

CWP-1028

FULL-WAVEFORM INVERSION OF TIME-LAPSE SEISMIC DATA USING PHYSICS-BASED AND  
DATA-DRIVEN TECHNIQUES

by  
Yanhua Liu

© Copyright by Yanhua Liu, 2023

All Rights Reserved

A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Geophysics).

Golden, Colorado

Date \_\_\_\_\_

Signed: \_\_\_\_\_

Yanhua Liu

Signed: \_\_\_\_\_

Dr. Ilya Tsvankin  
Thesis Advisor

Golden, Colorado

Date \_\_\_\_\_

Signed: \_\_\_\_\_

Dr. Paul Sava  
Professor and Department Head  
Department of Geophysics

## ABSTRACT

Time-lapse (4D) full-waveform inversion (FWI) is an advanced seismic technique that can enable accurate estimation of changes in the subsurface properties, such as fluid saturation or reservoir depletion, by utilizing amplitude and phase information in seismic data. However, most existing 4D FWI research is limited to isotropic and, often, acoustic media, which hinders its application to realistic subsurface models. In this thesis, I develop an efficient 4D FWI algorithm for elastic transversely isotropic media with a vertical (VTI) and tilted (TTI) symmetry axis. In addition, I employ a “source-independent” technique to mitigate the influence of errors in the source wavelet on the results of time-lapse FWI. Furthermore, the thesis presents machine-learning techniques with uncertainty quantification to efficiently perform real-time monitoring with high spatial resolution.

First, I extend the methodology of time-lapse FWI to elastic VTI media. The algorithm is tested on multicomponent and pressure data using three common time-lapse strategies: the parallel-difference (PD), sequential-difference (SD), and double-difference (DD) techniques. The multiscale approach is adopted to mitigate cycle-skipping. Synthetic tests show that the proposed methodology can reconstruct localized time-lapse parameter variations with sufficient spatial resolution. The DD strategy produces the most accurate results for clean and repeatable time-lapse data because it directly inverts the data difference for the parameter changes.

VTI algorithms become inadequate in the presence of an even moderate symmetry-axis tilt. Therefore, next the time-lapse FWI methodology is extended to 2D tilted TI models. The symmetry-axis tilt is incorporated into the modeling code and computation of the inversion gradients by rotating the stiffness tensor using the Bond transformation. Comparison between the TTI and VTI algorithms confirms that incorporating tilt improves the accuracy of the inverted medium parameters, especially when the reservoir is located in a dipping layer. In addition, I discuss the influence of several common nonrepeatability issues on the time-lapse inversion results for TTI media.

FWI requires an accurate estimate of the source wavelet, which is both time consuming and often challenging for field-data applications. The so-called “source-independent” (SI) technique is designed to reduce the influence of the employed source wavelet on the FWI results. Therefore, I incorporate the convolution-based SI technique into the developed 4D FWI algorithm for both VTI and TTI models. The SI method substantially reduces the dependence of the estimated parameters on the accuracy of the source wavelet and guides the inversion toward the global minimum of the objective function even for a strongly distorted wavelet and noisy data.

Then the developed 4D FWI methodology that incorporates the SI technique is applied to the time-lapse streamer data from Pyrenees oil/gas field in offshore Australia. Despite the pronounced nonrepeatability in the baseline and monitor surveys, the developed algorithm successfully reconstructs the velocity variations in the reservoir caused by hydrocarbon production. The case study demonstrates the importance of accounting for anisotropy and elasticity in time-lapse inversion and confirms the effectiveness of the SI technique when the source wavelet is distorted.

To alleviate the ill-posedness and high computational cost of FWI, I propose an efficient “hybrid” time-lapse workflow that combines physics-based FWI and data-driven machine-learning (ML) inversion. The scarcity of the available training data is addressed by developing a new data-generation technique that operates with physics constraints. The proposed approach is validated on a synthetic CO<sub>2</sub>-sequestration model based on the Kimberlina reservoir in California. A large volume of high-quality and physically realistic training data, generated by the algorithm, proves to be critically and efficiently important in accurately characterizing the CO<sub>2</sub> movement in the reservoir.

The deterministic neural network described above, however, yields only one prediction for a certain input, which may not properly reflect the distribution of the entire testing data, especially when those data are out-of-distribution. To estimate the entire distribution of the target variable along with the prediction accuracy, I incorporate the Simultaneous Quantile Regression method into the developed convolutional neural network. Testing on the Kimberlina data demonstrates the accuracy of the obtained uncertainty estimates, even if the testing data are distorted due to problems in the field-data acquisition. In addition, the proposed novel data-augmentation method can further improve the spatial resolution of the determined time-lapse velocity field and reduce the prediction error.

## TABLE OF CONTENTS

ABSTRACT . . . . .	iii
LIST OF FIGURES . . . . .	x
LIST OF TABLES . . . . .	xviii
ACKNOWLEDGMENTS . . . . .	xix
CHAPTER 1 INTRODUCTION . . . . .	1
1.1 Thesis Contributions . . . . .	3
CHAPTER 2 METHODOLOGY OF TIME-LAPSE ELASTIC FULL-WAVEFORM INVERSION FOR VTI MEDIA . . . . .	5
2.1 Introduction . . . . .	6
2.2 Methodology of time-lapse FWI for VTI media . . . . .	6
2.2.1 Parallel-difference method . . . . .	7
2.3 Sequential-difference method . . . . .	7
2.3.1 Double-difference method . . . . .	8
2.4 Synthetic Examples . . . . .	8
2.4.1 Multicomponent data . . . . .	10
2.4.2 Pressure data . . . . .	12
2.4.3 Noisy data . . . . .	14
2.5 Conclusions . . . . .	17
2.6 Acknowledgments . . . . .	17
CHAPTER 3 SOURCE-INDEPENDENT TIME-LAPSE FULL-WAVEFORM INVERSION FOR VTI MEDIA . . . . .	18
3.1 Introduction . . . . .	18
3.2 Methodology of source-independent time-lapse FWI . . . . .	20
3.2.1 Source-independent methodology . . . . .	21
3.2.2 Implementation of time-lapse FWI . . . . .	22
3.3 Synthetic Examples . . . . .	22

3.3.1	VTI graben model . . . . .	23
3.3.1.1	Parallel- and sequential-difference methods . . . . .	25
3.3.1.2	Double-difference method . . . . .	25
3.3.1.3	Influence of noise . . . . .	28
3.3.1.4	Influence of the reference trace and time window . . . . .	28
3.3.2	VTI Marmousi model . . . . .	29
3.4	Conclusions . . . . .	34
3.5	Acknowledgments . . . . .	34
CHAPTER 4 TIME-LAPSE FULL-WAVEFORM INVERSION FOR ELASTIC TTI MEDIA . . . . .		35
4.1	Introduction . . . . .	35
4.2	Methodology of source-independent time-lapse FWI for TTI media . . . . .	37
4.2.1	FWI for anisotropic media . . . . .	37
4.2.2	Strategies of time-lapse FWI . . . . .	38
4.3	Synthetic Example . . . . .	38
4.3.1	2D BP TTI model . . . . .	38
4.3.1.1	Influence of tilt . . . . .	41
4.3.1.2	Influence of nonrepeatable noise . . . . .	44
4.3.1.3	Influence of source wavelet . . . . .	44
4.3.1.4	Influence of geometry nonrepeatability . . . . .	46
4.3.1.5	Influence of water statics . . . . .	48
4.3.1.6	Influence of overburden changes . . . . .	48
4.4	Conclusions . . . . .	49
4.5	Acknowledgments . . . . .	51
CHAPTER 5 ELASTIC TIME-LAPSE FWI FOR ANISOTROPIC MEDIA: A PYRENEES CASE STUDY . . . . .		52
5.1	Introduction . . . . .	52
5.2	Methodology of source-independent full-waveform inversion . . . . .	53
5.2.1	FWI objective functions . . . . .	53

5.2.2	Source-independent FWI . . . . .	54
5.2.3	Forward modeling and model updating for FWI . . . . .	54
5.3	Field-data application . . . . .	55
5.3.1	Location and geology . . . . .	55
5.3.2	Preprocessing . . . . .	55
5.3.3	Initial model . . . . .	57
5.3.4	Wavelet estimation . . . . .	58
5.3.5	Inversion of baseline data . . . . .	60
5.3.5.1	Elastic isotropic FWI . . . . .	60
5.3.5.2	Influence of wavelet distortion . . . . .	62
5.3.5.3	Elastic FWI for VTI media . . . . .	63
5.3.6	Inversion of monitor data . . . . .	64
5.4	Conclusions . . . . .	65
5.5	Acknowledgments . . . . .	65
CHAPTER 6 JOINT PHYSICS-BASED AND DATA-DRIVEN TIME-LAPSE SEISMIC INVERSION: MITIGATING DATA SCARCITY . . . . .		75
6.1	Introduction . . . . .	75
6.2	Review of physics-based and data-driven inversions . . . . .	77
6.2.1	Physics-based full-waveform inversion . . . . .	77
6.2.2	Data-driven InvNet-VelSat . . . . .	78
6.3	Methodology of time-lapse processing . . . . .	79
6.3.1	Strategy 1: Physics-based time-lapse FWI . . . . .	79
6.3.2	Strategy 2: Data-driven time-lapse inversion . . . . .	79
6.3.3	Strategy 3: Hybrid approach . . . . .	80
6.3.3.1	Velocity-generation method . . . . .	81
6.3.3.2	Synthetic CO <sub>2</sub> -saturation models . . . . .	81
6.4	Synthetic Examples . . . . .	82
6.4.1	Kimberlina data set . . . . .	82

6.4.2	Physics-based time-lapse FWI . . . . .	84
6.4.3	Data-driven time-lapse inversion . . . . .	85
6.4.4	Hybrid strategy . . . . .	86
6.4.4.1	Hybrid strategy vs. physics-based method . . . . .	89
6.4.4.2	Hybrid strategy vs. data-driven strategy . . . . .	89
6.4.5	Generalization and robustness: Hybrid strategy vs. data-driven strategy . . . . .	91
6.4.5.1	Additional monitoring tests . . . . .	94
6.4.5.2	Influence of noise . . . . .	94
6.5	Conclusions . . . . .	94
6.6	Acknowledgments . . . . .	96
6.7	Data and materials availability . . . . .	96
CHAPTER 7 ENHANCED PREDICTION ACCURACY WITH UNCERTAINTY		
QUANTIFICATION IN MONITORING CO <sub>2</sub> SEQUESTRATION USING		
CONVOLUTIONAL NEURAL NETWORKS . . . . .		
7.1	Introduction . . . . .	97
7.2	Method . . . . .	99
7.2.1	Simultaneous Quantile Regression . . . . .	100
7.2.2	Network architecture of InvNet_UQ . . . . .	101
7.2.3	Uncertainty-guided data augmentation . . . . .	102
7.3	Results . . . . .	103
7.3.1	Kimberlina data set . . . . .	103
7.3.2	Implementation details . . . . .	103
7.3.3	Noise-free data . . . . .	104
7.3.4	Influence of random noise . . . . .	105
7.3.5	Influence of missing traces . . . . .	106
7.3.6	Influence of absent low-frequency data . . . . .	108
7.3.7	Test of uncertainty-guided data augmentation . . . . .	108
7.4	Conclusions . . . . .	109

7.5	Data and Codes Availability . . . . .	111
7.6	Acknowledgments . . . . .	112
7.7	Author contributions statement . . . . .	112
7.8	Additional information . . . . .	112
CHAPTER 8 CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK . . . . .		115
8.1	Summary . . . . .	115
8.2	Recommendations for future work . . . . .	117
REFERENCES . . . . .		119
APPENDIX A INVERSION GRADIENTS FOR THE SOURCE-INDEPENDENT FWI . . . . .		128
APPENDIX B WAVE PROPAGATION IN 2D TILTED TI MEDIA . . . . .		133
APPENDIX C PERMISSIONS FOR CHAPTERS . . . . .		135

## LIST OF FIGURES

Figure 2.1	Parameters of the baseline model with a grid size of $10 \times 10$ m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (d) the S-wave vertical velocity ( $V_{S0}$ ), (g) the P-wave horizontal velocity ( $V_{hor,P}$ ), (j) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ), and (m) the density ( $\rho$ ). The initial baseline model of: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The actual time-lapse differences for (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	9
Figure 2.2	Normalized parameter errors for the final baseline model reconstructed from the multicomponent data: (a) $V_{P0}$ , (b) $V_{S0}$ , (c) $V_{hor,P}$ , (d) $V_{nmo,P}$ and (e) $\rho$ . . . . .	10
Figure 2.3	Normalized errors of the time-lapse parameter variations obtained from the multicomponent data using three different approaches. Only the parameters $V_{P0}$ , $V_{S0}$ and $\rho$ are changed. The parallel-difference method: (a) $V_{P0}$ , (d) $V_{S0}$ , (g) $V_{hor,P}$ , (j) $V_{nmo,P}$ , and (m) $\rho$ . The sequential-difference method: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The double-difference method: (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	11
Figure 2.4	Normalized errors of the time-lapse parameter variations obtained from the multicomponent data. Only the parameter $V_{S0}$ is changed. The parallel-difference method: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $V_{nmo,P}$ . The sequential-difference method: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $V_{nmo,P}$ . The double-difference method: (g) $V_{P0}$ , (h) $V_{S0}$ , and (i) $V_{nmo,P}$ . . . . .	12
Figure 2.5	Baseline models reconstructed from the pressure data: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $V_{nmo,P}$ . Normalized errors of the time-lapse parameter variations reconstructed from the pressure data. Only the parameters $V_{P0}$ , $V_{S0}$ , and $\rho$ are changed. The parallel-difference method: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $V_{nmo,P}$ . The sequential-difference method: (g) $V_{P0}$ , (h) $V_{S0}$ , and (i) $V_{nmo,P}$ . The double-difference method: (j) $V_{P0}$ , (k) $V_{S0}$ , and (l) $V_{nmo,P}$ . . . . .	13
Figure 2.6	Parameters of the time-lapse model produced by three different strategies from the multicomponent data (SNR = 16). The parallel-difference method: (a) $V_{P0}$ , (d) $V_{S0}$ , (g) $V_{hor,P}$ , (j) $V_{nmo,P}$ , and (m) $\rho$ . The sequential-difference method: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The double-difference method: (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	15
Figure 2.7	Parameters of the time-lapse model produced by three different strategies from the noisy multicomponent data (SNR = 8). The parallel-difference method: (a) $V_{P0}$ , (d) $V_{S0}$ , (g) $V_{hor,P}$ , (j) $V_{nmo,P}$ , and (m) $\rho$ . The sequential-difference method: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The double-difference method: (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	16
Figure 3.1	Source wavelets used in the synthetic examples: (a) the Ricker wavelet with a central frequency of 10 Hz (actual wavelet), (b) the distorted ‘‘Ricker’’ wavelet with a central frequency of 17 Hz (Wavelet 1), and (c) the spike (Wavelet 2). The frequency spectra of (d) the actual wavelet [see plot (a)] and (e) Wavelet 1 [see plot (b)]. . . . .	23
Figure 3.2	Parameters of the baseline graben model with a grid size of $10 \times 10$ m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (d) the S-wave vertical velocity ( $V_{S0}$ ), (g) the P-wave horizontal velocity ( $V_{hor,P}$ ), (j) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ), and (m) the density ( $\rho$ ). The actual time-lapse differences for (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The initial baseline model of: (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	24

Figure 3.3	Time-lapse parameter variations obtained by conventional FWI with the actual wavelet using the parallel-difference method (benchmark results): (a) $V_{P0}$ , (b) $V_{S0}$ , (c) $V_{hor,P}$ , (d) $V_{nmo,P}$ , and (e) $\rho$ . . . . .	24
Figure 3.4	Baseline models estimated by conventional FWI using Wavelet 1: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . . . . .	25
Figure 3.5	Time-lapse parameter variations reconstructed by the source-independent algorithm with Wavelet 1. The parallel-difference method: (a) $V_{P0}$ , (d) $V_{S0}$ , (g) $V_{hor,P}$ , (j) $V_{nmo,P}$ , and (m) $\rho$ . The sequential-difference method: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The double-difference method: (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	26
Figure 3.6	Time-lapse parameter variations estimated by the source-independent algorithm with the spike wavelet (Wavelet 2) using the parallel-difference method: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . . . . .	26
Figure 3.7	Time-lapse variations of the velocity $V_{P0}$ reconstructed by the source-independent algorithm using the double-difference method. FWI is performed with Wavelet 1, whereas the data for the inverted baseline model ( $\mathbf{d}_b^{sim}$ ) are generated with: (a) Wavelet 1, (b) a signal that has the shape of Wavelet 1 but correct frequency, (c) the actual wavelet. . . . .	27
Figure 3.8	Inverted baseline models and time-lapse variations obtained from noisy data (signal-to-noise ratio is 16). Inverted baseline models obtained by conventional FWI with the actual wavelet [(a) $V_{P0}$ , (b) $V_{S0}$ , (c) $\rho$ ] and by the source-independent FWI with Wavelet 1 [(d) $V_{P0}$ , (e) $V_{S0}$ , (f) $\rho$ ]. Time-lapse variations obtained by conventional FWI with the actual wavelet [(g) $V_{P0}$ , (h) $V_{S0}$ , (i) $\rho$ ] and by the source-independent FWI with Wavelet 1 [(j) $V_{P0}$ , (k) $V_{S0}$ , (l) $\rho$ ]. . . . .	29
Figure 3.9	Velocity $V_{P0}$ estimated by the source-independent algorithm from the baseline data for the graben model using the spike wavelet (Wavelet 2) and two time windows: (a) window 1 [plot(c)] and (b) window 2 [plot(d)]. . . . .	30
Figure 3.10	Baseline parameters of the modified VTI Marmousi model with a grid size of $10 \times 10$ m: (a) $V_{P0}$ , (d) $V_{S0}$ , (g) $V_{hor,P}$ , (j) $V_{nmo,P}$ , and (m) $\rho$ . The initial baseline model of: (b) $V_{P0}$ , (e) $V_{S0}$ , (h) $V_{hor,P}$ , (k) $V_{nmo,P}$ , and (n) $\rho$ . The actual time-lapse differences for (c) $V_{P0}$ , (f) $V_{S0}$ , (i) $V_{hor,P}$ , (l) $V_{nmo,P}$ , and (o) $\rho$ . . . . .	31
Figure 3.11	Time-lapse variations for the Marmousi model obtained by conventional FWI using the parallel-difference method with the actual wavelet (benchmark results): (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . . . . .	32
Figure 3.12	Baseline Marmousi models produced by conventional FWI using a wavelet with only a shape distortion: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . . . . .	32
Figure 3.13	Time-lapse parameter variations for the Marmousi model obtained by the source-independent algorithm with Wavelet 1. The parallel-difference method: (a) $V_{P0}$ , (c) $V_{S0}$ , and (e) $\rho$ . The sequential-difference method: (b) $V_{P0}$ , (d) $V_{S0}$ , and (f) $\rho$ . . . . .	33

Figure 4.1	Parameters of the baseline BP TTI model with a grid size of $10 \times 10$ m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (b) the S-wave vertical velocity ( $V_{S0}$ ), (c) the density ( $\rho$ ), (d) the P-wave velocity in the isotropy plane ( $V_{hor,P}$ ), and (e) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ). The initial baseline model of: (f) $V_{P0}$ , (g) $V_{S0}$ , (h) $\rho$ , (i) $V_{hor,P}$ , and (j) $V_{nmo,P}$ . The actual time-lapse differences in (k) $V_{P0}$ , (l) $V_{S0}$ , and (m) $\rho$ . Note that there are no time-lapse changes in $V_{hor,P}$ and $V_{nmo,P}$ . . . . .	40
Figure 4.2	Tilt of the symmetry axis estimated by migrating the baseline data using the actual baseline model. . . . .	41
Figure 4.3	Time-lapse parameter variations obtained from noise-free multicomponent data. The parallel-difference method: (a) $V_{P0}$ , (b) $V_{S0}$ , (c) $\rho$ , (d) $V_{hor,P}$ , and (e) $V_{nmo,P}$ . The sequential-difference method: (f) $V_{P0}$ , (g) $V_{S0}$ , (h) $\rho$ , (i) $V_{hor,P}$ , and (j) $V_{nmo,P}$ . The double-difference method: (k) $V_{P0}$ , (l) $V_{S0}$ , (m) $\rho$ , (n) $V_{hor,P}$ , and (o) $V_{nmo,P}$ . . . . .	42
Figure 4.4	Shot gathers of the vertical particle velocity for: (a) the actual TTI medium, and (b) the corresponding VTI medium. Plot (c) shows the difference between plots (a) and (b). . . . .	42
Figure 4.5	Time-lapse changes inside the reservoir located in a dipping segment of the anticline. The actual time-lapse changes: (a) $V_{P0}$ , (b) $V_{S0}$ , (c) $\rho$ . The time-lapse variations estimated by the VTI [(d) $V_{P0}$ , (e) $V_{S0}$ , (f) $\rho$ , (g) $V_{hor,P}$ , and (h) $V_{nmo,P}$ ] and TTI [(i) $V_{P0}$ , (j) $V_{S0}$ , (k) $\rho$ , (l) $V_{hor,P}$ , and (m) $V_{nmo,P}$ ] algorithms. . . . .	43
Figure 4.6	Time-lapse parameter variations obtained from noisy baseline and monitor data (the signal-to-noise ratio is 15) for the model in Figure 4.1. The parallel-difference strategy: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . The sequential-difference strategy: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ . The double-difference strategy: (h) $V_{P0}$ , (i) $V_{S0}$ , and (j) $\rho$ . . . . .	44
Figure 4.7	Source wavelets used in the synthetic examples: (a) the Ricker wavelet with a central frequency of 10 Hz (actual wavelet) and (b) the distorted “Ricker” wavelet with a central frequency of 17 Hz (Wavelet 1). The frequency spectra of: (c) the actual wavelet [see plot (a)] and (d) Wavelet 1 [see plot (b)]. . . . .	45
Figure 4.8	Baseline models estimated using Wavelet 1 by the conventional FWI [(a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ ] and the source-independent FWI [(d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ ]. . . . .	45
Figure 4.9	Time-lapse parameter variations reconstructed by the source-independent algorithm using Wavelet 1. The parallel-difference method: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . The sequential-difference method: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ . . . . .	46
Figure 4.10	Difference between the vertical particle-velocity components from the monitor and baseline surveys: (a) the perfectly repeatable data (referred to as the actual difference) and the data with (b) geometry nonrepeatability, (c) water statics, and (d) overburden changes. The actual differences caused by the reservoir changes are circled in blue. The noise caused by nonrepeatability is marked by the red arrows. . . . .	47
Figure 4.11	Time-lapse parameter variations of the velocity $V_{P0}$ estimated from the data with geometry nonrepeatability. The results obtained with the actual acquisition geometries: (a) the parallel-difference method, (b) the sequential-difference method, and (c) the double-difference method. The results obtained using the baseline geometry in the inversion of both the baseline and monitor data: (d) the parallel-difference method, (e) the sequential-difference method, and (f) the double-difference method. . . . .	47

Figure 4.12	Time-lapse parameter variations estimated from the data with “water statics.” The parallel-difference method: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ . The sequential-difference method: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ . The double-difference method: (g) $V_{P0}$ , (h) $V_{S0}$ , and (i) $\rho$ . . . . .	49
Figure 4.13	Time-lapse parameter variations estimated from the data with overburden changes. The actual time-lapse variations in (a) $V_{P0}$ and (b) $V_{S0}$ . The parallel-difference strategy: (c) $V_{P0}$ and (d) $V_{S0}$ . The sequential-difference strategy: (e) $V_{P0}$ and (f) $V_{S0}$ . The double-difference strategy: (g) $V_{P0}$ and (h) $V_{S0}$ . . . . .	50
Figure 5.1	Location of the Pyrenees 4D survey (after Cai et al., 2019). . . . .	56
Figure 5.2	Geometries of the time-lapse seismic surveys. The orange and yellow lines mark the 2D baseline and monitor surveys, respectively. . . . .	57
Figure 5.3	Seismic shot gather after low-bandpass filtering with a high cut-off frequency of (a) 2Hz, (b) 3Hz, (c) 4Hz, and (d) 5Hz. The white dashed circle shows the seismic signal. . . . .	58
Figure 5.4	Parameters of the Inpex VTI tomography model with a grid size of $12.5 \times 12.5$ m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (b) the S-wave vertical velocity ( $V_{S0}$ ), and (c) the density ( $\rho$ ). The smoothed initial model for FWI: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ . . . . .	59
Figure 5.5	(a) Observed and simulated direct wave. (b) The Wiener filter obtained by mapping the simulated data to observed data. . . . .	60
Figure 5.6	(a) Initial guess wavelet, (b) final wavelet after applying the Wiener filter (used for FWI), and (c) the difference. . . . .	61
Figure 5.7	Medium parameters estimated by isotropic FWI using different objective functions. The $L_2$ -norm objective function: (a) $V_{P0}$ , (b) $V_{S0}$ , and (c) $\rho$ ; the source-independent (SI) objective function: (d) $V_{P0}$ , (e) $V_{S0}$ , and (f) $\rho$ ; the global-correlation (GC) objective function: (h) $V_{P0}$ , (i) $V_{S0}$ , and (j) $\rho$ . The arrows point to the low-velocity layer and the ellipses mark high-frequency artifacts. . . . .	62
Figure 5.8	Vertical profiles of the P-wave vertical velocity $V_{P0}$ . The velocity obtained from the well logs is marked by the blue line and the initial velocity used in FWI by the red line. The inversion results are marked by the yellow ( $L_2$ -norm objective function), violet (SI objective function), and green (GC objective function) lines. . . . .	66
Figure 5.9	Quality of data matching for shot 144 at $x \approx 1$ km. The recorded traces (“field data”) are interleaved between the traces simulated using (a) the initial model and (b, c, d) the models obtained by the isotropic FWI applied with the (b) $L_2$ -norm objective function, (c) source-independent objective function, and (d) global-correlation objective function. The arrows point to the areas where the GC objective function produces a better match with the field data. . . . .	67
Figure 5.10	P-wave velocity ( $V_{P0}$ ) estimated by isotropic FWI with a distorted wavelet using different objective functions: (a) the $L_2$ -norm objective function, (b) the global-correlation (GC) objective function, and (c) the source-independent (SI) objective function. . . . .	68
Figure 5.11	Medium parameters estimated by the VTI FWI algorithm with the global-correlation objective function: (a) $V_{P0}$ , (b) $V_{S0}$ , (c) $V_{hor,P}$ , (d) $V_{nmo,P}$ , and (e) $\rho$ . . . . .	69

Figure 5.12	Thomsen parameters obtained from the inverted velocities (Figure 5.11) using equations 5.7 and 5.8: (a) $\epsilon$ and (b) $\delta$ . . . . .	69
Figure 5.13	Vertical profiles of the P-wave vertical velocity $V_{P0}$ . The velocity obtained from the sonic log is marked by the blue line and the initial velocity used in FWI by the red line. The inversion results are marked by the yellow (VTI FWI) and violet (isotropic FWI) lines. . . . .	70
Figure 5.14	Quality of data matching for shot 144 at $x \approx 1$ km. The recorded traces are interleaved between the traces simulated using the models obtained by the (a) isotropic FWI and (b) VTI FWI, both applied with the GC objective function. The arrows point to the areas where the VTI algorithm produces a better match with the field data. . . . .	71
Figure 5.15	Medium parameters estimated by the VTI FWI algorithm with the highest frequency equal to 19 Hz: (a) $V_{P0}$ , (b) $\epsilon$ , and (c) $\delta$ . Plot (d) shows the vertical profiles of the P-wave vertical velocity $V_{P0}$ . The velocity obtained from the sonic log is marked by the blue line. The inversion results obtained using the highest frequencies equal to 12 Hz and 19 Hz are marked by the yellow and orange lines, respectively. . . . .	72
Figure 5.16	Quality of data matching for shot 144 at $x \approx 1$ km. The recorded traces are interleaved between the traces simulated using the models obtained by the VTI FWI with the GC objective function using the highest frequencies equal to (a) 12 Hz and (b) 19 Hz. . . . .	73
Figure 5.17	Seismograms for shot 144 after applying Ormsby filtering (3-12 Hz) to the (a) baseline and (b) monitor data. . . . .	73
Figure 5.18	Time-lapse parameter variations obtained by the VTI FWI algorithm using a multiscale approach. The results for the frequency bands 3-5 Hz [(a) $V_{P0}$ , (b) $\epsilon$ , and (c) $\delta$ ], 3-9 Hz [(d) $V_{P0}$ , (e) $\epsilon$ , and (f) $\delta$ ], and 3-12 Hz [(h) $V_{P0}$ , (i) $\epsilon$ , and (j) $\delta$ ]. The arrows point to the “sandwich”-shape 4D response. . . . .	74
Figure 6.1	Network architecture of InvNet-VelSat. The encoder, built with five convolution layers and maximum pooling, extracts high-level features from the input seismic data to the latent space and reduces the dimensions of the feature map. Each of the two decoders consists of five deconvolution layers and translates the latent variable to the velocity (output 1) and saturation (output 2) models. . . . .	79
Figure 6.2	Workflow of the hybrid time-lapse strategy that combines physics-based FWI and data-driven InvNet-VelSat. FWI is employed to invert the baseline data and obtain the background velocity distribution for generating synthetic training samples. The trained InvNet-VelSat is applied to the monitor survey to predict both the velocity and saturation models. The temporal velocity changes are computed by subtracting the inverted baseline velocity model from the predicted monitor model. . . . .	81
Figure 6.3	Workflow of the velocity- and saturation-simulation algorithms. The dashed orange oval marks the prior information and the empirical relationship between the velocity and saturation obtained from well logs. The solid green oval marks the inverted baseline velocity field used for generating training velocity models, and the sonic and saturation logs used for obtaining saturation. The solid blue oval marks the synthetic velocity and saturation models generated for training. . . . .	82
Figure 6.4	Visualization of the actual 4D velocity model. The baseline and monitor surveys are simulated for the same 2D vertical profiles before and during (or after) CO <sub>2</sub> injection. The time-lapse changes are encircled in red. . . . .	83

Figure 6.5	3D geometry of the Kimberlina data set. The injection well and three monitor wells are denoted by the red and orange dots, respectively, in the $[x, y]$ -plane. The red line in the $[x, y]$ -plane shows the projection of the baseline and monitor surveys. The red arrows in the $[y, z]$ -plane point to three structures used for evaluating the effectiveness of time-lapse strategies: the low velocity layer above the reservoirs (Zone 1), the three CO <sub>2</sub> reservoirs (Zone 2), and the high-velocity dipping layer below the reservoirs (Zone 3). . . . .	83
Figure 6.6	P-wave velocity of the (a) baseline Kimberlina model and (b) monitor model in year 20 with a grid size of $10 \times 10$ m. CO <sub>2</sub> saturation for the (d) baseline and (e) monitor model in year 20. The actual time-lapse changes of the (c) P-wave velocity and (f) saturation. . . . .	84
Figure 6.7	Initial P-wave velocity for the inversion of the (c) baseline and (d) monitor data. P-wave velocity obtained by the physics-based FWI: (d) baseline velocity, (e) monitor velocity, and (f) time-lapse velocity variations. The dashed red lines on plot (d) mark the reservoir boundaries. . . . .	85
Figure 6.8	Actual baseline (a) velocity and (d) saturation models. The monitor (b) velocity and (e) saturation models predicted by the data-driven time-lapse inversion. The time-lapse variations of (c) velocity and (f) saturation. . . . .	86
Figure 6.9	Actual P-wave velocity model (first row) in (a) year 1 (b) year 5, (c) year 20, and (d) year 130. The second row shows the corresponding generated velocity samples. The actual saturation model (third row) in (i) year 1, (j) year 5, (k) year 20, and (l) year 130. The fourth row shows the corresponding generated saturation samples. . . . .	88
Figure 6.10	Well logs and empirical relationship between the velocity and CO <sub>2</sub> saturation for the injection well in year 20. The red and blue lines on plot (c) are two linear fitting equations separated according to the CO <sub>2</sub> saturation level. . . . .	90
Figure 6.11	Baseline (a) velocity and (d) saturation models inverted by physics-based FWI. The monitor (b) velocity and (e) saturation models predicted by the hybrid time-lapse inversion. The time-lapse variations of (c) velocity and (f) saturation. . . . .	92
Figure 6.12	(a) Actual time-lapse seismogram (the difference between the monitor data in year 21 and baseline data in year 0). Time-lapse seismogram produced by (b) the physics-based FWI, (c) the data-driven method, and (d) the hybrid strategy. Artifacts after the first arrivals are either encircled in purple or marked by red arrows. . . . .	92
Figure 6.13	Actual time-lapse velocity variations for (a) year 1, (b) year 5, and (c) year 130. Plots (d), (e), and (f) show the predicted time-lapse velocity variations. The actual time-lapse saturation variations for (h) year 1, (i) year 5, and (g) year 130. Plots (k), (l), and (m) show the predicted time-lapse saturation variations. . . . .	93
Figure 6.14	Temporal variations in the P-wave velocity obtained from noise-contaminated data (SNR = 10) by (a) the data-driven method and (b) the hybrid strategy. Plots (c) and (d) show the corresponding predicted time-lapse variations in the CO <sub>2</sub> saturation. . . . .	95
Figure 7.1	Flowchart of the method. (a) InvNet_UQ is trained and tested to estimate the velocity and the corresponding uncertainty. (b) The uncertainty-guided data-augmentation method is applied to the predicted velocity models; then the network is retrained with the augmented data. . . . .	100

Figure 7.2	Visualization of the actual 2D models sliced from the 3D Kimberlina time-lapse velocity model along the $y$ -direction. Three important geologic structures are marked by red arrows. . . . .	104
Figure 7.3	Actual P-wave velocity in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. The P-wave velocity obtained from noise-free seismic data by (e-h) InversionNet and (i-l) InvNet_UQ. . . . .	105
Figure 7.4	Difference between the actual velocity models and those predicted by InvNet_UQ in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. Plots (e-h) are the corresponding uncertainty maps. Plots (i-l) show the asymmetric uncertainty estimated from $(p_{upper} - p_{median}) - (p_{median} - p_{lower})$ , where $p_{upper}$ and $p_{lower}$ are the predictions at quantiles $\tau = 0.975$ and $\tau = 0.025$ , respectively, and $p_{median}$ is the median prediction at $\tau = 0.5$ . . . . .	106
Figure 7.5	Correlation map between the absolute prediction error (MAE) and the corresponding uncertainty. The Pearson correlation is 0.727 (p-value is $1.0 \times 10^{-10}$ ), which shows a strong positive correlation between the absolute prediction error and the uncertainty. . . . .	107
Figure 7.6	Seismic shot gather for the testing velocity model in year 20 (a) without noise, and (b) with Gaussian noise (the signal-to-noise ratio SNR is 10). . . . .	108
Figure 7.7	P-wave velocity obtained from seismic data with SNR=10 by (a-d) InversionNet and (e-h) InvNet_UQ. Plots (i-l) show the corresponding uncertainty computed by InvNet_UQ. . . . .	109
Figure 7.8	Comparison of InversionNet and InvNet_UQ in terms of the mean-square error (MSE) vs. signal-to-noise ratio (SNR). . . . .	110
Figure 7.9	Seismic shot gather for the testing velocity model in year 20 (a) without missing data, and (b) with 30% of data missing. . . . .	110
Figure 7.10	P-wave velocity obtained from the testing seismic data with 30% of the traces missing by (a-d) InversionNet and (e-h) InvNet_UQ. Plots (i-l) show the corresponding uncertainty from InvNet_UQ. . . . .	111
Figure 7.11	Seismic shot gather for the testing velocity model in year 20: (a) without a high bandpass filter, and (b) with the Butterworth filter that has a cut-off frequency at 3 Hz. Plot (c) shows the frequency spectrum of the data. The blue dashed line on plot (c) marks the frequencies removed by the filter. . . . .	112
Figure 7.12	P-wave velocity obtained from the testing seismic data using a high bandpass filter with a cut-off frequency of 3 Hz by (a-d) InversionNet and (e-h) InvNet_UQ. Plots (i-l) show the corresponding uncertainty from InvNet_UQ. . . . .	113
Figure 7.13	Velocity model in year 1: (a) actual and (b) predicted. Plots (c) - (f) show the augmented data from the prediction. . . . .	113
Figure 7.14	P-wave velocity predicted by InvNet_UQ trained using the augmented data in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. Plots (e)-(h) show the corresponding uncertainty maps. . . . .	114
Figure C.1	Permission from <i>Journal of Seismic Exploration</i> . . . . .	135

Figure C.2 Permission from *Geophysics* highlighted in yellow. . . . . 136

## LIST OF TABLES

Table 6.1	Computational cost of the monitor inversion using the physics-based and hybrid strategies. .	91
Table 7.1	Comparison of the MSE values for the velocity distribution predicted by InversionNet and InvNet_UQ applied without data augmentation (columns 2 & 3) and with data augmentation (column 4). . . . .	107

## ACKNOWLEDGMENTS

I want to extend my heartfelt appreciation to Professor Ilya Tsvankin, my advisor, for his unwavering support and invaluable guidance throughout my Ph.D. journey. Despite my status as an undergraduate student with limited research and geophysics knowledge, Dr. Ilya graciously accepted me into the program and has been a constant source of inspiration. His exceptional seismology courses, particularly his lectures on seismic anisotropy, have been the most enlightening and advantageous experiences during my time at CSM. I have always held great admiration for his unwavering dedication and commitment to research, teaching, and writing.

I extend heartfelt thanks to my esteemed thesis committee members: Prof. Ali Tura, Prof. Jeffrey Shragge, Prof. Mahadevan Ganesh, Prof. Michael Wakin, and Dr. Youzuo Lin. Their valuable insights, comments, and suggestions on my comprehensive exams and thesis have played a pivotal role in shaping my research. I am especially grateful to Prof. Mahadevan Ganesh for his parallel computing course, which has proven immensely useful in my Ph.D. studies. Additionally, I would like to express my sincere appreciation to Dr. Youzuo Lin for providing me with the opportunity to work with his team for nearly two years, during which I gained invaluable knowledge about machine-learning-based inversion.

I would like to acknowledge the fellow students at CWP who generously shared their time, brainstormed ideas with me, and accompanied me on hiking adventures. Their camaraderie and support have made my academic journey truly enjoyable. I thank Dr. Sagar Singh (currently at Nvidia) for his help with developing and implementing the elastic anisotropic FWI algorithm. Furthermore, I would like to extend my gratitude to the Department of Geophysics and the CWP staff, especially Michelle, Noelle, and Lynn, for their invaluable assistance with administrative tasks. I am also indebted to the CWP sponsors for their generous support of my research.

My 4-year PhD journey were pleasant because of great friends like Mengxia Ren, Xingyuan Zhao, Haipeng Li, Peiyao Li, Zhuo Liu, Mengli Zhang, Donglin Zhu and Youfang Liu, who providing me with many laughs and great time. Big thanks to my best friend, Chi Zhang, who cheered me on, celebrated every milestone, and stood by me through thick and thin.

Last but certainly not least, I am eternally grateful to my family (my parents, my sister and brother) in China for their unwavering love and support. Their unshakeable faith in me has been a constant source of motivation, and I owe them a debt of gratitude that words cannot adequately express.

# CHAPTER 1

## INTRODUCTION

Time-lapse (4D) full-waveform inversion (FWI) is a powerful technique for high-resolution monitoring of temporal changes in the subsurface using baseline and monitor seismic data (e.g., Lumley, 2010; Pevzner et al., 2017; Smith and Tsvankin, 2013, 2016). This method has proven successful in optimizing both hydrocarbon production from oil/gas reservoirs and CO<sub>2</sub> injection in carbon-sequestration projects.

In time-lapse FWI, baseline and monitor data sets acquired at different times are inverted to estimate 4D subsurface changes. Various approaches have been proposed to obtain sufficiently accurate time-lapse results. The parallel-difference (PD) strategy (Plessix et al., 2010), for example, involves independent inversion of the baseline and monitor data sets. Alternatively, the sequential-difference (SD) strategy (Routh et al., 2012) facilitates convergence of the monitor inversion by starting it with the estimated baseline model. Assuming good data repeatability, the double-difference (DD) approach (Denli and Huang, 2009; Watanabe et al., 2004) directly inverts the data difference for the time-lapse parameter variations.

Maharramov and Biondi (2014) propose to invert the baseline and monitor data simultaneously using a cross-updating algorithm. To reduce the dependence of the estimated time-lapse changes on the accuracy of the inverted baseline model, the central-difference strategy (CD) averages the parameter variations produced by forward (baseline to monitor) and reverse (monitor to baseline) applications of the SD technique (Zhou and Lumley, 2021a). However, the two latter approaches double the computational time compared to the SD and DD methods without a significant improvement in resolution.

Extending published FWI methods designed for isotropic media (Egorov et al., 2017; Hicks et al., 2016; Yang et al., 2016) to more realistic anisotropic models increases the computational cost and introduces trade-offs (cross-talk) between medium parameters. Analysis of the radiation (sensitivity) patterns of the medium parameters provides useful insights into the potential trade-offs and helps in choosing an optimal parameterization to mitigate them (Kamath and Tsvankin, 2016; Singh and Tsvankin, 2020). Here, I employ a velocity-based parameterization because all velocity parameters and their gradients have the same units and similar magnitudes.

Accurate determination of the source wavelet is a critical factor in the success of FWI. Errors in the source signature (in its shape, frequency, or amplitude) can impede the matching of the observed and simulated data, resulting in distorted inversion results (Luo et al., 2014; Pratt, 1999; Song et al., 1995; Yuan et al., 2014). A conventional way to estimate the source wavelet is by using water-bottom and direct arrivals. However, this approach requires additional preprocessing steps including deghosting and multiple

suppression (Warner et al., 2013). Alternatively, the source wavelet can be estimated iteratively during the inversion along with the medium parameters (Pratt, 1999; Song et al., 1995; Xu et al., 2006). Inversion that includes the source wavelet, however, is computationally intensive, introduces additional trade-offs, and requires an accurate initial approximation for the source wavelet (Xu et al., 2006).

A more practical alternative is the so-called “source-independent” method designed to reduce the influence of the source signature on the inversion results. Deconvolution-based trace normalization employed in the frequency domain helps remove information about the source wavelet from both the observed and simulated data (Choi and Min, 2012; Choi et al., 2005; Lee and Kim, 2003; Seo et al., 2005; Zhou and Greenhalgh, 2003). However, the resolution achieved by this deconvolution-based method strongly depends on the signal-to-noise ratio.

Choi and Alkhalifah (2011) propose a convolution-based source-independent (SI) technique and implement it in the time domain. Their approach seeks to minimize the difference between two data sets obtained by convolving the observed and simulated data with a reference trace from the other data set. The convolution and cross-correlation operations can produce additional noise, but it can be suppressed by applying a time window to the reference traces (Zhang et al., 2016). Here, the time-domain SI technique is incorporated into the developed time-lapse FWI algorithm for elastic VTI and TTI media.

Accurate estimation of subsurface changes from time-lapse surveys is often challenging due to nonrepeatability (NR) issues. Such NR factors, as geometry changes, water statics, and source-signature variations can distort the data difference and the inversion results (Jervis et al., 2018; Lumley, 2001a). Permanent acquisition systems have been proposed to make the acquisition geometry more consistent and mitigate 4D artifacts. Also 4D methodologies have been modified to correct for the velocity variations in the water layer in time-lapse inversion (Lecerf et al., 2022). Smith and Tsvankin (2013) analyze the stress changes caused by the reservoir production in the surrounding rocks. Such changes can provide valuable information for reservoir monitoring, but can also lead to misinterpretation of time-lapse data. The impact of different NR issues on acoustic time-lapse parameter inversion and on the performance of 4D methods is studied by Zhou and Lumley (2021b). Here, analysis of the nonrepeatability-related errors is extended to elastic TI media.

In recent years, data-driven machine-learning methods for seismic inversion have gained popularity due to the increasing computational power and the emergence of deep neural networks (ArayaPolo et al., 2017; Cao and Roy, 2017; Dahlke et al., 2016; Hale, 2013b; Li et al., 2021; Yuan et al., 2019). In contrast to conventional inverse problems, where the distribution shift between the training and testing data can pose a challenge, time-lapse monitoring of CO<sub>2</sub> injection is relatively immune to this issue because the medium outside the reservoir can be assumed to remain almost unchanged. Furthermore, uncertainty evaluation

has been introduced to quantify the prediction accuracy in such monitoring applications. Tang et al. (2022) propose a 3D recurrent R-U-Net surrogate model to predict CO<sub>2</sub> saturation from synthetic seismic data and estimate the uncertainty using the CNN (convolutional neural network)-PCA (principle-component analysis) model with rejection sampling. Likewise, Um et al. (2022) employ a U-Net network to estimate CO<sub>2</sub> saturation and model uncertainty using the Monte Carlo dropout method and a bootstrap aggregation method. Here, I develop a neural network to predict both velocity and saturation models, and employ the Simultaneous Quantile Regression method for uncertainty estimation.

### 1.1 Thesis Contributions

In Chapter 2, I develop a time-lapse FWI algorithm for 2D elastic VTI media. The FWI code is a modified version of the software from the public package by Köhn et al. (2012). The gradient of the objective function is derived using the adjoint-state method and a nonlinear conjugate-gradient technique is employed to update the model parameters. The developed algorithm can accurately reconstruct localized time-lapse parameter variations with sufficient spatial resolution for data with moderate noise, even in the absence of ultra-low frequencies (i.e., below 2 Hz). The study was published in the *Journal of Seismic Exploration* (Liu and Tsvankin, 2021).

FWI requires an accurate estimate of the source signature. However, as discussed above, obtaining the source wavelet either directly from the data or during the inversion is time-consuming and error-prone. In Chapter 3, I employ the time-domain “source-independent” technique in the developed time-lapse FWI algorithm for VTI media. Tests on the 2D graben and Marmousi VTI models demonstrate that the developed source-independent algorithm improves the accuracy of the estimated time-lapse variations compared with the conventional  $L_2$ -norm-based FWI, even for significantly distorted source signals (e.g., the spike wavelet). This technique is particularly important in time-lapse processing because it can handle the nonrepeatability of the source wavelet. This work was published in *Geophysics* (Liu and Tsvankin, 2022).

Chapter 4 addresses the limitations of VTI algorithms in the presence of even moderate symmetry-axis tilt by generalizing the time-lapse FWI methodology for 2D TI media with a tilted symmetry axis (TTI). The tilt model is estimated from the structural dips and is updated after completing the inversion for each frequency band. Testing on the BP TTI model demonstrates the improvements in reconstructing subsurface changes achieved by taking the tilt into account. I also investigate the influence of several common nonrepeatability issues on the inversion results and incorporate the source-independent (SI) technique into the TTI algorithm. This work is under review in *Geophysics*.

In Chapter 5, the developed time-lapse methodology and the SI technique are applied to 4D data from the Pyrenees oil/gas field in Australia. Because the geologic structure of the field is relatively flat, the VTI algorithm provides sufficiently accurate results. The masking technique that tapers the parameter gradient in the water is extended to a nonflat water bottom for this case study. Three different objective functions are used in the baseline inversion, with the global-correlation (GC) one providing the highest-quality results. The estimated parameter variations reveal the reservoir changes likely caused by gas coming out of solution and the replacement of gas with oil. Comparison with the output of the isotropic 4D FWI algorithm confirms the importance of taking anisotropy into account in time-lapse inversion. This work is being submitted to *Geophysics*.

To make inversion of time-lapse seismic data more efficient, in Chapter 6 I develop a convolutional neural network (CNN) that predicts the spatial and temporal variations in both velocity and saturation. This network helps devise a “hybrid” workflow for real-time monitoring that combines physics-based FWI and data-driven CNN. The data scarcity problem is mitigated by a novel data-generation technique that employs physics-based constraints. The developed hybrid methodology can predict the variations in velocity and saturation simultaneously with similar or higher spatial resolution compared with time-lapse FWI, even in the presence of realistic noise in the data. This work was published in *Geophysics* (Liu et al., 2022).

Chapter 7 focuses on uncertainty quantification (UQ), which provides valuable information about the prediction accuracy of time-lapse inversion. I incorporate Simultaneous Quantile Regression (SQR) into the CNN developed in Chapter 6 to estimate the conditional distributions of all pertinent quantiles using the pinball loss. The resulting uncertainty map is obtained from the prediction interval, which can be asymmetric with respect to the predicted mean value - a feature that cannot be achieved by traditional UQ methods. The proposed network exhibits higher robustness and generalizability than the conventional CNN that uses an  $L_1$ -norm objective function. This work is under review in *Geophysics*.

Finally, Chapter 8 presents the thesis conclusions and recommendations for future work.

## CHAPTER 2

### METHODOLOGY OF TIME-LAPSE ELASTIC FULL-WAVEFORM INVERSION FOR VTI MEDIA

A paper published\* in *Journal of Seismic Exploration*

Yanhua Liu<sup>1,2</sup>, Ilya Tsvankin<sup>2</sup>

Time-lapse seismic processing can provide important information about the variations of reservoir properties during hydrocarbon production and CO<sub>2</sub> injection. High-resolution results for time-lapse seismic can potentially be obtained from full-waveform inversion (FWI), but most existing time-lapse FWI methods are limited to isotropic and, often, acoustic media. Extension of these techniques to more realistic anisotropic elastic models is hampered by the trade-offs between the medium parameters and significantly increased computational cost. Here, we develop a time-lapse FWI algorithm for VTI (transversely isotropic with a vertical symmetry axis) media and evaluate several strategies of applying it to multicomponent and pressure data. The adjoint-state method and a nonlinear conjugate-gradient technique are employed to derive the gradient of the objective function and update the model parameters. We test the algorithm on a relatively simple VTI graben model using the parallel-difference, sequential-difference and double-difference time-lapse methods. The results confirm the ability of the proposed technique to reconstruct localized time-lapse parameter variations in anisotropic media with sufficient spatial resolution. The double-difference approach proves to be more accurate than the other methods in reconstructing the time-lapse variations from noise-free multicomponent data. When FWI operates with clean pressure data, the parallel-difference method is generally more accurate than the other techniques, especially in estimating the shear-wave vertical velocity  $V_{S0}$ . For multicomponent and pressure data contaminated with realistic noise, the double-difference method produces large errors in the temporal variations of the VTI parameters. The parallel-difference technique outperforms its sequential-difference counterpart in reconstructing the time-lapse variations inside the target zone from the noisy data while the latter approach performs better in suppressing the false artifacts outside the “reservoir”. The tests also demonstrate that including more information in time-lapse FWI does not always improve the inversion results, likely due to the increased multimodality of the objective function.

---

\*Reprinted with permission of Journal of Seismic Exploration, 2021, 30, 257-270.

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Department of Geophysics, Colorado School of Mines, Golden, CO, USA

## 2.1 Introduction

Time-lapse (4D) seismic has become a common tool for optimizing hydrocarbon reservoir production and  $CO_2$  injection (e.g., Lumley, 2010; Pevzner et al., 2017; Smith and Tsvankin, 2013, 2016). Processing of time-lapse data can provide important information about the variations of the reservoir properties, such as pressure and fluid saturation.

Full-waveform inversion has been successfully applied to velocity analysis and reservoir characterization (Asnaashari et al., 2015; Singh et al., 2018; Vigh et al., 2014) and can potentially provide estimates of time-lapse parameter variations with high spatial resolution. Among the proposed strategies of time-lapse FWI are the parallel-difference (Plessix et al., 2010), sequential-difference, and double-difference methods (Denli and Huang, 2009; Watanabe et al., 2004), as well as joint (Alemie and Sacchi, 2016) and simultaneous (Maharramov and Biondi, 2014) inversion techniques. The parallel-difference method uses the same initial model for the baseline and monitor inversions, while the sequential-difference method employs the inverted baseline data to build the initial model for the inversion of the monitor data. The double-difference technique directly inverts the time-lapse data difference starting with the inverted baseline model.

However, most published time-lapse FWI algorithms are limited to isotropic and, sometimes, acoustic media. Extension of these methods to more realistic anisotropic models involves serious challenges, including the much higher computational cost and the trade-offs between multiple parameters needed to describe anisotropic formations (Kamath and Tsvankin, 2016; Singh et al., 2018).

Here, we present a time-lapse FWI methodology for VTI media and test it on a relatively simple graben model, where the “reservoir” is located in thin dipping layers. Three most common time-lapse approaches are applied to estimate the temporal variations of the VTI parameters using both multicomponent and pressure data. Analysis of the inversion results reveals the advantages and shortcomings of the employed strategies in reconstructing the time-lapse VTI model.

## 2.2 Methodology of time-lapse FWI for VTI media

The goal of time-lapse FWI is to estimate the temporal variations of the medium parameters using seismic data before (baseline survey) and during or after (monitor survey) hydrocarbon production or  $CO_2$  injection. We parameterize VTI media by  $V_{P0}$  (P-wave vertical velocity),  $V_{S0}$  (S-wave vertical velocity),  $V_{hor,P}$  (P-wave horizontal velocity),  $V_{nmo,P}$  (P-wave normal-moveout velocity from a horizontal reflector), and  $\rho$  (density). These parameters fully describe the properties of P- and SV- waves for vertical transverse isotropy. The advantages of this notation for elastic FWI are discussed by Kamath et al. (2017).

The velocities  $V_{\text{hor,P}}$  and  $V_{\text{nm0,P}}$  are expressed through the Thomsen parameters  $\varepsilon$  and  $\delta$  as follows (Thomsen, 1986; Tsvankin, 2012):

$$V_{\text{hor,P}} = V_{\text{P0}} \sqrt{1 + 2\varepsilon}, \quad (2.1)$$

$$V_{\text{nm0,P}} = V_{\text{P0}} \sqrt{1 + 2\delta}. \quad (2.2)$$

Typically, a conventional FWI algorithm is applied to the baseline data, while the monitor survey can be processed using several different approaches. Here, we focus on the parallel-, sequential-, and double-difference methods, which share the

same  $L_2$ -norm objective function (e.g., Tarantola, 1984) for the inversion of the baseline (subscript b) data:

$$S_b(\mathbf{m}) = \frac{1}{2} \left\| \mathbf{W}_b \left[ \mathbf{d}_b^{\text{sim}}(\mathbf{m}_b) - \mathbf{d}_b^{\text{obs}} \right] \right\|^2. \quad (2.3)$$

where  $\mathbf{d}_b^{\text{sim}}$  is the data simulated for the baseline model  $\mathbf{m}_b$ ,  $\mathbf{d}_b^{\text{obs}}$  is the observed data, and  $\mathbf{W}_b$  is the weighting data-misfit operator.

To simulate multicomponent seismic data, the 2D elastic wave equation for arbitrarily heterogeneous VTI media is solved with a fourth-order finite-difference algorithm. Model updating based on the objective function in equation 2.3 is performed with the conjugate-gradient technique, and FWI is implemented using a multiscale approach.

The three strategies examined here differ in the way they handle the monitor survey or the time-lapse data difference, as described in more detail below.

### 2.2.1 Parallel-difference method

In the parallel-difference approach (Plessix et al., 2010), the baseline and monitor inversions are performed independently but with the same initial model. Then the time-lapse variation in each parameter is obtained by subtracting the inversion results for the two surveys. If the errors in the inverted baseline and monitor models are similar (for example, have the same sign), the time-lapse model produced by the parallel-difference method can remain sufficiently accurate.

### 2.3 Sequential-difference method

The sequential-difference strategy (Asnaashari et al., 2012) uses the inverted baseline model as the initial model in FWI of the monitor data. Because the time-lapse variation is largely confined to the reservoir, the baseline model should provide a good approximation for the parameters obtained from the monitor data outside the target area. Then, as in the parallel-difference method, the inverted parameters for the baseline survey are subtracted from those for the monitor survey.

By using the inverted baseline model as the initial model for the monitor reconstruction, the sequential-difference method speeds up the convergence of the updating algorithm and partially mitigates parameter trade-offs and problems caused by local minima of the objective function. However, this approach relies on an accurate reconstruction of the baseline model, which might be problematic, for example, if the baseline data are noisy.

### 2.3.1 Double-difference method

In the double-difference method (Denli and Huang, 2009; Watanabe et al., 2004), FWI is first applied to the baseline data  $\mathbf{d}_b^{\text{obs}}$  to obtain the inverted baseline model  $\mathbf{m}_b^{\text{inv}}$ . Next, the so-called “composite” data set  $\mathbf{d}_{\text{com}}$  is generated by adding the observed time-lapse data  $\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}$  to the wavefield simulated for the inverted baseline model. Then FWI operates directly with the time-lapse response by minimizing the difference  $\Delta \mathbf{d}$  between the simulated monitor data  $\mathbf{d}_m^{\text{sim}}$  and the “composite” data”  $\mathbf{d}_{\text{com}}$ , starting from the inverted baseline model:

$$\Delta \mathbf{d} = (\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}) - (\mathbf{d}_m^{\text{sim}} - \mathbf{d}_b^{\text{sim}}) = \mathbf{d}_{\text{com}} - \mathbf{d}_m^{\text{sim}}, \quad (2.4)$$

$$\mathbf{d}_{\text{com}} = \mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}} + \mathbf{d}_b^{\text{sim}}. \quad (2.5)$$

Finally, the time-lapse parameter variations are obtained by subtracting the inverted baseline model from the reconstructed “composite” (i.e., monitor) model.

## 2.4 Synthetic Examples

The proposed time-lapse FWI algorithm is tested on a 2D VTI model that includes a graben structure (Figure 2.1a, d, g, j, and m). The time-lapse model for the monitor survey is obtained by reducing the baseline velocities  $V_{P0}$  and  $V_{S0}$  in the target area (i.e., in the dipping layer segments) by approximately 13%, and density  $\rho$  by 10% (Figure 2.1c, f, i, l, and o). The initial baseline parameters are computed by Gaussian smoothing of the actual parameter distributions (Figure 2.1b, e, h, k, and n).

The elastic wavefield is excited by 116 shots (which represent point explosions) placed with a constant increment along a horizontal line at a depth of 40 m. The source signal is a Ricker wavelet with a central frequency of 10 Hz. We employ 400 receivers evenly distributed along the horizontal line at a depth of 100 m (Figure 2.1a).

Following Singh et al. (2019), FWI is performed using a multiscale approach with four frequency bands (2-5 Hz, 2-8 Hz, 2-13 Hz, 2-19 Hz). The low-frequency data (0-2 Hz), which usually are difficult to record in the field, are not included in the inversion. The accuracy of the three time-lapse methods is evaluated by computing the normalized errors in the temporal variation of each parameter at all grid points.

