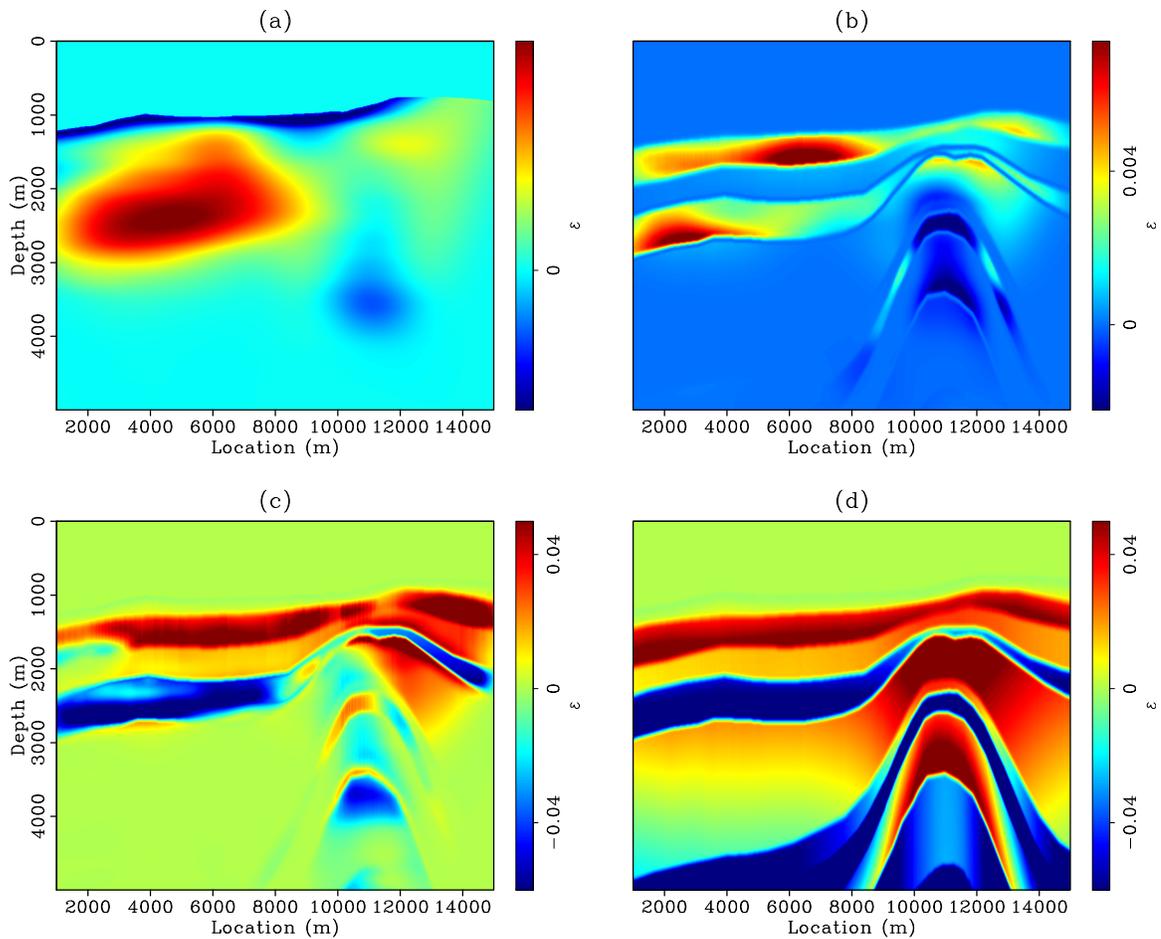


Figure 3.15: Updates in ϵ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ updates for comparison. [CR] chap3/. cvn-eupdt-21

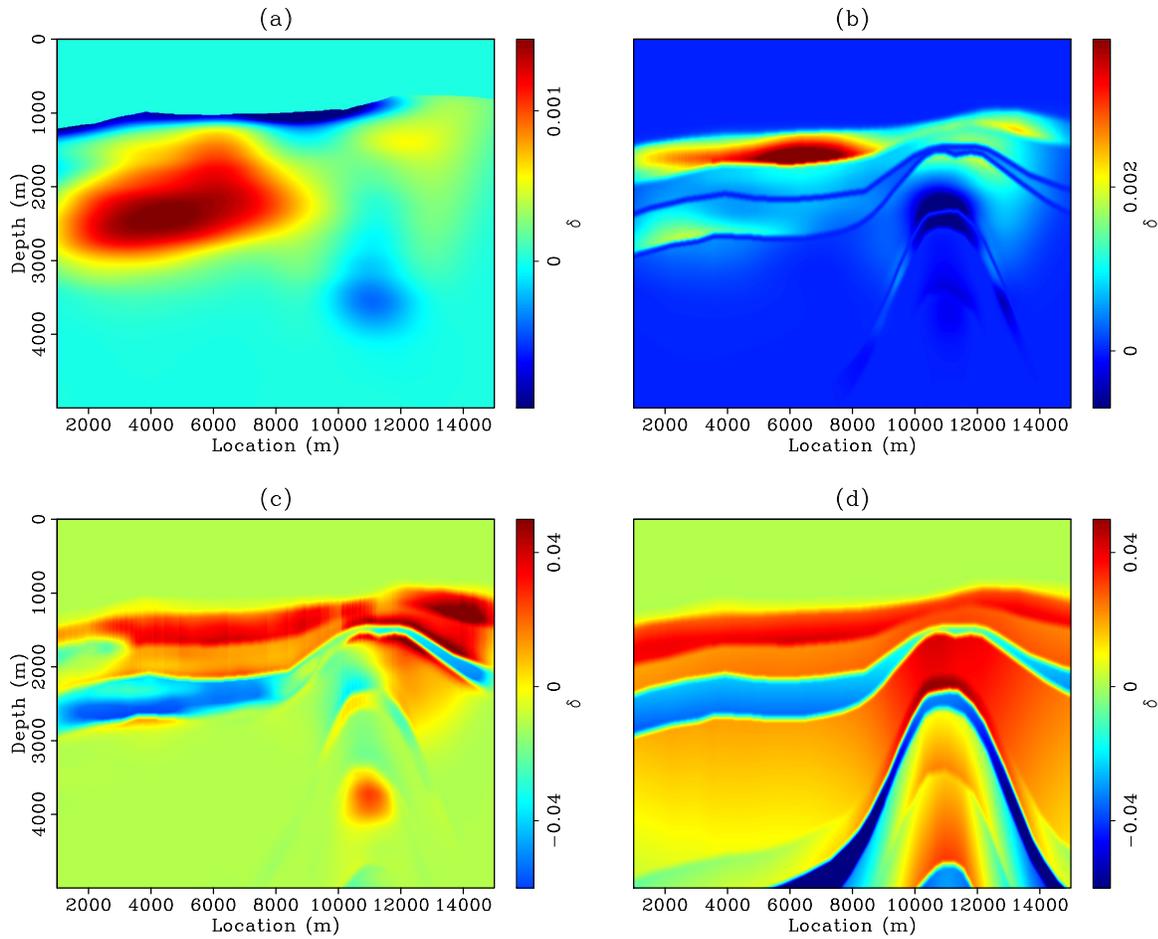


Figure 3.16: Updates in δ after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ updates for comparison. [CR] chap3/. cvn-dupdt-21

even when the full preconditioning scheme is used. One example is highlighted by the ellipses in the figures. In Figure 3.17(c), a low velocity wedge can be seen when compared with the true velocity model in Figure 3.17(d). In the corresponding region, a thinner isotropic sand layer has been observed from the inverted anisotropic parameters (Figures 3.18(c) and 3.19(c)) when compared with the true models (Figures 3.18(d) and 3.19(d)). The low velocity anomaly and the high anisotropic anomalies, although of different shapes, compensate each other and focus the seismic data associated with this area.

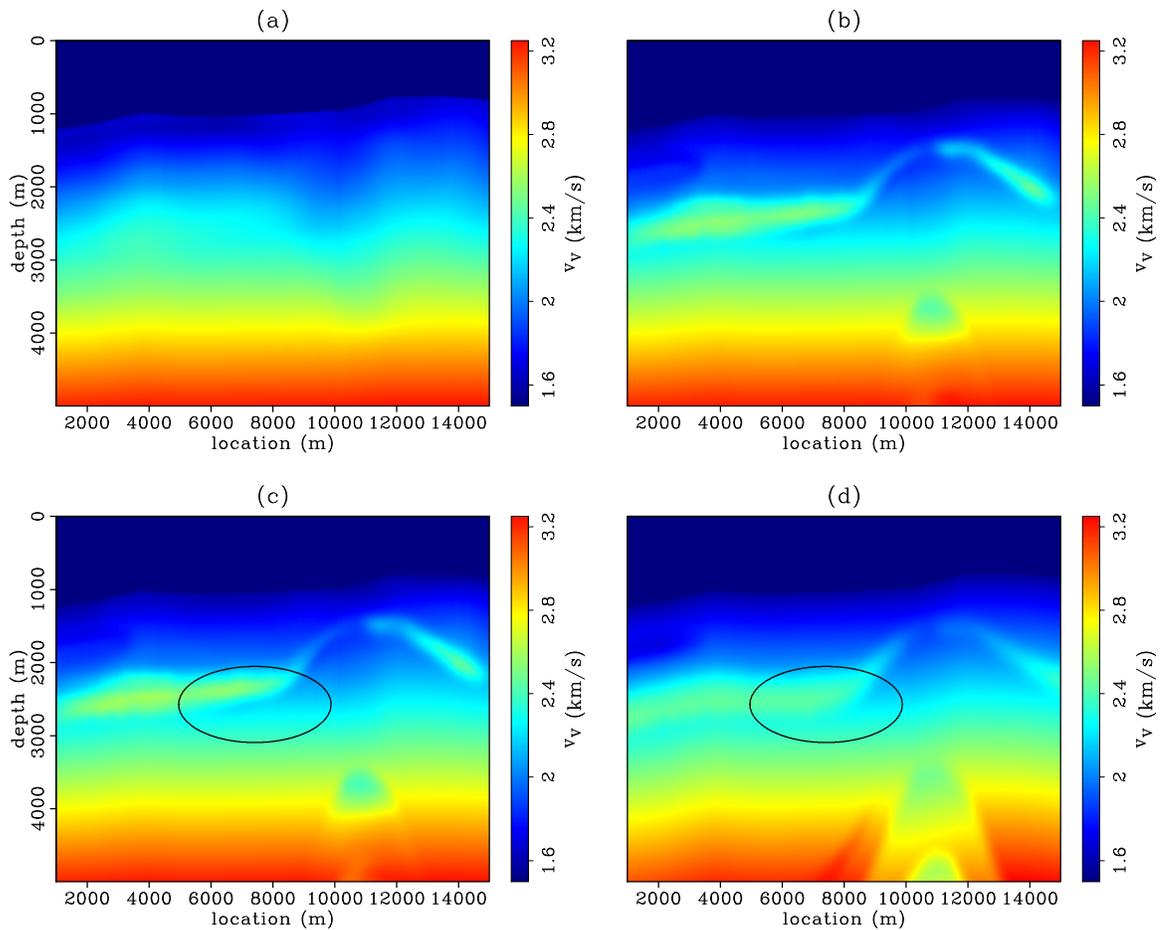


Figure 3.17: Inverted vertical velocity model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true vertical velocity model for comparison. [CR] `chap3/. cvn-final-vp`

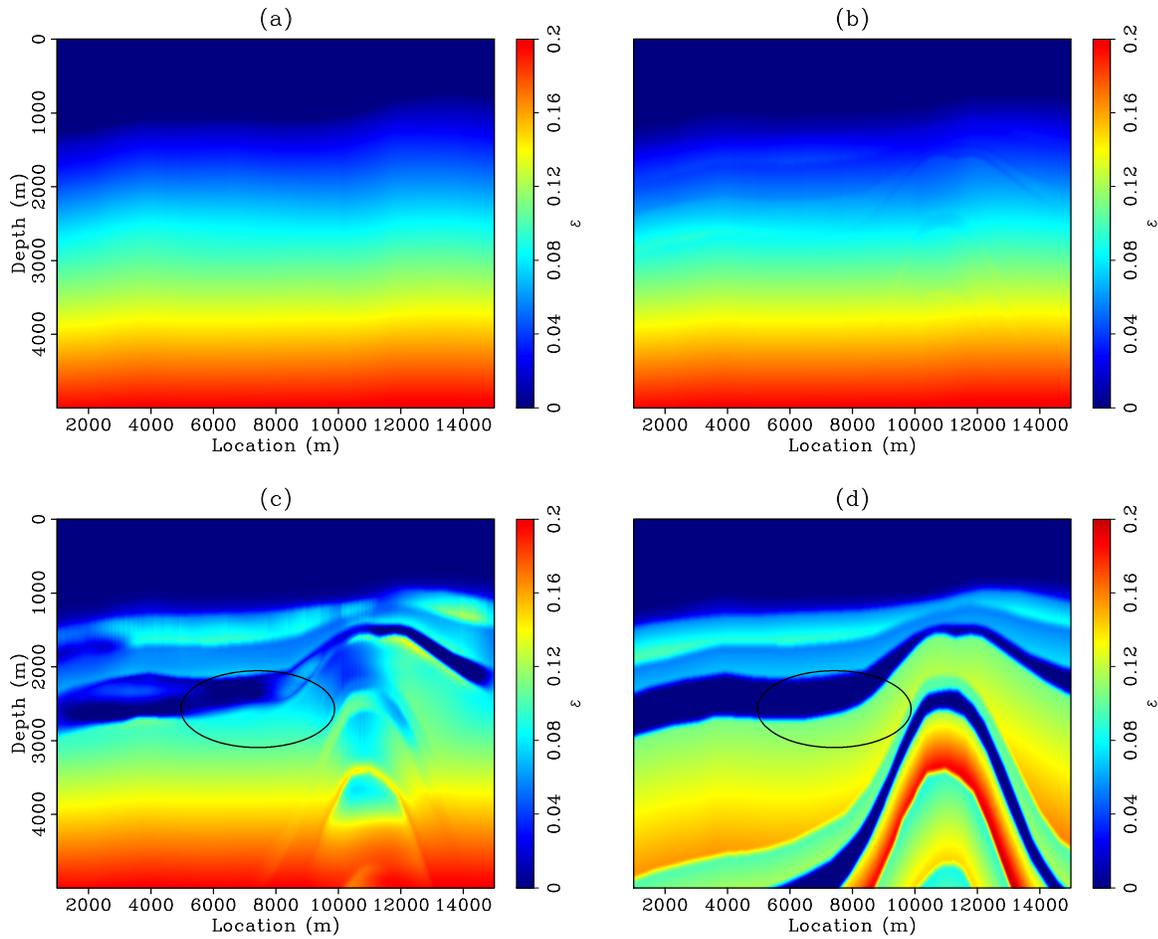


Figure 3.18: Inverted ϵ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true ϵ model for comparison. [CR] chap3/. cvn-final-eps

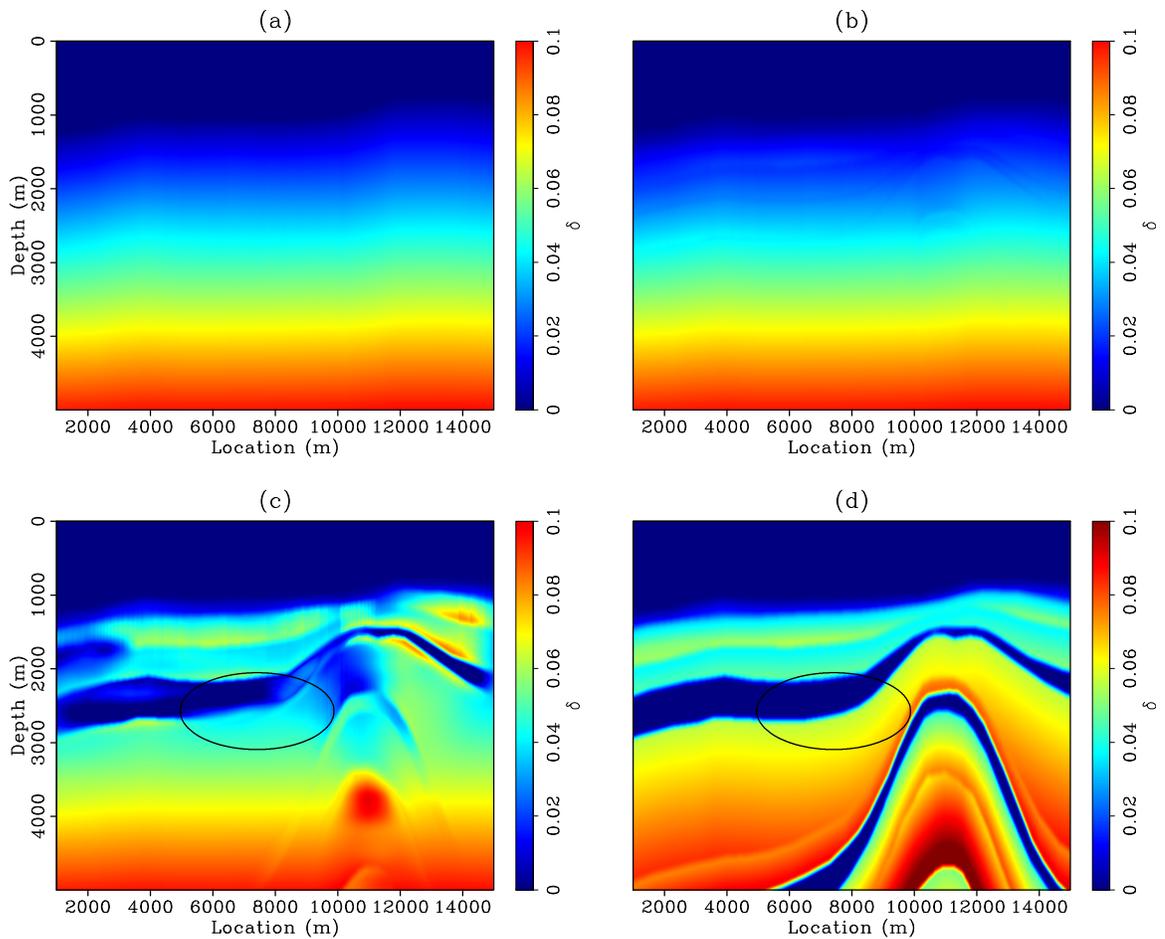


Figure 3.19: Inverted δ model after twenty iterations with (a) no preconditioning, (b) diagonal preconditioning and (c) full preconditioning. Panel (d) shows the true δ model for comparison. [CR] chap3/. cvn-final-del

Figure 3.20 shows the comparison between the final stacked images produced by different preconditioning schemes with the true stacked image. All three inverted images are greatly improved from the initial stacked image in Figure 3.9(a) with better focusing and better defined depths of the reflectors. Despite the significant differences in the inverted velocity and anisotropic models from three preconditioning schemes, the effective imaging velocities along the wavepaths are very similar.

Figure 3.21 shows the comparison of the final ADCIGs with different preconditioning schemes. Compared with the initial ADCIGs in Figure 3.10(a), the inversion flattens the angle domain events no matter which preconditioning scheme is applied. The three inversion tests give very similar results in the center of the section between $x = 7.5\text{km}$ and $x = 10.5\text{km}$ where the reflectors are illuminated by the data with rich angles. When the subsurface is less well illuminated by the data, for example in the regions highlighted by the circles, the full preconditioning scheme produces flatter events, stronger amplitudes, and higher angle coverage.

The WEMVA data fitting objective function can be evaluated at each imaging point. I compare the initial objective function map with the inverted objective function map with different preconditioning schemes in Figure 3.22. The warmer color indicates higher data fitting error, whereas the cooler color indicates lower data fitting error. In the initial objective function map, the error is small in the shallow region (above 2.5km) but it gets stronger with depth. As shown by the inverted objective function map, the data fitting error has been significantly reduced in all three cases, especially in the regions between $x = 1.5\text{km}$ and $x = 4\text{km}$ and between $x = 10\text{km}$ and $x = 12\text{km}$.

To compare the objective function maps quantitatively (the visual comparison is difficult due to the similarity), I subtract the objective function map in Figure 3.22(d) from Figure 3.22(b) and plot the difference in Figure 3.23(a). Similarly, the difference between Figures 3.22(d) and 3.22(c) is plotted in Figure 3.23(b). The region in red denotes the area where the inversion with full preconditioning fits the data better than the other two tests, and the region in blue denotes the reverse. The inversion with full preconditioning does not guarantee a better fit for the data globally, but

it does show better fits locally (Figure 3.22). Furthermore, the differences among different preconditioning schemes increase with depth as the data constraint becomes weaker with depth.

Figure 3.24(a) plots the value of the data fitting objective function (sum of the objective function map in Figure 3.22) as a function of iteration for three different preconditioning tests. The objective function value is normalized with respect to the objective function value of the image migrated with the true models. All three inversion tests reduce the objective function value to about 30% of the initial value. During the 20 iterations, the objective function values with the full covariance matrix are almost always the lowest among three tests; however, the differences among the three objective values at each iteration are small. The curvature of the objective function curves suggests that the inversion tests with diagonal or full preconditioning have converged at iteration 20, whereas the inversion test without any preconditioning may need more iterations to fully converge.

Figure 3.24(b) plots the normalized length of the gradient as a function of iteration for three inversion tests. Although not guaranteed to be monotonically decreasing, the length of the gradient should have a decreasing trend and should approach zero as the inversion converges, as shown by the curves with diagonal or full preconditioning. On the other hand, the gradient when no preconditioning is applied still has significant sizes at the last few iterations. This confirms with the objective function curve (magenta line in Figure 3.24(a)) that more iterations may be needed to improve the convergence.

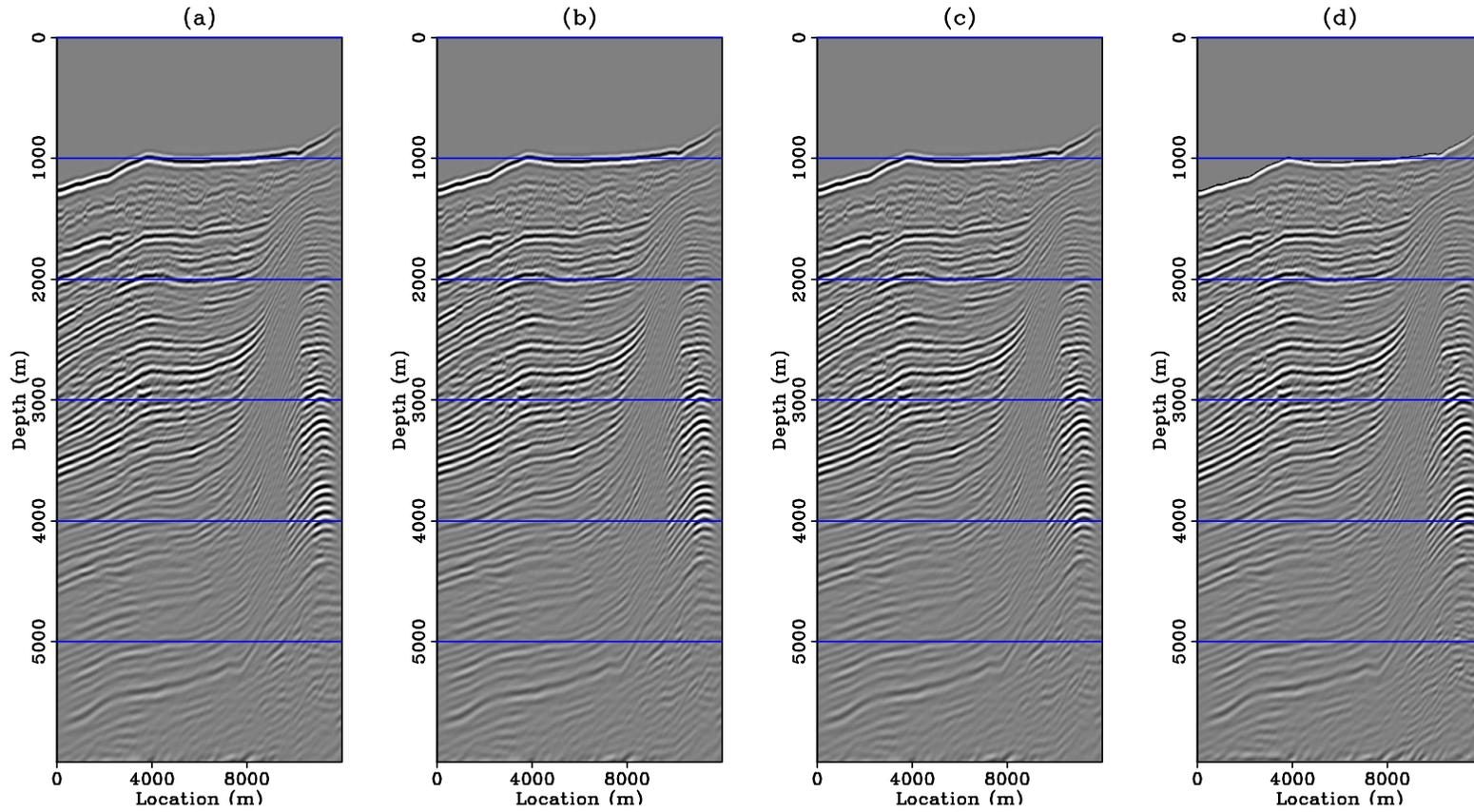


Figure 3.20: Final stacked images with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Panel (d) shows the true stacked image. [CR] chap3/. cvn-imag-21

DISCUSSION

The nonlinear and underdetermined nature of the anisotropic model building problem leads to multiple VTI models that fit the same surface seismic data equally well. These models are referred to as equiprobable models (Yang et al., 2012; Osypov et al., 2008). The key element in this chapter is the inclusion of the covariance matrix among the VTI parameters so that these equiprobable models can be differentiated from the perspective of the geological and lithological environment.

At first glance, one may have concerns about the reliability of the “extra” information from the preconditioning. However, our tests using different preconditioning on the synthetic example show that if the VTI parameters are well constrained by the data in a certain region, different preconditioning schemes will not change the solutions to those parameters when the convergence has been achieved in all the inversion tests. For those strongly constrained parameters, proper preconditioning simply improves the convergence and enhances the resolution. The “extra” information mostly affects the weakly constrained parameters, especially in the poorly illuminated region. In these cases, the solutions to the model building problem can be dramatically different. It is possible that the rock physics model is inaccurate, which eventually leads to erroneous VTI models in the less constrained region. However, the inverted VTI models are always consistent with the input rock physics information. Therefore, quick evaluation and modification of the models are possible when more accurate rock physics information is available.

