RICE UNIVERSITY

Frequency-Dependent Traveltime Tomography and Full Waveform Inversion for Near-Surface Seismic Refraction Data

by

Jianxiong Chen

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Colin A. Zelt, Chair
Professor, Earth Science

Brandon Dugan
Associate Professor, Earth Science

Fenglin Niu
Professor, Earth Science

William W. Symes
Noah Harding Professor, Computational and Applied Mathematics

HOUSTON, TEXAS
February 2016
ABSTRACT

Frequency-dependent traveltime tomography and full waveform inversion for near-surface seismic refraction data

by

Jianxiong Chen

I demonstrate the utility and benefits of a combined use of frequency-dependent traveltime tomography (FDTT) and full waveform inversion (FWI) to estimate the near-surface seismic velocity that contains wavelength- and sub-wavelength-scale features. FDTT is fundamentally different from conventional ray-theory infinite-frequency traveltime tomography (IFTT) methods in the calculation of a frequency-dependent traveltime using wavelength-dependent velocity smoothing (WDVS). I justify the use of WDVS in FDTT for calculating a frequency-dependent traveltime by comparing the calculated traveltimes with that picked from acoustic synthetics, showing that compared to the conventional infinite-frequency traveltimes calculated based on ray-theory, the frequency-dependent traveltimes calculated using WDVS can better match the picked traveltimes.

In the combined workflow of FDTT and FWI, FDTT provides a long-wavelength background seismic velocity model as the starting model, and then FWI introduces wavelength- and sub-wavelength-scale features that allow for direct geologic interpretation of the velocity models. I apply this workflow to seismic data generated by a near-surface realistic synthetic velocity model representing a geologic setting consisting of unconsolidated sediment overlying faulted bedrock, successfully imaging the key
model features, a thin low-velocity layer in the sediments, a steep bedrock offset and a steeply dipping low-velocity fault zone. I then apply this workflow to 2D P- and SH-waves collected in 2011 at Rice campus with a known target of a buried concrete tunnel. The P- and SH-wave models image the top part of the tunnel as a high-velocity anomaly. The P-wave models also image the air in the void space of the tunnel as a low-velocity anomaly.

As a comparison, conventional IFTT is also applied in a combined workflow with FWI. The comparisons of the inverted models show that both IFTT and FDTT models can serve as adequate starting models for FWI, but FDTT is favored over IFTT because: 1) The FDTT models better recover the magnitude of the velocity anomalies, and 2) The FDTT model serves as a better starting model for FWI, which results in a more accurate FWI velocity estimation with better recovery of the magnitude and location of the key features, particularly in the absence of usable low frequency data.
Acknowledgments

I would like to sincerely thank my advisor Dr. Colin Zelt for his support through the years. Without his advice, expertise, and encouragement, this research and dissertation would not have happened. I would also like to thank my committee members, Dr. Fenglin Niu, Dr. Brandon Dugan, and Dr. William Symes for sharing their thoughts and feedbacks during my research progresses.

I would like to thank Rice Earth Science alumni Dr. Priyank Jaiswal, Dr. Fuchun Gao, and Geoff Chambers for sharing their insights and expertise in full wave inversion and their experiences in using the waveform tomography code, that is kindly provided and made available by Dr. Gerhard Pratt and his group to the community.

I would like to thank the graduate students who voluntarily worked on the 2D Rice seismic experiments and the used equipment was kindly provided by PASSCAL.

I would like to thank all the department staff members for their management support. Clinton Heider and Mary Cochran saved me a lot of time by solving the computing system problems timely.

Finally I would like to thank my wife Zhiwei Qian and entirely family members for their enduring love that supports me through all the adventures.
# Table of Contents

Acknowledgments ............................................................................................................................... iv

Table of Contents ............................................................................................................................... v

List of Figures ........................................................................................................................................ vii

Introduction ........................................................................................................................................... 1

Comparison of Full Wavefield Synthetics with frequency-dependent traveltimes calculated using wavelength-dependent velocity smoothing ................................................................. 6

2.1. Abstract ......................................................................................................................................... 6

2.2. Introduction ................................................................................................................................ 7

2.3. Wavelength-Dependent Velocity Smoothing .............................................................................. 9

2.4. Experiment Setup ......................................................................................................................... 13

2.5. Results ......................................................................................................................................... 18

2.6. Discussion and Conclusions ....................................................................................................... 24

2.7. Acknowledgements ...................................................................................................................... 25

Application of frequency-dependent traveltime tomography and full waveform inversion to realistic near-surface seismic refraction data ................................................................. 26

3.1. Abstract ......................................................................................................................................... 26

3.2. Introduction ................................................................................................................................ 27

3.3. Model and Data ............................................................................................................................ 32

3.4. Methods ...................................................................................................................................... 34

3.4.1. Frequency-Dependent Traveltime Tomography ................................................................. 35

3.4.2. Frequency Domain Full Waveform Inversion ....................................................................... 37

3.5. Results ......................................................................................................................................... 41

3.6. Discussion ................................................................................................................................... 47

3.7. Conclusions ................................................................................................................................ 51

3.8. Acknowledgements ...................................................................................................................... 52

Detecting a known near-surface target through application of frequency-dependent traveltime tomography and full waveform inversion to P- and SH-wave seismic refraction data ......................................................................................................................... 53

4.1. Abstract ......................................................................................................................................... 53
4.2. Introduction .................................................................54
4.3. Site and Experiment .......................................................57
4.4. Methods ...........................................................................63
  4.4.1. Frequency-Dependent Traveltime Tomography .................63
  4.4.2. Frequency Domain Full Waveform Inversion .....................66
4.5. Synthetic P-Wave Results ....................................................72
4.6. Real P-Wave Results ........................................................76
4.7. Real Sh-Wave Results .......................................................81
4.8. Discussion .......................................................................85
4.9. Conclusion .......................................................................95
4.10. Acknowledgements ........................................................95

Conclusions ...........................................................................97

References .............................................................................100
List of Figures

Figure 2.1. Cartoon showing different behavior of ray paths and wave paths (from Zelt and Chen, 2016) ........................................................................................................ 10

Figure 2.2. Examples of velocity smoothing weights. ......................................................... 12

Figure 2.3. Velocity models .................................................................................................. 16

Figure 2.4. Source spectra. .................................................................................................. 16

Figure 2.5. Comparisons traveltimes with picked first arrivals ........................................... 18

Figure 2.6. Comparisons traveltimes with picked first arrivals ........................................... 19

Figure 2.7. Comparisons traveltimes with picked first arrivals ........................................... 20

Figure 2.8. Comparisons traveltimes with the noisy acoustic synthetics ......................... 22

Figure 3.1. True model......................................................................................................... 30