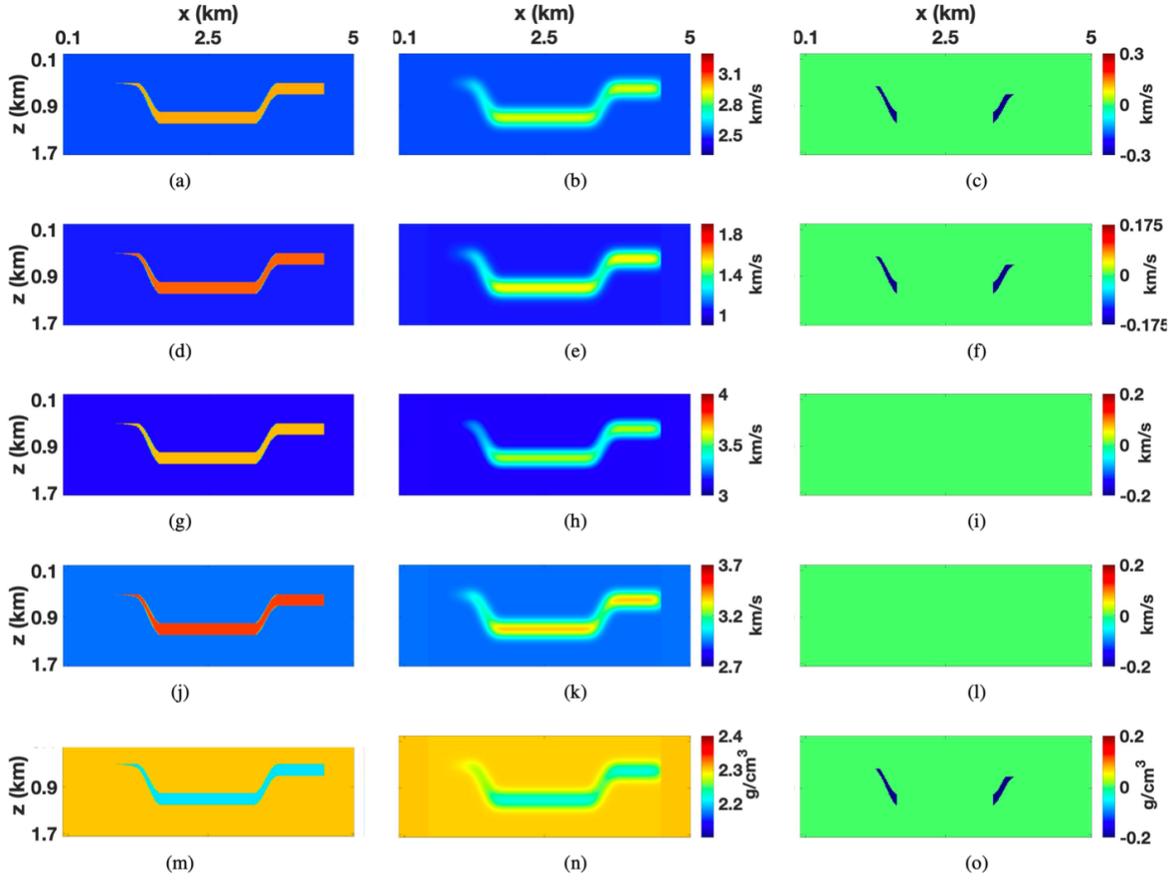


Figure 2.1 Parameters of the baseline model with a grid size of  $10 \times 10$  m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (d) the S-wave vertical velocity ( $V_{S0}$ ), (g) the P-wave horizontal velocity ( $V_{hor,P}$ ), (j) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ), and (m) the density ( $\rho$ ). The initial baseline model of: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The actual time-lapse differences for (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

### 2.4.1 Multicomponent data

First, the noise-free multicomponent data for the baseline model are inverted using the anisotropic FWI algorithm described above. While the overall accuracy of parameter estimation is satisfactory, there are noticeable errors at and near the boundaries of the graben structure due to edge effects (Figure 2.2).

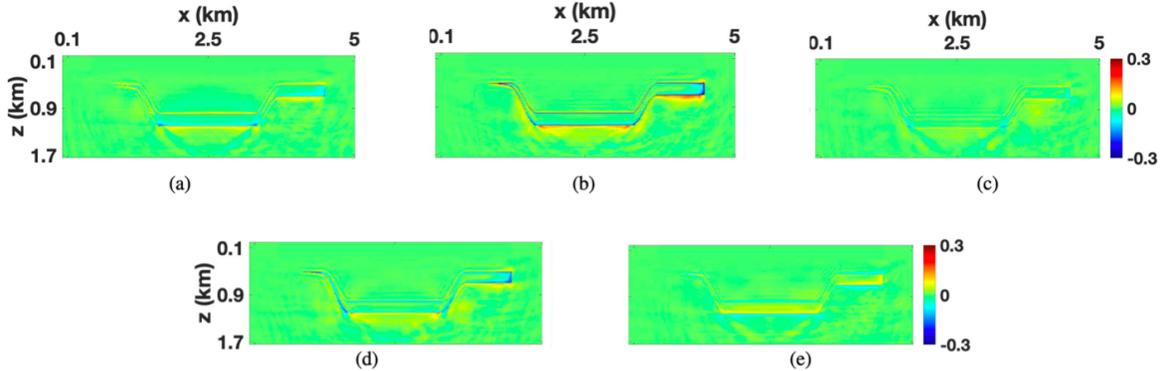


Figure 2.2 Normalized parameter errors for the final baseline model reconstructed from the multicomponent data: (a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $V_{hor,P}$ , (d)  $V_{nmo,P}$  and (e)  $\rho$ .

Then we apply the parallel-, sequential- and double-difference methods to evaluate the time-lapse variations in the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  (the NMO and horizontal velocities are held constant). In our application of the sequential-difference method below, the inverted baseline model for the 2-5 Hz frequency range is used as the initial model for the monitor inversion. Similar to the baseline inversion, all three techniques (Figure 2.3) produce errors in the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  near the boundaries of the target area. These errors are caused primarily by the jumps in  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  across the boundaries of the target for the monitor survey and for the “composite” data used in the double-difference method.

Outside the target area, the double-difference method (Figure 2.3c, f, and o) yields more accurate estimates of  $V_{P0}$ ,  $V_{S0}$  and  $\rho$  (i.e., it shows no changes) than the other methods (Figure 2.3a, d, m, b, e, and n) because it is specifically designed for inverting the time-lapse variations. The three methods reconstruct the time-lapse changes of  $V_{P0}$ ,  $V_{S0}$  and  $\rho$  inside the target zone with similar accuracy.

Although there are no temporal variations in the velocities  $V_{hor,P}$  and  $V_{nmo,P}$ , all methods generate false time-lapse anomalies in them (Figure 2.3g, j, h, k, i, and l) because of the parameter trade-offs. As is the case for the inversion results for  $V_{P0}$ ,  $V_{S0}$  and  $\rho$ , the false anomalies in  $V_{hor,P}$  and  $V_{nmo,P}$  (Figure 2.3i and l) are less pronounced (and mainly located near the target boundaries) in the output of the double-difference method. In contrast, the errors in  $V_{hor,P}$  and  $V_{nmo,P}$  produced by the parallel- and sequential-difference approaches are observed not only near the target boundaries, but also beneath the graben structure. The double-difference method, however, generates a pronounced false perturbation in

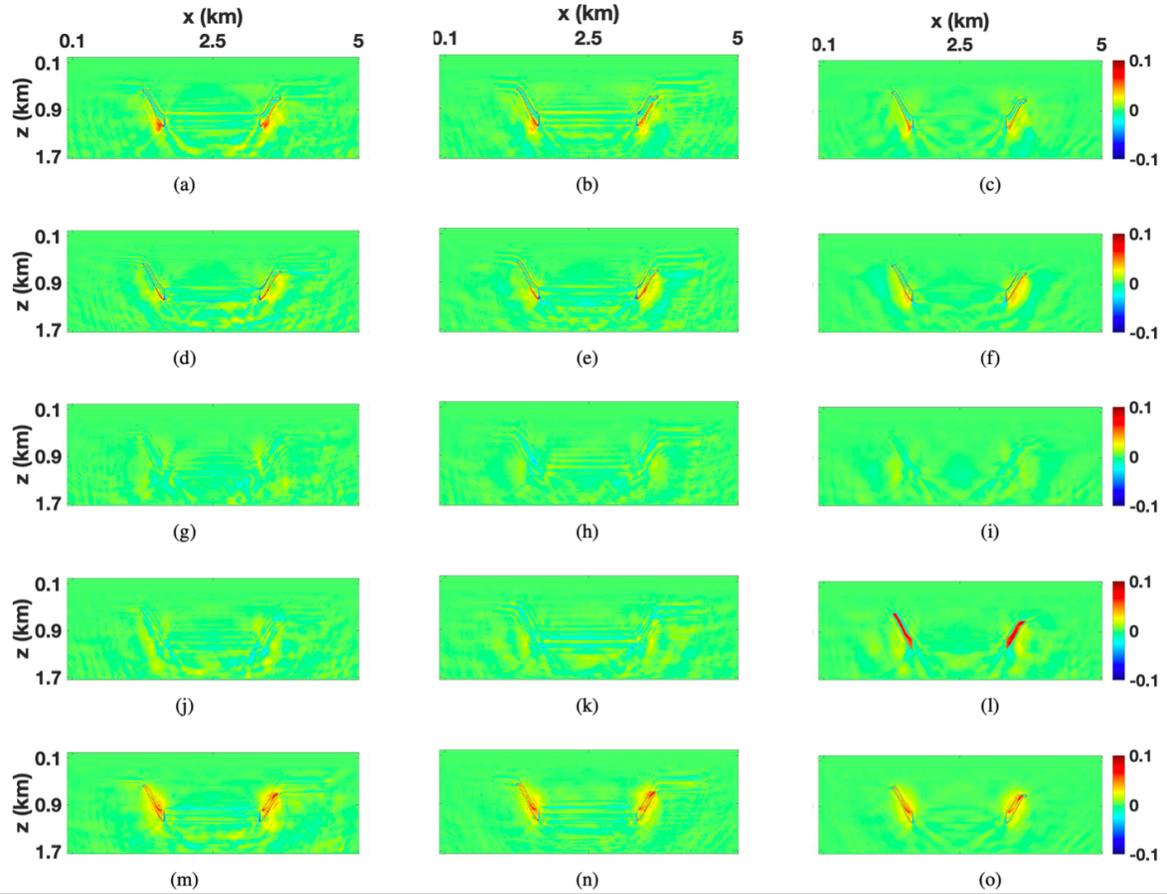


Figure 2.3 Normalized errors of the time-lapse parameter variations obtained from the multicomponent data using three different approaches. Only the parameters  $V_{P0}$ ,  $V_{S0}$  and  $\rho$  are changed. The parallel-difference method: (a)  $V_{P0}$ , (d)  $V_{S0}$ , (g)  $V_{hor,P}$ , (j)  $V_{nmo,P}$ , and (m)  $\rho$ . The sequential-difference method: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The double-difference method: (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

$V_{\text{nm},P}$  inside the target zone (Figure 2.3l).

To analyze the leakage between the temporal variations of different parameters, next we change only one parameter ( $V_{P0}$ ,  $V_{S0}$ , or  $\rho$ ) at a time. Here, we discuss just the test for the perturbation in the velocity  $V_{S0}$ ; the results for the parameters  $V_{P0}$  and  $\rho$  are generally similar. Interestingly, the time-lapse inversion errors in this test (Figure 2.4) are larger compared to the previous experiment in which we changed three parameters. Apparently, for this model the distortions related to the temporal variations in several parameters partially compensate one another.

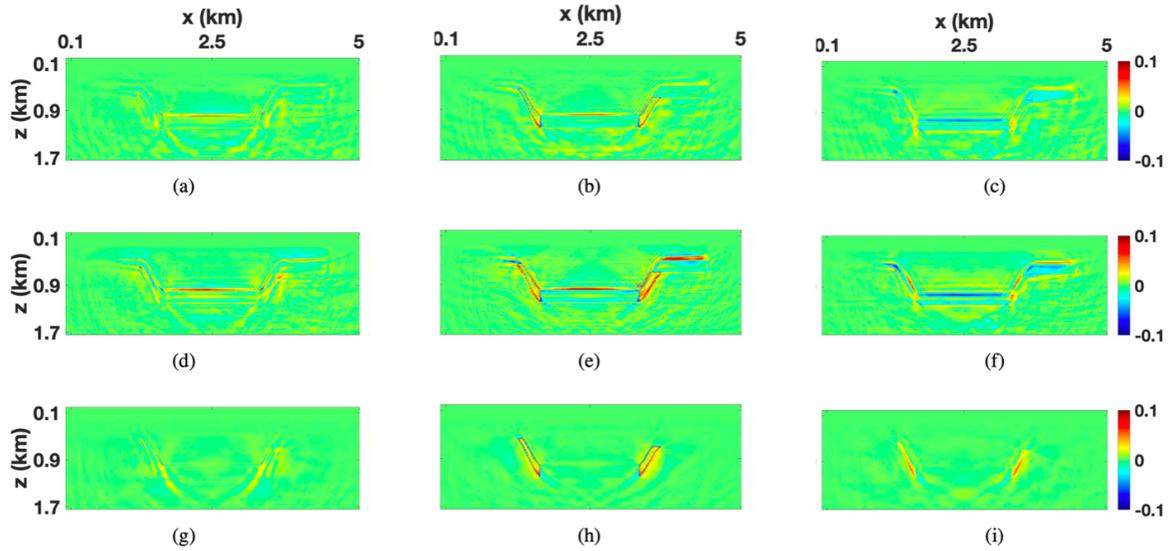


Figure 2.4 Normalized errors of the time-lapse parameter variations obtained from the multicomponent data. Only the parameter  $V_{S0}$  is changed. The parallel-difference method: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $V_{\text{nm},P}$ . The sequential-difference method: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $V_{\text{nm},P}$ . The double-difference method: (g)  $V_{P0}$ , (h)  $V_{S0}$ , and (i)  $V_{\text{nm},P}$ .

As in the previous test (Figure 2.3), the double-difference method produces the most accurate results outside the target area for all parameters (where they remain unchanged), as well as for  $V_{\text{hor},P}$  inside the “reservoir” (Figure 2.4). Still, there is a significant false anomaly in the velocity  $V_{\text{nm},P}$  (Figure 2.4i) in the output of the double-difference algorithm. A similar anomaly in the velocity  $V_{\text{nm},P}$  in the previous test (Figure 2.3l) is apparently caused by the temporal variation in  $V_{S0}$  because it does not appear when we change just the velocity  $V_{P0}$  or density  $\rho$ .

#### 2.4.2 Pressure data

Because pressure recordings contain less information than the displacement field (in particular, about shear waves), the resolution of the inverted parameters (Figure 2.5) is lower compared to that for the multicomponent data. However, the absence of the false time-lapse anomaly in  $V_{\text{nm},P}$  (Figure 2.5l) in the

results of the double-difference method indicates that FWI of pressure data may be less influenced by the parameter trade-offs. Although adding multicomponent data typically increases parameter resolution, it can also make the objective function more multimodal, which hinders convergence toward the global minimum of the objective function. This observation is also made by Kamath and Tsvankin (2013) who performed elastic FWI for horizontally layered VTI media.

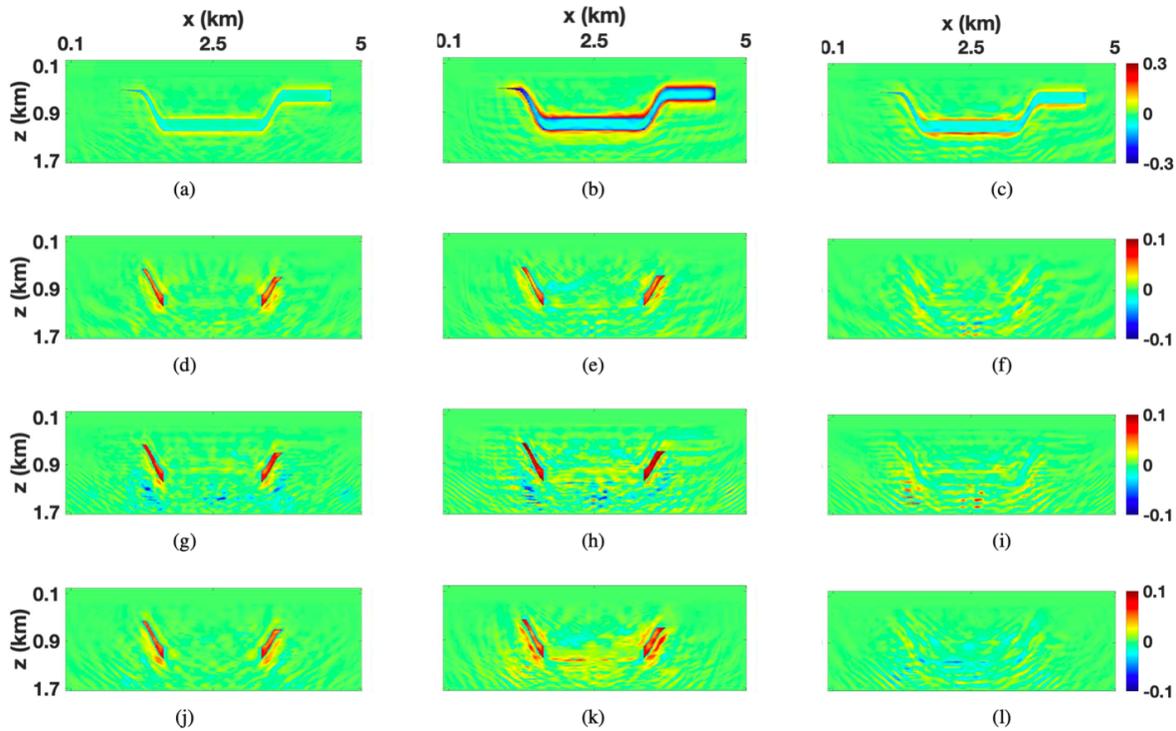


Figure 2.5 Baseline models reconstructed from the pressure data: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $V_{nmo,P}$ . Normalized errors of the time-lapse parameter variations reconstructed from the pressure data. Only the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  are changed. The parallel-difference method: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $V_{nmo,P}$ . The sequential-difference method: (g)  $V_{P0}$ , (h)  $V_{S0}$ , and (i)  $V_{nmo,P}$ . The double-difference method: (j)  $V_{P0}$ , (k)  $V_{S0}$ , and (l)  $V_{nmo,P}$ .

Note that the algorithm was able to reconstruct the graben structure in the fields of all medium parameters (including density, which is not shown in Figure 2.5) because FWI operates with the entire elastic wavefield and the acquisition geometry in our experiment covers a wide range of offsets. However, the spatial resolution at and below the graben boundaries is noticeably lower than that obtained from the multicomponent data. As a result, the time-lapse variations estimated by the sequential-difference method (Figure 2.5g, h, and i) have a lower resolution compared to the other two methods. This is explained by the strong reliance of the sequential-difference approach on the accuracy of the inverted baseline model.

Predictably, for the baseline, monitor, and “composite” data, the results of the pressure inversion are less accurate than those for multicomponent data. However, the errors in the recovered baseline and monitor models are similar and, therefore, partially cancel each other in the output of the parallel-difference method. Therefore, the parallel-difference method (Figure 2.5d and e) outperforms its double-difference counterpart (Figure 2.5j and k) in suppressing false anomalies in  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$ . The advantages of the parallel-difference method are especially obvious in the time-lapse model of  $V_{S0}$  because this parameter is not well constrained by FWI of the pressure data (Figure 2.5b). The artifacts in the time-lapse models of the velocities  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  (which are unchanged) for the parallel- and double-difference methods have close magnitudes, but different spatial distributions (Figure 2.5f and l). On the whole, the accuracy of these two methods for  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  is comparable, but the parallel-difference method yields better estimates of the changes in  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$ .

### 2.4.3 Noisy data

Next, the multicomponent and pressure data are contaminated with Gaussian noise to evaluate the robustness of FWI with the different time-lapse strategies. The double-difference method is essentially designed to invert the data difference (equations 2.4 and 2.5), for which the relative magnitude of the noise is much higher than for each data set separately. Hence, the time-lapse models produced by the double-difference method are distorted much more significantly compared to the two other techniques, and the actual time-lapse variations can be hardly identified from the inversion results (Figure 2.6 and Figure 2.7). Predictably, the resolution of the time-lapse variations reconstructed by all three methods decreases with the reduction in the signal-to-noise ratio (SNR; see Figure 2.3, Figure 2.6 and Figure 2.7). With increasing level of noise, the amplitude of the time-lapse variations inside the “reservoir” for the parameters  $V_{P0}$ ,  $\rho$ , and, especially,  $V_{S0}$  generally decrease (i.e., the changes in these parameters are underestimated). At the same time, the false anomalies in these parameters as well as in the velocities  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  outside the target area become more pronounced (Figure 2.6 and Figure 2.7).

The time-lapse changes inside the target zone reconstructed by the parallel-difference method are more accurate than those obtained by the sequential-difference method. In contrast, the latter method performs slightly better outside the “reservoir” (Figure 2.6 and Figure 2.7). As in the previous test, the time-lapse models produced by all three methods have a lower resolution for the noisy pressure records than for the noisy multicomponent data.

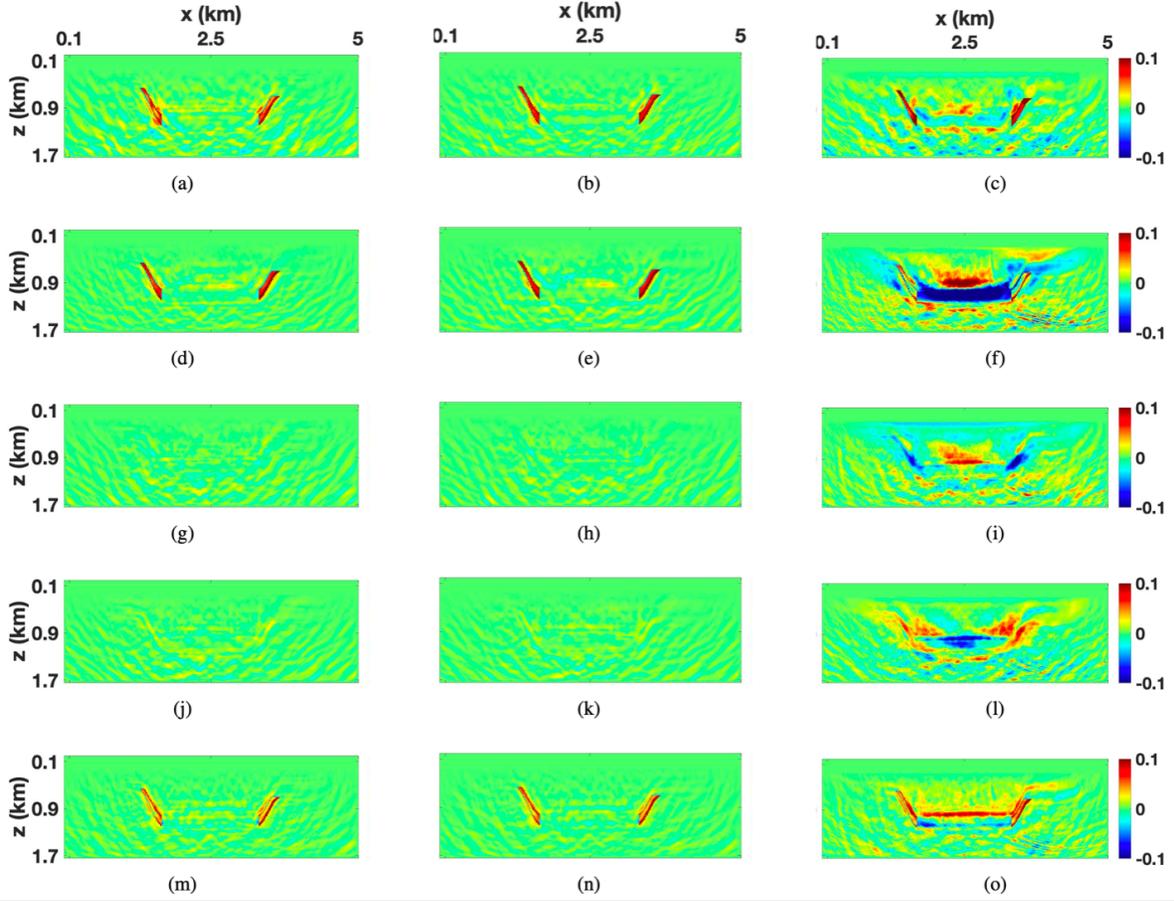


Figure 2.6 Parameters of the time-lapse model produced by three different strategies from the multicomponent data (SNR = 16). The parallel-difference method: (a)  $V_{P0}$ , (d)  $V_{S0}$ , (g)  $V_{hor,P}$ , (j)  $V_{nmo,P}$ , and (m)  $\rho$ . The sequential-difference method: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The double-difference method: (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

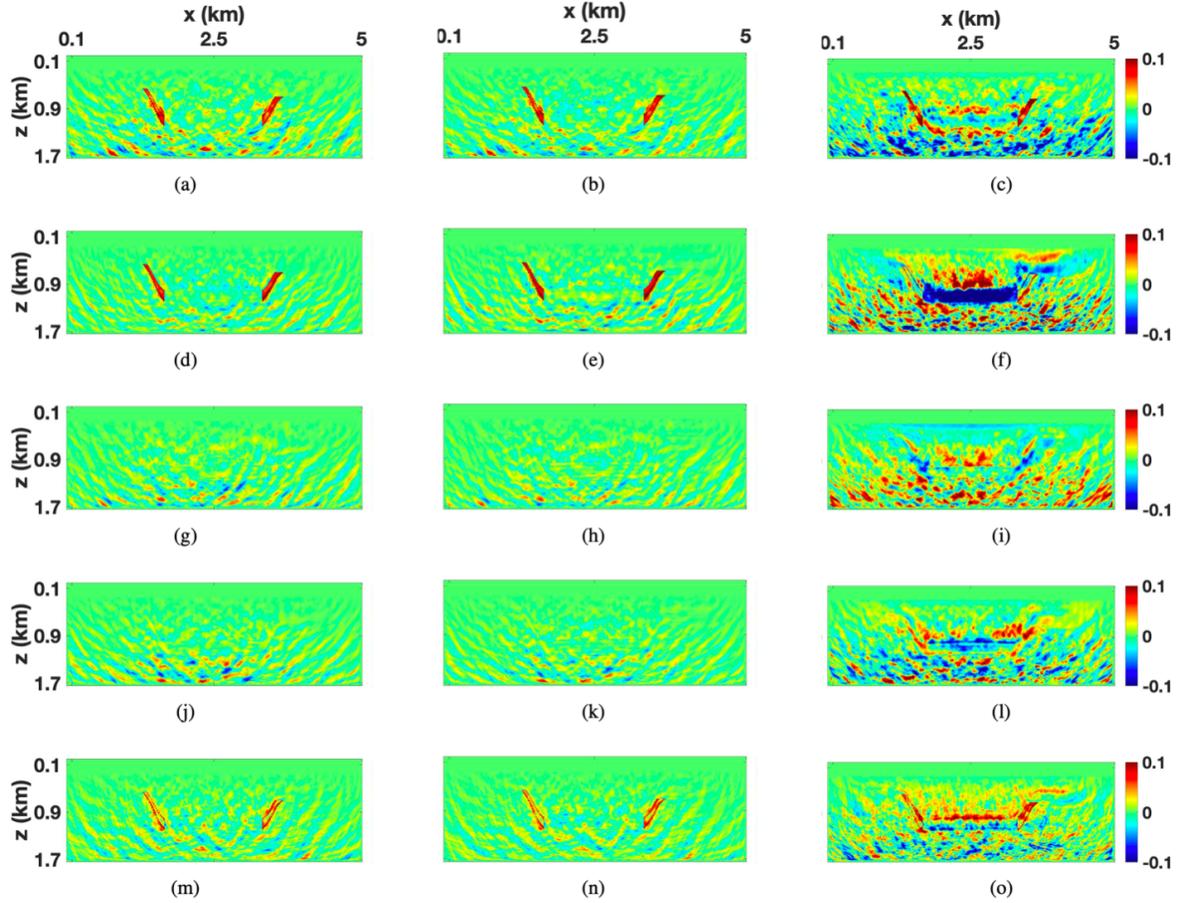


Figure 2.7 Parameters of the time-lapse model produced by three different strategies from the noisy multicomponent data ( $\text{SNR} = 8$ ). The parallel-difference method: (a)  $V_{P0}$ , (d)  $V_{S0}$ , (g)  $V_{\text{hor},P}$ , (j)  $V_{\text{nmo},P}$ , and (m)  $\rho$ . The sequential-difference method: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{\text{hor},P}$ , (k)  $V_{\text{nmo},P}$ , and (n)  $\rho$ . The double-difference method: (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{\text{hor},P}$ , (l)  $V_{\text{nmo},P}$ , and (o)  $\rho$ .

## 2.5 Conclusions

We extended elastic time-lapse FWI to VTI models which are typical, for example, for unconventional shale reservoirs. To resolve the temporal parameter variations, we employed three different time-lapse strategies previously proposed for isotropic media. Application of these methods to a VTI graben model showed that none of them significantly outperforms the other techniques in reconstructing the time-lapse variations of the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  (the P-wave NMO and horizontal velocities were held constant). For noise-free multicomponent data, the double-difference method produces the fewest artifacts outside the target zone because it focuses specifically on the time-lapse response. Still, a pronounced false anomaly in  $V_{\text{nmo,P}}$  in the output of that method shows its susceptibility to parameter trade-offs (in this case, between the velocities  $V_{S0}$  and  $V_{\text{nmo,P}}$ ).

When FWI operates with noise-free pressure data, the parallel-difference method yields the most accurate results for  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  due to the similarity between the errors for the baseline and monitor models. The parallel-difference and double-difference techniques generate comparable artifacts in the time-lapse variations of  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  (which are unchanged) caused by parameter trade-offs. The sequential-difference method, which is more sensitive to the errors in the inverted baseline parameters, does not perform well for the pressure data, which produce a low-resolution baseline model.

It is interesting that the false time-lapse anomaly in  $V_{\text{nmo,P}}$ , which has a large magnitude for the multicomponent data, is much less pronounced when using the pressure records. Although multicomponent data help increase parameter resolution compared with the pressure inversion, they apparently make the objective function more multimodal, which hinders the convergence toward its global minimum.

When the data are contaminated by realistic Gaussian noise, the double-difference method fails to reconstruct the time-lapse variations because of the low signal-to-noise ratio of the difference between the monitor and baseline data. The parallel-difference method generally performs better than the sequential-difference method in reconstructing the time-lapse variations inside the target area but produces more artifacts outside the “reservoir”.

## 2.6 Acknowledgments

We thank the members of the A(Anisotropy)-Team at the Center for Wave Phenomena (CWP) at Colorado School of Mines for useful discussions. We are grateful to Sagar Singh (CWP) for his kind help and suggestions. This work is supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.

## CHAPTER 3

### SOURCE-INDEPENDENT TIME-LAPSE FULL-WAVEFORM INVERSION FOR VTI MEDIA

A paper published\* in *Geophysics*

Yanhua Liu<sup>1,2</sup>, Ilya Tsvankin<sup>2</sup>

Time-lapse full-waveform inversion can provide high-resolution information about changes in the reservoir properties during hydrocarbon production and  $CO_2$  injection. However, the accuracy of the estimated source wavelet, which is critically important for time-lapse FWI, is often insufficient for field-data applications. The so-called “source-independent” FWI is designed to reduce the influence of the source wavelet on the inversion results. We incorporate the convolution-based source-independent technique into a time-lapse FWI algorithm for VTI (transversely isotropic with a vertical symmetry axis) media. The gradient of the modified FWI objective function is obtained from the adjoint-state method. The algorithm is tested on a model with a graben structure and the modified VTI Marmousi model using three time-lapse strategies (the parallel-difference, sequential-difference, and double-difference methods). The results confirm the ability of the developed methodology to reconstruct the localized time-lapse parameter variations even for a strongly distorted source wavelet. The algorithm remains robust in the presence of moderate noise in the input data but the accuracy of the estimated time-lapse changes depends on the model complexity.

#### 3.1 Introduction

Because seismic signatures are sensitive to the changes (e.g., in pressure and saturation) inside the reservoir, seismic time-lapse monitoring has been widely used for optimizing hydrocarbon production and  $CO_2$  injection (Lumley, 2010; Pevzner et al., 2017; Smith and Tsvankin, 2013). Full-waveform inversion (FWI) is an established tool for high-resolution velocity analysis and has been applied to reservoir characterization (Asnaashari et al., 2015; Li et al., 2021; Singh et al., 2018; Vigh et al., 2014; Zhang and Alkhalifah, 2020). FWI iteratively updates the medium parameters by minimizing the misfit between the observed and simulated seismic data.

Unlike conventional time-lapse methods, FWI operates with both the phase and amplitude of seismic waves, which can potentially increase the resolution of the inverted time-lapse parameter variations. However, FWI requires an accurate estimate of the source wavelet. Errors in the source signature (e.g., in

---

\*Reprinted with permission of *Geophysics*, 2022, 87, no. 1, R111-R122.

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Department of Geophysics, Colorado School of Mines, Golden, CO, USA

its shape, frequency, or amplitude) can hinder matching of the simulated and observed data and distort the inversion results (Luo et al., 2014; Pratt, 1999; Song et al., 1995; Warner et al., 2013; Yuan et al., 2014). This issue is particularly important for time-lapse FWI because of the commonly observed non-repeatability of the source signature between the baseline and monitor surveys. In particular, seasonal changes in the near surface may produce dramatic changes in the source signal (Jervis et al., 2018).

One way to reconstruct the source wavelet in FWI is to iteratively estimate it during the inversion along with the medium parameters (Pratt, 1999; Song et al., 1995; Xu et al., 2006). However, this method incurs a substantial computational cost, introduces additional trade-offs, and requires an accurate initial approximation for the source wavelet (Xu et al., 2006).

A more practical alternative is the so-called “source-independent” method designed to reduce the influence of the source signature on the inversion results. Deconvolution-based trace normalization can be employed in the frequency domain to remove information about the source wavelet from both the recorded and modeled data (Choi and Min, 2012; Lee and Kim, 2003; Seo et al., 2005; Zhou and Greenhalgh, 2003). Choi et al. (2005) propose to define the objective function in the frequency domain by multiplying the data with the corresponding reference trace. However, the resolution achieved by this deconvolution-based method strongly depends on the signal-to-noise ratio.

Choi and Alkhalifah (2011) define the convolution-based source-independent FWI objective function in the time domain. The first step is to choose a reference trace from both the observed and simulated data sets. Then the simulated data are convolved with the reference trace from the observed data, and the observed data are convolved with the reference trace from the simulated data. The modified FWI objective function is designed to minimize the difference between these two convolved data sets. One issue with this approach is that the convolution and cross-correlation operations tend to generate additional noise, which can be suppressed by applying a time window to the reference traces beforehand (Zhang et al., 2016). Wang et al. (2018) employ the source-independent objective function to estimate the location and excitation times. Bai and Tsvankin (2019) extend the time-domain source-independent waveform inversion to attenuation estimation using reflection or transmission data from VTI media. Wang et al. (2020) use the source-independent FWI to mitigate cycle skipping caused by the unknown source signature of passive events and apply their algorithm to a field data set.

Liu and Tsvankin (2021) develop a time-lapse FWI algorithm for VTI media and test it on synthetic data using three common time-lapse strategies (Asnaashari et al., 2015): the parallel-difference (Plessix et al., 2010), sequential-difference, and double-difference techniques (Denli and Huang, 2009; Watanabe et al., 2004). The parallel-difference method uses the same initial model for the baseline and monitor FWI, whereas the sequential-difference method inverts the baseline data to build the initial model for the

monitor inversion. The double-difference technique directly estimates the time-lapse parameter variations from the difference between the monitor and baseline data sets.

As mentioned above, time-lapse FWI is particularly sensitive to the accuracy of the source wavelet. Note that even moderate wavelet-related distortions in the inversion of the baseline and monitor data can lead to large percentage errors in the estimated time-lapse parameter variations. Most existing time-lapse processing algorithms either assume the source wavelet to be known or iteratively estimate it during the inversion.

Here, we extend the convolution-based source-independent method to time-lapse FWI of reflection data from VTI media. This extension involves adaptation of the source-independent technique for different time-lapse strategies, in particular for the double-difference method.

We begin by discussing the methodology of the source-independent FWI technique and outline its application to time-lapse seismic. Analysis of the corresponding objective function is followed by a brief review of the three above-mentioned time-lapse strategies. The derivation of the inversion gradients in terms of the VTI parameters is given in the appendix A. The proposed algorithm is applied to the reconstruction of the time-lapse parameter variations in a VTI graben model and in a modified version of the VTI Marmousi model. The inversion results are used to evaluate the robustness of the source-independent algorithm implemented with different time-lapse strategies. We also discuss the influence of the reference trace, time window, and noise on the inverted time-lapse variations.

### 3.2 Methodology of source-independent time-lapse FWI

FWI of time-lapse seismic data generally involves the inversion of baseline and monitor surveys. The baseline survey is often acquired before hydrocarbon production or  $CO_2$  injection, and the monitor survey during or after production/injection. Typically, FWI is first applied to the baseline data to obtain the baseline model. Then the monitor survey is processed using different approaches according to the chosen time-lapse strategy. The subtraction of the inverted baseline model from the monitor model yields the time-lapse parameter variations.

Conventional algorithms use the  $L_2$ -norm objective function (e.g., Tarantola, 1984) for the inversion of the baseline (subscript b) data:

$$S_b(\mathbf{m}_b) = \frac{1}{2} \left\| \left[ \mathbf{d}_b^{\text{sim}}(\mathbf{m}_b) - \mathbf{d}_b^{\text{obs}} \right] \right\|^2 = \frac{1}{2} \left\| \left[ \mathbf{G}^{\text{sim}} * \mathbf{s}^{\text{sim}} - \mathbf{G}^{\text{obs}} * \mathbf{s}^{\text{obs}} \right] \right\|^2, \quad (3.1)$$

where  $\mathbf{d}_b^{\text{sim}}$  is the data simulated for the baseline model  $\mathbf{m}_b$ ,  $\mathbf{d}_b^{\text{obs}}$  is the observed baseline data, the symbol “ $*$ ” denotes convolution, and  $\mathbf{G}^{\text{sim}}$  and  $\mathbf{G}^{\text{obs}}$  are the simulated and actual Green’s functions, respectively.

### 3.2.1 Source-independent methodology

To reduce the influence of the source wavelet, Choi and Alkhalifah (2011) and Zhang et al. (2016) propose the following convolution-based source-independent objective function in the time domain:

$$\begin{aligned}
S_b(\mathbf{m}_b) &= \frac{1}{2} \left\| \left[ \mathbf{d}_b^{\text{sim}}(\mathbf{m}_b) * (W \mathbf{d}_{\text{ref}}^{\text{obs}}) - \mathbf{d}_b^{\text{obs}} * (W \mathbf{d}_{\text{ref}}^{\text{sim}}) \right] \right\|^2 \\
&= \frac{1}{2} \left\| \left[ \mathbf{G}^{\text{sim}} * \mathbf{s}^{\text{sim}} * W \mathbf{G}_{\text{ref}}^{\text{obs}} * \mathbf{s}^{\text{obs}} - \mathbf{G}^{\text{obs}} * \mathbf{s}^{\text{obs}} * W \mathbf{G}_{\text{ref}}^{\text{sim}} * \mathbf{s}^{\text{sim}} \right] \right\|^2 \\
&= \frac{1}{2} \left\| \left[ \tilde{\mathbf{G}}^{\text{s}} * \tilde{\mathbf{s}}^{\text{c}} - \tilde{\mathbf{G}}^{\text{o}} * \tilde{\mathbf{s}}^{\text{c}} \right] \right\|^2,
\end{aligned} \tag{3.2}$$

where  $W = W(t)$  is a chosen time window, and  $\mathbf{d}_{\text{ref}}$  denotes the reference trace.  $\mathbf{G}_{\text{ref}}^{\text{obs}}$  and  $\mathbf{G}_{\text{ref}}^{\text{sim}}$  are the Green's functions for the reference traces,  $\tilde{\mathbf{G}}^{\text{s}}$  and  $\tilde{\mathbf{G}}^{\text{o}}$  are the Green's functions for the new convolution-based simulated and observed data, and  $\tilde{\mathbf{s}}^{\text{c}}$  is the new source wavelet. The expressions for  $W(t)$ ,  $\tilde{\mathbf{G}}^{\text{s}}$ ,  $\tilde{\mathbf{G}}^{\text{o}}$ , and  $\tilde{\mathbf{s}}^{\text{c}}$  are:

$$\text{Time window : } W(t) = \begin{cases} 1 & t_l \leq t \leq t_h; \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

$$\text{New Green's functions : } \begin{cases} \tilde{\mathbf{G}}^{\text{s}} = \mathbf{G}^{\text{sim}} * W \mathbf{G}_{\text{ref}}^{\text{obs}}; \\ \tilde{\mathbf{G}}^{\text{o}} = \mathbf{G}^{\text{obs}} * W \mathbf{G}_{\text{ref}}^{\text{sim}}. \end{cases} \tag{3.4}$$

$$\text{New source wavelet : } \tilde{\mathbf{s}}^{\text{c}} = \mathbf{s}^{\text{sim}} * \mathbf{s}^{\text{obs}}. \tag{3.5}$$

$t_l$  and  $t_h$  are the boundaries of the time window applied to the reference trace.

The new simulated and observed data in equation 3.2 could be expressed as the convolutions of the new Green's functions ( $\tilde{\mathbf{G}}^{\text{s}}$  and  $\tilde{\mathbf{G}}^{\text{o}}$ ) and the source wavelet ( $\tilde{\mathbf{s}}^{\text{c}}$ ). Because the wavelet  $\tilde{\mathbf{s}}^{\text{c}}$  is the same for the simulated and observed data, the deviation of the estimated source signature from the actual wavelet is theoretically eliminated from the modified objective function.

The time window  $W(t)$  is used to mitigate the artifacts caused by the convolution and cross-correlation operations. To save computing time and ensure efficient noise suppression, the time window should be as short as possible but has to contain sufficient information about the source wavelet. Zhang et al. (2016) suggest that the lower cut-off time ( $t_l$ ) should be before the first arrival, and the window should contain at least one full waveform, such as the P-wave direct arrival. Here, we define the time window using the Butterworth filter:

$$W(t) = \frac{1}{1 + \left(\frac{t-t_l}{t_h}\right)^{2n}} = \frac{1}{1 + \left(\frac{t-t_l}{t_h}\right)^{30}}, \tag{3.6}$$

where  $n$  is the order of the filter ( $n = 15$  in our examples).

### 3.2.2 Implementation of time-lapse FWI

We parameterize VTI media by the velocities  $V_{P0}$  (P-wave vertical velocity),  $V_{S0}$  (S-wave vertical velocity),  $V_{\text{hor,P}}$  (P-wave horizontal velocity),  $V_{\text{nmo,P}}$  (P-wave normal-moveout velocity from a horizontal reflector), and density  $\rho$  (Alkhalifah and Plessix, 2014; Kamath and Tsvankin, 2016; Singh et al., 2020; Tsvankin, 2012; Zhang et al., 2018). Multicomponent data are simulated by solving the 2D wave equation for elastic, arbitrarily heterogeneous VTI media with a fourth-order finite-difference algorithm (Singh et al., 2020). All five VTI parameters are updated simultaneously using the gradients derived in Appendix A.

The time-lapse strategies employed here use the same objective function for the baseline inversion but differ in handling the monitor survey (or the time-lapse data difference). In the parallel-difference approach (Plessix et al., 2010), the baseline and monitor inversions are performed independently but with the same initial model. The time-lapse model produced by this method can remain sufficiently accurate when the errors in the inverted baseline and monitor models are similar. The sequential-difference strategy (Asnaashari et al., 2012) uses the inversion of the baseline data to build the initial model for FWI of the monitor survey. This facilitates the convergence of the monitor inversion due to the similarity between the baseline and monitor surveys.

In the double-difference method (Denli and Huang, 2009; Waldhauser and Ellsworth, 2020), the monitor inversion operates directly on the difference between the monitor and baseline data ( $\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}$ ). The monitor model is obtained by minimizing the difference  $\Delta \mathbf{d}$  between the simulated monitor data  $\mathbf{d}_m^{\text{sim}}$  and the “composite” data  $\mathbf{d}_{\text{com}}$ , starting from the inverted baseline model:

$$\Delta \mathbf{d} = (\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}) - (\mathbf{d}_m^{\text{sim}} - \mathbf{d}_b^{\text{sim}}) = \mathbf{d}_{\text{com}} - \mathbf{d}_m^{\text{sim}}, \quad (3.7)$$

$$\mathbf{d}_{\text{com}} = \mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}} + \mathbf{d}_b^{\text{sim}}, \quad (3.8)$$

where  $\mathbf{d}_b^{\text{sim}}$  is the data simulated for the inverted baseline model.

Because the double-difference method operates on the data difference, the geometries of the baseline and monitor surveys should be similar and the amplitude of the data difference should be above the noise level.

### 3.3 Synthetic Examples

The developed source-independent time-lapse FWI algorithm is tested on a VTI graben model and the modified VTI Marmousi model. The synthetic data for both models are generated with the Ricker wavelet (Figure 3.1a). The source-independent algorithm is applied for two distorted wavelets (Figure 3.1b and Figure 3.1c) used in the inversion. FWI is implemented for multicomponent surface data (vertical and horizontal velocities) using the multiscale approach with four frequency bands starting at 2 Hz (Singh et al., 2020).

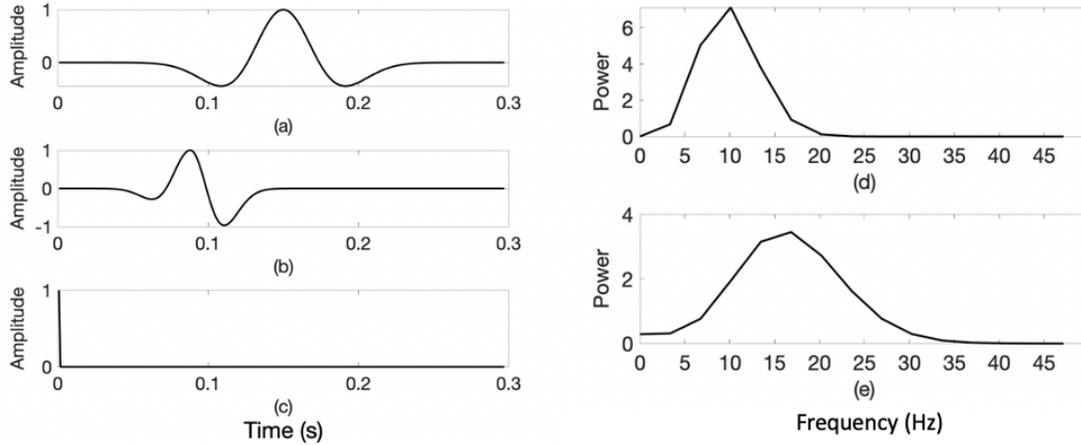


Figure 3.1 Source wavelets used in the synthetic examples: (a) the Ricker wavelet with a central frequency of 10 Hz (actual wavelet), (b) the distorted “Ricker” wavelet with a central frequency of 17 Hz (Wavelet 1), and (c) the spike (Wavelet 2). The frequency spectra of (d) the actual wavelet [see plot (a)] and (e) Wavelet 1 [see plot (b)].

### 3.3.1 VTI graben model

The elastic wavefield is excited by 58 shots (point explosions) placed with a constant increment (80 m) along a horizontal line at a depth of 40 m (Figure 3.2a). We employ 400 receivers evenly distributed with an increment of 10 m along the horizontal line at a depth of 100 m (Figure 3.2a). The medium parameters for the monitor survey (Liu and Tsvankin, 2021) are obtained by reducing the baseline vertical velocities  $V_{P0}$ ,  $V_{S0}$ , and density  $\rho$  in the target area (i.e., in the dipping layer segments) by approximately 10% (Figure 3.2). The initial baseline models (Figure 3.2c, Figure 3.2f, Figure 3.2i, Figure 3.2l, and Figure 3.2o) are computed by Gaussian smoothing of the actual parameter distributions with a standard deviation of 10. The benchmark time-lapse results (Figure 3.3) are obtained by performing FWI with the actual wavelet following the parallel-difference strategy. The time-lapse parameter variations inside the “reservoir” are well estimated and there are no significant artifacts outside the target zone. The artifacts near the boundaries of the graben structure are caused primarily by edge (smoothing) effects (Schmidt, 2005; Zhang and Zhang, 2012) in the  $L_2$ -norm objective function.

Next, the conventional FWI algorithm is applied to the baseline data using a distorted source wavelet (Wavelet 1). The incorrect source signature completely corrupts the inversion results, and even the graben structure itself is barely visible (Figure 3.4). It is clear that conventional FWI is unable to reconstruct the time-lapse parameter changes for the distorted wavelet.

Then we apply the source-independent FWI algorithm using Wavelet 1. After estimating the baseline model, three time-lapse methods are used to reconstruct the parameter variations (note that the P-wave

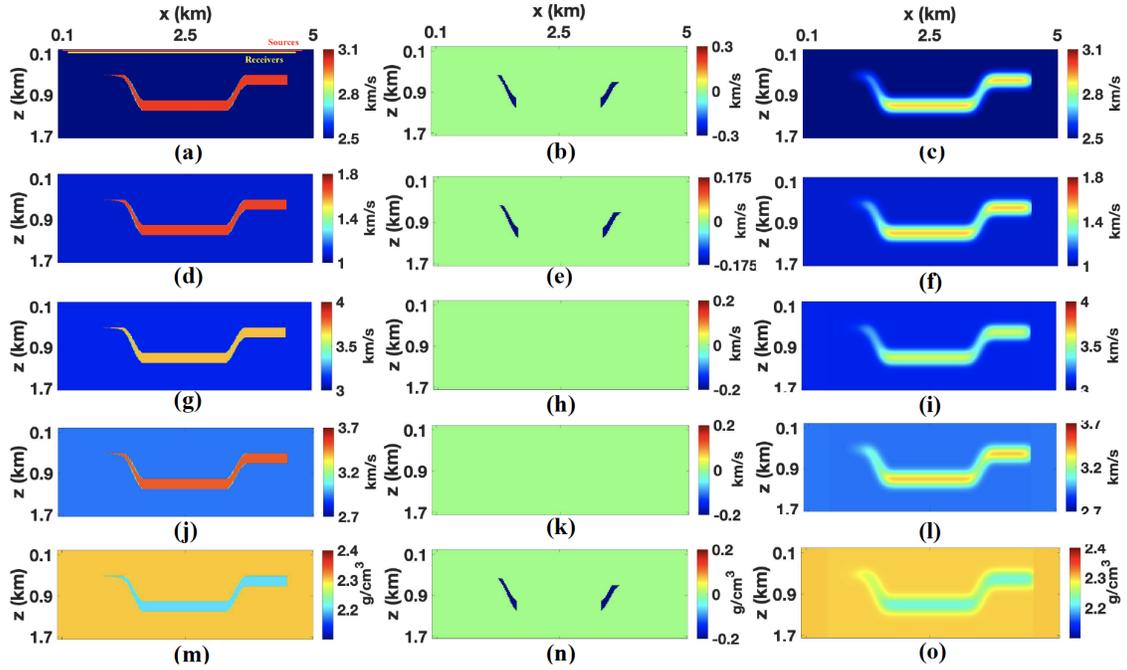


Figure 3.2 Parameters of the baseline graben model with a grid size of  $10 \times 10$  m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (d) the S-wave vertical velocity ( $V_{S0}$ ), (g) the P-wave horizontal velocity ( $V_{hor,P}$ ), (j) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ), and (m) the density ( $\rho$ ). The actual time-lapse differences for (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The initial baseline model of: (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

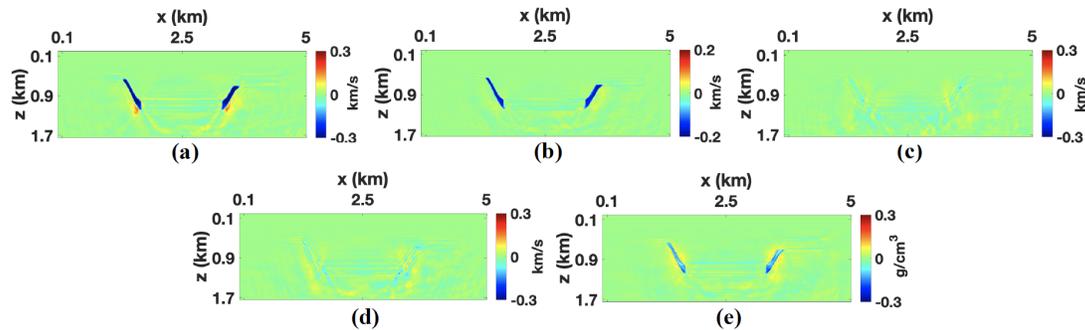


Figure 3.3 Time-lapse parameter variations obtained by conventional FWI with the actual wavelet using the parallel-difference method (benchmark results): (a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $V_{hor,P}$ , (d)  $V_{nmo,P}$ , and (e)  $\rho$ .

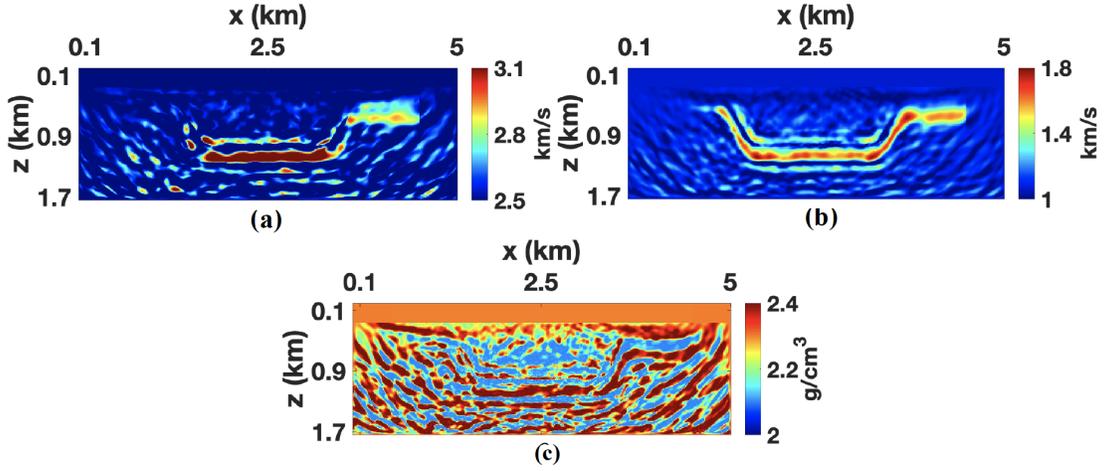


Figure 3.4 Baseline models estimated by conventional FWI using Wavelet 1: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ .

horizontal and NMO velocities are held constant). All five VTI parameters are updated simultaneously during the inversion. The selected reference traces for both the vertical and horizontal particle velocities correspond to the receiver located closest to the source, which ensures high fidelity of the source signal.

### 3.3.1.1 Parallel- and sequential-difference methods

The time-lapse changes estimated by the source-independent FWI using the parallel- and sequential-difference strategies are sufficiently close to the benchmark results (Figure 3.2), although the amplitude of the time-lapse anomalies is slightly underestimated. Also, our method generates some false time-lapse anomalies in the velocities  $V_{hor,P}$  and  $V_{nmo,P}$  because of the parameter trade-offs (Figure 3.5g, Figure 3.5h, Figure 3.5j, Figure 3.5k). The performance of the parallel-difference (Figure 3.5a, Figure 3.5d, and Figure 3.5m) and sequential-difference (Figure 3.5b, Figure 3.5e, and Figure 3.5n) methods in reconstructing the time-lapse variations in  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  and suppressing the artifacts is comparable. The robustness of our algorithm applied with the parallel-difference method is further evaluated in Figure 3.6 for the spike wavelet that represents an extreme shape distortion. Still, the time-lapse changes of the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  are reconstructed with sufficient resolution similar to those for Wavelet 1 (Figure 3.5a, Figure 3.5d, and Figure 3.5m). Clearly, our source-independent algorithm can handle even extreme frequency and shape distortions in the source wavelet for high-quality multicomponent data.

### 3.3.1.2 Double-difference method

Although the double-difference method does not produce significant false anomalies, the reconstructed temporal variations in  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  have the wrong sign (Figure 3.5c, Figure 3.5f, and Figure 3.5o). In

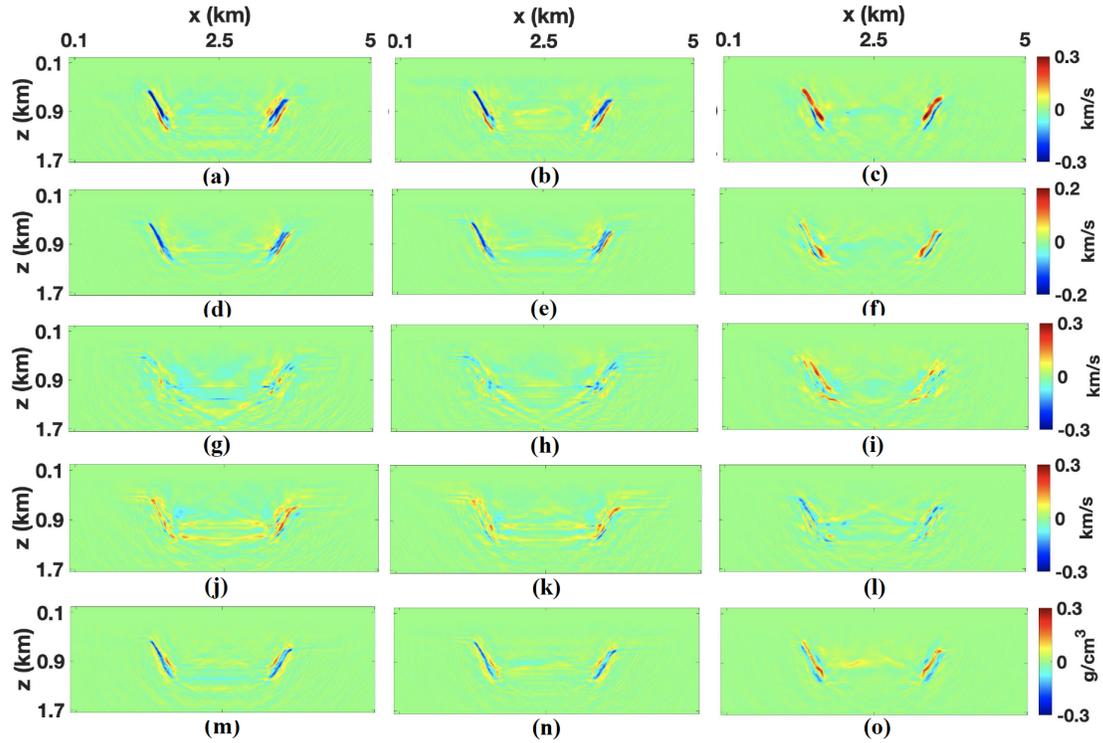


Figure 3.5 Time-lapse parameter variations reconstructed by the source-independent algorithm with Wavelet 1. The parallel-difference method: (a)  $V_{P0}$ , (d)  $V_{S0}$ , (g)  $V_{hor,P}$ , (j)  $V_{nmo,P}$ , and (m)  $\rho$ . The sequential-difference method: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The double-difference method: (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

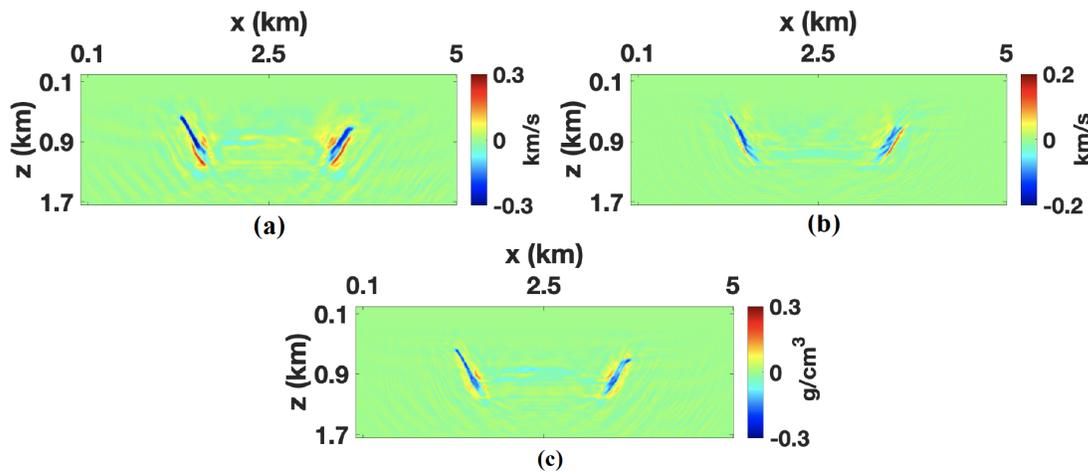


Figure 3.6 Time-lapse parameter variations estimated by the source-independent algorithm with the spike wavelet (Wavelet 2) using the parallel-difference method: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ .

contrast to the other two strategies, this method operates with the “composite” data ( $\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}} + \mathbf{d}_b^{\text{sim}}$ ) generated by adding the wavefield simulated for the inverted baseline model to the actual time-lapse data difference (equations 3.7 and 3.8). The frequency distortion of Wavelet 1 leads to a phase mismatch between the simulated baseline data  $\mathbf{d}_b^{\text{sim}}$  and the observed wavefield. The resulting degradation of the “composite” data set prevents the source-independent algorithm from producing a sufficiently accurate monitor model.