As shown by the inversion results, even when the exact full preconditioning was applied in the tests, ambiguities among the VTI parameters cannot be completely resolved. Two potential reasons may explain the residual ambiguity. First, the preconditioning scheme proposed in this chapter utilizes the multivariate Gaussian distribution assumption among the VTI parameters. When the VTI parameters do not follow the Gaussian distribution, the proposed formulation may not provide a sufficient description of the correlation among them. Moreover, the covariance matrix among the VTI parameters is estimated based on the current VTI model and/or a

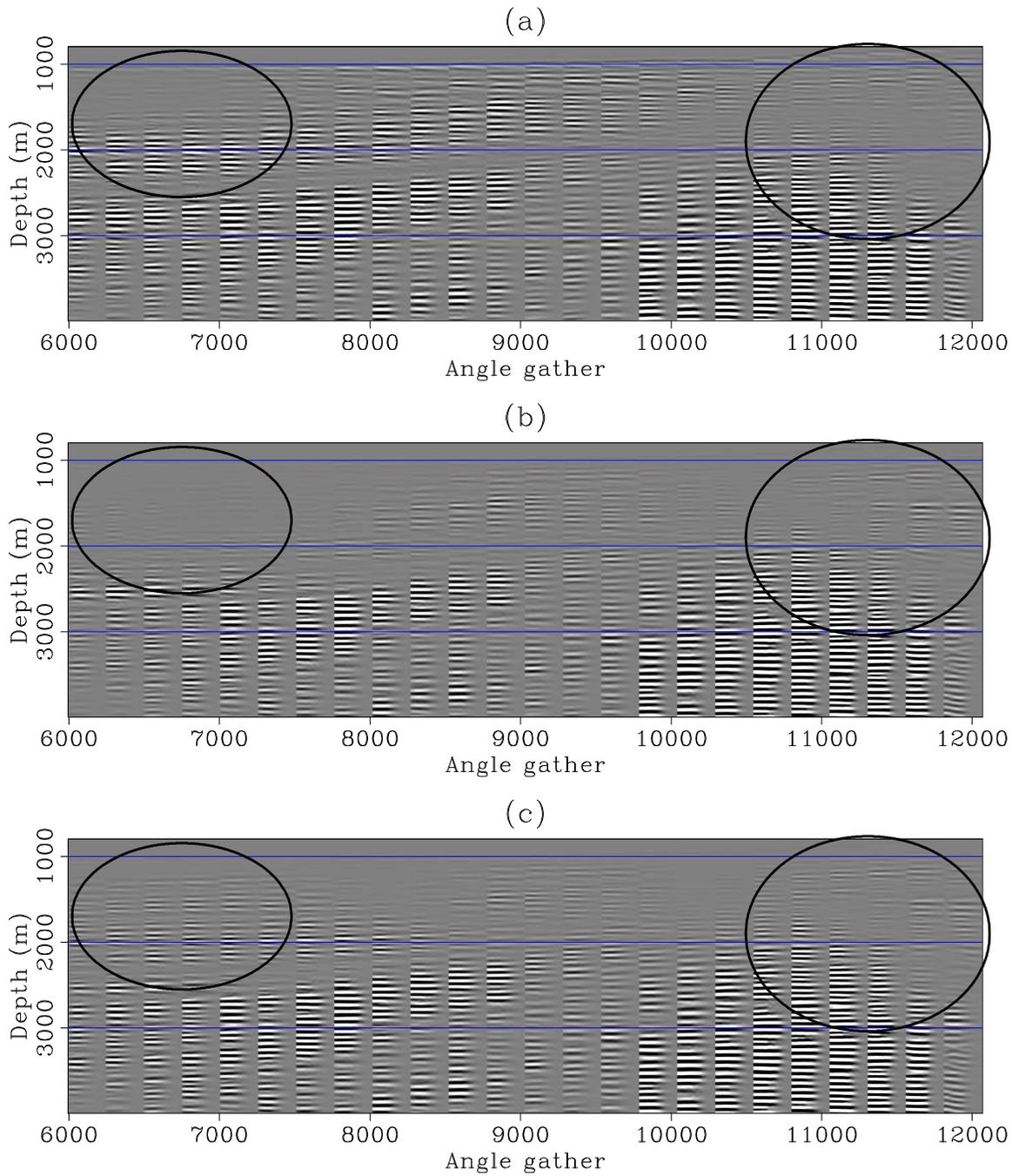


Figure 3.21: Final ADCIGs with (a) no preconditioning, (b) diagonal preconditioning, and (c) full preconditioning. Notice the improved angle coverage in the ellipse on the left and the improved flatness on the right. [CR] chap3/. cvn-fnl-aimg

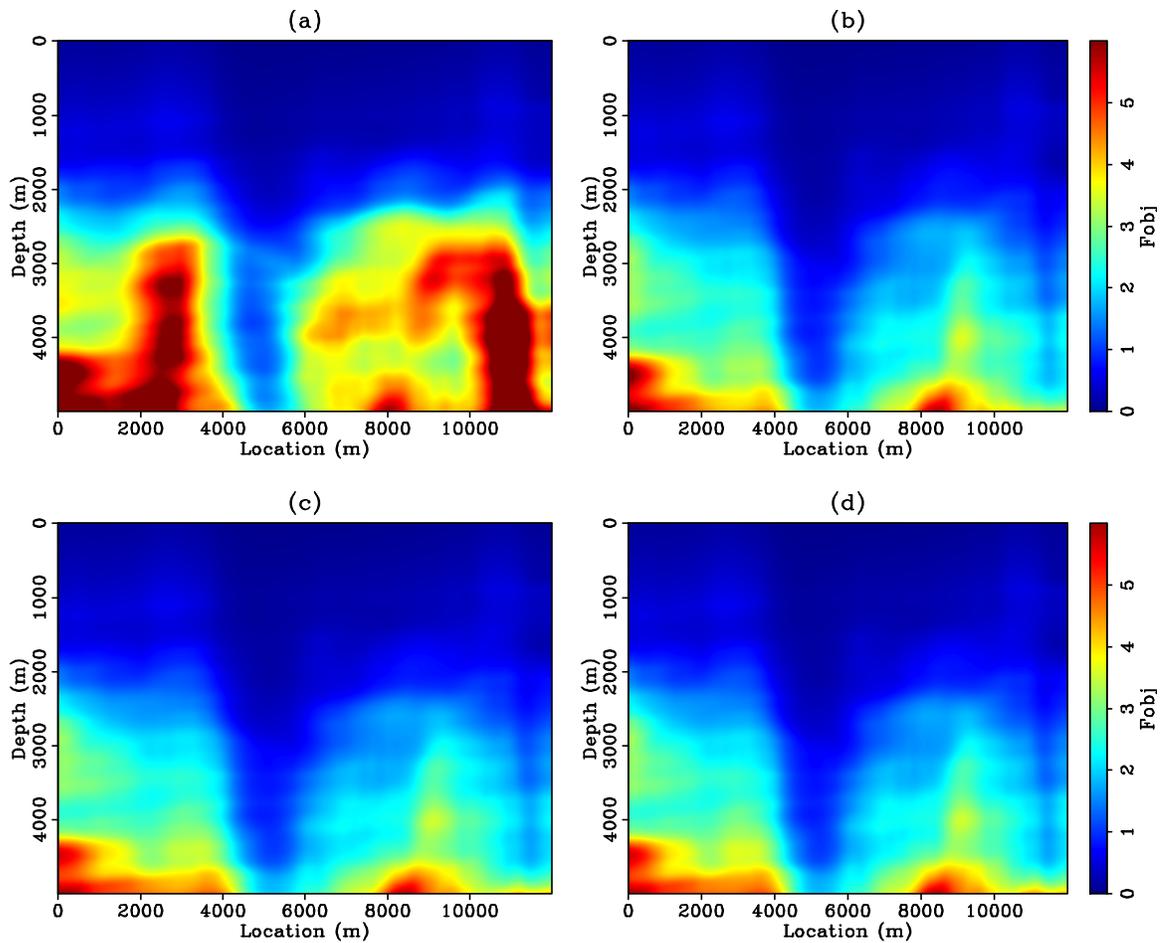


Figure 3.22: Comparison of the initial objective function map (a), with the final objective function map with (b) no preconditioning, (c) diagonal preconditioning, and (d) full preconditioning. [CR] chap3/. objmap

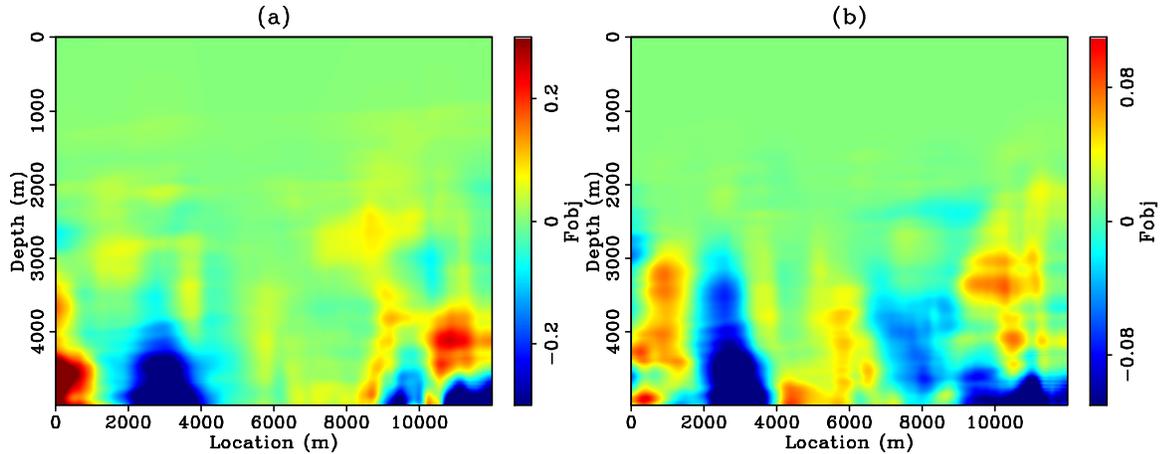


Figure 3.23: Differences between the objective function maps. (a) Difference between Figure 3.22(d) and 3.22(b). (b) Difference between Figure 3.22(d) and 3.22(c). [CR] chap3/. objmapdiff

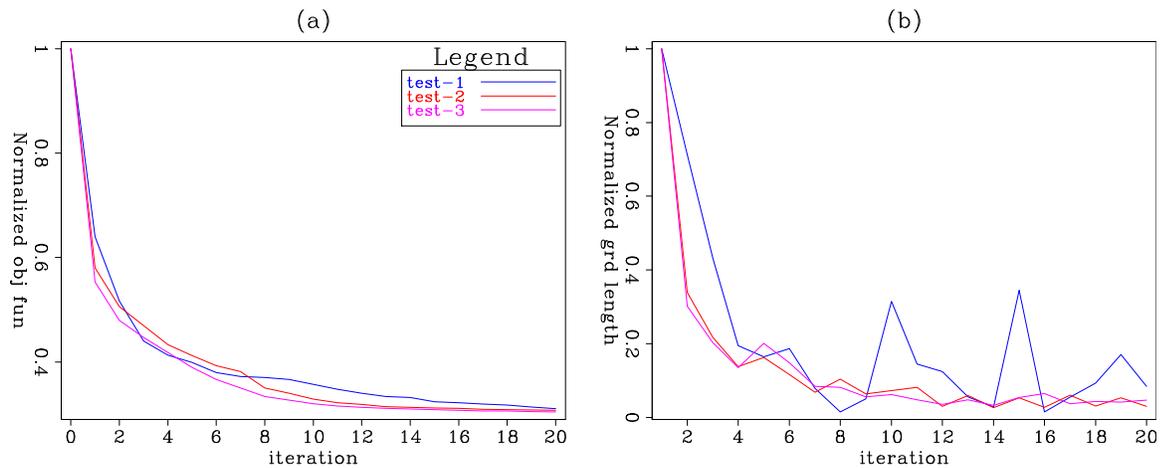


Figure 3.24: Objective function value of the data fitting goal in (a) and the size of the gradient in (b) as a function of the iteration. Blue line: no preconditioning; red line: diagonal preconditioning; magenta line: full preconditioning. [CR] chap3/. converge

previous lithological interpretation. The conditional probability distribution might have been changed as the inversion updates the VTI parameters. Therefore, reevaluation of the covariance matrix might be necessary after a few nonlinear iterations to better mitigate the ambiguities.

CONCLUSION

This chapter presents a preconditioned anisotropic WEMVA scheme to better constrain the anisotropic model building process. The proposed preconditioning method includes lithological information in order to guide the inversion towards a plausible geological and rock physical solution. Numerical examples on a 2-D synthetic dataset show that when proper preconditioning is applied during inversion, the inversion achieves the best resolution from the first iteration. By utilizing the cross-correlation between the VTI parameters, the inversion also correctly resolves the less constrained anisotropic parameter. Therefore, a properly preconditioned WEMVA inversion provides a reliable tool for anisotropic model building.

ACKNOWLEDGEMENTS

The author thanks BP for providing the synthetic anisotropic model. I acknowledge the Stanford Center for Computational Earth and Environmental Science for providing computing resources.

Chapter 4

3D Field Data Test

In this chapter, I test the rock physics constrained anisotropic WEMVA methodology on a 3-D Gulf of Mexico dataset. Based on the well logs and the previously inverted lithological interpretations, I perform stochastic rock physics modeling to sample the possible ranges of the anisotropic parameters. These modeling results are then summarized by an average model and a cross-parameter covariance matrix under the multivariate Gaussian assumption. When inverting the surface seismic data using the anisotropic WEMVA method, I start from the average model and regularize the inversion using the geological dips and the cross-parameter covariance. The inverted VTI model simultaneously explains the surface seismic data (minimizes the anisotropic WEMVA objective function), follows the geological structures, and is consistent with the rock physics principles. The migration image based on the inverted models shows higher resolution and clear definition of the dipping sedimentary layers around the salt in the shallow region. The improved VTI model also yields a better focused image at depth.

INTRODUCTION

Earth model building is an underdetermined and hence challenging inverse problem, especially at the exploration stage. In general, the input information can be obtained

from geological information, surface seismic data, and well log measurements (Woodward et al., 2014). Surface seismic data have the best compromise between accuracy and coverage among the three types of information. Geological knowledge covers large regions, whereas well logs provide accurate high resolution information only at sparse locations. Therefore, most of the current information integration practices occur after seismic imaging and structural interpretation. First, seismic images are stretched vertically according to the well markers. Then, borehole analysis is extrapolated from the well location to the rest of the region based on the seismic images and the underlying geological assumptions. This conventional workflow does not include a feedback loop to verify if the modified seismic images honor the kinematics of the seismic data. Therefore, inconsistencies may be introduced by the sequential evaluations of the data.

In this chapter, I examine a surface seismic dataset that was acquired offshore Gulf of Mexico (GoM) by Schlumberger Multiclient. Migration images based on isotropic Earth models are obtained from the seismic data and the reservoir properties are interpreted based on the well logs following a conventional data integration workflow. However, both the well logs and the lithological interpretations suggest that the sedimentary basin consists of large portion of shales, which contradicts the isotropic Earth assumption used in seismic imaging. Furthermore, the migration image manifests steep dips in both in-line and cross-line directions where the sediments were pushed up by salt intrusion. Wavepaths reflected from these structures can be highly sensitive to the anisotropic parameters. Consequently, the inconsistencies in the isotropic Earth model may lead to defocused image and/or incorrect positioning of the reflectors.

To take advantage of the complementary information in all data, I propose to integrate geological and rock physics information during the seismic inversion when building an anisotropic Earth model. As introduced in the previous chapter, I include the geological information as dip filters and rock physics information as cross-parameter covariances (Li et al., 2011; Woodward et al., 2014; Zdraveva et al., 2014). By evaluating all objectives simultaneously, I hope to resolve anisotropic Earth models that

are consistent with all the available data.

Another challenge of the field application is the computational cost associated with the large volume of data. To overcome this difficulty, various data reduction methods have been proposed. Whitmore (1995); Shen and Symes (2008) and Tang et al. (2008) proposed to assemble point sources into a smaller number of plane-wave sources. However, this strategy leaves artifacts in the model space which can translate into false model updates by certain sensitive WEMVA objective functions. Guerra et al. (2009) proposed to select isolated image points in the subsurface to create a smaller dataset and evaluate the accuracy of the model in a certain targeted area. However, the resolution of this method is limited by the density of the selected image points. Large number of densely sampled image points with corresponding subsurface gathers are needed for higher resolution, which increases the computational cost. Tang and Biondi (2011) proposed to model a smaller dataset using the Born modeling method with the prestack subsurface offset domain common image gathers to maintain the velocity information in the original dataset. This method is effective in reducing the data size and hence the computational cost when the target area is significantly smaller than the original study area. It assumes a correct velocity model above the target area. At the exploration stage, a reliable anisotropic model is needed for the whole area and the initial model is often far from accurate (isotropic vs. anisotropic). Therefore, this target-oriented method cannot be applied. Instead, I take advantage of the redundancy of the data when computing the gradient by randomly removing two thirds of all shots in each iteration (Friedlander and Schmidt, 2012). We show that the gradients obtained by the reduced number of shots contain the same update information with limited under-stacked noise and unbalance amplitude effects. These artifacts can be suppressed by changing the random set of shots through WEMVA iterations.

This chapter contains two parts. In the first part, I perform stochastic rock physics modeling based on the well log measurements and the interpretation results output from a previous seismic processing workflow. This exercise enables us to explore the possible range of the anisotropic models. I obtain the background (averaged)

anisotropic model to initialize the WEMVA inversion and the cross-parameter covariance matrix to regularize the inversion.

In the second part, I invert a subset of the 3-D GoM seismic data set based on the model and the constraints from the stochastic rock physics modeling. Inversion results show improvements with higher resolution and better definition of the sedimentary structure around the salt.

STOCHASTIC ROCK PHYSICS MODELING FOR SHALE ANISOTROPY

Many rock physics models have been proposed to describe shale anisotropy, considering both the intrinsic mineral anisotropy and the particle alignment during compaction. In this thesis, I combine the rock physics models proposed by Bachrach (2010a) and Bandyopadhyay (2009). I model a sedimentary rock using two distinctive rock components: the anisotropic shale and the isotropic sand. Shale anisotropy comes from three aspects: intrinsic anisotropy from smectite and illite, the compaction effect on the mineral alignment, and the transition from smectite to illite due to compaction and temperature. I model the sand using pure quartz, which is assumed isotropic. To combine the sand component with the shale component, I consider two effective medium models: the suspension model and the lamination model. In the suspension model, sand (quartz) is modeled as inclusions in the anisotropic shale background. In the lamination model, sand and shale are modeled as a laminated system.

Details of the rock physics modeling are presented in Appendix B. I summarize the process of anisotropic rock physics modeling in Algorithm 2. To model the rock physics properties, models for earth temperature, porosity and shale percentage are needed. I estimate these models from both well logs and previous lithological inversion results. A few key parameters are also needed. These parameters are highly uncertain even with well log measurements. Therefore, I capture the uncertainties in these parameters by randomly sampling a uniform distribution for each parameter at each

rock physics realization. The ranges of the uniform distribution for each parameter are determined based on the previous studies.

By varying the key parameters of the rock physics model, I explore different shale rock scenarios with varying initial compaction states, compaction rates, temperature gradients, and temperature windows for smectite-to-illite transition. As a result, I capture the uncertainties in the rock physics modeling and obtain an ensemble of models instead of a single model. These models are the sources of the prior rock physics knowledge.

Algorithm 2 Stochastic rock physics modeling workflow

for $n = 1 \cdots N_{\text{real}}$ **do**

 Initialize the modeling parameters.

 Compute the percentage of illite in the rock given a temperature model.

 Compute the average stiffness coefficients for smectite and illite, given a porosity model.

 Compute the volumetric percentage for each of the mineral phase, given a volumetric percentage of shale.

 Compute the effective stiffness coefficients using suspension model and lamination model.

 Compute the VTI models from the effective stiffness coefficients.

end for

Figure 4.1(a) shows the input models for the rock physics modeling. I derive the shale content from gamma ray measurements by a simple linear stretch. I filter out the high frequencies of these models since they are beyond the seismic resolution. Figure 4.1(b) shows two instances of the modeling results. The first and third panels in Figure 4.1(b) show the ϵ and δ models using the lamination model. The second and fourth panels in Figure 4.1(b) show the ϵ and δ models using the inclusion model. In general, anisotropic parameters predicted by both models correlate with the shale content in the well log. We observe that different rock physics models produce greater differences in δ model than in ϵ model.

Besides the well log measurements, an initial isotropic processing work flow has been applied on the field data by Schlumberger. Seismic inversion results, such as, P-wave velocity, shale content, and porosity are provided. The comparison between the

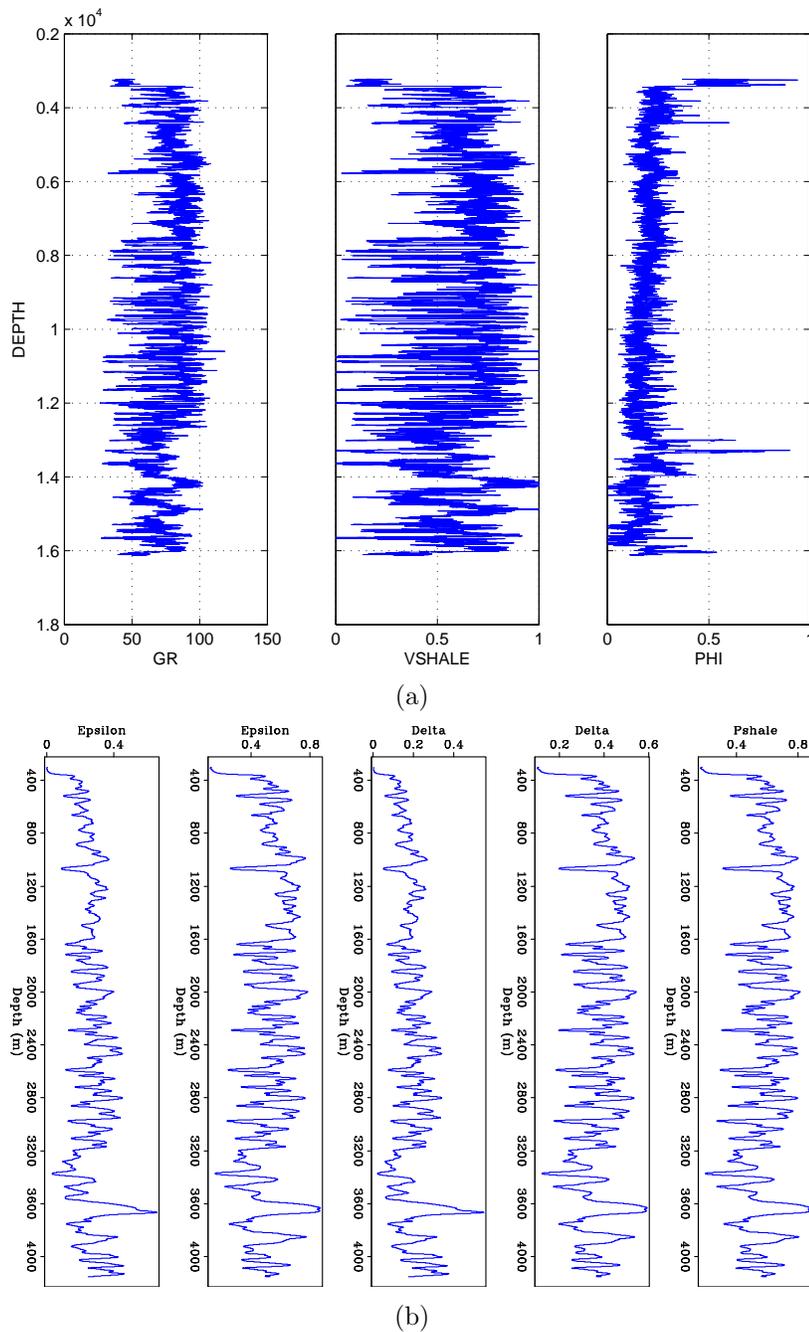


Figure 4.1: (a): Input of the rock physics modeling from well log measurements. From left to right: gamma ray measurements, deduced shale contents from gamma ray and porosity measurements. (b): One instance of the rock physics modeling experiment. From left to right, panels are ϵ profile from the lamination model and inclusion model; δ profile from the lamination model and inclusion model; shale content at the well location. [ER] chap4/. wellinput,cmpmodel

porosity inversion result and the well log measurements at the well location suggests that the porosity inversion from seismic is unreliable. The unreliability is probably due to the lack of low frequencies in the seismic data. However, the inversion result for shale content is comparable to the well log measurements. Therefore, I use the provided shale content inversion cube (Figure 4.2(a)) but model a smoothly varying porosity trend (Figure 4.2(b)) from the provided P-wave velocity. The high shale content estimates indicate that the subsurface is rich in clay minerals at all depths, although the shale content slightly increases with depth. The low porosity region below 5 km in Figure 4.2(b) highlights a shale-rich basin. The resolution of the shale content corresponds to a much larger scale (order of tens of meters) than the inclusion rock model (order of centimeters). Therefore, I use the lamination model to estimate the shale anisotropy.

Using the input in the three-dimensional space, I am able to model the anisotropic parameters in 3-D based on the rock physics models. Figure 4.2(c) and Figure 4.2(d) show an example of the ϵ and δ model. The maximum resolution of the modeled elastic constants is similar to the resolution of the interpreted shale content. Nonetheless, users can control the resolution of the anisotropic models by smoothing the modeling results. Given the chosen rock physics model, both anisotropic parameters ϵ and δ are correlated with the shale content and are inversely correlated with the porosity.

More importantly than producing a single modeling result, the stochastic rock physics modeling also allows us to explore the possible ranges of the anisotropic parameters. At a particular location in the subsurface, I obtain an ensemble of vertical velocity, ϵ , and δ values while sampling the distribution of the key parameters, as shown in Figure 4.3. Assuming the three parameters v_v , ϵ , and δ follow a multivariate Gaussian distribution, I summarize their variance with a 3×3 matrix with 6 independent elements. This cross-covariance matrix is estimated at each subsurface location.

Figures 4.4(a) to 4.4(c) show the diagonal elements in the covariance matrix. The strong lateral variations in vertical velocity variance (Figure 4.4(a)) and in ϵ variance show that their uncertainties are strongly correlated with the lithology. The

δ variance (Figure 4.4(c)) show less lateral variations, indicating that parameter δ is mainly controlled by compaction and mineral transition.

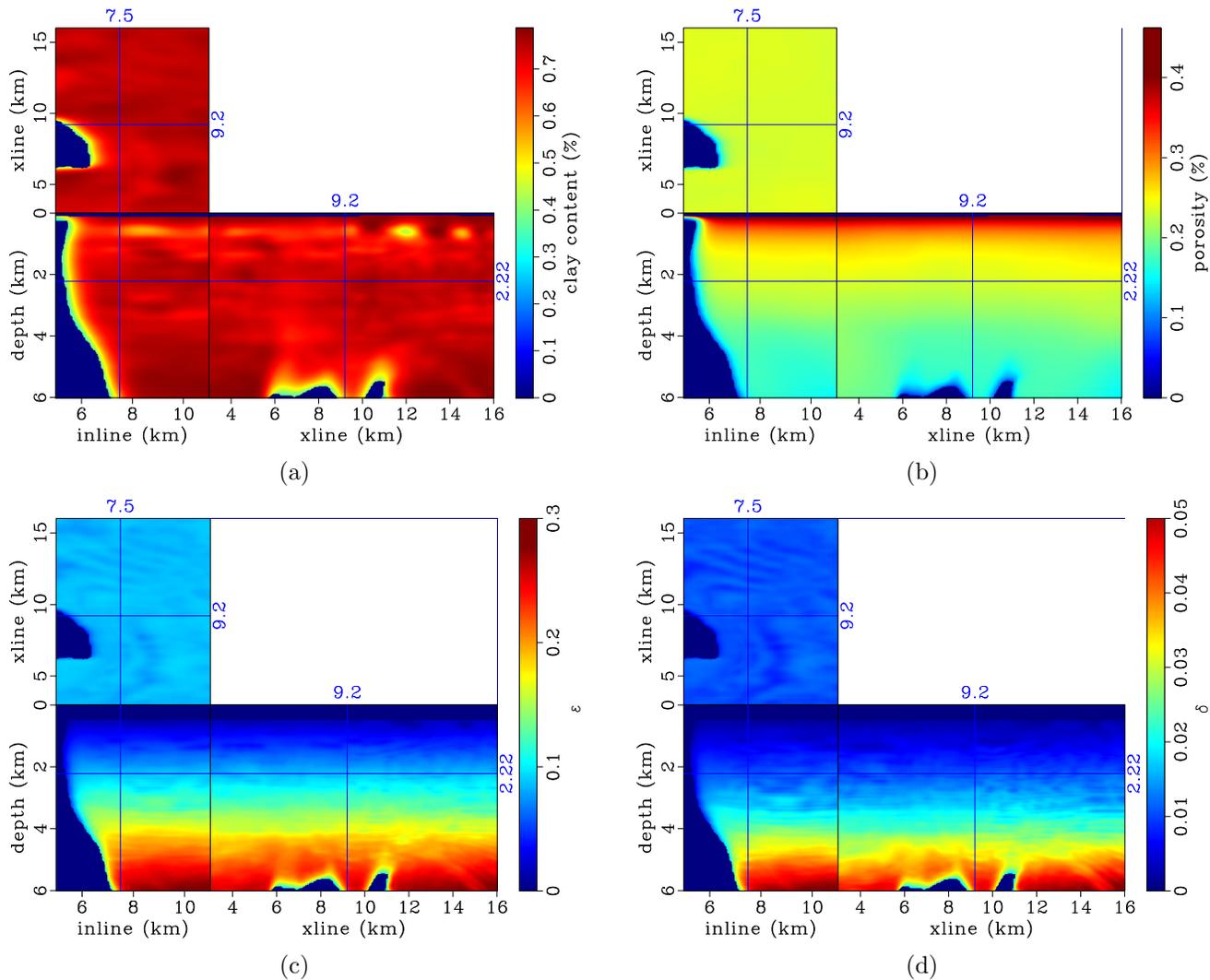


Figure 4.2: Interpreted shale content (a) and modeled porosity (b). Average ϵ model (c) and average δ model (d). Both models are obtained from the stochastic rock physics modeling experiment. [ER]

chap4/. pctlrend,phitrend,epstrend,deltrend

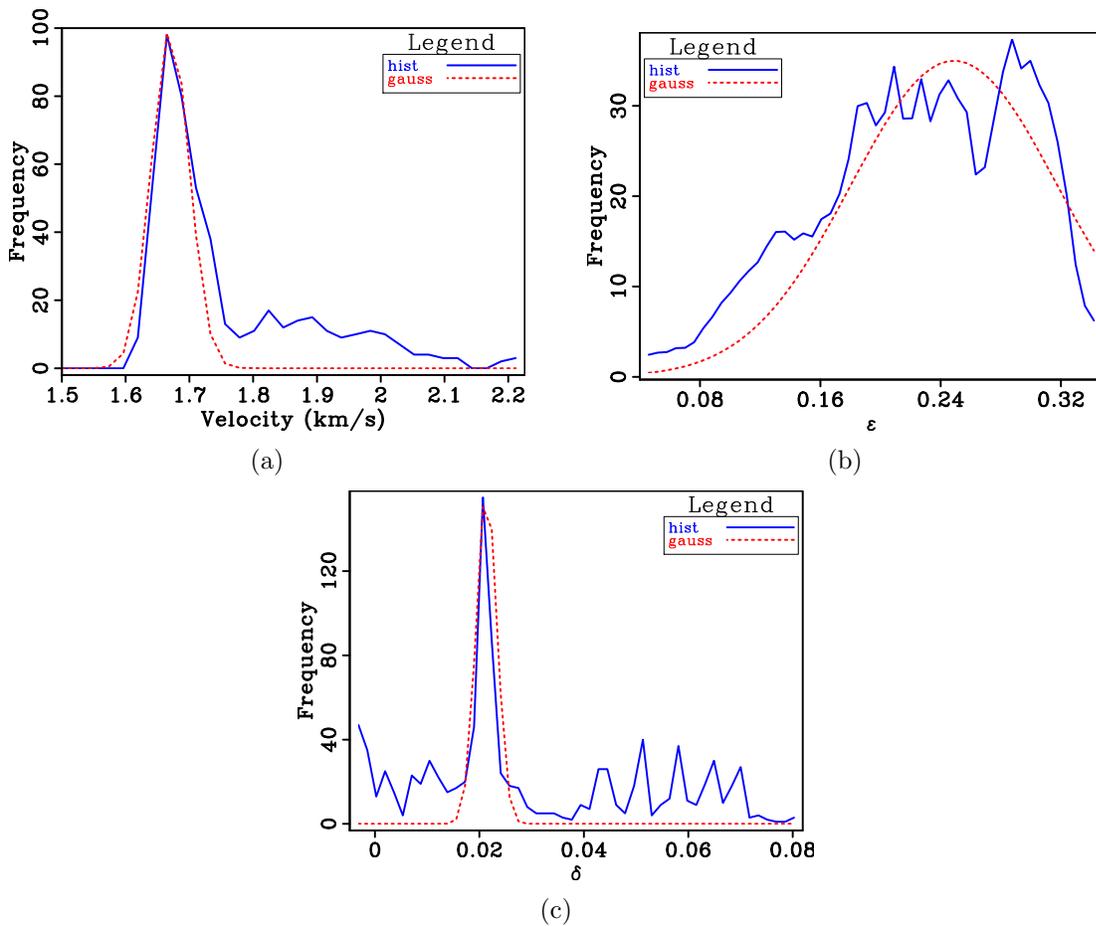


Figure 4.3: Stochastic rock physics modeling results for (a) vertical velocity; (b) ϵ ; and (c) δ . The solid line in each panel denotes the numerical histogram for each model variable. The dashed line denotes the approximated multivariate Gaussian distribution based on the numerical realizations. [ER] chap4/. vp0hist,epshist,delhist