Figure 3.2. Representative synthetic acoustic seismograms. .............................................. 33

Figure 3.3. True and estimated models from this study and from SAGEEP 2011 blind test (Zelt et al., 2013) ........................................................................................................ 41

Figure 3.4. Comparison of a horizontal slice at Z=25 m ...................................................... 42

Figure 3.5. Full waveform inversion (FWI) estimated models ............................................ 43

Figure 3.6. Comparison of horizontal slices through models .............................................. 45

Figure 3.7. Representative synthetic acoustic seismograms .............................................. 46

Figure 4.1. A schematic view of the 2-D survey .................................................................. 58

Figure 4.2. Two component seismic receiver .................................................................... 59

Figure 4.3. P-wave data examples ..................................................................................... 61

Figure 4.4. SH-wave data examples .................................................................................. 62
Figure 4.5. Shot-receiver pairs versus reciprocal first-arrival time difference. ........................................................................................................................................66

Figure 4.6. P-wave traveltime tomography models from synthetic data........72

Figure 4.7. P-wave full waveform inversion (FWI) models from synthetic data. ........................................................................................................................................74

Figure 4.8. P-wave traveltime tomography models from field data.. ............76

Figure 4.9. Wavepaths for 250 Hz for P-wave FDTT model. .........................77

Figure 4.10. P-wave full waveform inversion (FWI) models from field data....78

Figure 4.11. One-dimensional velocity profiles.............................................80

Figure 4.12. SH-wave traveltime tomography models from field data..........81

Figure 4.13. SH-wave full waveform inversion (FWI) models from field data.83

Figure 4.14. Wavepaths for 250Hz for SH-wave FDTT model......................84

Figure 4.15. P- and SH-wave source signatures...........................................89

Figure 4.16 P-wave data examples.................................................................90

Figure 4.17 SH-wave data examples............................................................91

Figure 4.18. P-wave frequency-domain data...............................................92

Figure 4.19. SH-wave frequency-domain data.............................................93
Chapter 1

Introduction

Near-surface geophysics is concerned with the use of quantitative methods to study the physical properties and particularly small-scale features in the shallow subsurface (<100-m depth). These studies address a number of near-surface environmental, engineering, archeological, neotectonic and resource exploration problems. Commonly used near-surface geophysical methods include but are not limited to seismic refraction, seismic reflection, ground-penetrating radar, electrical and electromagnetic techniques, and well logging. These methods and their results are separately or jointly used and interpreted according to the geological situation, required resolution, and available budget (Steeples, 2001; Pelton, 2005; Reynolds, 2011).

In the 1970’s, seismologists adapted the tomography methodology from the medical community to different scale of geophysical studies (e.g., Dziewonski et al., 1977). Since then and until recently, ray-theory, an infinite-frequency
approximation of wave propagation has been used to model the traveltimes, which serves as the forward modeling component of traveltime tomography because of its computational efficiency. Compared with global-scale or crustal-scale studies, where the size of anomalies is typically much larger than that of the seismic wavelength, in the near surface (<100-m depth), the size of anomalies is often comparable with or smaller than the seismic wavelength (e.g., Gao et al., 2007), theoretically invalidating the infinite-frequency assumption of ray theory.

In this thesis, I introduce a traveltime tomography method that takes frequency content in the seismic waves into consideration. This frequency-dependent traveltime tomography (FDTT) (Zelt and Chen, 2016) is fundamentally different from conventional ray-theory-based infinite-frequency traveltime tomography (IFTT) methods in the calculation of a frequency-dependent traveltime using wavelength-dependent velocity smoothing (WDVS). In this thesis, I demonstrate the ability and benefits of the use of FDTT in a combined workflow with full waveform inversion (FWI) to estimate the subsurface seismic velocity, where FDTT provides a long-wavelength background seismic velocity model and then FWI introduces wavelength- and sub-wavelength-scale features that allow for direct geologic interpretation.

For a successful FWI application, the accuracy of the starting model is crucial and has been one of the challenges due to the extreme non-linearity of inverting the waveform data. The workflow of using a lower resolution but more robust method like ray-theory-based infinite-frequency traveltime tomography (IFTT) methods to
provide a starting model for FWI is well documented in the publications (e.g., Pratt et al.; Gao et al. 2006, 2007; 2002, Gao, Priyank et al., 2008, 2009) as a common practice. One of this thesis’ contributions is to demonstrate that FDTT can serve as a better starting model for FWI than conventional ray-theory-based IFTT methods.

Zelt and Chen (2016) thoroughly present the FDTT methodology and with both synthetic and field data examples demonstrate its ability of achieving better velocity estimation over IFTT in terms of a more accurate magnitude recovery and higher spatial resolution. One of FDTT’s key modifications to conventional IFTT is forward modeling a frequency-dependent traveltime using wavelength-dependent velocity smoothing (WDVS). WDVS is based on the wavelength-smoothing technique developed by Lomax (1994) within the context of a ray-tracing algorithm, but it is implemented in Zelt and Chen (2016) within the wavefront-tracking algorithm of Vidale (1988). WDVS does not have a theoretical framework mathematically deduced from wave equation, but it is designed to model the behaviors that are consistent with finite-frequency wave propagation. Because of the key role that WDVS plays in FDTT for calculating a frequency-dependent traveltime, I dedicate a whole chapter, Chapter 2, in this thesis to justify the use of WDVS for modeling a frequency-dependent traveltime through examples showing its frequency-dependent behavior that are consistent with finite-frequency wave propagation, and by comparing the frequency-dependent traveltimes with that from synthetic seismographs. Chapter 2 is based on a submitted manuscript to Journal of Environmental and Engineering Geophysics (JEEG) (Chen and Zelt, 2016).
Chapter 3 presents the results from the workflow of FDTT and FWI applied to the data generated by a benchmark synthetic velocity model consisting of unconsolidated sediment overlying faulted bedrock, representing a realistic near-surface geological setting. Chapter 3 is based on a paper in press by the JEEG (Chen and Zelt, 2016), and it serves as a follow-up of an earlier paper (Zelt et al., 2013) in JEEG, when and where the used realistic synthetic velocity model in Chapter 3 was first published. Compared to only characterizing the large-scale (>wavelength) structures as shown by ten participants in a blind test using eight different inversion algorithms to invert traveltimes (Zelt et al., 2013), the combination of FDTT and FWI in my thesis demonstrates its ability to characterize both the large-scale and wavelength-scale structures in near-surface studies. The FWI inverted models in Chapter 3 reveal wavelength-scale features that allow for direct interpretations of the key features in the true model: a low-velocity zone, a bedrock offset, and a dipping fault. Also, Chapter 3 fully demonstrates the benefits of using FDTT over IFTT through comparisons. I use both FDTT and IFTT to separately provide the starting model for FWI, and test the influences of starting FWI frequency as a variable for either case. The comparisons show that FDTT is preferable to IFTT in two aspects: 1) The FDTT model better recovers the magnitude of the velocity anomalies. 2) The FDTT model serves as a better starting model for FWI, which results in a more accurate FWI velocity estimation, particularly in the absence of usable low frequency data.