To verify this conjecture, in Figure 3.7b we use a wavelet which has the shape of Wavelet 1 but the correct central frequency (the same as for the actual wavelet) to simulate the baseline seismogram and generate the “composite” data. The monitor inversion is still performed using the source-independent algorithm with Wavelet 1. In this case, the reconstructed parameter variations in the target area have the correct sign (Figure 3.7b) and are sufficiently close to the benchmark results. Evidently, the double-difference method has to be applied with a wavelet that has a sufficiently accurate frequency to properly estimate the temporal parameter variations. If the data for the inverted baseline model are simulated using the actual wavelet (FWI is still performed with Wavelet 1), the double-difference method (Figure 3.7c) produces the time-lapse changes in the velocity  $V_{P0}$  with higher resolution than the other two methods (Figure 3.5a and Figure 3.5b).

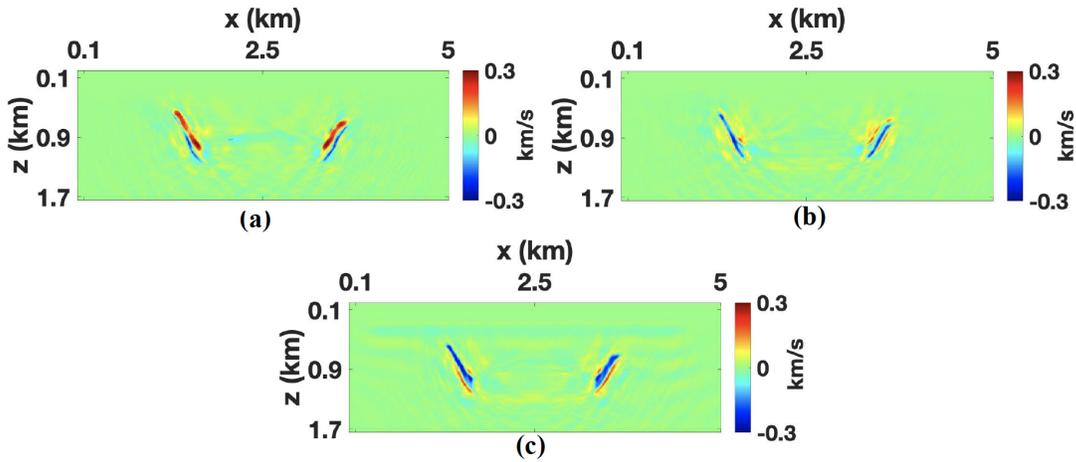


Figure 3.7 Time-lapse variations of the velocity  $V_{P0}$  reconstructed by the source-independent algorithm using the double-difference method. FWI is performed with Wavelet 1, whereas the data for the inverted baseline model ( $\mathbf{d}_b^{\text{sim}}$ ) are generated with: (a) Wavelet 1, (b) a signal that has the shape of Wavelet 1 but correct frequency, (c) the actual wavelet.

One possible solution to at least partially resolve the issue with errors in the simulated baseline data generated with a distorted wavelet is to apply dynamic seismic warping. This method has been proposed for time-shift estimation in time-lapse processing by matching traces from the baseline and monitor surveys

using the criterion of minimal dissimilarity (Hale, 2013a; Holschuh et al., 2014; Li et al., 2019; Rickett et al., 2007; Venstad, 2014). We will investigate this method as part of our future research in anisotropic time-lapse FWI.

### 3.3.1.3 Influence of noise

Next, the data are contaminated with Gaussian noise that has the signal-to-noise ratio equal to 16, a realistic value for field data. We employ only the parallel-difference method due to its general robustness for noisy data (Liu and Tsvankin, 2021). Here we show only the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  reconstructed by conventional FWI using the actual source wavelet (benchmark results) and by the source-independent algorithm using the two distorted wavelets.

The baseline models (especially  $V_{P0}$ ) inverted by the source-independent algorithm using Wavelet 1 are somewhat distorted, likely due to the noise amplification in the convolution and cross-correlation operations. However, the source-independent algorithm (Figure 3.8d and Figure 3.8e) surprisingly produces a slightly more accurate reconstruction of the baseline velocities in the “reservoir” compared to conventional FWI applied with the actual wavelet (Figure 3.8a, Figure 3.8b). Still, the time-lapse variations of  $V_{P0}$  estimated by our algorithm (Figure 3.8j) have a somewhat lower resolution than the benchmark section (Figure 3.8g). Nevertheless, the changes in the velocity  $V_{S0}$  reconstructed by the source-independent FWI (Figure 3.8k) are close to those obtained with the actual wavelet (Figure 3.8h). There is no significant difference between the accuracy of the time-lapse results produced with Wavelet 1 and the spike wavelet (Wavelet 2; not shown). This test indicates that our algorithm not only can handle significant wavelet distortions, but also reconstruct the temporal parameter variations for data contaminated with substantial noise.

### 3.3.1.4 Influence of the reference trace and time window

The objective function for the source-independent FWI (equation 3.2) is sensitive to the choice of the reference trace and time window. Our results show that employing a reference trace that represents the average of several near-offset traces produces smaller errors compared to using the zero-offset trace; the worst results are obtained using a far-offset trace. This is not surprising because of the higher data fidelity close to the source (Choi and Alkhalifah, 2011; Xu et al., 2006; Zhang et al., 2016). An arrival-time (normal-moveout) correction needs to be applied to the near-offset traces to ensure their in-phase stacking in generating the averaged reference trace. Although NMO stretch can lead to a loss of resolution, it is usually negligible for small offset-to-depth ratios.

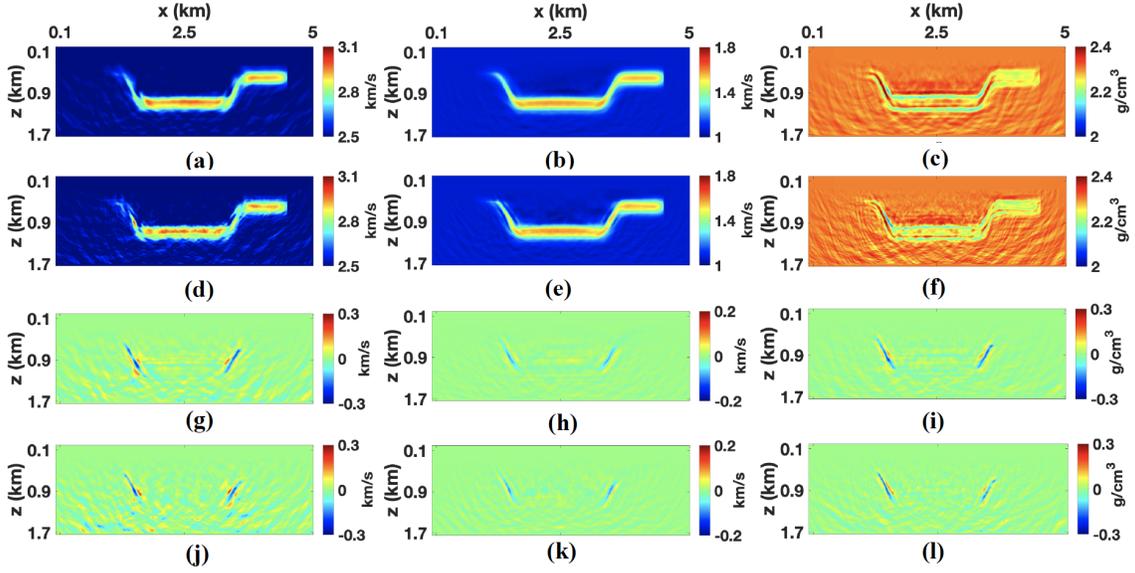


Figure 3.8 Inverted baseline models and time-lapse variations obtained from noisy data (signal-to-noise ratio is 16). Inverted baseline models obtained by conventional FWI with the actual wavelet [(a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $\rho$ ] and by the source-independent FWI with Wavelet 1 [(d)  $V_{P0}$ , (e)  $V_{S0}$ , (f)  $\rho$ ]. Time-lapse variations obtained by conventional FWI with the actual wavelet [(g)  $V_{P0}$ , (h)  $V_{S0}$ , (i)  $\rho$ ] and by the source-independent FWI with Wavelet 1 [(j)  $V_{P0}$ , (k)  $V_{S0}$ , (l)  $\rho$ ].

As mentioned above, time windowing mitigates the artifacts (noise) caused by the convolution and cross-correlation operations. However, the reference trace may not contain enough information about the source signal if the window is too narrow. Figure 3.9 shows the velocity  $V_{P0}$  reconstructed from the baseline data for the graben model. The source-independent FWI is applied with the spike wavelet and two different time windows. If the window does not capture the entire first arrival (Figure 3.9c), the graben structure is poorly resolved because the reference trace does not provide sufficient information about the wavelet (Figure 3.9a). The section in Figure 3.9b is reconstructed with higher resolution because the time window is closer to optimal, as it includes only the first arrival (Figure 3.9d). This observation agrees with the conclusions of Zhang et al. (2016). Therefore, the time window in our tests is chosen to contain the full waveform of the first arrival (such as the direct P-wave).

### 3.3.2 VTI Marmousi model

Next, we apply our algorithm to the modified VTI Marmousi model (Figure 3.10), with data simulated using OBC-style acquisition. The source/receiver geometry is the same as that for the graben model but the receivers are at a depth of 230 m, which corresponds to the water bottom. The time-lapse parameters for the monitor survey are obtained by reducing the baseline vertical velocities  $V_{P0}$ ,  $V_{S0}$  and density  $\rho$  in the target area (between the depths of 870 m and 1100 m) by 15% (Figure 3.10c, Figure 3.10f, and

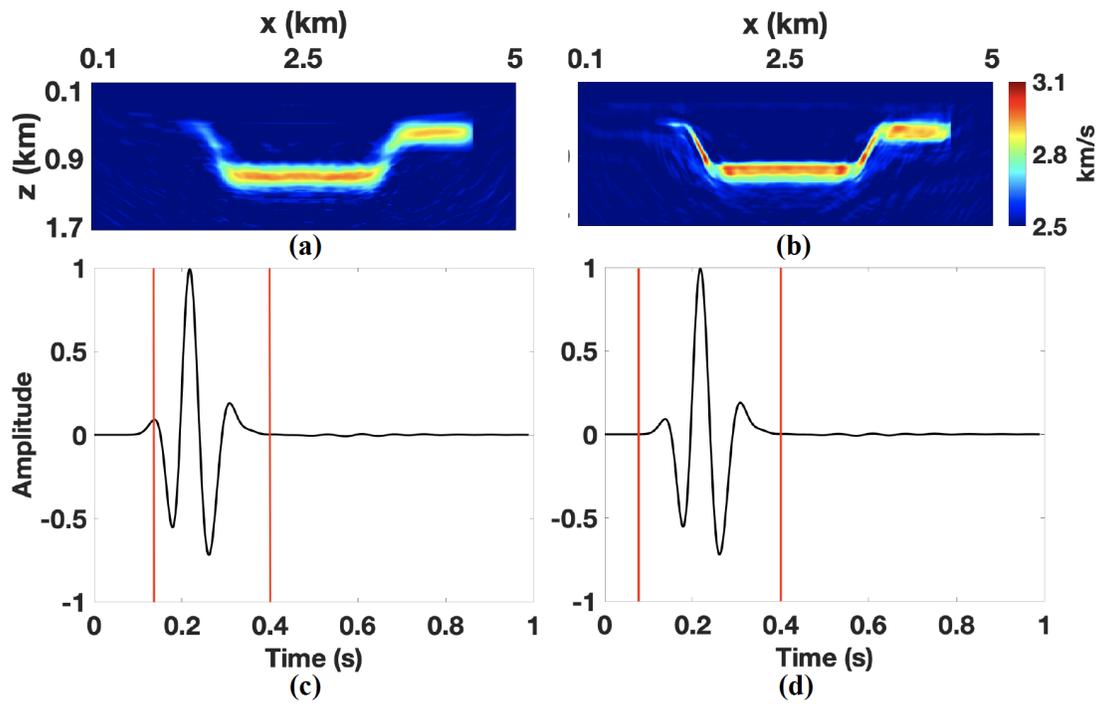


Figure 3.9 Velocity  $V_{P0}$  estimated by the source-independent algorithm from the baseline data for the graben model using the spike wavelet (Wavelet 2) and two time windows: (a) window 1 [plot(c)] and (b) window 2 [plot(d)].

Figure 3.10o). The reference trace used in the source-independent algorithm is recorded 70 m away from the source.

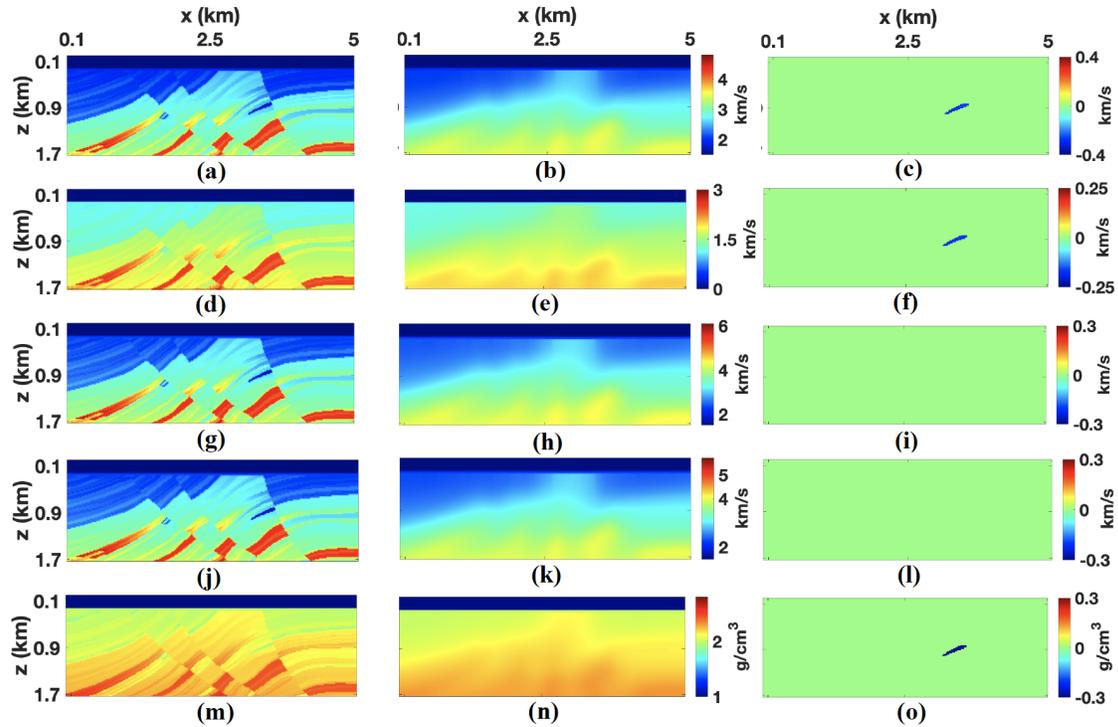


Figure 3.10 Baseline parameters of the modified VTI Marmousi model with a grid size of  $10 \times 10$  m: (a)  $V_{P0}$ , (d)  $V_{S0}$ , (g)  $V_{hor,P}$ , (j)  $V_{nmo,P}$ , and (m)  $\rho$ . The initial baseline model of: (b)  $V_{P0}$ , (e)  $V_{S0}$ , (h)  $V_{hor,P}$ , (k)  $V_{nmo,P}$ , and (n)  $\rho$ . The actual time-lapse differences for (c)  $V_{P0}$ , (f)  $V_{S0}$ , (i)  $V_{hor,P}$ , (l)  $V_{nmo,P}$ , and (o)  $\rho$ .

As before, elastic FWI is applied to the vertical and horizontal particle-velocity components. After generating the observed data using the actual wavelet, the benchmark parameter changes are obtained by the conventional FWI algorithm using the parallel-difference method with the actual wavelet (Figure 3.11). Despite the model complexity and small thickness of the “reservoir,” the time-lapse variations are reconstructed with sufficiently high resolution.

Next, we perform conventional FWI of the baseline data using Wavelet 1. However, the inversion becomes unstable and does not converge, likely due to the increased sensitivity to wavelet distortions for this structurally complex model. Then we employ a wavelet with the same shape distortion as Wavelet 1 but correct central frequency. Still, the geologic structure is not properly resolved in the inverted baseline models (Figure 3.12), and it is obvious that conventional FWI is incapable of reconstructing the time-lapse changes with the inaccurate wavelet.

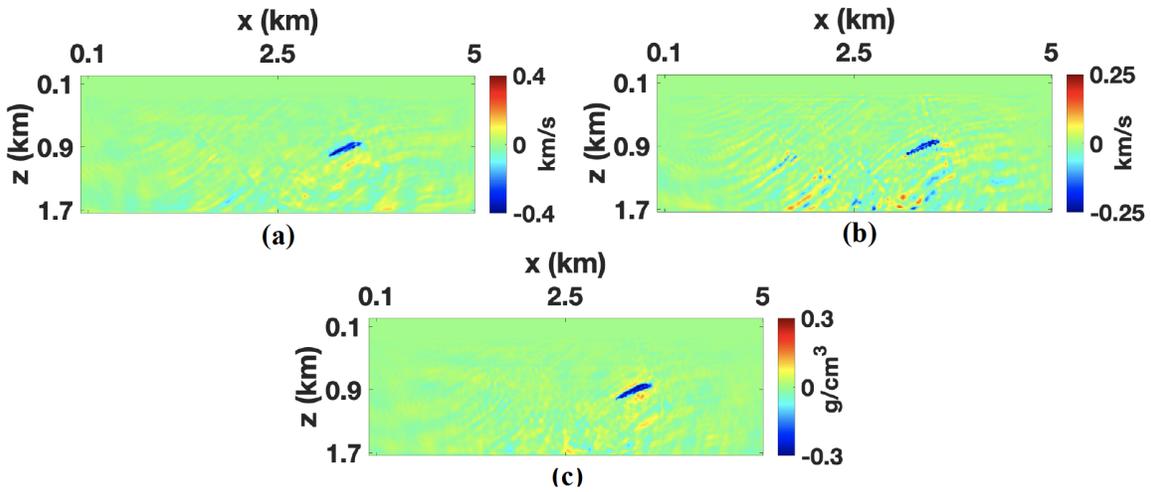


Figure 3.11 Time-lapse variations for the Marmousi model obtained by conventional FWI using the parallel-difference method with the actual wavelet (benchmark results): (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ .

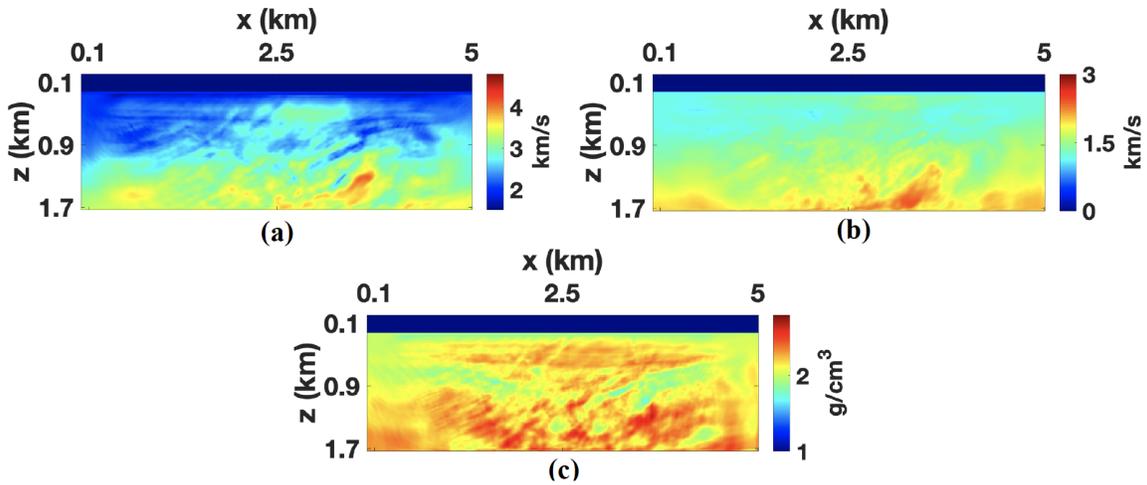


Figure 3.12 Baseline Marmousi models produced by conventional FWI using a wavelet with only a shape distortion: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ .

Because of the problems with the double-difference method discussed above, the source-independent algorithm is applied only with the other two strategies using the same distorted wavelet (Wavelet 1). In contrast to the graben model, the sequential-difference method (Figure 3.13b, Figure 3.13d, and Figure 3.13f) reconstructs the temporal variations with higher resolution and fewer artifacts than the parallel-difference method (Figure 3.13a, Figure 3.13c, and Figure 3.13e).

Clearly, the performance of these methods varies with the model complexity. Because FWI is implemented using a local optimization technique, the updating algorithm generally converges toward the minimum of the objective function closest to the initial model. The convolution operations in the source-independent method make the shape of the objective function more complex, and the pronounced heterogeneity of the Marmousi model increases the inversion nonlinearity. Starting with the inverted baseline model (which is relatively close to the monitor model) in the sequential-difference method improves the convergence toward the monitor model, which yields more accurate time-lapse variations (Figure 3.13b, Figure 3.13d, and Figure 3.13f).

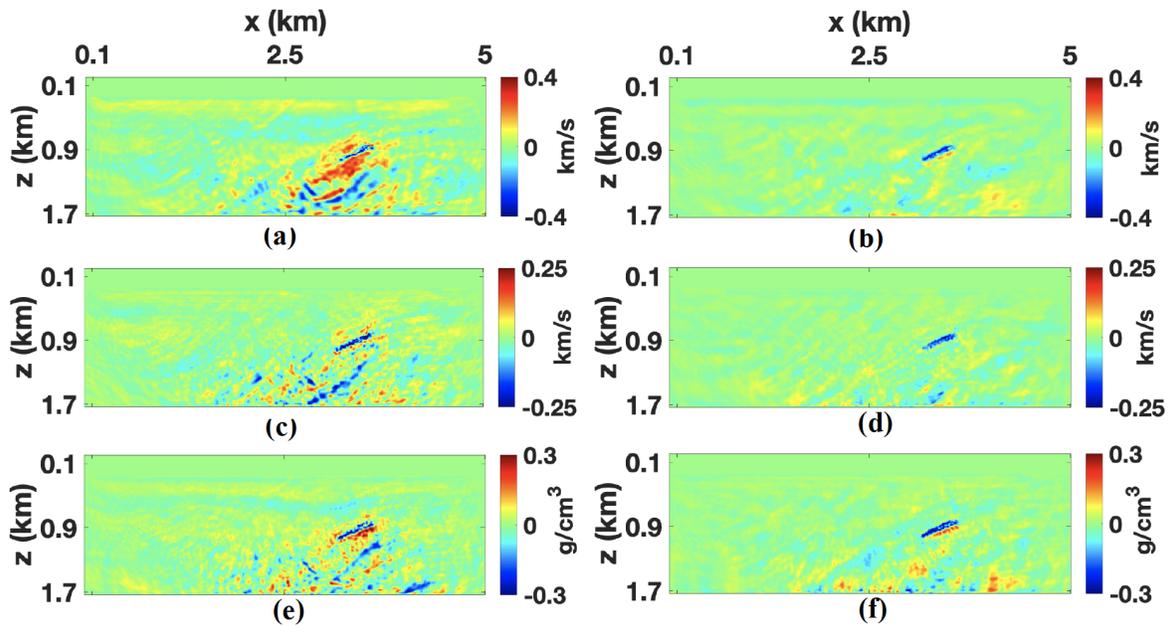


Figure 3.13 Time-lapse parameter variations for the Marmousi model obtained by the source-independent algorithm with Wavelet 1. The parallel-difference method: (a)  $V_{P0}$ , (c)  $V_{S0}$ , and (e)  $\rho$ . The sequential-difference method: (b)  $V_{P0}$ , (d)  $V_{S0}$ , and (f)  $\rho$ .

Comparison of the output of the source-independent FWI applied with the sequential-difference method (Figure 3.13b, d, f) and the benchmark results (Figure 3.11) shows that despite the combination of the significantly distorted wavelet and pronounced heterogeneity, our algorithm reconstructs the time-lapse

variations with acceptable resolution. The somewhat lower accuracy of the source-independent technique (as well as additional artifacts) for the Marmousi model as compared to the graben structure is likely due to the increased medium complexity. In particular, multiple reflections, which could be captured by the time window used in the convolution operations, may hinder the convergence of the algorithm.

### **3.4 Conclusions**

We implemented a “source-independent” time-lapse FWI algorithm for VTI media in the time domain. The modified FWI objective function is designed to mitigate the dependence of the inverted parameters on the accuracy of the source wavelet by employing two additional data sets obtained by convolution operations. The reference traces used in the convolutions should be recorded near the source to ensure high fidelity of the source signature. The convolution time window is chosen to include the entire first arrival (e.g., the direct P-wave) to ensure that the reference trace contains sufficient source information.

The synthetic examples demonstrate that the developed source-independent algorithm can accurately reconstruct the time-lapse variations even for significantly distorted source signals (such as the spike wavelet). The testing also confirms the ability of the method to deal with noisy data and strongly heterogeneous media, although the results for the Marmousi model are somewhat inferior to those for the simpler graben structure. The source-independent technique is particularly important in time-lapse processing because it can handle the non-repeatability of the source wavelet.

### **3.5 Acknowledgments**

We thank the members of the Anisotropy-Team at the Center for Wave Phenomena (CWP) at Colorado School of Mines for useful discussions. This work is supported by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.

## CHAPTER 4

### TIME-LAPSE FULL-WAVEFORM INVERSION FOR ELASTIC TTI MEDIA

Submitted to *Geophysics*

Yanhua Liu<sup>1,2</sup>, Ilya Tsvankin<sup>2</sup>

Time-lapse (4D) full-waveform inversion (FWI) of seismic data can help estimate the subsurface changes due to hydrocarbon production and CO<sub>2</sub> injection. Previously, we have developed a 4D FWI algorithm for VTI (transversely isotropic with a vertical symmetry axis) media. However, the VTI algorithm fails to accurately reconstruct the 4D variations in the presence of dipping anisotropic layers that often cause a tilt of the symmetry axis. Here, we extend the time-lapse FWI methodology to 2D TI media with a tilted symmetry axis (TTI). The symmetry axis is assumed to be orthogonal to the reflectors, so its orientation can be estimated from migrated depth images. The symmetry-axis tilt is incorporated into the modeling code and computation of the FWI gradients by rotating the stiffness tensor using the Bond transformation. The proposed algorithm is tested on the BP TTI model using three different time-lapse strategies. If the 4D data are repeatable, the parameter changes are reconstructed with sufficient accuracy in the presence of moderate noise. We also incorporate the “source-independent” FWI technique to mitigate the influence of errors in the estimated source wavelet and address the wavelet nonrepeatability in time-lapse data. In addition, we discuss the influence of several common nonrepeatability issues on the time-lapse inversion results. Testing on the BP model shows that the parallel-difference time-lapse method is more sensitive to nonrepeatability-related problems than the other employed 4D strategies.

#### 4.1 Introduction

Time-lapse (4D) seismic processing is an important tool for monitoring subsurface changes caused by hydrocarbon production and CO<sub>2</sub> injection (Lumley, 2010; Pevzner et al., 2017; Smith and Tsvankin, 2013). Full-waveform inversion (FWI) can estimate the medium parameters with high spatial resolution by minimizing the phase and amplitude differences between the observed and simulated data (Li et al., 2021; Singh et al., 2018; Vigh et al., 2014; Zhang and Alkhalifah, 2020). Therefore, time-lapse FWI, which involves the inversion of the baseline and monitor data, can be efficiently used to monitor 4D changes in the reservoir (Asnaashari et al., 2015; Li et al., 2021; Lumley, 2001b).

Plessix et al. (2010) develop the parallel-difference (PD) 4D FWI strategy that inverts the baseline and monitor data independently using the same initial model. Another workflow (Routh et al., 2012) uses the

---

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Department of Geophysics, Colorado School of Mines, Golden, CO, USA

inversion of the baseline survey to build the initial model for inverting the monitor data (sequential-difference strategy; SD), which improves the convergence of the monitor inversion. To increase the sensitivity of FWI to the changes inside the reservoir, Watanabe et al. (2004) and Denli and Huang (2009) suggest to directly invert the difference between the monitor and baseline data for the time-lapse parameter variations (double-difference strategy; DD). Zhou and Lumley (2021a) average the parameter variations produced by forward (baseline to monitor) and reverse (monitor to baseline) applications of the sequential-difference strategy to reduce the dependence of the estimated time-lapse changes on the accuracy of the inverted baseline model (central-difference strategy; CD). However, the CD approach doubles the computational time compared to the SD and DD methods without a significant improvement in resolution.

Most existing 4D FWI methods are limited to isotropic (often acoustic) media, which limits their applicability. Liu and Tsvankin (2021) extend the methodology of 4D FWI to VTI (transversely isotropic with a vertical symmetry axis) media and implement it with three different time-lapse strategies mentioned above (PD, SD, and DD). However, transversely isotropic layers (such as shales) may be dipping, which leads to a tilted symmetry axis (e.g., Tsvankin, 2012). For example, uptilted shale layers near salt domes produce an effective tilted TI (TTI) model, often with a large inclination of the symmetry axis.

As shown by a number of synthetic and field-data studies, VTI algorithms become inadequate in the presence of an even moderate symmetry-axis tilt. For example, Behera and Tsvankin (2009) demonstrate that application of VTI velocity-analysis and imaging methods to data from TTI media may lead to significant misfocusing of reflectors and errors in parameter estimation. Wang and Tsvankin (2013) develop 2D P-wave reflection tomography for TTI media and test it on field data from Volve field in the North Sea. Their results show that taking the symmetry-axis tilt into account significantly improves migrated images. Singh et al. (2021) apply an elastic TTI FWI algorithm to the Volve data set and incorporate facies constraints to regularize the inversion. However, to our knowledge, 4D FWI for TTI media has not been discussed in the literature.

The increased number of independent parameters for tilted transverse isotropy makes the objective function more multimodal than that for VTI media, and the model-updating algorithm is more likely to get trapped in local minima. The ill-posedness of FWI can be mitigated by adding regularization terms and prior constraints to the objective function (Asnaashari et al., 2013; Zhang et al., 2018). Another problem is the trade-offs between the TTI parameters that can significantly distort the estimated model. Sensitivity analysis can provide valuable insights into such trade-offs and optimal parameterization for different types of input data (Alkhalifah and Plessix, 2014; Kamath and Tsvankin, 2016; Rusmanugroho et al., 2017; Singh and Tsvankin, 2020).

One of the challenges in FWI is its high sensitivity to the accuracy of the estimated source wavelet (Luo et al., 2014; Pratt, 1999; Song et al., 1995; Warner et al., 2013; Yuan et al., 2014). This issue is particularly significant for time-lapse FWI because the wavelet can change between the baseline and monitor surveys. Liu and Tsvankin (2022) incorporate a source-independent (SI) technique (Bai and Tsvankin, 2019; Choi and Alkhalifah, 2011; Zhang et al., 2016) into 4D FWI for VTI media and demonstrate its effectiveness in mitigating the influence of errors in the source wavelet on the obtained time-lapse parameter variations.

Another common issue in time-lapse seismic is the nonrepeatability (NR) of the baseline and monitor surveys. Ideally, the difference between the monitor and baseline data is caused only by the subsurface changes. However, the data difference can be distorted by a number of other factors. Zhou and Lumley (2021b) discuss the influence of several NR issues on time-lapse parameter inversion and the performance of 4D methods. However, their analysis is limited to isotropic acoustic media with known density, an inadequate model for many case studies.

Here, we extend the previously developed time-lapse VTI FWI algorithm to TTI media, which makes it suitable for a wider range of subsurface models. We begin by discussing FWI methodology for anisotropic media and its application to time-lapse seismic data. Then we introduce a TTI FWI algorithm and implement it using three 4D strategies discussed above (PD, SD, and DD). The proposed methodology is tested on synthetic data for the section of the BP TTI model that includes an anticline. Then the source-independent method is incorporated into the TTI extension of the FWI algorithm to reduce the impact of the distortions in the source signature. In addition, we analyze the influence of several other common nonrepeatability issues on 4D FWI and the performance of the employed time-lapse strategies.

## 4.2 Methodology of source-independent time-lapse FWI for TTI media

### 4.2.1 FWI for anisotropic media

Full-waveform inversion is designed to minimize the difference between the observed and simulated data, typically using the  $L_2$ -norm objective function  $S(\mathbf{m})$  (e.g., Tarantola, 1984):

$$S(\mathbf{m}) = \frac{1}{2} \left\| \left[ \mathbf{d}^{\text{sim}}(\mathbf{m}) - \mathbf{d}^{\text{obs}} \right] \right\|^2, \quad (4.1)$$

where  $\mathbf{d}^{\text{sim}}$  is the data simulated for the model  $\mathbf{m}$  and  $\mathbf{d}^{\text{obs}}$  is the observed data. Here, we minimize this objective function for multicomponent data from 2D TTI media.

We describe P- and SV-waves in TTI media using a velocity parameterization that was shown to reduce the trade-offs between the TI parameters (Kamath and Tsvankin, 2016). Also, velocities (and their gradients) share the same units, which facilitates the optimization process. This notation includes the P- and S-wave velocities in the symmetry-axis direction ( $V_{P0}$  and  $V_{S0}$ ), the P-wave velocity in the isotropy

plane, which is orthogonal to the symmetry axis ( $V_{\text{hor,P}}$ ), and the P-wave normal-moveout velocity from a horizontal reflector in the corresponding VTI medium ( $V_{\text{nmo,P}}$ ). These four parameters along with density and the symmetry-axis tilt  $\theta$  fully describe the signatures of P- and SV-waves in the vertical plane that contains the symmetry axis.

Following Kamath and Tsvankin (2016) and Liu and Tsvankin (2021), the adjoint-state method is used to calculate the gradient of the objective function  $S(\mathbf{m})$  with respect to the model parameters. We employ a nonlinear conjugate-gradient method to iteratively update the medium parameters:

$$\frac{\partial S}{\partial \mathbf{m}} = - \left[ \frac{\partial \mathbf{d}^{\text{sim}}}{\partial \mathbf{m}} \right] \mathbf{q}, \quad (4.2)$$

where  $\mathbf{q}$  is the residual wavefield.

Forward modeling is based on the finite-difference solution of the elastic wave equation for heterogeneous TTI media:

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left[ c_{ijkl} \frac{\partial u_k}{\partial x_l} \right] = f_i, \quad (4.3)$$

where  $\mathbf{u}$  is the particle displacement,  $\rho$  is the density,  $c_{ijkl}$  is the stiffness tensor, and  $\mathbf{f}$  is the density of the body forces. The wave equations for TTI and VTI media are described in more detail in Appendix B.

## 4.2.2 Strategies of time-lapse FWI

Time-lapse (4D) FWI operates with the baseline and monitor data sets. Here, we focus on three common 4D FWI strategies discussed in the introduction: the parallel-difference (PD; Plessix et al., 2010), sequential-difference (SD; Asnaashari et al., 2012), and double-difference (DD; Denli and Huang, 2009; Waldhauser and Ellsworth, 2020) methods. To operate directly with the observed ( $\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}$ ) and simulated ( $\mathbf{d}_m^{\text{sim}} - \mathbf{d}_b^{\text{sim}}$ ) data difference, the DD method generates the ‘‘composite’’ data  $\mathbf{d}_{\text{com}}$ :

$$\Delta \mathbf{d} = (\mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}}) - (\mathbf{d}_m^{\text{sim}} - \mathbf{d}_b^{\text{sim}}) = \mathbf{d}_{\text{com}} - \mathbf{d}_m^{\text{sim}}, \quad (4.4)$$

$$\mathbf{d}_{\text{com}} = \mathbf{d}_m^{\text{obs}} - \mathbf{d}_b^{\text{obs}} + \mathbf{d}_b^{\text{sim}}, \quad (4.5)$$

where the subscripts ‘‘b’’ and ‘‘m’’ refer to the baseline and monitor data, respectively. The DD method estimates the parameters of the monitor model by minimizing  $\Delta \mathbf{d}$  from equation 4.4; the initial monitor model is obtained by inverting the baseline data.

## 4.3 Synthetic Example

### 4.3.1 2D BP TTI model

The proposed time-lapse FWI algorithm is applied to the section of the 2D BP TTI model (Figure 4.1a-e), which contains an anticline. The model is  $4.5 \times 1.8$  km with a grid size of  $10 \times 10$  m;

the data are excited by 56 shots placed at the water surface with a constant interval of 80 m. There are 435 receivers on the sea floor at a depth of 210 m, which emulates ocean-bottom acquisition. The maximum tilt for this section of the BP model does not exceed  $35^\circ$  (Figure 4.2).

To obtain the initial model (Figure 4.1f-j), a Gaussian smoothing filter with a standard deviation of  $70 \times 70$  m is applied to the actual parameter fields. Reflection tomography and/or borehole information can help build the initial model in field-data applications. The monitor model is generated by reducing the baseline parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$  in the target area (the thin horizontal layer) by 10% (Figure 4.1k, Figure 4.1l, and Figure 4.1m).

The input data are the vertical and horizontal particle velocities simulated using the elastic wave equation for TTI media. FWI updates all parameters (the velocities and density) simultaneously. A multiscale approach with four frequency bands is implemented to mitigate the cycle-skipping problem in FWI. Data below 2 Hz are excluded from the inversion because such low frequencies are usually unavailable in the field.

To estimate the tilt of the symmetry axis, we conduct depth migration for the initial baseline model and set the axis orthogonal to the reflectors (Singh et al., 2021; Wang and Tsvankin, 2013). Because the tilt for the BP model is mild, we also use a VTI migration code to generate the depth image and estimate the dips, which yields the results that are similar to those for the TTI migration. An alternative way to generate the depth image from the velocity model itself, proposed by Davy et al. (2021), also works well for our model. Because the symmetry-axis tilt is mild and does not change significantly over iterations, we update the tilt only after each inversion stage (frequency band).

As mentioned above, we apply 4D FWI with the parallel-difference (PD), sequential-difference (SD), and double-difference (DD) strategies. The inverted baseline model is applied to the monitor inversion of the SD and DD strategies without any smoothing. For the first test, the monitor data are acquired with perfect repeatability. The results (used below as the benchmark) show that all three strategies reconstruct 4D variations with sufficient spatial resolution (Figure 4.3). Because of the parameter trade-offs, the time-lapse variations of  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  (which are unchanged) do not completely vanish. The DD strategy reconstructs the 4D variations with the highest resolution and generates the fewest errors outside the reservoir because the inversion of the monitor data operates with the actual 4D data difference. The SD and DD strategies are more computationally efficient and yield fewer artifacts than PD because the monitor inversion starts from the estimated baseline model.

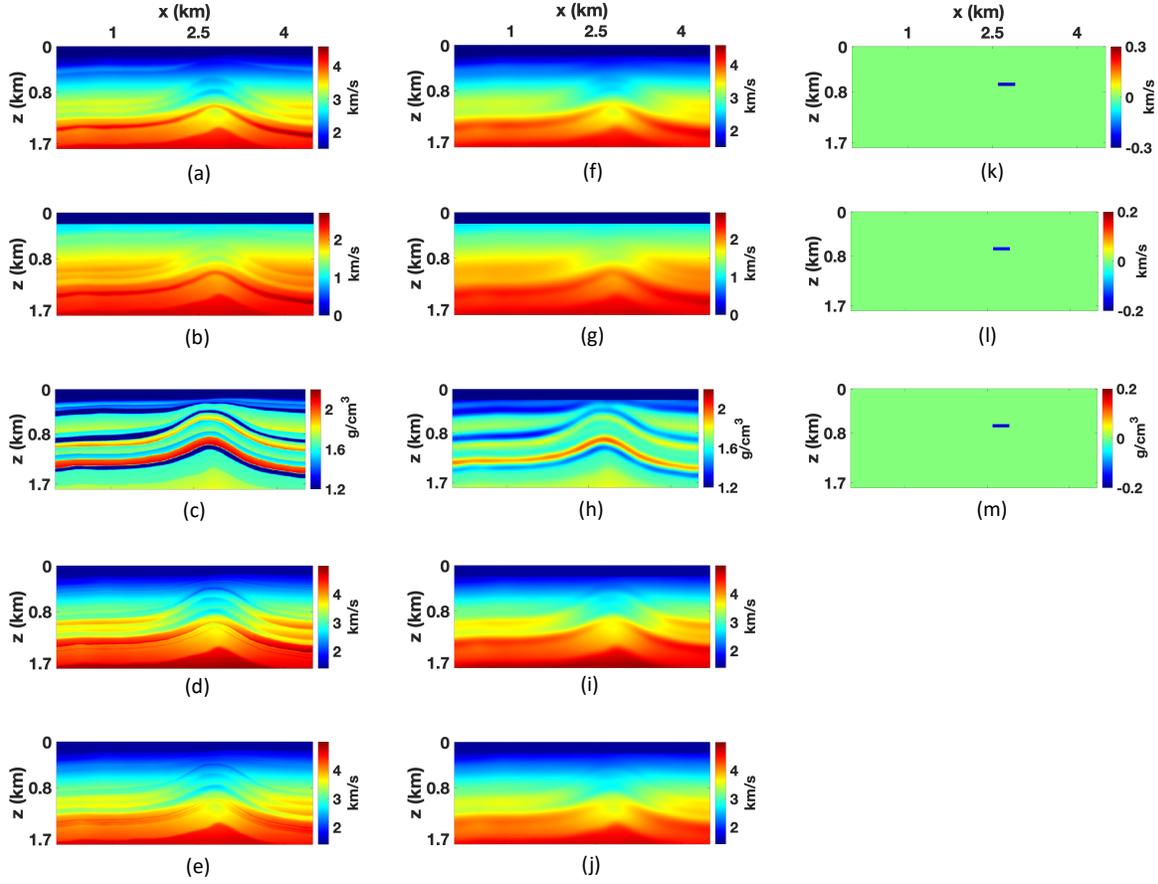


Figure 4.1 Parameters of the baseline BP TTI model with a grid size of  $10 \times 10$  m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (b) the S-wave vertical velocity ( $V_{S0}$ ), (c) the density ( $\rho$ ), (d) the P-wave velocity in the isotropy plane ( $V_{hor,P}$ ), and (e) the P-wave normal-moveout velocity ( $V_{nmo,P}$ ). The initial baseline model of: (f)  $V_{P0}$ , (g)  $V_{S0}$ , (h)  $\rho$ , (i)  $V_{hor,P}$ , and (j)  $V_{nmo,P}$ . The actual time-lapse differences in (k)  $V_{P0}$ , (l)  $V_{S0}$ , and (m)  $\rho$ . Note that there are no time-lapse changes in  $V_{hor,P}$  and  $V_{nmo,P}$ .

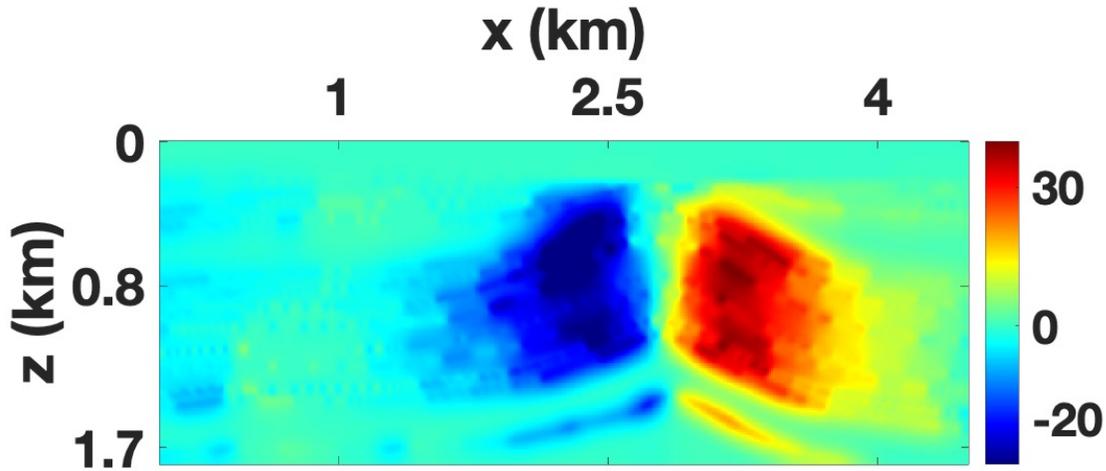


Figure 4.2 Tilt of the symmetry axis estimated by migrating the baseline data using the actual baseline model.

#### 4.3.1.1 Influence of tilt

To evaluate the influence of the symmetry-axis tilt on the wavefield, we compare the vertical particle velocity for the actual baseline model of  $V_{P0}$ ,  $V_{S0}$ ,  $V_{hor,P}$ ,  $V_{nmo,P}$ , and  $\rho$  simulated using TTI and VTI algorithms (Figure 4.4). The reflected, diving, and converted waves computed for the two models match reasonably well for offsets smaller than 2 km because the tilt is relatively mild. The influence of tilt increases with offset  $x$  and becomes more significant for waves that cross the dipping flanks of the anticline ( $x \geq 2$  km) where the tilt is largest. Because the reservoir is subhorizontal (Figure 4.2), the VTI algorithm still resolves the parameter changes with acceptable accuracy.

Next, we design a different monitor model where the target layer is dipping and embedded in the left flank of the anticline (Figure 4.5a-c). After simulating the data for the actual TTI model, we conduct time-lapse FWI using the SD strategy. The 4D TTI algorithm accurately estimates the parameter changes in the reservoir (Figure 4.5i-m). In contrast, the size and magnitude of the time-lapse parameter variations (especially in  $\rho$ ) is underestimated if the symmetry-axis tilt is ignored in FWI (Figure 4.5d-h). In addition, the VTI inversion generates significant false anomalies in all parameters along and below the anticline (red arrows). Evidently, the inadequacy of the VTI model leads to parameter cross-talk that produces false anomalies in the velocities  $V_{hor,P}$  and  $V_{nmo,P}$  (which are held constant). Therefore, it is essential to apply the TTI algorithm if the symmetry axis is tilted inside and around the reservoir, even if the tilt is relatively mild.

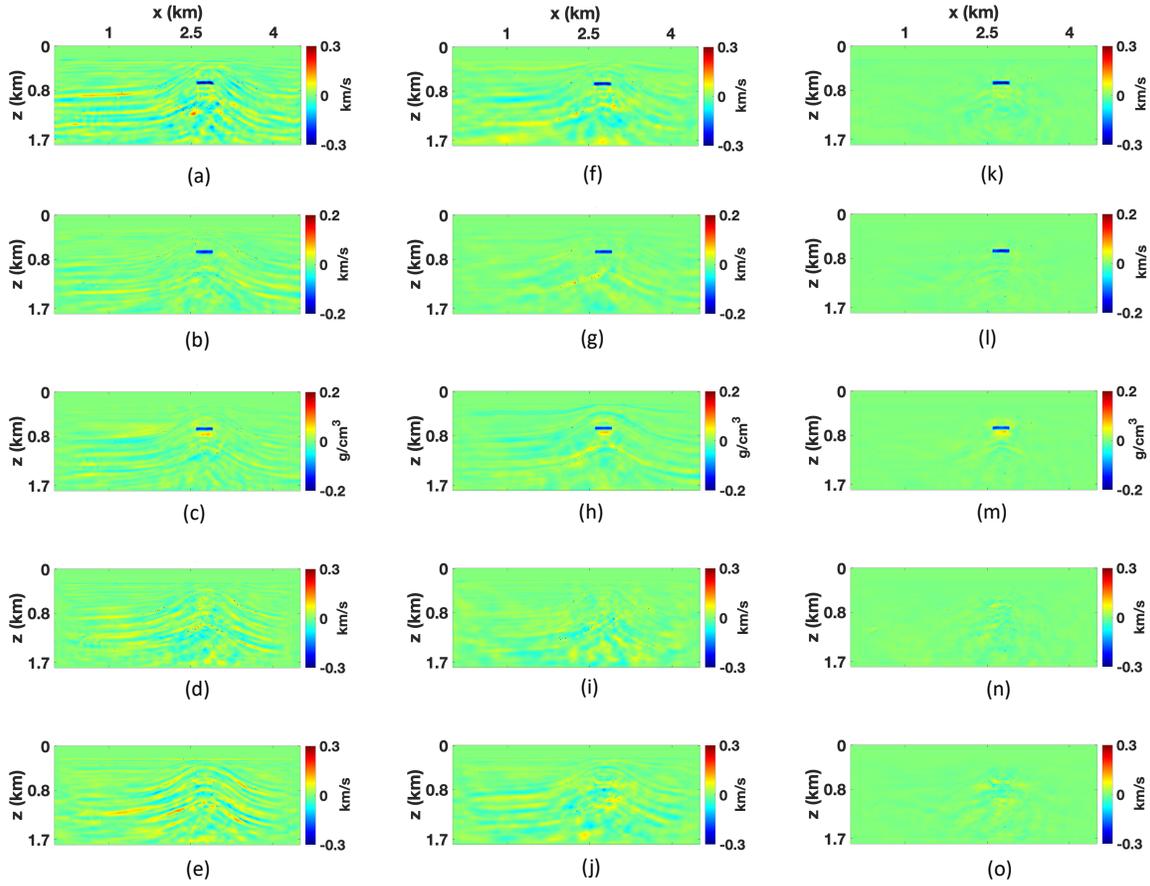


Figure 4.3 Time-lapse parameter variations obtained from noise-free multicomponent data. The parallel-difference method: (a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $\rho$ , (d)  $V_{hor,P}$ , and (e)  $V_{nmo,P}$ . The sequential-difference method: (f)  $V_{P0}$ , (g)  $V_{S0}$ , (h)  $\rho$ , (i)  $V_{hor,P}$ , and (j)  $V_{nmo,P}$ . The double-difference method: (k)  $V_{P0}$ , (l)  $V_{S0}$ , (m)  $\rho$ , (n)  $V_{hor,P}$ , and (o)  $V_{nmo,P}$ .

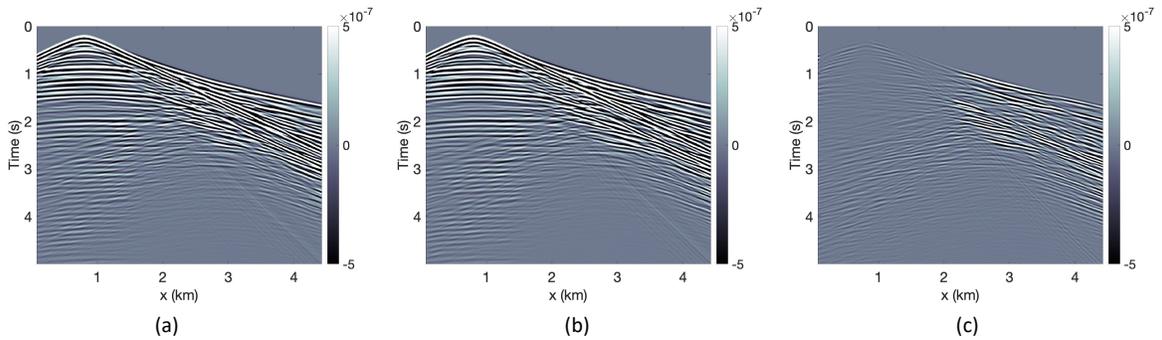


Figure 4.4 Shot gathers of the vertical particle velocity for: (a) the actual TTI medium, and (b) the corresponding VTI medium. Plot (c) shows the difference between plots (a) and (b).

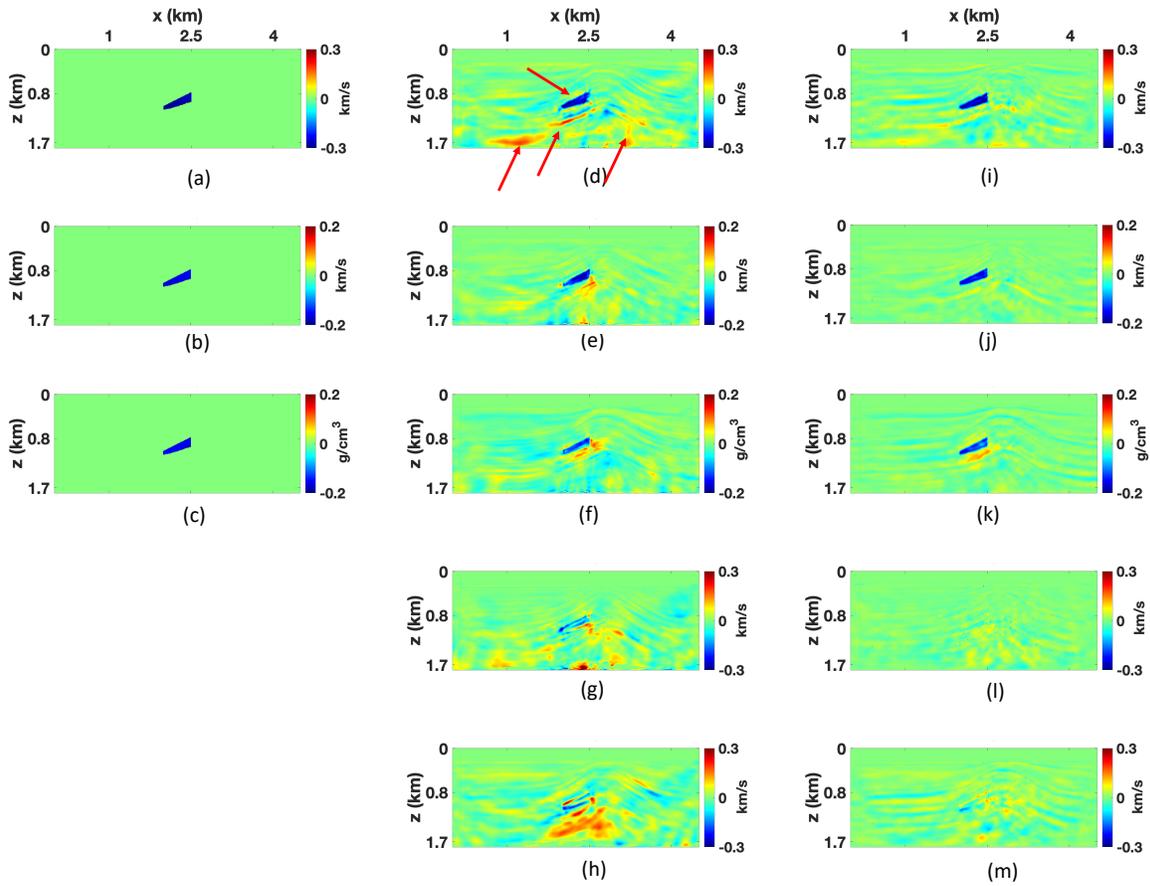


Figure 4.5 Time-lapse changes inside the reservoir located in a dipping segment of the anticline. The actual time-lapse changes: (a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $\rho$ . The time-lapse variations estimated by the VTI [(d)  $V_{P0}$ , (e)  $V_{S0}$ , (f)  $\rho$ , (g)  $V_{hor,P}$ , and (h)  $V_{nmo,P}$ ] and TTI [(i)  $V_{P0}$ , (j)  $V_{S0}$ , (k)  $\rho$ , (l)  $V_{hor,P}$ , and (m)  $V_{nmo,P}$ ] algorithms.

### 4.3.1.2 Influence of nonrepeatable noise

To test the robustness of our TTI method for noisy data, we add nonrepeatable random Gaussian noise with the signal-to-noise ratio (SNR) equal to 15 to both the baseline and monitor data for the model in Figure 4.1.