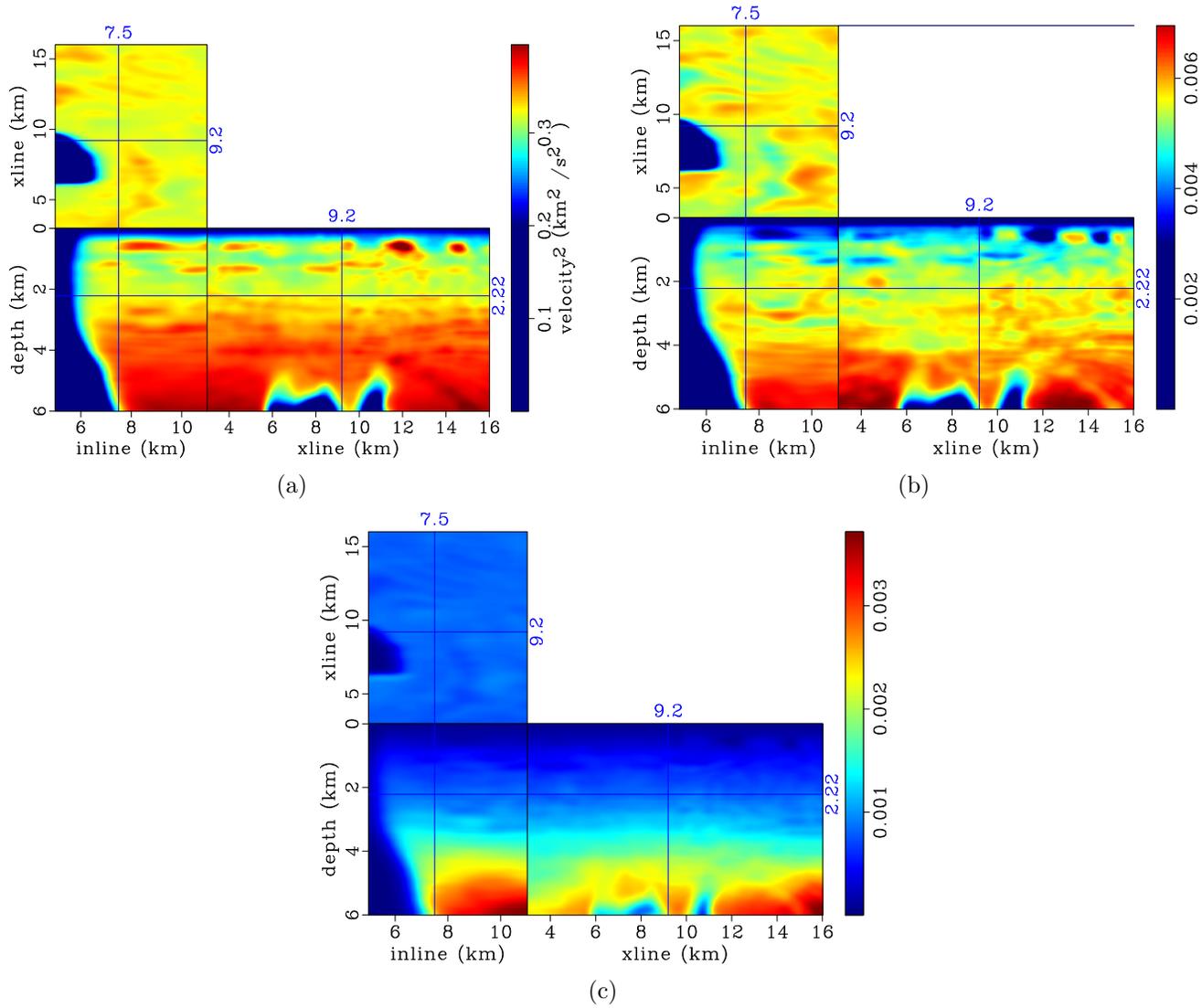


Figure 4.4: Diagonal elements of the covariance matrix. (a): Variance of vertical velocity; (b): Variance of ϵ ; and (c): Variance of δ . [ER] chap4/. covancvv,covancee,covancdd

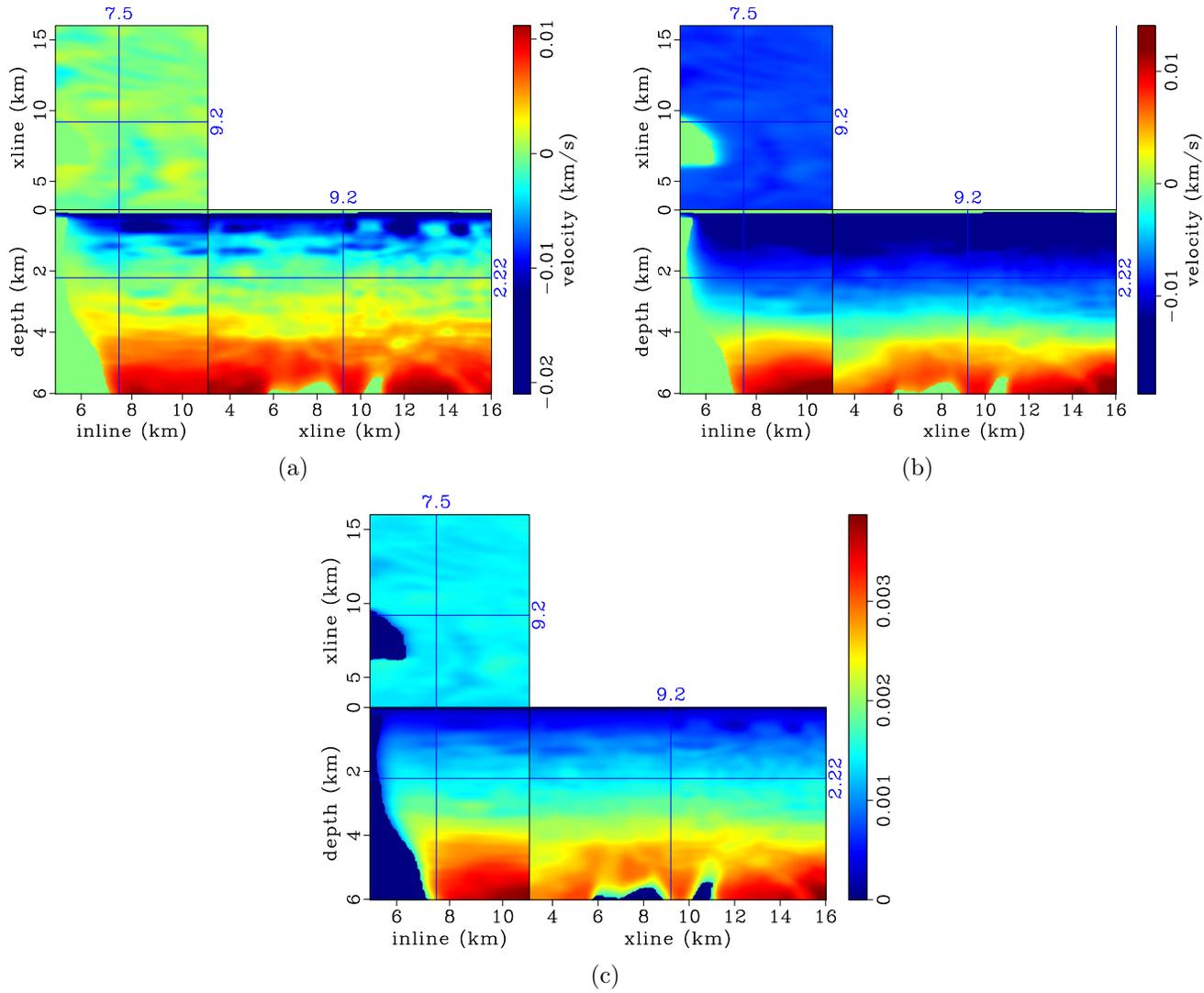


Figure 4.5: Off-diagonal elements of the covariance matrix. (a): Covariance between vertical velocity and ϵ ; (b): Covariance between vertical velocity and δ ; and (c): Covariance between ϵ and δ . [ER]

chap4/. covancve,covancvd,covanced

Figures 4.5(a) to 4.5(c) show the off-diagonal elements in the covariance matrix. The cross-covariances between v_0 and ϵ (Figure 4.5(a)) as well as between v_0 and δ suggest that velocity and anisotropy are negatively correlated in the shallow region but positively correlated in the deep region. This correlation can be explained by rock properties. In the shallow region, high velocity correlates with low anisotropic sand; whereas in the deeper region, high velocity can be caused by mineral diagenesis from smectite to illite, which is also highly anisotropic (Vernik and Liu, 1997). Covariance between ϵ and δ (Figure 4.5(c)) shows positive correlations for all depths.

ANISOTROPIC WEMVA ON 3-D FIELD DATA IN THE GULF OF MEXICO

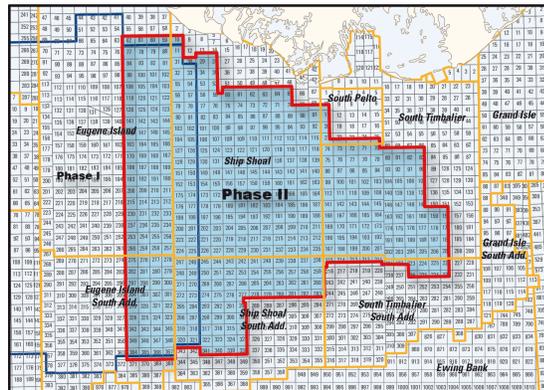
The most important piece of information of the subsurface comes from the seismic data. In this section, I discuss the preprocessing and the inversion of the seismic data to build reliable anisotropic models.

Acquisition and preprocessing

The data were acquired offshore Louisiana in the Gulf of Mexico. Figure 4.6 plots the region that the acquisition covered. Both the pressure and the vertical displacement were measured using sensors laid out along ocean bottom cables (OBC). I was provided with the P-Z summed data in which only the up-going wavefield was retained.

I work with a subset of the 3-D dataset due to the limited computational resources. Figure 4.7 shows the source and receiver locations of the study area given by the acquisition. The source lines are orthogonal to the receiver cables to produce an even midpoint fold. The sources are spaced 400 m in-line and 50 m cross-line (with respect to the receiver lines). The receivers are spaced 50 m in-line and 600 m cross-line. Due to the relatively sparse receiver locations, I use reciprocity to exchange the sources and receivers in the migration.

Figure 4.6: Area covered by the E-Dragon Phase II acquisition. Ocean bottom cables were laid out on the sea bed under shallow water with an average depth of 120 m. [NR] chap4/. dragonarea



I group the receivers onto a 50 m×600 m grid and the sources onto a 50 m×50 m grid. Figure 4.8 shows the acquisition pattern for a single receiver, and Figure 4.9 shows the corresponding receiver gather. In the in-line direction, large acquisition holes are consistent with the sparse in-line source sampling. In the better sampled cross-line direction, I notice strong dispersive arrivals with high amplitudes after 3.5 s. These low frequency surface wave events dominate the receiver gathers in the later time.

To remove the surface waves, I apply a bandpass filter to select the reflection energy between 3 Hz and 25 Hz. The upper bound of the frequency band is limited by the available computational resources. I also apply a first arrival muting and a $t^{2.5}$ time weighting on the receiver gather to boost up the later arrivals. The receiver gather after preprocessing is shown in Figure 4.10. I regard the receiver gather as a shot gather and use the implicit finite different implementation of the VTI one-way wave-equation (Shan, 2008) to perform shot-profile migration.

Initial anisotropic models and gradients

An initial ray-based tomography has been performed by Schlumberger MultiClient. Figure 4.11(a) shows the provided isotropic velocity model. Figures 4.11(b) to 4.11(d) show the initial anisotropic model. The initial ϵ model (Figure 4.11(c)) and δ model (Figure 4.11(d)) are the averaged models from the stochastic rock physics modeling.

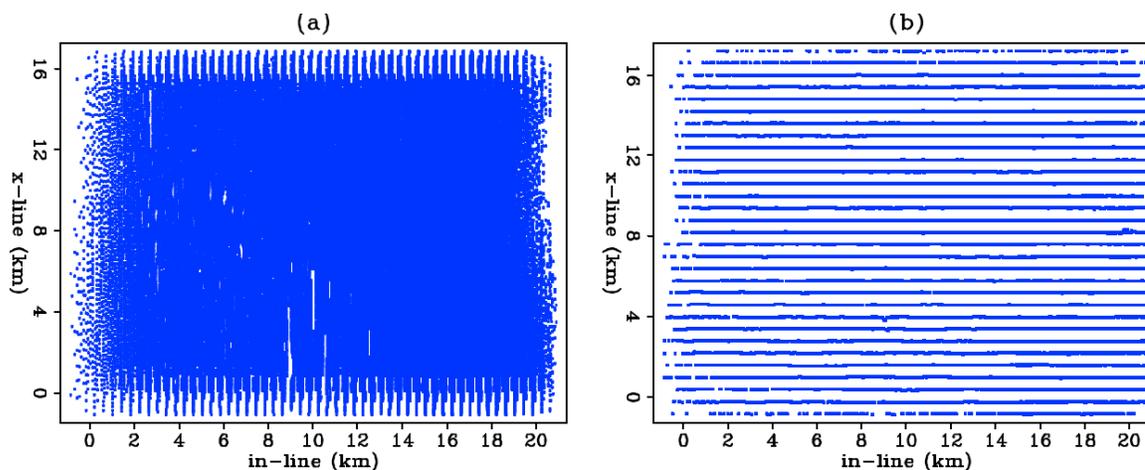
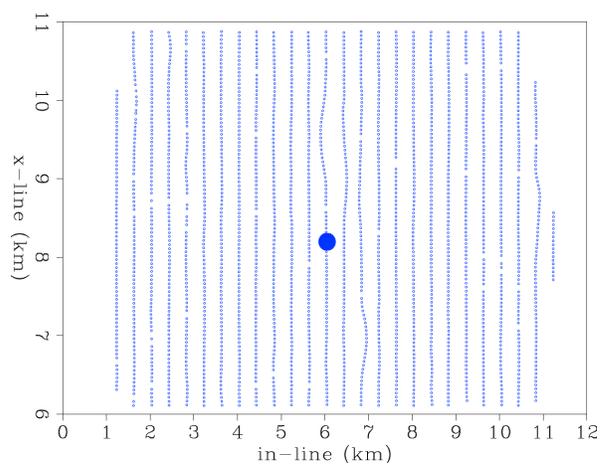


Figure 4.7: Cross-spread geometry of the GoM dataset. Panel (a): Source locations; Panel (b): Receiver locations. The source lines and the receiver lines are perpendicular to each other. I define the direction along the receiver lines the in-line direction and the perpendicular direction cross-line direction. [ER] chap4/. patch-geo

Figure 4.8: Acquisition pattern for one particular receiver. Source locations are denoted by "o" and the receiver location is denoted as the solid circle. For any given receiver, the maximum in-line offset is about 6 km and the maximum cross-line offset is about 3 km. [ER] chap4/. 1rec-geo



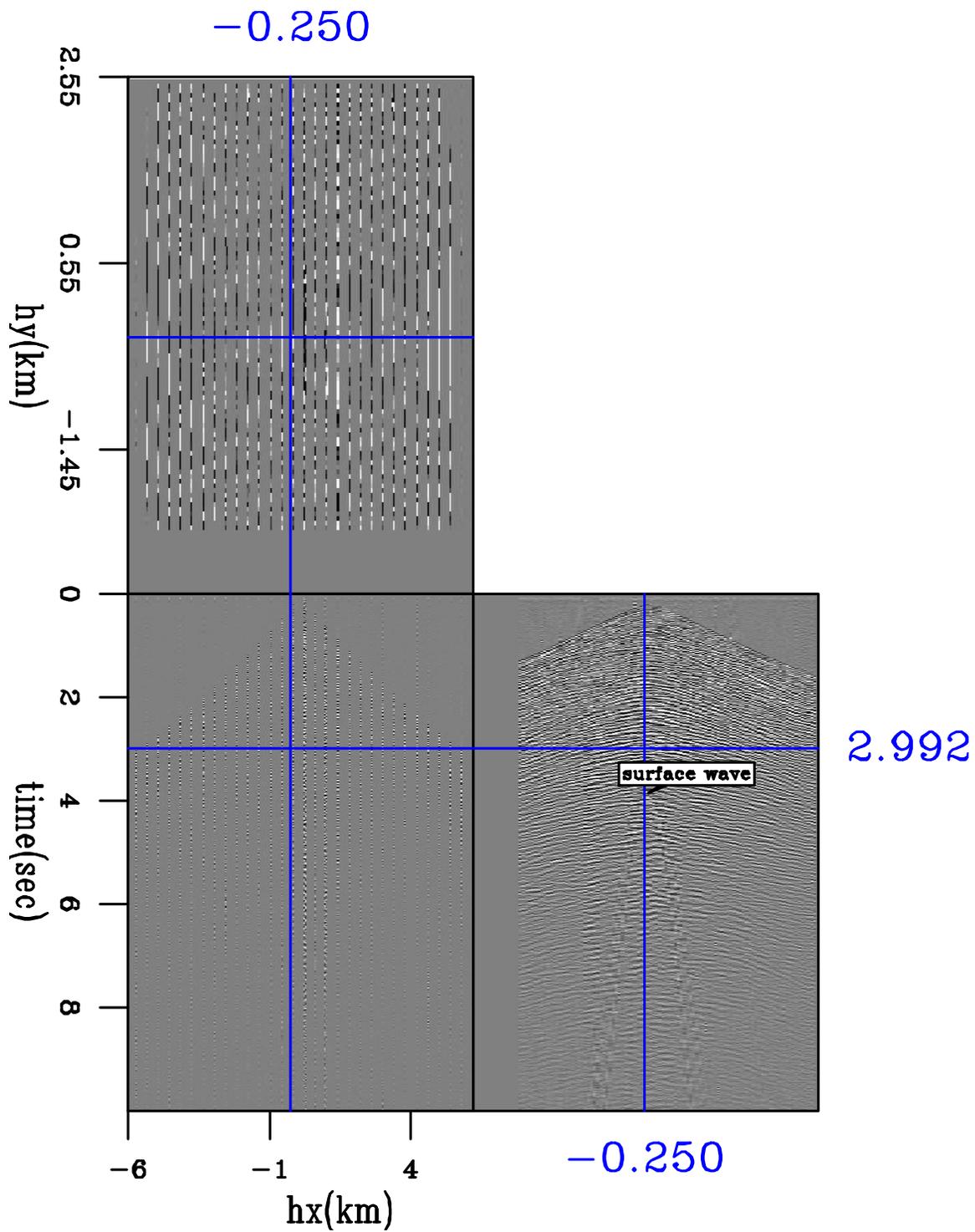


Figure 4.9: One receiver gather shows the dense cross-line sampling and very coarse in-line sampling. The slow arrivals after 3.5 s are the low frequency surface wave energies. [ER] `chap4/. recgth1`

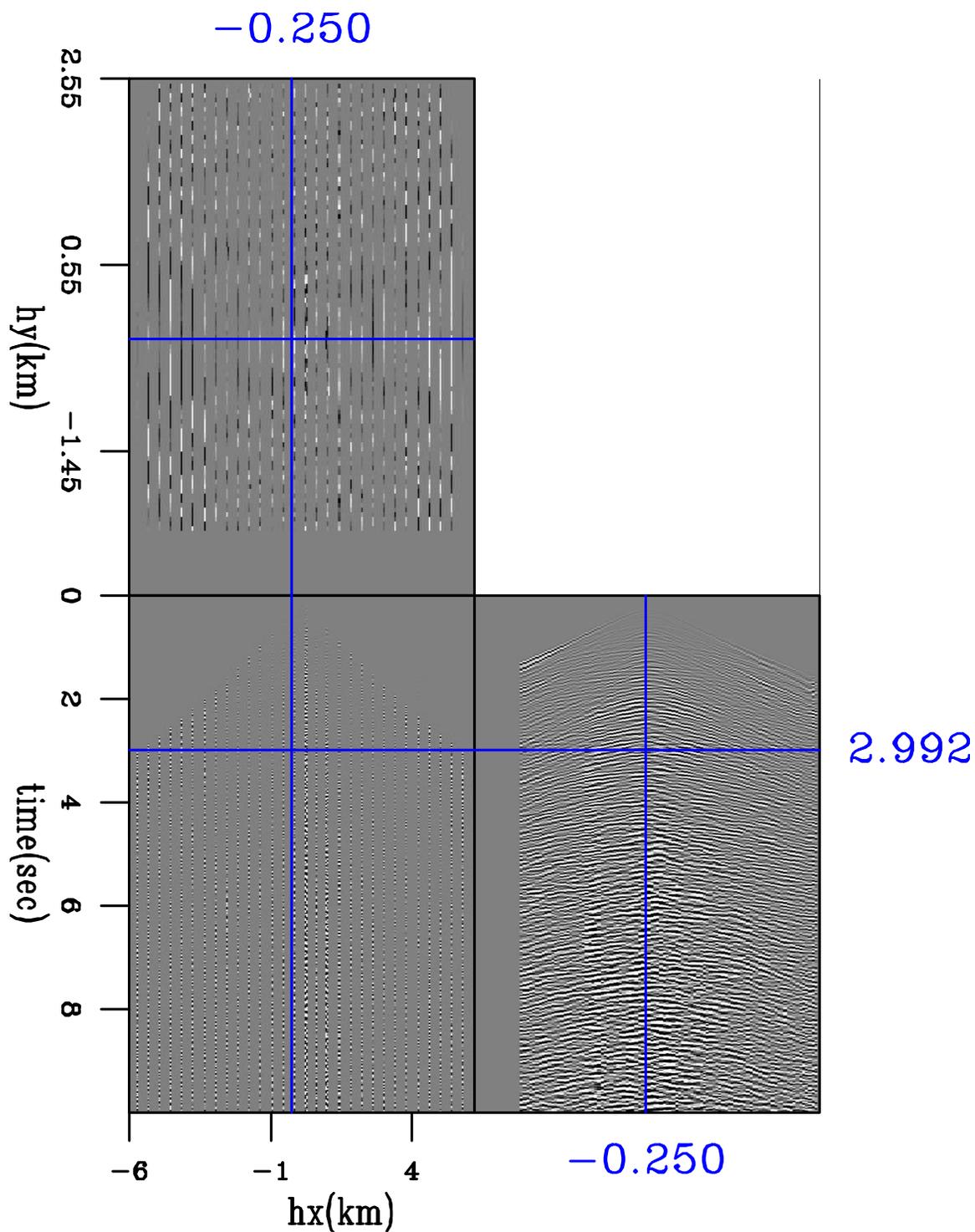


Figure 4.10: Receiver gather in Figure 4.9 after bandpass, muting and time weighting. Surface waves have been attenuated. Reflection events become more continuous especially in the later time. [ER] chap4/. recgth1-proc

The salt body and the water column are considered isotropic. It is reasonable to assume that the tomographic isotropic velocity (Figure 4.11(a)) best estimates the NMO velocity in a VTI parameterization. Therefore, I compute the anisotropic vertical velocity model 4.11(b) from the δ model (Figure 4.11(d)) and the isotropic NMO velocity (Figure 4.11(a)). The vertical velocity is slightly smaller than the NMO velocity due to a positive δ model.

Figure 4.12 and Figure 4.13 compare the images between the isotropic migration and the anisotropic migration at two different locations. Compared with Figure 4.12(a), the depths of the flat reflectors in the anisotropic migration image (Figure 4.12(b)) remain roughly the same due to a small δ value. Nevertheless, the focusing of these shallow flat events are improved by the anisotropic model. The depth slice of the anisotropic migration image shows better coherence and higher resolution on the discontinuities across the fault (labeled 1). The dipping reflectors in the in-line direction (highlighted by oval) and the strong near-horizontal reflector above 6 km (labeled 2) in the cross-line direction are also more continuous in the anisotropic migration image. Comparison between Figure 4.13(a) and Figure 4.13(b) shows similar qualities in both migration images. The crossing events and broken reflectors below 4 km in the cross-line direction indicate better models are needed to properly focus the image. These images demonstrate a typical case where two different models, with completely different geological implications, can explain the seismic traveltimes equally well.

To update the anisotropic model, I use the anisotropic WEMVA scheme (Chapter 3). However, the computational and memory requirement to build the subsurface offset common image gathers (SOCIGs) and hence the angle domain common image gathers (ADCIGs) are prohibitively expensive. Furthermore, the migration image based on the initial VTI model is relatively well focused. Therefore, I assume the initial model is close enough to the true solution, and drop the differential semblance term in the objective function.

The anisotropic WEMVA objective function I minimize in this chapter reads as

follows:

$$S(\mathbf{m}) = -\frac{1}{2} \left\| \sum_{\theta} \mathbf{I}(\mathbf{x}, \theta) \right\| + \frac{\alpha}{2} (\mathbf{m} - \mathbf{m}_{\text{prior}})^T \mathbf{C}_M^{-1} (\mathbf{m} - \mathbf{m}_{\text{prior}}), \quad (4.1)$$

where the first term defines the “data fitting” objective by maximizing the stacking power, and the second term defines the “model regularization” objective by including the geological and the rock physics information through the initial model $\mathbf{m}_{\text{prior}}$ and the covariance matrix \mathbf{C}_M . Parameter α balances the relative weights between the two objectives. During implementation, the stack image $\sum_{\theta} \mathbf{I}(\mathbf{x}, \theta)$ is simply computed using the cross-correlation imaging condition at zero subsurface offset in both in-line and cross-line directions.

Each anisotropic WEMVA iteration has three main computation-intensive steps: migration, gradient computation, and the line search process. I summarize the computational cost during each anisotropic WEMVA iteration in Table 4.1. Computation time for each process is measured on a single node with OpenMP parallelization over 16 processors. Assuming 60 nodes are always available, it takes 56 hours to finish one WEMVA iteration.

Table 4.1 shows that the most expensive processing step is the gradient computation. To reduce the computation cost, I take advantage of the redundancy in shots during the gradient computation. During each iteration, I randomly select one-third of the shots to perform the tomographic gradient calculations. The inaccuracy introduced by the subsampling is insignificant for two reasons: first, the subsampling artifacts are dominated by higher wavenumber components, especially in the horizontal direction. These artifacts are mitigated by the structural smoothing process; second, artifacts introduced in the current iteration are incoherent with the inaccuracies in the next iterations when another set of random shots are used. Therefore, the stacking of the updates over iterations will also suppress the artifacts in the gradients.

The subsampling strategy is not applied for the migration and the line search steps although it would greatly reduce the computational cost. The reasons for not applying the subsampling strategy are two fold. First, dense sampling in shots are needed to

fully collapse the migration smiles due to the sparse cross-spread acquisition geometry. Second, the stacking power objective function is very sensitive to the artifacts in the migration image, which could be translated into unrealistic model updates.

To demonstrate the redundancy of the shots when computing the gradients, I show the gradient in vertical velocity using all the shots in Figure 4.14(a). The gradient has been smoothed along the structure dips estimated from the initial migration image. A fast velocity anomaly centered at $x = 9$ km and $y = 15$ km dominates the gradient.

The gradient using only a third of the shots is shown in Figure 4.15(a). Due to the reduced number of stacked shots, the amplitudes of the gradients are reduced. Nonetheless, the structures of the updates are remarkably similar except for minimal amplitude shifts. Therefore, I can safely reduce the number of shots in gradient computation without introducing significant errors. Similar comparison and conclusions can be made for the gradients of ϵ and δ (compare Figures 4.14(b) with 4.15(b), and Figures 4.14(c) with 4.15(c)).

I summarize the non-linear conjugate gradient optimization workflow in Algorithm 3. With the random shot reduction (shown between parenthesis in Table 4.1), the total time for one WEMVA iteration is reduced to 36 hours.

Process	Time/Shot (min)	# of Shots
Migration	7	3500
Gradient computation	30	3500 (1167)
Line search	21	3500

Table 4.1: Computational cost for each WEMVA iteration.