Chapter 4 presents the FDTT and FWI results from the 2D P- and SH-wave data collected at Rice targeting at a known concrete tunnel filled with air. The P- and
SH-wave models image the top part of the tunnel at the correct location at a depth of 1.6 m as a high-velocity anomaly. The P-wave models also image the air in the void space of the tunnel as a low-velocity anomaly. The dimensions of these recovered tunnel features are at the wavelength and sub-wavelength scales. Similarly with the benchmark test in Chapter 3, IFFT is applied for comparisons and the observations from the inverted velocity models favor the use of FDTT over IFFT. Chapter 4 is based on a submitted manuscript to Geophysics (Chen et al., 2016).

In summary, the near-surface seismic studies are unique and challenging in that the seismic wavelength is comparable or smaller than the scale of interested anomalies, contrary to the ray-theory as is commonly assumed. I apply a combined workflow of FDTT and FWI. FDTT inverts first-arrival times by calculating frequency-dependent traveltimes. FWI inverts the waveforms of early arrivals by solving the acoustic wave or SH-wave equation. The presented models reveal wavelength- and sub-wavelength-scale structures that allow for direct geologic interpretation from the velocity itself. There is no obstacle preventing the use of FDTT in crustal- and global-scale of seismic studies, but the frequency-dependent effects are most significant in the near surface.
Chapter 2

Comparison of Full Wavefield Synthetics with frequency-dependent traveltimes calculated using wavelength-dependent velocity smoothing

2.1. Abstract

Ray-theory-based traveltime calculation that assumes infinitely-high-frequency wave propagation is likely to be invalid in the near-surface (upper tens of meters) due to the relatively large seismic wavelength compared with the total travel path lengths and the scale of the near-surface velocity heterogeneities. The wavelength-dependent velocity smoothing (WDVS) algorithm calculates a frequency-dependent, first-arrival traveltime by assuming that using a wavelength-smoothed velocity model and conventional ray theory is equivalent to using the original unsmoothed model and a frequency-dependent
calculation. This chapter presents comparisons of WDVS-calculated traveltimes with band-limited full wavefield synthetics including the results from 1) different velocity models, 2) different frequency spectra, 3) different values of a free parameter in the WDVS algorithm, and 4) different levels of added noise to the synthetics. The results show that WDVS calculates frequency-dependent traveltimes that are generally consistent with the first arrivals from band-limited full wavefield synthetics. Compared to infinite-frequency traveltimes calculated using conventional ray theory, the WDVS frequency-dependent traveltimes are more consistent with the first arrivals picked from full wavefield synthetics in terms of absolute time and trace-to-trace variation. The results support the use of WDVS as the forward modeling component of a tomographic inversion method, or any seismic method that involves modeling first-arrival traveltimes.

2.2. Introduction

Ray-theory-based travelt ime calculation (e.g., Vidale, 1988) and ray tracing (e.g., Julian and Gubbins, 1977; Um and Thurber, 1987) are efficient methods to approximately simulate seismic wave propagation assuming high frequencies (equivalently small wavelengths), that can serve as the basis for seismic tomography (e.g., Zelt and Barton, 1998) and seismic migration (e.g., Gray and May, 1994). Compared to global- and crustal-scale studies, the high-frequency assumption is likely to be less valid, and the resulting effects more significant, in near-surface seismic studies (upper tens of meters) due to the relatively large seismic wavelength compared with the total travel path lengths and the scale of near-surface velocity heterogeneities (e.g., Gao et al., 2007). This motivates the desire to model the finite-frequency behavior of wave propagation within
the cost-effective framework of traveltime calculation or ray tracing, without solving the wave equation.

Zelt and Chen (2016) present a frequency-dependent traveltime tomography method to invert picked first arrivals for seismic velocity estimation whose key feature is calculating frequency-dependent traveltimes using wavelength-dependent velocity smoothing (WDVS) for forward modeling. WDVS is implemented in a wavefront-tracking algorithm (Vidale, 1988, 1990) but is based on the wavelength-smoothing technique applied by Lomax (1994) in the context of ray tracing. Both WDVS and the wavelength-smoothing technique are designed to simulate the frequency-dependent behavior of finite-frequency wave propagation, although they are not formally derived from wave or ray theory. Zelt and Chen (2016) focus on presenting the calculation method of frequency-dependent traveltimes and their use in tomographic inversion for improved velocity estimation. This chapter focuses on justifying the use of WDVS to effectively model frequency-dependent traveltimes through comparisons with band-limited full wavefield synthetics.

Similar comparisons have been done by other researchers to justify their frequency-dependent traveltimes either produced from a traveltime calculation algorithm (e.g., Biondi, 1997; Hogan and Margrave, 2007) or resulting from a ray tracing algorithm (e.g., Washbourne et. al., 2008; Protosov et al., 2011; Yarman et al., 2013). In these studies the calculated traveltimes are compared with full wavefield synthetics by either displaying the calculated traveltimes on wavefield snapshots for a visual comparison without explicitly picking the traveltimes from the wavefield snapshots (Biondi, 1997; Hogan and Margrave, 2007; Yarman et al., 2013), or by displaying the calculated traveltimes on
seismograms for a visual comparison with (Washbourne et. al., 2008) or without (Protosov et al., 2011) explicitly picking the first arrivals from the seismograms. In this chapter the calculated traveltimes are compared with both the synthetic seismograms and the picked first arrivals from the synthetics (representing the “true” traveltimes).

The purpose of this chapter is to validate the use of WDVS as the forward modeling component of traveltime tomography or for other seismic methods that involve modeling first-arrival traveltimes. Comparisons of the WDVS traveltimes with the synthetics include results from 1) different velocity models, 2) different frequency spectra, 3) different values of a free parameter in the WDVS algorithm, and 4) different levels of added noise to the synthetics.

2.3. **Wavelength-Dependent Velocity Smoothing**

This section summarizes the key steps of implementing WDVS in a wavefront-tracking algorithm to calculate a frequency-dependent traveltime; the reader is referred to Zelt and Chen (2016) for a complete description of the background, motivation, and implementation of the WDVS algorithm including examples illustrating the frequency-dependent behavior of the traveltimes and travel paths.
Figure 2.1. Cartoon showing different behavior of ray paths and wave paths (from Zelt and Chen, 2016). (a) A ray path is unperturbed but a wave path is deflected by a wavelength-scale velocity anomaly that is within one wavelength of the path. (b) A ray path is strongly perturbed but a wave path is not affected by a small velocity anomaly on the path. The wavelength-dependent velocity smoothing (WDVS) algorithm assumes the extent of the wave path sensitivity to velocity heterogeneities is determined by the local seismic wavelength.