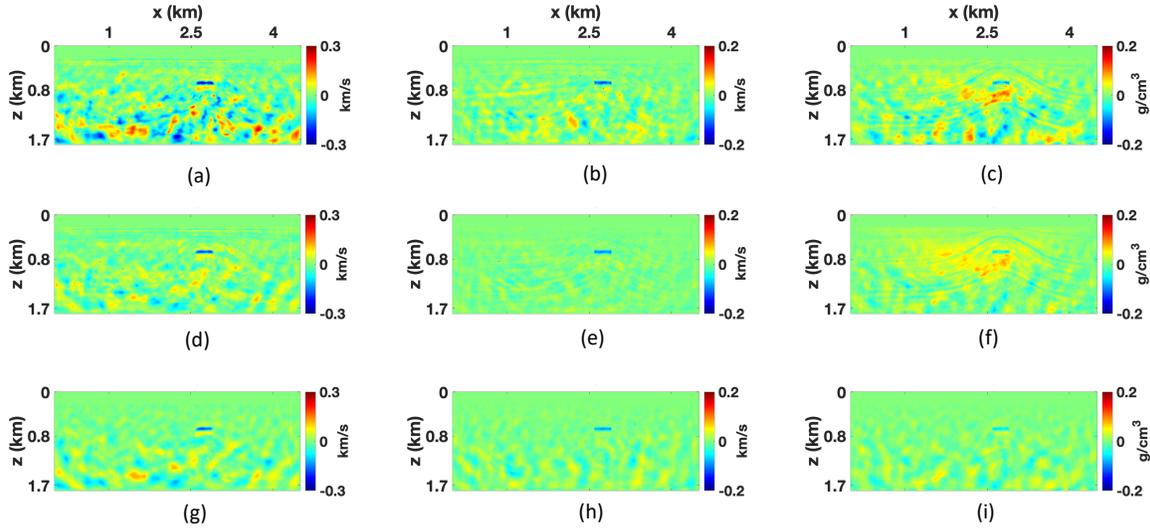


Figure 4.6 Time-lapse parameter variations obtained from noisy baseline and monitor data (the signal-to-noise ratio is 15) for the model in Figure 4.1. The parallel-difference strategy: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ . The sequential-difference strategy: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ . The double-difference strategy: (g)  $V_{P0}$ , (h)  $V_{S0}$ , and (i)  $\rho$ .

Figure 4.6 shows the inversion results for the parameters  $V_{P0}$ ,  $V_{S0}$ , and  $\rho$ . As expected, the added noise degrades the reconstruction of the 4D variations (especially in  $\rho$ ) and enhances parameter trade-offs. Similar to the noise-free test, the SD and DD strategies produce relatively few artifacts, whereas the noise-induced distortions are more significant in the output of the PD strategy. These observations are consistent with the results of Liu and Tsvankin (2021) for VTI media.

### 4.3.1.3 Influence of source wavelet

In 4D seismic processing, the source wavelet is usually estimated from the baseline data after matching monitor and baseline records. However, wavelet extraction is time-consuming and error-prone, and the wavelets for the baseline and monitor data sets may differ. Errors in the wavelet distort the FWI results because they hinder matching of the observed and simulated data.

To analyze the influence of signal distortions, we carry out the inversion using a wavelet (Wavelet 1, Figure 4.7b) that substantially differs from the actual one (Figure 4.7a) in both phase and amplitude.

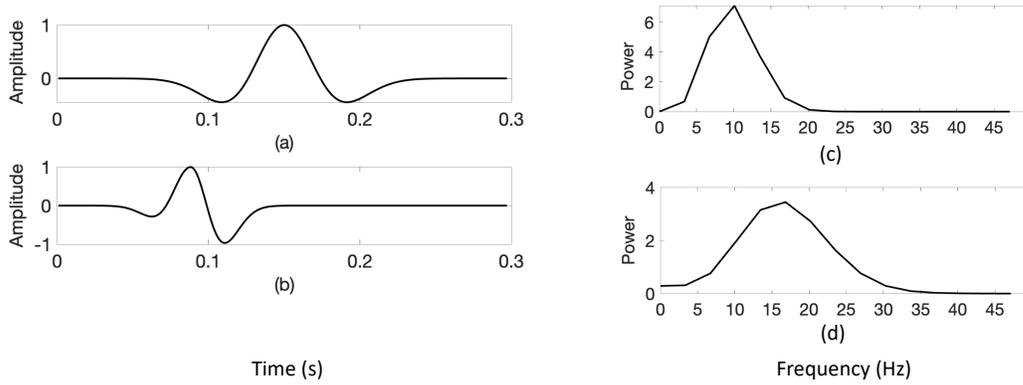


Figure 4.7 Source wavelets used in the synthetic examples: (a) the Ricker wavelet with a central frequency of 10 Hz (actual wavelet) and (b) the distorted “Ricker” wavelet with a central frequency of 17 Hz (Wavelet 1). The frequency spectra of: (c) the actual wavelet [see plot (a)] and (d) Wavelet 1 [see plot (b)].

First, we apply FWI to the baseline data using Wavelet 1. Because of the wavelet distortion, the model-updating algorithm could not converge to the global minimum of the objective function and reconstruct the anticline and the reservoir (Figure 4.8a-c). It is clear that conventional FWI would not be able to estimate the time-lapse changes due to errors in the wavelet.

To reduce the dependence of the inversion results on the accuracy of the source signal, we follow Liu and Tsvankin (2022) and implement the convolution-based source-independent (SI) objective function in the time domain (Choi and Alkhalifah, 2011). The SI algorithm (Figure 4.8d-f) dramatically improves the accuracy of the inverted parameters and reconstructs the baseline model with a resolution close to that of the benchmark sections (compare Figure 4.8d-f and Figure 4.8a-c).

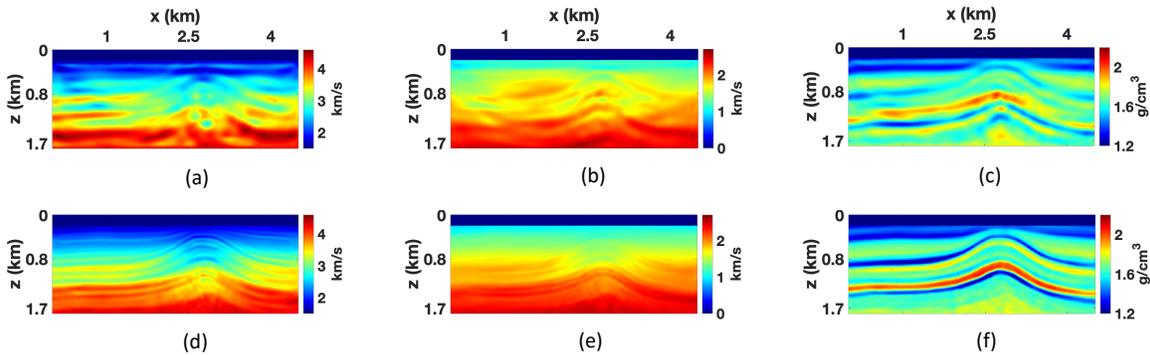


Figure 4.8 Baseline models estimated using Wavelet 1 by the conventional FWI [(a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ ] and the source-independent FWI [(d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ ].

Next we perform the monitor inversion using the source-independent algorithm with the PD and SD strategies and compute the time-lapse parameter variations. The DD method is not employed here because of the phase mismatch between the data simulated for the inverted baseline model using the distorted wavelet and the observed data (Liu and Tsvankin, 2022). Despite underestimating the time-lapse changes inside the reservoir, the output of the PD and SD methods is comparable to the benchmark sections, which is consistent with the results of Liu and Tsvankin (2022) for VTI media. Similar to the benchmark test, the SD workflow recovers the time-lapse changes with a higher resolution and fewer artifacts than the PD strategy (Figure 4.9). The SI methodology remains sufficiently accurate for noisy data with moderate levels of noise.

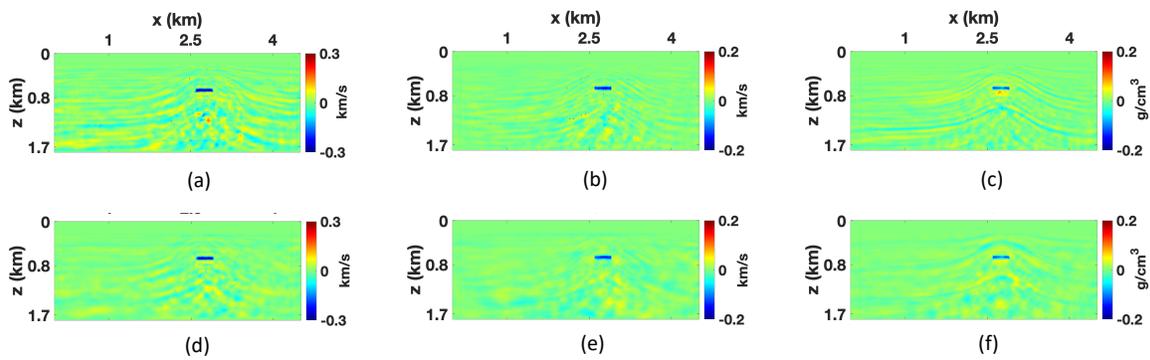


Figure 4.9 Time-lapse parameter variations reconstructed by the source-independent algorithm using Wavelet 1. The parallel-difference method: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ . The sequential-difference method: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ .

#### 4.3.1.4 Influence of geometry nonrepeatability

The geometries of the baseline and monitor surveys are seldom perfectly repeatable, even if the ocean-bottom nodes (OBN) are equipped with GPS devices. Preprocessing can reduce the impact of the geometry NR but it is difficult to eliminate it completely. To explore the influence of the geometry changes on 4D FWI, we move each source in the monitor survey up and to the right by 10 m. Due to this source displacement, the data difference exhibits changes near and after the first arrival, which are unrelated to the actual time-lapse anomalies (Figure 4.10b).

First, 4D FWI is applied using the actual acquisition geometries. Not surprisingly, the shift in the source positions does not significantly influence the reconstruction of the time-lapse changes with all three time-lapse strategies (Figure 4.11a-c) because the baseline and monitor inversions are performed with the actual geometries. Although the DD strategy operates on the data difference, the error caused by the geometry changes cancels out in equation 4.4.

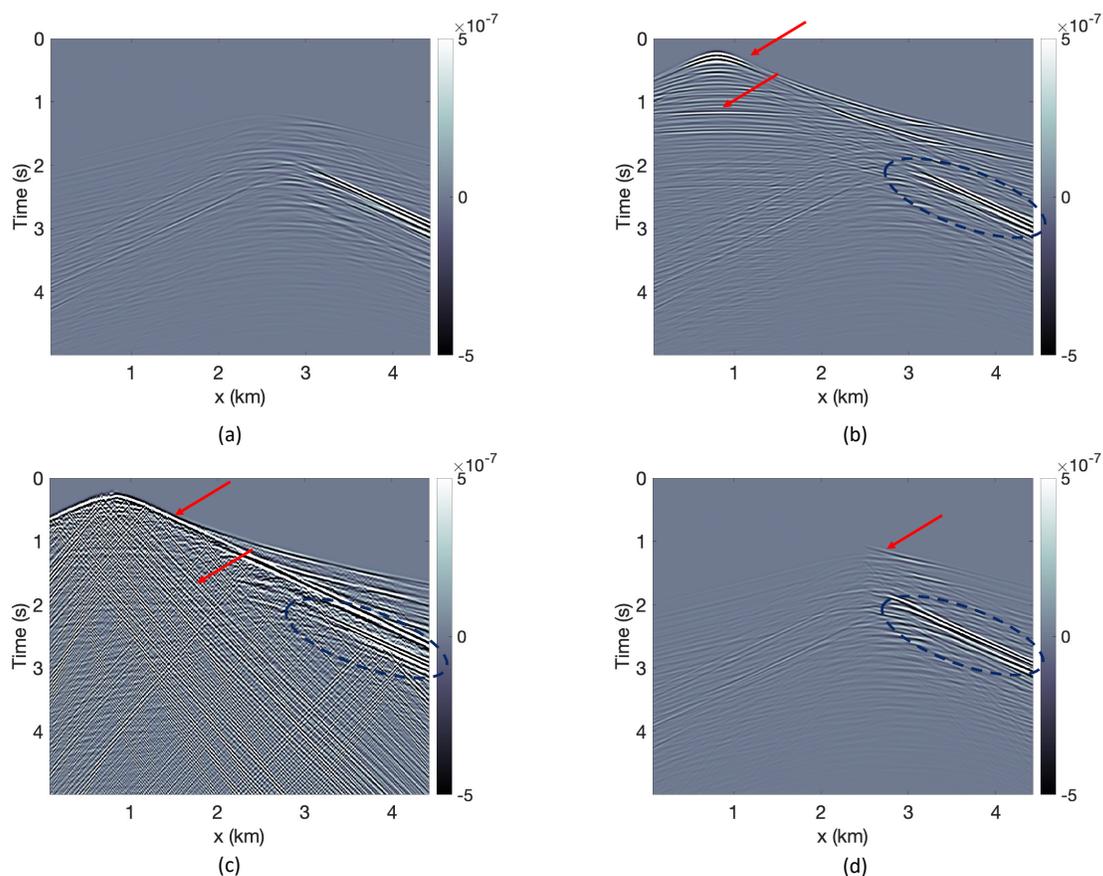


Figure 4.10 Difference between the vertical particle-velocity components from the monitor and baseline surveys: (a) the perfectly repeatable data (referred to as the actual difference) and the data with (b) geometry nonrepeatability, (c) water statics, and (d) overburden changes. The actual differences caused by the reservoir changes are circled in blue. The noise caused by nonrepeatability is marked by the red arrows.

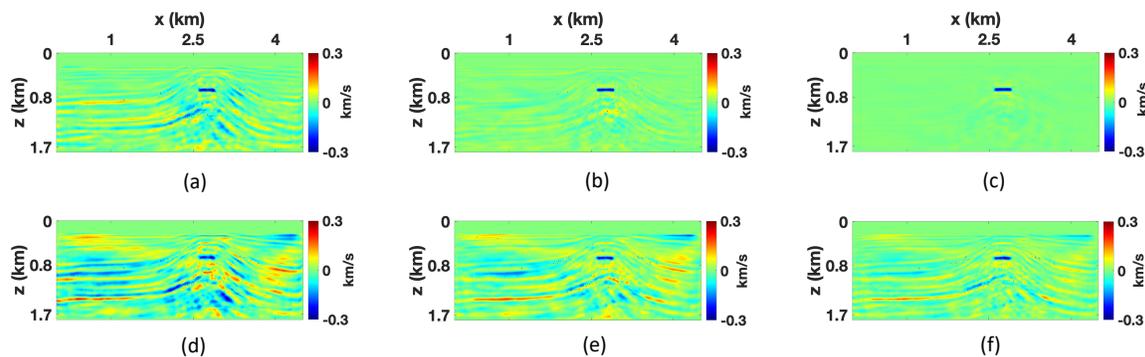


Figure 4.11 Time-lapse parameter variations of the velocity  $V_{P0}$  estimated from the data with geometry nonrepeatability. The results obtained with the actual acquisition geometries: (a) the parallel-difference method, (b) the sequential-difference method, and (c) the double-difference method. The results obtained using the baseline geometry in the inversion of both the baseline and monitor data: (d) the parallel-difference method, (e) the sequential-difference method, and (f) the double-difference method.

Next, to evaluate the influence of geometry errors on the reconstructed 4D parameter changes, the baseline data are inverted using the actual geometry, whereas the monitor inversion is performed with the geometry of the baseline survey. Figure 4.11d-f show that all three time-lapse strategies produce significant artifacts in the 4D results because of the time shifts between the simulated and observed monitor data caused by the geometry errors. In particular, these time shifts lead to mispositioning of the reflectors due to the mismatch between these two data sets. Note that the PD strategy starts from a smoothed initial model in the monitor inversion, which gives the model-updating algorithm more flexibility to misplace reflectors. Therefore, the parameter variations produced by the PD strategy (Figure 4.11d) are distorted more significantly by the geometry error compared to the two other strategies (Figure 4.11e and f), which agrees with the observation of Zhou and Lumley (2021b).

#### 4.3.1.5 Influence of water statics

The water velocity may change between the baseline and monitor surveys, and the corresponding time shifts for some deep-water fields can reach 8 ms (Lecerf et al., 2022). Such “water statics” is often corrected by applying time shifts trace-by-trace before the inversion. However, those corrections may not be adequate for time-lapse FWI if the parameter changes are relatively small (Borges et al., 2022). In the next test, we perturb the water velocity of the monitor model to generate time shifts up to 3 ms in the monitor data. Both the baseline and monitor inversions are performed with the water velocity of the baseline model.

Figure 4.10c shows that the water velocity variations produce noticeable changes in the wavefields above the target area. Because of the incorrect water velocity for the monitor inversion, all three time-lapse strategies generate artifacts and underestimate the size of the reservoir and the amplitude of the parameter changes (Figure 4.12). Predictably, the parameters  $V_{P0}$  and  $\rho$  are distorted more than  $V_{S0}$  because there are no S-waves in the water. Similar to the geometry NR test above, the PD strategy yields the 4D results with the most artifacts due to the smoothed initial model in the monitor inversion. The SD and DD strategies produce comparable results, which indicates that these two approaches better handle moderate NR-related distortions.

#### 4.3.1.6 Influence of overburden changes

Hydrocarbon production and CO<sub>2</sub> injection reduce the pore pressure inside the reservoir, which causes stress changes in the surrounding rocks (e.g., Holt et al., 2016; Smith and Tsvankin, 2012). These overburden and underburden changes produce time shifts that can provide useful information for reservoir monitoring (e.g., Smith and Tsvankin, 2013). Precise modeling of stress-related velocity variations is beyond the scope of this paper, so here we simulate such overburden changes by simply increasing the

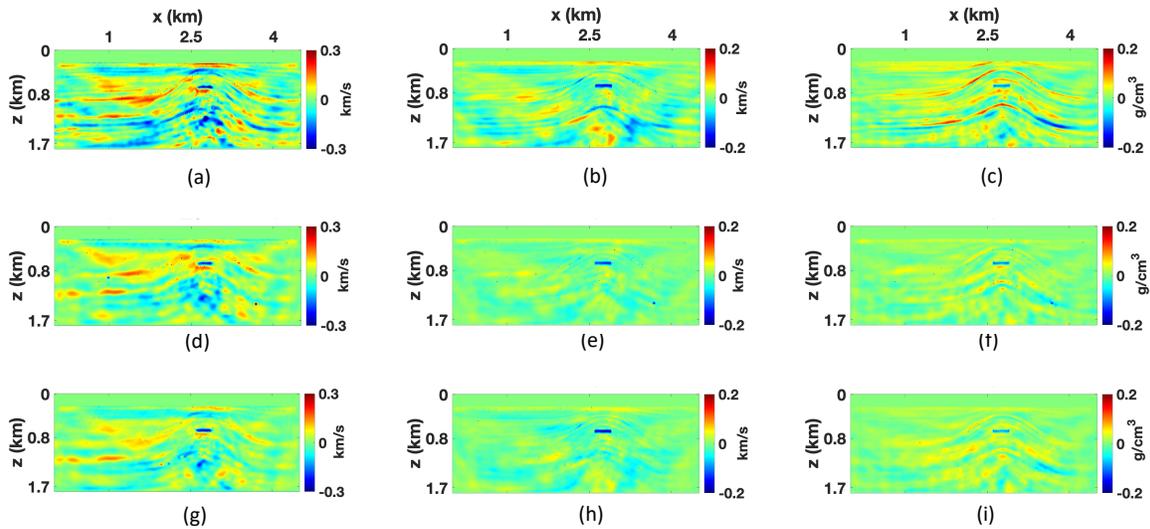


Figure 4.12 Time-lapse parameter variations estimated from the data with “water statics.” The parallel-difference method: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ . The sequential-difference method: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ . The double-difference method: (g)  $V_{P0}$ , (h)  $V_{S0}$ , and (i)  $\rho$ .

baseline velocities  $V_{P0}$  and  $V_{S0}$  above the reservoir by up to 4% for the monitor survey (Figure 4.13a and Figure 4.13b).

The influence of the overburden changes is visible in the difference between the monitor and baseline data (Figure 4.10d). Still, likely because the amplitude of the overburden velocity changes is smaller than those in the reservoir, all three strategies are able to reconstruct the time-lapse variations in the reservoir and overburden with sufficient accuracy (Figure 4.13). Indeed, FWI is generally capable of estimating the entire model, although relatively large overburden anomalies could potentially mask those in the reservoir. Similar to the repeatability tests above, the DD strategy (Figure 4.13g and Figure 4.13h) reconstructs the parameter changes with the highest resolution and fewest artifacts. This test also illustrates the sensitivity of our algorithm to relatively small changes in the overburden for input data with a low noise level.

#### 4.4 Conclusions

We extended the previously developed 4D FWI methodology from VTI to TTI media, which makes it suitable for a wide range of subsurface structures. The algorithm is tested on the anticlinal section of the BP TTI model using three difference time-lapse strategies: the parallel- (PD), sequential- (SD), and double-difference (DD) workflows. To mitigate cycle-skipping, a multiscale technique with four frequency bands is employed in both the baseline and monitor inversions. The velocity parameters responsible for P- and SV-wave propagation in TTI models are updated (along with density) simultaneously at each iteration. The tilt of the symmetry axis is estimated from the structure dips by setting the axis orthogonal

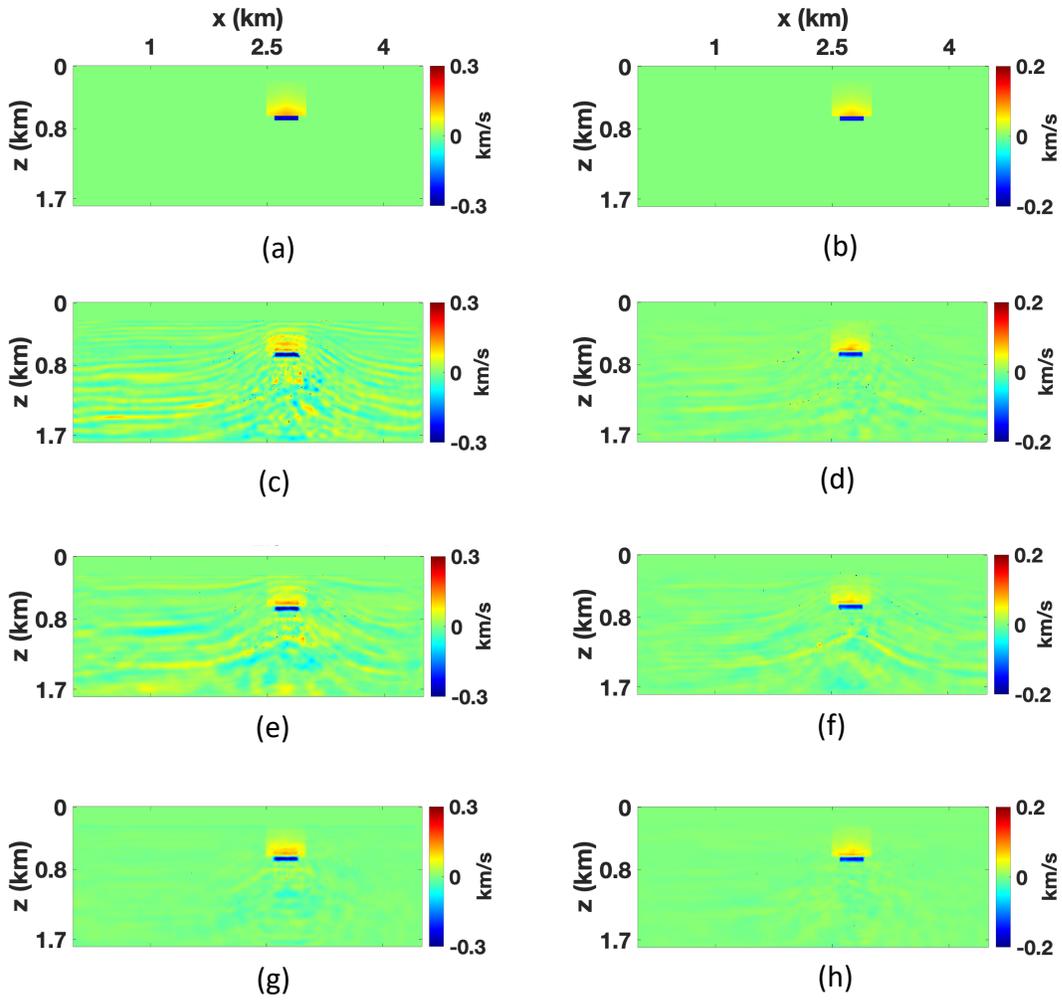


Figure 4.13 Time-lapse parameter variations estimated from the data with overburden changes. The actual time-lapse variations in (a)  $V_{P0}$  and (b)  $V_{S0}$ . The parallel-difference strategy: (c)  $V_{P0}$  and (d)  $V_{S0}$ . The sequential-difference strategy: (e)  $V_{P0}$  and (f)  $V_{S0}$ . The double-difference strategy: (g)  $V_{P0}$  and (h)  $V_{S0}$ .

to reflectors in migrated images. The tilt model is updated after completing the inversion for each frequency band. If the baseline and monitor data are repeatable, our algorithm reconstructs the time-lapse variations of the TTI parameters with high spatial resolution, even in the presence of moderate noise.

We also incorporated the source-independent (SI) technique into FWI to mitigate the influence of errors in the estimated source wavelet on the time-lapse results. The SI FWI algorithm proved capable of reconstructing 4D parameter changes with sufficient accuracy even for a strongly distorted wavelet and noisy data. Employed with the parallel-difference and sequential-difference strategies, the SI FWI algorithm can overcome the source-wavelet nonrepeatability (NR) issue in field-data applications.

In addition, we analyzed the impact of other common NR problems, such as geometry changes, water statics, and time-lapse velocity variations in the overburden. In the test with different source locations for the baseline and monitor surveys, the time-lapse FWI that operated with the actual acquisition geometries accurately estimated the 4D changes. Water statics was shown not to have a significant influence on the inverted parameter variations, unless the corresponding time shifts are comparable to those caused by the reservoir changes. If stress-induced velocity changes in the overburden do not dominate the time-lapse response, FWI is capable of resolving the 4D parameter variations in the entire model including the reservoir.

Our testing also demonstrates that NR issues are best handled with the SD and DD strategies, whereas the results obtained with the PD workflow are generally inferior. The PD strategy operates with a smoothed initial model, for which the NR-induced time shifts between the simulated and observed data may lead to significant distortions, such as reflector mispositioning. The SD and DD strategies generally produce comparable 4D results, with the latter reconstructing the time-lapse changes with the fewer artifacts for data with a low level of noise and high repeatability. Note that the performance of the SD and DD strategies may vary with the model structure, the size of the reservoir, and the amplitude of the time-lapse changes. Therefore, in field applications, it is advisable to test both strategies on a subset of the data prior to processing the entire survey.

#### **4.5 Acknowledgments**

We thank the members of the Anisotropy-Team at the Center for Wave Phenomena (CWP) in the Geophysics Department at Colorado School of Mines for useful discussions. This work is supported by a joint project with INPEX and the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.

## CHAPTER 5

### ELASTIC TIME-LAPSE FWI FOR ANISOTROPIC MEDIA: A PYRENEES CASE STUDY

Submitting to *Geophysics*

Yanhua Liu<sup>1,2</sup>, Ilya Tsvankin<sup>2</sup>, Shogo Masaya<sup>3</sup>, and Masanori Tani<sup>3</sup>

In the context of reservoir monitoring, time-lapse (4D) full-waveform inversion (FWI) of seismic data has the potential to estimate reservoir changes with high resolution. However, most existing field-data applications are carried out with isotropic, and often acoustic FWI algorithms. Here, we apply a time-lapse FWI methodology for TI media with a vertical symmetry axis (VTI) to offshore streamer data acquired at Pyrenees field in Australia. We explore different objective functions, including those based on global correlation (GC) and designed to mitigate errors in the source signature (SI, or source-independent). The GC objective function, which utilizes mostly phase information, produces the most accurate inversion results by overcoming the difficulties associated with amplitude matching of the synthetic and field data. The SI FWI algorithm is generally more robust in the presence of distortions in the source wavelet than the other two methods, but its application to field data is hampered by reliance on amplitude matching. Taking anisotropy into account provides a better fit to the recorded data, especially at far offsets. The 4D response obtained by FWI shows time-lapse variations in the P-wave vertical velocity likely caused by the reservoir gas coming out of solution and the replacement of gas with oil.

#### 5.1 Introduction

Time-lapse (4D) seismic data have been extensively utilized in reservoir management to monitor subsurface changes caused by oil and gas production (Lumley, 2010; Pevzner et al., 2017; Smith and Tsvankin, 2013). Time-lapse full-waveform inversion (FWI) of baseline and monitor surveys is particularly well-suited for monitoring 4D changes in the reservoir (Asnaashari et al., 2015; Li et al., 2021; Lumley, 2001b).

Different strategies have been proposed to optimize the inversion process and increase its sensitivity to reservoir changes. One of them is the parallel-difference (PD) FWI method (Plessix et al., 2010), where the baseline and monitor data are inverted independently using the same initial model. Routh et al. (2012) present the sequential-difference (SD) strategy, which uses the baseline inversion to build the initial model for inverting the monitor data. This approach has been shown to improve the convergence of the monitor

---

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Center for Wave Phenomena, Colorado School of Mines

<sup>3</sup>INPEX corporation

inversion. The sensitivity of FWI to the time-lapse reservoir changes can be increased by applying the double-difference (DD) strategy developed by Watanabe et al. (2004) and Denli and Huang (2009). The DD method directly inverts the difference between the monitor and baseline data for the time-lapse parameter variations.

Yang et al. (2016) apply 4D FWI to an OBC (ocean-bottom-cable survey) from the Valhall field in the North Sea and show that the DD strategy produces the most accurate and interpretable reservoir changes. Hicks et al. (2016) conduct acoustic 4D FWI for two vintages of the permanent Grane reservoir-monitoring system in the North Sea and recover the P-wave velocity changes related to gas replacing oil during hydrocarbon production. Egorov et al. (2017) estimate the time-lapse changes related to CO<sub>2</sub> injection at Otway field in Australia by applying 4D elastic FWI to walkaway VSP (vertical seismic profiling) data.

Application of elastic anisotropic 4D FWI to field data involves considerable challenges due to the multimodal character of the objective function and parameter trade-offs. Liu and Tsvankin (2021) extend the methodology of time-lapse FWI to VTI media and, in a sequel paper (Liu and Tsvankin, 2022), incorporate a source-independent technique to mitigate the influence of errors in the source wavelet. Their work highlights the importance of accounting for anisotropy in time-lapse FWI to achieve accurate and robust inversion results, especially in complex geologic settings.

Here, the time-lapse FWI algorithm proposed by Liu and Tsvankin (2022) is applied to data from Pyrenees field in western Australia. First, we discuss the FWI methodology for anisotropic media using different objective functions and outline its applications to time-lapse seismic data. Basic information about the Pyrenees field and data acquisition is presented next. Then the FWI algorithm for VTI media is tested on the baseline data and the results are compared with those produced by the isotropic inversion. We also evaluate the impact of different objective functions on the inversion results. Finally, preliminary estimates of the time-lapse parameter variations are obtained using the sequential-difference strategy.

## 5.2 Methodology of source-independent full-waveform inversion

### 5.2.1 FWI objective functions

Full-waveform inversion is designed to minimize the data misfit, typically using the  $L_2$ -norm objective function  $E(\mathbf{m})$  (e.g., Tarantola, 1984):

$$E(\mathbf{m}) = \frac{1}{2} \left\| \left[ \mathbf{d}^{\text{sim}}(\mathbf{m}) - \mathbf{d}^{\text{obs}} \right] \right\|^2, \quad (5.1)$$

where  $\mathbf{d}^{\text{obs}}$  is the observed data and  $\mathbf{d}^{\text{sim}}$  is the data simulated for the model  $\mathbf{m}$ . Here, we minimize this objective function for multicomponent data from isotropic and VTI media.

The  $L_2$ -norm objective function is designed to optimize matching both the amplitude and phase of the observed and simulated wavefield. However, amplitude matching can be challenging for field data due to the complexity of wave-propagation phenomena in the subsurface. To address this issue, Choi and Alkhalifah (2012) proposed a normalized global-correlation (GC) objective function  $G(\mathbf{m})$  that reduces the dependence of data fitting on amplitude:

$$G(\mathbf{m}) = - \sum_{s=1}^{n_s} \frac{\sum_r \int_0^T \mathbf{d}^{\text{obs}} \mathbf{d}^{\text{sim}}(\mathbf{m}) dt}{\sqrt{\sum_r \int_0^T (\mathbf{d}^{\text{obs}})^2 dt} \sqrt{\sum_r \int_0^T (\mathbf{d}^{\text{sim}})^2 dt}} = - \sum_{s=1}^{n_s} \bar{\mathbf{d}}^{\text{obs}} \bar{\mathbf{d}}^{\text{sim}}. \quad (5.2)$$

Here,  $\bar{\mathbf{d}}^{\text{obs}} = \mathbf{d}^{\text{obs}} / \|\mathbf{d}^{\text{obs}}\|$  and  $\bar{\mathbf{d}}^{\text{sim}} = \mathbf{d}^{\text{sim}} / \|\mathbf{d}^{\text{sim}}\|$  denote the normalized observed and simulated data, respectively,  $n_s$  is the number of sources,  $r$  is the trace number, and  $T$  is the total recording time. The approach based on equation 5.2 has shown promising results in improving the convergence and robustness of FWI in field-data applications (Singh et al., 2021).

### 5.2.2 Source-independent FWI

Errors in the estimation of the source wavelet can significantly distort the FWI results, and this problem can be especially serious for time-lapse data. To reduce the influence of the source wavelet on FWI, Choi and Alkhalifah (2011), Zhang et al. (2016), and Liu and Tsvankin (2022) propose the following convolution-based source-independent (SI) objective function in the time domain:

$$\begin{aligned} S(\mathbf{m}) &= \frac{1}{2} \left\| \left[ \mathbf{d}^{\text{sim}}(\mathbf{m}) * \left( W \mathbf{d}_{\text{ref}}^{\text{obs}} \right) - \mathbf{d}^{\text{obs}} * \left( W \mathbf{d}_{\text{ref}}^{\text{sim}} \right) \right] \right\|^2 \\ &= \frac{1}{2} \left\| \left[ \mathbf{G}^{\text{sim}} * \mathbf{s}^{\text{sim}} * W \mathbf{G}_{\text{ref}}^{\text{obs}} * \mathbf{s}^{\text{obs}} - \mathbf{G}^{\text{obs}} * \mathbf{s}^{\text{obs}} * W \mathbf{G}_{\text{ref}}^{\text{sim}} * \mathbf{s}^{\text{sim}} \right] \right\|^2 \\ &= \frac{1}{2} \left\| \left[ \tilde{\mathbf{G}}^{\text{s}} * \tilde{\mathbf{s}}^{\text{c}} - \tilde{\mathbf{G}}^{\text{o}} * \tilde{\mathbf{s}}^{\text{c}} \right] \right\|^2, \end{aligned} \quad (5.3)$$

where  $W = W(t)$  is the chosen time window that includes the first arrival, and  $\mathbf{d}_{\text{ref}}$  denotes the reference trace.  $\mathbf{G}_{\text{ref}}^{\text{obs}}$  and  $\mathbf{G}_{\text{ref}}^{\text{sim}}$  are the Green's functions for the reference traces,  $\tilde{\mathbf{G}}^{\text{s}}$  and  $\tilde{\mathbf{G}}^{\text{o}}$  are the Green's functions for the newly computed convolution-based simulated and observed data, and  $\tilde{\mathbf{s}}^{\text{c}}$  is the new source wavelet. Note that because the objective function in equation 5.3 is based on the  $L_2$ -norm, the SI technique may still suffer from the amplitude mismatch between the observed and simulated data.

### 5.2.3 Forward modeling and model updating for FWI

The wavefield is simulated by solving the elastic wave equation for anisotropic heterogeneous media using a finite-difference scheme:

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left[ c_{ijkl} \frac{\partial u_k}{\partial x_l} \right] = f_i, \quad (5.4)$$

where  $\mathbf{u}$  is the particle displacement,  $\rho$  is the density,  $c_{ijkl}$  is the stiffness tensor, and  $\mathbf{f}$  is the density of the body forces.

Following Kamath and Tsvankin (2016) and Liu and Tsvankin (2021), the adjoint-state method is used to calculate the gradient of the objective function with respect to the model parameters. We employ a nonlinear conjugate-gradient method for iterative parameters updating. The derivatives of the objective function can be found as:

$$-\left[\frac{\partial \mathbf{d}^{\text{sim}}}{\partial \mathbf{m}}\right] \mathbf{q}, \quad (5.5)$$

where  $\mathbf{q}$  is the residual wavefield.

### 5.3 Field-data application

#### 5.3.1 Location and geology

The Pyrenees field is located in the Exmouth Subbasin, offshore western Australia, where the water depth is around 200 m (Figure 5.1). The reservoir is located about 1100 m below the sea surface. Oil and gas production from the Pyrenees field commenced in 2010 following several exploration efforts (Scibiowski et al., 2005). Studies of the area stratigraphy show that the depositional system can be characterized as a wave-dominated delta (Woodall and Stark, 2013).

Prior to production, a dedicated baseline survey was conducted over the Pyrenees fields in 2005 using conventional streamers (Woodall and Stark, 2013). The first time-lapse monitor survey was acquired in 2013, three years after production commenced in 2010, using PGS GeoStreamer cable, which has geophones as well as hydrophones, with airguns. However, as shown in Figure 5.2, the acquisition geometries of these two surveys are different, with the largest distance between the baseline and monitor receiver locations exceeding 280 m. Moreover, the available monitor data were subject to P-up separation to eliminate the second hydrophone ghost, and have a different frequency spectrum and amplitude scale compared to the baseline pressure data. Therefore, the double-difference time-lapse FWI method, which requires high repeatability, is not applicable to this case study. These issues highlight the challenges in performing robust time-lapse seismic analysis of the Pyrenees data and emphasize the need for advanced methodologies that can properly handle the data complexity.

#### 5.3.2 Preprocessing

To apply our 2D FWI algorithm, we extract a line from the 3D data volume by selecting inline #1722 (marked by the orange and yellow line in Figure 5.2). Both the baseline and monitor data have a relatively high signal-to-noise ratio, so only minimal preprocessing is applied to enhance the useful signal. The three primary preprocessing steps include: (1) noise attenuation to suppress the linear noise in the data; (2) zero-phase low-cut Ormsby filtering to remove low-frequency noise below 3 Hz, and (3) zero-phase low-pass Ormsby filtering to retain frequencies below 12 Hz, which is the maximum frequency used in FWI.

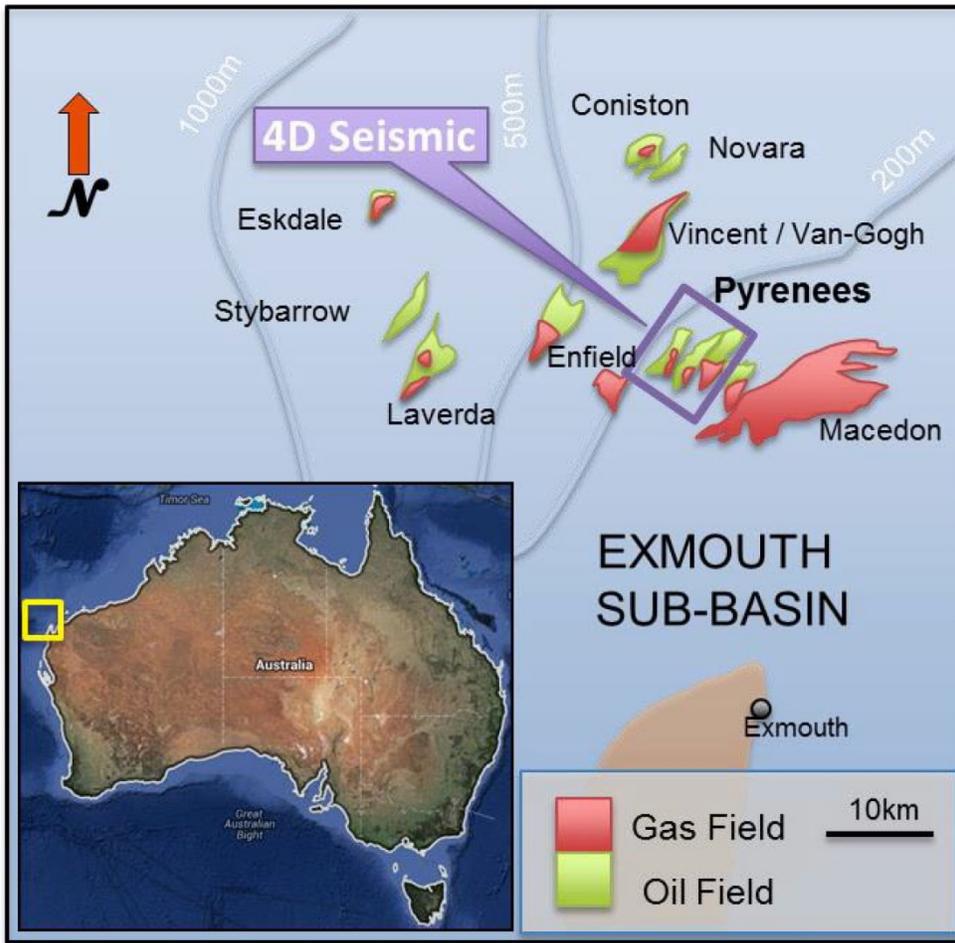


Figure 5.1 Location of the Pyrenees 4D survey (after Cai et al., 2019).

The inclusion of low-frequency data is critical for obtaining accurate FWI results. A careful selection of the lowest frequency in FWI is essential to prevent cycle-skipping and mitigate the influence of noise. For that purpose, low-pass filters with high-cut frequencies of 2, 3, 4, and 5 Hz were applied to denoised raw shot gathers in the time domain (Figure 5.3). For the 2-Hz filter, the signal is somewhat discernible, but the near- and mid-offset traces are significantly contaminated by linear noise. On the other hand, for the 3- and 5-Hz filters, coherent and higher-amplitude signals become observable, which is indicative of an improved signal-to-noise ratio. Therefore, FWI is applied with a low-cut filter that has an edge frequency of 3 Hz. To reduce the computational cost, the maximum frequency was set to 12 Hz.

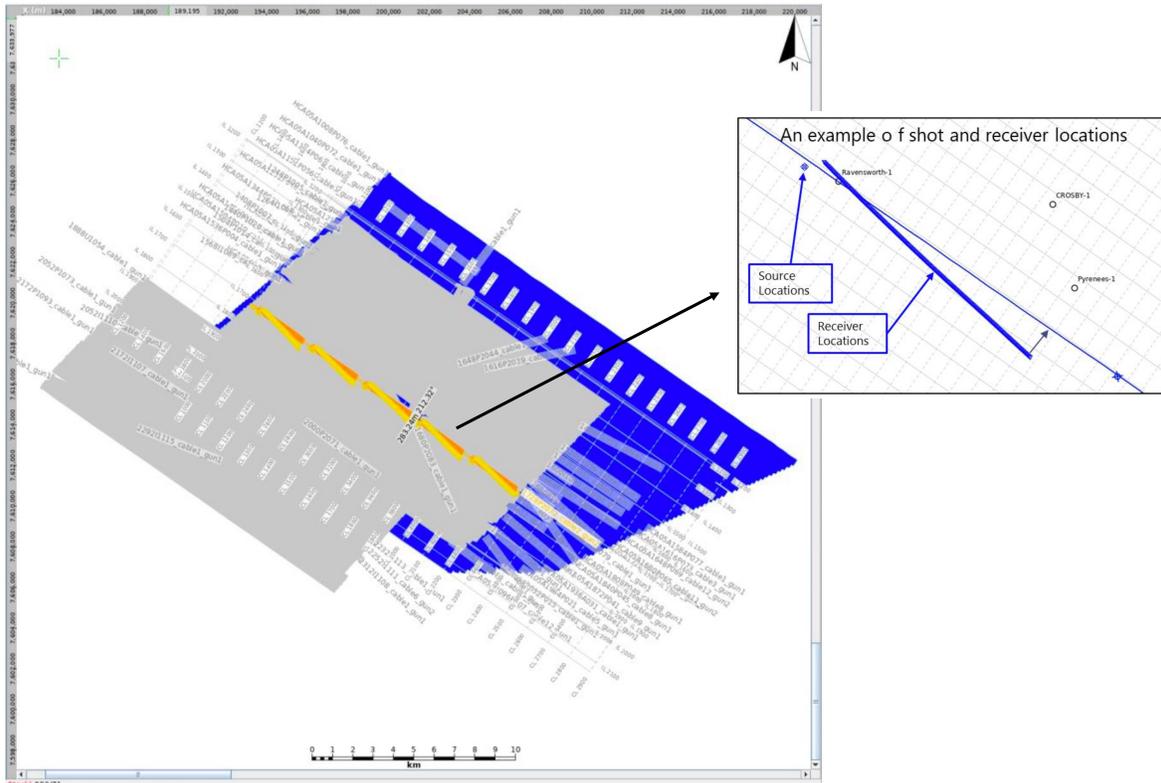


Figure 5.2 Geometries of the time-lapse seismic surveys. The orange and yellow lines mark the 2D baseline and monitor surveys, respectively.

The employed FWI algorithm operates with all wave types recorded with the chosen time window including refracted and reflected arrivals. Multiples and surface ghosts are also used in FWI and simulated by the forward-modeling code. Elasticity and anisotropy (transverse isotropy) are taken into account to make the algorithm more suitable for field data, as discussed below.

### 5.3.3 Initial model

The P-wave vertical velocity ( $V_{P0}$ ) was built based on well logs and the result of reflection tomography for VTI media (Figure 5.4a) in the previous seismic data processing, which are published by Geoscience Australia. The Thomsen parameters  $\epsilon$  and  $\delta$  estimated in the processing are small, with maximum values less than 0.05. The shear-wave vertical velocity ( $V_{S0}$ ) (Figure 5.4b) is obtained by extrapolating the ratio of the P-wave and S-wave velocities in a well, which is about 1 km away from the 2D line (not shown here). The density (Figure 5.4c) is obtained from the P-wave velocity by applying Gardner's (Gardner et al., 1974) equation for sedimentary formations:

$$\rho = 310 V_{P0}^{0.25}, \quad (5.6)$$

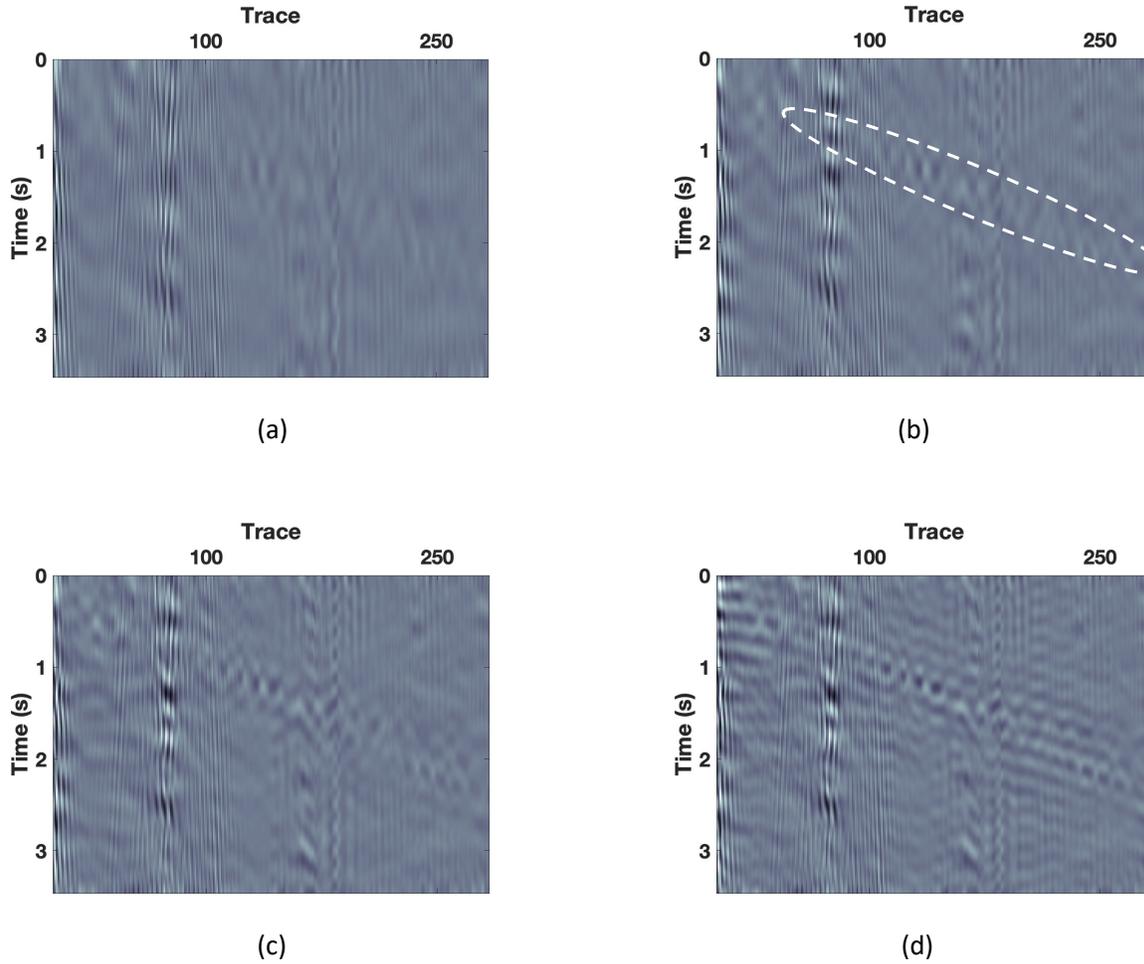


Figure 5.3 Seismic shot gather after low-bandpass filtering with a high cut-off frequency of (a) 2Hz, (b) 3Hz, (c) 4Hz, and (d) 5Hz. The white dashed circle shows the seismic signal.

where  $\rho$  is in  $\text{kg/m}^3$ , and  $V_{P0}$  is in  $\text{m/s}$ . Note that equation 5.6 is used only to compute the initial density model, which is updated during FWI using the corresponding inversion gradient (Liu and Tsvankin, 2022). The Gaussian filter with a kernel of  $62.5 \text{ m} \times 62.5 \text{ m}$  is applied to the parameters in Figure 5.4a-b to obtain the smooth initial model (Figure 5.4d-f).

#### 5.3.4 Wavelet estimation

Accurate estimation of the source signature is of utmost importance in achieving successful FWI results. Errors in the wavelet can hamper the model-updating process (e.g., it can get trapped in a local minimum of the objective function) and cause distortions in the subsurface model (Luo et al., 2014; Warner et al., 2013; Yuan et al., 2014). This issue is particularly critical in time-lapse FWI where the source

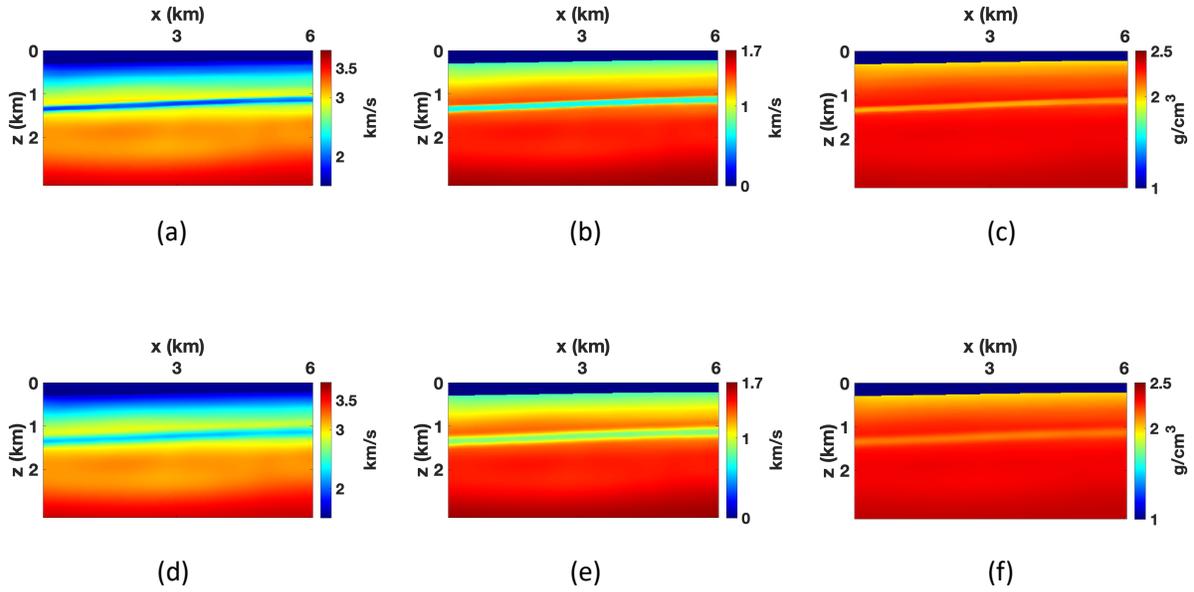


Figure 5.4 Parameters of the Inpex VTI tomography model with a grid size of  $12.5 \times 12.5$  m: (a) the P-wave vertical velocity ( $V_{P0}$ ), (b) the S-wave vertical velocity ( $V_{S0}$ ), and (c) the density ( $\rho$ ). The smoothed initial model for FWI: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ .

wavelet can be different for the baseline and monitor surveys (Liu and Tsvankin, 2022).

A common approach for wavelet estimation involves using water-bottom reflections and direct arrivals, which requires such preprocessing steps as deghosting and multiple removal. Alternatively, Song et al. (1995) and Pratt (1999) suggest estimating the source wavelet during the inversion process. However, their methodology entails additional trade-offs that may compromise the quality of the inversion.

Here, we adopt a method for source-signature estimation proposed by Robinson and Treitel (1967). Their filter-based approach eliminates the need for deghosting and is particularly advantageous for deep-water surveys, in which the direct waves does do not interfere with water-bottom arrivals (e.g., reflections and multiples).

The inversion is initiated with a Ricker wavelet that has a central frequency of 8 Hz. The water velocity model is constructed based on the results of P-wave tomography provided by Inpex. Next, forward modeling is employed to simulate the direct wave using the initial wavelet. The direct arrival is extracted from the nearest-offset (250 m) trace.

Matching of the observed and simulated first arrivals makes it possible to generate a Wiener filter (Figure 5.5b), which serves as a corrective mechanism to update the initial wavelet (Robinson and Treitel, 1967). Finally, the Wiener filter is applied to the initial wavelet, resulting in the wavelet (Figure 5.6b)

employed in FWI.

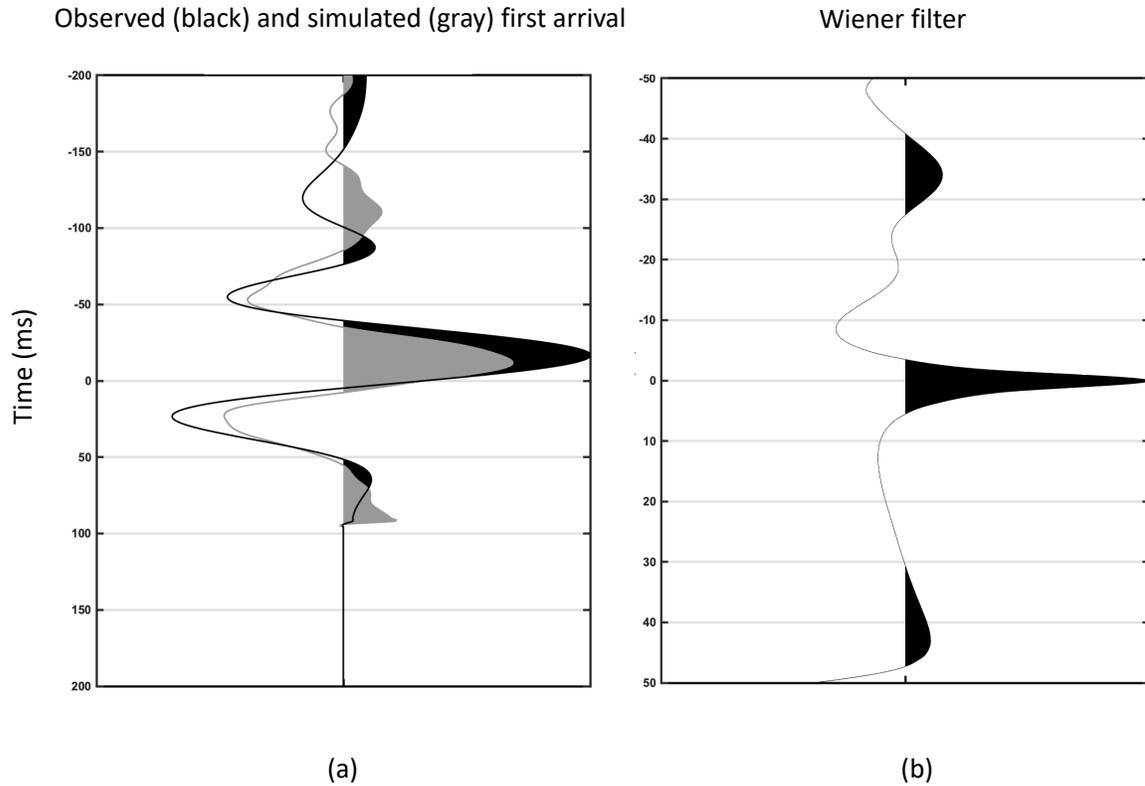


Figure 5.5 (a) Observed and simulated direct wave. (b) The Wiener filter obtained by mapping the simulated data to observed data.

### 5.3.5 Inversion of baseline data

Because the data are acquired in 3D geometry, the sources and receivers for the chosen 2D line set do not belong to a straight line. Therefore, a source/receiver projection onto the line is necessary before performing 2D FWI. The source wavelet for FWI is estimated by the Wiener filter proposed by Robinson and Treitel (1967), as described above. FWI operates in the frequency range from 3-12 Hz. All medium parameters including density are updated independently using the nonlinear conjugate-gradient method.