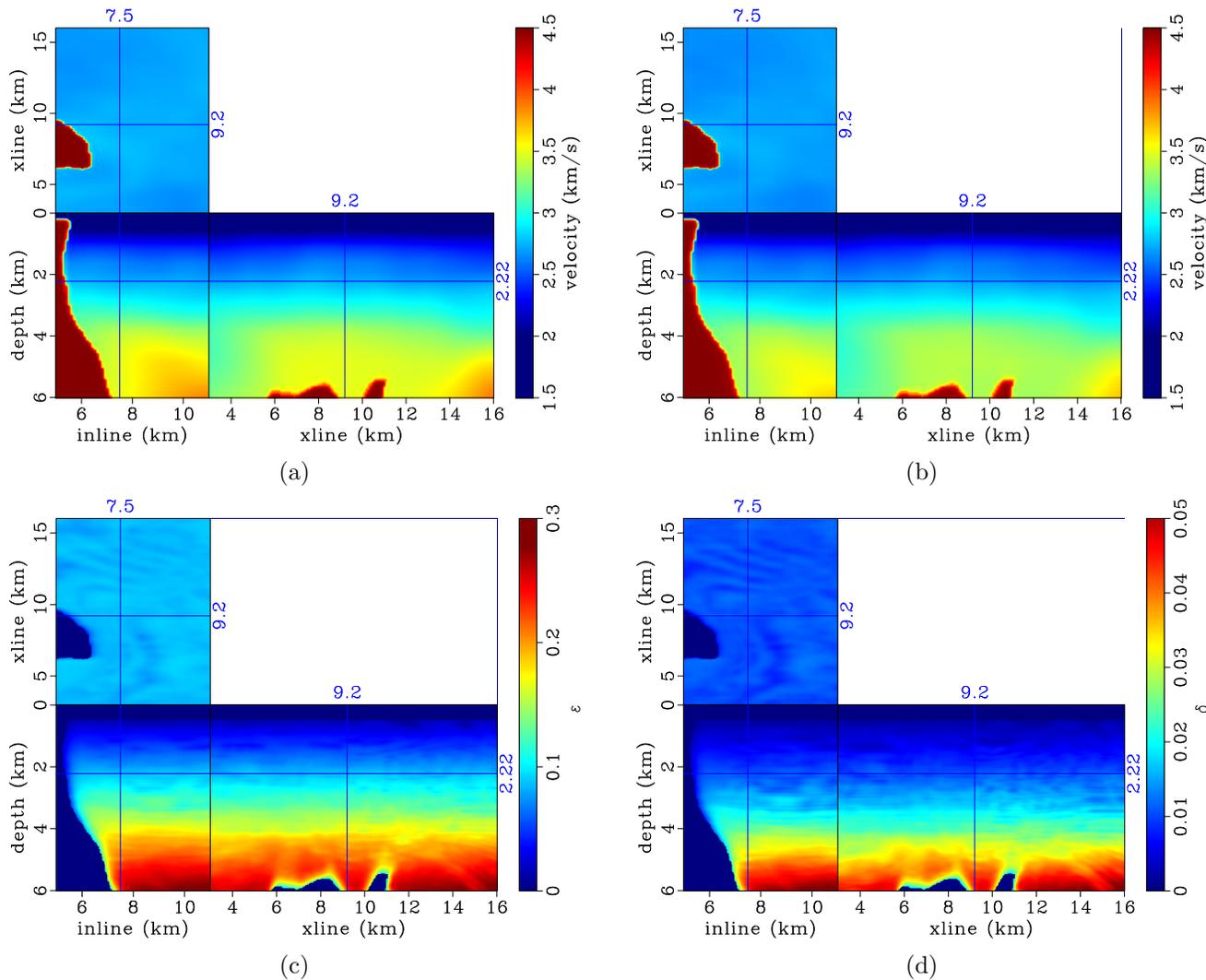


Figure 4.11: (a): Isotropic velocity model provided by ray-based tomography. (b): Vertical velocity model, (c): initial ϵ model and (d): initial δ model from the rock physics modeling. [CR]

chap4/. iso-vel,v0-iter0,eps-iter0,delta-iter0

Algorithm 3 Optimization algorithm

initialize the model: \mathbf{m}_0 compute the migrated image: I_0

randomly select 1/3 of all shots

compute the gradient: \mathbf{g}_0 precondition the gradient: $\mathbf{g}_s^0 = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_0$ initialize the search direction: $\mathbf{p}_0 = -\mathbf{g}_s^0$ **for** $k = 1 \cdots N_k$ **do**perform a line search: optimize λ , argmin $J(\mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1})$ update the velocity model: $\mathbf{m}_k = \mathbf{m}_{k-1} + \lambda\mathbf{p}_{k-1}$ compute the migrated image: I_k

randomly select 1/3 of all shots

compute the gradient: \mathbf{g}_k precondition the gradient: $\mathbf{g}_s^k = \mathbf{S}\Sigma\Sigma^*\mathbf{S}^*\mathbf{g}_k$ find the search direction: $\mathbf{p}_k = -\mathbf{g}_k + \frac{(\mathbf{g}_s^k)^T(\mathbf{g}_s^k - \mathbf{g}_s^{k-1})}{(\mathbf{g}_s^{k-1})^T\mathbf{g}_s^{k-1}}$ **end for**

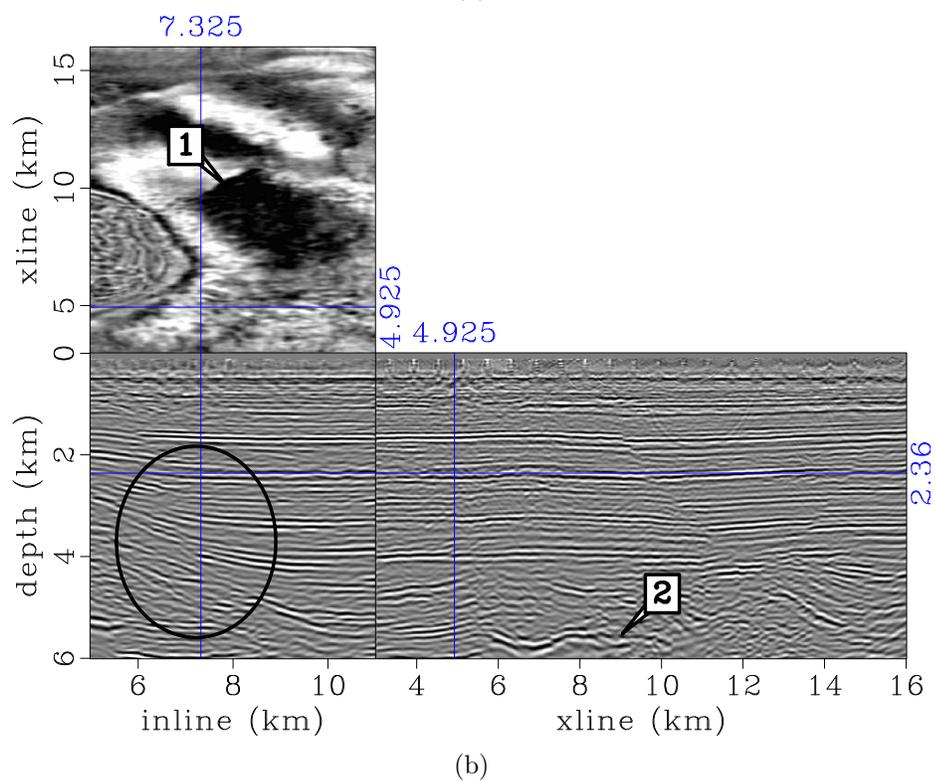
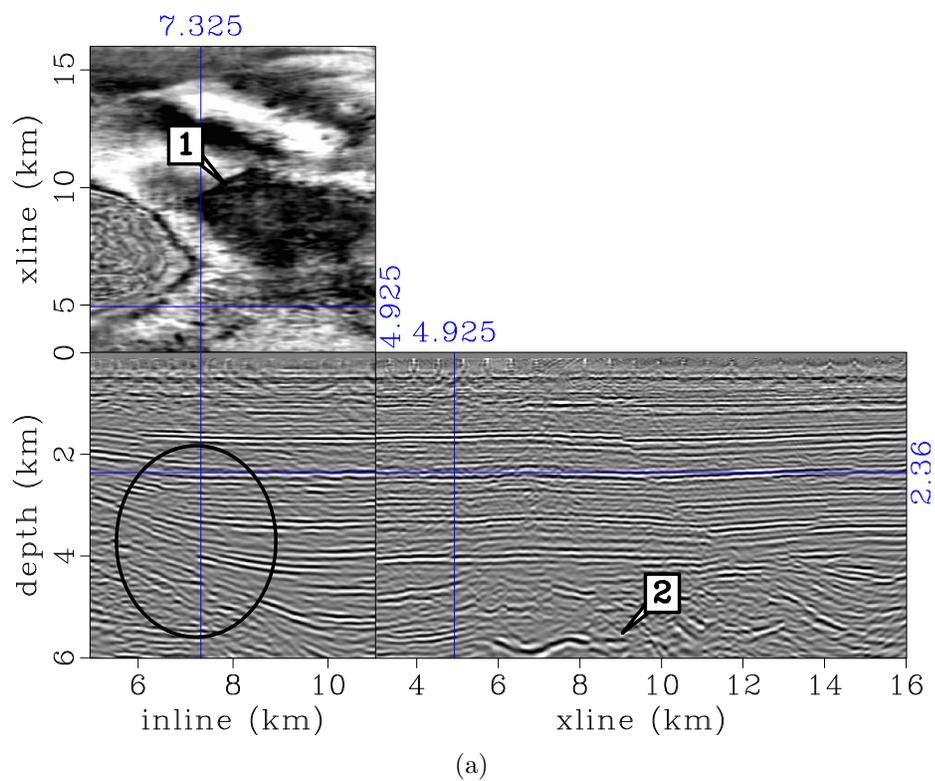


Figure 4.12: (a): Isotropic migration image using the isotropic velocity in Figure 4.11(a). (b): Anisotropic migration image using the models in Figure 4.11(b) to 4.11(d). [CR] chap4/. iso-img2,ani-img2

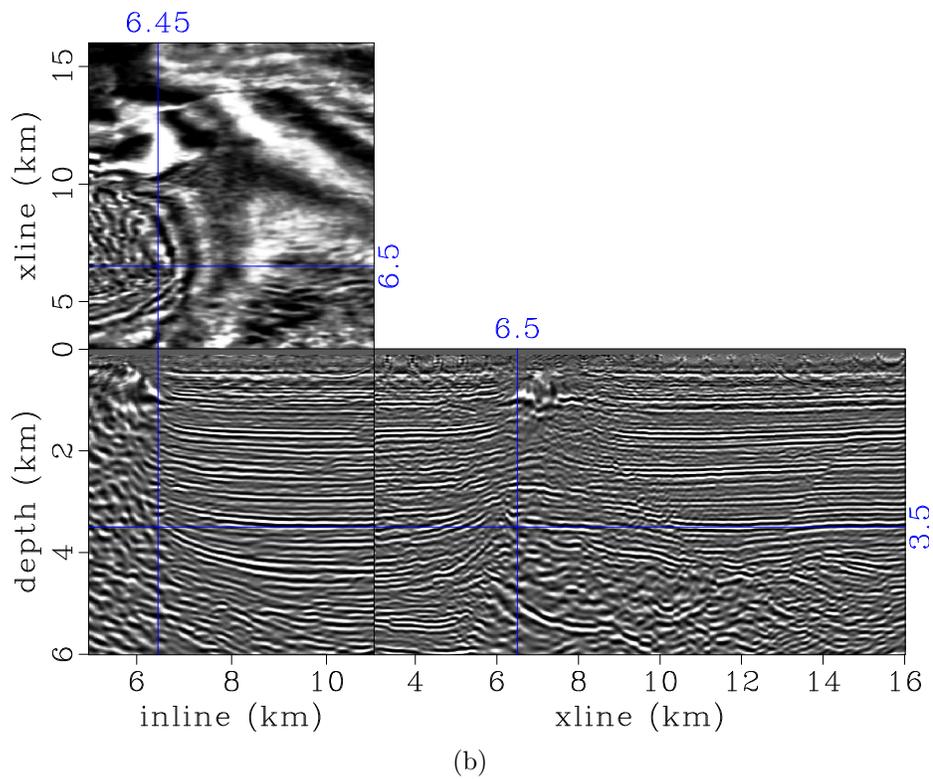
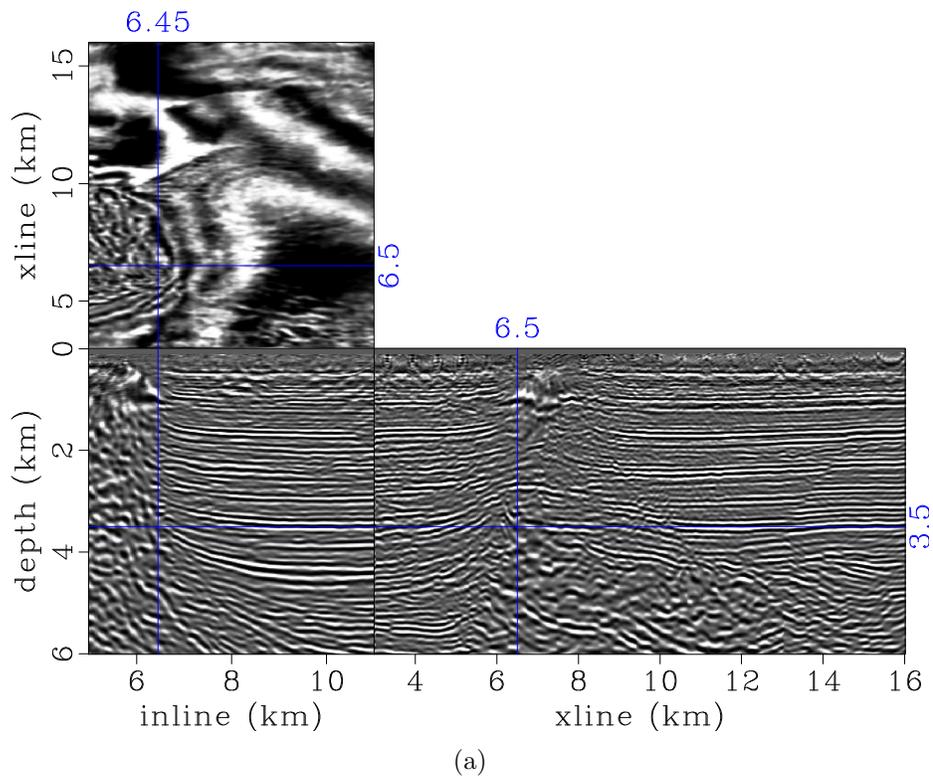


Figure 4.13: Same as Figure 4.12 but at a different location. [CR]

chap4/. iso-img1,ani-img1

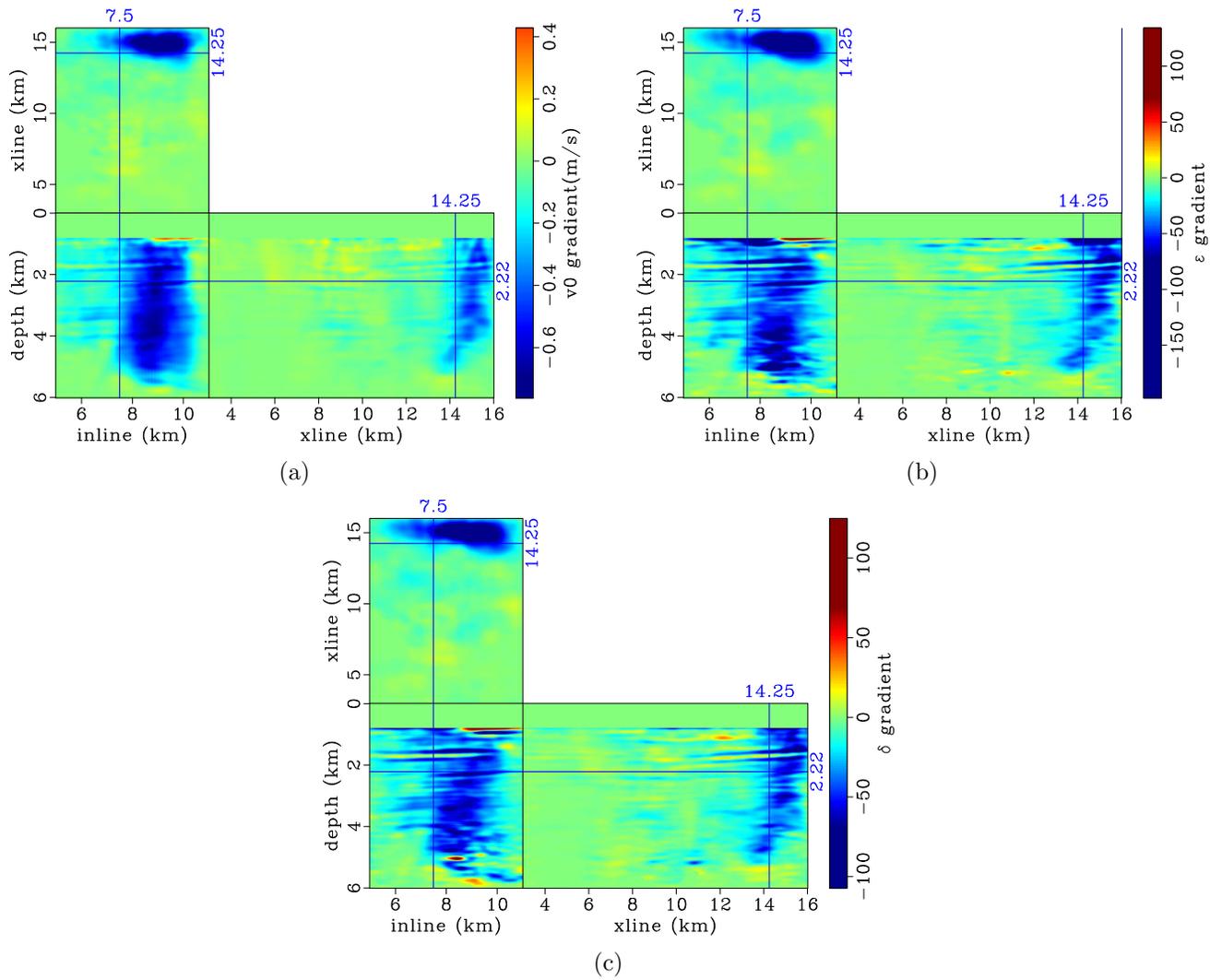


Figure 4.14: First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ using all shots. [CR]
 chap4/. gvel-all-raw,geps-all-raw,gdel-all-raw

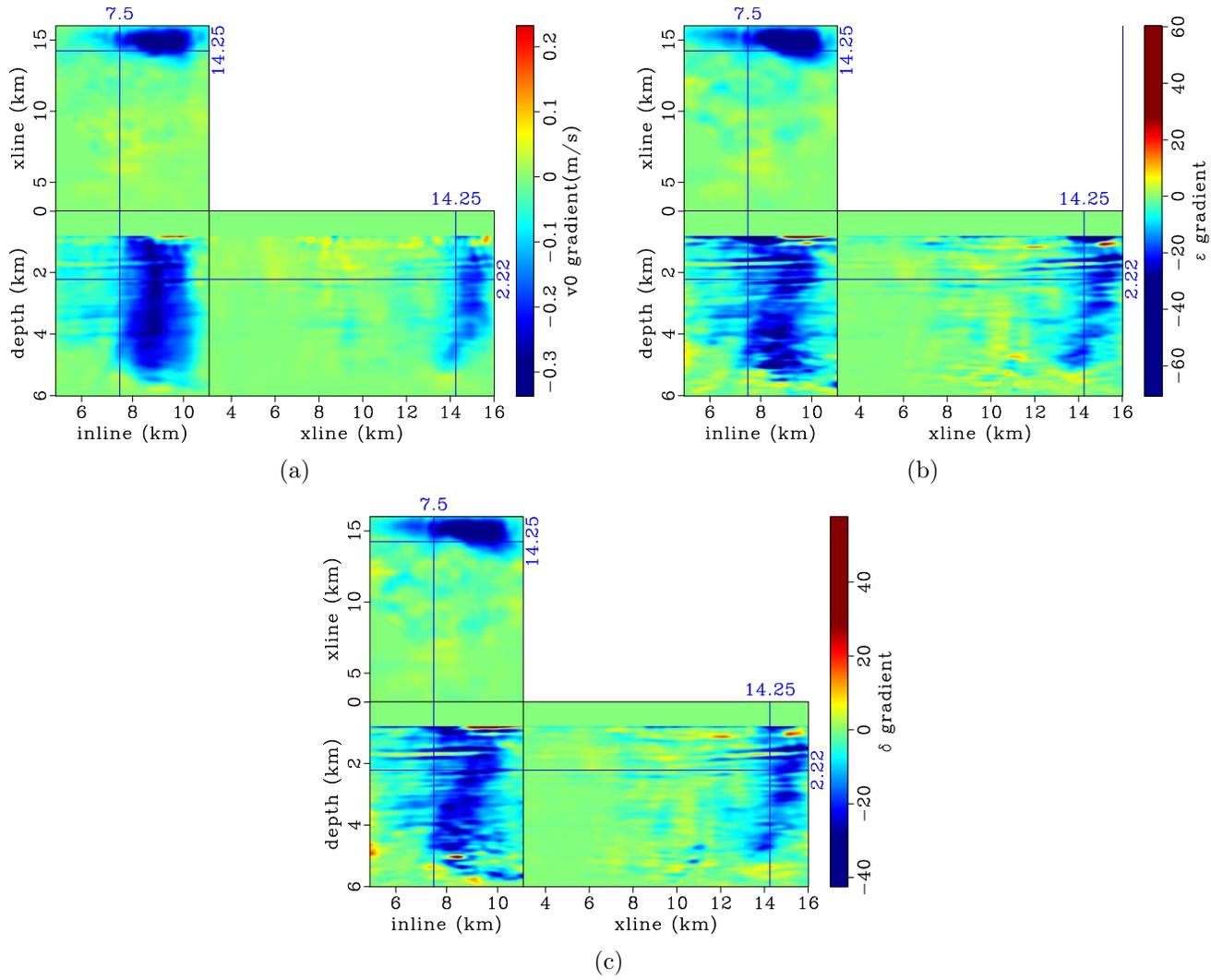


Figure 4.15: First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ using one-third of all the shots. [CR] `chap4/. gvel-third-row,geps-third-row,gdel-third-row`

Compared with the gradient in vertical velocity (Figure 4.15(a)), the gradients in ϵ (Figure 4.15(b)) and δ (Figure 4.15(c)) show similar structures and point to the same update directions. These updates are determined by the WEMVA tomographic operator so that they increase the stacking power constructively. However, it is unrealistic for a shallow layer to have both positive velocity and positive anisotropy based on the lithological inversion results and our rock physics prior knowledge.

To honor the rock physics prior information, we apply the cross-parameter covariance matrix (Figures 4.4 and 4.5) in the preconditioning. As a result, the gradient direction in ϵ (Figure 4.16(b)) and δ (Figure 4.16(c)) in the shallow region has been reversed based on the negative correlation between velocity and anisotropy. These preconditioned gradients (Figure 4.16) are used in a Polak-Ribière nonlinear conjugate gradient scheme (Wright and Nocedal, 1999) to calculate the update direction.

I use a three-point parabolic interpolation scheme (Box et al., 1969) to conduct a line search. The line search algorithm evaluates the objective function with three different step lengths. Then a parabola is fitted to find the minimum and its corresponding step length. This line search scheme is exact when the objective function is parabolic near its bottom. However, the shape of the objective function is not guaranteed parabolic in the case of nonlinear inversion. Therefore, I compare the four trial step lengths with the objective function in the previous iteration, keep the lowest objective function, and update the model with the corresponding step length.

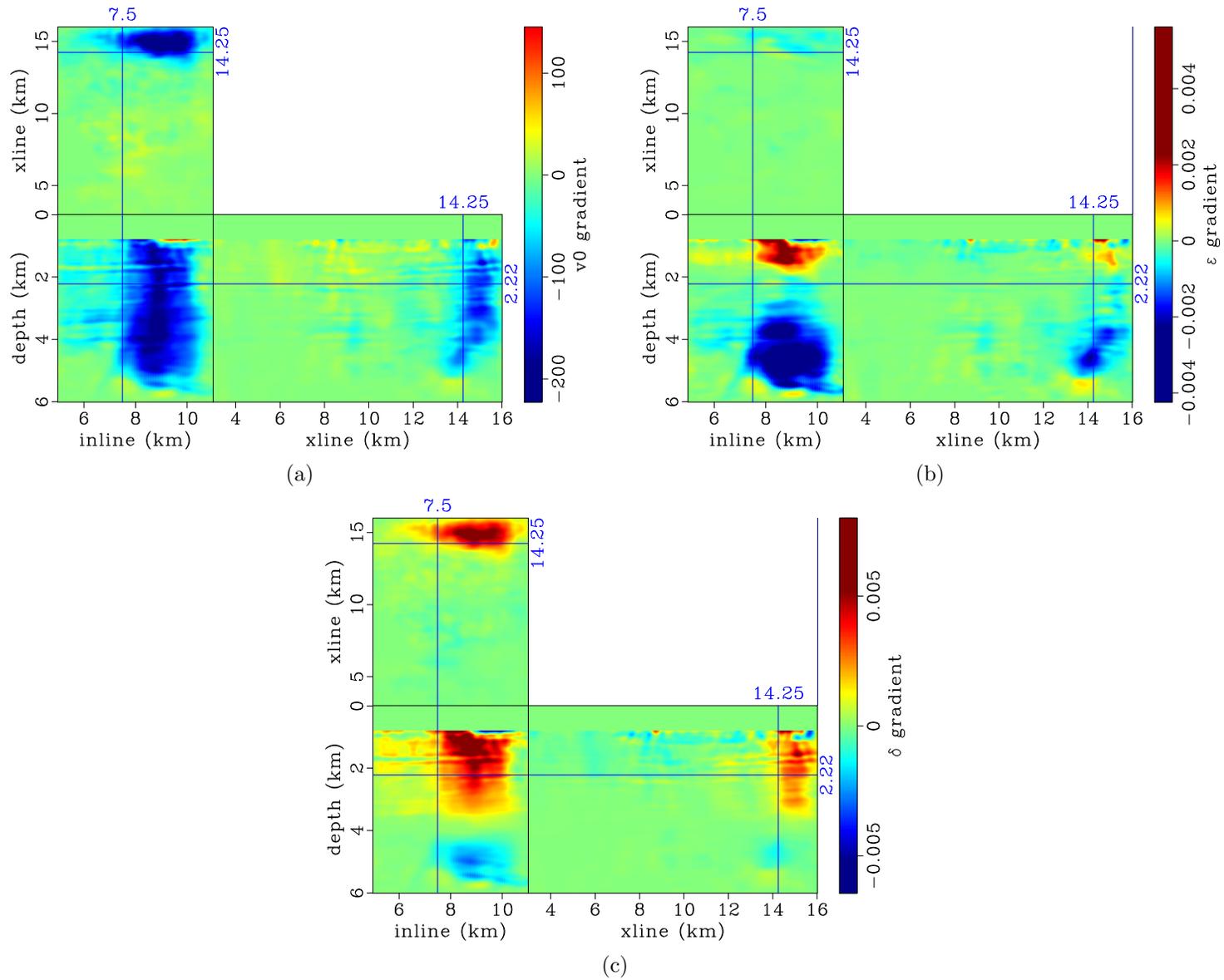


Figure 4.16: First anisotropic WEMVA gradient of (a) vertical velocity, (b) ϵ , and (c) δ after rock physics preconditioning. [CR] `chap4/. gvel-third-prc,geps-third-prc,gdel-third-prc`

Inversion results

In this subsection, I present the inversion results after ten nonlinear anisotropic WEMVA iterations. Both the normalized objective function curve (Figure 4.17(a)) and the normalized gradient magnitude curve (Figure 4.17(b)) demonstrate a fast convergence in the first few iterations. The objective function curve flattens out after seven iterations. Although the inversion still updates the VTI model in the last three iterations, these updates do not affect the focusing of the seismic image.

Figure 4.18 shows the total updates in the VTI model after ten nonlinear iterations. The inversion identifies a positive velocity anomaly on the footwall of the normal fault, which can be caused by the regional stress distribution. To consistently produce a high velocity rock in the shallow region, the regularization scheme reduces the anisotropy in the corresponding region, indicating a stronger sand-dominated sedimentary environment. In the deeper region, both velocity and anisotropy are positively updated due to the strong diagenesis in the shale. Significant updates are also shown around the salt body where the initial ray-based tomography failed to update.

It is interesting to note that the negative correlation between velocity and δ extends deeper than the negative correlation between velocity and ϵ . However impossible to be obtained from the seismic data inversion, it is consistent with the rock physics prior information. On average there are 5% positive updates in vertical velocity and more than 10% updates in both ϵ and δ .

Figure 4.19 shows the updated VTI model. All three VTI parameters are smooth along the structure dip direction, which agrees with our geological assumptions. All three VTI parameters are consistent with the rock physics model I choose to apply, because I constrain both the initial model and the model updates during each iteration.

Figures 4.20 and 4.21 compare the migration images with the initial model with the migration image with the updated model at a cross-line location near the salt flank. Reflectors in both in-line and cross-line directions are strongly dipping. Highlighted

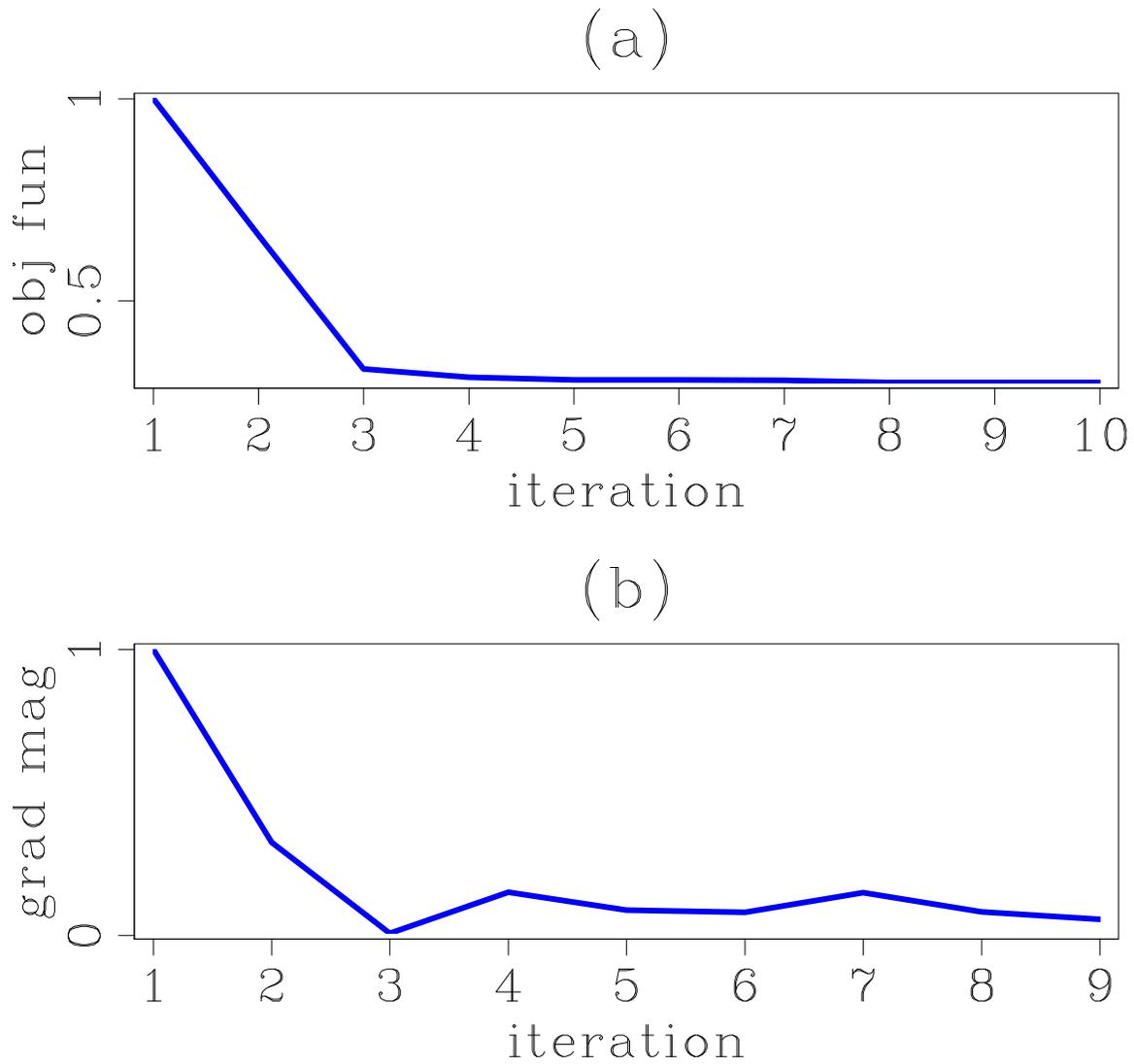


Figure 4.17: (a): Normalized objective function curve. (b): Normalized gradient magnitude curve. Both curves show fast convergence in the first few iterations. [CR] chap4/. dragon-converge

by circles labeled 2 and 3, these reflectors are broken in the initial migration image, making the interpretation of the sediments around the salt body a challenging task. The updated model revealed a much clearer image around the salt. Reflectors in both areas are better imaged with more continuity and better resolution. The improved resolution is also shown on the depth slice. Flag 1 points out a fault which is blurred in the initial migration image but distinctly imaged with the updated model.

Similar improvements can be observed on the other side of the salt flank. The sediment reflectors are ambiguous from the initial migration image (Figure 4.22), as highlighted by the labeled circles. The truncation effects of the sedimentary layers against the salt body are better imaged with the improved model (Figure 4.23), providing higher confidence for salt body interpretation. Moreover, a set of faults caused by the salt intrusion (label 1) is much better focused on the depth slice with the improved anisotropic model.

Figure 4.24 and 4.25 cut through the image cube near the strongest model updates. On the in-line section, the strong reflectors are imaged deeper and more continuous with a higher velocity (circle 2). On the cross-line section, two normal faults (one is highlighted by circle 3) are much more distinct due to the improved focusing. On the depth slice, the resolution of the sediments around the salt body are much higher in the updated image.

In Figures 4.26 to 4.28, I compare the angle-domain common image gathers before and after update. The initial ADCIGs are reasonably flat, only with minor residual moveouts. Nevertheless, the flatness of the ADCIGs has been improved with the updated VTI model. Without changing depth, the ADCIGs show increased angle coverage in the shallow region, mostly due to the improved definition of the Thomsen parameters ϵ and δ . Overall, the increased flatness and angle coverage produces a better focused image. These angle events can certainly help the interpretation and the AVA analysis.