WDVS is based on the wavelength-smoothing technique that models finite-frequency wave propagation within the context of ray tracing (Lomax, 1994). Figure 1 illustrates the expected behavior of WDVS in the context of ray/wave paths. A wave path represents the travel path of finite-frequency wave propagation, i.e., a frequency-dependent ray path. WDVS and the wavelength-smoothing technique of Lomax (1994) assume that a wave path will be more influenced by heterogeneities closer to the path than those farther away from the path, and the extent of sensitivity should be about one wavelength of the seismic wave, while a ray path representing infinitely high frequency wave propagation is only sensitive to heterogeneities on the path. As a result, the ray and wave paths can behave significantly different when the size of the velocity anomaly is
comparable (Fig. 1(a)) or smaller (Fig. 1(b)) than the wavelength, which are likely the cases in the near-surface.

WDVS is implemented in two steps: 1) Smoothing the velocity model according to the local seismic wavelength using a cosine-squared weighting function with a geometrical correction factor specific to 2-D and 3-D models. 2) Using the smoothed model in a conventional ray-theory-based wavefront-tracking algorithm (eikonal solver) under the assumption that the traveltime in the smoothed model is equivalent to the traveltime that would be calculated using a frequency-dependent simulation in the original unsmoothed model. Zelt and Chen (2016) use the 2-D (Vidale, 1988) and 3-D (Vidale, 1990) wavefront-tracking algorithms, with modifications to account for large velocity gradients (Hole and Zelt, 1995), to solve the eikonal equation for traveltimes on a fine square grid using finite-difference operators with the velocity model specified at each node of the grid.
Figure 2. Examples of velocity smoothing weights applied by WDVS at 200 Hz for a single node (white circle) in 2-D and 3-D versions of a linear velocity gradient (modified after Zelt and Chen, 2016). For the 3-D model, weights are only shown in the vertical plane in which the central node (white circle) is located. The central node is located where it is expected to have an average velocity of 1000 m/s (horizontal dashed line in the velocity profile on the left). The calculated $V_{ave}$ based on the smoothing weights and the unsmoothed model is indicated. A uniform grid with a 0.05 m node spacing is used; gray dots indicate every 10th node in the grid in the horizontal and vertical directions. Contour interval is 0.1. The slight difference between the weights is due to a geometrical correction term described in the text.

Figure 2 shows examples of the velocity smoothing weights applied by WDVS for a single node in 2-D and 3-D versions of a linear velocity gradient. The frequency (200 Hz) and the velocity gradient are realistic for the near-surface. All nodes within one period of the central node, extending outward in all directions in 2-D or 3-D from it have a non-zero weight. At each node, the weight is a product of the cosine-squared weighting function and a geometrical correction term. The cosine-squared weighting function has a
maximum value of one at the central node, decreasing to zero at one period away from the center. The geometrical correction term is $1/v^p$, where $v$ is the velocity of the node, and $p=2$ for 2-D models, and $p=3$ for 3-D models; see Zelt and Chen (2016) for details. This geometrical correction term accounts for the fact that one wavelength from the central node is longer and includes more velocity nodes in the higher velocity regions than in the lower velocity regions. Without this correction the higher velocity nodes will be over-represented. In both the 2-D and 3-D cases in Fig. 2, the smoothed velocity for the central node should be 1000 m/s. The 3-D case has a 0.6% error due to the imprecision of the geometrical term when using a discretized model.

As Zelt and Chen (2016) point out, the Vidale algorithms calculate traveltimes corresponding to a frequency whose wavelength is on the order of the model node spacing, and normally the grid is fine enough so that these traveltimes are effectively infinitely high frequency traveltimes. In the rest of this chapter, the traveltimes calculated using an unsmoothed model are referred to as infinite-frequency traveltimes.

2.4. Experiment Setup

A time-domain finite-difference acoustic wave simulator (Symes et al., 2011) generates the synthetics using a plane wave composed of two-excursion keuper wavelets (Brenders and Pratt, 2007). The plane wave travels vertically through a 2-D velocity model containing randomly distributed velocity anomalies, and is sampled by a horizontal array of 50 receivers, evenly-spaced, 45 meters away (Fig. 3). Wavelets with dominant frequencies of 50, 100, and 200 Hz (Fig. 4) are tested. Three random velocity models (Fig. 3) are used by adding random velocity anomalies to a constant background
velocity model (1000 m/s) and applying a low-pass filter. For both the acoustic simulator and the wavefront-tracking eikonal solver, the velocity models are parameterized using a square grid with a node spacing of 0.1 m. Adding random noise that has the same bandwidth as the noise-free synthetics produced the noisy synthetics. The root-mean-square (RMS) amplitude of the added noise is proportional to the root-mean-square amplitude of the early arrivals of the noise-free synthetics. The source spectra, the magnitude of the velocity anomalies, the relative size of the velocity anomalies compared to the seismic wavelengths, and the amplitudes of added noise to the synthetics are realistic for near-surface studies.
Figure 2.3. Velocity models with randomly distributed velocity perturbations up to \(\pm 50\%\) from the constant background model (1000 m/s). The three models were produced using different seed numbers for initializing a random number generator. (a) Seed-300 model has an RMS perturbation of 113 m/s from the constant background model. (b) Seed-200 model has an RMS perturbation of 120 m/s. (c) Seed-120 model has an RMS perturbation of 131 m/s. A horizontal plane wave source is introduced into the models at \(z=5\) m for simulating the acoustic wavefield synthetics. An array of 50 receivers with an even 2-m spacing is positioned at \(z=50\) m. The green lines at the bottom of each model represent the 20-, 8-, 10-, 4-, 5-, and 2-m wavelengths, corresponding to 50-, 125-, 100-, 250-, 200-, and 500-Hz waves in the background velocity.