#### 5.3.5.1 Elastic isotropic FWI

First, we apply elastic isotropic FWI with three different objective functions (equations 5.1-5.3) to the baseline data. As expected, the inverted models (Figure 5.7) have a higher spatial resolution than the initial parameter distributions. The global-correlation (GC) objective function produces a more continuous

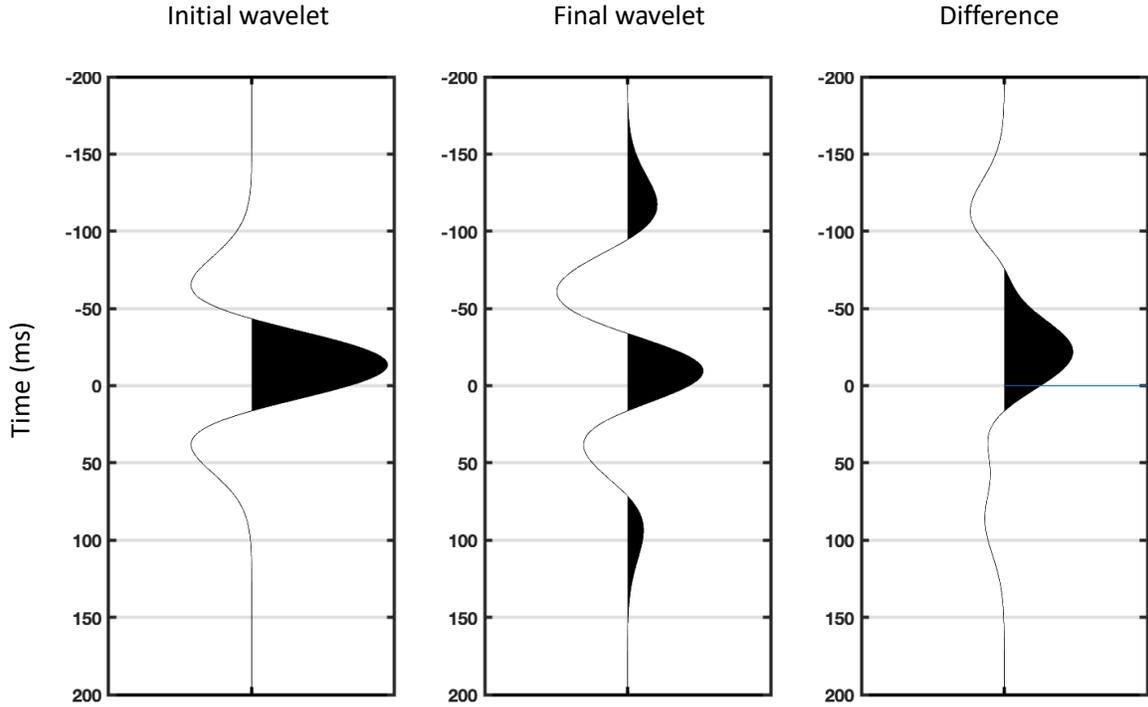


Figure 5.6 (a) Initial guess wavelet, (b) final wavelet after applying the Wiener filter (used for FWI), and (c) the difference.

low-velocity layer at a depth of 1.1 km, which corresponds to the reservoir location (indicated by the white arrows). The high-frequency artifacts below 2 km (circled in white) generated by the  $L_2$ -norm objective function are significantly suppressed by applying the source-independent (SI) and GC functions. The periodic artifacts in the deeper part of the section (below 1.2 km), which are especially strong for the inverted density, are likely caused by the difficulties in matching the multiples.

It should be noted that the initial  $\rho$  model is obtained from Gardner's equation, and there is a larger uncertainty in this parameter compared to the velocities. Furthermore, FWI is carried out without constraints, making it even more challenging to accurately estimate density from the surface seismic data.

The vertical profiles of the velocity  $V_{P0}$  show that the GC objective function produces the velocity that is closer to the sonic log than that estimated by the two other objective functions (Figure 5.8). In particular, both  $L_2$ -norm-based objective functions underestimate the velocity in the overburden, likely due to the difficulty of matching the amplitudes of the observed and simulated data. The inverted models obtained with all three objective functions reduce the data misfit compared to that for the initial model

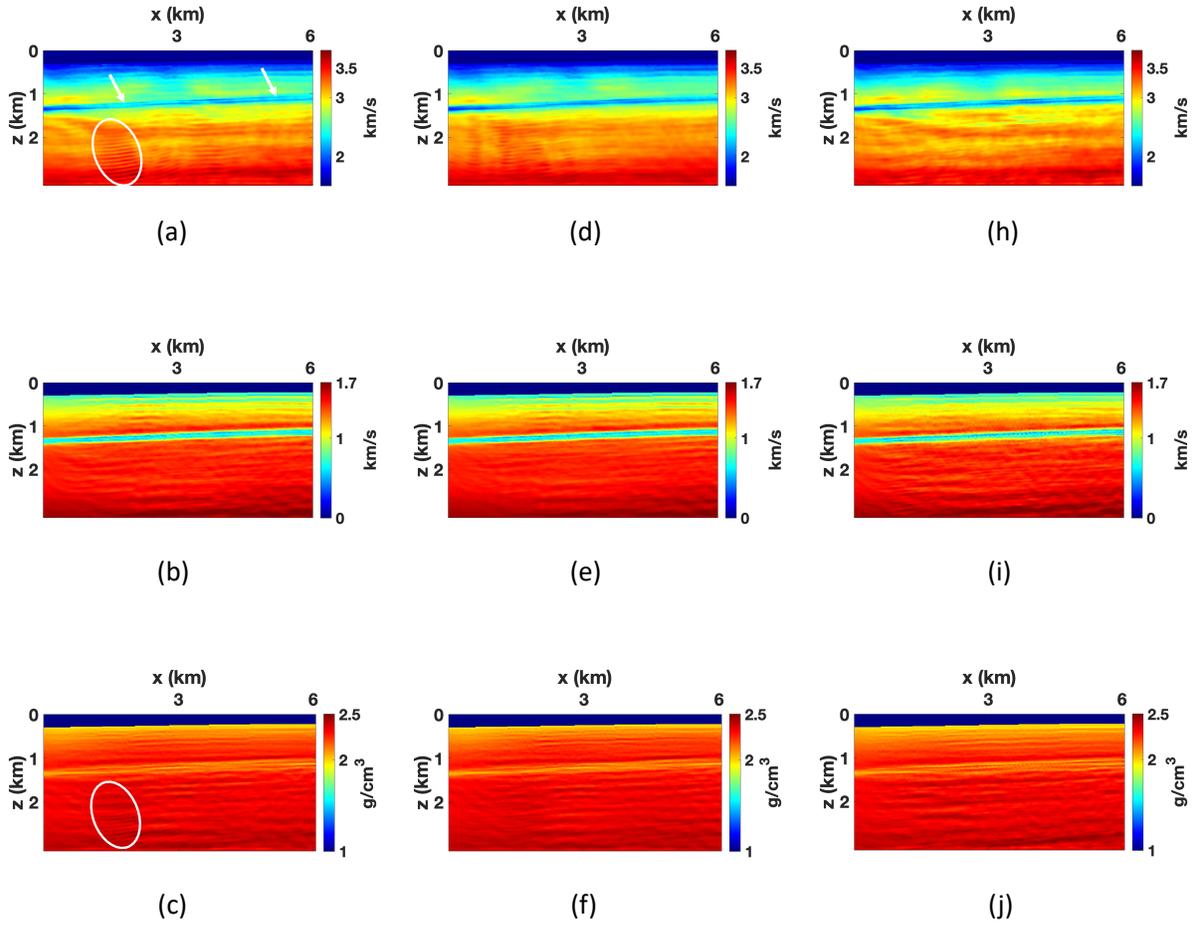


Figure 5.7 Medium parameters estimated by isotropic FWI using different objective functions. The  $L_2$ -norm objective function: (a)  $V_{P0}$ , (b)  $V_{S0}$ , and (c)  $\rho$ ; the source-independent (SI) objective function: (d)  $V_{P0}$ , (e)  $V_{S0}$ , and (f)  $\rho$ ; the global-correlation (GC) objective function: (h)  $V_{P0}$ , (i)  $V_{S0}$ , and (j)  $\rho$ . The arrows point to the low-velocity layer and the ellipses mark high-frequency artifacts.

(Figure 5.9a). The best fit to the data (see the white row), however, is provided by the model reconstructed using the GC objective function.

### 5.3.5.2 Influence of wavelet distortion

The satisfactory results produced by the GC objective function indicate that the estimated wavelet may be sufficiently accurate. To investigate the influence of errors in the wavelet on the FWI results, we distort the central frequency and shape of the filtered wavelet (not shown here).

The inversion of the baseline data shows that both  $L_2$ -norm (Figure 5.10a) and GC (Figure 5.10b) objective functions are significantly influenced by the wavelet errors. FWI perform with these objective functions cannot reconstruct the continuous low-velocity layer, and the inverted parameters are noisy. In

contrast, the SI FWI (Figure 5.10c) estimates the medium parameters with resolution similar to that in the previous test, which is consistent with the results of Liu and Tsvankin (2022). However, application of the SI technique to this data set is hampered by difficulties in amplitude matching, as discussed above.

### 5.3.5.3 Elastic FWI for VTI media

Asaka et al. (2016) have identified the presence of anisotropy (transverse isotropy) during their velocity-model building process for a field, offshore Western Australia. They emphasized the importance of incorporating anisotropic inversion techniques for accurate reservoir characterization. Therefore, next we apply the FWI algorithm for VTI media developed by Liu and Tsvankin (2021, 2022). Because the structure dips in the area are mild, the VTI model should be adequate for our purposes (Figure 5.7).

We parameterize vertical transverse isotropy media by  $V_{P0}$  (P-wave vertical velocity),  $V_{S0}$  (S-wave vertical velocity),  $V_{\text{hor,P}}$  (P-wave horizontal velocity),  $V_{\text{nmo,P}}$  (P-wave normal-moveout velocity from a horizontal reflector), and  $\rho$  (density) (Kamath et al., 2017). The velocities  $V_{\text{hor,P}}$  and  $V_{\text{nmo,P}}$  are expressed through the Thomsen parameters  $\epsilon$  and  $\delta$  as follows (Thomsen, 1986; Tsvankin, 2012):

$$V_{\text{hor,P}} = V_{P0} \sqrt{1 + 2\epsilon}, \quad (5.7)$$

$$V_{\text{nmo,P}} = V_{P0} \sqrt{1 + 2\delta}. \quad (5.8)$$

Because reflection tomography published by Geoscience Australia produced  $\epsilon$  and  $\delta$  with a maximum value close to 0.04, the initial model is isotropic. We employ the GC objective function, which provided the highest accuracy in the isotropic inversion discussed above. The anisotropic (VTI) FWI algorithm produces a more continuous low-velocity layer (marked by the white arrow) in the sections of all four velocities along with reduced artifacts below the reservoir (Figure 5.11). We also computed the Thomsen parameters  $\epsilon$  and  $\delta$  from the inverted velocities using equations 5.7 and 5.8, respectively (Figure 5.12). Our results show that the maximum value of  $\epsilon$  is about 0.15 and  $\delta$  around 0.085, which is consistent with the observations of Asaka et al. (2016), but indicates that the previous tomographic model likely underestimates the magnitude of anisotropy.

Taking anisotropy into account yields a more accurate P-wave velocity field. Indeed, the velocity  $V_{P0}$  reconstructed by the VTI algorithm closely matches the sonic log, with higher values above and within the reservoir and lower values below it (Figure 5.13). The comparison of the observed and simulated data (Figure 5.14) shows a good agreement for shallow events, especially at the far offsets (see the white arrows), where the offset-to-depth ratio is close to 1.5. The VTI algorithm does not generate significant parameter updates below 2 s, probably due to the limited offset-to-depth ratio.

The inverted velocity  $V_{P0}$  deviates from the sonic log at depths from 900-1000 m (Figure 5.13). As mentioned above, FWI was applied using the frequency range from 3-12 Hz. Including higher frequencies in the inversion may improve the accuracy of the inverted parameters, especially in the shallow part of the model. Therefore, we increase the highest frequency in VTI FWI to 19 Hz; the initial model is obtained by FWI using the top frequency of 12 Hz. Although the change in  $V_{P0}$  is not significant (Figure 5.15a and d), there is a larger update in the anisotropy parameters  $\epsilon$  and  $\delta$  (Figure 5.15b and c). As a result, the match between the observed and simulated data (Figure 5.16) is substantially improved, especially for zero-offset times less than 2 s.

### 5.3.6 Inversion of monitor data

As mentioned above, the available monitor data were subject to P-up separation, which changes the frequency spectrum. The central frequency of the monitor data is lower than that of the baseline data (Figure 5.17), especially at the near and far offsets. In contrast to the baseline data, the frequency of the monitor data is the highest in the mid-offset range (from traces 90 to 150).

We employ the sequential-difference (SD) strategy for time-lapse analysis because of its efficiency and robustness compared to the parallel-difference method (Liu and Tsvankin, 2021). Because of the inconsistent frequency content of the monitor data at different offsets, the field and modeled data are not likely to match if only one frequency band (e.g., 3-12 Hz) is used in FWI. Therefore, a multiscale inversion approach is implemented with three frequency bands: 3-5 Hz, 3-9 Hz, and 3-12 Hz. The multiscale inversion also helps mitigate the cycle-skipping issue common in FWI.

Figure 5.18 shows the time-lapse variations of the velocity  $V_{P0}$  and parameters  $\epsilon$  and  $\delta$  estimated by the SD strategy. “Sandwich”-shape 4D changes are observed in the velocity  $V_{P0}$  obtained for the frequency range 3-5 Hz (Figure 5.18a-c), where a low-velocity layer is embedded between two high-velocity horizons (marked by the black arrow). The main 4D response is the velocity decrease (softening) of the reservoir likely caused by gas coming out of solution. The velocity increase (hardening) is also observed above the reservoir, which could be due to oil replacing gas (Cai et al., 2019; Duncan et al., 2015).

A similar pattern is observed in the variations of the parameters  $\epsilon$  and  $\delta$ . However, false 4D anomalies appear in  $V_{P0}$  below 1.5 km and in  $\epsilon$  around the water bottom (marked by the dashed black rectangles). Although these artifacts are less pronounced for the frequency range 3-9 Hz (Figure 5.18d-f), they get amplified when the maximum frequency goes up to 12 Hz (Figure 5.18h-j). This enhancement of artifacts is likely due to the forward modeling algorithm in FWI being unable to simulate the inconsistent frequency content of the monitor data.

## 5.4 Conclusions

We applied a previously developed 4D FWI methodology for VTI media to a field data set from the Pyrenees field in Australia, which involved baseline and monitor surveys with different geometries and frequency spectra. Inversion of the baseline data showed that the global-correlation objective function produces better results than the  $L_2$ -norm and source-independent (SI) objective functions which are sensitive to difficulties in amplitude matching. The superior performance of the global-correlation objective function was ensured by the sufficiently accurate estimation of the source wavelet. If the wavelet is significantly distorted, FWI should be performed with the SI objective function. Taking vertical transverse isotropy into account led to a better match between the sonic log and the P-wave vertical velocity and reduced the misfit between the observed and simulated data, particularly at the far offsets. Increasing the highest frequency in FWI from 12 to 19 Hz improves the match between the observed and simulated data, especially at shallow depths (for zero-offset times less than 2 s).

We also presented preliminary results of applying time-lapse FWI to the monitor data. The time-lapse inversion was hampered by nonrepeatability issues and the inconsistent frequency content of the monitor records. Still, application of the sequential-difference time-lapse strategy revealed the velocity variations inside and around the reservoir caused by hydrocarbon production. The time-lapse results indicate that the main 4D response is the softening of the reservoir caused by gas coming out of the solution and the hardening effect above the reservoir caused by the replacement of gas with oil.

## 5.5 Acknowledgments

We thank the members of the Anisotropy-Team at the Center for Wave Phenomena (CWP) in the Geophysics Department at Colorado School of Mines for useful discussions. This work was supported by a joint project with INPEX and by the Consortium Project on Seismic Inverse Methods for Complex Structures at CWP.

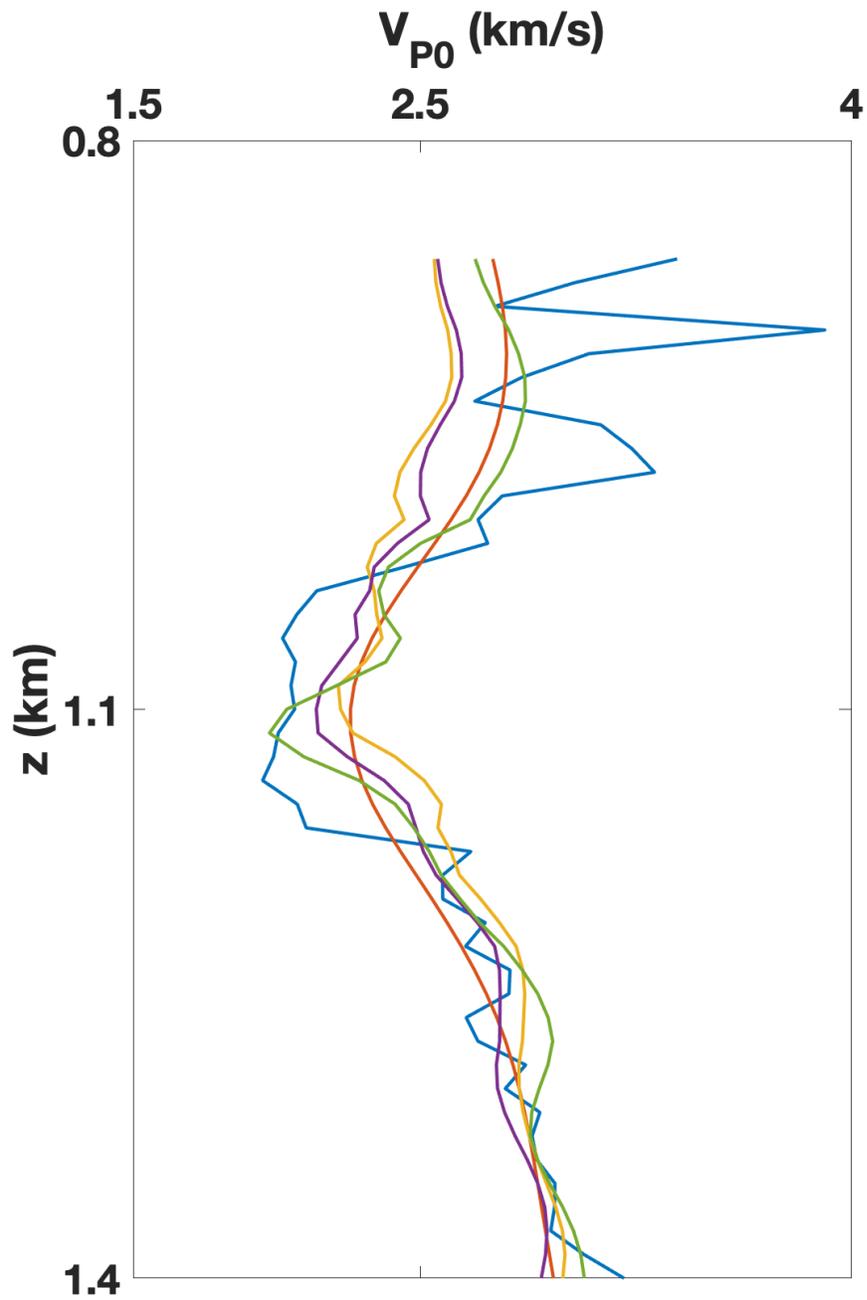


Figure 5.8 Vertical profiles of the P-wave vertical velocity  $V_{P0}$ . The velocity obtained from the well logs is marked by the blue line and the initial velocity used in FWI by the red line. The inversion results are marked by the yellow ( $L_2$ -norm objective function), violet (SI objective function), and green (GC objective function) lines.

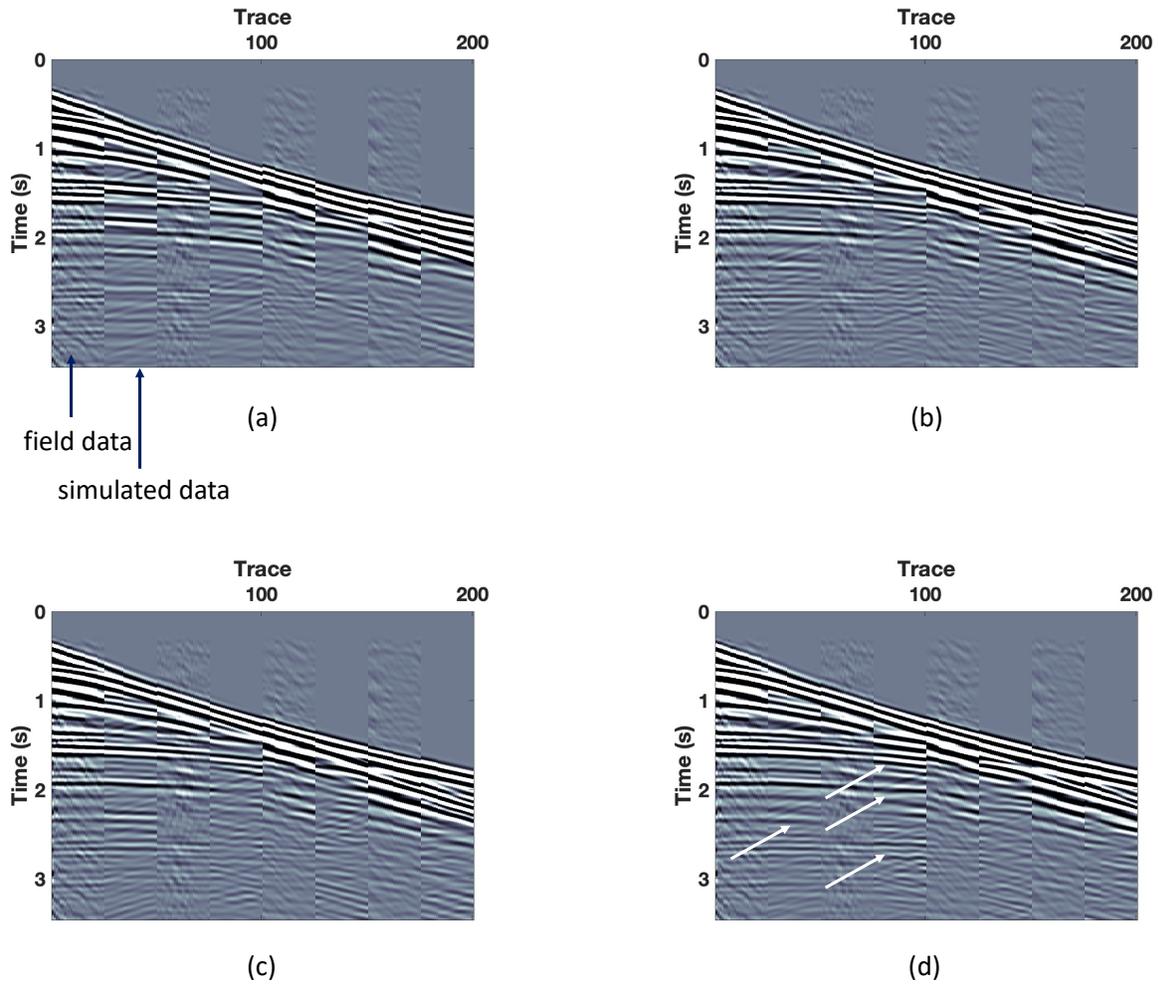


Figure 5.9 Quality of data matching for shot 144 at  $x \approx 1$  km. The recorded traces (“field data”) are interleaved between the traces simulated using (a) the initial model and (b, c, d) the models obtained by the isotropic FWI applied with the (b)  $L_2$ -norm objective function, (c) source-independent objective function, and (d) global-correlation objective function. The arrows point to the areas where the GC objective function produces a better match with the field data.

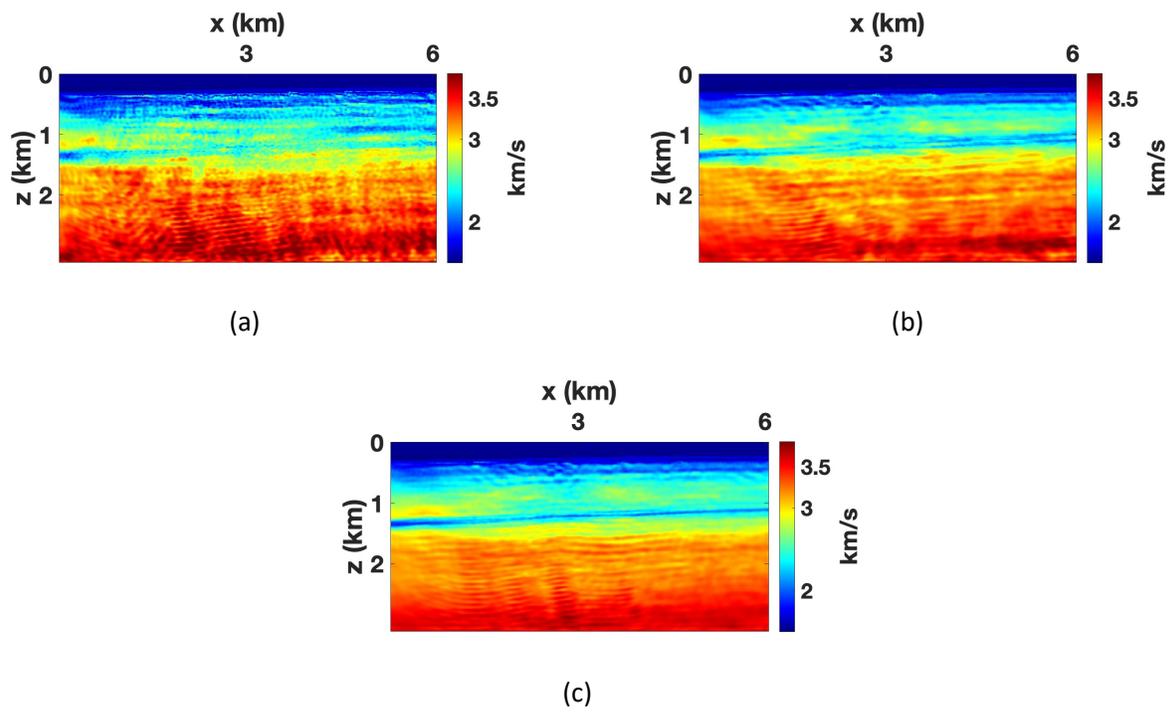


Figure 5.10 P-wave velocity ( $V_{P0}$ ) estimated by isotropic FWI with a distorted wavelet using different objective functions: (a) the  $L_2$ -norm objective function, (b) the global-correlation (GC) objective function, and (c) the source-independent (SI) objective function.

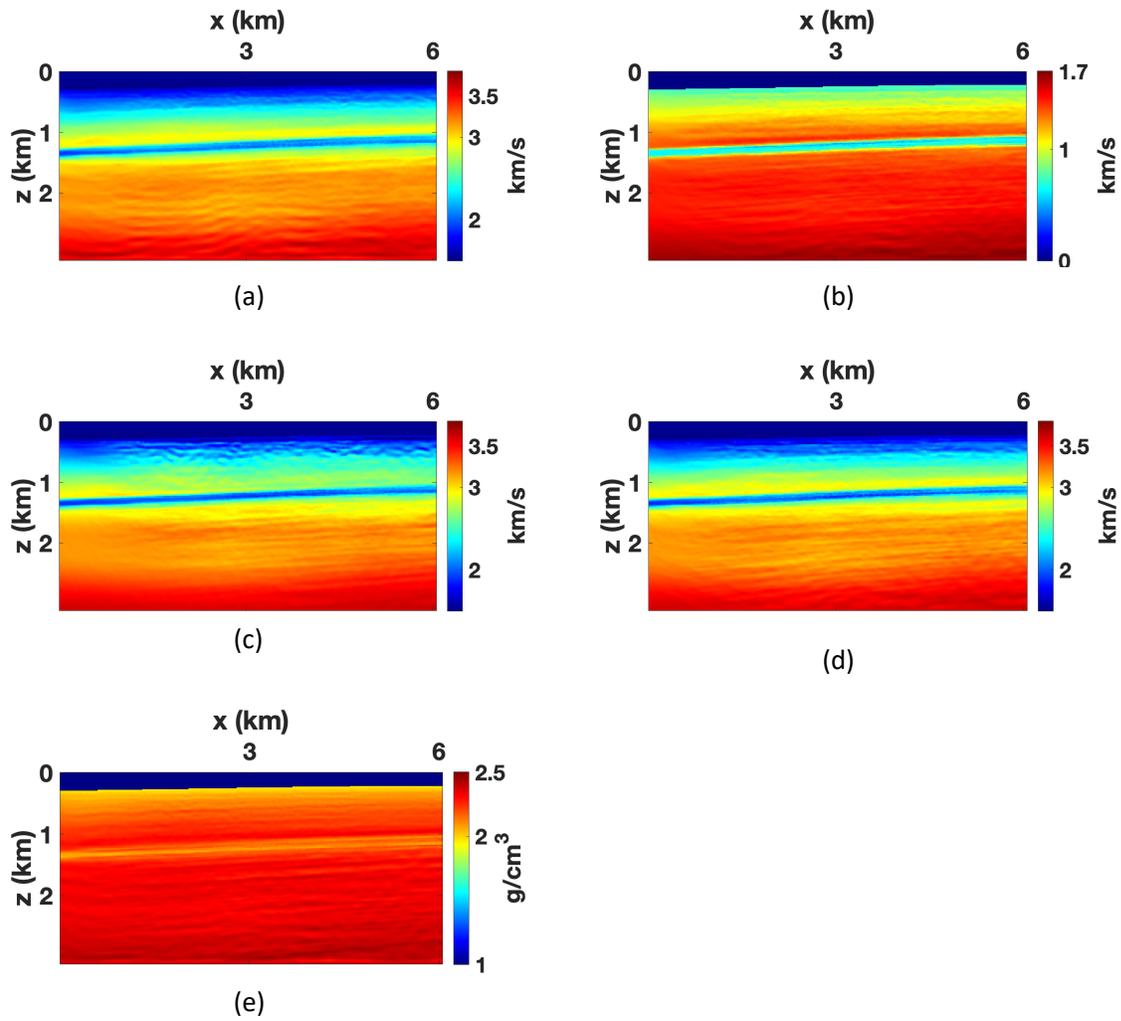


Figure 5.11 Medium parameters estimated by the VTI FWI algorithm with the global-correlation objective function: (a)  $V_{P0}$ , (b)  $V_{S0}$ , (c)  $V_{hor,P}$ , (d)  $V_{nmo,P}$ , and (e)  $\rho$ .

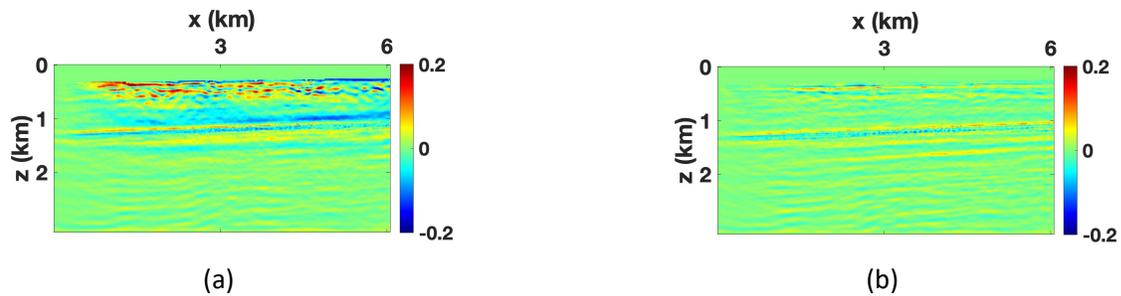


Figure 5.12 Thomsen parameters obtained from the inverted velocities (Figure 5.11) using equations 5.7 and 5.8: (a)  $\epsilon$  and (b)  $\delta$ .

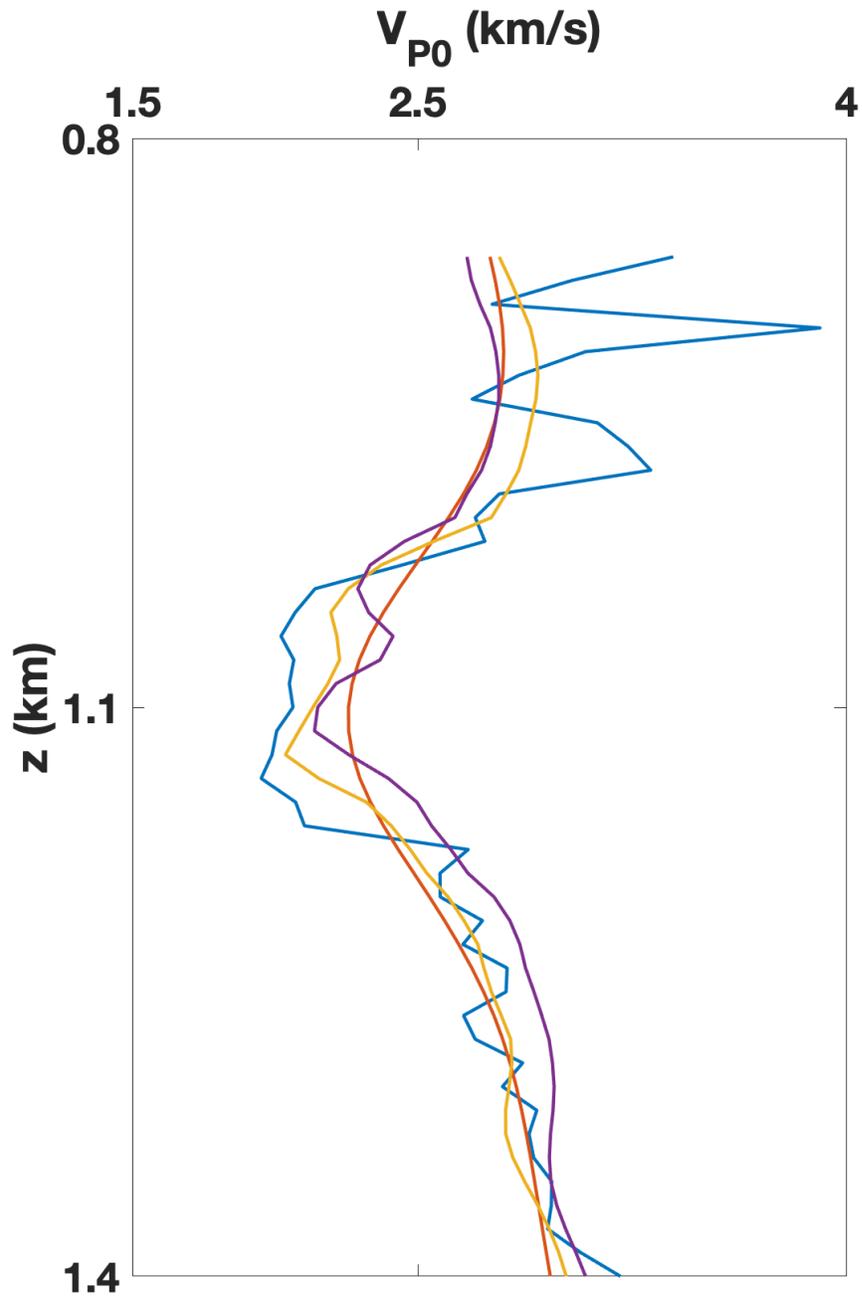


Figure 5.13 Vertical profiles of the P-wave vertical velocity  $V_{P0}$ . The velocity obtained from the sonic log is marked by the blue line and the initial velocity used in FWI by the red line. The inversion results are marked by the yellow (VTI FWI) and violet (isotropic FWI) lines.

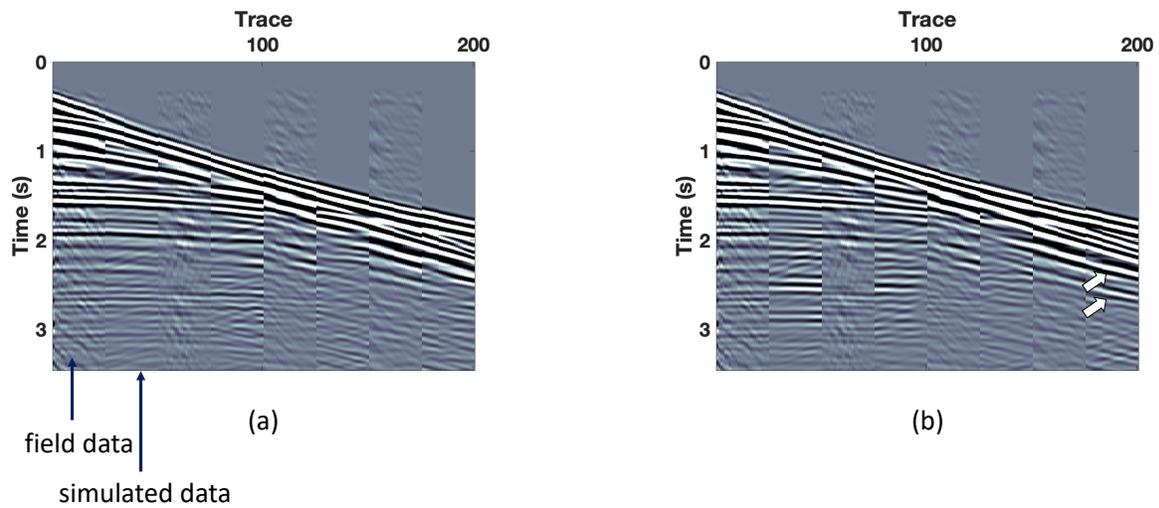


Figure 5.14 Quality of data matching for shot 144 at  $x \approx 1\text{km}$ . The recorded traces are interleaved between the traces simulated using the models obtained by the (a) isotropic FWI and (b) VTI FWI, both applied with the GC objective function. The arrows point to the areas where the VTI algorithm produces a better match with the field data.

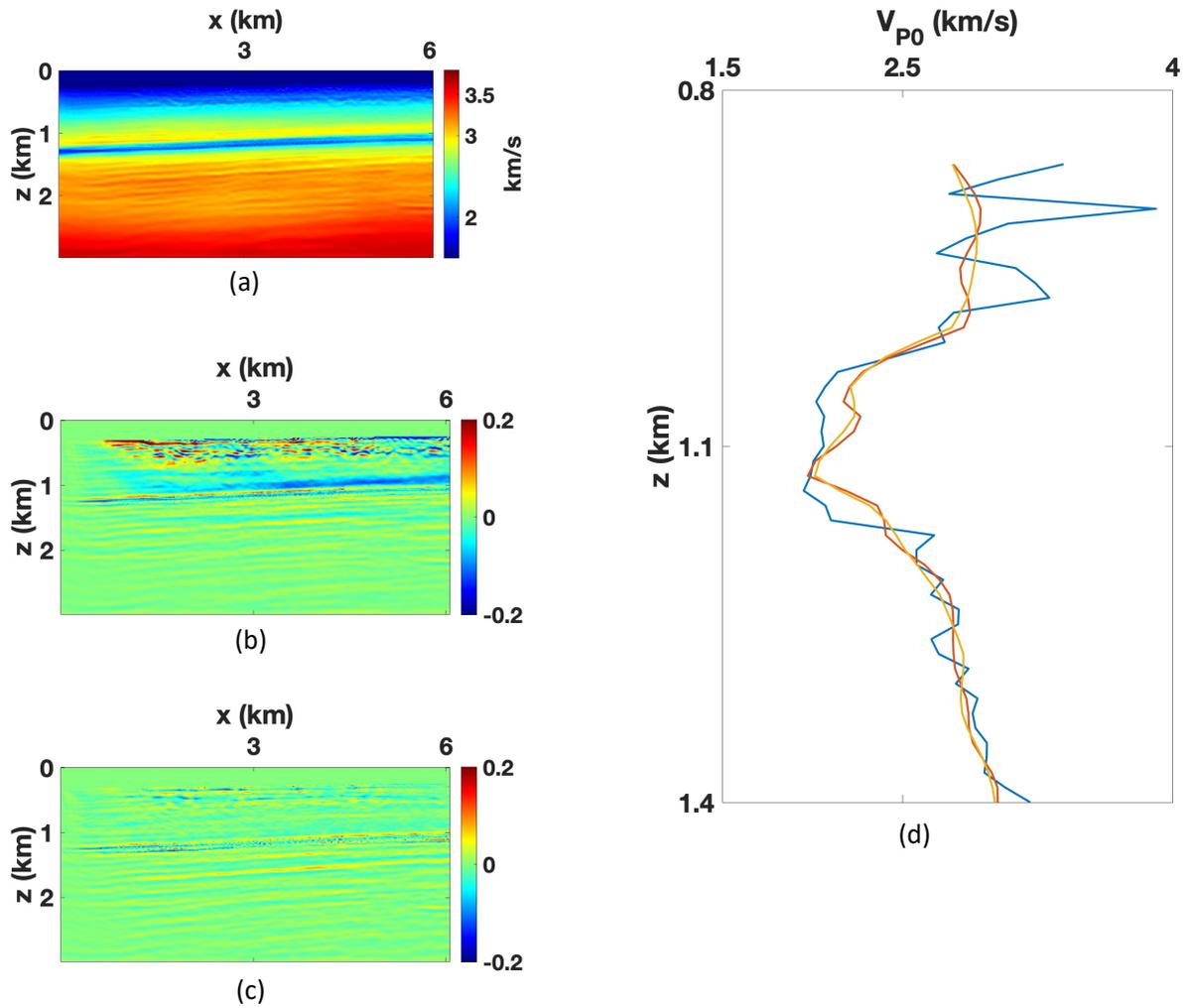


Figure 5.15 Medium parameters estimated by the VTI FWI algorithm with the highest frequency equal to 19 Hz: (a)  $V_{P0}$ , (b)  $\epsilon$ , and (c)  $\delta$ . Plot (d) shows the vertical profiles of the P-wave vertical velocity  $V_{P0}$ . The velocity obtained from the sonic log is marked by the blue line. The inversion results obtained using the highest frequencies equal to 12 Hz and 19 Hz are marked by the yellow and orange lines, respectively.

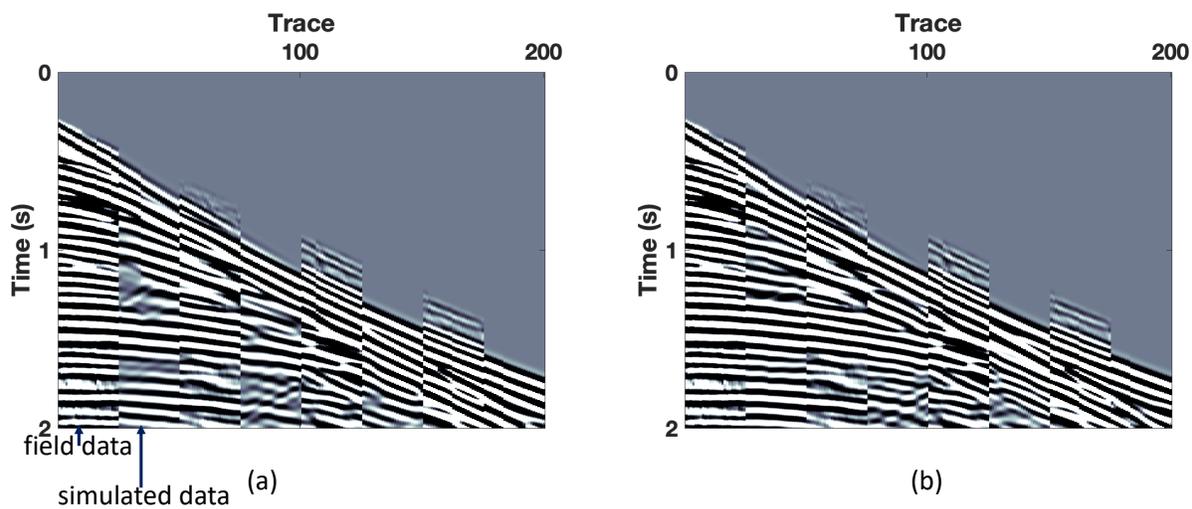


Figure 5.16 Quality of data matching for shot 144 at  $x \approx 1$  km. The recorded traces are interleaved between the traces simulated using the models obtained by the VTI FWI with the GC objective function using the highest frequencies equal to (a) 12 Hz and (b) 19 Hz.

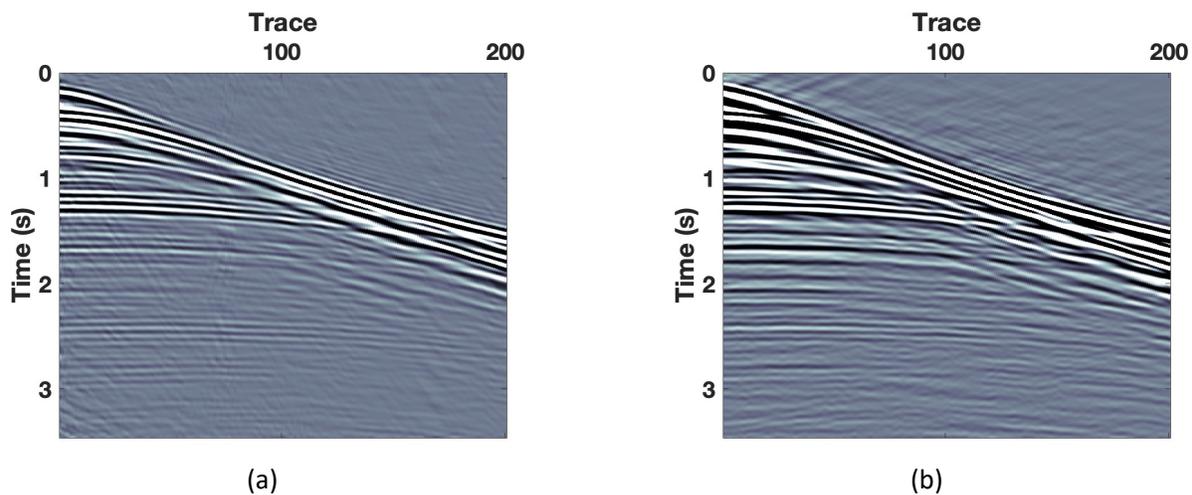


Figure 5.17 Seismograms for shot 144 after applying Ormsby filtering (3-12 Hz) to the (a) baseline and (b) monitor data.

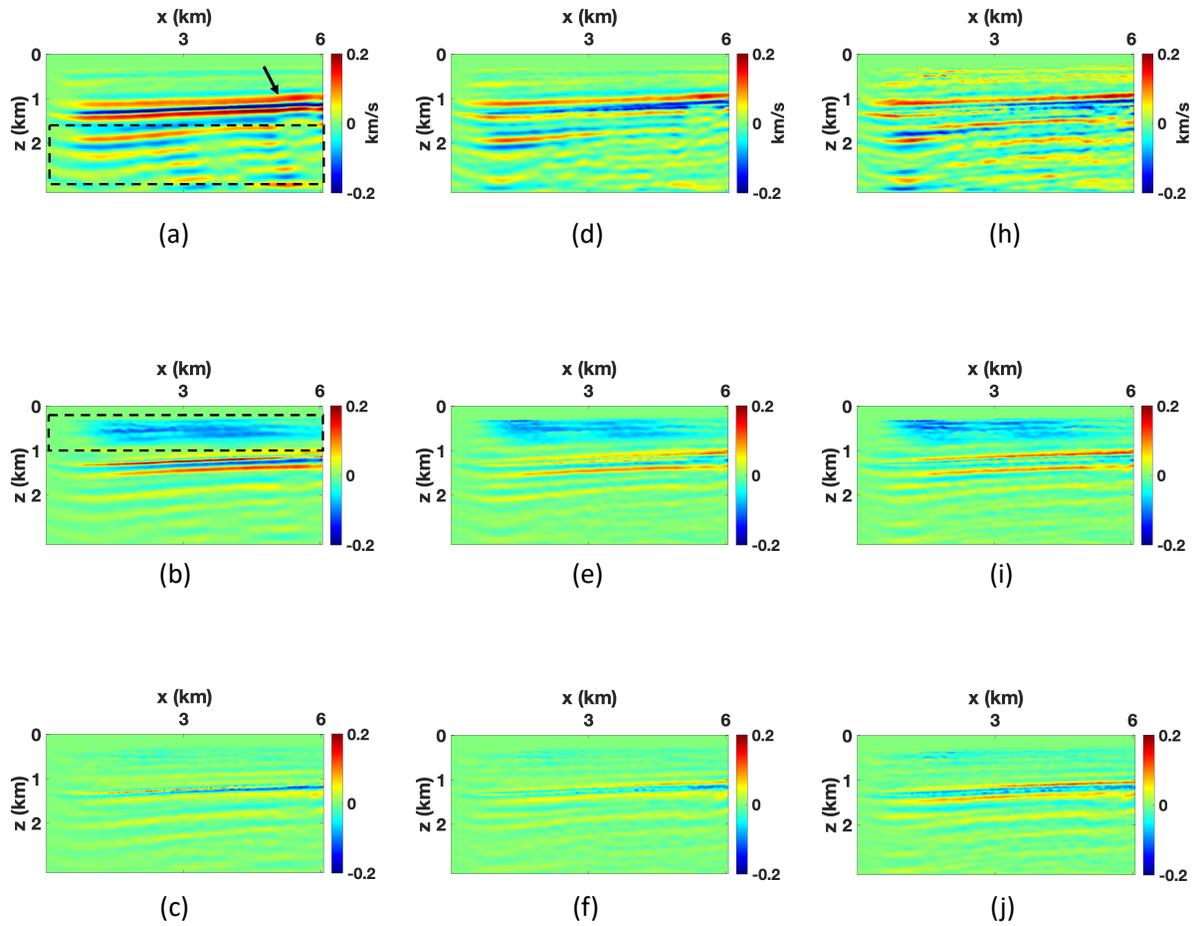


Figure 5.18 Time-lapse parameter variations obtained by the VTI FWI algorithm using a multiscale approach. The results for the frequency bands 3-5 Hz [(a)  $V_{P0}$ , (b)  $\epsilon$ , and (c)  $\delta$ ], 3-9 Hz [(d)  $V_{P0}$ , (e)  $\epsilon$ , and (f)  $\delta$ ], and 3-12 Hz [(h)  $V_{P0}$ , (i)  $\epsilon$ , and (j)  $\delta$ ]. The arrows point to the “sandwich”-shape 4D response.

## CHAPTER 6

# JOINT PHYSICS-BASED AND DATA-DRIVEN TIME-LAPSE SEISMIC INVERSION: MITIGATING DATA SCARCITY

A paper published\* in *Geophysics*

Yanhua Liu<sup>1,2,3</sup>, Shihang Feng<sup>2</sup>, Ilya Tsvankin<sup>3</sup>, David Alumbaugh<sup>4</sup>, and Youzuo Lin<sup>2</sup>

In carbon capture and sequestration (CCS), developing rapid and effective imaging techniques is crucial for real-time monitoring of the spatial and temporal dynamics of CO<sub>2</sub> propagation during/after injection. With continuing improvements in computational power and data storage, data-driven techniques based on machine learning (ML) have been effectively applied to seismic inverse problems. In particular, ML helps alleviate the ill-posedness and high computational cost of full-waveform inversion (FWI). However, such data-driven inversion techniques require massive high-quality training data sets to ensure prediction accuracy, which hinders their application to time-lapse monitoring of CO<sub>2</sub> sequestration. We propose an efficient “hybrid” time-lapse workflow that combines physics-based FWI and data-driven ML inversion. The scarcity of the available training data is addressed by developing a new data-generation technique with physics constraints. The method is validated on a synthetic CO<sub>2</sub>-sequestration model based on the Kimberlina storage reservoir in California. The proposed approach is shown to synthesize a large volume of high-quality, physically realistic training data, which is critically important in accurately characterizing the CO<sub>2</sub> movement in the reservoir. The developed hybrid methodology can also simultaneously predict the variations in velocity and saturation and achieve high spatial resolution in the presence of realistic noise in the data.

### 6.1 Introduction

The importance of CO<sub>2</sub> sequestration has increased in the past decade due to the global warming caused by greenhouse gases. Rapidly imaging and monitoring the spatial and temporal dynamics of CO<sub>2</sub> migration ensures that the supercritical CO<sub>2</sub> is injected at the correct location and does not leak out of the reservoir (e.g., Lumley, 2010; Pevzner et al., 2017). Time-lapse seismic can represent an effective tool for monitoring CO<sub>2</sub> injected into an aquifer (Ajo-Franklin et al., 2013; Furre et al., 2017; Pevzner et al., 2017; Raknes et al., 2015; Zhang et al., 2012).

---

\*Reprinted with permission of *Geophysics*, 2023, 88, no. 1, K1-K12.

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Earth and Environmental Sciences Division, Los Alamos National Laboratory

<sup>3</sup>Center for Wave Phenomena, Colorado School of Mines

<sup>4</sup>Earth and Environmental Sciences, Lawrence Berkeley National Laboratory

Full-waveform inversion (FWI) of seismic data has been widely utilized for velocity analysis and, under favorable circumstances, reservoir characterization (Asnaashari et al., 2015; Fabien-Ouellet et al., 2017; Singh et al., 2018; Vigh et al., 2014). FWI can potentially provide estimates of the time-lapse parameter variations with high spatial resolution. Time-lapse FWI (TLFWI) strategies proposed for monitoring hydrocarbon production and CO<sub>2</sub> injection (Asnaashari et al., 2015; Liu and Tsvankin, 2021; Queiβer and Singh, 2013) include the parallel-difference, sequential-difference, and double-difference techniques. The parallel-difference method (Plessix et al., 2010) uses the same initial model for the baseline and monitor inversions, whereas the sequential-difference method (Asnaashari et al., 2012) inverts the baseline data to build the initial model for the monitor inversion. The double-difference technique (Denli and Huang, 2009; Lin and Huang, 2015; Watanabe et al., 2004) estimates the time-lapse parameter variations by inverting the difference between the monitor and baseline data sets. Liu and Tsvankin (2021, 2022) extend these time-lapse FWI techniques to realistic elastic anisotropic media. Huang and Zhu (2020) implement uncertainty estimation to the acoustic time-lapse FWI results.

However, FWI suffers from ill-posedness, cycle-skipping, and high computational cost because of the nonlinearity of the inverse problem (Feng et al., 2021; Feng and Schuster, 2019; Virieux and Operto, 2009; Wu and Lin, 2019a; Zhang and Gao, 2021; Zhang and Alkhalifah, 2020). In addition, FWI cannot directly estimate the CO<sub>2</sub> saturation, though empirical and statistical petrophysics relationships have to be employed as a post inversion step to reconstruct the saturation distribution from the inverted velocity field (Ali and Al-Shuhail, 2018; Grana and Rossa, 2010).

Recently, with the increased computational power and the revitalization of deep neural networks, significant research and development efforts have been put into data-driven ML methods for seismic imaging (ArayaPolo et al., 2017; Cao and Roy, 2017; Dahlke et al., 2016; Hale, 2013b; Li et al., 2021; Yuan et al., 2019). For example, Um et al. (2022) propose a network (U-net) to reconstruct CO<sub>2</sub> saturation and compare the results of the uncertainty analysis with the Monte-Carlo dropout and bagging methods. Such methods aim to approximate the inversion operator by neural networks based on the universal approximation theorem (Zhang and Lin, 2020). Data-driven techniques can potentially produce a more accurate reconstruction of subsurface structure than physics-based FWI in a significantly shorter time, which is essential for real-time monitoring.

However, the accuracy of data-driven inversion depends on the generalization ability of the employed networks and the quality and volume of the available training data (Yang et al., 2021; Zhang and Gao, 2021; Zhang and Alkhalifah, 2020). The generalization ability refers to the performance of a network on testing data sets that are significantly different from the training data. Considering that the medium outside the reservoir does not substantially change during injection, the distribution shift between the

testing monitor data and the training data should not be a significant issue in time-lapse monitoring.

However, obtaining high-quality training data can be challenging for typical CO<sub>2</sub> sequestration projects. To address this data scarcity problem, Renán et al. (2020) develop an active learning strategy by incorporating the wave equation to expand the training data set. Yang et al. (2021) present a convolutional neural network to augment training data using such physics information as the wave equation and time-lapse variations of P-wave velocity. Despite some encouraging results, these two methods still require extensive training data in advance, which limits their applicability. Yuan et al. (2019) randomly perturb the inverted baseline model of the reservoir produced by physics-based FWI to generate training data that can be used to predict the time-lapse velocity changes. However, because the data generation is random, the training data set can be physically unrealistic and results in incorrect predictions.

We begin by discussing the methodology of physics-based and data-driven inversions and then outline their application to time-lapse seismic data. To overcome the aforementioned issues related to data scarcity, we propose a “hybrid” time-lapse strategy that combines physics-based FWI with data-driven inversion (called “InvNet-VelSat”). The developed data-generation algorithm simulates high-quality velocity and saturation training models utilizing the available physics information and prior knowledge (i.e., the inverted baseline model and well logs). The data-generation method and the proposed hybrid approach are tested on the Kimberlina reservoir model, assuming that only the baseline and one monitor data set are available. The reconstructed time-lapse variations are compared with those produced by the pure physics-based and data-driven techniques. The generalization ability and robustness of our algorithm is validated by applying it to other monitor surveys and noisy seismic data. Finally, we discuss the impact of the prior information and the quality of the simulated training data on the reconstructed time-lapse variations.

## 6.2 Review of physics-based and data-driven inversions

### 6.2.1 Physics-based full-waveform inversion

Full-waveform inversion iteratively minimizes the difference between the observed and simulated data in the process of updating the subsurface model. Acoustic models may represent a suitable first-order approximation for CO<sub>2</sub> monitoring because shear-wave velocity is not expected to be influenced by supercritical CO<sub>2</sub> (Carcione et al., 2006; Mavko and Mukerji, 1998; Queiβer and Singh, 2010). For acoustic isotropic media, FWI operates with shot gathers of the observed ( $\mathbf{p}^{\text{obs}}$ ) and simulated ( $\mathbf{p}^{\text{sim}}$ ) pressure field.

The L<sub>2</sub>-norm FWI objective function  $S(\mathbf{m})$  is then defined as (e.g., Tarantola, 1984):

$$S(\mathbf{m}) = \frac{1}{2} \|\mathbf{p}^{\text{sim}}(\mathbf{m}) - \mathbf{p}^{\text{obs}}\|^2, \quad (6.1)$$

where  $\mathbf{m}$  includes the gridded P-wave velocity and density models.

Due to the high nonlinearity of FWI, the model-updating procedure can be trapped in local minima. Therefore, the initial model needs to be in the basin of convergence near the global minimum.

The acoustic wave equation used here to simulate seismic data has the form:

$$\frac{1}{V_p^2(\mathbf{x})} \frac{\partial^2 p(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} - \nabla^2 p(\mathbf{x}, t; \mathbf{x}_s) = \delta(\mathbf{x} - \mathbf{x}_s) f(t), \quad (6.2)$$

where  $V_p$  is the P-wave velocity,  $p$  is pressure wavefield,  $f(t)$  is the source signal, “ $\nabla^2$ ” denotes the Laplacian operator, and  $\mathbf{x}_s$  and  $\mathbf{x}$  are the source and receiver locations, respectively. The gradient of the objective function with respect to the model parameters is computed from the adjoint-state method. A nonlinear conjugate-gradient algorithm is employed for updating the velocity:

$$V_p^{k+1} = V_p^k - \alpha_k \nabla S(V_p^k), \quad (6.3)$$

where  $\alpha_k$  is the step length at the  $k$ th iteration.