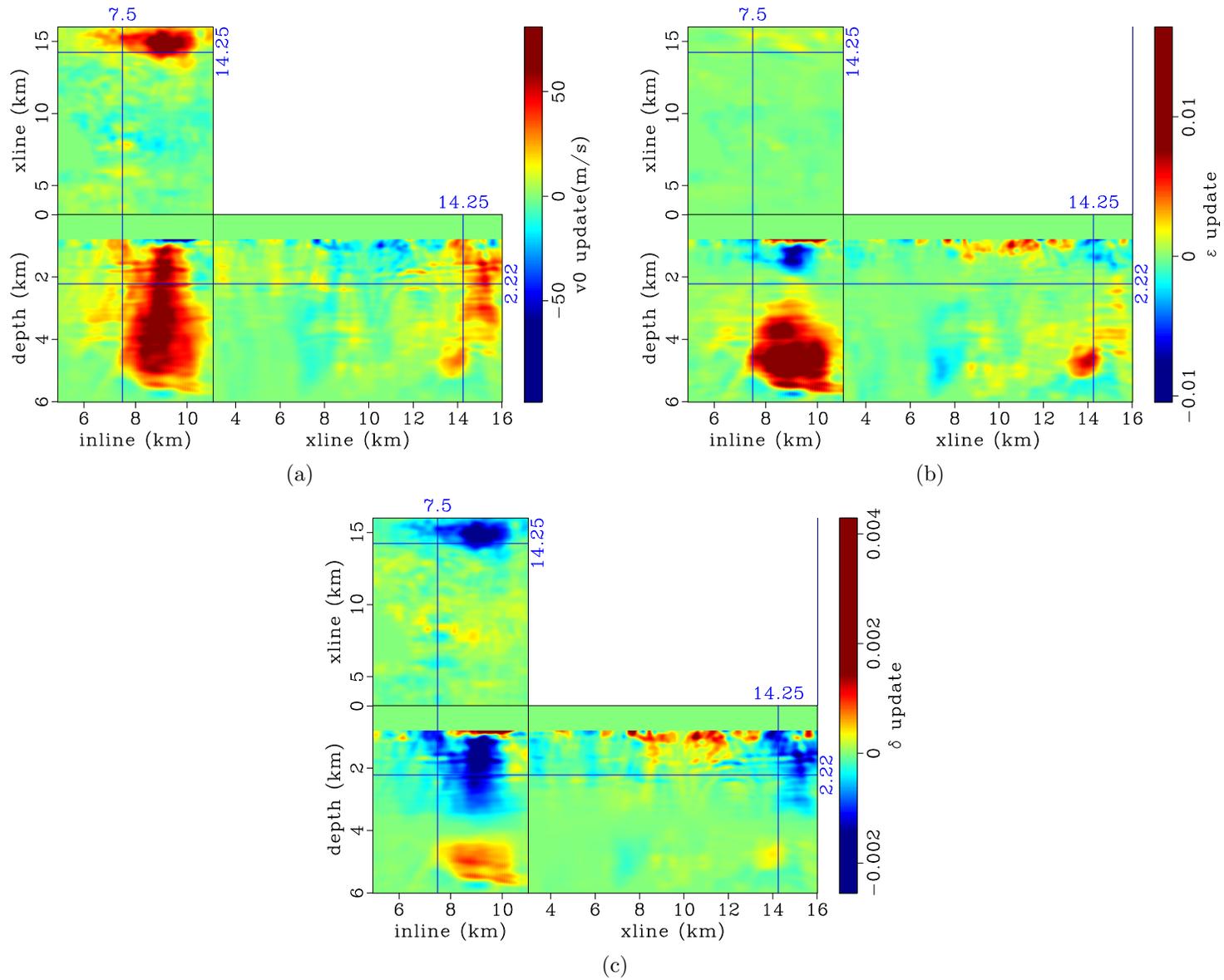


Figure 4.18: Total updates after ten anisotropic WEMVA iterations in (a): vertical velocity model, (b): ϵ model and (c): δ model after ten anisotropic WEMVA iterations. [CR] chap4/. vel-updt-all,eps-updt-all,del-updt-all

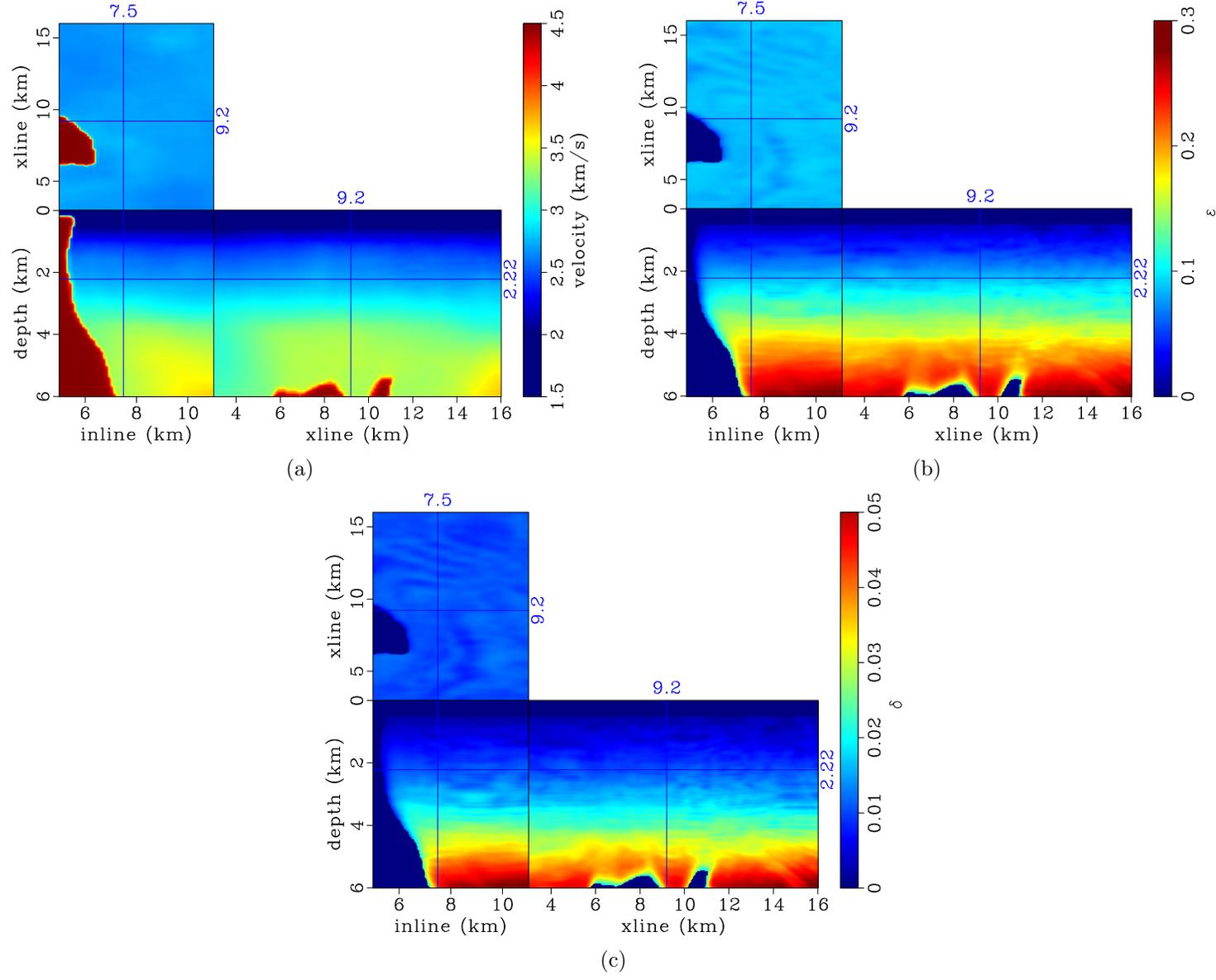


Figure 4.19: (a): Vertical velocity model, (b): ϵ model and (c): δ model after ten anisotropic WEMVA iterations.
 [CR] chap4/. v0-final,eps-final,delta-final

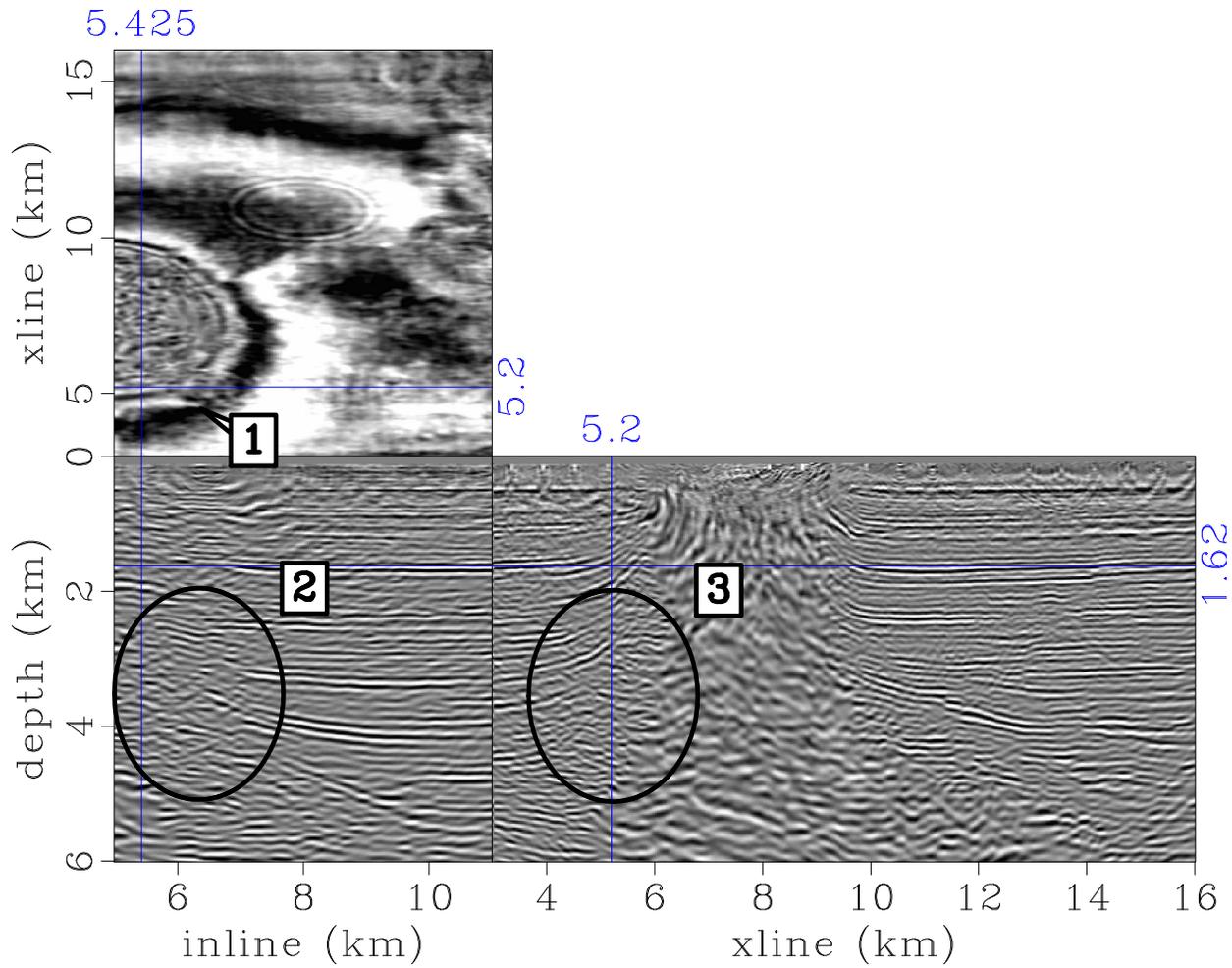


Figure 4.20: Migration image using the initial anisotropic model. Labels 2 and 3 highlights areas with broken reflectors in the in-line and cross-line direction, respectively. [CR] chap4/. ani-iter0-1

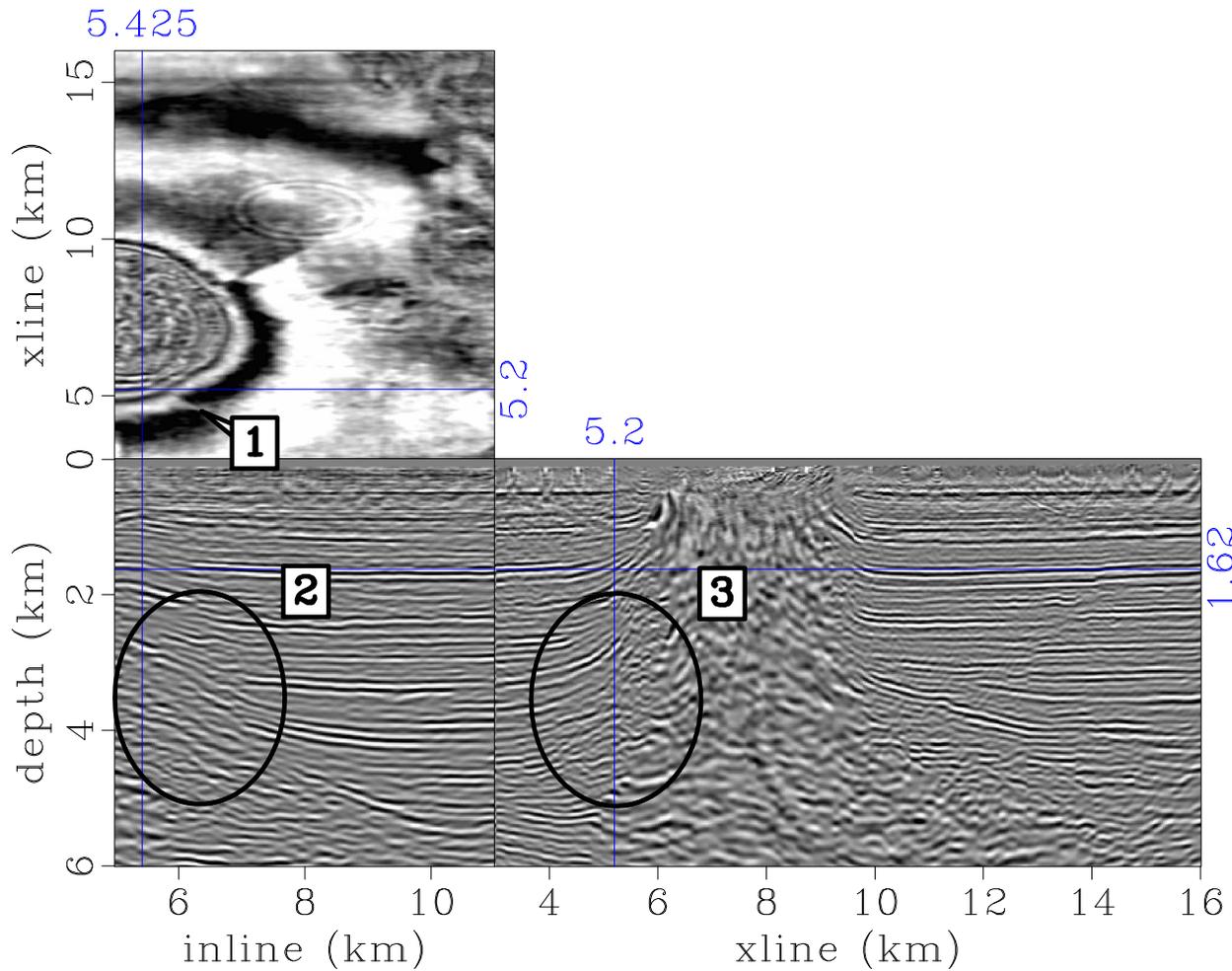


Figure 4.21: Migration image using the updated anisotropic model after ten iterations. Comparing to Figure 4.20, label 1 points out a distinct fault near the salt body which is much better imaged with the updated model. Dipping Reflectors in both circles 2 and 3 are more continuous. [CR] chap4/. ani-iter10-1

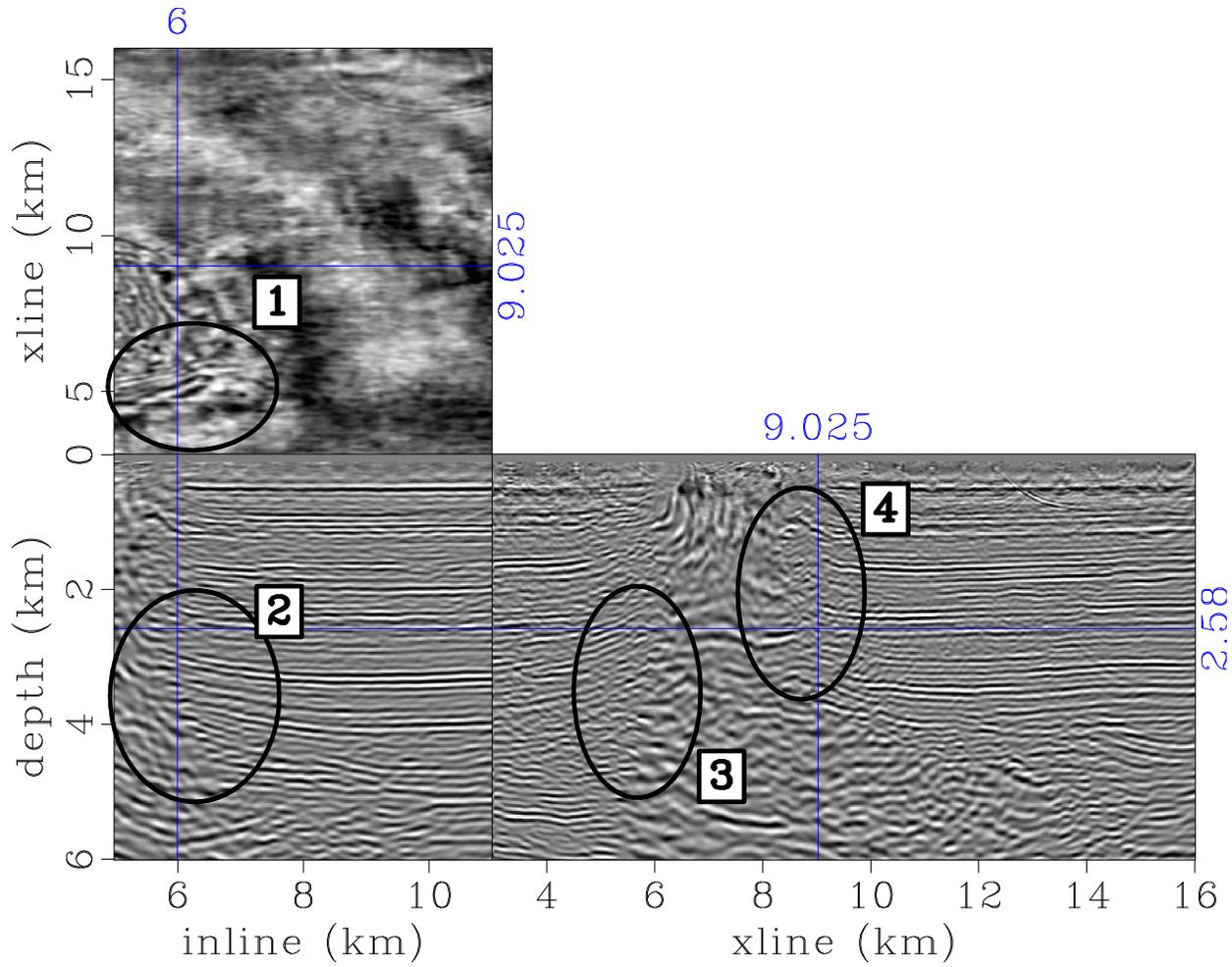


Figure 4.22: Migration image using the initial anisotropic model. Sediments truncation and salt flank are ambiguous in circle 2, 3, and 4. [CR] `chap4/. ani-iter0-2`

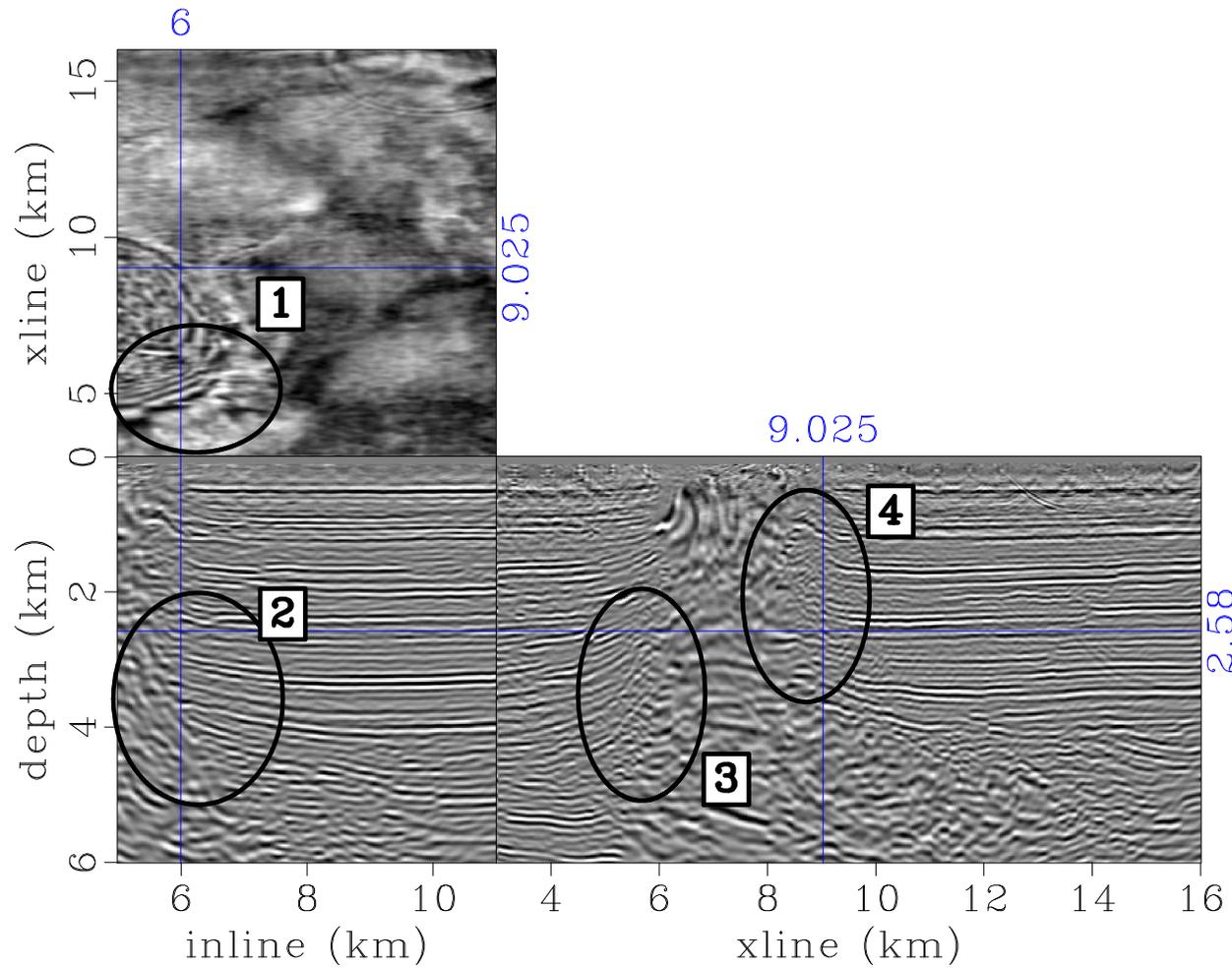


Figure 4.23: Migration image using the updated anisotropic model. Compared to Figure 4.22, sediments truncation and salt flank are much better imaged. On the depth slice, a set of faults (circle 1) caused by the salt intrusion are better focused. [CR] chap4/. ani-iter10-2

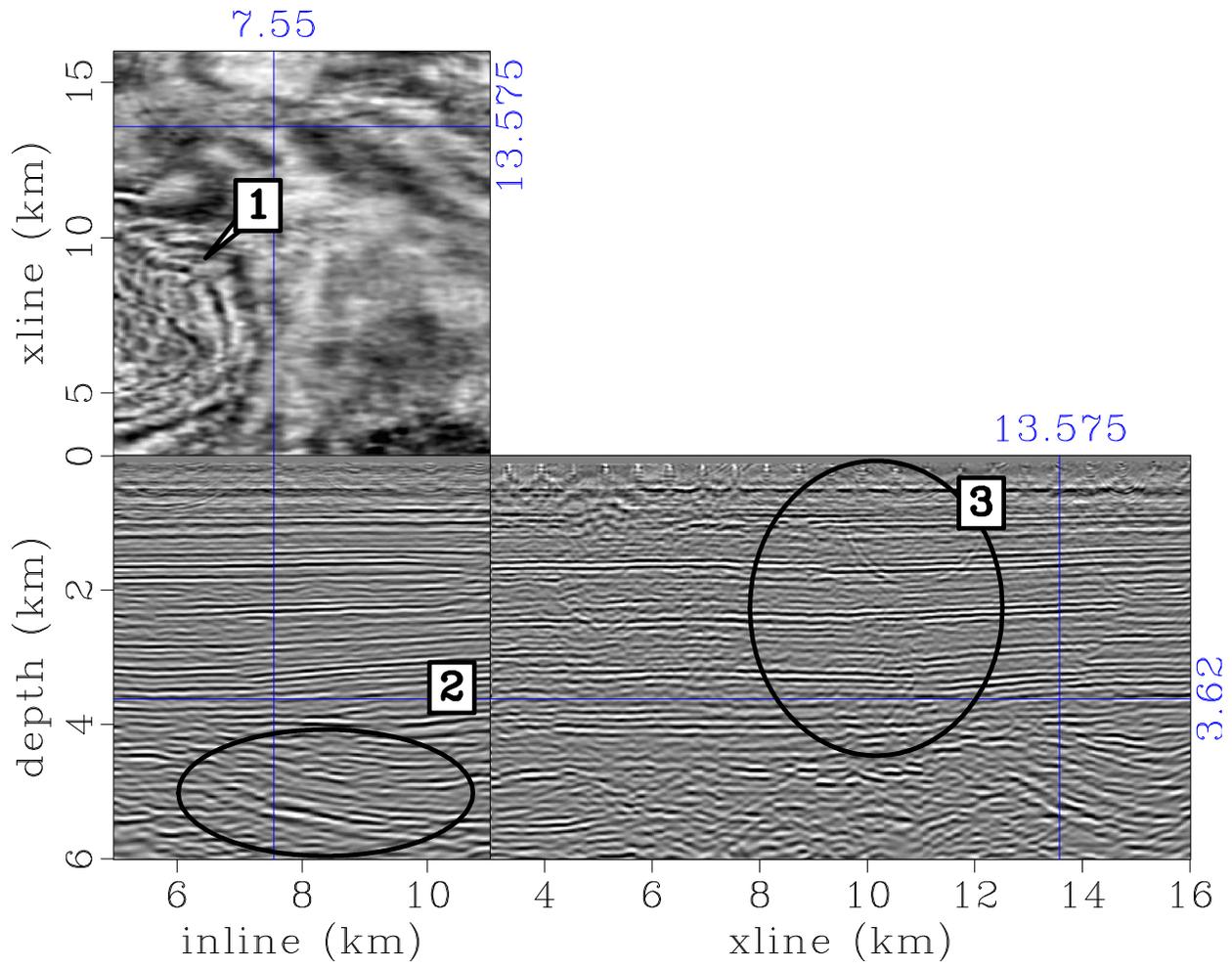


Figure 4.24: Migration image using the initial anisotropic model. [CR] chap4/. ani-iter0-3

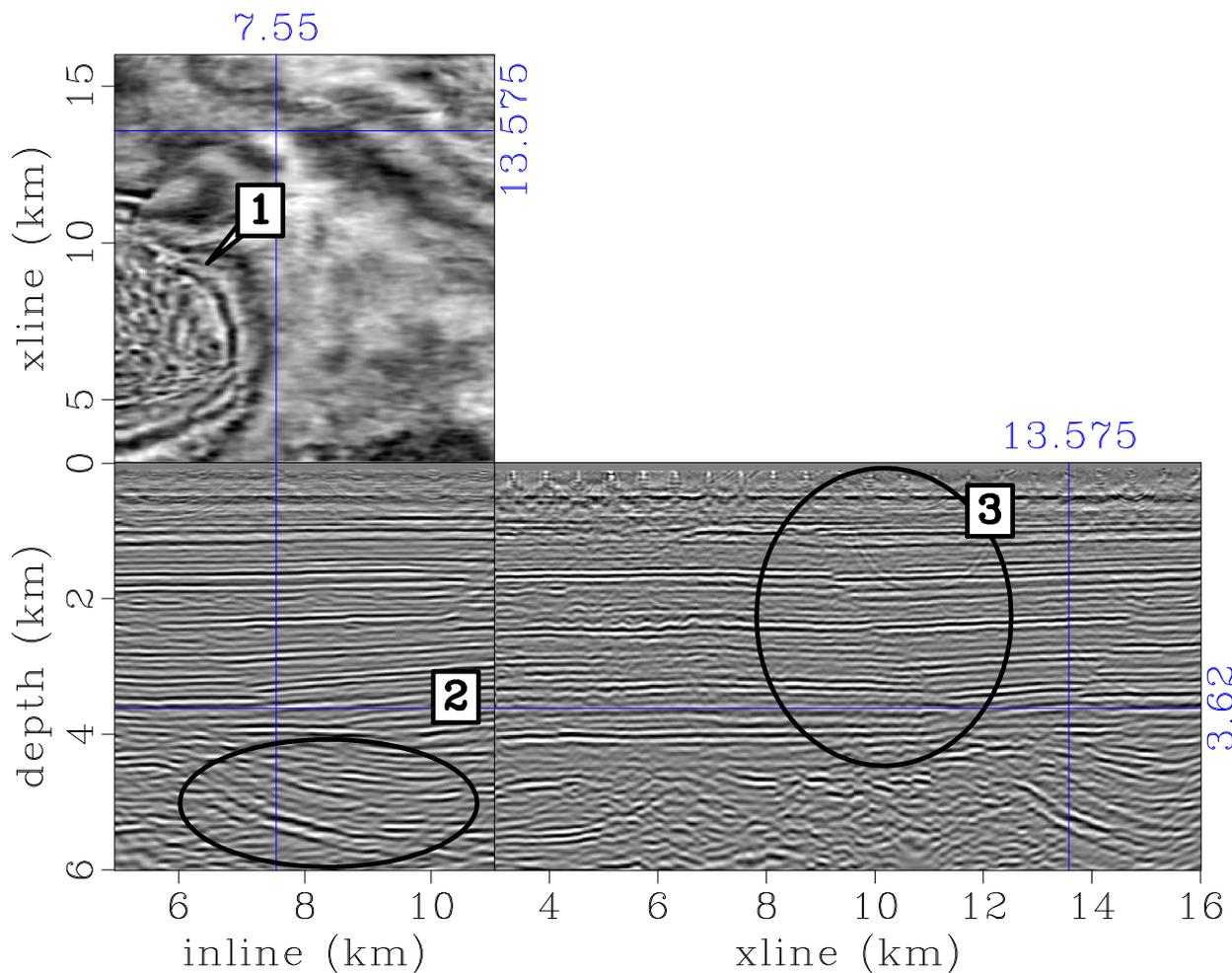


Figure 4.25: Migration image using the updated anisotropic model. Compare to Figure 4.24, flag 1 points out the much higher resolution dipping reflector around the salt. The reflections in circle 2 are imaged deeper with better continuity. Two normal faults on the cross-line section are distinctly imaged due to the improved focusing. Circle 3 highlights one of the two faults. [CR] chap4/. ani-iter10-3

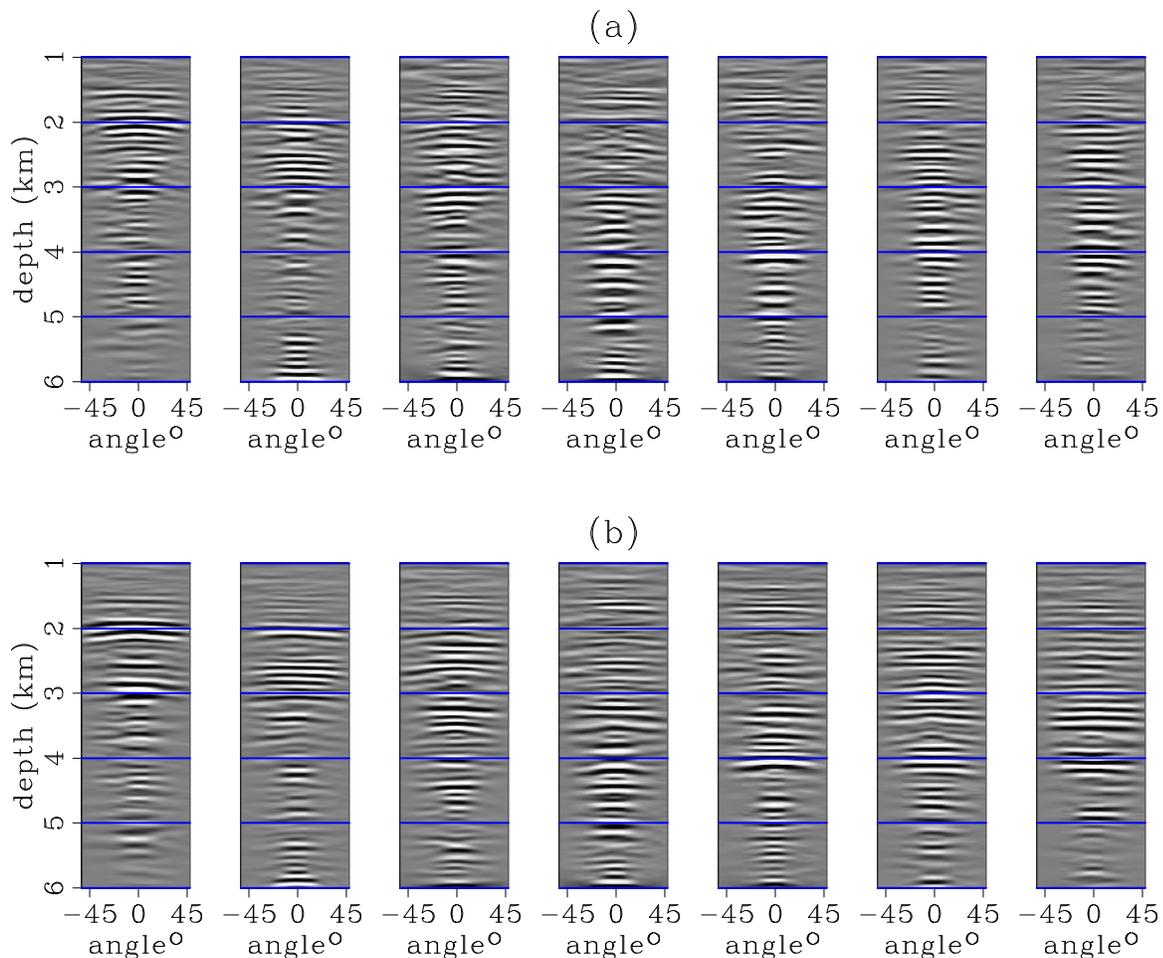


Figure 4.26: ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 5$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. Although the shallow curvatures are roughly unchanged, the updated ADCIGs show significantly improved angle coverage at medium depths (between 2 km and 4 km).