Zelt and Chen (2016) suggest two ways to pick the first arrivals of real data depending on the signal-to-noise ratio. When the data have a high signal-to-noise ratio, the onsets are picked; when the data have a low-signal-to-noise ratio, the picks are based on a trace-to-trace correlation of the peaks/troughs in the first-arrival waveform and made by a certain time advance from the peaks/troughs toward the onsets that would be

Figure 2.4. Source spectra of a two-excursion keuper wavelet (Brenders and Pratt, 2007) with dominant frequencies of 50, 100 and 200 Hz are used in the acoustic wavefield modeling; for modeling frequency-dependent traveltimes using WDVS, high-end frequencies of 125, 250 and 500 Hz, respectively, are used.
predicted without noise. In this chapter, to avoid biases from manual picks, in the comparisons of the WDVS traveltimes with the noise-free synthetics, the onsets of the synthetics are picked based on an amplitude threshold that is 1% of the RMS amplitude of the early arrivals; in the comparisons of the WDVS traveltimes with the noisy synthetics, the first arrivals are not explicitly picked. The extent of the early arrivals is defined by a time window based on the dominant-frequency of the wavelet that is used to generate the synthetics: 40, 20 and 10 ms for the 50-, 100- and 200-Hz-wavelet synthetics, respectively, each starting from the calculated infinite-frequency traveltimes.

Both the infinite-frequency and frequency-dependent traveltimes are calculated for the comparisons. For using WDVS to model the first arrivals picked from band-limited real data, Zelt and Chen (2016) suggest that the selected modeling frequency should be based on the data spectra and signal-to-noise ratio: the highest frequency in the spectra is selected to model the first arrivals picked from high signal-to-noise ratio data; the dominant frequency is selected to model the first arrivals picked from low signal-to-noise ratio data. In this chapter, the high-end frequencies (125 Hz, 250 Hz, 500 Hz) of each wavelet spectra (Fig. 4) are selected for modeling the first arrivals of the noise-free synthetics; and the dominant frequencies of each wavelet spectra are selected for modeling the first arrivals of the noisy synthetics.

To quantitatively compare the calculated traveltimes with the picked first arrivals from the synthetics, a minimum RMS difference is calculated to compare two sets of traveltimes. The minimum RMS difference is calculated by bulk shifting the two sets of traveltimes by subtracting the mean of each set. As a result, the minimum RMS
difference excludes the effect of a bulk time shift between the two sets of traveltimes but reflects the trace-to-trace travelt ime variations.

### 2.5. Results

![Figure 2.5](image)

Figure 2.5. Comparisons of infinite-frequency traveltimes from conventional ray theory (red lines) and frequency-dependent traveltimes from WDVS (blue lines) with picked first arrivals from noise-free synthetic traces using the 3 models in Fig. 3 and the source spectra with a dominant frequency of 100 Hz (Fig. 4) (green lines). The RMS differences between the infinite-frequency/frequency-dependent traveltimes and the picked first arrivals from the synthetics are indicated in the figures. (a), (c), (e) compare the calculated traveltimes with the picked arrivals with
the acoustic synthetics as the background, using Seed-300, Seed-200, Seed-120 models, respectively. (b), (d), (f) compare the calculated traveltimes with the picked first arrivals with an expanded vertical axis, corresponding to that in (a), (c), (e), respectively.

Figure 2.6. Comparisons of infinite-frequency traveltimes from conventional ray theory (red lines) and frequency-dependent traveltimes from WDVS (blue lines) with picked first arrivals from noise-free synthetic traces using the Seed-300 model in Fig. 3 and different source spectra in Fig. 4 (green lines). The RMS differences between the infinite-frequency/frequency-dependent traveltimes and the picked first arrivals from the synthetics are indicated in the figures. (a), (c), (e) compare the calculated traveltimes with the picked first arrivals with the acoustic synthetics as
the background, using source wavelets with a dominant frequency of 50, 100, 200 Hz, respectively. (b), (d), (f) compare the calculated traveltimes with the picked first arrivals with an expanded vertical axis, corresponding to that in (a), (c), (e), respectively.

Figure 2.7. Comparisons of frequency-dependent traveltimes calculated from WDVS using three different $L_{\text{max}}$ values of 0.5, 1, and 2 (gray lines), with the picked first arrivals from noise-free synthetic traces using the Seed-300 model in Figure 3 and the source spectra with a dominant frequency of 100 Hz in Figure 4 (black line). The RMS differences between the frequency-dependent traveltimes and the picked first arrivals from the synthetics are indicated in the figure.
Figure 2.8. Comparisons of infinite-frequency traveltimes from conventional ray theory (red lines) and frequency-dependent traveltimes from WDVS (blue lines) with the noisy acoustic synthetics using the Seed-300 model in Fig. 3 and the source spectra with a dominant frequency of 100 Hz in Fig. 4. (a) Noise-free synthetics. (b) 10% noise synthetics. (c) 25% noise synthetics. The random noise added to the synthetics in (b) and (c) has the same bandwidth as the synthetics and its RMS amplitude is based on the RMS amplitude of the noise-free early arrivals in a 20-ms window after the infinite-frequency traveltime of each trace. The two sets of noise are only different in amplitude; they represent the same sequence of random numbers.

Figure 5 presents the comparisons of the infinite-frequency and frequency-dependent WDVS traveltimes with the picked first arrivals, resulting from the 3 different models (Fig. 3) and the source spectra with dominant and high-end frequencies of 100 and 250 Hz, respectively (Fig. 4). In all cases, the 250 Hz WDVS-calculated traveltimes are more consistent with the picked first arrivals compared to the infinite-frequency traveltimes in terms of absolute time and the trace-to-trace traveltime variations as indicated by a smaller RMS difference between the frequency-dependent traveltimes and the picked traveltimes.

Figure 6 presents the comparisons of the infinite-frequency and frequency-dependent WDVS traveltimes with the picked first arrivals, resulting from 3 source spectra with different frequencies (Fig. 4) and the Seed-300 model (Fig. 3). In each case the results show that the WDVS-calculated frequency-dependent traveltimes are more consistent with the picked first arrivals compared to the infinite-frequency traveltimes in terms of absolute time and trace-to-trace traveltime variations. The results also show the
frequency dependence of both the picked first arrivals and WDVS-calculated traveltimes. As the frequency increases, the picked and frequency-dependent travel curves become rougher and approach the infinite-frequency traveltime curve. As expected, the frequency-dependent effects are less significant for higher frequencies. Also, given the characteristic length scale of the velocity anomalies in the model, the infinite-frequency, 500 Hz and picked traveltimes are nearly the same, as expected.

Figure 7 presents a comparison of the frequency-dependent WDVS traveltimes calculated using three different values of a free parameter in the WDVS algorithm, $L_{\text{max}}$, with the picked first arrivals, resulting from the Seed-300 model (Fig. 3) and the source spectra with dominant and high-end frequencies of 100 and 250 Hz, respectively (Fig. 4). In the WDVS algorithm, the value of $L_{\text{max}}$ equals the number of periods from a central model node over which the velocity is averaged for that node (Zelt and Chen, 2016). The results show that among 0.5, 1, and 2, the best $L_{\text{max}}$ value is 1 in that its corresponding frequency-dependent traveltimes are most consistent with the picked first arrivals in terms of absolute time, and similarities in trace-to-trace traveltime variations. When the $L_{\text{max}}$ value is 0.5, the frequency-dependent traveltimes are biased fast; when the $L_{\text{max}}$ value is 2, the frequency-dependent traveltimes are biased slow.