### 6.2.2 Data-driven InvNet-VelSat

In contrast to physics-based FWI methods, data-driven inversion aims to form a pseudoinverse operator  $F$ , which performs mapping from the input to the model (target) domain (i.e., from the shot gathers  $\mathbf{p}$  to the medium parameters  $\mathbf{m}$ )

$$\mathbf{m} \approx F(\mathbf{p}). \quad (6.4)$$

The universal approximation theorem (Hornik et al., 1990) allows the neural network to approximate most complex functions given sufficient training data. In particular, InversionNet is an end-to-end convolutional neural network (CNN) developed by Wu and Lin (2019a) and based on the above theorem. The input is the recorded pressure and the output is the P-wave velocity model. The CNN consists of an encoder and a decoder. The encoder contains five convolution blocks followed by maximum pooling to extract high-level features (representing the input) from the pressure data and significantly reduce the data dimension. Each convolution block includes a convolution layer, batch normalization, and a LeakyReLU activation function. The decoder, which comprises five deconvolution blocks, translates these features into the velocity model.

Since saturation is closely related to CO<sub>2</sub> diffusion in the subsurface, we propose a modified network (InvNet-VelSat, Figure 6.1) to simultaneously predict both velocity and CO<sub>2</sub> saturation. The structures of the encoder and decoder for the velocity prediction in InversionNet are incorporated into InvNet-VelSat. The decoder added for predicting the saturation is parallel to that for the velocity prediction. The Mean Absolute Error (MAE) is employed to measure the training and testing loss and to preserve sharp

boundaries in the predicted models.

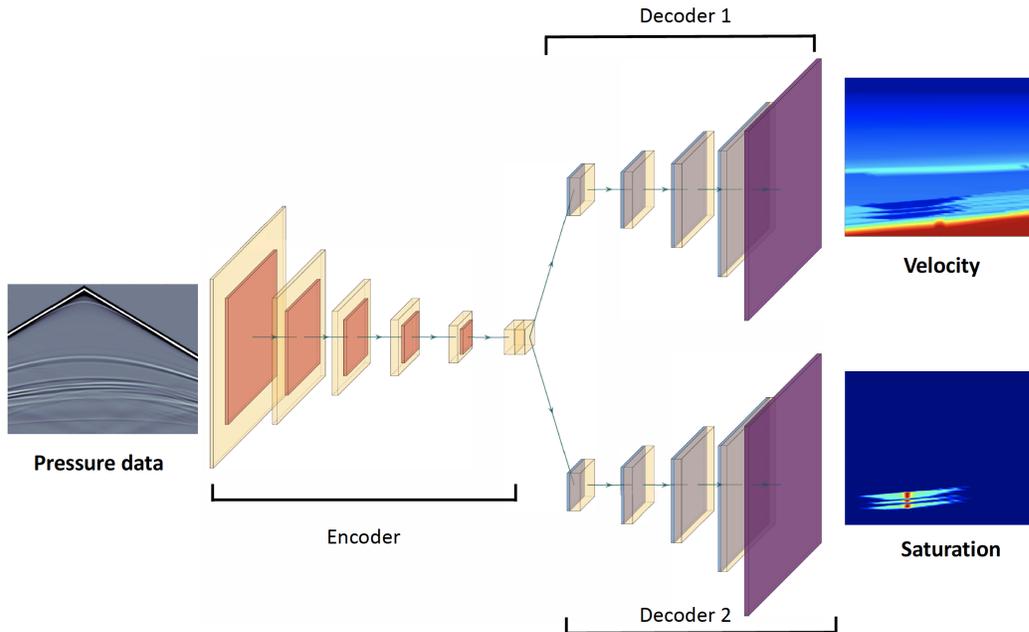


Figure 6.1 Network architecture of InvNet-VelSat. The encoder, built with five convolution layers and maximum pooling, extracts high-level features from the input seismic data to the latent space and reduces the dimensions of the feature map. Each of the two decoders consists of five deconvolution layers and translates the latent variable to the velocity (output 1) and saturation (output 2) models.

### 6.3 Methodology of time-lapse processing

The difference between the medium parameters reconstructed from the monitor and baseline surveys reflects the temporal changes in the subsurface properties caused by the injected CO<sub>2</sub>. Next, we discuss three different time-lapse strategies involving physics-based FWI and data-driven InvNet-VelSat.

#### 6.3.1 Strategy 1: Physics-based time-lapse FWI

In time-lapse FWI based on the parallel-difference method (Plessix et al., 2010), the baseline seismic data are inverted starting from a certain initial model. The same initial model is then used in the monitor inversion. The parameter changes are calculated by subtracting the inverted baseline parameters from those of the monitor model.

#### 6.3.2 Strategy 2: Data-driven time-lapse inversion

The data-driven algorithm uses the InvNet-VelSat for the monitor inversion. The network is first trained on extensive pre-existing training data set including the baseline survey. Once fully trained, InvNet-VelSat predicts both velocity and saturation from the observed monitor data. Since the baseline

models are assumed to be known in advance, the time-lapse variations represent the difference between the predicted monitor and the actual baseline parameters.

InvNet-VelSat has several advantages over FWI. First, FWI suffers from cycle-skipping (convergence) problems caused by local minima of the objective function (e.g., Virieux and Operto, 2009). Mitigating cycle skipping requires a sufficiently accurate initial model that is not always available. Second, FWI is computationally expensive (Virieux and Operto, 2009; Wu and Lin, 2019a), while well-trained InvNet-VelSat can predict the subsurface properties with sufficient resolution in seconds. The high efficiency of InvNet-VelSat is especially important in real-time monitoring during oil/gas production and CO<sub>2</sub> sequestration.

Furthermore, InvNet-VelSat can simultaneously predict velocity and saturation, while FWI has to convert the reconstructed velocity field into saturation via empirical petrophysics equations Xue et al. (2009). In addition, the background model stays almost unchanged outside the reservoirs, which mitigates the generalization issue in applying InvNet-VelSat to time-lapse processing. Although the velocity model can be dramatically altered at the near surface where seasonal changes and precipitation can dramatically alter the near surface velocity structure, it is outside the scope of this work and needs additional investigation. However, all data-driven inversions, including InvNet-VelSat, assume that a large amount of high-quality data (i.e., sufficiently similar to the testing data) is available for training purposes.

### 6.3.3 Strategy 3: Hybrid approach

The proposed hybrid approach (Figure 6.2), which combines the physics-based and data-driven methods, does not require pre-existing training data. First, we apply FWI to the baseline survey to reconstruct the baseline velocity model. For real-time monitoring of CO<sub>2</sub> sequestration, we process the monitor survey with InvNet-VelSat. To generate a sufficient volume of high-quality training data, the inverted baseline velocity is perturbed in the CO<sub>2</sub> reservoir, whereas the rest of velocity model remains unchanged. Hence, the prediction capability of InvNet-VelSat largely depends on the accuracy of the inverted baseline model, which is used to generate the training data.

To ensure sufficient accuracy of the baseline full-waveform inversion for field data, preprocessing should include careful application of the statics correction, robust estimation of the source wavelet, building of an appropriate initial model for FWI, etc. (Liu et al., 2013). For example, time shifts should be added to each trace as part of the statics correction to compensate for the weather changes and/or the difference in the depth of the sources and receivers. Our method is described in more detail below.

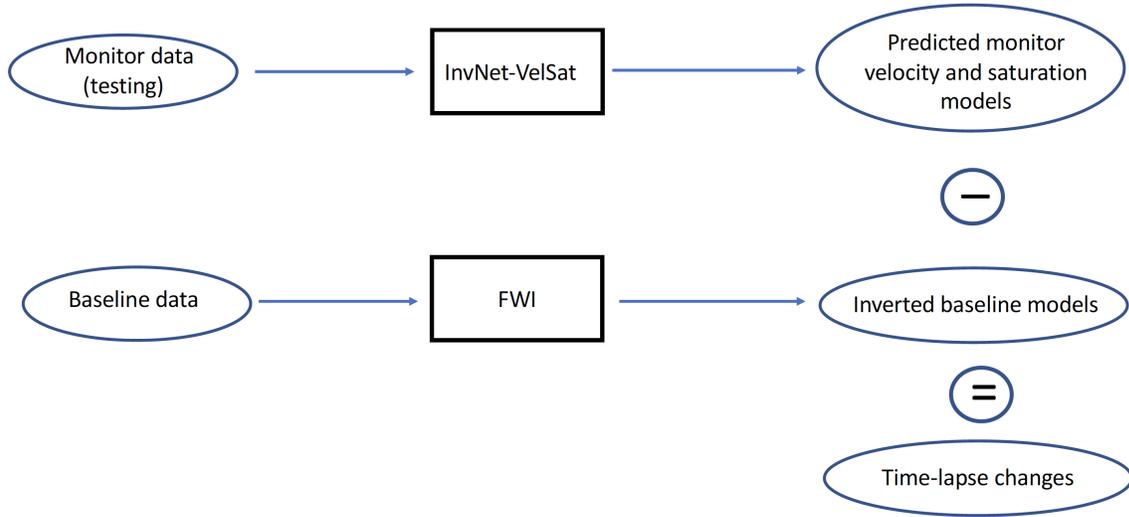


Figure 6.2 Workflow of the hybrid time-lapse strategy that combines physics-based FWI and data-driven InvNet-VelSat. FWI is employed to invert the baseline data and obtain the background velocity distribution for generating synthetic training samples. The trained InvNet-VelSat is applied to the monitor survey to predict both the velocity and saturation models. The temporal velocity changes are computed by subtracting the inverted baseline velocity model from the predicted monitor model.

### 6.3.3.1 Velocity-generation method

We employ the following prior knowledge when generating synthetic velocity models (Figure 6.3).

1. The subsurface structure before the injection is obtained from physics-based FWI, which is assumed to provide sufficient spatial resolution. The reconstructed baseline velocity model is changed only inside the reservoir to obtain the training velocity models for the model prediction from the monitor data.

2. CO<sub>2</sub> migration in the reservoir obeys certain laws of physics. Because supercritical CO<sub>2</sub> is less dense than water, the injected CO<sub>2</sub> should first accumulate near the top of the reservoir and then diffuse along the reservoir boundary.

3. Well logs (i.e., sonic, saturation, density, and electrical resistivity) from the injection and monitor wells are used to provide geologic and geophysical information, and thus constrain the inversion process.

The physics-related information used in our data-generation method typically can be obtained in practice. Therefore, the proposed hybrid approach and data-generation algorithm should be applicable to time-lapse field data.

### 6.3.3.2 Synthetic CO<sub>2</sub>-saturation models

Although the relationship between P-wave velocity and CO<sub>2</sub> saturation can be inferred from empirical equations based on laboratory and field experiments (Gassmann, 1951; Mavko and Mukerji, 1995), this

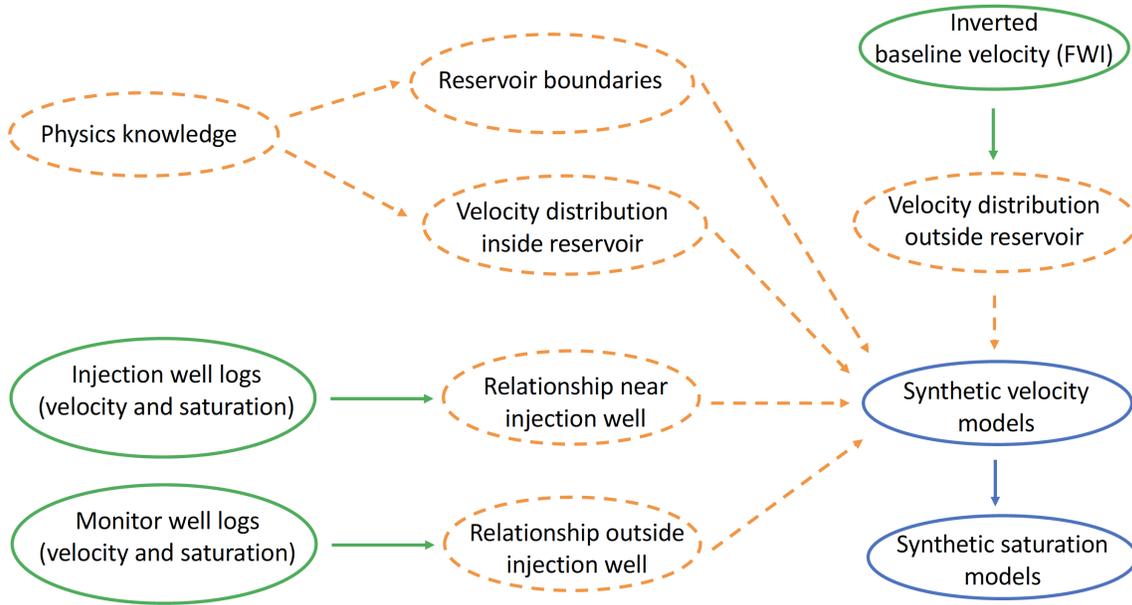


Figure 6.3 Workflow of the velocity- and saturation-simulation algorithms. The dashed orange oval marks the prior information and the empirical relationship between the velocity and saturation obtained from well logs. The solid green oval marks the inverted baseline velocity field used for generating training velocity models, and the sonic and saturation logs used for obtaining saturation. The solid blue oval marks the synthetic velocity and saturation models generated for training.

approach relies on the geologic similarity between the field of interest and those used to derive the equations. Here, we obtain the relationship between velocity and saturation from the available well logs using linear fitting equations (Figure 6.3).

Because the saturation level in the vicinity of the injection well is higher than that near the monitor wells, we separate the reservoir into two parts — the injection zone and the monitor zone. The velocity-saturation plots from the Kimberlina well logs indicate a linear relationship between these two parameters, but the slope changes between the low- and high-saturation parts. Therefore, we approximate the saturation-velocity relationship with two linear functions based on the saturation level.

## 6.4 Synthetic Examples

### 6.4.1 Kimberlina data set

The physics-based, data-driven, and hybrid strategies are applied to monitor and predict  $\text{CO}_2$  migration using a synthetic data set. (We have been unable so far to find publicly available field data suitable for our purposes.) The test is performed on the 4D Kimberlina reservoir data generated by several institutions (Alumbaugh et al., 2021) as part of the U.S. Department of Energy “SMART Initiative” (U.S. Department of Energy, 2029). This data set is generated from multiple  $\text{CO}_2$  injection reservoir simulations

for a model similar to the geologic structure of a commercial-scale carbon-sequestration reservoir at the Kimberlina site in the southern San Joaquin Basin (Wagoner, 2009). The Kimberlina time-lapse data can be used to evaluate the effectiveness and robustness of different geophysical techniques for monitoring CO<sub>2</sub> migration.

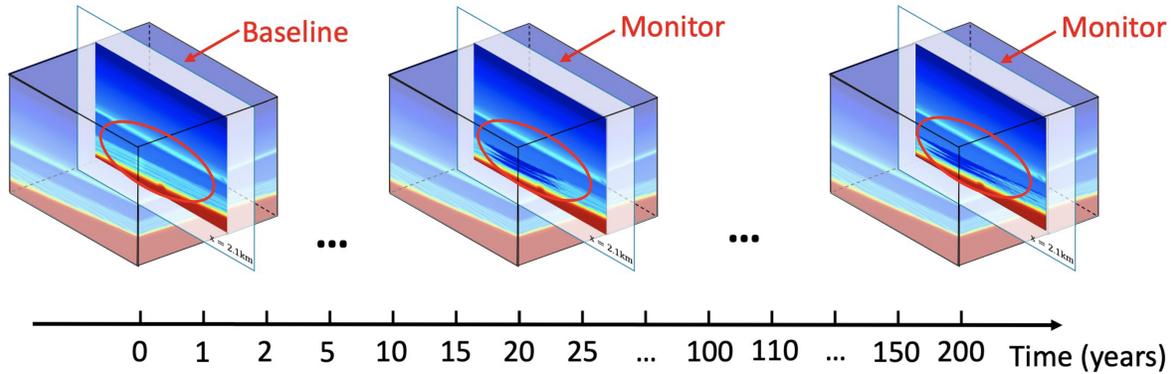


Figure 6.4 Visualization of the actual 4D velocity model. The baseline and monitor surveys are simulated for the same 2D vertical profiles before and during (or after) CO<sub>2</sub> injection. The time-lapse changes are encircled in red.

The data include 29 3D P-wave velocity and CO<sub>2</sub>-saturation models simulated for over 200 years with a grid size of 10 × 10 × 10 m (601 × 601 × 351 points; Figure 6.4). We slice each 3D model along the  $y$ -axis with a spatial interval of 100 m starting from  $x = 0$  to obtain 53 2D samples, which yields 1,537 velocity and saturation models (referred to as the “actual data”). In addition, time-lapse multiphysics well logs have been synthesized for four hypothetical well locations (Figure 6.5). These CO<sub>2</sub>-saturation, density, sonic velocity, and resistivity logs are available for 0, 1, 2, 5, 10, 15, and 20 years after the start of the CO<sub>2</sub> injection. Note that these four wells are not confined to the same vertical plane.

Figure 6.5 3D geometry of the Kimberlina data set. The injection well and three monitor wells are denoted by the red and orange dots, respectively, in the  $[x, y]$ -plane. The red line in the  $[x, y]$ -plane shows the projection of the baseline and monitor surveys. The red arrows in the  $[y, z]$ -plane point to three structures used for evaluating the effectiveness of time-lapse strategies: the low velocity layer above the reservoirs (Zone 1), the three CO<sub>2</sub> reservoirs (Zone 2), and the high-velocity dipping layer below the reservoirs (Zone 3).

The 2D slice containing the injection well in year 0 (i.e., before the injection) is treated as the baseline model. The monitor data are acquired in the same 2D vertical plane to map the movement of the injected CO<sub>2</sub> and compute the time-lapse parameter variations. The temporal velocity changes (Figure 6.5) are

heterogeneous and vary between  $x = -0.5$  km and  $x = 0$ . The saturation changes in the reservoir are also heterogeneous and vary between zero and unity. Both velocity and saturation models have a grid size of  $10 \times 10$  m.

There are three important geologic structures in the velocity model (Figure 6.5): the low-velocity layer immediately above the reservoir region (Zone 1), the three reservoirs themselves (Zone 2), and the high-velocity dipping layer beneath the reservoirs (Zone 3). Note that the shallow low-velocity horizon is not a water layer.

#### 6.4.2 Physics-based time-lapse FWI

The synthetic acoustic wavefield is excited by 37 shots (point explosions) placed with a constant increment (80 m) along a horizontal line at a depth of 10 m. The source signal is the Ricker wavelet with a central frequency of 10 Hz. We employ 294 receivers evenly distributed with an increment of 20 m along the horizontal line 20 m deep. The initial baseline and monitor models are identical (Figure 6.7a and Figure 6.7b) and computed by Gaussian smoothing of the actual baseline velocity field (Figure 6.6a) with a standard deviation of 15. Note that the  $\text{CO}_2$  plumes are not present in the initial model. FWI is applied to the simulated pressure recordings using a multiscale approach with four frequency bands starting from 2 Hz (2-5, 2-8, 2-13, 2-20 Hz; Liu and Tsvankin, 2021; Singh et al., 2020).

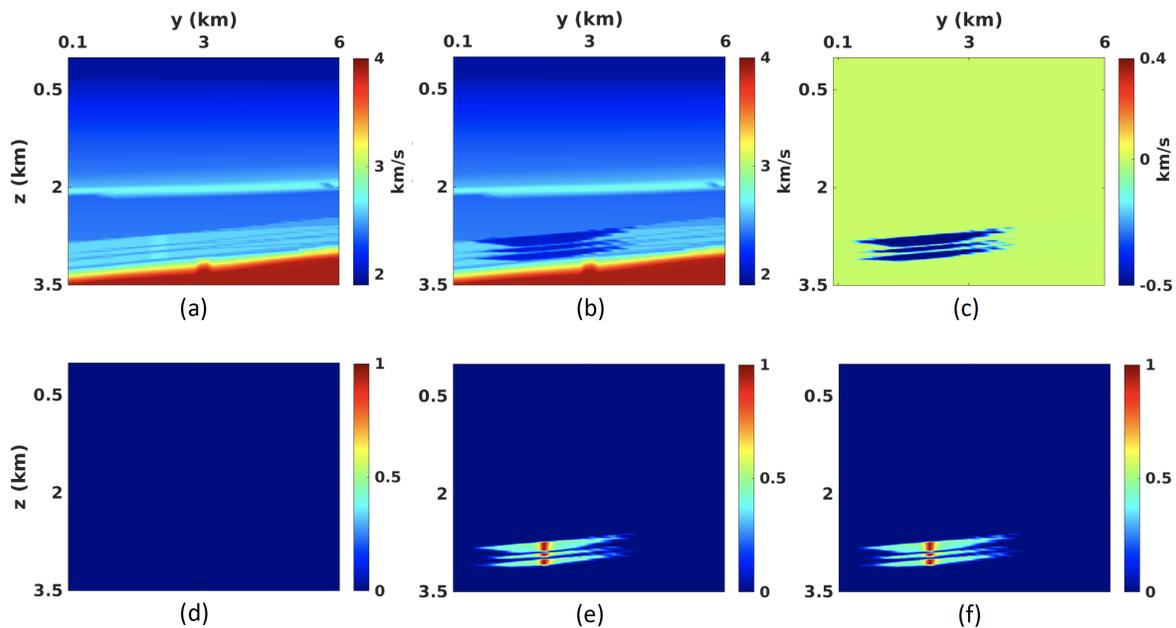


Figure 6.6 P-wave velocity of the (a) baseline Kimberlina model and (b) monitor model in year 20 with a grid size of  $10 \times 10$  m.  $\text{CO}_2$  saturation for the (d) baseline and (e) monitor model in year 20. The actual time-lapse changes of the (c) P-wave velocity and (f) saturation.

Figure 6.7d and Figure 6.7e show that FWI reconstructs the baseline and monitor velocity distribution with sufficient accuracy. Particularly, the “bump” in Zone 3 and the low-velocity region in Zone 1 (Figure 6.7d and Figure 6.7e) are well resolved in both the baseline and monitor models. There are inversion errors (Figure 6.7f) at depth, especially near the CO<sub>2</sub> plumes because the monitor inversion, which starts from the initial baseline model, could not adequately resolve the reservoirs. In addition, there are errors at the reservoir boundaries (i.e., in the thin nonpermeable layers) caused primarily by edge (smoothing) artifacts in the  $L_2$ -norm objective function (Schmidt, 2005; Zhang and Zhang, 2012).

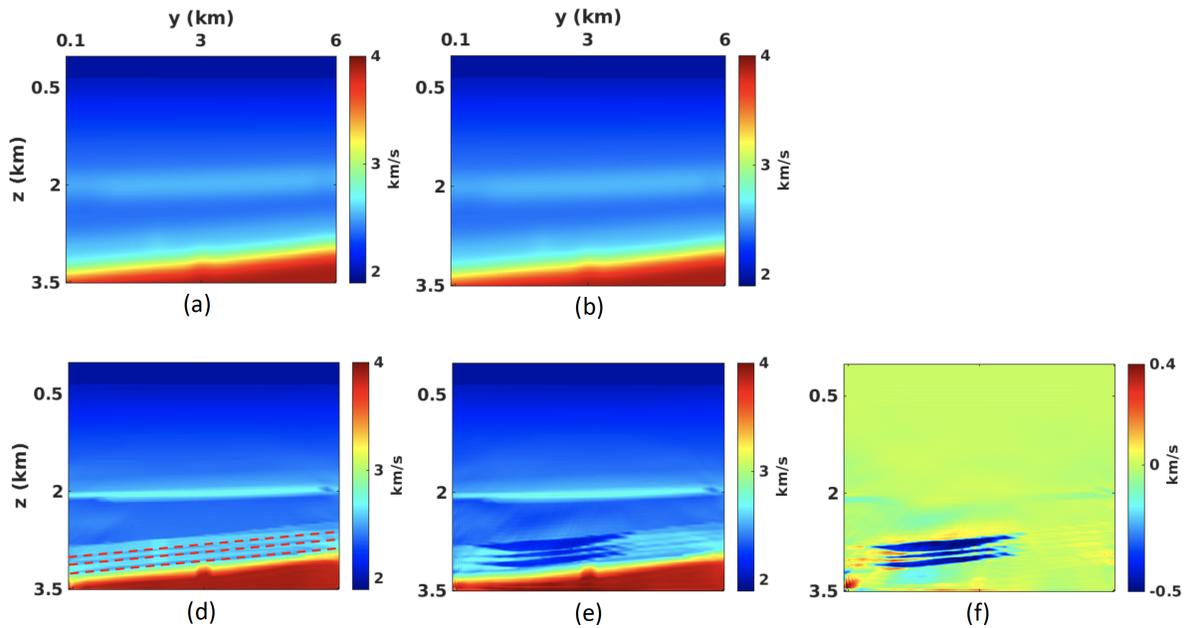


Figure 6.7 Initial P-wave velocity for the inversion of the (c) baseline and (d) monitor data. P-wave velocity obtained by the physics-based FWI: (d) baseline velocity, (e) monitor velocity, and (f) time-lapse velocity variations. The dashed red lines on plot (d) mark the reservoir boundaries.

### 6.4.3 Data-driven time-lapse inversion

The data-driven strategy assumes that the entire Kimberlina data set (1,537 velocity and saturation samples) is available. The seismic data are simulated using the same survey configuration as in the previous physics-based FWI test. The monitor survey in year 20 provides the testing data. The remaining actual data for over 200 years are randomly divided into the training (1,436 samples) and validation (100 samples) sets. The loss of the training test converges at around 0.05 after 2,212 epochs. Then the testing monitor data are inverted by the trained network.

Figure 6.8 shows that both the velocity and saturation for the monitor survey are generally predicted with acceptable accuracy. However, the shallow low-velocity layer, the low-velocity anomaly in Zone 1, and

the small “bump” in Zone 3 are not well reconstructed either due to the limitations of the training data set. Specifically, the 3D Kimberlina model is laterally heterogeneous, and the 2D velocity distribution varies for different vertical profiles. Only a relatively small fraction of the velocity models contains such geologic features as the “bump.” Thus, it is challenging for InvNet-VelSat to capture those features in the training data and predict them from the monitor samples.

Those limitations of the training data lead to errors in the background saturation model (Figure 6.8e and Figure 6.8f), which deviates from the actual saturation changes (Figure 6.6f). In addition, the saturation near the injection well is underestimated, which can be explained by the insensitivity of seismic acoustic velocity to high CO<sub>2</sub> saturation (Kim et al., 2010; Xue et al., 2009) and again the incompleteness of the training data.

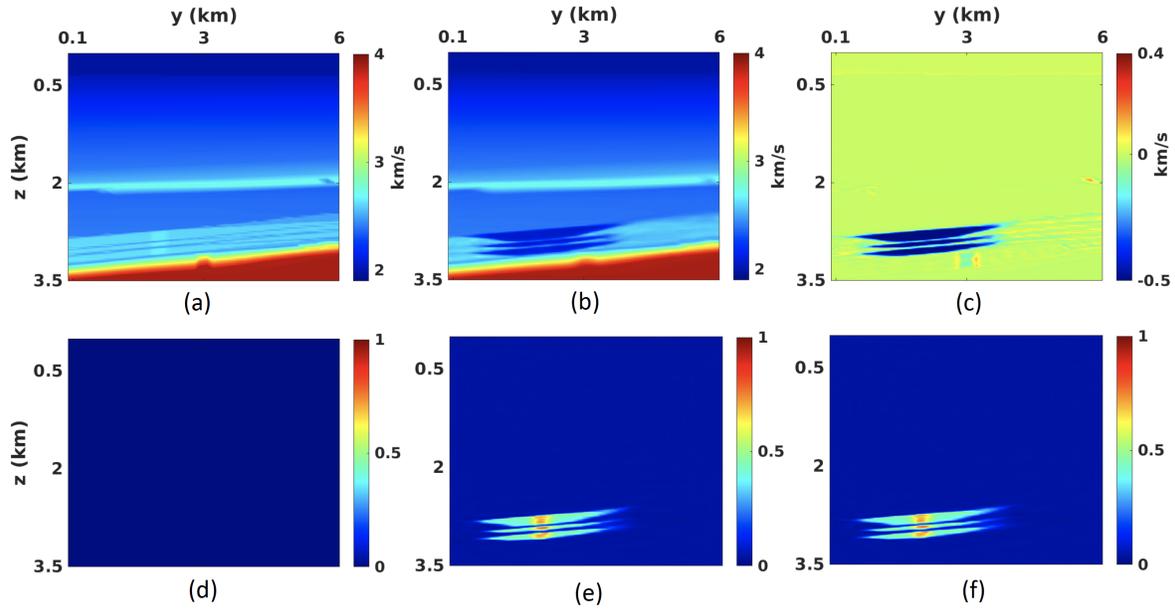


Figure 6.8 Actual baseline (a) velocity and (d) saturation models. The monitor (b) velocity and (e) saturation models predicted by the data-driven time-lapse inversion. The time-lapse variations of (c) velocity and (f) saturation.

#### 6.4.4 Hybrid strategy

According to the proposed hybrid strategy, FWI is first employed to reconstruct the baseline model (Figure 6.7d). To obtain synthetic velocity models, we first identify three aquifer layers and their top and bottom boundaries from the recovered baseline velocity (Figure 6.7d) and well logs. We then keep the velocity outside these layers unchanged because the temporal and spatial velocity variations are assumed to be confined to the reservoirs. Well logs provide the velocity profiles in the wells. We identify

the maximum and minimum velocity values in the reservoirs for all years from sonic well logs. The velocity in the reservoirs is perturbed within that range for generating a large number of velocity models. After obtaining the training velocity models, linear fitting equations are employed to build the relationship between the velocity and saturation at the injection well using the available well logs. The velocity-saturation relationship away from the injection well is found from linear fitting equations using the monitor-well logs. Note that these relationships may not be applicable to other data sets.

Reservoir simulation, required to accurately model the shape of CO<sub>2</sub> plumes, involves such parameters as permeability and viscosity, which are not available to us. Therefore, we approximate CO<sub>2</sub> movement in the aquifer layers by defining a transportation velocity function. According to this function, CO<sub>2</sub> moves parallel to the upper reservoir boundary with a speed that exponentially decreases with depth. Hence, buoyancy-driven CO<sub>2</sub> moves faster in the upper reservoir (Sigfusson et al., 2015):

$$V_c^d = V_c^{d_0} e^{-c(d-d_0)}, \quad (6.5)$$

where  $V_c^d$  is the velocity of CO<sub>2</sub> migration at depth  $d$ ,  $d_0$  is the depth of the reservoir top boundary, and  $c$  is a constant that controls the rate of decrease of  $V_c^d$ . Note that the transportation function is based on the structure of the Kimberlina reservoir and, therefore, may need to be modified for other data sets.

The well logs show that the P-wave velocity in the plumes varies between 1.9 km/s and 2.5 km/s. To simplify the velocity distribution, we ignore its lateral variation between the wells. A total of 28,000 training data samples are generated for different velocities ranging from 2.0 km/s to 2.3 km/s with an increment of 0.1 km/s.

Comparison with the actual velocity models (Figure 6.9a-Figure 6.9d) demonstrates that the synthetic velocity samples (Figure 6.9e-Figure 6.9h) are physically realistic and capture the spatial and temporal dynamics of CO<sub>2</sub> migration inside the reservoirs. However, the contours of the CO<sub>2</sub> plumes may not be accurately simulated over long periods of time (i.e., after year 20) because the physics knowledge becomes more limited with time.

Using the generated velocity models, we obtain synthetic CO<sub>2</sub> saturation samples (Figure 6.9m-Figure 6.9p) via empirical relationships between saturation and velocity obtained from the well logs (Figure 6.10c). This procedure is illustrated for the injection well in year 20 in Figure 6.10a and Figure 6.10b. Comparison with the actual values (Figure 6.9i-Figure 6.9l) proves that the synthetic saturation models obtained using the proposed method reflect the movement of the injected CO<sub>2</sub> with sufficient accuracy.

Predictably, the results become less accurate with time (especially, after year 20) for the reasons listed below: 1. The CO<sub>2</sub> plumes have similar contours in velocity and saturation. Hence, the synthetic

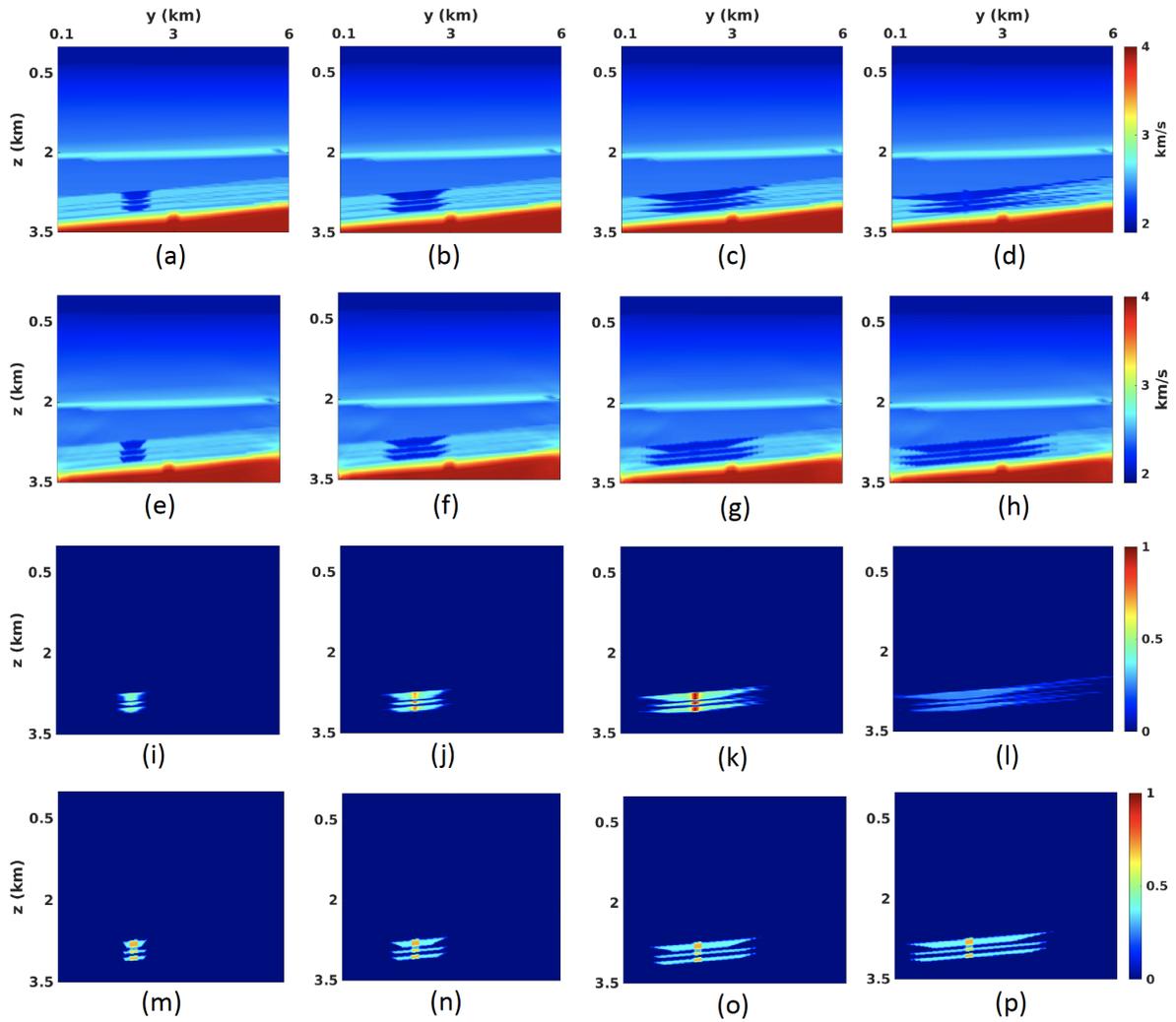


Figure 6.9 Actual P-wave velocity model (first row) in (a) year 1 (b) year 5, (c) year 20, and (d) year 130. The second row shows the corresponding generated velocity samples. The actual saturation model (third row) in (i) year 1, (j) year 5, (k) year 20, and (l) year 130. The fourth row shows the corresponding generated saturation samples.

saturation samples could not properly capture the plumes after year 20 because the corresponding velocity contours were inaccurate. 2. The high-saturation anomaly near the injection well is somewhat distorted because the velocity and CO<sub>2</sub> saturation profiles from the well logs are not adequately represented by the linear fitting equations. 3. The simulated velocity distribution inside the CO<sub>2</sub> plumes is homogeneous (i.e., it does not account for spatial variations), which reduces the accuracy of the synthetic saturation model.

Then the 28,000 synthetic data samples are randomly divided into the training (27,500) and validation (500) sets. After 140 epochs, the training loss flattens out at around 0.0013. Then, we test the trained neural network on the actual Kimberlina data that have the same acquisition geometry as the baseline survey. Note that the network employed in the hybrid strategy converges faster and produces a smaller loss than the pure data-driven method because of the higher quality of the training data. Both the velocity and saturation for the monitor survey in year 20 are well predicted by the trained neural network.

#### 6.4.4.1 Hybrid strategy vs. physics-based method

InvNet-VelSat can reconstruct the velocity (Figure 6.11c) and saturation (Figure 6.11f) from the monitor seismic data simultaneously and with sufficient resolution, which is not possible for the physics-based strategy. In general, our method outperforms the physics-based FWI (Figure 6.7f), which is confirmed by comparing the MAE values for the estimated velocity and by visually inspecting the results. The MAE for the velocity field reconstructed by our approach is 0.0084, whereas that of the physics-based method is 0.056. The CO<sub>2</sub> plumes and the high-velocity zone underneath are reconstructed with a higher resolution using the hybrid method (Figure 6.11c). However, the boundaries are positioned less accurately due to the limited physics information and simulation errors in the velocity-generation algorithm.

Because the baseline inversion is conducted by the physics-based FWI for both strategies, we compare the computational efficiency only for the monitor survey (Table Table 6.1). Although the training time of the hybrid approach (about 25.3 h) is significantly longer than the running time of the physics-based method (about 8.5 h), the testing (inversion) for one survey using the hybrid strategy takes only 14 s. Hence, the hybrid strategy is much more efficient than the physics-based method when processing multiple seismic surveys in real-time monitoring.

#### 6.4.4.2 Hybrid strategy vs. data-driven strategy

Compared to the velocity model predicted by the pure data-driven method (Figure 6.8b and Figure 6.8c), the hybrid approach better captures the small “bump” and the low-velocity anomaly (Figure 6.11b and Figure 6.11c). In addition, the hybrid algorithm (Figure 6.11c) succeeds in reconstructing the nonpermeable thin layers between the reservoirs without the false anomalies produced

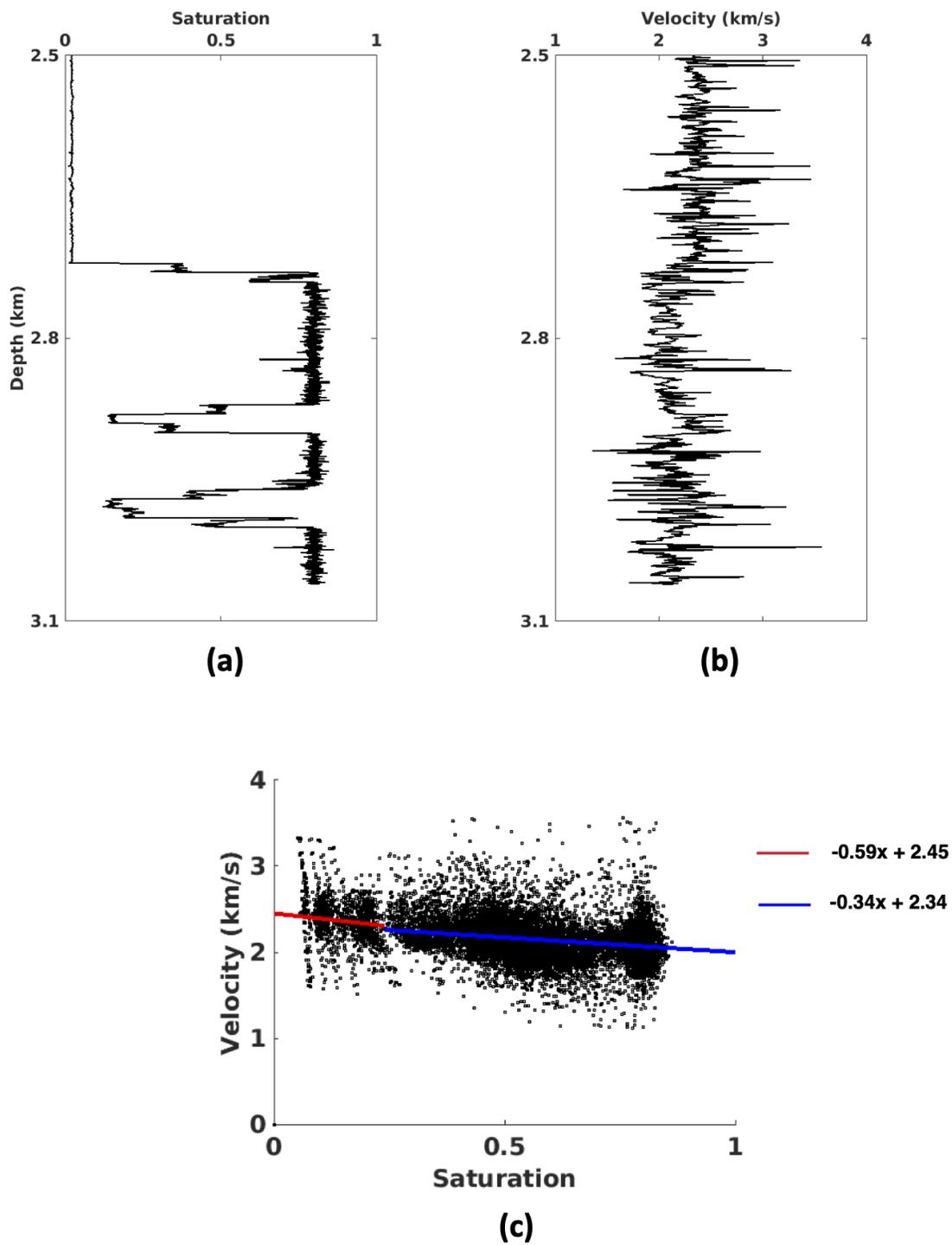


Figure 6.10 Well logs and empirical relationship between the velocity and CO<sub>2</sub> saturation for the injection well in year 20. The red and blue lines on plot (c) are two linear fitting equations separated according to the CO<sub>2</sub> saturation level.

Table 6.1 Computational cost of the monitor inversion using the physics-based and hybrid strategies.

	CPU/GPU	Number of (#) nodes	# Processors	# Training seismic surveys
Physics-based method	CPU	3	50	1
Hybrid	GPU	2	64	27,500
	# Iterations	Running time	Training time	Testing time per survey
Physics-based method	300	8.5 h	N/A	8.5 h
Hybrid	140	N/A	25.3 h	14 s

by the data-driven method (Figure 6.8c). These improvements are due to the more representative synthetic training data set generated using our method. Specifically, by leveraging a well-constrained background velocity distribution (the region outside the reservoir), the network can better capture essential geologic features from the training set and estimate the background velocity using the monitor surveys.

On the other hand, the hybrid strategy produces larger errors at the boundaries of the CO<sub>2</sub> plumes because of the inaccurate boundary delineation in the synthetic models (see above). Primarily due to the distortions at the boundaries, the overall MAE for the time-lapse velocity variations produced by our approach (0.0084) is larger than that of the data-driven method (0.0062).

Because the injected CO<sub>2</sub> is confined to the reservoir region, saturation outside the reservoir is supposed to vanish (Figure 6.6f), as correctly predicted by the hybrid method (Figure 6.11f). In contrast, the data-driven strategy yields nonzero saturation values (ranging between 0.02 and 0.06) in that area (Figure 6.8f). Although the contours of the plumes and the high-saturation zone near the injection well are not accurately estimated by the hybrid approach, its MAE for the entire saturation model (0.0075) is much smaller compared to that of the data-driven method (0.039).

We simulate the inverted/predicted baseline and monitor data and compute the time-lapse seismograms (Figure 6.12) by subtracting these two data sets. Compared to the actual record (Figure 6.12a), all three methods (Figure 6.12b-Figure 6.12d) produce artifacts near the first arrival due to the errors in the inverted time-lapse variations (Figure 6.7f, Figure 6.8c, and Figure 6.11c). In addition, because of the inaccurate parameter estimation outside the reservoirs, the physics-based FWI (Figure 6.12b) and data-driven method (Figure 6.12c) yield more intensive false reflection events after the first arrival than the hybrid approach. Hence, the proposed hybrid strategy outperforms the other two methods in the data domain.

#### 6.4.5 Generalization and robustness: Hybrid strategy vs. data-driven strategy

Generalizability and robustness of a neural network indicate whether it is actually learning rather than simply memorizing the input-output relationship.

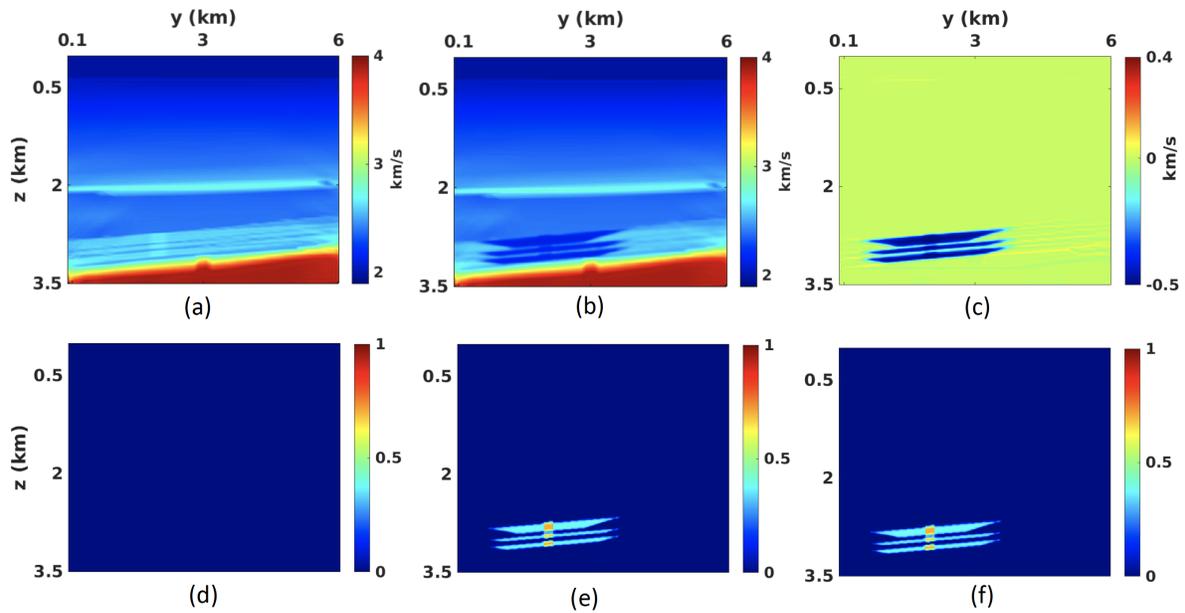


Figure 6.11 Baseline (a) velocity and (d) saturation models inverted by physics-based FWI. The monitor (b) velocity and (e) saturation models predicted by the hybrid time-lapse inversion. The time-lapse variations of (c) velocity and (f) saturation.

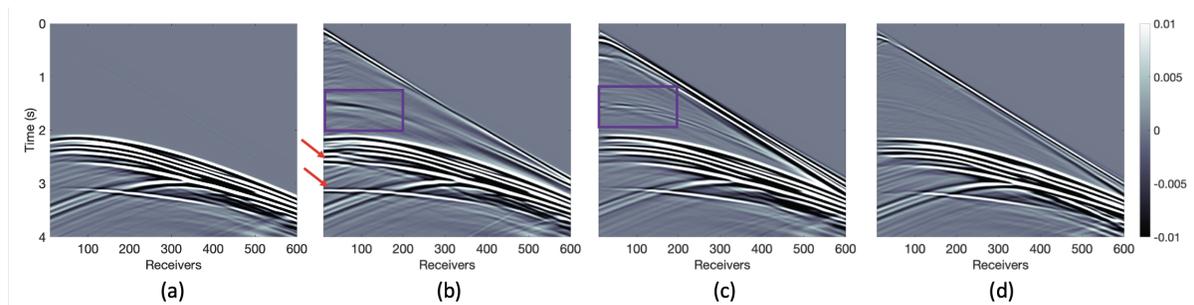


Figure 6.12 (a) Actual time-lapse seismogram (the difference between the monitor data in year 21 and baseline data in year 0). Time-lapse seismogram produced by (b) the physics-based FWI, (c) the data-driven method, and (d) the hybrid strategy. Artifacts after the first arrivals are either encircled in purple or marked by red arrows.

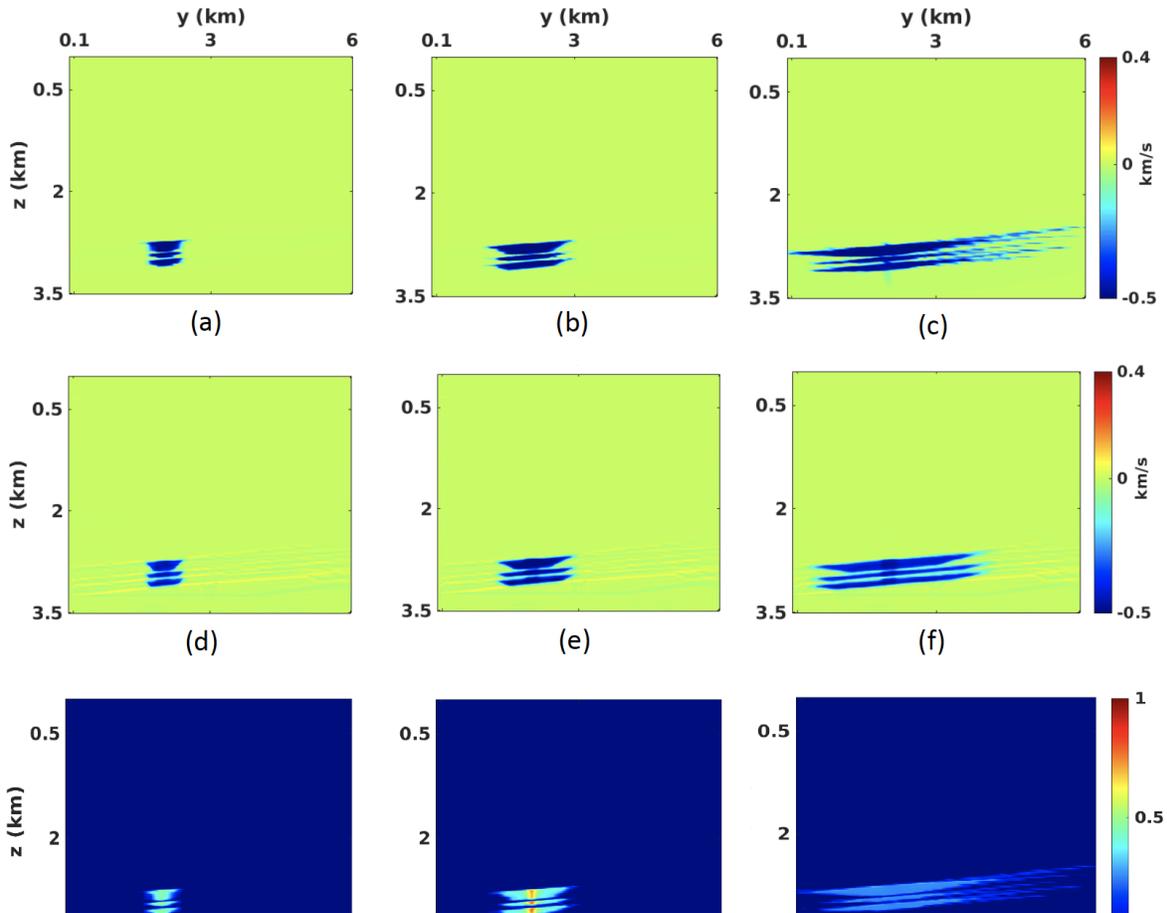


Figure 6.13 Actual time-lapse velocity variations for (a) year 1, (b) year 5, and (c) year 130. Plots (d), (e), and (f) show the predicted time-lapse velocity variations. The actual time-lapse saturation variations for (h) year 1, (i) year 5, and (g) year 130. Plots (k), (l), and (m) show the predicted time-lapse saturation variations.

#### 6.4.5.1 Additional monitoring tests

To find out if the proposed hybrid approach is generalizable, we apply it to the remaining monitor surveys (not used in the training or validation). The three testing 2D monitor surveys (for years 1, 5, and 130) are acquired along the same line as the baseline survey. The testing is performed using InvNet-VelSat without any further fine-tuning. The velocity and saturation both inside and outside the three CO<sub>2</sub> plumes are reconstructed with sufficient accuracy (Figure 6.13). The differences between the actual and inverted temporal changes in all predicted monitor models are observed mostly at the boundaries. As expected, however, the prediction error increases with time.

#### 6.4.5.2 Influence of noise

Next, the testing monitor data for year 20 are contaminated with Gaussian noise that has the signal-to-noise ratio (SNR) equal to 10. The noise is added only to the testing data, whereas InvNet-VelSat is still trained on noise-free samples.

Predictably, both the data-driven and hybrid strategies produce more errors in the velocity and saturation changes estimated from the noise-contaminated data (Figure 6.14). For the velocity model obtained by the data-driven method, distortions are observed in the low-velocity layer above the CO<sub>2</sub> plumes (Zone 1), in the plumes themselves (Zone 2), and near the boundary between the reservoir and the deep high-velocity horizons (Zone 3). As a result, the corresponding MAE increases to 0.054, which is much larger than that for the noise-free data (0.0062).

In contrast, the hybrid strategy reconstructs Zones 1 and 3 with sufficient resolution, with the errors (MAE=0.013) mostly concentrated near the boundaries of the CO<sub>2</sub> plumes in Zone 2. Evidently, our method is more robust for noisy data than the data-driven approach. Likewise, the hybrid strategy predicts the saturation from the noisy pressure recordings with higher accuracy (MAE=0.0085) than the data-driven method (MAE=0.043) by taking advantage of the large volume of high-quality training data.

### 6.5 Conclusions

We developed a time-lapse inversion workflow that does not require pre-existing training data by combining physics-based full-waveform inversion (FWI) and data-driven neural networks. First, FWI is applied to the baseline survey to estimate the background velocity model. Then data-driven inversion (InvNet-VelSat) is employed to predict the velocity and CO<sub>2</sub> saturation from the monitor data set. The training data for InvNet-VelSat are simulated with a data-generation method based on the reconstructed background velocity model and physics knowledge (i.e., well logs), which is expected to be available for typical CO<sub>2</sub>-sequestration projects. Because these training models are simulated using the

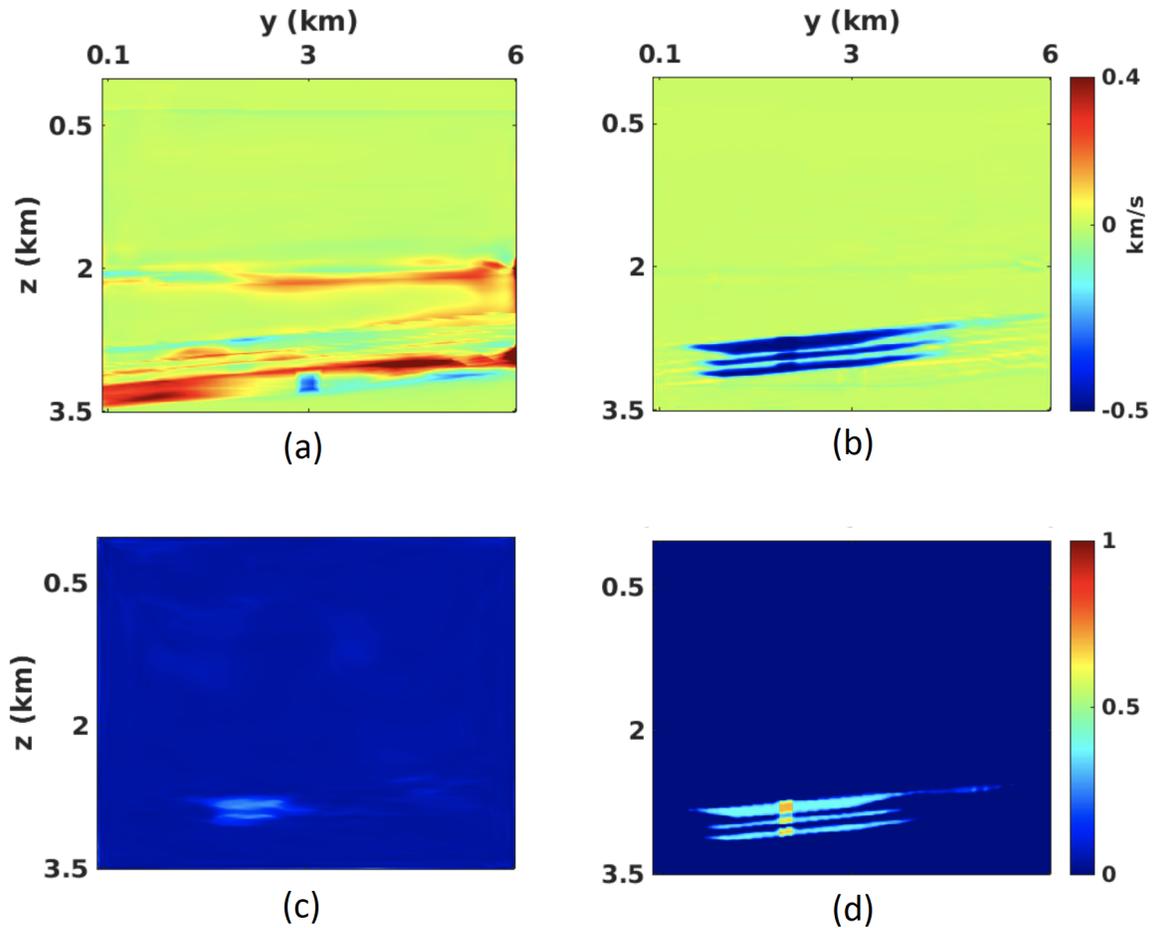


Figure 6.14 Temporal variations in the P-wave velocity obtained from noise-contaminated data ( $\text{SNR} = 10$ ) by (a) the data-driven method and (b) the hybrid strategy. Plots (c) and (d) show the corresponding predicted time-lapse variations in the  $\text{CO}_2$  saturation.

inverted baseline data, the proposed workflow relies on the accuracy of the employed FWI algorithm.

The hybrid strategy is tested on realistic synthetic data from the Kimberlina reservoir. The training samples (i.e., the velocity and saturation models) simulated by our data-generation method adequately capture the spatial and temporal dynamics of CO<sub>2</sub> movement. The accuracy of the time-lapse variations predicted by the hybrid method is comparable to or even higher compared to the much more time-consuming physics-based FWI. Testing on noisy data for different time intervals illustrates the robustness and generalization capability of the hybrid strategy, whose performance is superior to that of the data-driven method. It should be emphasized that the hybrid strategy does not require a large volume of training data, which should facilitate its application to monitoring CO<sub>2</sub> sequestration.