[CR] chap4/. adcig-x-y5000

DISCUSSION

Earth model building is a highly underdetermined problem and uncertainties propagate from each step of the model building process. In this study, I include the uncertainties of the rock physics modeling by sampling the distribution of the input

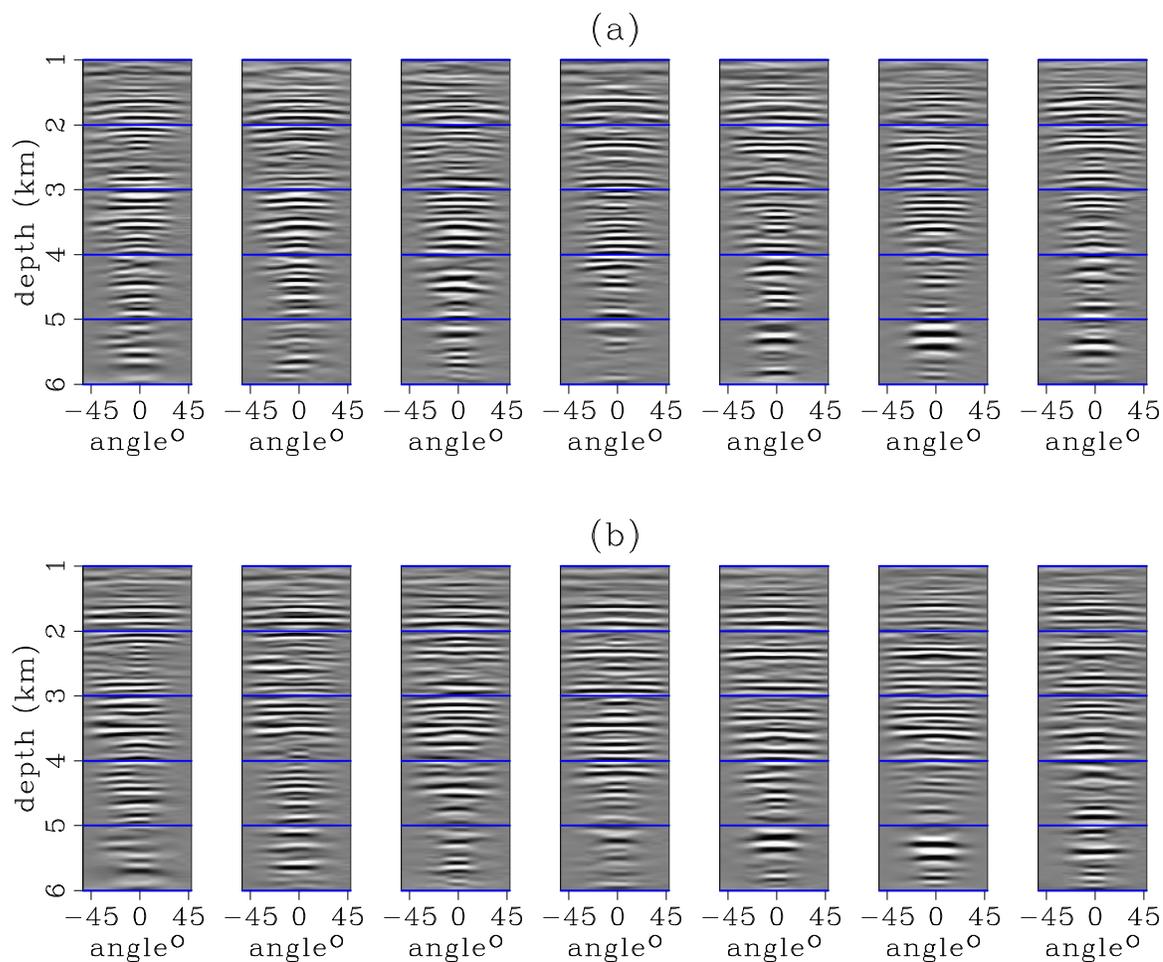


Figure 4.27: ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 13.5$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. The downward curvatures in the shallow gathers (above 2 km) are flattened in the updated ADCIGs. Angle coverages of the ADCIGs at medium depths (between 2 km and 4 km) are improved, especially for the horizon at roughly 3.5 km depth. In the deep section, the updated ADCIGs show wider angle coverage and higher resolution.

[CR] chap4/. adcig-x-y13500

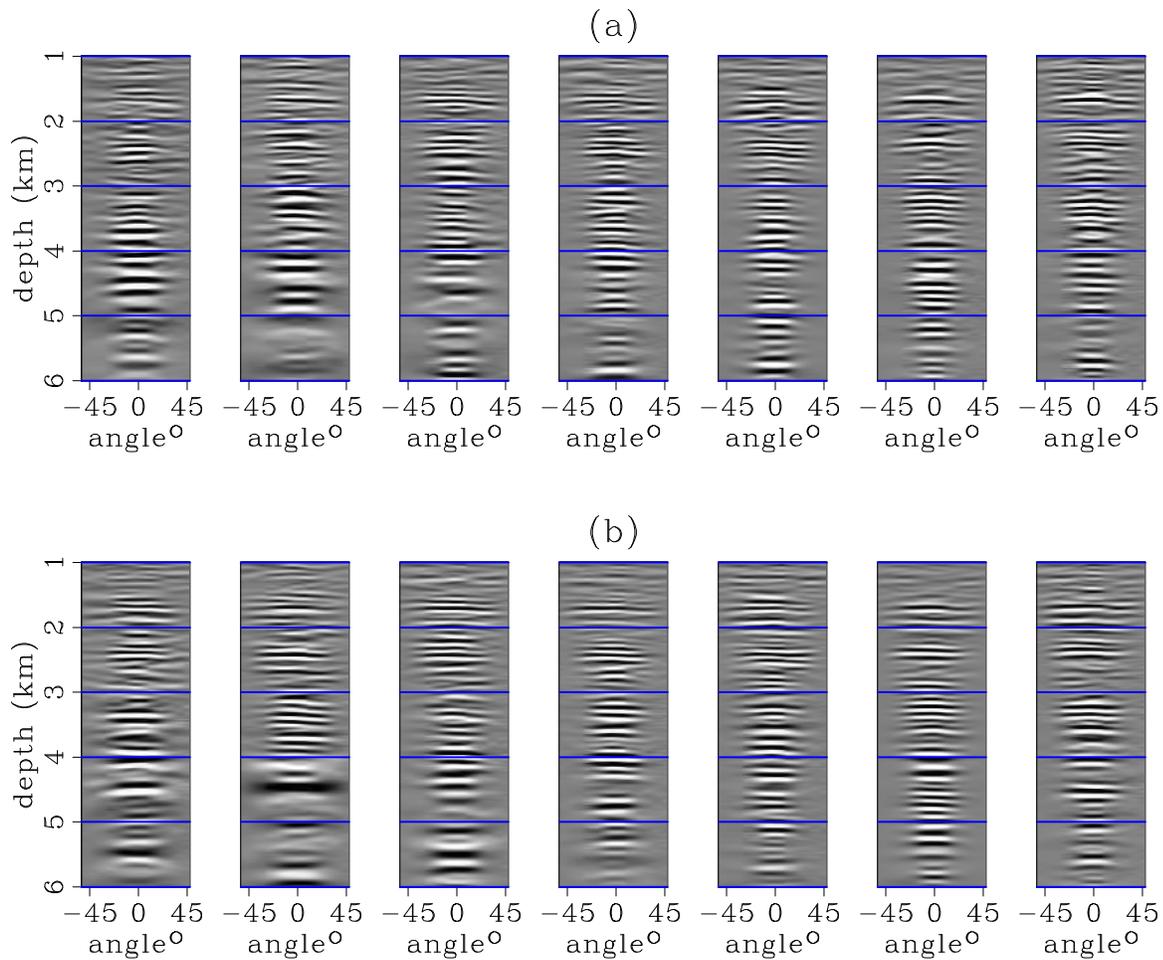


Figure 4.28: ADCIGs using the initial model (a) and the updated model (b). ADCIGs are extracted from $y = 10$ km and $x = 6, 6.5, 7, 7.5, 8, 8.5, 9$ km from left to right. Updated ADCIGs show strong improvements below 4 km (more significantly on the three left panels) due to the cumulative effects of anisotropy along the wavepaths.

[CR] chap4/. adcig-y-y10000

parameters and by using two different effective medium models. However, there are certainly more uncertainties with different geological scenarios which have been neglected. Moreover, the seismic data inversion is also nonunique. The current inversion practice of producing a single solution lacks uncertainty analysis, which can be highly valuable for decision making and risk management.

Using previous seismic inversion results to constrain the seismic model building in the next iteration helps us close the loop linking seismic data with reservoir modeling. Traditional processing from seismic data to a reservoir model does not include feedback; and the seismic data modeled from the inverted reservoir model usually do not match the field data. This study provides a step toward building a closed loop from exploration to production. With more iterations, Earth models that are consistent with all the available data could be obtained and their uncertainties should be evaluated.

CONCLUSIONS

In this chapter, I tested the rock physics constrained anisotropic WEMVA method on a 3-D field data set. I first build the rock physics constraints using stochastic rock physics modeling and then utilize these constraints during the anisotropic WEMVA inversion to produce reliable VTI models as well as better focused images.

Stochastic rock physics modeling is a powerful tool to model the shale anisotropy and to explore the possible ranges of the anisotropic parameters. By sampling the distributions of the key parameters, I include their uncertainties and produce an ensemble of anisotropic models which are realizable by rock physics modeling. The field data example demonstrates that it is reasonable to assume a multivariate Gaussian distribution and summarize the random variables using the mean and the covariance matrix. Migration images based on the mean model demonstrate that the rock physics modeling produces good initial anisotropic models for seismic imaging. A 3-D prior distribution model is made possible by utilizing the lithological inversion results of a previous seismic processing workflow.

By constraining the anisotropic WEMVA with the geological and rock physics covariance, I feed the prior information from the rock physics to the seismic data inversion, which significantly improves the convergence. The inverted VTI model not only explains the reflection data (flattens the gathers), but also follows the basic geological and rock physics principles. The 3-D example in this chapter demonstrates that anisotropic WEMVA can further improve the subsurface model and further focus the subsurface image especially for the steeply dipping reflectors around the salt body.

ACKNOWLEDGEMENTS

The author thanks Schlumberger MultiClient for providing the field dataset. This paper includes data supplied by IHS Energy Log Services; Copyright (2013) IHS Energy Log Services Inc. I acknowledge the Stanford Center for Computational Earth and Environmental Science for providing computing resources.

Chapter 5

Conclusions

In this dissertation, I present an integrated technique to build an anisotropic Earth model that satisfies seismic, geological, and rock physics constraints. In order to achieve this goal, I develop three key novel algorithms and one workflow for anisotropic model building.

First, I extend the isotropic wave-equation migration velocity analysis scheme to the anisotropic medium. I utilize the defocused energy to evaluate the accuracy of the background model and to provide meaningful updates in all parameters. I use a combined DSO operator and stacking power operator, which shares the merits of the global convergence of the DSO operator and the high resolution of the stacking power operator. Using an optimized finite differencing one-way wavefield extrapolation, the cost of the anisotropic WEMVA is on the same order as the isotropic WEMVA.

Second, I develop a new rock physics model for shale anisotropy that models both sand-shale lamination and sand as inclusions in the shale background material. This new model allows testing of multiple geological settings and rock physics scenarios. As a result, the uncertainties and the correlations in the anisotropic models are better captured.

Third, I develop a new regularization scheme to integrate geological and rock physics information to better constrain the seismic inversion. I summarize both types

of information into the covariance matrix. The spatial component, defined by the geological knowledge, and the cross-parameter component, defined by the rock physics modeling, are considered independent of each other. This formulation allows me to modify the unconstrained anisotropic WEMVA workflow by gradient preconditioning, a process whose cost is negligible.

Finally, I develop a workflow to best utilize all the available information. Start an anisotropic model building process by translating the well log and/or other available lithological information into anisotropic models using stochastic rock physics modeling. Obtain the average anisotropic model and the cross-parameter covariance matrix. Estimate the spatial covariance matrix using the initial migration image. Set up the surface seismic data inversion using the average anisotropic model as the initial solution and the covariance matrix as the constraints.

I test the algorithms and the workflow developed in this dissertation extensively on synthetic, 2-D field, and 3-D field datasets. The inversion yields models that simultaneously explain the seismic data and satisfy the geological and rock physics constraints. The corresponding migration images show faults with higher resolution, better defined depths, and better imaged steeply dipping reflectors.

Appendix A

Implicit finite differencing

Assuming the S-wave velocity is much slower than the P-wave velocity, we can approximate the dispersion relationship for VTI media as follows (Shan, 2009):

$$S_z = \sqrt{\frac{1 - (1 + 2\epsilon)S_r^2}{1 - 2(\epsilon - \delta)S_r^2}}, \quad (\text{A.1})$$

where $S_z = \frac{k_z}{\omega/v_v}$, $S_r = \frac{k_r}{\omega/v_v}$ with $k_r = \sqrt{k_x^2 + k_y^2}$, ω is the angular frequency, v_v is the vertical velocity, and ϵ and δ are the Thomsen parameters.

Shan (2009) suggests that the exact dispersion relationship A.1 can be approximated by a rational function $R_{n,m}(S_r)$:

$$R_{n,m}(S_r) = \frac{P_n(S_r)}{Q_m(S_r)}, \quad (\text{A.2})$$

where

$$P_n(S_r) = \sum_{i=0}^n a_i S_r^i \quad (\text{A.3})$$

and

$$Q_m(S_r) = \sum_{i=0}^m b_i S_r^i. \quad (\text{A.4})$$

Moreover, when the polynomials in equations A.3 and A.4 are of the same degree, namely $m = n$, dispersion relationship A.2 can be further split as follows:

$$S_z = 1 - \sum_{i=1}^n \frac{\alpha_i S_r^2}{1 - \beta_i S_r^2}. \quad (\text{A.5})$$

The coefficients α_i and β_i can be obtained by solving the least-square problem below:

$$\min \sum_{S_r} \left(\sqrt{\frac{1 - (1 + 2\epsilon) S_r^2}{1 - 2(\epsilon - \delta) S_r^2}} - \left(1 - \sum_{i=1}^n \frac{\alpha_i S_r^2}{1 - \beta_i S_r^2} \right) \right)^2. \quad (\text{A.6})$$

The tables for coefficients α and β for ϵ ranging from 0 to 0.2 and δ ranging from -0.004 to 0.2 are shown in Figure A.1. In general, coefficient α is more sensitive to the change in δ than to the change in ϵ . Coefficient β has similar sensitivities to both ϵ and δ .

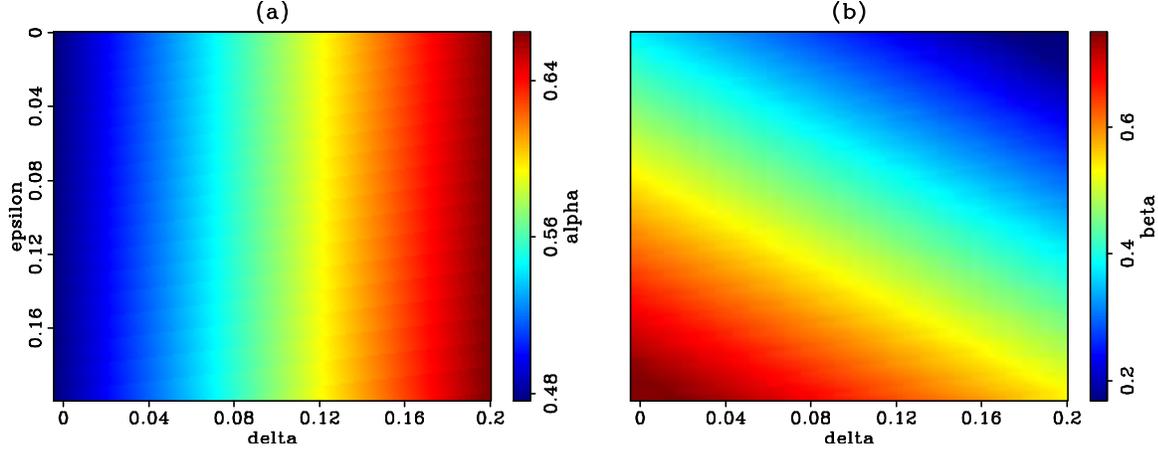


Figure A.1: (a) Table for α and (b) table for β at discrete ϵ and δ values. [ER] `append1/.coef`

In the downward extrapolation, the wavefield at the next depth (P_{z+1}) can be computed from the wavefield at the current depth (P_z) according to the following equation:

$$P_{z+1} = P_z e^{ik_z dz}, \quad (\text{A.7})$$

where $i = \sqrt{-1}$, dz is the extrapolation distance in depth, and k_z can be obtained from the first-order approximation of the dispersion relation A.5:

$$k_z = \frac{w}{v_v} \left(1 - \frac{\alpha \frac{k_r^2}{(w/v_v)^2}}{1 - \beta \frac{k_r^2}{(w/v_v)^2}} \right). \quad (\text{A.8})$$

Dispersion relation A.8 can be further simplified to polynomials using Taylor expansion:

$$\begin{aligned} k_z &= \frac{w}{v_v} \left(1 - \alpha \frac{k_r^2}{(w/v_v)^2} \left(1 + \beta \frac{k_r^2}{(w/v_v)^2} \right) \right) \\ &= \frac{w}{v_v} \left(1 - \alpha \frac{k_r^2}{(w/v_v)^2} - \alpha\beta \frac{k_r^4}{(w/v_v)^4} \right). \end{aligned} \quad (\text{A.9})$$

Therefore, the perturbed wavefield is

$$\Delta P_{z+1} = e^{ik_z dz} i dz P_z \Delta k_z, \quad (\text{A.10})$$

with

$$\Delta k_z = \frac{\partial k_z}{\partial v_v} \Delta v_v + \frac{\partial k_z}{\partial \epsilon} \Delta \epsilon + \frac{\partial k_z}{\partial \delta} \Delta \delta, \quad (\text{A.11})$$

$$\frac{\partial k_z}{\partial v_v} = -\frac{w}{v_v^2} \left(1 + \alpha \frac{k_r^2}{(w/v_v)^2} + 3\alpha\beta \frac{k_r^4}{(w/v_v)^4} \right), \quad (\text{A.12})$$

$$\frac{\partial k_z}{\partial \epsilon} = -\frac{w}{v_v} \left(\frac{\partial \alpha}{\partial \epsilon} \frac{k_r^2}{(w/v_v)^2} + \left(\frac{\partial \alpha}{\partial \epsilon} \beta + \alpha \frac{\partial \beta}{\partial \epsilon} \right) \frac{k_r^4}{(w/v_v)^4} \right), \quad (\text{A.13})$$

and

$$\frac{\partial k_z}{\partial \delta} = -\frac{w}{v_v} \left(\frac{\partial \alpha}{\partial \delta} \frac{k_r^2}{(w/v_v)^2} + \left(\frac{\partial \alpha}{\partial \delta} \beta + \alpha \frac{\partial \beta}{\partial \delta} \right) \frac{k_r^4}{(w/v_v)^4} \right). \quad (\text{A.14})$$

Since the finite difference parameters α and β are obtained by optimization, the derivatives in Equation A.13 and Equation A.14 are obtained numerically by taking derivatives along the ϵ and δ axis in Figure A.1. The tables of the derivatives of the coefficients with respect to the anisotropic parameters are shown in Figure A.2.

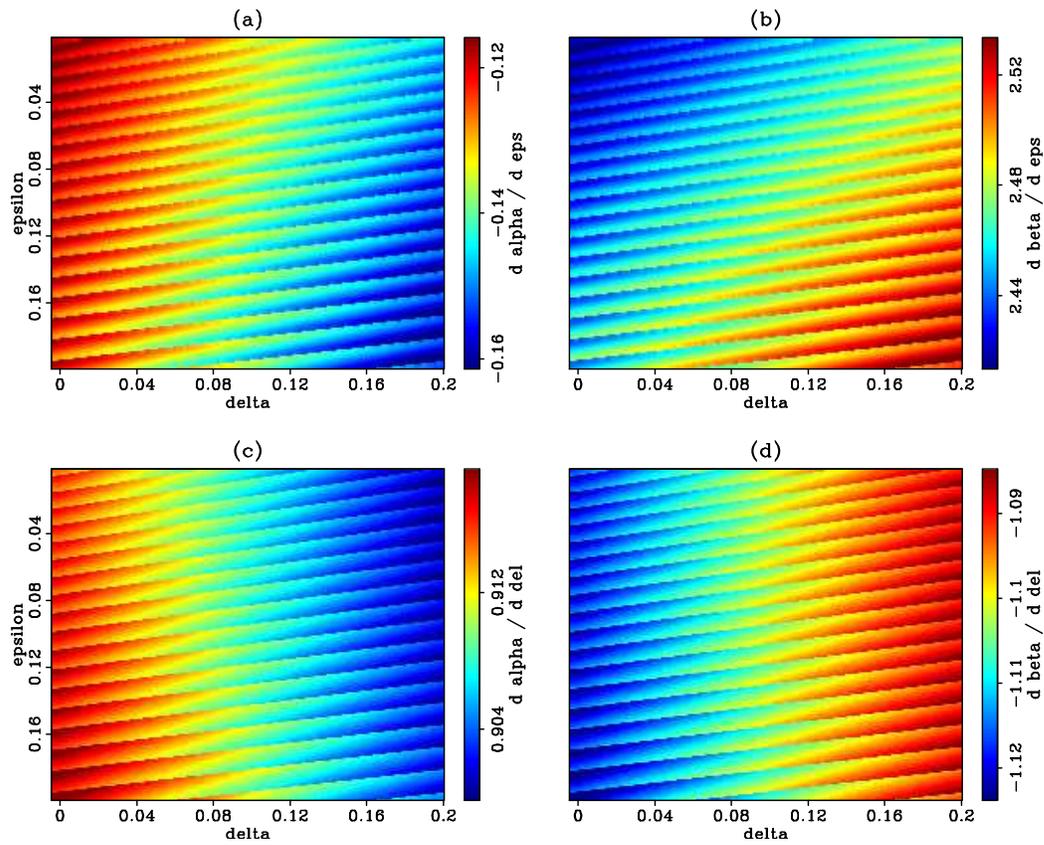


Figure A.2: Derivatives of the coefficients with respect to ϵ and δ ; (a) table of $\frac{\partial\alpha}{\partial\epsilon}$; (b) table of $\frac{\partial\beta}{\partial\epsilon}$; (c) table of $\frac{\partial\alpha}{\partial\delta}$, and (d) table of $\frac{\partial\beta}{\partial\delta}$. [ER] `append1/. dr-coef`

Appendix B

Rock physics modeling

I model the shale anisotropy from three aspects: mineral anisotropy of the rock constituents, compaction effect on the particle alignment, and the transition from smectite to illite due to compaction and temperature. To combine the anisotropic shale (clay) and the isotropic sand (quartz), I investigate two different effective-media models: the suspension model which and the lamination model. The suspension model simulates quartz as spherical inclusions in the clay background. The lamination model assumes thin layering effects between sand and shale.

Mineral anisotropy

We assume shales have three end-member mineral constituents: smectite, illite and quartz. The elastic properties are listed in Table B.1. The values of the smectite elasticity are the anisotropic elasticity values for a Cretaceous shale (Hornby et al., 1994). These approximated values are used as an end member when pure shale is fully compacted. Anisotropic elasticity for illite (muscovite) comes from work by Wenk et al. (2007). Similar to smectite, although the elasticity of a quartz crystal may be anisotropic, we assume isotropic quartz to approximate pure sand as an end-member in the rock.

Mineral	ρ (g/cc)	v_v (km/s)	v_s (km/s)	ϵ	δ	γ
Smectite	2.4	3.075	1.5	0.255	-0.05	0.48
Illite	2.4	4.94	2.6	1.02	0.	1.68
Quartz	2.65	6.0	4.0	0.	0.	0.

Table B.1: End-member mineral elastic properties.

Smectite to illite transition

The transition from smectite to illite is a common mineralogical reaction during burial diagenesis of shale. Many studies (e.g., Hower et al. (1976)) have shown that this transition reaction is controlled by the temperature in the subsurface. In this paper, we follow the work of Bachrach (2010a) to calibrate the percentage of illite P_I to the temperature T as follows:

$$P_I(T) = 0.5 + 0.5 \tanh\left(\frac{T - T_t}{2\sigma_t}\right), \quad (\text{B.1})$$

with T_t as the transition temperature and σ_t as the length of the transition window. Reference values $T_t = 58^\circ\text{C}$ and $\sigma_t = 60^\circ\text{C}$ are from the work of Freed and Peacor (1989).

Preferred orientation distribution of clay mineral

Preferred orientation of the clay minerals is another important factor to the shale anisotropy (Hornby et al., 1994; Sayers, 2004). When initially deposited, mineral domains are oriented in random directions. In this case, even though an individual mineral domain can be anisotropic, the effective medium with randomly oriented domains is isotropic. During maximum compaction, all the mineral domains are fully aligned, which produces the effective medium with maximum anisotropy.

According to Bandyopadhyay (2009), the Voigt averaged stiffness coefficients C_{ij}^a

are

$$\begin{aligned}
C_{11}^a &= L + 2M + \frac{4\sqrt{2}}{105}\pi^2(2\sqrt{5}a_3W_{200} + 3a_1W_{400}); \\
C_{33}^a &= L + 2M - \frac{16\sqrt{2}}{105}\pi^2(\sqrt{5}a_3W_{200} - 2a_1W_{400}); \\
C_{12}^a &= L - \frac{4\sqrt{2}}{315}\pi^2(2\sqrt{5}(7a_2 - a_3)W_{200} - 3a_1W_{400}); \\
C_{13}^a &= L + \frac{4\sqrt{2}}{315}\pi^2(\sqrt{5}(7a_2 - a_3)W_{200} - 12a_1W_{400}); \\
C_{44}^a &= M - \frac{2\sqrt{2}}{315}\pi^2(\sqrt{5}(7a_2 + a_3)W_{200} + 24a_1W_{400}); \\
C_{66}^a &= \frac{\langle C_{11} - C_{12} \rangle}{2},
\end{aligned} \tag{B.2}$$

where

$$\begin{aligned}
a_1 &= C_{11} + C_{33} - 2C_{13} - 4C_{44}; \\
a_2 &= C_{11} - 3C_{12} + 2C_{13} - 2C_{44}; \\
a_3 &= 4C_{11} - 3C_{33} - C_{13} - 2C_{44}; \\
L &= \frac{1}{15}(C_{11} + C_{33} + 5C_{12} + 8C_{13} - 4C_{44}); \\
M &= \frac{1}{30}(7C_{11} + 2C_{33} - 5C_{12} - 4C_{13} + 12C_{44}),
\end{aligned} \tag{B.3}$$

with C_{ij} as the stiffness coefficients of the individual domain. The exponents W_{200} and W_{400} define the compaction rate. We use the porosity as an indicator for compaction (Bachrach, 2010a).

$$\begin{aligned}
W_{200}(\phi) &= W_{200}^{\max}(1 - \phi/\phi_0)^m, \\
W_{400}(\phi) &= W_{400}^{\max}(1 - \phi/\phi_0)^n,
\end{aligned} \tag{B.4}$$

with ϕ as the porosity at depth and ϕ_0 as the critical porosity. The choice of exponents m and n has not been well studied. We refer to Bachrach (2010a) and let both parameters vary between 0.5 and 2.