Figure 8 presents the comparisons of the infinite-frequency and frequency-dependent WDVS traveltimes with the noisy synthetics, resulting from the Seed-300 model (Fig. 3) and the source spectra with dominant and high-end frequencies of 100 and 250 Hz, respectively (Fig. 4). Both 100 and 250 Hz WDVS traveltimes are calculated and noise-free, 10% and 25% noisy synthetics are presented. As the noise level increases, both the infinite-frequency and high-end (250 Hz) frequency-dependent traveltimes seem
less realistic, \textit{i.e.}, less likely to be consistent with the first arrivals that one would pick from the noisy traces.

\textbf{2.6. Discussion and Conclusions}

This chapter presents a series of comparisons of WDVS-calculated traveltimes with band-limited full wavefield synthetics, showing that WDVS can calculate frequency-dependent traveltimes that are significantly more consistent with the picked first arrivals from the band-limited full wavefield synthetics than conventional ray theory infinite-frequency traveltimes. This suggests that WDVS captures some of the frequency-dependent behavior of finite-frequency wave propagation. The presented examples are realistic for near-surface seismic wave propagation in terms of the source spectra, the magnitude of the velocity anomalies, the relative size of the velocity anomalies compared to the wavelengths, and the noise level. Compared to infinite-frequency traveltimes, the WDVS-calculated traveltimes are more consistent with the picked first arrivals (considered the “true” traveltimes) in terms of absolute time and trace-to-trace traveltime variations. These results justify the use of WDVS for modeling frequency-dependent traveltimes that can serve as the forward modeling component of a tomographic inversion method (Zelt and Chen, 2016).

The extra step when using WDVS to calculate a frequency-dependent traveltime, compared to using conventional ray theory, is to specify a frequency for the velocity smoothing. The tests in this chapter select the frequency based on the spectra of the source wavelet and the noise level in the synthetics: the high-end frequency is selected for modeling the first arrivals of noise-free synthetics; the dominant frequency is selected
for modeling the first arrivals of noisy synthetics. In dealing with real data, one should analyze the spectra of the early arrivals and the level of noise to select a proper frequency using the results presented here for synthetic data as a guideline (e.g., Zelt and Chen, 2016; Chen et al., 2016).

Adjusting the $L_{max}$ value (Fig. 7) has the same effect as varying the modeling frequency. Using an $L_{max}$ value smaller than 1 is equivalent to using a proportionately higher frequency; using an $L_{max}$ value larger than 1 is equivalent to using a proportionately lower frequency. The results presented here suggest that using an $L_{max}$ value of 1 means that the appropriate modeling frequencies for high and low signal-to-noise ratio data are the high-end and dominant frequencies, respectively.

The WDVS algorithm, including its use in frequency-dependent traveltime tomography (Zelt and Chen, 2016), is most applicable to near-surface studies of P- and S-wave first-arrival-time data where the typical seismic wavelength is large relative to the total path lengths and the length-scale of velocity anomalies (e.g., Chen et al., 2016). WDVS is also applicable to exploration- and crustal-scale data, but the effects will be less significant (e.g., Kiser et al., 2015).

### 2.7. Acknowledgements

This research was funded by National Science Foundation grant EAR-1056073. We thank William Symes for providing the acoustic simulation code.
Chapter 3

Application of frequency-dependent traveltime tomography and full waveform inversion to realistic near-surface seismic refraction data

3.1. Abstract

We present a synthetic test that uses a workflow consisting of a new frequency-dependent traveltime tomography (FDTT) method to provide a starting model for full waveform inversion (FWI) for near-surface seismic velocity estimation from refraction data. Commonly-used ray-theory-based traveltime tomography methods may not be valid in the near surface given the likelihood of relatively large seismic wavelengths compared to the length scales of heterogeneities that are possible in the near surface. FDTT makes use of the frequency content in the seismic waves in both the forward and inverse modeling steps. In this application to a near-surface benchmark model, the results show
that FDTT can better recover the magnitude of velocity anomalies than infinite-frequency (ray-theory) traveltime tomography (IFTT). FWI can fail by converging to a local minimum when there is an absence of sufficiently low frequency data and an accurate starting model, either of which if present can provide long-wavelength constraints on the inverted velocity model. Both IFTT and FDTT models can serve as adequate starting models for FWI. However, FWI produces significantly better results starting from the FDTT model as compared to the IFTT model when low frequency data are not available. The final FWI models provide wavelength-scale structures allowing for direct geologic interpretation from the velocity model itself, demonstrating the effectiveness of FDTT and FWI in near-surface studies given the modest experiment and data requirements of refraction surveys.

3.2. Introduction

Seismic methods, including tomography, are used to address a number of near-surface environmental, engineering, archeological, neotectonic and resource exploration problems (Steeples, 2001; Pelton, 2005). Ray theory is commonly assumed to analyze the arrival times of seismic waves for velocity model estimation (e.g., Zhao and Xu, 2010; Ramachandran et al., 2011; Baumann-Wilke et al., 2012). Compared with global-scale or crustal-scale studies, where the size of anomalies is typically much larger than that of the seismic wavelength, in the near surface (<100m depth), the size of anomalies is often comparable with or smaller than the seismic wavelength (e.g., Gao et al., 2007), invalidating the infinite-frequency assumption of ray theory.

Zelt and Chen (2013) present a form of frequency-dependent traveltime
tomography (FDTT) that takes frequency content into consideration in both the forward and inverse modeling steps. It represents a small modification to ray theory in which the velocity model is pre-smoothed and the width of the sensitivity kernels is extended, both of which are according to the local seismic wavelength (Zelt and Chen, 2013). The advantage of FDTT over ray-theory-based infinite-frequency traveltime tomography (IFTT) is more significant for near-surface data because of the relatively longer seismic wavelengths and smaller heterogeneities. Applications of FDTT to realistic synthetic and real data reveal better results than IFTT, especially in terms of recovery of the true magnitude of velocity anomalies (Zelt and Chen, 2013). As a result, the model obtained from FDTT has the potential to provide more accurate information for geologic interpretation and can be used as a more effective starting model for full waveform inversion (FWI) (Pratt, 1999).
In this chapter we test FDTT and FWI using a synthetic dataset generated from a realistic near-surface velocity model (Fig. 1(a)) that was used previously in a blind test of first-arrival-time inversion and tomography methods (Zelt et al., 2013). The estimated velocity models from ten participants in the blind test using eight different inversion algorithms are generally consistent in their large-scale (> wavelength) features, although showing only smooth expressions of the true model’s key features (Fig. 1(b)). The deeper part of the model (> 35m) is not as well resolved as the shallower part given the ray coverage (Fig. 1(c)). Since the original release of the traveltime data and the velocity model in the blind test in 2011, there have been studies (e.g., Rohdewald, 2014; Stoyer, 2012) using different methods analyzing the traveltimes to improve the inverted model resolution.