## **6.6 Acknowledgments**

This work was funded by the U.S. Department of Energy (DOE) Office of Fossil Energy’s Carbon Storage Research Program via the Science-Informed Machine Learning to Accelerate Real Time Decision Making for Carbon Storage (SMART-CS) Initiative, and also co-funded by the Laboratory Directed Research and Development (LDRD) Program under the project # 20210542MFR at Los Alamos National Laboratory. Y. Liu and I. Tsvankin will also acknowledge the support of the sponsors of the Center for Wave Phenomena (CWP) at Colorado School of Mines.

## **6.7 Data and materials availability**

Data associated with this research are available and can be obtained by contacting the corresponding author.

## CHAPTER 7

# ENHANCED PREDICTION ACCURACY WITH UNCERTAINTY QUANTIFICATION IN MONITORING CO<sub>2</sub> SEQUESTRATION USING CONVOLUTIONAL NEURAL NETWORKS

Submitted to *Geophysics*

Yanhua Liu<sup>1,2,3</sup>, Xitong Zhang<sup>4</sup>, Ilya Tsvankin<sup>3</sup>, and Youzuo Lin<sup>2</sup>

Monitoring changes inside a reservoir in real time is crucial for the success of CO<sub>2</sub> injection and long-term storage. Machine learning (ML) is well-suited for real-time CO<sub>2</sub> monitoring because of its computational efficiency. Additionally, the subsurface outside the reservoir is assumed to be unchanged, which reduces the sensitivity of the ML application to the generalization issue. However, most existing ML implementations yield only one prediction (i.e., the expectation) for a given input, which may not properly reflect the distribution of the testing data, if it has a shift with respect to that of the training data. The Simultaneous Quantile Regression (SQR) method can estimate the entire conditional distribution of the target variable of a neural network via pinball loss. Here, we incorporate this technique into seismic inversion for purposes of CO<sub>2</sub> monitoring. The uncertainty map is then calculated pixel by pixel from a particular prediction interval around the median. We also propose a novel data-augmentation method by sampling the uncertainty to further improve prediction accuracy. The developed methodology is tested on synthetic Kimberlina data, which are created by the Department of Energy and based on a CO<sub>2</sub> capture and sequestration (CCS) project in California. The results prove that the proposed network can estimate the subsurface velocity rapidly and with sufficient resolution. Furthermore, the computed uncertainty quantifies the prediction accuracy. The method remains robust even if the testing data are distorted due to problems in the field data acquisition. Another test demonstrates the effectiveness of the developed data-augmentation method in increasing the spatial resolution of the estimated velocity field and in reducing the prediction error.

### 7.1 Introduction

Since the rise of industrialization in the 18th century, human activities have increased the volume of atmospheric CO<sub>2</sub> by 50%, thus increasing the global temperature. Capturing industrial CO<sub>2</sub> at its various sources and injecting it into geologic formations for long-term storage (sequestration) is one of the most promising methods to combat global warming.

---

<sup>1</sup>Primary researcher and author.

<sup>2</sup>Earth and Environmental Sciences Division, Los Alamos National Laboratory

<sup>3</sup>Center for Wave Phenomena, Colorado School of Mines

<sup>4</sup>Computational Mathematics, Science and Engineering, Michigan State University

Seismic data provide valuable information about the subsurface, and can be used to monitor CO<sub>2</sub> injection. Both laboratory and field data confirm that P-wave velocity decreases with CO<sub>2</sub> saturation, especially for relatively low saturation levels (Kim et al., 2010). Seismic images can also help delineate CO<sub>2</sub> plumes inside the reservoir. Therefore, seismic data can be used to monitor CO<sub>2</sub> injection and storage in subsurface structures and, potentially, detect small leakages of CO<sub>2</sub> (Furre et al., 2017; Lumley, 2010; Pevzner et al., 2017).

Data-driven neural networks are known to be suitable for monitoring the movement of CO<sub>2</sub> using time-lapse seismic data. As the medium properties outside the reservoir often remain almost unchanged during CO<sub>2</sub> injection, the potential distribution shift between the testing monitor data and the training data typically is not significant in time-lapse monitoring. This implies that the issue of generalization, often encountered when neural networks are applied to a single data set, is expected to have a smaller impact for time-lapse data (Liu et al., 2022).

Data-driven neural networks have been applied to monitor the CO<sub>2</sub> movement using time-lapse seismic data. For example, Li et al. (2021) develop a fully-connected neural network to map the relationship between time-lapse seismic data and the velocity changes caused by the injected CO<sub>2</sub>. Feng et al. (2021) propose spatio-temporal neural-network-based models with long short-term memory (LSTM) structure to monitor and forecast the CO<sub>2</sub> storage at Sleipner field in the North Sea. Liu et al. (2022) use a hybrid time-lapse strategy that combines physics-based FWI and data-driven neural network to monitor the CO<sub>2</sub> movement in the reservoir. The main advantage of data-driven inversion is its efficiency during the application stage (after training), which is essential in real-time monitoring of CO<sub>2</sub> sequestration (Zhang and Lin, 2020). An extensive overview of data-driven seismic inversion methods can be found in Lin et al. (2022).

However, deterministic neural networks assume that the mapping learned by the network is accurate, which is not always the case, and produce only the learned output for a given input. On the other hand, uncertainty quantification (UQ) not only describes predictive distributions over outputs for given inputs, but also indicates whether the model is confident about the prediction. The uncertainty can be divided into two categories based on its sources: epistemic (model) and aleatoric (data) uncertainty (Tagasovska and Lopez-Paz, 2019). Epistemic uncertainty is produced by the neural network itself: its architecture, training procedures, the number of samples, etc. It can be mitigated by collecting more representative training data, which helps improve testing performance (Ren et al., 2021). Data uncertainty describes the variance of the conditional distribution of a prediction for given input features. In contrast to epistemic uncertainty, data uncertainty cannot be reduced by modifying model architecture, training algorithms, or collecting more data under the same experimental conditions because the noise distribution in seismic data

cannot be considered constant.

Uncertainty evaluation has been used in seismic monitoring of CO<sub>2</sub> injection to quantify prediction accuracy. Chen et al. (2018) propose to monitor CO<sub>2</sub> leakage by multivariate adaptive regression splines and to measure the prior and posterior uncertainty using the percentile estimation from Monte-Carlo simulations. Tang et al. (2022) develop a 3D recurrent R-U-Net surrogate model to predict CO<sub>2</sub> saturation of a synthetic sequestration region and quantify the uncertainty based on rejection sampling using the CNN (convolutional neural network)-PCA (principle-component analysis) model. Um et al. (2022) introduce a U-Net network to estimate CO<sub>2</sub> saturation and model uncertainty with two UQ methods (i.e., the Monte Carlo dropout method and a bootstrap aggregating method).

However, the uncertainty quantification methods mentioned above either require an extra sampling step during testing or use the predictive variance/disagreement to represent the uncertainty. The sampling step can be time-consuming, if physics simulation is involved. Moreover, using just one variable (e.g., variance) as the uncertainty indicator assumes the uncertainty to be symmetric around the estimated prediction, which is not always true. For example, the uncertainty becomes asymmetric, if the noise inherent in the input data is sampled from a skew-normal distribution.

The Simultaneous Quantile Regression (SQR) is proposed by Tagasovska and Lopez-Paz (2019) to quantify the uncertainty of a 1D regression problem. SQR is designed to train a model that makes predictions at different quantile levels utilizing pinball loss. The uncertainty map is computed by subtracting the predictions at two selected quantile levels around the predicted median values. Hence, SQR does not need extra sampling for uncertainty estimation and can produce asymmetric confidence intervals.

Here, we incorporate SQR into a CNN designed to estimate the uncertainty of the velocity distribution estimated from seismic data. To improve the prediction accuracy, we propose a novel data-augmentation method that operates with the prediction and calculated uncertainty. We begin by discussing the methodology of SQR and the architecture of the proposed network (InvNet\_UQ). Then InvNet\_UQ is tested on the Kimberlina model based on a CO<sub>2</sub> injection project in California. The performance of InvNet\_UQ is evaluated by comparing the predictions with those of a deterministic neural network from both the perfect and distorted testing data. The robustness of the computed uncertainty is verified by its comparison with the prediction error. Finally, the effectiveness of the data-augmentation method is evaluated by comparing the error of the predicted velocity model before and after applying the augmentation.

## 7.2 Method

Figure 7.1 illustrates the developed workflow using the proposed probabilistic convolutional neural network (InvNet\_UQ). First, InvNet\_UQ is trained and applied to predict the velocity distributions at all

quantile levels from the testing data. Meanwhile, the uncertainty map is calculated from two predictions at different quantile levels. Next, uncertainty-guided data augmentation is applied to all predicted velocity models from the testing data. Finally, the augmented data are added to the existing training data for retraining the network.

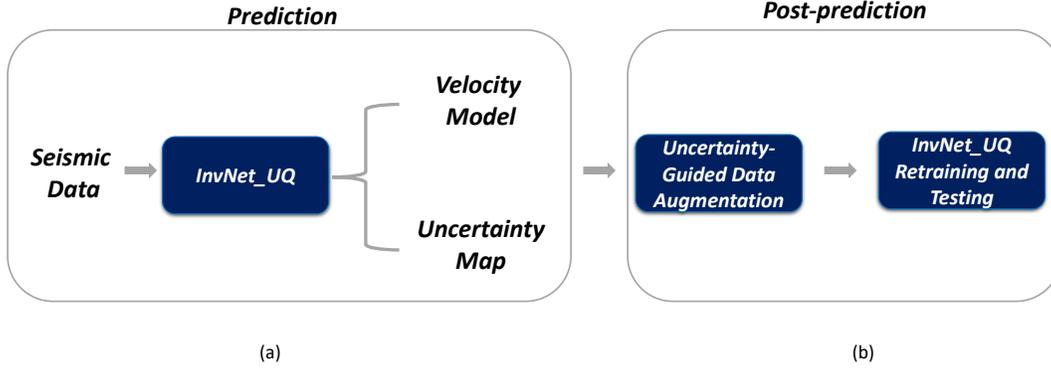


Figure 7.1 Flowchart of the method. (a) InvNet.UQ is trained and tested to estimate the velocity and the corresponding uncertainty. (b) The uncertainty-guided data-augmentation method is applied to the predicted velocity models; then the network is retrained with the augmented data.

### 7.2.1 Simultaneous Quantile Regression

Uncertainty estimation is essential for seismic inversion for two reasons. First, seismic data contain different kinds of noise caused by the recording equipment, ambient disturbances, etc. Second, inversion is often ill-posed, which means that the solution can be nonunique or correspond to a local minimum of the objective function. Here, we estimate the uncertainty using the probabilistic neural based on the SQR proposed by Tagasovska and Lopez-Paz (2019).

To solve the regression problem, we implement a neural network as a function  $\hat{y} = \hat{f}_\tau(x)$  to approximate the relationship between the input  $x$  and output  $y$ . To identify the network parameters, we minimize the mean-square error (MSE) between the prediction ( $\hat{y}$ ) and actual ( $y$ ) values:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = E(y - \hat{y})^2. \quad (7.1)$$

Therefore, the prediction ( $\hat{y}$ ) represents the conditional mean (expectation) of the prediction for a given input.

However, the expectation alone does not accurately reflect the data distribution. Therefore, quantile regression is proposed to analyze the prediction in a specific quantile ( $\tau$ ) (Koenker and Hallock, 2001). The main advantage of QR over linear regression (LR) is that QR explores different values of the response variable (instead of only the average) and, therefore, delivers a more complete description of the

relationships between the variables. Moreover, QR is more robust than LR in the presence of outliers and influential points (Zablotski, 2023). One way to build such a model approximating the conditional quantile distribution function  $y = F^{-1}(\tau|X = x)$  is to minimize the pinball loss  $(l_\tau(y, \hat{y}))$ :

$$l_\tau(y, \hat{y}) = \begin{cases} \tau(y - \hat{y}) & \text{if } y - \hat{y} \geq 0, \\ (1 - \tau)(\hat{y} - y) & \text{else.} \end{cases} \quad (7.2)$$

Indeed,

$$E[l_\tau(y, \hat{y})] = (\tau - 1) \int_{-\infty}^{\hat{y}} (y - \hat{y}) dF(y) + \tau \int_{\hat{y}}^{\infty} (y - \hat{y}) dF(y), \quad (7.3)$$

where  $F(y) = P(Y \leq y)$  is the strictly monotonic cumulative distribution function of the target variable  $Y$  taking real values  $y$ . Consequently,  $F^{-1}(\tau) = \inf\{y : F(y) \geq \tau\}$  denotes the quantile distribution function of the same variable  $Y$  for all quantile levels  $0 \leq \tau \leq 1$ .

With pinball loss, Tagasovska and Lopez-Paz (2019) propose employing Simultaneous Quantile Regression (SQR) to estimate all the quantile levels simultaneously by solving the following equation:

$$\hat{f} \in \arg \min_f \frac{1}{n} \sum_{i=1}^n E_{\tau \sim U[0,1]} [l_\tau(f(x_i, \tau), y_i)], \quad (7.4)$$

where  $(x_i, y_i)$  are identically and independently distributed feature-target pairs drawn from the unknown probability distribution  $P(X, Y)$ . Note that, unlike Bayesian neural networks,  $P(X, Y)$  is not necessarily a Gaussian distribution. Then the estimated data uncertainty can be computed from the  $(1 - \alpha)$  prediction interval around the median:

$$u_\alpha(x^*) := \hat{f}(x^*, 1 - \frac{\alpha}{2}) - \hat{f}(x^*, \frac{\alpha}{2}), \quad (7.5)$$

where  $\alpha$  is the significance level.

### 7.2.2 Network architecture of InvNet\_UQ

Following Wu and Lin (2019b), we use an encoder-decoder-based CNN to approximate the relationship between the input seismic data and the output velocity model. There are seven convolutional blocks with a kernel size of  $3 \times 3$  followed by a  $2 \times 2$  max pooling. Each block consists of a convolutional layer, batch normalization, and a tanh activation function. The high-dimensional data are reduced to a  $1024 \times 1 \times 1$  array in the latent space. These features are upsampled to the output size by eight decoder blocks containing a deconvolutional layer, batch normalization, and a LeakyRelu activation function. The end of the decoder is a central-cropping layer, which crops the output of the deconvolutional blocks to the desired size.

For the benchmark network (InversionNet), the loss function is the  $L_1$ -norm of the difference between the prediction and the actual model. The input data size is  $6 \times 1,000 \times 200$ , where the first number (6) is

the number of channels (shot gathers). The locations and the number of shots can be changed, as long as the configuration of the input layer (i.e., the number of channels) is adjusted accordingly. To maintain consistency with the previous study by Wu and Lin (2019b), we keep the number of shots at six. For each shot gather, there are 200 receivers recording for 4 s with a time interval of 4 ms. The output is a velocity model with a size of  $351 \times 601$ .

InvNet\_UQ shares the same structure with InversionNet, but its loss function is computed using equation 7.4. To compute the SQR loss, first we randomly select six quantile levels ( $\tau$ ) for each training epoch, and the selected  $\tau$  is then fixed for every batch in the epoch. Then the computed six quantile losses are averaged to obtain the final loss. To incorporate quantile level into the neural network, every  $\tau$  is extended to a tensor with the exact size of the seismic data and then added to the input data, so that the number of input channels of InvNet\_UQ is increased to seven.

### 7.2.3 Uncertainty-guided data augmentation

Most existing UQ methods produce the variance of the prediction as the uncertainty map and abandon it after evaluating the prediction accuracy. In contrast, we propose a data-augmentation method to demonstrate the potential of the estimated uncertainty in improving the network performance. The idea of the method is to sample the uncertainty map  $u_a(x^*)$  and add it to the predicted velocity  $\hat{f}(x^*)$  to obtain new velocity models  $\hat{y}_{aug}(x^*)$ :

$$\hat{y}_{aug}(x^*) := \hat{f}(x^*) + w \times u_a(x^*), \quad (7.6)$$

where  $w$  is the weight matrix applied to the computed uncertainty. Ideally, one needs to sample the uncertainty of every pixel with every weight in the range from 0 to 1 to find the actual velocity model. However, this procedure is impossible because it requires an infinite number of weight matrices to fully cover the range  $[0, 1]$ . To simplify the sampling process and ensure the effectiveness of the data-augmentation method, the weights should be randomly selected and have a minimum of 0 and maximum of 1.

An acoustic forward-modeling algorithm simulates seismic data for these new velocity models with the same settings used to generate the initial data. Then we add these new seismic-velocity pairs to the existing training data and retrain the network. Finally, the retrained network is applied to the testing data to estimate the velocity distribution. Because the new velocity models are generated using the prediction and the estimated uncertainty, they provide extra information that the model does not learn from the training data set. Moreover, the inclusion of these new models makes the training data more representative of the testing data distribution.

## 7.3 Results

### 7.3.1 Kimberlina data set

The proposed neural network with UQ is applied to monitoring and predicting CO<sub>2</sub> migration using the synthetic Kimberlina data set. The Kimberlina reservoir model, generated by several institutions (Alumbaugh et al., 2021) as part of the U.S. Department of Energy “SMART Initiative” (U.S. Department of Energy, 2029), was built to simulate a potential commercial-scale geologic carbon storage in the Southern San Joaquin Basin of California (Wagoner, 2009), 30 km northwest of Bakersfield, CA, USA. The Kimberlina model has been released to the public to evaluate the effectiveness and robustness of different geophysical techniques for monitoring CO<sub>2</sub> migration.

There is a total of 29 3D time-lapse P-wave velocity models ( $601 \times 601 \times 351$  grid points) simulated for over 200 years. We slice the 3D velocity model along the  $y$  axis to obtain the corresponding 2D models. The 2D velocity models from different slices have distinct velocity distributions, while those from the same slice (Figure 7.2) share the velocity field outside the reservoir, but inside the reservoir the velocity distribution is changed by the injected CO<sub>2</sub>. The synthetic acoustic wavefield is excited by six shots (point explosions) placed with a constant increment (1 km) at the surface of the model. The source signal is the Ricker wavelet with a central frequency of 10 Hz. We employ 200 receivers evenly distributed along the horizontal line at a depth of 20 m.

### 7.3.2 Implementation details

The entire Kimberlina data set (1,537 velocity samples) is randomly divided into training (80%) and testing (20%) data. InverionNet and InvNet\_UQ are trained with the same data until they converge. Then these networks are applied to the testing data in different scenarios to predict the velocity distribution.

The implementation details are identical for both networks. The Adam optimizer is applied to update the parameters; the batch size is six for the computational efficiency. The initial learning rate is set to 0.0001 in the beginning and decreased to 10% of the previous value after each 15 training epochs. The proposed model has approximately 30 million parameters. Both networks are implemented on Pytorch with two Nvidia GTX 1080 Ti GPUs. The seismic data and the velocity models are normalized before being processed by the network.

To evaluate the predictions, we focus on three critical geologic structures (Figure 7.2; pointed by red arrows): the low-velocity layer immediately above the reservoir region (Zone 1), the three reservoirs (Zone 2), and the high-velocity dipping layer beneath the reservoirs (Zone 3).

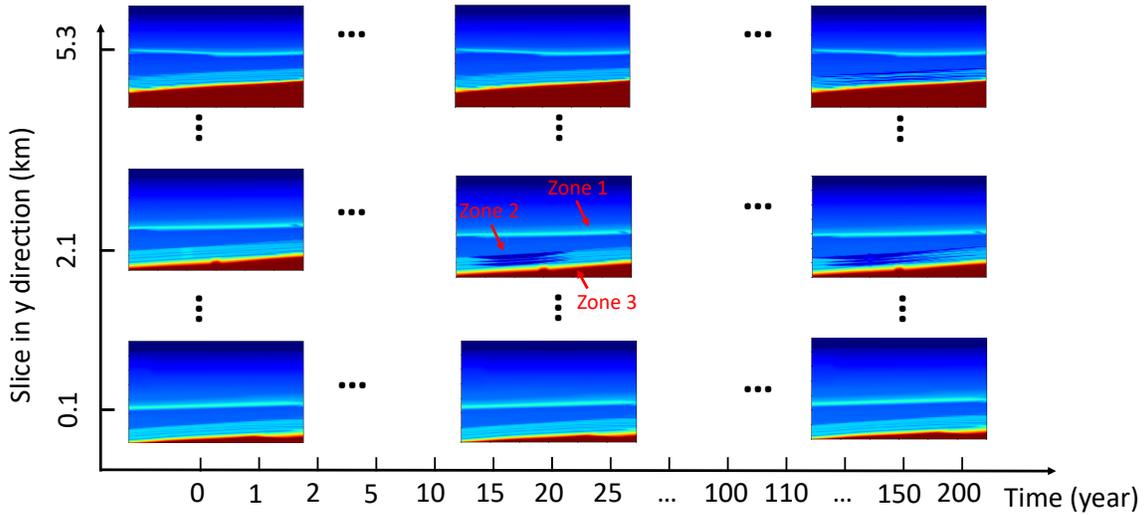


Figure 7.2 Visualization of the actual 2D models sliced from the 3D Kimberlina time-lapse velocity model along the  $y$ -direction. Three important geologic structures are marked by red arrows.

### 7.3.3 Noise-free data

First, we test the trained networks on the original (perfect) testing data. The predictions (Figure 7.3) show that most features of the velocity model can be predicted by both InversionNet (Figure 7.3e-h) and InvNet\_UQ (Figure 7.3i-l) with sufficient accuracy, although there exists an area with larger errors. For example, compared with the actual model (Figure 7.3), both networks cannot accurately reconstruct the small plumes at year 1 (Figure 7.3a) because of the lack of training data at the beginning of the injection when the plumes are small. InversionNet tends to underestimate the size of the plumes (Figure 7.3e), while InvNet\_UQ is more likely to overestimate them (Figure 7.3i).

In addition, both methods could not reconstruct the bumps in the high-velocity dipping layer (circled by a dashed red line) and the anomaly in the low-velocity horizon above the reservoirs (circled by a solid red line) because of the lack of training data for these structures.

For the uncertainty map, the 95% prediction interval around the median is calculated from equation 7.5 by setting the prediction interval  $\alpha = 0.05$ . We verify the uncertainty predicted by InvNet\_UQ (Figure 7.4e-h) by comparing it with the error map (Figure 7.4a-d). High uncertainty is observed around the low-velocity area above the reservoirs, especially near the anomaly (Zone 1) and the reservoir (Zone 2), as well as near the small bump in Zone 3, which matches the area with the most significant errors. The difference between the two confidence intervals, calculated from  $(p_{\text{upper}} - p_{\text{median}}) - (p_{\text{median}} - p_{\text{lower}}) \neq 0$  (Figure 7.4i-l), demonstrates the usefulness of the asymmetric

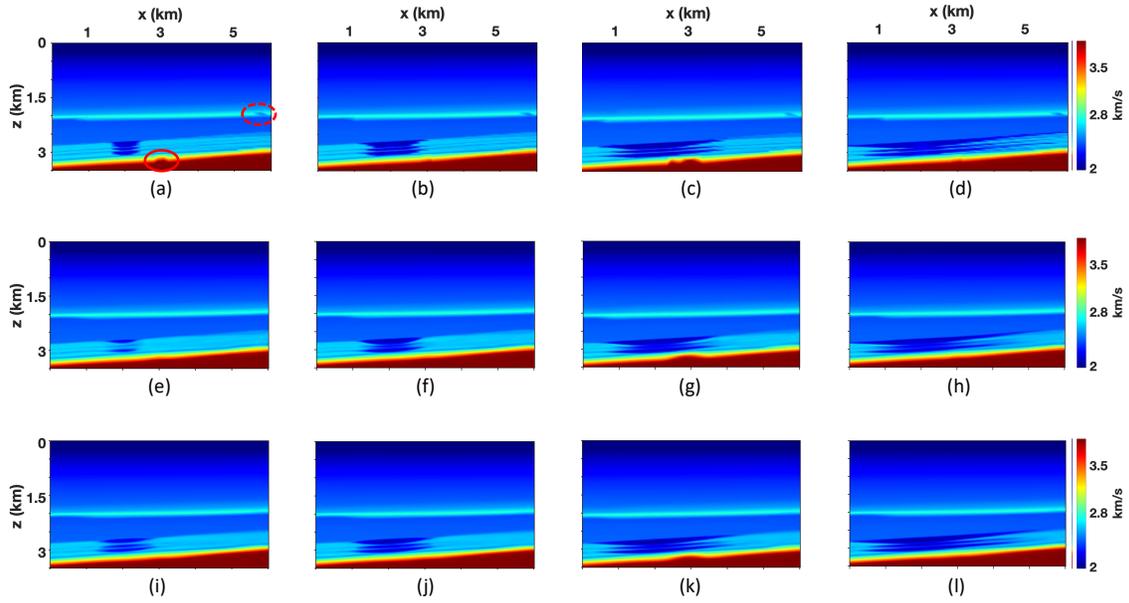


Figure 7.3 Actual P-wave velocity in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. The P-wave velocity obtained from noise-free seismic data by (e-h) InversionNet and (i-l) InvNet\_UQ.

uncertainty prediction, which cannot be provided by conventional UQ methods.

We also calculate the Pearson correlation between the absolute prediction error and the uncertainty and present the corresponding scatter plot in Figure 7.5. The Pearson correlation is 0.727, and the p-value is  $1.0 \times 10^{-10}$ , which shows a strong positive correlation between the two variables. This correlation implies that the uncertainty is an accurate indicator of the prediction confidence, and indicates that InvNet\_UQ can produce an uncertainty map sufficient to facilitate decision-making.

### 7.3.4 Influence of random noise

Next, the testing data are contaminated with Gaussian noise with the signal-to-noise ratio equal to 10 and 20. The noise is added only to the testing data, whereas the neural networks are still trained on noise-free samples. Here, we show only the prediction results for (SNR=10), which is more realistic for field data. Compared with the clean data (Figure 7.6a) where reflection events are apparent, reflections are not clearly visible in the noisy data (Figure 7.6b). Note that testing on noisy data is performed without fine-tuning the network.

Predictably, both networks produce more errors in the velocity field estimated from the noise-contaminated data (Figure 7.7). For the velocity model obtained by InversionNet (Figure 7.7a-d), significant distortions are observed in the low-velocity layer above the CO<sub>2</sub> plumes (Zone 1), in the plumes themselves (Zone 2), and near the boundary between the reservoir and the deep high-velocity

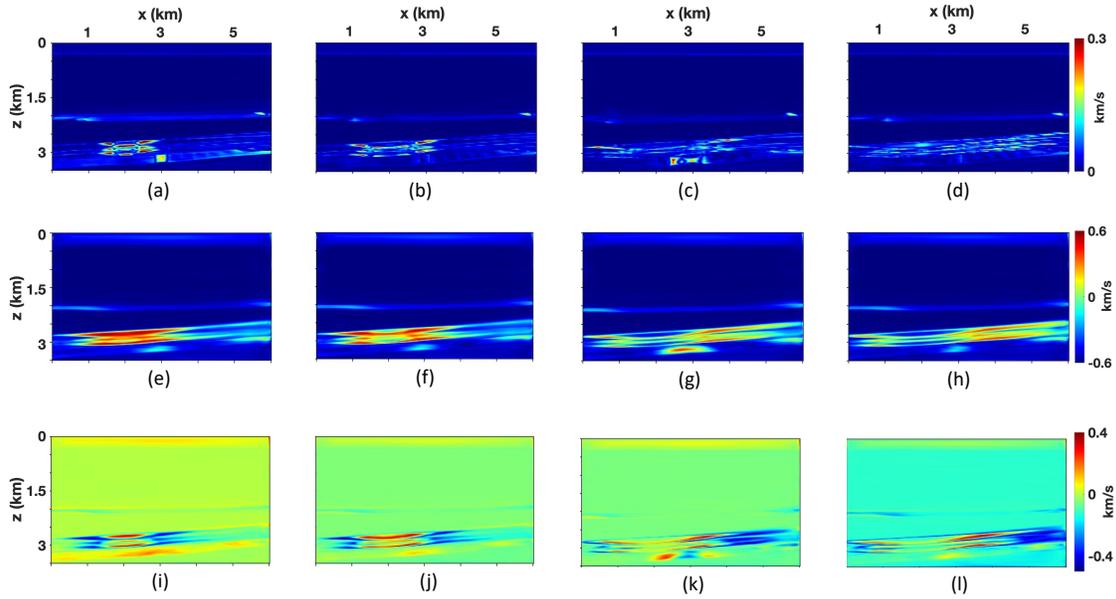


Figure 7.4 Difference between the actual velocity models and those predicted by InvNet\_UQ in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. Plots (e-h) are the corresponding uncertainty maps. Plots (i-l) show the asymmetric uncertainty estimated from  $(p_{\text{upper}} - p_{\text{median}}) - (p_{\text{median}} - p_{\text{lower}})$ , where  $p_{\text{upper}}$  and  $p_{\text{lower}}$  are the predictions at quantiles  $\tau = 0.975$  and  $\tau = 0.025$ , respectively, and  $p_{\text{median}}$  is the median prediction at  $\tau = 0.5$ .

horizons (Zone 3). The  $\text{CO}_2$  plumes are barely visible, which would complicate monitoring  $\text{CO}_2$  injection.

In contrast, InvNet\_UQ (Figure 7.7e-h) reconstructs Zones 1 and 3 with sufficient resolution. The errors are mainly concentrated near the  $\text{CO}_2$  plumes in Zone 2, which is consistent with the calculated uncertainty map (Figure 7.7i-l). In addition, the noise does not significantly distort the predictions, which confirms that the proposed network remains robust for at least moderate noise levels.

The MSE (Figure 7.8) of the predictions by InversionNet jumps from  $7.7 \times 10^{-4}$  for the clean data to  $1.5 \times 10^{-2}$  for the noisy data (SNR=10). In contrast, there is only a slight increase in the MSE of InvNet\_UQ (i.e., from  $8.0 \times 10^{-4}$  to  $1.9 \times 10^{-3}$ ). Evidently, our method can handle noisy data with a realistic signal-to-noise ratio, and the noise does not distort the uncertainty quantification.

### 7.3.5 Influence of missing traces

Missing traces are typical in seismic surveys due to unavailable receivers and low SNR on some records. The commonly used remedies are reconstruction of the missing traces using interpolation and compressive sensing. However, accurate reconstruction usually takes a significant amount of time and is computationally costly. Therefore, next we test the effectiveness of InvNet\_UQ in dealing with the missing-data problem by randomly blocking 30% traces (Figure 7.9b) for every testing sample. The

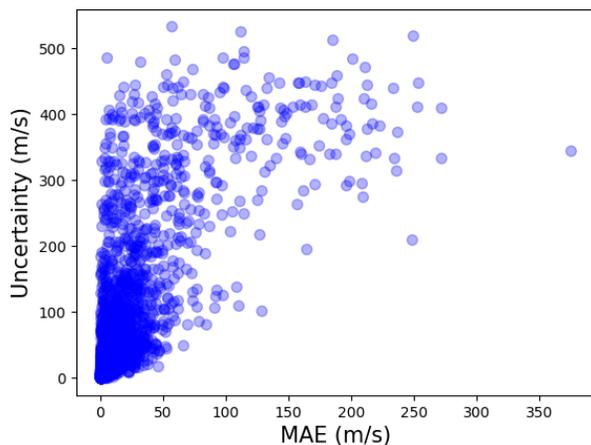


Figure 7.5 Correlation map between the absolute prediction error (MAE) and the corresponding uncertainty. The Pearson correlation is 0.727 (p-value is  $1.0 \times 10^{-10}$ ), which shows a strong positive correlation between the absolute prediction error and the uncertainty.

networks are trained with the complete seismic data. Hence, testing data that miss traces can be referred to as an “out-of-distribution” case.

Compared with the predicted results using the entire data set, both methods produce distortions in the velocity model due to the missing traces. InversionNet (Figure 7.10a-d) cannot predict the CO<sub>2</sub> plume, while InvNet\_UQ (Figure 7.10e-h) still reconstructs these plumes with acceptable resolution. In addition, InvNet\_UQ is able to estimate the bottom right part of the velocity model, which is substantially distorted by InversionNet. Table 7.1 shows that our method is more robust and reduces MSE by about 30% compared with InversionNet.

The uncertainty for the upper part of the model (Figure 7.10i-l; marked by the red arrow) caused by the missing trace is significantly larger than that for the original data. Similar to the previous tests, the reservoirs generally have a higher uncertainty than other parts of the model.

Table 7.1 Comparison of the MSE values for the velocity distribution predicted by InversionNet and InvNet\_UQ applied without data augmentation (columns 2 & 3) and with data augmentation (column 4).

Conditions/Methods	InversionNet	InvNet_UQ	Augmented InvNet_UQ
Noise-free testing data	$7.7 \times 10^{-4}$	$8.0 \times 10^{-4}$	$6.0 \times 10^{-4}$
Testing data with 30% missing traces	$8.0 \times 10^{-3}$	$5.6 \times 10^{-3}$	$9.0 \times 10^{-4}$
Testing data with absence of low frequencies	$9.3 \times 10^{-4}$	$8.0 \times 10^{-4}$	$6.0 \times 10^{-4}$

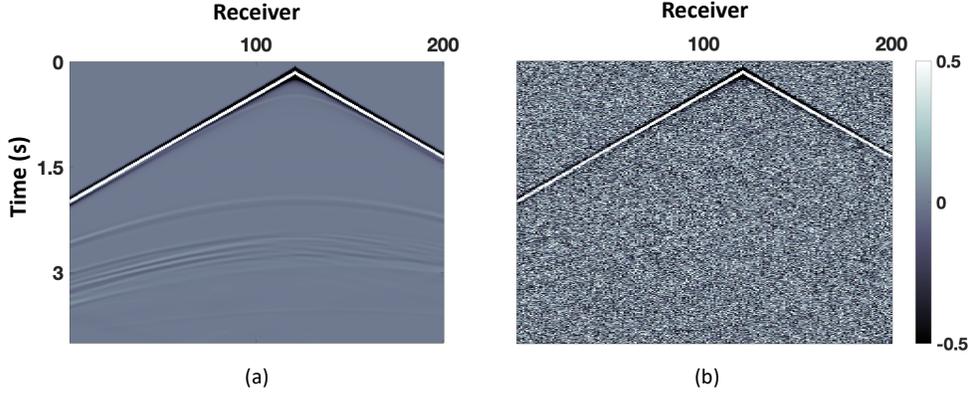


Figure 7.6 Seismic shot gather for the testing velocity model in year 20 (a) without noise, and (b) with Gaussian noise (the signal-to-noise ratio SNR is 10).

### 7.3.6 Influence of absent low-frequency data

Ultra-low-frequency seismic data are essential for the success of FWI, but they are seldom acquired in the field. Therefore, it is important to test the robustness of our algorithm in the absence of low frequencies (less than 3 Hz). We apply the Fourier transform to the perfect testing data and remove frequencies below 3 Hz with a Butterworth filter. Note that both neural networks are trained on the perfect training data that include the full frequency range.

Compared with the benchmark results, the velocity models produced by InversionNet are noticeably distorted (Figure 7.12a-d), even though the low frequencies do not dramatically influence the input seismic data (Figure 7.11). Specifically, InversionNet fails to reconstruct small CO<sub>2</sub> plumes, which could be indicative of CO<sub>2</sub> leakages. InvNet\_UQ (Figure 7.12e-h) can still predict the plumes in the early stages of the injection with the resolution similar to that of the benchmark sections (Figure 7.3i-l), as revealed by the computed uncertainty map. The MSE also demonstrates that InvNet\_UQ ( $8.0 \times 10^{-4}$ ) remains robust compared to the perfect data and achieves higher resolution than InversionNet ( $9.3 \times 10^{-4}$ ).

### 7.3.7 Test of uncertainty-guided data augmentation

We augment the velocity models predicted from the perfect data (see equation 7.6). The uncertainty map is sampled with different weights,  $w=[0.1, 0.3, 0.5, 0.7, 0.9]$ . Compared to the prediction (Figure 7.13b), the augmented data (Figure 7.13c-f) contain small CO<sub>2</sub> plumes and bumps in Zone 3 (marked by the red arrows), which improves the prediction of these features. These newly generated data are used together with the existing training data to retrain InvNet\_UQ. Then the trained network is applied to the testing data. The training process is the same as in the previous tests and is described in the subsection

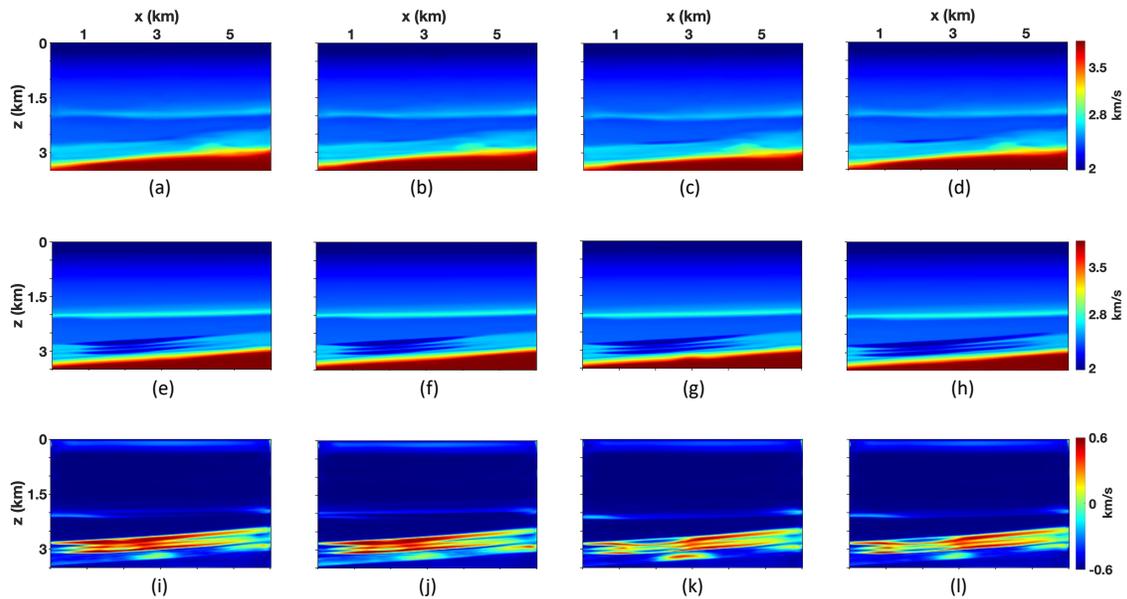


Figure 7.7 P-wave velocity obtained from seismic data with SNR=10 by (a-d) InversionNet and (e-h) InvNet\_UQ. Plots (i-l) show the corresponding uncertainty computed by InvNet\_UQ.

*Implementation details.*

The testing loss of the perfect data shows that training on the augmented data reduces the MSE of the testing data from  $8.0 \times 10^{-4}$  to  $6.0 \times 10^{-4}$ , which demonstrates the effectiveness of our data augmentation method. The detailed comparison between the velocity map (Figure 7.14a-d) and the benchmark result (Figure 7.3i-l) shows that the network trained with the augmented data can improve the reconstruction of the velocity model. In particular, with the augmented data, the shape and amplitude of the predicted CO<sub>2</sub> plumes are closer to those for the actual models, even in earlier years when the plumes are relatively small. In addition, the small bump in the high-velocity zone is captured more accurately. The calculated uncertainty further illustrates that the InvNet\_UQ has a much higher accuracy in predicting the CO<sub>2</sub> plumes (Zone 2), the low-velocity layer (Zone 3), the high-velocity dipping layer (Zone 1), and the upper part of the velocity model. Likewise, the ability of the proposed data-augmentation method to improve the prediction accuracy is demonstrated in the above tests for data with missing traces and missing low frequencies (Table 7.1).

**7.4 Conclusions**

We developed a convolutional neural network (CNN) with Simultaneous Quantile Regression (SQR) to predict the time-lapse velocity model and estimate the uncertainty map from the input seismic data. SQR is designed to build a network that estimates the conditional distributions of all pertinent quantiles

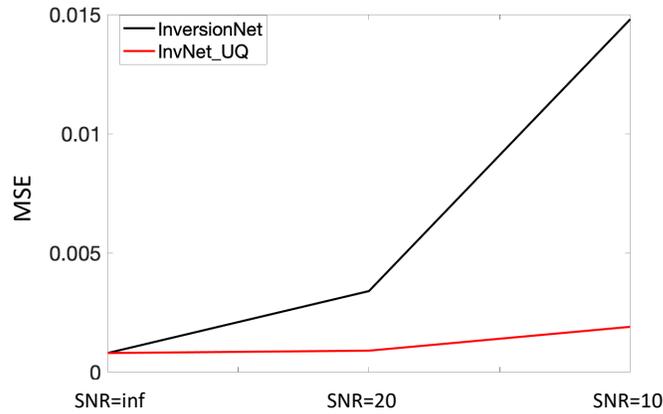


Figure 7.8 Comparison of InversionNet and InvNet\_UQ in terms of the mean-square error (MSE) vs. signal-to-noise ratio (SNR).

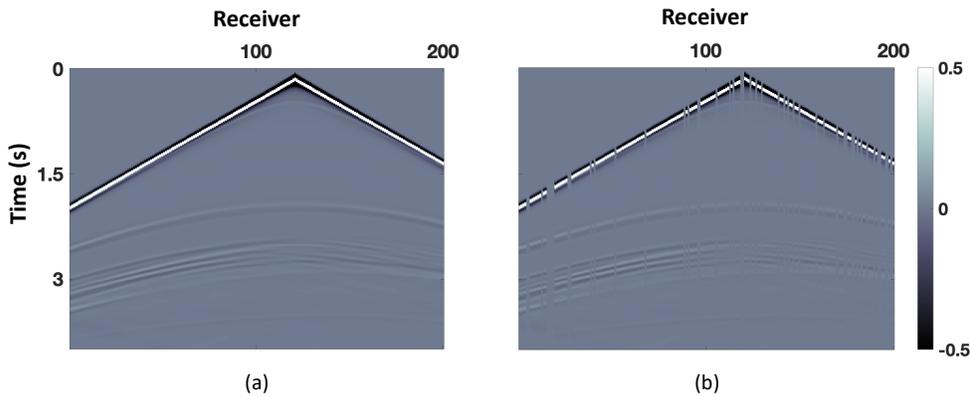


Figure 7.9 Seismic shot gather for the testing velocity model in year 20 (a) without missing data, and (b) with 30% of data missing.

utilizing the pinball loss. Then the uncertainty map is obtained from the prediction interval.

In contrast to most conventional UQ methods, the uncertainty calculated here is related to the predicted velocity and has a physical meaning (i.e., the confidence interval). To fully utilize the computed uncertainty, we propose to employ data augmentation by sampling the uncertainty, adding the newly sampled data to the existing data set, and retraining the network. To benchmark the developed method, the pinball loss is replaced by the L1-norm objective function to form a deterministic neural network (InversionNet). Compared to InversionNet, quantile-regression-based InvNet\_UQ makes no assumptions about the distribution of the target variable and is more robust in the presence of outlying observations.

We test InvNet\_UQ on realistic synthetic data from the Kimberlina reservoir and compare the results with those from InversionNet. The results for clean data demonstrate that InvNet\_UQ can predict the

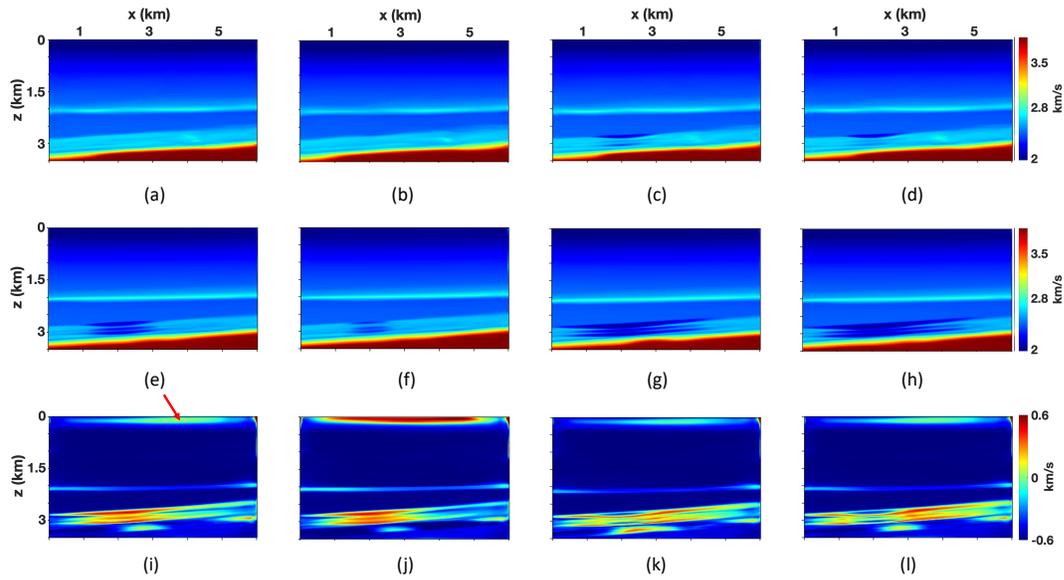


Figure 7.10 P-wave velocity obtained from the testing seismic data with 30% of the traces missing by (a-d) InversionNet and (e-h) InvNet\_UQ. Plots (i-l) show the corresponding uncertainty from InvNet\_UQ.

velocity model with sufficient accuracy. The estimated uncertainty map is consistent with the error map, which implies that the uncertainty can be used to evaluate the prediction accuracy. The tests on data with noise, missing traces, etc. show that InvNet\_UQ remains robust and produces acceptable velocity maps, whereas the output of InversionNet is substantially distorted. This implies that the proposed method is applicable to field data and can facilitate confident decision-making. Retraining the network using the augmented data from the computed uncertainty improves the prediction accuracy, especially in areas with larger parameter variations including the CO<sub>2</sub> plumes.

The developed neural network and data-segmentation method could be applied to other types of reservoirs, such as those containing hydrogen, as long as the seismic data are sufficiently sensitive to the fluid changes. However, in that case retraining would be needed due to the generalization issue common for data-driven neural networks.

## 7.5 Data and Codes Availability

Kimberlina data set can be downloaded from the U.S. Department of Energy - Energy Data exchange Website (<https://edx.netl.doe.gov/dataset/kimberlina-1-2-ccus-geophysical-models-and-synthetic-data-sets>). InversionNet codes are released and can be downloaded from OpenFWI Website (<https://openfwi-lanl.github.io/>).

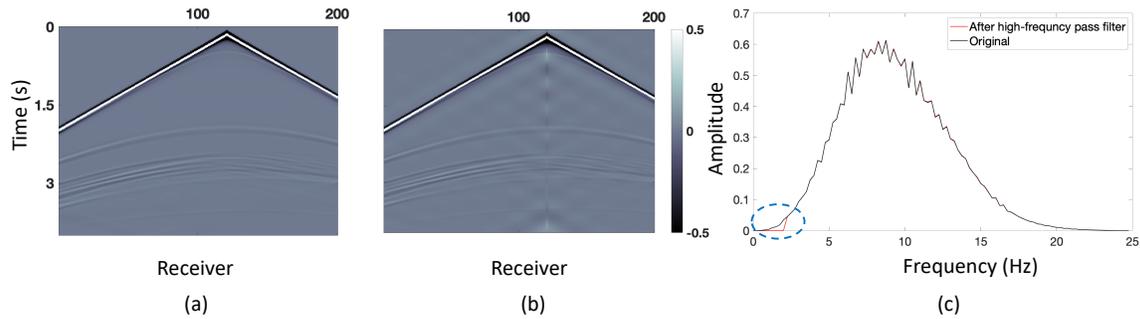


Figure 7.11 Seismic shot gather for the testing velocity model in year 20: (a) without a high bandpass filter, and (b) with the Butterworth filter that has a cut-off frequency at 3 Hz. Plot (c) shows the frequency spectrum of the data. The blue dashed line on plot (c) marks the frequencies removed by the filter.

## 7.6 Acknowledgments

This work was funded by the U.S. Department of Energy (DOE) Office of Fossil Energy’s Carbon Storage Research Program via the Science-Informed Machine Learning to Accelerate Real Time Decision Making for Carbon Storage (SMART-CS) Initiative. Y. Liu and I. Tsvankin also acknowledge the support of the sponsors of the Center for Wave Phenomena (CWP) in the Department of Geophysics at Colorado School of Mines.

## 7.7 Author contributions statement

Y. Lin proposed and supervised the research. Y. Liu and X. Zhang conceived the design of the experiment(s). Y. Liu conducted the experiment(s). All authors contributed ideas to the project, analyzed the results, and reviewed the manuscript.

## 7.8 Additional information

**Competing interests** The authors declare no competing interests.

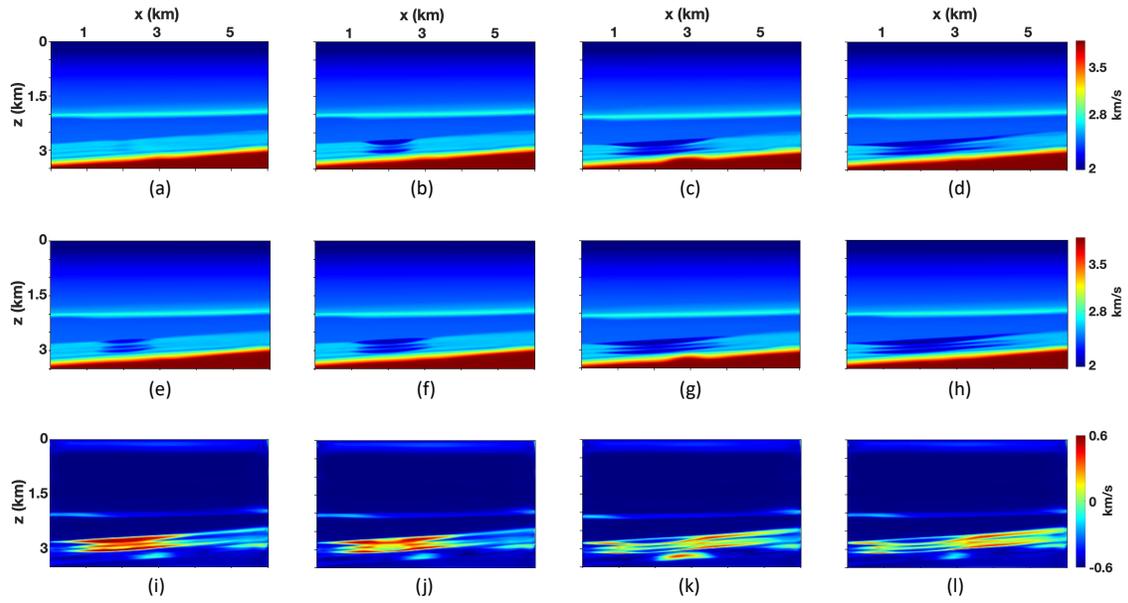


Figure 7.12 P-wave velocity obtained from the testing seismic data using a high bandpass filter with a cut-off frequency of 3 Hz by (a-d) InversionNet and (e-h) InvNet\_UQ. Plots (i-l) show the corresponding uncertainty from InvNet\_UQ.

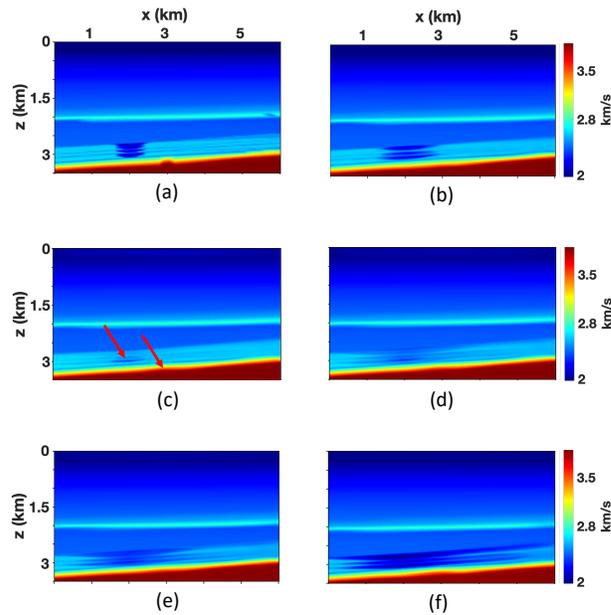


Figure 7.13 Velocity model in year 1: (a) actual and (b) predicted. Plots (c) - (f) show the augmented data from the prediction.

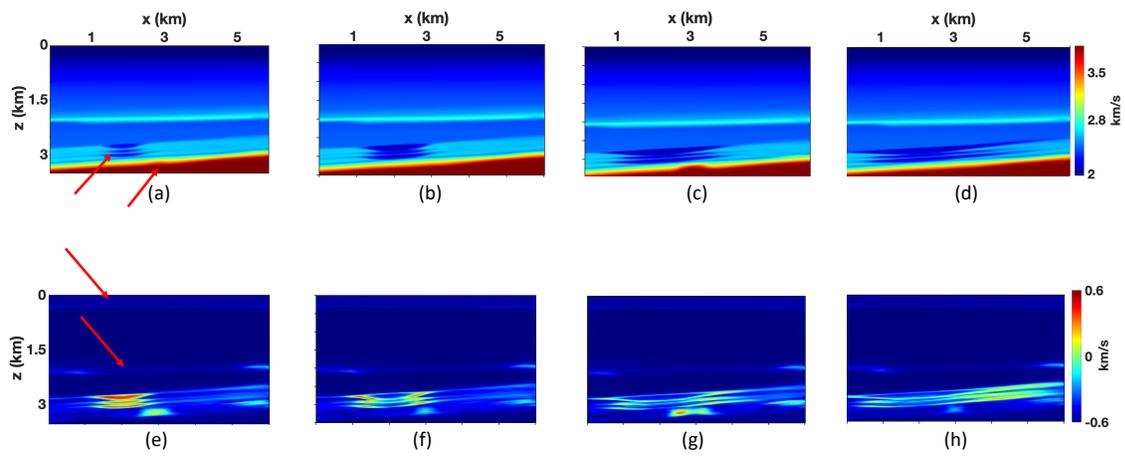


Figure 7.14 P-wave velocity predicted by InvNet\_UQ trained using the augmented data in (a) year 1, (b) year 5, (c) year 20, and (d) year 150. Plots (e)-(h) show the corresponding uncertainty maps.

## CHAPTER 8

### CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

This thesis presents novel algorithms for inverting time-lapse seismic data using physics-based and data-driven techniques, with a particular focus on elastic anisotropic media. The developed methods enable high-resolution reconstruction of time-lapse parameter changes for realistic subsurface models and are applicable to both hydrocarbon production and CO<sub>2</sub> sequestration. This chapter provides a summary of the key findings of the thesis and recommendations for future research.

#### 8.1 Summary

Most existing time-lapse (4D) full-waveform inversion (FWI) methodologies are limited to isotropic and, often, acoustic media, which reduces their applicability. In Chapter 2, I extended time-lapse FWI to 2D elastic VTI (transversely isotropic with a vertical symmetry axis) models and implemented it with three strategies: the parallel-, sequential-, and double-difference methods (PD, SD, and DD, respectively). Testing on realistic synthetic models showed that the developed algorithm can accurately reconstruct the time-lapse variations even without ultra-low frequencies in the data. For low-noise multicomponent data, the double-difference (DD) strategy produces the fewest artifacts outside the target zone because it focuses specifically on the time-lapse data response. Still, a pronounced false anomaly in the P-wave normal-moveout velocity shows the susceptibility of the DD approach to parameter trade-offs (in this case, between the S-wave vertical and P-wave normal-moveout velocities).

FWI requires an accurate estimate of the source wavelet because wavelet distortions may hinder the matching of the observed and simulated data. Therefore, in Chapter 3 I incorporated the “source-independent” (SI) technique, designed to mitigate the dependence of the inverted parameters on the accuracy of the source wavelet, into the time-lapse FWI algorithm for VTI media. Synthetic tests confirm that the SI algorithm helps accurately reconstruct the time-lapse variations even for significantly distorted source signals (such as the spike wavelet). Robust application of the SI technique requires using near-offset reference traces and limiting the size of the time window so that it includes only the direct wave. The testing also confirms that the method remains robust for noisy data and strongly heterogeneous media, although the results for Marmousi model are somewhat inferior to those for simpler structures.

To handle a wider range of subsurface structures, in Chapter 4 the 4D FWI methodology is extended to tilted TI media (TTI). The algorithm is tested on the anticlinal section of the BP TTI model using the three time-lapse strategies mentioned above. The algorithm reconstructs the time-lapse variations of the

TTI parameters with high spatial resolution when the baseline and monitor data are sufficiently repeatable, even in the presence of moderate noise. The SI technique successfully mitigates the impact of distortions in the estimated wavelet on the inversion results. I also discussed the influence of different nonrepeatability (NR) issues on the output of time-lapse FWI. NR-related problems are best handled with the SD and DD strategies because the PD strategy operates with a smoothed, low-resolution initial model, for which NR-induced time shifts between the simulated and observed data may lead to reflector mispositioning, etc. The nonrepeatability tests provide valuable information for field-data applications of time-lapse FWI.

The developed time-lapse FWI algorithm for elastic TI media (Chapter 3) was also applied to a 4D streamer data set from Pyrenees field in Australia. Because the dip of the reflectors is mild, the model was assumed to be VTI. Comparison of the results of the isotropic and VTI FWI reveals the improvement achieved by accounting for anisotropy in reconstructing the subsurface model. The increase in spatial resolution is especially significant in the reservoir area and the overlying layers. The inverted time-lapse parameter variations provide useful information about the reservoir production. This case study confirms the effectiveness of the developed 4D FWI methodology, which performed well despite serious nonrepeatability issues with the Pyrenees data.

To reduce the computational time and alleviate the ill-posedness of FWI, in Chapter 6 I proposed a time-lapse workflow (hybrid strategy) that combines FWI of the baseline data with a convolutional neural network (CNN; InvNet-VelSat) in monitoring of CO<sub>2</sub> injection. FWI is used to estimate the baseline model and then InvNet-VelSat is applied to the monitor data. To avoid employing a large volume of training data required by most machine-learning (ML) techniques, I developed a data-generation algorithm by perturbing the inverted baseline model under constraints provided by well logs and physics laws. Testing on the synthetic Kimberlina data set confirms that the data-generation algorithm adequately captures the spatial and temporal dynamics of CO<sub>2</sub> movement in the reservoir. In addition to predicting the saturation model, the hybrid strategy predicts the velocity field with the accuracy comparable to that of time-lapse FWI.