Suspension model: Anisotropic differential effective medium

One way to model sandy shales is to model quartz as an inclusion in the clay background. We use the anisotropic differential effective medium (DEM) method (Hornby et al., 1994) as implemented by Bandyopadhyay (2009). This process begins with an effective background shale with $\phi = 50\%$ modeled by self-consistent approximation (Berryman, 1980b,a), and it models the effective properties at other values of ϕ by successive operations of removing an infinitesimal subvolume of host material and replacing it with a corresponding subvolume of quartz. At each successive increment of components, the previous step is taken as the host material.

Lamination model: Backus average

At depth, seismic wavelengths can get as large as a few hundred meters. These long-wavelength seismic waves cannot resolve individual layers, but instead they interact with the subsurface as a single averaged medium. Elastic properties of an effective medium composed of fine-scale laminations of sand and shale can be described by the Backus average. We assume the sand layer contains pure sand and the shale layer contains pure smectite and illite. Backus (1962) showed that the elastic constants of the effective medium can be obtained by the elastic constants of the individual layers as follows:

$$\begin{aligned}
 C_{11} &= \langle c_{13}/c_{33} \rangle^2 / \langle 1/c_{33} \rangle - \langle C_{13}^2 \rangle + \langle c_{11} \rangle; & (B.5) \\
 C_{12} &= C_{11} - \langle c_{11} \rangle + \langle c_{12} \rangle; \\
 C_{13} &= \langle c_{13}/c_{33} \rangle^2; \\
 C_{33} &= \langle 1/c_{33} \rangle^{-1}; \\
 C_{44} &= \langle 1/c_{44} \rangle^{-1},
 \end{aligned}$$

where $\langle . \rangle$ indicates the averages of the enclosed properties weighted by their volumetric proportions. The volumetric proportions for each lithological components are calculated from the shale content and the percentage of illite. These enclosed

properties can be averaged values over orientation distribution functions for shales.

Workflow

Our anisotropic rock physics modeling follows the workflow described here:

- Compute the percentage of illite in the rock, given a temperature model.
- Compute the average stiffness coefficients for smectite and illite, given a porosity model.
- Compute the volumetric percentage for each mineral phase, given a volumetric percentage of shale.
- Compute the stiffness coefficients using the suspension model or the lamination model.

At each instance of the modeling, the key parameters: ϕ_0 , exponents m and n , T_t , and σ_t are varied within a certain range. Therefore, an assembly of models are obtained. These models will be the source of the prior rock physics covariance.

Appendix C

RTM-based WEMVA for VTI models

FIRST-ORDER TWO-WAY VTI WAVE-EQUATION

The first-order two-way vertical transverse isotropic (VTI) wave-equation can be derived from Hooke's law and Newton's law using Thomson anisotropy parameters (ϵ , δ) and setting shear wave velocity to $c_s = 0$ (Duveneck et al., 2008). The first-order system reads as follows:

$$\begin{aligned}\rho\partial_t v_x &= -\partial_x p_H \\ \rho\partial_t v_y &= -\partial_y p_H \\ \rho\partial_t v_z &= -\partial_z p_V \\ \frac{1}{\rho c^2}\partial_t p_V &= -\sqrt{(1+2\delta)}(\partial_x v_x + \partial_y v_y) - \partial_z v_z + f_V \\ \frac{1}{\rho c^2}\partial_t p_H &= -(1+2\epsilon)(\partial_x v_x + \partial_y v_y) - \sqrt{(1+2\delta)}\partial_z v_z + f_H\end{aligned}\tag{C.1}$$

where ρ is the density, c is the velocity, (v_x, v_y, v_z) is the particle velocity vector, and p_V and p_H are pressure in the vertical and horizontal directions, respectively. The

source term f_V and f_H are defined by the source wavelet $w(t)$ as follows:

$$f_V(t) = f_H(t) = \int_{-\infty}^t w(\tau) d\tau. \quad (\text{C.2})$$

When $\rho = 1$, $\epsilon = 0$ and $\delta = 0$, the first-order system C.1 is equivalent to the familiar isotropic acoustic second-order wave-equation:

$$\frac{1}{c^2} \partial_t^2 p - \nabla p = w. \quad (\text{C.3})$$

For simplicity, we can rewrite system C.1 in a matrix-vector notation:

$$\mathbf{L}(c)\mathbf{p} = \mathbf{f}, \quad (\text{C.4})$$

where $\mathbf{p} = (v_x, v_y, v_z, p_V, p_H)^T$, $\mathbf{f} = (0, 0, 0, f_V, f_H)^T$, and

$$\mathbf{L} = \begin{pmatrix} \partial_t & 0 & 0 & 0 & \partial_x \\ 0 & \partial_t & 0 & 0 & \partial_y \\ 0 & 0 & \partial_t & \partial_z & 0 \\ \sqrt{1+2\delta}\partial_x & \sqrt{1+2\delta}\partial_y & \partial_z & \frac{1}{c^2}\partial_t & 0 \\ (1+2\epsilon)\partial_x & (1+2\epsilon)\partial_y & \sqrt{1+2\delta}\partial_z & 0 & \frac{1}{c^2}\partial_t \end{pmatrix}. \quad (\text{C.5})$$

VTI REVERSE-TIME MIGRATION IMAGING CONDITION

Traditionally, the subsurface image is often considered as the first gradient of a full waveform inversion (FWI) objective function with single scattered data with respect to velocity. We are going to derive the VTI reverse-time migration imaging condition according to the same criteria.

We define the FWI objective function as

$$J_w = \frac{1}{2} \langle d - d_{\text{est}}, d - d_{\text{est}} \rangle, \quad (\text{C.6})$$

where d_{est} is the data estimated from the current model, which is sampled from wavefield \mathbf{p} , and d is the recorded data.

For the first iteration, $d_{\text{est}} = 0$. Therefore, the first gradient in velocity is:

$$\begin{aligned}\nabla_c J_{\mathbf{w}} &= \left(\frac{\partial \mathbf{p}}{\partial c} \right)^* d \\ &= (-\mathbf{L}^{-1} \frac{\partial \mathbf{L}}{\partial c} \mathbf{L}^{-1} f)^* d.\end{aligned}\tag{C.7}$$

Now, we introduce the receiver vector field $\mathbf{q} = (u_x, u_y, u_z, q_V, q_H)^T$, which is the solution of the following equation:

$$\mathbf{L}^*(c)\mathbf{q} = \mathbf{f}'.\tag{C.8}$$

The equivalent source term in equation C.8 is defined as $\mathbf{f}' = (0, 0, 0, f'_V, f'_H)^T$, where $f'_V = f'_H = d$. From equation C.5, we have

$$\frac{\partial \mathbf{L}}{\partial c} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{2}{c^3} \partial_t & 0 \\ 0 & 0 & 0 & 0 & -\frac{2}{c^3} \partial_t \end{vmatrix},\tag{C.9}$$

If we plug equation C.8 and C.9 into equation C.7 and ignore the velocity dependence, we arrive at the imaging condition as follows:

$$\begin{aligned}I &= \mathbf{p}^* \mathbf{M}^* \mathbf{q}, \\ &= (\mathbf{M}\mathbf{p})^* \mathbf{q},\end{aligned}\tag{C.10}$$

where

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2}\partial_t & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2}\partial_t \end{pmatrix}. \quad (\text{C.11})$$

The explicit form of this imaging condition for acoustic RTM is:

$$I = \int_0^{t_{max}} \frac{1}{2} ((\partial_t p_H)q_H + (\partial_t p_V)q_V) dt. \quad (\text{C.12})$$

The scaling factor of $\frac{1}{2}$ is chosen to make sure that when $p_H = p_V$, equation C.12 reduces to the isotropic cross-correlation imaging condition (Claerbout, 1987). For the purpose of velocity analysis, we often work with extended images and generalized imaging conditions. Similarly, we define our subsurface-offset-domain common-image gathers (SODCIGs) \mathbf{I} as a column vector:

$$\mathbf{I} = [I_{-\mathbf{h}_{max}}, I_{-\mathbf{h}_{max}+\Delta\mathbf{h}}, \dots, I_{\mathbf{0}}, \dots, I_{\mathbf{h}_{max}-\Delta\mathbf{h}}, I_{\mathbf{h}_{max}}]^*, \quad (\text{C.13})$$

where \mathbf{h} is the half-subsurface offset, which ranges from $-\mathbf{h}_{max}$ to \mathbf{h}_{max} with an increment of $\Delta\mathbf{h}$. For each element $I_{\mathbf{h}}$, the extended imaging condition is as follows (Sava and Formel, 2006):

$$I_{\mathbf{h}} = (\mathbf{S}_{+\mathbf{h}}\mathbf{p})^*\mathbf{M}^*(\mathbf{S}_{-\mathbf{h}}\mathbf{q}), \quad (\text{C.14})$$

where $\mathbf{S}_{+\mathbf{h}}$ is a shifting operator which shifts the wavefield by an amount of $+\mathbf{h}$ in the \mathbf{x} direction. Notice that $(\mathbf{S}_{+\mathbf{h}})^* = \mathbf{S}_{-\mathbf{h}}$.

MIGRATION VELOCITY ANALYSIS GRADIENTS

In this section, we derive the MVA gradients of objective function C.15 by two different approaches: the adjoint method from the perturbation theory, and the Lagrangian augmented function.

WEMVA is a non-linear inversion process that aims to find the velocity model that minimizes the residual field $\Delta\mathbf{I}$ in the image space. Without losing any generality, we define our objective function by DSO (Shen and Symes, 2008) in the subsurface offset domain:

$$J = \frac{1}{2} \sum_{\mathbf{h}} \langle \mathbf{h}I_{\mathbf{h}}, \mathbf{h}I_{\mathbf{h}} \rangle. \quad (\text{C.15})$$

Although we do not use this DSO objective function in the example, the derivation follows the same logic, and readers can easily substitute their desired image-space objective function into the derivation.

Adjoint method from the perturbation theory

A perturbation, δc , of the velocity model c , induces a perturbation $\delta\mathbf{p}$ in the source wavefield vector \mathbf{p} , a perturbation $\delta\mathbf{q}$ in the receiver wavefield vector \mathbf{q} , a perturbation $\delta\mathbf{I}$ in the extended image cube \mathbf{I} , and hence a perturbation δJ in the objective function J . To the first order and using the chain rule, δJ and δc have the following relationship:

$$\delta J = \sum_{\mathbf{h}} \frac{\partial J}{\partial I_{\mathbf{h}}} \frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial c} \delta c + \sum_{\mathbf{h}} \frac{\partial J}{\partial I_{\mathbf{h}}} \frac{\partial I_{\mathbf{h}}}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial c} \delta c. \quad (\text{C.16})$$

Now, we can define the gradient by the back-projection of a unit perturbation in the objective function:

$$\begin{aligned} \nabla_c J &= \sum_{\mathbf{h}} \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial c} \right)^* + \sum_{\mathbf{h}} \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \frac{\partial I_{\mathbf{h}}}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial c} \right)^* \\ &= (\nabla_c J)_1 + (\nabla_c J)_2. \end{aligned} \quad (\text{C.17})$$

We analyze the first term in equation C.17 in detail, and the second term follows the same reasoning.

$$\begin{aligned}
(\nabla_c J)_1 &= \sum_{\mathbf{h}} \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \frac{\partial \mathbf{p}}{\partial c} \right)^* \\
&= \sum_{\mathbf{h}} \left(\frac{\partial \mathbf{p}}{\partial c} \right)^* \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \right)^* \\
&= \sum_{\mathbf{h}} \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c} \right)^* \mathbf{L}^{-*} \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \right)^* \\
&= \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c} \right)^* \mathbf{L}^{-*} \sum_{\mathbf{h}} \left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \right)^* \left(\frac{\partial J}{\partial I_{\mathbf{h}}} \right)^* \tag{C.18}
\end{aligned}$$

where

$$\left(\frac{\partial J}{\partial I_{\mathbf{h}}} \right)^* = \mathbf{h}^* \mathbf{h} I_{\mathbf{h}}, \tag{C.19}$$

and

$$\begin{aligned}
\left(\frac{\partial I_{\mathbf{h}}}{\partial \mathbf{p}} \right)^* &= (\mathbf{S}_{+\mathbf{h}})^* \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}) \\
&= \mathbf{S}_{-\mathbf{h}} \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}). \tag{C.20}
\end{aligned}$$

By plugging equations C.19 and C.20 into equation C.18, we can rewrite equation C.18 explicitly as follows:

$$\left(\frac{\partial J}{\partial c} \right)_1 = \mathbf{p}^* \left(-\frac{\partial \mathbf{L}}{\partial c} \right)^* \mathbf{L}^{-*} \sum_{\mathbf{h}} \mathbf{S}_{-\mathbf{h}} \mathbf{M}^* (\mathbf{S}_{-\mathbf{h}} \mathbf{q}) \mathbf{h}^2 I_{\mathbf{h}}. \tag{C.21}$$

Similarly, we can obtain the explicit form for the second term in equation C.17:

$$\left(\frac{\partial J}{\partial c} \right)_2 = \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial c} \right) \mathbf{L}^{-1} \sum_{\mathbf{h}} \mathbf{S}_{+\mathbf{h}} \mathbf{M} (\mathbf{S}_{+\mathbf{h}} \mathbf{p}) \mathbf{h}^2 I_{\mathbf{h}}. \tag{C.22}$$

By substituting equation C.21 and equation C.22 for the corresponding terms in equation C.17, we now have derived the explicit form for the DSO gradient.

Lagrangian augmented functional method

We are now going to use the recipe with the augmented functional that Plessix (2006) provided to derive the image-space DSO gradient. First, we form the Lagrangian augmented functional, \mathcal{L} :

$$\mathcal{L}(\mathbf{p}, \mathbf{q}, I_{\mathbf{h}}, \lambda, \mu, \gamma_{\mathbf{h}}, c) = \quad (\text{C.23})$$

$$\begin{aligned} & \sum_{\mathbf{h}} \frac{1}{2} \langle \mathbf{h} I_{\mathbf{h}}, \mathbf{h} I_{\mathbf{h}} \rangle \\ & + \langle \lambda, \mathbf{f} - \mathbf{L}(c)\mathbf{p} \rangle \\ & + \langle \mu, \mathbf{f}' - \mathbf{L}^*(c)\mathbf{q} \rangle \\ & + \sum_{\mathbf{h}} \langle \gamma_{\mathbf{h}}, (\mathbf{S}_{+\mathbf{h}}\mathbf{p})^* \mathbf{M}^*(\mathbf{S}_{-\mathbf{h}}\mathbf{q}) - I_{\mathbf{h}} \rangle. \end{aligned} \quad (\text{C.24})$$

Then, the adjoint state equations are obtained by taking the derivative of \mathcal{L} with respect to state variables \mathbf{p} , \mathbf{q} and $I_{\mathbf{h}}$:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{p}} = -\mathbf{L}^*(c)\lambda + \sum_{\mathbf{h}} (\mathbf{S}_{+\mathbf{h}})^* \mathbf{M}^*(\mathbf{S}_{-\mathbf{h}}\mathbf{q}) \gamma_{\mathbf{h}} = \mathbf{0}, \quad (\text{C.25})$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\mathbf{L}(c)\mu + \sum_{\mathbf{h}} (\mathbf{S}_{-\mathbf{h}})^* \mathbf{M}(\mathbf{S}_{+\mathbf{h}}\mathbf{p}) \gamma_{\mathbf{h}} = \mathbf{0}, \quad (\text{C.26})$$

$$\frac{\partial \mathcal{L}}{\partial I_{\mathbf{h}}} = -\gamma_{\mathbf{h}} + \mathbf{h}^2 I_{\mathbf{h}} = 0, \forall \mathbf{h}. \quad (\text{C.27})$$

Equation C.25, C.26, C.27 are the adjoint-state equations. Variables $\lambda = (\lambda_x, \lambda_y, \lambda_z, \lambda_V, \lambda_H)^T$ and $\mu = (\mu_x, \mu_y, \mu_z, \mu_V, \mu_H)^T$ are the adjoint-state fields and the solution of the adjoint-state equations C.25 and C.26. Variable $\gamma_{\mathbf{h}}$ is the scaled image slice at the subsurface offset \mathbf{h} .

Now, the gradient of the objective function C.15 with respect to velocity is:

$$\begin{aligned} \nabla_c J &= \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \mathbf{c}} \mathbf{p} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \mathbf{c}} \mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial \mathbf{c}} \mathbf{p} \right)^* \lambda + \left(-\frac{\partial \mathbf{L}^*}{\partial \mathbf{c}} \mathbf{q} \right)^* \mu, \end{aligned} \quad (\text{C.28})$$

If we combine equations C.25, C.26, and C.27 with equation C.28, we will arrive at the same solution as in the previous section.

Extension to update anisotropic parameters

The extension from isotropic model updates to anisotropic updates is straightforward. Built on the derivations in the last section, we can easily obtain the gradients for anisotropic parameters ϵ and δ as follows:

$$\begin{aligned}\nabla_{\epsilon} J &= \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \epsilon} \mathbf{p} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \epsilon} \mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial \epsilon} \mathbf{p} \right)^* \lambda + \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial \epsilon} \right) \mu,\end{aligned}\tag{C.29}$$

$$\begin{aligned}\nabla_{\delta} J &= \left\langle \lambda, -\frac{\partial \mathbf{L}}{\partial \delta} \mathbf{p} \right\rangle + \left\langle \mu, -\frac{\partial \mathbf{L}^*}{\partial \delta} \mathbf{q} \right\rangle \\ &= \left(-\frac{\partial \mathbf{L}}{\partial \delta} \mathbf{p} \right)^* \lambda + \mathbf{q}^* \left(-\frac{\partial \mathbf{L}}{\partial \delta} \right) \mu,\end{aligned}\tag{C.30}$$

where

$$\frac{\partial \mathbf{L}}{\partial \epsilon} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2\partial_x & 2\partial_y & 0 & 0 & 0 \end{vmatrix},\tag{C.31}$$

$$\frac{\partial \mathbf{L}}{\partial \delta} = \begin{vmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \frac{\partial_x}{\sqrt{1+2\delta}} & \frac{\partial_y}{\sqrt{1+2\delta}} & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial_z}{\sqrt{1+2\delta}} & 0 & 0 \end{vmatrix}.\tag{C.32}$$

It is well-known that δ is the parameter most poorly constrained by surface seismic. Therefore, in our study, we assume that δ is obtained from well logs or seismic-well ties, and we invert only for velocity and ϵ .

Physical interpretation and implementation of the DSO gradient

In this subsection, we interpret each term in the DSO gradient formulation, and provide the readers with some hints for implementation. We find the Lagrangian formulation is easy to interpret, and readers can clearly relate the corresponding terms to the adjoint formulation. We will only discuss the physical meaning and the implementation for the first term in the gradient (Equations C.28 and C.25). Then, similar reasoning can be argued using reciprocity.

First, for each image slice in the subsurface-offset domain $I_{\mathbf{h}}$, we compute a weighted image $\gamma_{\mathbf{h}}$ using equation C.27. Then, we move on to equation C.25. We can rearrange the independent and commutable operators as follows:

$$\mathbf{L}^*(c)\lambda = \sum_{\mathbf{h}} \mathbf{S}_{-\mathbf{h}} (\mathbf{S}_{-\mathbf{h}} \mathbf{M}^* \mathbf{q}) \gamma_{\mathbf{h}}. \quad (\text{C.33})$$

Operator \mathbf{M}^* corresponds to differentiating q_V and q_H once reversely in time and setting u_x, u_y , and u_z fields to zero. The directions of propagation and differentiation in time of wavefield \mathbf{q} are the same. Therefore, we can compute the time derivative during the same process as the propagation. Then we shift the reverse-time derivative \mathbf{q} by $-\mathbf{h}$ in \mathbf{x} , and multiply it with the weighted image $\gamma_{\mathbf{h}}$. This product is shifted again by $-\mathbf{h}$. Finally, we sum over the contributions from all subsurface-offset image slices to get an effective source term \mathbf{f}_p . Next, we solve equation C.33 for λ backward in time, using \mathbf{f}_p as the source.

In equation C.28, $-\frac{\partial \mathbf{L}}{\partial c}$ is a sparse matrix with non-zero elements only for p_V and

p_H . We can therefore write everything out explicitly:

$$(\nabla_c J)_1 = \int_0^{t_{max}} \frac{2}{c^3} [(\partial_t p_H)\lambda_H + (\partial_t p_V)\lambda_V] dt. \quad (C.34)$$

The explicit forms for the complete gradients are:

$$\begin{aligned} \nabla_c J &= \int_0^{t_{max}} \frac{2}{c^3} [(\partial_t p_H)\lambda_H + (\partial_t p_V)\lambda_V] dt \\ &+ \int_0^{t_{max}} \frac{2}{c^3} [(\partial_t q_H)\mu_H + (\partial_t q_V)\mu_V] dt \end{aligned} \quad (C.35)$$

and

$$\begin{aligned} \nabla_\epsilon J &= - \int_0^{t_{max}} [(\partial_x v_x)\lambda_H + (\partial_y v_y)\lambda_H] dt \\ &+ \int_0^{t_{max}} [(\partial_x q_H)\mu_x + (\partial_y q_H)\mu_y] dt. \end{aligned} \quad (C.36)$$

Preconditioning the DSO gradient

Velocity model building is a highly underdetermined and nonlinear problem, therefore prior knowledge of the subsurface is needed to define a plausible subsurface model. In the formulation of Tarantola (1984), prior information is included as the covariance and the mean of the model. In this study, we assume the initial model we use is the mean, and the covariance of the model has two independent components: spatial covariance and collocated cross-parameter covariance (Li et al., 2011). Instead of regularizing the inversion using Tarantola (1984), we use a preconditioning scheme (Claerbout, 2009): smoothing filtering to approximate square-root of the spatial covariance, and a standard-deviation matrix to approximate the square-root of the cross-parameter covariance.

Mathematically, the preconditioned model perturbation $d\mathbf{n}$ of the subsurface is defined as follows:

$$d\mathbf{m} = \mathbf{B}\Sigma d\mathbf{n}, \quad (C.37)$$

where $\mathbf{m} = [c \ \epsilon]^T$. The smoothing operator \mathbf{B} is a diagonal matrix:

$$\mathbf{B} = \begin{vmatrix} \mathbf{B}_c & 0 \\ 0 & \mathbf{B}_\epsilon \end{vmatrix}. \quad (\text{C.38})$$

with different smoothing operators for velocity and ϵ , according to the geological information in the study area. The standard deviation matrix Σ :

$$\Sigma = \begin{vmatrix} \sigma_{cc} & \sigma_{c\epsilon} \\ \sigma_{\epsilon c} & \sigma_{\epsilon\epsilon} \end{vmatrix}. \quad (\text{C.39})$$

can be obtained by rock-physics modeling and/or lab measurements (Bachrach et al., 2011; Li et al., 2011).

We call \mathbf{n} the preconditioning variable, and it relates to the original model \mathbf{m} as follows:

$$\mathbf{m} = \mathbf{B}\Sigma\mathbf{n} + (\mathbf{m}_0 - \mathbf{B}\Sigma\mathbf{n}_0), \quad (\text{C.40})$$

where \mathbf{n}_0 and \mathbf{m}_0 are the initial models in preconditioned space and physical space, respectively. Now, the gradient of the objective function C.15 with respect to this preconditioning variable \mathbf{n} is

$$\begin{aligned} \nabla_{\mathbf{n}}J &= \left(\frac{\partial\mathbf{m}}{\partial\mathbf{n}}\right)^* \nabla_{\mathbf{m}}J \\ &= \Sigma^* \mathbf{B}^* \nabla_{\mathbf{m}}J, \end{aligned} \quad (\text{C.41})$$

where $\nabla_{\mathbf{m}}J = [\nabla_c J \ \nabla_\epsilon J]^T$.

In a steepest-descent inversion framework, the initial preconditioning model \mathbf{n}_0 is obtained by minimizing the following objective function:

$$J_{\text{init}} = \frac{1}{2} \langle \mathbf{m}_0 - \mathbf{B}\Sigma\mathbf{n}_0, \mathbf{m}_0 - \mathbf{B}\Sigma\mathbf{n}_0 \rangle. \quad (\text{C.42})$$

For the i_{th} iteration

$$\mathbf{n}_{i+1} = \mathbf{n}_i + \alpha_i \nabla_{\mathbf{n}}J, \quad (\text{C.43})$$

$$\begin{aligned}
\mathbf{m}_{i+1} &= \mathbf{B}\Sigma\mathbf{n}_{i+1} \\
&= \mathbf{B}\Sigma\mathbf{n}_i + \alpha_i\mathbf{B}\Sigma\nabla_{\mathbf{n}}J \\
&= \mathbf{m}_i + \alpha_i\mathbf{B}\Sigma\Sigma^*\mathbf{B}^*\nabla_{\mathbf{m}}J.
\end{aligned} \tag{C.44}$$

Equation C.44 suggests an interesting consideration in the context of nonlinear inversion: left-multiplying the gradient with a (semi)positive-definite matrix is equivalent to preconditioning with the square-root of the matrix; thus, the resulting direction is still a descent direction (Claerbout, 2009).

NUMERICAL TEST

In this section, we present the tests of our method on a synthetic VTI Marmousi model. We perform the tests in three steps: objective-function test, single parameter inversion, and joint inversion for two parameters.

Objective function test

We test our method on a VTI Marmousi model. First, synthetic Born data is generated using the models in Figures C.1 and C.2. The maximum offset is 3 km. Next, we perturb the ϵ model (Figure C.2(a)) by a very smooth $\delta\epsilon$ field, as shown in Figure C.3(a). We change the perturbation from -50% to 50% of the true ϵ model, and calculate the corresponding objective function.

Ideally, we would like to choose an objective function that reaches a local minimum at the correct model and is quadratic around the correct model, so that a gradient-based inversion scheme is guaranteed to converge. Based on the results, we choose an angle domain objective function instead of the DSO objective function (Equation C.15):

$$J = \frac{1}{2}\langle \mathbf{DRI}, \mathbf{DRI} \rangle, \tag{C.45}$$

where \mathbf{R} is the Radon transform operator, and \mathbf{D} is the derivative operator along the

ray-parameter axis.

As shown in Figure C.3(b), the angle-domain objective function has a minimum at the correct epsilon model, and it has a semi-quadratic shape with respect to the model perturbation. Therefore, this objective function is a good measure of the error in the anisotropic model. The tilting effect toward negative ϵ perturbation is caused by the limited acquisition geometry. This effect is negligible for velocity perturbation because velocity has a first-order effect on the flatness of the angle gather; ϵ 's effect is second-order. We can increase the acquisition offset to mitigate this tilting effect and help the inversion.

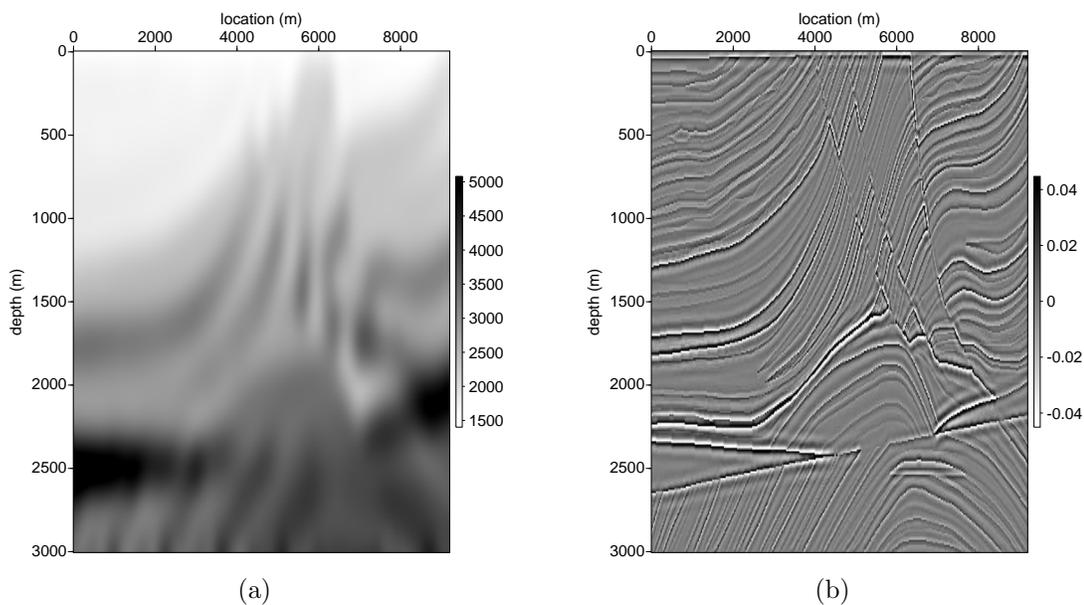


Figure C.1: Smooth velocity model (a) and reflectivity model (b) used to generate the synthetic Born data. [CR] `append3/. vtrue,ref`

Single parameter inversion

In this subsection, we invert for the anisotropic parameter ϵ alone. In this test, we model the synthetic data using very smooth ϵ (Figure C.4(a)) and δ (Figure C.4(b)) models as suggested by many field applications. To better constrain the inversion for

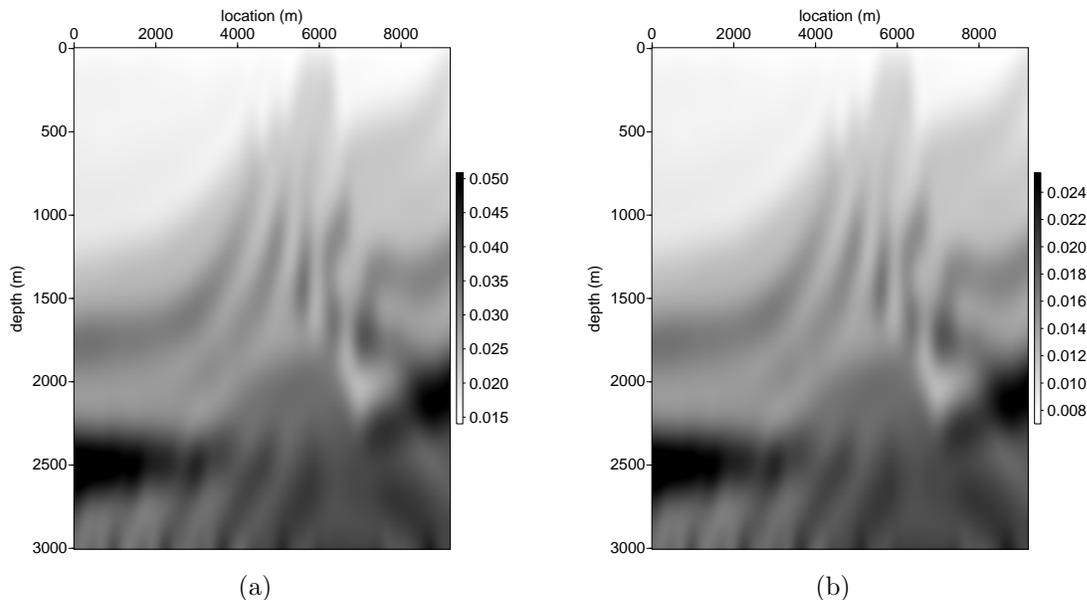


Figure C.2: The ϵ model (a) and δ model (b) used to generate the synthetic Born data. [CR] `append3/. eps,del`

ϵ , we also increase the maximum offset in the acquisition to 6 km.