The primary goal of this chapter is to serve as a follow-up and an extension to all the previous discussions on this realistic synthetic model by generating and inverting waveform data. The focus of this chapter is to test the ability of the combination of FDTT and FWI to characterize the large-scale and wavelength-scale structures in near-surface studies, compared to only characterizing the large-scale structures as shown by the previous studies that just used traveltime data.
Previous applications of FWI to near-surface field data (e.g., Smithyman et al., 2009; Adamczyk et al. 2014) commonly present synthetic tests as a way of assessing FWI. These tests use accurate background velocity models as the starting model for FWI to recover velocity anomalies with regular shapes (e.g., a checkerboard model). Using an accurate background model as the starting model means that the researchers assume the suitability of using a model from traveltime tomography (TT) as the starting model for FWI.

Brenders and Pratt (2007a) presented blind tests for a realistic crustal-scale model using a workflow including TT to obtain the starting model for FWI. In a following paper, Brenders and Pratt (2007b) tested the influence of adjusting different input parameters on FWI, but they did not test the influence of using different starting models for FWI.

This chapter uses the realistic synthetic near-surface velocity model in Zelt et al. (2013) and extends the discussions by incorporating waveform data and using a combined workflow of FDTT and FWI. The results and extended discussions in this chapter (1) confirm the suitability of using IFTT/FDTT to provide the starting model for FWI, (2) demonstrate the ability of FDTT to produce a more accurate starting velocity model for FWI to mitigate the lack of low frequency data, (3) show the improvement in model estimation using a combined strategy of FDTT and FWI, and (4) promote the use of FDTT and FWI in near-surface studies given the modest experiment and data requirements of refraction surveys over conventional reflection surveys.
3.3. Model and Data

The true velocity model (Fig. 1(a)) represents a geologic setting consisting of unconsolidated sediment overlying faulted bedrock (Zelt et al., 2013). The key features and potential targets for the seismic inversions in this model are: (1) a thin low-velocity layer in the sediments ~5-m deep between 12.5 m and 112.5 m lateral position, (2) a steep bedrock offset of ~12 m centered at 95 m lateral position, and (3) a steeply dipping (~35°) low-velocity fault zone in the bedrock centered at 185 m lateral position and 20-m depth. The surface topography is flat.
Figure 3.2. Representative synthetic acoustic seismograms from two end shots. The grey dots indicate the manually-picked first arrival times. A reducing velocity (Vr) of 2000 m/s is used to amplify subtle changes in the apparent velocities of both the near- and far-offset traces within the smallest possible time window. The seismic waveforms in the shaded areas were not used for waveform inversion as a result of the time window excluding later arrivals.
For the blind test presented in Zelt et al. (2013), the inverted data were traveltimes calculated using wavelength-dependent velocity smoothing (Zelt and Chen, 2013) corresponding to 100 Hz for 101 shots and 100 receivers. In this chapter, waveform data are needed for FWI but not available from the blind test. Acoustic waves are simulated in the frequency domain (Pratt 1999), transformed to the time domain for manual first-arrival picking as the input for traveltime tomography (Fig. 2), and later time-windowed (see Methods section) as a realistic processing step before transforming back to the frequency domain for waveform inversion. In this study, first-arrival picks and waveforms from 25 sources are used as compared to calculated times from 101 sources in Zelt et al. (2013); the smaller number of sources corresponds to a more realistic seismic experiment in terms of field work effort and cost. This reduction was also motivated by a result in Zelt et al. (2013) that also used a quarter of the shots to do the inversion and the resulting model fit the data from all 101 shots equally well. We have used 25 evenly spaced sources between 0 and 300 m, and 96 evenly spaced receivers between 1.5625 and 298.4375 m. The dominant frequency of the source wavelet is 80 Hz and the energy is down about 20 dB at 2 and 250 Hz. This spectra is consistent with that of a realistic near-surface seismic survey (e.g., Doll et al., 1998).

3.4. Methods

This chapter uses frequency-dependent traveltime tomography (FDTT) (Zelt and Chen, 2013) and full waveform inversion (FWI) (Pratt et al., 1998; Pratt 1999) in a combined workflow where FDTT provides a starting model for FWI. Both methods solve a nonlinear inverse problem through a local decent method starting from an initial
velocity model and iteratively updating it to reduce the differences between the modeled data and observed data. FDTT inverts first-arrival times by calculating a frequency-dependent traveltime. FWI inverts the waveforms of the early arrivals by solving the acoustic wave equation in the frequency domain. In this section, we review the important components of the FDTT and FWI methodologies, but refer the readers to Zelt and Chen (2013), Pratt et al. (1998) and Pratt (1999) for more details. In the following equations, bold lowercase letters denote column vectors, and bold capitals are second-order matrices. The operations between the quantities are matrix multiplication unless otherwise explicitly stated.

3.4.1. **Frequency-Dependent Traveltime Tomography**

The forward modeling component of FDTT consists of a wavefront-tracking algorithm that solves the 2D eikonal equation on a square grid of velocity nodes using finite difference operators (Vidale, 1988), with modifications to allow for large velocity gradients (Hole and Zelt, 1995) and to take frequency into account (Zelt and Chen, 2013). The modification to calculate frequency-dependent traveltimes involves pre-smoothing the velocity model using a local cosine-squared, wavelength-dependent operator determined by the chosen frequency; this approach is called wavelength-dependent velocity smoothing (WDVS).

The observed data are the first arrivals that are picked corresponding to the onset of seismic energy of each seismic trace (Fig. 2) using a semi-automated scheme whereby a few picks were made interactively, and the intervening picks were determined automatically using a cross-correlation scheme (Zelt, 1999), resulting in a total of 2400
picks. With noise-free synthetic data, as in this study, the onsets can be picked relatively accurately. As a result, the picks are interpreted as high frequency traveltimes, and thus the high end of the data spectrum is chosen for forward modeling during FDTT. We used 250 Hz corresponding to the frequency where the source wavelet energy is down ~20 dB. Each pick is assigned a 1 ms uncertainty, corresponding to a quarter of a period at 250 Hz, to be used in the inversion process for the stopping criteria based on an appropriate data misfit between the calculated first arrivals (modeled data) and the picked arrivals (observed data) (Zelt, 1999).