Finally, in Chapter 7, Simultaneous Quantile Regression (SQR) is incorporated into the CNN developed in Chapter 6 with the goal of predicting the time-lapse velocity model and estimating the uncertainty map from seismic data. The SQR-based CNN improves the prediction accuracy for the Kimberlina data set, especially when the testing data are noisy, have missing traces, and/or do not include low frequencies. To fully utilize the obtained uncertainty estimates, I proposed to employ data augmentation by sampling the uncertainty, adding the newly sampled data to the existing data set, and retraining the network. Synthetic testing confirms the effectiveness of this approach in increasing the spatial resolution of the estimated velocity field and in reducing the prediction error.

## 8.2 Recommendations for future work

Synthetic tests show that the developed time-lapse FWI algorithm can be potentially used to update all medium parameters of arbitrarily anisotropic media including density. However, anisotropic multiparameter inversion (especially if density is included) is problematic in practice because of parameter trade-offs and sensitivity to noise. Regularization and various constraints have proved to be an effective tool in reducing the nonlinearity of the FWI objective function. Therefore, I recommend to supplement the FWI methodology by adding constraints, such as geologic information and regularization of the data and/or the model, to guide the model-updating algorithm toward the global minimum of the objective function.

The presented FWI algorithm uses a staggered finite-difference method for wavefield modeling. However, the sources and receivers have to be placed on the fixed-size grid, which may be problematic for field data. The mispositioning of the sources and receivers may hinder the matching of the observed and simulated data and distort the inversion results. More sophisticated approaches (e.g., multigrid methods) can be applied in wavefield simulation to handle realistic acquisition geometries.

Whereas the algorithms discussed here are 2D, field data are usually acquired in 3D geometry. Wave propagation outside the vertical incidence plane might lead to errors in 2D modeling and inversion. Although out-of-plane corrections reduce the amplitude and phase differences between 2D and 3D data, lateral heterogeneity could still hinder field applications of 2D algorithms. Therefore, it is desirable to extend the presented methodology to 3D models and also account for azimuthal anisotropy.

The source-independent (SI) technique can handle errors and nonrepeatability of the source wavelet when the baseline and monitor data are inverted separately, as is done in the parallel- and sequential-difference strategies. For the double-difference (DD) strategy, the “composite” data used in the monitor inversion include the difference between the baseline and monitor data sets. The nonrepeatability of the source wavelet distorts the “composite” data and, therefore, the monitor inversion. Hence, more sophisticated approaches to generate the “composite” data are needed to reduce the influence of the source wavelet on the output of the DD method.

Because the objective function of the SI-based FWI is based on the  $L_2$ -norm, model updating is sensitive to the amplitude matching of the observed and simulated data, which is often problematic in the field due to complicated wave-propagation phenomena that are not accounted for by the employed models. Therefore, more advanced objective functions, such as the ones that normalize the data trace by trace need to be explored in field-data application of the SI FWI.

I have tested the developed time-lapse FWI algorithm for VTI media on the 4D Pyrenees field data. However, the accuracy of the reconstructed parameter variations is difficult to verify in the absence of

independent information about the subsurface. In addition, only the sequential-difference time-lapse strategy was applied because of the poor data repeatability. Application to higher-quality, more repeatable data sets should help better evaluate the robustness of the proposed 4D FWI methodology, as well as of the hybrid inversion that involves CNNs.

The data-driven time-lapse inversion methods developed in Chapters 6 and 7 have been validated on the synthetic acoustic Kimberlina data. However, the presented networks can handle only acoustic isotropic media and, therefore, may not be applicable to most field data sets. Extension to elastic and anisotropic media should make the hybrid time-lapse method more practical.

## REFERENCES

- Ajo-Franklin, J., J. Peterson, J. Doetsch, and T. Daley (2013). High-resolution characterization of a CO<sub>2</sub> plume using crosswell seismic tomography: Cranfield, MS, USA. *International Journal of Greenhouse Gas Control* 18, 497–509.
- Alemie, W. and M. Sacchi (2016). Joint reparametrized time-lapse full-waveform inversion. In *SEG Technical Program Expanded Abstracts*, pp. 1309–1314.
- Ali, A. and A. A. Al-Shuhail (2018). Characterizing fluid contacts by joint inversion of seismic P-wave impedance and velocity. *Journal of Petroleum Exploration and Production Technology* 8(1), 117–130.
- Alkhalifah, T. and R. É. Plessix (2014). A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study. *Geophysics* 79(3), R91–R101.
- Alumbaugh, D., M. Commer, D. Crandall, E. Gasperikova, S. Feng, W. Harbert, Y. Li, Y. Lin, S. Manthila Samarasinghe, and X. Yang (2021). Development of a multi-scale synthetic data set for the testing of subsurface CO<sub>2</sub> storage monitoring strategies. In *American Geophysical Union (AGU 2021)*.
- ArayaPolo, M., J. Jennings, A. Adler, and T. Dahlke (2017). Automated fault detection without seismic processing. *The Leading Edge* 36(3), 208–214.
- Asaka, M., T. Sekine, and K. Furuya (2016). Geologic cause of seismic anisotropy: A case study from offshore western australia. *The Leading Edge* 35(8), 662–668.
- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux (2012). Time-lapse imaging using regularized fwi: a robustness study. In *SEG Technical Program Expanded Abstracts*, pp. 1–5.
- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux (2013). Regularized seismic full waveform inversion with prior model information. *Geophysics* 78(2), 1–49.
- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux (2015). Time-lapse seismic imaging using regularized full-waveform inversion with a prior model: Which strategy? *Geophysical Prospecting* 63(1), 78–98.
- Bai, T. and I. Tsvankin (2019). Source-independent waveform inversion for attenuation estimation in anisotropic media. *Geophysical Prospecting* 67(9), 2343–2357.
- Behera, L. and I. Tsvankin (2009). Migration velocity analysis for tilted transversely isotropic media. *Geophysical Prospecting* 57(1), 13–26.
- Bond, W. L. (1943). The mathematics of the physical properties of crystals. *The Bell System Technical Journal* 22(1), 1–72.
- Borges, F., M. Muzzette, L. E. Queiroz, B. Pereira-Dias, R. Dias, and A. Bulcão (2022). Analysis of water velocity changes in time-lapse ocean bottom acquisitions - a synthetic 2d study in santos basin, offshore brazil. *Journal of Applied Geophysics* 197, 104521.

- Cai, J., G. Duncan, A. Goody, and K. Kostas (2019). Extract 4d signal from two streamer surveys with very different acquisition geometry by least-squares migration over pyrenees fields. In *SEG Technical Program Expanded Abstracts*, pp. 5265–5269.
- Cao, J. and B. Roy (2017). Time-lapse reservoir property change estimation from seismic using machine learning. *The Leading Edge* 36(3), 234–238.
- Carcione, J., S. Picotti, D. Gei, and G. Rossi (2006, 01). Physics and seismic modeling for monitoring co<sub>2</sub> storage. *Pure and Applied Geophysics - PURE APPL GEOPHYS* 163, 175–207.
- Chen, B., D. R. Harp, Y. Lin, E. H. Keating, and R. J. Pawar (2018). Geologic CO<sub>2</sub> sequestration monitoring design: a machine learning and uncertainty quantification based approach. *Applied energy* 225, 332–345.
- Choi, Y. and T. Alkhalifah (2011). Source-independent time-domain waveform inversion using convolved wavefields: Application to the encoded multisource waveform inversion. *Geophysics* 76(5), R125–R134.
- Choi, Y. and T. Alkhalifah (2012). Application of multi-source waveform inversion to marine streamer data using the global correlation norm. *Geophysical Prospecting* 60(4), 748–758.
- Choi, Y. and D.-J. Min (2012). Source-independent elastic waveform inversion using a logarithmic wavefield. *Journal of Applied Geophysics* 76, 13–22.
- Choi, Y., C. Shin, D. J. Min, and T. Ha (2005). Efficient calculation of the steepest descent direction for source-independent seismic waveform inversion: An amplitude approach. *Journal of Computational Physics* 208(2), 455–468.
- Dahlke, T., M. Araya-Polo, C. Zhang, C. Frogner, and T. Poggio (2016). Predicting geological features in 3D seismic data. In *3D Deep Learning Workshop*.
- Davy, R. G., L. Frahm, R. Bell, R. Arai, D. H. N. Barker, S. Henrys, N. Bangs, J. Morgan, and M. Warner (2021). Generating high-fidelity reflection images directly from full-waveform inversion: Hikurangi subduction zone case study. *Geophysical Research Letters* 48(19), e2021GL094981.
- Denli, H. and L. Huang (2009). Double-difference elastic waveform tomography in the time domain. In *SEG Technical Program Expanded Abstracts*, pp. 2302–2306.
- Duncan, G., J. Cai, K. Kostas, T. Perrett, M. Florez, J. Stewart, and S. Kuzmin (2015). 4D Seismic over the Pyrenees Fields. *ASEG Extended Abstracts 2015*(1), 1–4.
- Egorov, A., R. Pevzner, A. Bóna, S. Glubokovskikh, V. Puzyrev, K. Tertyshnikov, and B. Gurevich (2017). Time-lapse full waveform inversion of vertical seismic profile data: Workflow and application to the CO<sub>2</sub>CRC Otway project. *Geophysical Research Letters* 44(14), 7211–7218.
- Fabien-Ouellet, G., E. Gloaguen, and B. Giroux (2017, 03). Time domain viscoelastic full waveform inversion. *Geophysical Journal International* 209(3), 1718–1734.
- Feng, S., L. Fu, Z. Feng, and G. T. Schuster (2021). Multiscale phase inversion for vertical transverse isotropic media. *Geophysical Prospecting*.
- Feng, S. and G. T. Schuster (2019). Transmission+reflection anisotropic wave-equation travelttime and waveform inversion. *Geophysical Prospecting* 67(2), 423–442.

- Feng, S., X. Zhang, B. Wohlberg, N. Symons, and Y. Lin (2021). Monitoring and forecasting CO<sub>2</sub> storage in the Sleipner area with spatio-temporal cnns. In *First International Meeting for Applied Geoscience & Energy Expanded Abstracts*, pp. 1686–1690.
- Furre, A. K., O. Eiken, H. Alnes, J. N. Vevatne, and A. F. Kiær (2017). 20 years of monitoring CO<sub>2</sub>-injection at Sleipner. *Energy Procedia* 114, 3916–3926.
- Gardner, G. H. F., L. W. Gardner, and A. R. Gregory (1974). Formation velocity and density—the diagnostic basics for stratigraphic traps. *Geophysics* 39(6), 770–780.
- Gassmann, F. (1951). Über die Elastizität poröser Medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 96, 1–23.
- Grana, D. and E. D. Rossa (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. *Geophysics* 75(3), O21–O37.
- Hale, D. (2013a). Dynamic warping of seismic images. *Geophysics* 78(2), S105–S115.
- Hale, D. (2013b). Methods to compute fault images, extract fault surfaces, and estimate fault throws from 3D seismic images. *Geophysics* 78(2), 33–43.
- Hicks, E., H. Hoerber, M. Houbiers, S. P. Lescoffit, A. Ratcliffe, and V. Vinje (2016). Time-lapse full-waveform inversion as a reservoir-monitoring tool — A North Sea case study. *The Leading Edge* 35(10), 850–858.
- Holschuh, N. D., C. Li, M. A. Meadows, and S. Dobbs (2014). Systems and methods for aligning a monitor seismic survey with a baseline seismic survey. U. S. Patent 14,565,117: Chevron USA Inc., 1–16.
- Holt, R., A. Bauer, and A. Bakk (2016). The importance of overburden stress path in assessment of stress dependence for 4D applications. In *European Association of Geoscientists & Engineers*, Number 1, pp. 1–5.
- Hornik, K., M. Stinchcombe, and H. White (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks* 3, 551–560.
- Huang, C. and T. Zhu (2020, 07). Towards real-time monitoring: data assimilated time-lapse full waveform inversion for seismic velocity and uncertainty estimation. *Geophysical Journal International* 223(2), 811–824.
- Jervis, M., A. Bakulin, and R. Smith (2018). Making time-lapse seismic work in a complex desert environment for CO<sub>2</sub> EOR monitoring — Design and acquisition. *The Leading Edge* 37(8), 598–606.
- Kamath, N. and I. Tsvankin (2013). Full-waveform inversion of multicomponent data for horizontally layered VTI media. *Geophysics* 78(5), WC113–WC121.
- Kamath, N. and I. Tsvankin (2016). Elastic full-waveform inversion for VTI media: Methodology and sensitivity analysis. *Geophysics* 81(2), C53–C68.
- Kamath, N., I. Tsvankin, and E. Díaz (2017). Elastic full-waveform inversion for VTI media: A synthetic parameterization study. *Geophysics* 82(5), C163–C174.
- Kim, J., Z. Xue, and T. Matsuoka (2010, 06). Experimental Study on CO<sub>2</sub> Monitoring and Saturation with Combined P-wave Velocity and Resistivity. Volume All Days of *SPE International Oil and Gas Conference and Exhibition in China*. SPE-130284-MS.

- Koenker, R. and K. F. Hallock (2001, December). Quantile regression. *Journal of Economic Perspectives* 15(4), 143–156.
- Köhn, D., D. D. Nil, A. Kurzmann, A. Przebindowska, and T. Bohlen (2012). On the influence of model parametrization in elastic full waveform tomography. *Geophysical Journal International* 191(1), 325–345.
- Lecerf, D., M. Lange, A. Oates, and J. Kumar (2022). Water column corrections: Joint water velocity inversion for 4d marine surveys. In *Second International Meeting for Applied Geoscience & Energy*, pp. 3404–3408.
- Lee, K. H. and H. J. Kim (2003). Source-independent full-waveform inversion of seismic data. *Geophysics* 68(6), 2010–2015.
- Li, C., M. Meadows, and T. Dygert (2019). Warping least-squares inversion for 4d velocity change: Gulf of Mexico case study. *Interpretation* 7(2), SB23–SB31.
- Li, D., S. Peng, Y. Guo, Y. Lu, and X. Cui (2021). CO<sub>2</sub> storage monitoring based on time-lapse seismic data via deep learning. *International Journal of Greenhouse Gas Control* 108, 103336.
- Li, N., Z. Yu, R. To, M. Wang, Y. Xie, and D. Dickinson (2021). 4D FWI using towed-streamer data: A case study near Laverda oil field. In *SEG Technical Program Expanded Abstracts*, pp. 622–626.
- Li, Y., T. Alkhalifah, and Q. Guo (2021). Target-oriented time-lapse waveform inversion using deep learning-assisted regularization. *Geophysics* 86, R485–R495.
- Lin, Y. and L. Huang (2015). Quantifying subsurface geophysical properties changes using double-difference seismic-waveform inversion with a modified total-variation regularization scheme. *Geophysical Journal International* 203, 2125–2149.
- Lin, Y., J. Theiler, and B. Wohlberg (2022). Physics-guided data-driven seismic inversion: Recent progress and future opportunities in full waveform inversion. *IEEE Signal Processing Magazine*.
- Liu, F., S. A. Morton, X. Ma, and S. Checkles (2013). Some key factors for the successful application of full-waveform inversion. *The Leading Edge* 32(9), 1124–1129.
- Liu, Q. and J. Tromp (2006). Finite-frequency kernels based on adjoint methods. *Bulletin of the Seismological Society of America* 96(6), 2383–2397.
- Liu, Y., S. Feng, I. Tsvankin, D. Alumbaugh, and Y. Lin (2022). Joint physics-based and data-driven time-lapse seismic inversion: Mitigating data scarcity. *Geophysics* 88, 1–53.
- Liu, Y. and I. Tsvankin (2021). Methodology of time-lapse elastic full-waveform inversion for VTI media. *Journal of Seismic Exploration* 30(3), 257–270.
- Liu, Y. and I. Tsvankin (2022). Source-independent time-lapse full-waveform inversion for anisotropic media. *Geophysics* 87(1), R111–R122.
- Lumley, D. (2001a). Time-lapse seismic reservoir monitoring. *Geophysics* 66, 50–53.
- Lumley, D. (2010). 4D seismic monitoring of CO<sub>2</sub> sequestration. *The Leading Edge* 29(February), 150–155.
- Lumley, D. E. (2001b). Time-lapse seismic reservoir monitoring. *Geophysics* 66(1), 50–53.

- Luo, C. M., S. X. Wang, and S. Y. Yuan (2014). Effect of inaccurate wavelet phase on prestack waveform inversion. *Journal of Applied Geophysics* 11, 479–488.
- Maharramov, M. and B. Biondi (2014). Joint full-waveform inversion of time-lapse seismic data sets. In *SEG Technical Program Expanded Abstracts*, pp. 954–959.
- Mavko, G. and T. Mukerji (1995). Seismic pore space compressibility and Gassmann’s relation. *Geophysics* 60(6), 1743–1749.
- Mavko, G. and T. Mukerji (1998). Bounds on low-frequency seismic velocities in partially saturated rocks. *Geophysics* 63(3), 918–924.
- Pevzner, R., M. Urosevic, D. Popik, V. Shulakova, K. Tertyshnikov, E. Caspari, J. Correa, T. Dance, A. Kepic, S. Glubokovskikh, S. Ziramov, B. Gurevich, R. Singh, M. Raab, M. Watson, T. Daley, M. Robertson, and B. Freifeld (2017). 4D surface seismic tracks small supercritical CO<sub>2</sub> injection into the subsurface: CO<sub>2</sub>CRC Otway Project. *International Journal of Greenhouse Gas Control* 63(March), 150–157.
- 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(2), 495–503.
- Plessix, R.-E., S. Michelet, H. Rynja, H. Kuehl, C. Perkins, J. Maag, and P. Hatchell (2010, 06). Some 3D applications of full waveform inversion. *EAGE Expanded Abstracts*.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model. *Geophysics* 64(3), 888–901.
- Queiβer, M. and S. Singh (2010). *Time lapse seismic monitoring of CO<sub>2</sub> sequestration at Sleipner using time domain 2D full waveform inversion*, pp. 2875–2879.
- Queiβer, M. and S. C. Singh (2013). Full waveform inversion in the time lapse mode applied to CO<sub>2</sub> storage at Sleipner. *Geophysical Prospecting* 61(3), 537–555.
- Raknes, E. B., W. Weibull, and B. Arntsen (2015). Seismic imaging of the carbon dioxide gas cloud at Sleipner using 3D elastic time-lapse full waveform inversion. *International Journal of Greenhouse Gas Control* 42, 26–45.
- Ren, P., Y. Xiao, X. Chang, P.-Y. Huang, Z. Li, B. B. Gupta, X. Chen, and X. Wang (2021). A survey of deep active learning. *ACM Computing Surveys* 54(9).
- Renán, R.-G., J. Yang, Y. Lin, J. Theiler, and B. Wohlberg (2020). Physics-consistent data-driven waveform inversion with adaptive data augmentation. *IEEE Geoscience and Remote Sensing Letters*, 1–5.
- Rickett, J., L. Duranti, T. Hudson, B. Regel, and N. Hodgson (2007). 4d time strain and the seismic signature of geomechanical compaction at genesis. *The Leading Edge* 26(5), 644–647.
- Robinson, E. A. and S. Treitel (1967). Principles of digital wiener filtering. *Geophysical Prospecting* 15(3), 311–332.
- Routh, P., G. Palacharla, I. Chikichev, and S. Lazaratos (2012). Full wavefield inversion of time-lapse data for improved imaging and reservoir characterization. In *SEG Technical Program Expanded Abstracts*, pp. 4361–4366.

- Rusmanugroho, H., R. Modrak, and J. Tromp (2017). Anisotropic full-waveform inversion with tilt-angle recovery. *Geophysics* 82(3), R135–R151.
- Schmidt, M. (2005). Least squares optimization with L1-norm regularization. Technical report.
- Scibiorski, J., M. Micenko, and D. Lockhart (2005). Recent discovery in the Pyrenees member, Exmouth Sub-basin: A new oil play fairway. *The APPEA Journal* 45(1), 233–252.
- Seo, T., C. Shin, and D. Min (2005). Efficient calculation of steepest descent direction of source signature independent waveform inversion of logarithmic wavefield. In *SEG Technical Program Expanded Abstracts*, pp. 1838–1841.
- Sigfusson, B., S. R. Gislason, J. M. Matter, M. Stute, E. Gunnlaugsson, I. Gunnarsson, E. S. Aradottir, H. Sigurdardottir, K. Mesfin, H. A. Alfredsson, D. Wolff-Boenisch, M. T. Arnarsson, and E. H. Oelkers (2015). Solving the carbon-dioxide buoyancy challenge: The design and field testing of a dissolved CO<sub>2</sub> injection system. *International Journal of Greenhouse Gas Control* 37, 213–219.
- Singh, S. and I. Tsvankin (2020). Sensitivity analysis of FWI for elastic orthorhombic media. In *SEG Technical Program Expanded Abstracts*, pp. 171–175.
- Singh, S., I. Tsvankin, and E. Z. Naeini (2020). Full-waveform inversion with borehole constraints for elastic VTI media. *Geophysics* 85, R553–R563.
- Singh, S., I. Tsvankin, and E. Z. Naeini (2021). Facies-based full-waveform inversion for anisotropic media: a north sea case study. *Geophysical Prospecting* 69(8-9), 1650–1663.
- Singh, S., I. Tsvankin, and E. Zahibi Naeini (2018). Bayesian framework for elastic full-waveform inversion with facies information. *The Leading Edge* 37(12), 924–931.
- Singh, S., I. Tsvankin, and E. Zahibi Naeini (2019). Bayesian approach to facies-constrained waveform inversion for VTI media. In *SEG Technical Program Expanded Abstracts*, pp. 1370–1374.
- Smith, S. and I. Tsvankin (2012). Coupled geomechanical and seismic modeling of compaction-induced traveltime shifts for multi-component data. *Geophysics* 77(6), 1687–1692.
- Smith, S. S. and I. Tsvankin (2013). Sensitivity of compaction-induced multicomponent seismic time shifts to variations in reservoir properties. *Geophysics* 78(5), T151–T163.
- Smith, S. S. and I. Tsvankin (2016). Inversion of multicomponent seismic time shifts for reservoir pressure and length: A feasibility study. *Geophysical Prospecting* 64(1), 83–101.
- Song, Z., P. R. Williamson, and R. G. Pratt (1995). Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part ii—inversion method, synthetic experiments and real-data results. *Geophysics* 60(3), 796–809.
- Tagasovska, N. and D. Lopez-Paz (2019). *Single-Model Uncertainties for Deep Learning*. Red Hook, NY, USA: Curran Associates Inc.
- Tang, M., X. Ju, and L. J. Durlofsky (2022). Deep-learning-based coupled flow-geomechanics surrogate model for CO<sub>2</sub> sequestration. *International Journal of Greenhouse Gas Control* 118, 103692.
- Tarantola, A. (1984). Linearized inversion of seismic reflection data. *Geophysical Prospecting* 32(6), 998–1015.

- Thomsen, L. (1986). Weak elastic anisotropy. *Geophysics* 51(10), 1954–1966.
- Tsvankin, I. (2012). *Seismic Signatures and Analysis of Reflection Data in Anisotropic Media, Third edition*. Society of Exploration Geophysicists.
- Um, E. S., D. Alumbaugh, Y. Lin, and S. Feng (2022). Real-time deep-learning inversion of seismic full waveform data for CO<sub>2</sub> saturation and uncertainty in geological carbon storage monitoring. *Geophysical Prospecting n/a(n/a)*.
- U.S. Department of Energy (2019 - 2029). Science-informed machine learning for accelerating real-time decisions in subsurface applications (smart) initiative.
- Venstad, J. M. (2014). Dynamic time warping — an improved method for 4D and tomography time shift estimation? *Geophysics* 79(5), R209–R220.
- Vigh, D., K. Jiao, D. Watts, and D. Sun (2014). Elastic full-waveform inversion application using multicomponent measurements of seismic data collection. *Geophysics* 79(2), R63–R77.
- Virieux, J. and S. Operto (2009). An overview of FWI in Exploration Geophysics. *Geophysics* 74(6), WCC127–WCC152.
- Wagoner, J. (2009, April). 3D geologic modeling of the southern San Joaquin basin for the Westcarb Kimberlina Demonstration Project- a status report. Technical report.
- Waldhauser, F. and W. Ellsworth (2020). A double-difference earthquake location algorithm: Method and application to the northern hayward fault, california. *Bulletin of the Seismological Society of America* 90, 1353–1368.
- Wang, H., Q. Guo, T. Alkhalifah, and Z. Wu (2020). Regularized elastic passive equivalent source inversion with full-waveform inversion: Application to a field monitoring microseismic data set. *Geophysics* 85, KS207–KS219.
- Wang, X. and I. Tsvankin (2013). Multiparameter tti tomography of p-wave reflection and vsp data. *Geophysics* 78(5), WC51–WC63.
- Wang, Z., W. Harbert, R. Dilmore, and L. Huang (2018, 09). Modeling of time-lapse seismic monitoring using CO<sub>2</sub> leakage simulations for a model CO<sub>2</sub> storage site with realistic geology: Application in assessment of early leak-detection capabilities. *International Journal of Greenhouse Gas Control* 76, 39–52.
- Warner, M., A. Ratcliffe, T. Nangoo, J. Morgan, A. Umpleby, N. Shah, V. Vinje, I. Štekl, L. Guasch, C. Win, G. Conroy, and A. Bertrand (2013). Anisotropic 3d full-waveform inversion. *Geophysics* 78(2), R59–R80.
- Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka (2004). Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring. In *SEG Technical Program Expanded Abstracts*, pp. 2323–2326.
- Woodall, M. and C. Stark (2013). The Stratigraphy of the Pyrenees Member of the Ravensworth, Crosby and Stickle Fields, Exmouth Sub-Basin, Northwest Australia – A Wave- and Storm-Dominated Shelf-Margin Delta. In *The Petroleum Exploration Society of Australia Symposium*.
- Wu, Y. and Y. Lin (2019a). InversionNet: An efficient and accurate data-driven full waveform inversion. *IEEE Transactions on Computational Imaging* 6, 419–433.

- Wu, Y. and Y. Lin (2019b). InversionNet: An efficient and accurate data-driven full waveform inversion. *IEEE Transactions on Computational Imaging* 6, 419–433.
- Xu, K., S. A. Greenhalgh, and M. Wang (2006). Comparison of source-independent methods of elastic waveform inversion. *Geophysics* 71(6), R91–R100.
- Xue, Z., J. W. Kim, S. Mito, K. Kitamura, and T. Matsuoka (2009). Detecting and monitoring CO<sub>2</sub> with P-wave velocity and resistivity from both laboratory and field scales. In *SPE International Conference on CO<sub>2</sub> Capture, Storage, and Utilization*. OnePetro.
- Yang, D., F. Liu, S. Morton, A. Malcolm, and M. Fehler (2016). Time-lapse full-waveform inversion with ocean-bottom-cable data: Application on Valhall field. *Geophysics* 81(4), R225–R235.
- Yang, Y., X. Zhang, Q. Guan, and Y. Lin (2021). Making invisible visible: Data-driven seismic inversion with physics-informed data augmentation. *CoRR abs/2106.11892*.
- Yuan, C., X. Zhang, X. Jia, and J. Zhang (2019). Time-lapse velocity imaging via deep learning. *Geophysical Journal International* 220(2), 1228–1241.
- Yuan, F., S. Wang, S. Yuan, J. Wang, J. Li, P. Shi, and Y. Liu (2014, 06). Influence of inaccurate wavelet amplitude on frequency-domain full waveform inversion. In *EAGE Conference Expanded Abstracts*, pp. 1–5.
- Zablotski, Y. (2023, Jan). Quantile regression as an useful alternative for ordinary linear regression. <https://yuzar-blog.netlify.app/posts/2022-12-01-quantileregression/>.
- Zhang, F., C. Juhlin, C. Cosma, A. Tryggvason, and R. G. Pratt (2012). Cross-well seismic waveform tomography for monitoring CO<sub>2</sub> injection: a case study from the Ketzin Site, Germany. *Geophysical Journal International* 189(1), 629–646.
- Zhang, Q., H. Zhou, Q. Li, H. Chen, and J. Wang (2016). Robust source-independent elastic full-waveform inversion in the time domain. *Geophysics* 81(2), R29–R44.
- Zhang, W. and J. Gao (2021). Deep-learning full-waveform inversion using seismic migration images. *IEEE Transactions on Geoscience and Remote Sensing*, 1–18.
- Zhang, X. and J. Zhang (2012). Edge preserving regularization for seismic traveltime tomography. In *SEG Technical Program Expanded Abstracts*, pp. 2830–2834.
- Zhang, Z. and T. Alkhalifah (2020). High-resolution reservoir characterization using deep learning-aided elastic full-waveform inversion: The north sea field data example. *Geophysics* 85(4), WA137–WA146.
- Zhang, Z., T. Alkhalifah, E. Z. Naeini, and B. Sun (2018). Multiparameter elastic full waveform inversion with facies-based constraints. *Geophysical Journal International* 213(3), 2112–2127.
- Zhang, Z. and Y. Lin (2020). Data-driven seismic waveform inversion: A study on the robustness and generalization. *IEEE Transactions on Geoscience and Remote Sensing* 58(10), 6900–6913.
- Zhou, B. and S. A. Greenhalgh (2003). Crosshole seismic inversion with normalized full-waveform amplitude data. *Geophysics* 68(4), 1320–1330.
- Zhou, W. and D. Lumley (2021a). Central-difference time-lapse 4d seismic full-waveform inversion. *GEOPHYSICS* 86(2), R161–R172.

Zhou, W. and D. Lumley (2021b). Nonrepeatability effects on time-lapse 4D seismic full-waveform inversion for ocean-bottom node data. *Geophysics* 86(4), R547–R561.

## APPENDIX A

### INVERSION GRADIENTS FOR THE SOURCE-INDEPENDENT FWI

Starting with Choi and Alkhalifah (2011), several publications discuss the “source-independent” FWI in the time domain, but none of these papers (e.g., Bai and Tsvankin, 2019; Zhang et al., 2016) presents explicit expressions for the inversion gradients. Following Kamath and Tsvankin (2016), we present an explicit derivation of the gradient of the source-independent FWI with respect to the VTI parameters. Note that the objective function and the generation of the back-propagated wavefield here are more complex.

To minimize the objective function (equation 3.2), the simulated wavefield  $\mathbf{d}^{\text{sim}}(\mathbf{x}_r, t)$  is generated using the wave equation for arbitrarily anisotropic, heterogeneous media:

$$\rho \frac{\partial^2 d_i^{\text{sim}}}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial d_k^{\text{sim}}}{\partial x_l} \right) = f_i, \quad (\text{A.1})$$

where  $\mathbf{d}$  is displacement field,  $\rho$  is the density,  $\mathbf{f}$  is the density of the body forces, and  $c_{ijkl}$  ( $i, j, k, l = 1, 2, 3$ ) are the stiffness coefficients.

The initial conditions for the displacement are:

$$\mathbf{d}^{\text{sim}}(\mathbf{x}, 0) = 0, \quad \frac{\partial \mathbf{d}^{\text{sim}}(\mathbf{x}, 0)}{\partial t} = 0, \quad (\text{A.2})$$

and the radiation boundary condition,

$$\mathbf{d}^{\text{sim}}(\mathbf{x}, t)|_{\mathbf{x} \rightarrow \infty} \rightarrow 0. \quad (\text{A.3})$$

Adopting the Lagrange multiplier method, we can define the Lagrangian  $\Delta$  as:

$$\begin{aligned} \Delta = & \frac{1}{2} \sum_r \int_0^T \|\mathbf{d}^{\text{sim}}(\mathbf{x}, t) * W \mathbf{d}_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - \mathbf{d}^{\text{obs}}(\mathbf{x}, t) * W d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t)\|^2 dt \\ & - \int_0^T \int_{\Omega} \lambda_i \left[ \rho \frac{\partial^2 d_i^{\text{sim}}}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial u_k}{\partial x_l} \right) - f_i \right] dV dt, \end{aligned} \quad (\text{A.4})$$

where  $\Omega$  is the integration domain (volume  $V$ ), and  $\lambda(\mathbf{x}, t)$  is the yet unknown Lagrangian multiplier. After intergration by parts and application of the Gaussian divergence theorem, the change in the Lagrangian

can be found in the following form:

$$\begin{aligned}
\delta\Delta = & - \int_0^T \sum_{r=1}^N \left[ d_i^{\text{sim}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - d_i^{\text{obs}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) \right] d_i^{\text{obs}}(\mathbf{x}, t) \delta(\mathbf{x} - \mathbf{x}_r) \delta d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) dt \\
& \int_0^T \int_{\Omega} \sum_{r=1}^N \left[ d_i^{\text{sim}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - d_i^{\text{obs}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) \right] d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) \delta(\mathbf{x} - \mathbf{x}_r) \delta d_i^{\text{sim}}(\mathbf{x}, t) dV dt \\
& - \int_0^T \int_{\Omega} \delta c_{ijkl} \frac{\partial d_k^{\text{sim}}}{\partial x_l} \frac{\partial \lambda_i}{\partial x_j} dV dt - \int_0^T \int_{\Omega} \left[ \rho \frac{\partial^2 \lambda_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial \lambda_k}{\partial x_l} \right) \right] \delta d_i^{\text{sim}} dV dt \\
& - \int_{\Omega} \left[ \rho \lambda_i \frac{\partial (\delta d_i^{\text{sim}})}{\partial t} - \rho \delta d_i^{\text{sim}} \frac{\partial \lambda_i}{\partial t} \right] \Big|_0^T dV \\
& + \int_0^T \int_{\partial\Omega} \lambda_i \left[ \delta c_{ijkl} \frac{\partial d_k^{\text{sim}}}{\partial x_l} + c_{ijkl} \frac{\partial (\delta d_k^{\text{sim}})}{\partial x_l} \right] n_j dS dt - \int_0^T \int_{\partial\Omega} \delta d_i^{\text{sim}} c_{ijkl} \frac{\partial \lambda_k}{\partial x_l} n_j dS dt,
\end{aligned} \tag{A.5}$$

where  $\partial\Omega$  is the surface of  $\Omega$ ,  $\mathbf{n}$  is the vector normal to  $\partial\Omega$ , and  $r = 1, 2, 3 \dots N$  denotes the receivers (Kamath and Tsvankin, 2016). Following Choi and Alkhalifah (2011), the second term of equation A.5 can be ignored because it operates with only the reference trace. The initial and boundary conditions for the perturbation in  $\mathbf{d}^{\text{sim}}(\mathbf{x}, t)$  are given by:

$$\delta \mathbf{d}^{\text{sim}}(\mathbf{x}, 0) = 0, \quad \frac{\partial \delta \mathbf{d}^{\text{sim}}(\mathbf{x}, 0)}{\partial t} = 0, \quad \delta \mathbf{d}^{\text{sim}}(\mathbf{x}, t)|_{\mathbf{x} \rightarrow \infty} \rightarrow 0. \tag{A.6}$$

Assuming that there are no parameter perturbations, the wavefield  $\boldsymbol{\lambda}$  at time  $T$  is subject to the following conditions:

$$\boldsymbol{\lambda}(\mathbf{x}, T) = 0, \quad \frac{\partial \boldsymbol{\lambda}(\mathbf{x}, T)}{\partial t} = 0, \tag{A.7}$$

and the radiation boundary condition,

$$\boldsymbol{\lambda}(\mathbf{x}, t)|_{\mathbf{x} \rightarrow \infty} \rightarrow 0. \tag{A.8}$$

Using equations A.7 and A.8, equation A.5 can be reduced to:

$$\begin{aligned}
\delta\Delta = & \int_0^T \int_{\Omega} \sum_{r=1}^N \left[ d_i^{\text{sim}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - d_i^{\text{obs}}(\mathbf{x}, t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) \right] d_i^{\text{obs}}(\mathbf{x}_f, t) \delta(\mathbf{x} - \mathbf{x}_r) \delta d_i^{\text{sim}} dV dt \\
& - \int_0^T \int_{\Omega} \delta c_{ijkl} \frac{\partial d_k^{\text{sim}}}{\partial x_l} \frac{\partial \lambda_i}{\partial x_j} dV dt - \int_0^T \int_{\Omega} \left[ \rho \frac{\partial^2 \lambda_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial \lambda_k}{\partial x_l} \right) \right] \delta d_i^{\text{sim}} dV dt.
\end{aligned} \tag{A.9}$$

According to Plessix (2006), the condition  $\frac{\partial \Delta}{\partial \boldsymbol{\lambda}} = 0$  leads to the so-called ‘‘state equations.’’ The adjoint state equations are obtained by setting  $\frac{\partial \Delta}{\partial \mathbf{d}^{\text{sim}}} = 0$ . Taking the derivative of equation A.9 leads to the

adjoint state equation,

$$\begin{aligned} \rho \frac{\partial^2 \lambda_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial \lambda_k}{\partial x_l} \right) &= \sum_{r=1}^N \left\{ \left[ d_i^{\text{sim}}(\mathbf{x}_r, t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - d_i^{\text{obs}}(\mathbf{x}_r, t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) \right] d^{\text{obs}}(\mathbf{x}_f, t) \right\} \\ &= \sum_{r=1}^N \left\{ d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) \otimes \left[ d_i^{\text{sim}}(\mathbf{x}_r, t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, t) - d_i^{\text{obs}}(\mathbf{x}_r, t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, t) \right] \right\}, \end{aligned} \quad (\text{A.10})$$

where  $\otimes$  denotes cross-correlation. Because  $\mathbf{d}^{\text{sim}}$  satisfies the wave equation,  $\Delta$  is equal to the objective function  $S_b$  from equation 3.2 (see equation A.4) (Plessix, 2006). Hence, the change  $\delta S_b$  due to the perturbations of the stiffness coefficients is given by:

$$\delta S_b = - \int_0^T \int_{\Omega} \delta c_{ijkl} \frac{\partial d_i^{\text{sim}}}{\partial x_j} \frac{\partial \lambda_k}{\partial x_l} dV dt. \quad (\text{A.11})$$

To simulate the Lagrangian multiplier, we follow Liu and Tromp (2006) in defining the ‘‘adjoint wavefield’’  $\psi$ :

$$\psi(\mathbf{x}, t) \equiv \boldsymbol{\lambda}(\mathbf{x}, T - t). \quad (\text{A.12})$$

The adjoint wavefield  $\psi$  satisfies the following wave equation with a reversed-time source function:

$$\rho \frac{\partial^2 \psi_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left( c_{ijkl} \frac{\partial \psi_k}{\partial x_l} \right) = \quad (\text{A.13})$$

$$\sum_{r=1}^N \left\{ d^{\text{obs}}(\mathbf{x}_f, T - t) \otimes \left[ d_i^{\text{sim}}(\mathbf{x}_r, T - t) * d_{\text{ref}}^{\text{obs}}(\mathbf{x}_f, T - t) - d_i^{\text{obs}}(\mathbf{x}_r, T - t) * d_{\text{ref}}^{\text{sim}}(\mathbf{x}_f, T - t) \right] \right\}. \quad (\text{A.14})$$

The initial conditions for  $\psi$  are

$$\psi(\mathbf{x}, 0) = 0, \quad \frac{\partial \psi(\mathbf{x}, 0)}{\partial t} = 0, \quad (\text{A.15})$$

and the boundary condition is

$$\psi(\mathbf{x}, t)|_{\mathbf{x} \rightarrow \infty} \rightarrow 0. \quad (\text{A.16})$$

From equations A.10 and A.11, we can derive the inversion gradient with respect to the stiffness coefficients (Kamath and Tsvankin, 2016):

$$\frac{\partial S}{\partial c_{ijkl}} = - \int_0^T \frac{\partial d_i^{\text{sim}}}{\partial x_j} \frac{\partial \psi_k}{\partial x_l} dt. \quad (\text{A.17})$$

The gradient of  $S$  with respect to the chosen parameters  $m_n$  can be obtained using the chain rule:

$$\frac{\partial S}{\partial m_n} = \sum_{ijkl} \frac{\partial S}{\partial c_{ijkl}} \frac{\partial c_{ijkl}}{\partial m_n}. \quad (\text{A.18})$$

For VTI media, we define  $m_1 = V_{P0}$ ,  $m_2 = V_{S0}$ ,  $m_3 = V_{\text{nmo,P}}$ ,  $m_4 = V_{\text{hor,P}}$ ,  $m_5 = V_{\text{hor,SH}} = V_{S0}\sqrt{1+2\gamma}$ , and  $m_6 = \rho$  (Singh et al., 2020). The stiffness elements can be expressed using the velocities as follows (e.g., Kamath and Tsvankin, 2016; Tsvankin, 2012):

$$\begin{aligned} C_{11} &= \rho V_{\text{hor,P}}^2, \\ C_{13} &= \rho \left[ \sqrt{(V_{P0}^2 - V_{S0}^2)(V_{\text{nmo,P}}^2 - V_{S0}^2)} - V_{S0}^2 \right], \\ C_{33} &= \rho V_{P0}^2, \\ C_{55} &= \rho V_{S0}^2, \\ C_{66} &= \rho V_{\text{hor,SH}}^2. \end{aligned} \tag{A.19}$$

Note that  $C_{66}$  is not estimated in our synthetic examples because the modeling algorithm simulates only P- and SV-waves, which are polarized in the vertical propagation plane. Equations A.17 and A.18 allow us to obtain the gradient for the model parameters:

$$\begin{aligned} \frac{\partial S}{\partial V_{P0}} &= -2\rho V_{P0} \int_0^T \left[ \frac{\partial \psi_x}{\partial z} \frac{\partial d_z^{\text{sim}}}{\partial z} \right. \\ &\quad \left. + \frac{q}{2} \left( \frac{\partial \psi_x}{\partial z} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_y^{\text{sim}}}{\partial y} \right) \right] dt, \end{aligned} \tag{A.20}$$

$$\begin{aligned} \frac{\partial S}{\partial V_{S0}} &= 2\rho V_{S0} \int_0^T \left[ \left(1 + \frac{q}{2} + \frac{1}{2q}\right) \left( \frac{\partial \psi_x}{\partial x} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_y^{\text{sim}}}{\partial y} \right) \right. \\ &\quad \left. - \left( \frac{\partial \psi_x}{\partial z} + \frac{\partial \psi_z}{\partial x} \right) \left( \frac{\partial d_x^{\text{sim}}}{\partial z} + \frac{\partial d_z^{\text{sim}}}{\partial x} \right) - \left( \frac{\partial \psi_y}{\partial z} + \frac{\partial \psi_z}{\partial y} \right) \left( \frac{\partial d_y^{\text{sim}}}{\partial z} + \frac{\partial d_z^{\text{sim}}}{\partial y} \right) \right] dt, \end{aligned} \tag{A.21}$$

$$\frac{\partial S}{\partial V_{\text{nmo,P}}} = -\frac{\rho V_{\text{nmo,P}}}{q} \int_0^T \left( \frac{\partial \psi_x}{\partial x} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_z^{\text{sim}}}{\partial z} \right) dt, \tag{A.22}$$

$$\frac{\partial S}{\partial V_{\text{hor,P}}} = -2\rho V_{\text{hor,P}} \int_0^T \left( \frac{\partial \psi_y}{\partial y} + \frac{\partial \psi_x}{\partial x} \right) \left( \frac{\partial d_y^{\text{sim}}}{\partial y} + \frac{\partial d_x^{\text{sim}}}{\partial x} \right) dt, \tag{A.23}$$

$$\begin{aligned} \frac{\partial S}{\partial V_{\text{hor,SH}}} &= -2\rho V_{\text{hor,S}} \int_0^T \left[ \left( \frac{\partial \psi_x}{\partial y} + \frac{\partial \psi_y}{\partial x} \right) \left( \frac{\partial d_x^{\text{sim}}}{\partial y} + \frac{\partial d_y^{\text{sim}}}{\partial x} \right) \right. \\ &\quad \left. - 2 \left( \frac{\partial \psi_y}{\partial y} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_x}{\partial x} \frac{\partial d_y^{\text{sim}}}{\partial y} \right) \right] dt, \end{aligned} \tag{A.24}$$

$$\begin{aligned} \frac{\partial S}{\partial \rho} &= - \int_0^T \left\{ V_{P0}^2 \left( \frac{\partial \psi_z}{\partial z} \frac{\partial d_z^{\text{sim}}}{\partial z} \right) + V_{\text{hor,P}}^2 \left( \frac{\partial \psi_x}{\partial x} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_y^{\text{sim}}}{\partial y} + \frac{\partial \psi_x}{\partial x} \frac{\partial d_y^{\text{sim}}}{\partial y} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_x^{\text{sim}}}{\partial x} \right) \right. \\ &\quad \left. + V_{S0}^2 \left[ \left( \frac{\partial \psi_x}{\partial z} + \frac{\partial \psi_z}{\partial x} \right) \left( \frac{\partial d_x^{\text{sim}}}{\partial z} + \frac{\partial d_z^{\text{sim}}}{\partial x} \right) + \left( \frac{\partial \psi_y}{\partial z} + \frac{\partial \psi_z}{\partial y} \right) \left( \frac{\partial d_y^{\text{sim}}}{\partial z} + \frac{\partial d_z^{\text{sim}}}{\partial y} \right) \right] \right. \\ &\quad \left. + V_{\text{hor,SH}}^2 \left[ \left( \frac{\partial \psi_x}{\partial y} + \frac{\partial \psi_y}{\partial x} \right) \left( \frac{\partial d_x^{\text{sim}}}{\partial y} + \frac{\partial d_y^{\text{sim}}}{\partial x} \right) - 2 \frac{\partial \psi_y}{\partial y} \frac{\partial d_x^{\text{sim}}}{\partial x} - 2 \frac{\partial \psi_x}{\partial x} \frac{\partial d_y^{\text{sim}}}{\partial y} \right] \right. \\ &\quad \left. + \left[ \sqrt{(V_{\text{nmo,P}}^2 - V_{S0}^2)(V_{P0}^2 - V_{S0}^2)} - V_{S0}^2 \right] \left[ \frac{\partial \psi_z}{\partial z} \frac{\partial d_x^{\text{sim}}}{\partial x} + \frac{\partial \psi_x}{\partial x} \frac{\partial d_z^{\text{sim}}}{\partial z} + \frac{\partial \psi_z}{\partial z} \frac{\partial d_y^{\text{sim}}}{\partial y} + \frac{\partial \psi_y}{\partial y} \frac{\partial d_z^{\text{sim}}}{\partial z} \right] \right. \\ &\quad \left. + v_x \psi_x + v_y \psi_y + v_z \psi_z \right\} dt, \end{aligned} \tag{A.25}$$

where

$$q = \sqrt{\frac{V_{\text{nm},P}^2 - V_{S_0}^2}{V_{P_0}^2 - V_{S_0}^2}}. \quad (\text{A.26})$$

Here,  $\mathbf{v}$  denotes the forward-propagated velocity field and  $\boldsymbol{\psi}$  denotes the back-propagated velocity field (Singh et al., 2020).

## APPENDIX B

### WAVE PROPAGATION IN 2D TILTED TI MEDIA

We consider a TTI medium with the symmetry axis confined to the  $[x, z]$ -plane of a Cartesian coordinate system. The  $[x, z]$ -plane is also assumed to coincide with the dip plane of the subsurface structures to make wave propagation two-dimensional. The wave equation for P- and SV-waves in such 2D TTI media can be written as:

$$\rho \frac{\partial v_x}{\partial t} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} + f_x, \quad (\text{B.1})$$

$$\rho \frac{\partial v_z}{\partial t} = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} + f_z, \quad (\text{B.2})$$

where  $\mathbf{v}$  is the particle velocity, and  $\boldsymbol{\sigma}$  is the stress tensor. The Hooke' law for this model takes the form:

$$\frac{\partial \sigma_{xx}}{\partial t} = d_{11} \frac{\partial v_x}{\partial x} + d_{13} \frac{\partial v_z}{\partial z} + d_{15} \left( \frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right), \quad (\text{B.3})$$

$$\frac{\partial \sigma_{zz}}{\partial t} = d_{13} \frac{\partial v_x}{\partial x} + d_{33} \frac{\partial v_z}{\partial z} + d_{35} \left( \frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right), \quad (\text{B.4})$$

$$\frac{\partial \sigma_{xz}}{\partial t} = d_{15} \frac{\partial v_x}{\partial x} + d_{35} \frac{\partial v_z}{\partial z} + d_{55} \left( \frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right), \quad (\text{B.5})$$

where  $d_{11}$ ,  $d_{13}$ ,  $d_{15}$ ,  $d_{33}$ ,  $d_{35}$ , and  $d_{55}$  denote the stiffness coefficients given by:

$$d_{11} = (c_{11} \cos^2 \theta + c_{11} \sin^2 \theta) \cos^2 \theta + (c_{13} \cos^2 \theta + c_{33} \sin^2 \theta) \sin^2 \theta + 4c_{44} \cos^2 \theta \sin^2 \theta, \quad (\text{B.6})$$

$$d_{13} = c_{11} \cos^2 \theta \sin^2 \theta + c_{13} \sin^4 \theta + c_{13} \cos^4 \theta + c_{33} \cos^2 \theta \sin^2 \theta - 4c_{44} \cos^2 \theta \sin^2 \theta, \quad (\text{B.7})$$

$$d_{15} = (c_{13} - c_{11}) \cos^3 \theta \sin \theta + (c_{33} - c_{13}) \cos \theta \sin^3 \theta + 2c_{44} \cos \theta \sin \theta (\cos^2 \theta - \sin^2 \theta), \quad (\text{B.8})$$

$$d_{33} = (c_{11} \sin^2 \theta + c_{13} \cos^2 \theta) \sin^2 \theta + (c_{33} \cos^2 \theta + c_{13} \sin^2 \theta) \cos^2 \theta + 4c_{44} \cos^2 \theta \sin^2 \theta, \quad (\text{B.9})$$

$$d_{35} = (c_{13} - c_{11}) \cos \theta \sin^3 \theta + (c_{33} - c_{13}) \cos^3 \theta \sin \theta - 2c_{44} \cos \theta \sin \theta (\cos^2 \theta - \sin^2 \theta), \quad (\text{B.10})$$

$$d_{55} = c_{44} (1 - 2 \sin^2 \theta)^2 + (c_{33} - c_{13}) \cos^2 \theta \sin^2 \theta - (c_{13} - c_{11}) \cos^2 \theta \sin^2 \theta, \quad (\text{B.11})$$

where  $\theta$  is the angle between the symmetry axis and the vertical, and  $c_{ij}$  is the stiffness tensor of the corresponding VTI medium. Equations B.6 to B.11 are obtained using the Bond (1943) transformation.

When  $\theta = 90^\circ$ , the medium becomes HTI (TI with a horizontal symmetry axis). Zero tilt ( $\theta = 0^\circ$ ) corresponds to the VTI model, for which equations B.1 to B.5 reduce to:

$$\rho \frac{\partial v_x}{\partial t} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} + f_x, \quad (\text{B.12})$$

$$\rho \frac{\partial v_z}{\partial t} = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} + f_z, \quad (\text{B.13})$$

$$\frac{\partial \sigma_{xx}}{\partial t} = c_{11} \frac{\partial v_x}{\partial x} + c_{13} \frac{\partial v_z}{\partial z}, \quad (\text{B.14})$$

$$\frac{\partial \sigma_{zz}}{\partial t} = c_{13} \frac{\partial v_x}{\partial x} + c_{33} \frac{\partial v_z}{\partial z}, \quad (\text{B.15})$$

$$\frac{\partial \sigma_{xz}}{\partial t} = c_{44} \left( \frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right). \quad (\text{B.16})$$

APPENDIX C  
PERMISSIONS FOR CHAPTERS

The permission to include work from Journal of Seismic Exploration shown in Figure C.1. The permission to include work from Geophysicalsis shown in Figure C.2. The URL for permissions in Geophysics is <https://seg.org/Publications/Policies-and-Permissions/Open-Access-Policy>.

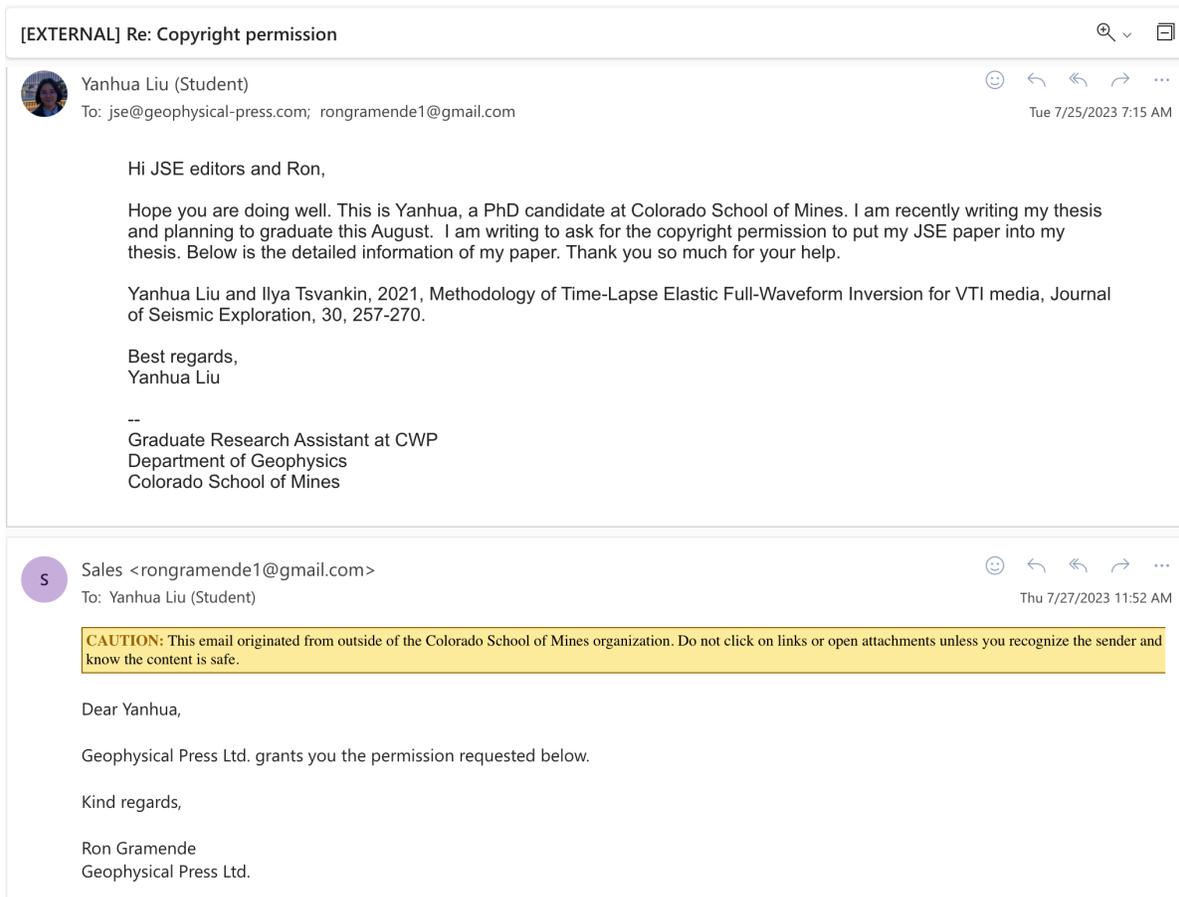


Figure C.1 Permission from *Journal of Seismic Exploration*.

## SEG Policy on Open-access Publishing

SEG provides authors a variety of open-access publishing options, including some "green" and "gold" open-access options.

SEG provides the world's premier vehicles for knowledge exchange in applied geophysics. The Society's publishing program is one such vehicle. To help attract the best scholarship and disseminate it to the widest possible audience through journals, meetings papers, and books, SEG provides authors a variety of open-access publishing options. These include some "green" open-access options that long have been features of the traditional subscription-based business model that has sustained the Society's publishing program.

In 2011, SEG went further by instituting a "gold" open-access option for its Journals, transforming them into "hybrid" open-access journals. In exchange for payment of an author publication charge (APC) of US\$2,500 plus payment of all applicable voluntary and mandatory page and color charges, SEG agreed to remove access barriers to full-text presentations of the author's paper on SEG publications sites. Standard copyright transfer was required.

Beginning with papers published in 2014, this gold open-access option is expanded with additional copyright and licensing options, and the policy also is extended to meetings papers and book chapters. These changes make all SEG publications compliant with requirements in place and emerging from a variety of research funders and their overseers, such as Research Councils UK (RCUK), the European Commission, and the U.S. Office of Science and Technology Policy. Also, some aspects of SEG's green open-access policy are refined.

In October 2015, the SEG Board of Directors approved a policy change removing the requirement of authors to pay all voluntary charges for open-access papers for papers submitted on or after 15 October 2015. Authors must pay mandatory charges and the \$2,500 open-access fee.

Here are features of open-access options available to authors, or in cases of works made for hire, their employers:

### Traditional publication (including green open access)

- No author publication charge (APC) is levied, although payment of mandatory page charges are assessed and payment of voluntary charges is requested. Relief from mandatory charges may be requested under SEG's hardship relief policy.
- Copyright is transferred to SEG.
- Authors/employers retain proprietary rights such as the right to patentable subject matter and the right to make oral presentation of the work with full citation and proper copyright acknowledgment.
- Authors/employers enjoy the right to prepare and hold copyright in derivative publications based on the paper provided that the derivative work is published subsequent to the official date of the original paper's publication by SEG.
- Authors/employers may post a final accepted version of the manuscript or the final SEG-formatted version (book chapters excluded) on authors' personal websites, employers' websites, or in institutional repositories operated and controlled exclusively by authors' employers provided that:
  1. the SEG-prepared version is presented without modification;
  2. copyright notice and a full citation appear with the paper;
  3. a link to the SEG version of record in the SEG Digital Library using Digital Object Identifier (DOI) permalinks is provided;
  4. the posting is noncommercial in nature, and the paper is made available to users without charge; and
  5. that notice be provided that use is subject to SEG terms of use and conditions.
- Authors/employers may not post their articles in an institutional repository or other site in which the content is required to carry or is implied as carrying a license contrary to SEG copyright and terms of use and terms of this policy.
- Authors/employers may post to their own websites a preprint (a version prior to SEG peer review) of a submitted manuscript provided that the posting is accompanied by prominent notice that the paper is under review for publication by SEG and the publication to which the manuscript has been submitted is identified. Upon publication of the paper by SEG, the author must replace any previously posted electronic versions of the paper with either (a) a full citation to the SEG-published work with a DOI permalink to the paper's abstract or (b) the final version of the manuscript or SEG-formatted version (book chapters excluded) subject to SEG conditions for such posting.
- **Authors may reuse all or part of their papers published with SEG in a thesis or dissertation that authors write and are required to submit to satisfy criteria of degree-granting institutions.**
- Authors/employers have the nonexclusive right, after publication by SEG, to give permission to third parties to republish print versions of the paper, or excerpts therefrom, without obtaining permission from SEG, provided that:
  1. the SEG-prepared version is not used for this purpose;
  2. the paper is not republished in another journal or book; and
  3. the third party does not charge a fee. Permission must be obtained from SEG for other republication of the paper.

Figure C.2 Permission from *Geophysics* highlighted in yellow.