Compared with the true ϵ model, our initial ϵ model (Figure C.5(a)) has negative perturbation of about 50% in the shallower part. Because a perfect velocity model is used in this case, the moveout at large angles is so small that it is almost undetectable to human eyes (Figure C.5(b)). However, our inversion scheme is very sensitive to the residual moveout and successfully updates the ϵ model in the correct direction. Figure C.6 shows the inverted ϵ model and the corresponding angle-domain common-image gathers after 40 iterations. By comparing with the initial angle gathers (Figure C.5(b)), we see that the slightly curving events at large angles are flattened and the inverted ϵ model is closer to the true one.

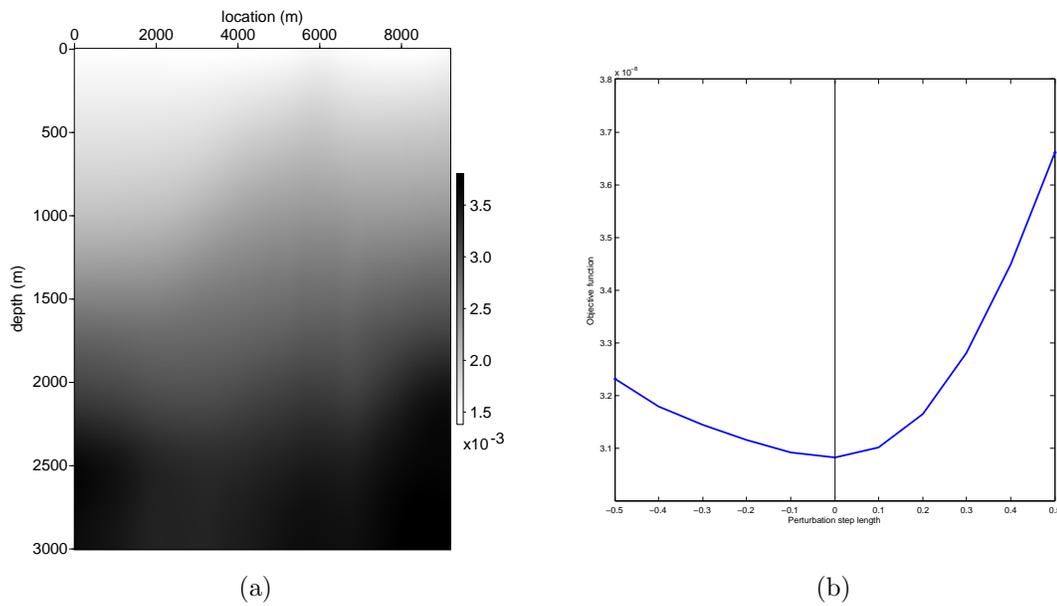


Figure C.3: (a) The $\delta\epsilon$ model to test the objective function. (b) Objective function vs. ϵ perturbation. The angle-domain objective function C.45 has a minimum at the correct epsilon model, and it has a semi-quadratic shape with respect to the model perturbation. [CR] `append3/.deltaeps,objcurve`

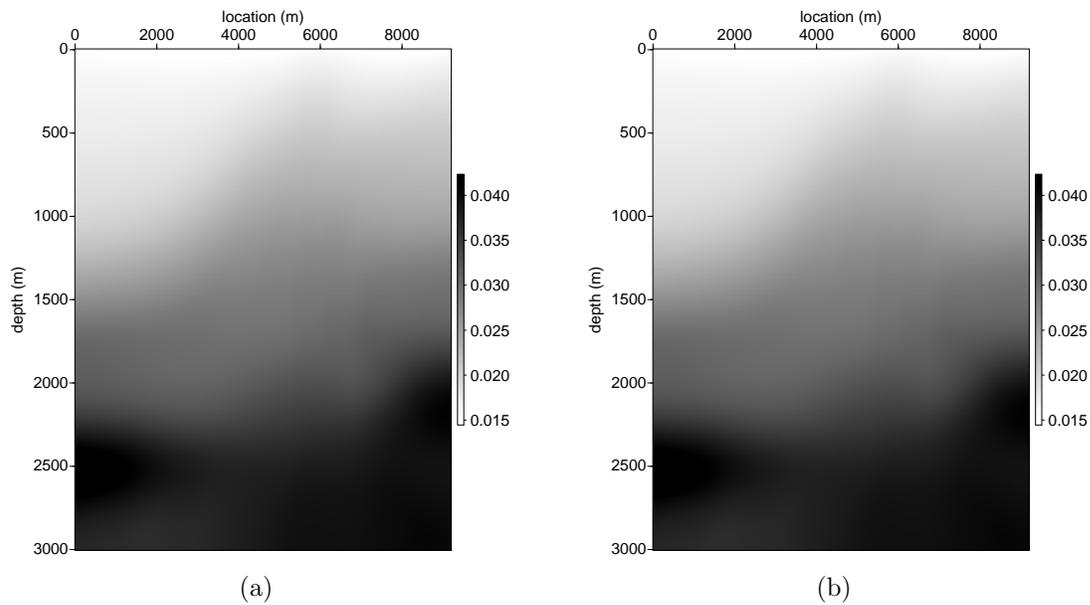


Figure C.4: (a) True ϵ model and (b) true δ model used to generate the synthetic data. [CR] `append3/. epssm,dltsm`

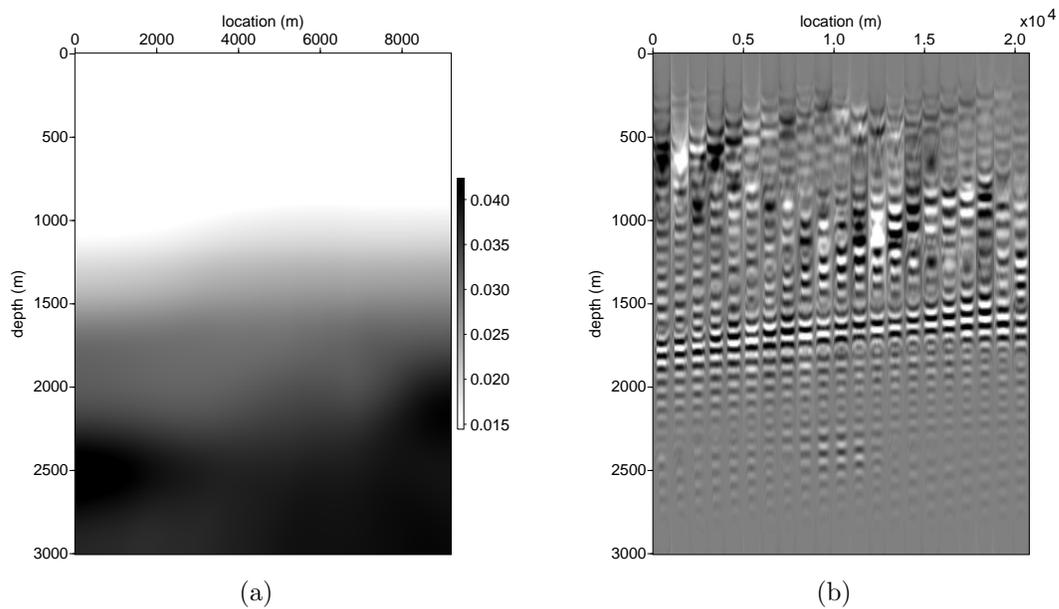


Figure C.5: (a) Initial ϵ model and (b) initial angle-domain common-image gathers. Gathers are taken at every 10th common image point from $x = 4$ km to $x = 8$ km. [CR] `append3/. initeps,init-e-image`

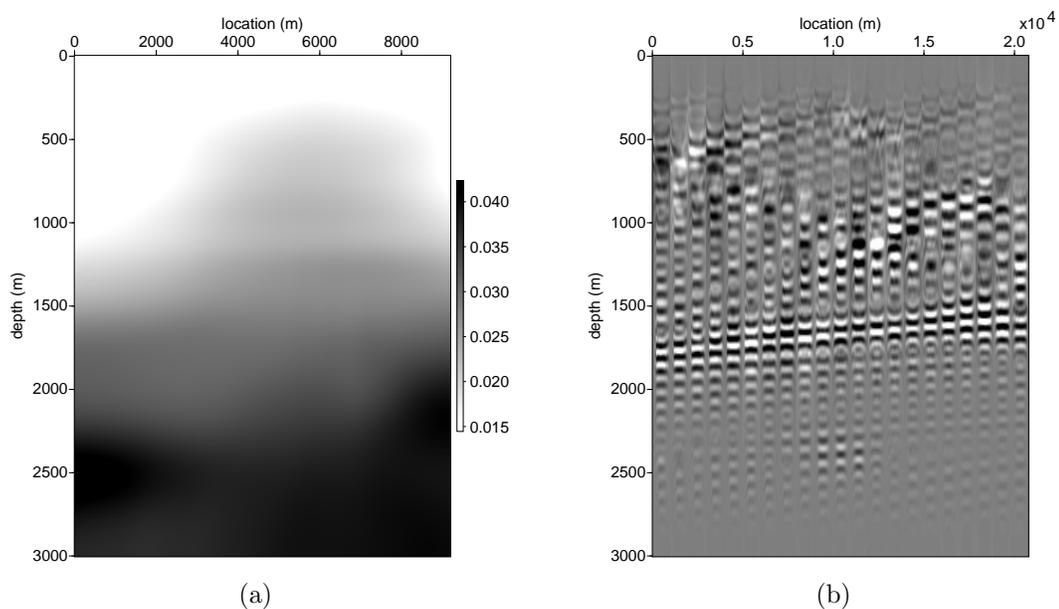


Figure C.6: (a) Inverted ϵ model and (b) final angle-domain common-image gathers using the inverted ϵ model in (a). Compared with Figure C.5(b), panel (b) shows more even energy across different angles. Gathers are taken at every 10th common image point from $x = 4$ km to $x = 8$ km. [CR] `append3/. inveps,inv-e-image`

Joint inversion for two parameters

The tests in the previous sections show that we have a reliable objective function and successful inversion results for a single parameter. However, joint inversion for more than one parameter for each grid in the subsurface is far more challenging because of the ambiguity between parameters. As a result, a preconditioning scheme using geological and rock-physics information is crucial for its success.

In this test, we use the same synthetic data as in the last section. Unlike in the last example where we use the perfect velocity model, the starting models for velocity and ϵ are both inaccurate. The initial velocity model and ϵ model are shown in Figures C.7(a) and C.5(a), respectively. The angle gathers generated using these initial models are shown in Figure C.7(b). Significant moveout in the angle-gather events indicates that the initial model is far from the true model. In fact, the initial velocity has a maximum of 15% error compared to the true velocity (Figure C.1(a)), while the initial ϵ is about 50% smaller than the true value in the shallow part of the model. The error in velocity has a large effect on the kinematics of the seismic wave, hence a larger effect on the flatness in the angle domain.

After 40 iterations, we obtain the inverted velocity and ϵ models shown in Figures C.8(a) and C.8(b). By comparing Figure C.8(a) with Figure C.1(a), we conclude that the inversion has successfully recovered the high-resolution vertical structure in the shallow part of the model. Due to the limited illumination, the steep structure in the deeper part of the model is not well resolved. Comparing Figure C.6(a) and Figure C.8(b), we notice that, because of the error in velocity, the inversion does not converge to the same solution. This is an indication that we have not completely resolved the ambiguity between velocity and ϵ .

Angle gathers generated by the inverted model are shown in Figure C.8(c). They are extracted from the same common-image points as in Figure C.7(b). The improved model flattens the gathers across the whole section. The low-frequency energy in the water is the commonly seen wave-path energy for RTM images.

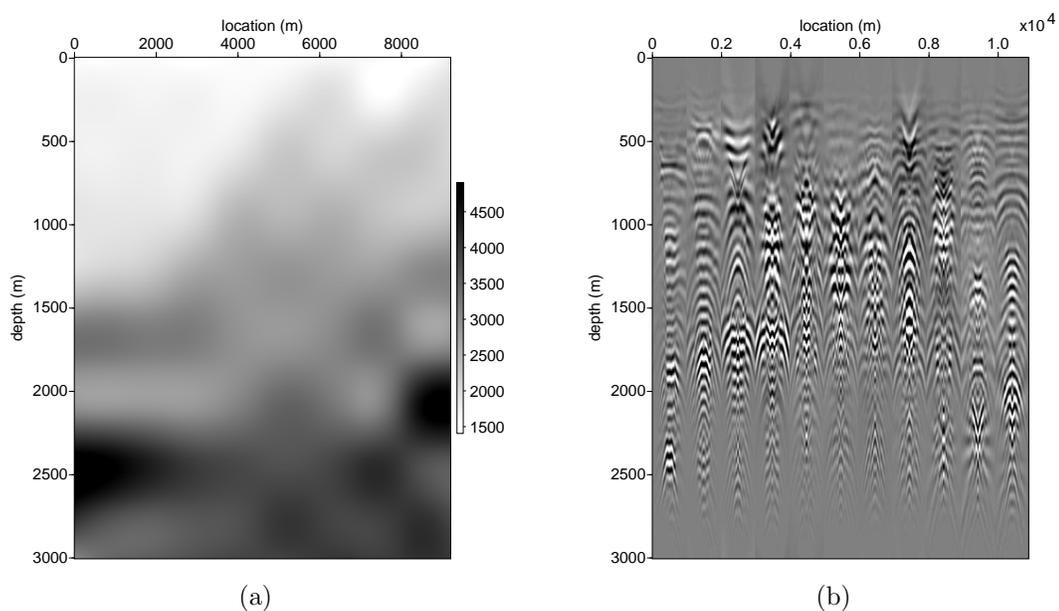


Figure C.7: Initial velocity model (a) and the angle gathers (b) obtained using the initial velocity model. The initial ϵ model is shown in Figure C.5(a). Model error causes significant curvatures in the angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [CR] `append3/. vini,angini`

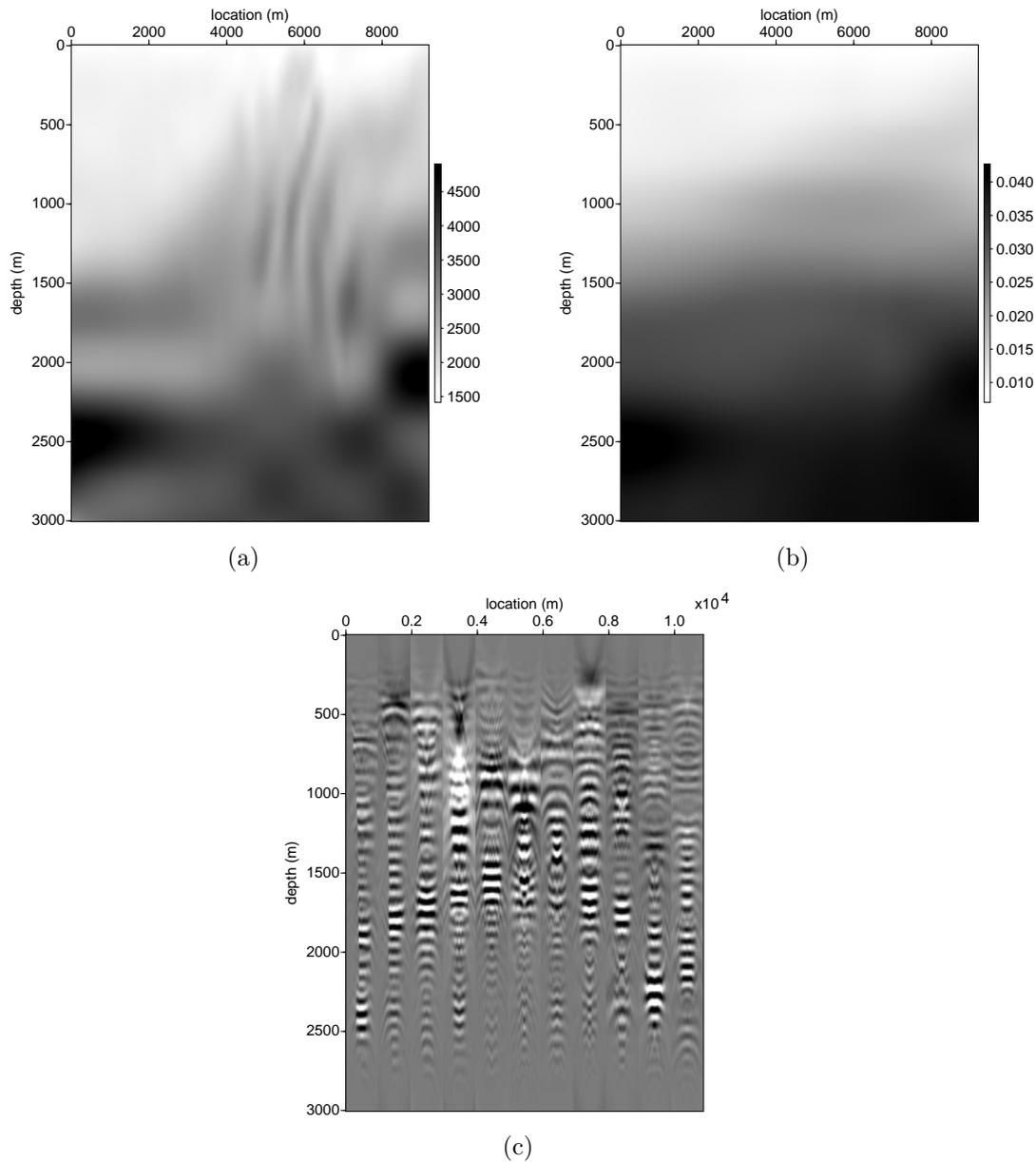


Figure C.8: Inverted velocity model (a) in m/s and ϵ model (b) after 40 iterations. Angle gathers (c) obtained by the inverted model. Angle gathers are extracted at the same CIP as those in Figure C.7(b). Improved velocity and ϵ flattens the corresponding angle gathers. Gathers are taken every 100 common image points from $x = 0$ km to $x = 9$ km. [CR] `append3/. vfinal,epsfinal,angfinal`

CONCLUSIONS

In this paper, we define an image-space inverse problem to solve for an optimized anisotropy model. To better describe the properties of the subsurface and the wave propagation, we use the first-order VTI two-way wave equation to compute our Green's function. Test results on the objective function show that flatness in the angle domain is a valid measure of both velocity and anisotropic parameter ϵ . The results of single-parameter and joint inversions demonstrate that we have successfully resolved the high-frequency structure in velocity and successfully improved the anisotropic model where we have enough angle coverage. The updated image has flattened the gathers in the angle domain.

Bibliography

- Alkhalifah, T. and I. Tsvankin, 1995, Velocity analysis for transversely isotropic media: *Geophysics*, **60**, 1550–1566.
- Bachrach, R., 2010a, Applications of deterministic and stochastic rock physics modeling to anisotropic velocity model building: *SEG Expanded Abstracts*, **29**, 2436–2440.
- , 2010b, Elastic and resistivity anisotropy of compacting shale: Joint effective medium modeling and field observations: *SEG Expanded Abstracts*, **29**, 2580–2584.
- Bachrach, R., Y. K. Liu, M. Woodward, O. Zradrova, Y. Yang, and K. Osypov, 2011, Anisotropic velocity model building using rock physics: Comparison of compaction trends and check-shot-derived anisotropy in the gulf of mexico: *SEG Expanded Abstract*, **30**, 207–211.
- Backus, G., 1962, Long-wave elastic anisotropy produced by horizontal layering: *Journal of Geophysical Research*, **76**, 4427–4440.
- Bakulin, A., M. Woodward, D. Nichols, K. Osypov, and O. Zdraveva, 2009, Can we distinguish TTI and VTI media?: *SEG Expanded Abstracts*, **28**, 226–230.
- , 2010a, Building tilted transversely isotropic depth models using localized anisotropic tomography with well information: *Geophysics*, **75**, 27–36.
- , 2010b, Localized anisotropic tomography with well information in VTI media: *Geophysics*, **75**, 37–45.
- Bandyopadhyay, K., 2009, Seismic anisotropy: geological causes and its implications: PhD thesis, Stanford University.
- Bear, L., T. Dickens, J. Krebs, J. Liu, and P. Traynin, 2005, Integrated velocity model estimation for improved positioning with anisotropic PSDM: *The Leading Edge*,

622–634.

- Berryman, J. G., 1980a, Long wavelength propagation in composite elastic media I: Ellipsoidal inclusions: *Journal of Acoustic Society of America*, **68**, 1820–1831.
- , 1980b, Long wavelength propagation in composite elastic media I: Spherical inclusions: *Journal of Acoustic Society of America*, **68**, 1809–1819.
- Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: *SEG Expanded Abstract*, **69**, 1723–1726.
- Box, M., D. Davies, W. H. Swann, and I. Australia, 1969, *Non-linear optimization techniques*: Oliver & Boyd Edinburgh.
- Cai, J., Y. He, Z. Li, B. Wang, and M. Guo, 2009, TTI/VTI anisotropy parameters estimation by focusing analysis, part i: theory: *SEG Expanded Abstracts*, **28**.
- Claerbout, J., 1987, *Imaging the earth's interior*: Blackwell Scientific Publications.
- Claerbout, J. F., 2009, *Image estimation by example*.
- Clapp, R., 2000, *Geologically constrained migration velocity analysis*: PhD thesis, Stanford University.
- Dræge, A., M. Jakobsen, and T. A. Johansen, 2006, Rock physics modeling of shale diagenesis: *Petroleum Geoscience*, **12**, 49–57.
- Duveneck, E., P. Milcik, P. M. Bakker, and C. Perkins, 2008, Acoustic VTI wave equations and their application for anisotropic reverse-time migration: *SEG Expanded Abstracts*, **27**.
- Fei, T. W. and C. L. Liner, 2008, Hybrid fourier finite-difference 3D depth migration for anisotropic media: *Geophysics*, **73**, S27.
- Fletcher, R., X. Du, and P. J. Fowler, 2009, Stabilizing acoustic reverse-time migration in TTI media: *SEG Expanded Abstracts*, **28**.
- Freed, R. L. and D. R. Peacor, 1989, Variability in temperature of the smectite/illite reaction in Gulf coast sediments: *Clay Minerals*, **24**, 171–180.
- Friedlander, M. P. and M. Schmidt, 2012, Hybrid deterministic-stochastic methods for data fitting: *SIAM Journal on Scientific Computing*, **34**, A1380–A1405.
- Gholami, Y., R. Brossier, S. Operto, A. Ribodetti, and J. Virieux, 2013, Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion? Part 1: Sensitivity and trade-off analysis: *Geophysics*, **78**, R81–R105.

- Guerra, C., Y. Tang, and B. Biondi, 2009, Wave-equation tomography using image-space phase-encoded data: SEP-report, **138**, 95.
- Helbig, K., 1956, Die ausbreitung elastischer Wellen in anisotropen Medien: Geophys. Prosp., **04**, 70–81.
- Hornby, B., D. Miller, C. Esmersoy, and P. Christie, 1995, Ultrasonic-to-seismic measurements of shale anisotropy in a North Sea well: SEG Expanded Abstracts, **14**, 17–21.
- Hornby, B., L. M. Schwartz, and J. A. Hudson, 1994, Anisotropic effective-medium modeling of the elastic properties of shales: Geophysics, **59**, no. 10, 1570–1583.
- Hower, J., E. Eslinger, M. Hower, and E. Perry, 1976, Mechanism of burial metamorphism of argillaceous sediments: 1. Mineralogical and chemical evidence: Bulletin of The Geological Society of America, **87**, 725–737.
- Kaufman, H., 1953, Velocity functions in seismic prospecting: Geophysics, **18**, 289–297.
- Li, Y., D. Nichols, K. Osypov, and R. Bachrach, 2011, Anisotropic tomography using rock physics constraints: 73rd EAGE Conference & Exhibition, Extended Abstracts.
- Marquering, H., F. A. Dahlen, and G. Nolet, 1998, Three-dimensional waveform sensitivity kernels: Geophysical Journal International, **132**, 521–534.
- , 1999, Three-dimensional sensitivity kernels for finite-frequency traveltimes: the banana-doughnut paradox: Geophysical Journal International, **137**, 805–815.
- McCollum, B. and F. Snell, 1932, Asymmetry of sound velocity in stratified formations: Physics (Journal of Applied Physics), **2**, 174–185.
- Osypov, K., D. Nichols, M. Woodward, O. Zdraveva, and C. E. Yarman, 2008, Uncertainty and resolution analysis for anisotropy tomography using iterative eigen-decomposition: SEG Expanded Abstracts, **27**, 3244–3249.
- Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.
- Plessix, R.-E. and H. Rynja, 2010, VTI full waveform inversion: a parameterization study with a narrow azimuth streamer data example: SEG Expanded Abstracts,

29.

- Polak, E., 1997, Optimization: Algorithms and consistent approximations: Springer-Verlag.
- Postma, G. W., 1955, Wave propagation in a stratified medium: *Geophysics*, **20**, 780–806.
- Rickett, J., 2000, Traveltime sensitivity kernels: Banana-doughnuts or just plain bananas?: *SEP-Report*, **103**, 63–71.
- Sarkar, D. and I. Tsvankin, 2003, Analysis of image gathers in factorized VTI media: *Geophysics*, **68**, 2016–2025.
- , 2004, Migration velocity analysis in factorized VTI media: *Geophysics*, **69**, 708–718.
- Sava, P. and B. Biondi, 2004a, Wave-equation migration velocity analysis-I: Theory: *Geophysical Prospecting*, **52**, 593–606.
- , 2004b, Wave-equation migration velocity analysis-II: Examples: *Geophysical Prospecting*, **52**, 607–623.
- Sava, P. and S. Formel, 2006, Generalized imaging conditions for wave equation migration: *CWP Report*, **524**.
- Sayers, C., 2004, Seismic anisotropy of shales: What determines the sign of Thomsen’s delta parameter?: *SEG Expanded Abstracts*, **23**, 103–106.
- , 2010, The effect of anisotropy on the Young’s moduli and Poisson’s ratios of shales: *SEG Expanded Abstracts*, **29**, 2606–2611.
- Shan, G., 2008, Imaging of steep reflectors in anisotropic media by wavefield extrapolation: PhD thesis, Stanford University.
- , 2009, Optimized implicit finite-difference and fourier finite-difference migration for VTI media: *Geophysics*, WCA189–WCA197.
- Shen, P., 2004, Wave-equation migration velocity analysis by differential semblance optimization: PhD thesis, Rice University.
- Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: *Geophysics*, **73**, VE49–VE59.
- Slotnick, M. M., 1936, On seismic computation with applications, I: *Geophysics*, **1**, 9–22.

- Soubaras, R. and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: *Geophysics*, **72**, U67–U73.
- Symes, W. W. and J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663.
- Takanashi, M. and I. Tsvankin, 2012, Migration velocity analysis for TI media in the presence of quadratic lateral velocity variation: *Geophysics*, **77**, U87–U96.
- Tang, Y. and B. Biondi, 2011, Target-oriented wavefield tomography using synthesized Born data: *Geophysics*, **97**, WB191–WB207.
- Tang, Y., C. Guerra, and B. Biondi, 2008, Image-space wave-equation tomography in the generalized source domain: SEP-report, **136**, 1–22.
- Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259–1266.
- Thomsen, L., 1986, Weak elastic anisotropy: *Geophysics*, **51**, 1954–1966.
- Tsvankin, I. and L. Thomsen, 1994, Nonhyperbolic reflection moveout in anisotropic media: *Geophysics*, **59**, 1290–1304.
- Vernik, L. and X. Liu, 1997, Velocity anisotropy of shales: A petrophysical study: *Geophysics*, **62**, 521–532.
- Wenk, H.-R., I. Lnardelli, H. Franz, K. Nihei, and S. Nakagawa, 2007, Preferred orientation and elastic anisotropy of illite-rich shale: *Geophysics*, **72**, E69–E75.
- Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismogram: PhD thesis, University of Tulsa.
- Woodward, M., Y. Yang, K. Osypov, R. Bachrach, D. Nichols, O. Zdraveva, Y. Liu, and A. Fournier, 2014, Geological and rock physics constraints in anisotropic tomography: 76th EAGE Conference & Exhibition, WS12–B01.
- Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, **57**, 15–26.
- Woodward, M. J., D. Nichols, O. Zdraveva, P. Whitfield, and T. Johns, 2008, A decade of tomography: *Geophysics*, **73**, VE5VE11.
- Wright, S. and J. Nocedal, 1999, Numerical optimization, volume **2**: Springer New York.
- Yang, Y., K. Osypov, R. Bachrach, M. Woodward, O. Zdraveva, Y. Liu, A. Fournier, and Y. You, 2012, Anisotropic tomography and uncertainty analysis with rock

- physics constraints: Green Canyon case study: SEG Expanded Abstract, 1–5.
- Yuan, J., X. Ma, S. Lin, and D. Lowrey, 2006, P-wave tomographic velocity updating in 3d inhomogeneous VTI media: SEG Expanded Abstracts, **25**.
- Zdraveva, O., M. Woodward, D. Nichols, and K. Osypov, 2014, Importance of cross-disciplinary constraints in anisotropic model building and updating: 76th EAGE Conference & Exhibition, WS12–B03.
- Zhang, Y. and H. Zhang, 2009, A stable TTI reverse time migration and its implementation: SEG Expanded Abstracts, **28**.
- Zhou, H., D. Pham, and S. Gray, 2004, Tomographic velocity analysis in strongly anisotropic TTI media: SEG Expanded Abstracts, **23**.
- Zhou, H., D. Pham, S. Gray, and B. Wang, 2003, 3-d tomographic velocity analysis in transversely isotropic media: SEG Expanded Abstracts, **22**.