The inverse step of FDTT is the same as that used in IFTT as described by Zelt and Barton (1998), except the linear system of equations relating traveltime residuals and slowness perturbations is less sparse for FDTT because of the wavelength-dependent width of the sensitivity kernels (Zelt and Chen, 2013). For IFTT, the sensitivity kernels are nonzero only in the model cells along the ray path.

Given the strong lateral variations in the true model, a best-fit model with minimal lateral (2-D) structure was determined from the picked traveltimes using the Zelt and Smith (1992) algorithm to serve as the starting model for IFTT and FDTT (Fig. 3(b)). For both IFTT and FDTT a regularized inversion algorithm is used to minimize the data misfit and model roughness to provide the smoothest model with a proper data misfit (Zelt and Barton, 1998). The objective function $\Phi$ of the regularized inversion is expressed as

$$\Phi(m) = \delta t^T C_d^{-1} \delta t$$

$$+ \lambda [\alpha (m^T C_h^{-1} m + s_z m^T C_v^{-1} m) + (1 - \alpha)(m - m_o)^T (m - m_o)]$$

(1)
where \( \mathbf{m} \) is the model vector; \( \mathbf{m}_0 \) is the starting model vector; \( \delta t \) is the traveltime data residual vector; \( \mathbf{C}_d \) is the data covariance matrix; \( \mathbf{C}_h \) and \( \mathbf{C}_v \) are the horizontal and vertical roughness matrices, respectively; \( \lambda \) is the trade-off parameter that weighs the relative importance of data misfit and model regulations; \( s_z \) determines the relative importance of maintaining vertical versus horizontal model smoothness; and \( \alpha \) weighs the relative importance of model smoothness and the absolute model perturbation. The superscript \( T \) represents transpose, and the superscript \( -1 \) represents matrix inverse. During inversions, \( s_z \) and \( \alpha \) are fixed, \( \lambda \) starts with a relatively large free-parameter value and is decreased automatically during the iterations updating the velocity model to introduce large-scale structure first and fine-scale structure in later iterations. At the end, \( \lambda \) is adjusted to yield a final model that provides a normalized chi-squared misfit value of one, quantitatively indicating that the picked arrivals have been fit to within the assigned uncertainties (Zelt, 1999).

### 3.4.2. Frequency Domain Full Waveform Inversion

The forward modeling of FWI solves the isotropic, 2-D, visco-acoustic wave equation in the frequency domain that can be expressed as

\[
\mathbf{S} \mathbf{u} = \mathbf{f} \quad (2)
\]

where the complex impedance matrix \( \mathbf{S} = \mathbf{K} - \omega^2 \mathbf{M} + i\omega \mathbf{C} \) is a linear combination of the stiffness matrix \( \mathbf{K} \), the mass matrix \( \mathbf{M} \), and the damping matrix \( \mathbf{C} \), and where \( \omega \) is the angular frequency. The vector \( \mathbf{f} \) represents the source signature, and the vector \( \mathbf{u} \) represents the seismic wavefield to be calculated. Equation (2) can be expressed as \( \mathbf{u} = \mathbf{S}^{-1} \mathbf{f} \), and it is discretized by a finite-difference method and numerically solved.
by LU decomposition (Pratt et al., 1998; Pratt 1999). An 80 Hz-dominant frequency Ricker wavelet is used as the source. The data are modeled from 2 to 400 Hz with a 2 Hz interval.

Pre-processing the data to account for the acoustic approximation is crucial for real data applications, for example, pre-inversion matching of the amplitude between the observed and modeled data (e.g., Brenders and Pratt, 2007a; Ravout et al., 2004). It is not needed for the synthetic test in this chapter since the same algorithm is used for forward modeling and inversion. Therefore, the inverted models in this chapter represent ideal results one can expect from FWI.

Another commonly used preprocessing procedure for real data is time windowing to exclude later-arrivals (e.g., shear-waves, surface-waves) that are not modeled by visco-acoustic modeling (e.g., Brenders and Pratt, 2007a). To honor the reality that only the early arrivals are typically used in FWI as a result of time windowing, an offset-dependent time window is applied to the modeled data in this study. The window is 50 ms at 0 m offset and linearly increases to 100 ms at 100 m offset, and is 100 ms beyond 100 m offset (Fig. 2). This window starts at the picked first-arrival time with a 10-ms cosine taper added before and after.

The outermost loop of the inversion is over frequency from low to high to mitigate the non-linearity of the seismic inverse problem. This implementation strategy is known as a multi-scale approach (Bunks et al., 1995; Sirgue and Pratt 2004). Effectively, low-wavenumber, large-scale structure will first be introduced to the model updates, and then high-wavenumber, fine-scale structure will improve the resolution in later iterations. The
choice of the starting low frequency and stopping high frequency follows a trial and error procedure according to a visual assessment of (1) the spectra and signal-to-noise ratio of the real data, (2) the updated model for its overall geological sensibility and artifacts, and (3) the predicted data for their similarities with the real data (Gao et al., 2006; Jaiswal et al., 2009). In this study we determined that the usable frequency range for FWI is 6 to 64 Hz. Below this range, the data have a poor consistency among different shots, which can be caused by the windowing operations; above this range, the data begin to introduce fine-scale artifacts especially to the shallow part of the model where there are relatively lower velocities and accordingly smaller seismic wavelengths. The data in the frequency domain are grouped with increasing frequency content. The first group is 6Hz by itself, the second group is 8 and 10 Hz, and the rest of the groups each contain three frequency components without skipping or overlapping from 12 to 64 Hz, resulting in 11 groups in total. For the inversion of each group of frequencies, the source signature is first treated as an unknown and is inverted from the observed data assuming a known velocity model provided by traveltime tomography. The source inversion is linear, i.e., requires only one iteration, and produces a good match for both the amplitude and phase of the true source wavelet [equation 17 of Pratt (1999)]. After that, the velocity model will be updated multiple times according to the stopping criterion assuming a known source signature. The source signature and the velocity model are iteratively updated as the frequency of each group increases.

The non-linear seismic waveform inversion problem is solved by a local decent method. The data misfit is defined by the $L_2$ norm of the data residual in the frequency domain so the objective function $\Phi$ is
where \( \mathbf{m} \) is the model vector; \( \delta \mathbf{d} \) is the waveform data residual vector. The superscript \( T \) represents matrix transpose, and the superscript * represents complex conjugate. The velocity model is iteratively updated according to the gradient of the data residual to minimize the misfit function. The data residual gradient gives the direction in which the misfit function decreases most for a given amount of change of the model. It is an iterative process and model updates stop when the misfit function diverges or when the largest model change among the entire grid is relatively small, less than 1% of the lowest velocity value in the model.
3.5. Results

Figure 3.3. True and estimated models from this study and from SAGEEP 2011 blind test (Zelt et al., 2013). a) True model. b) Starting model for IFTT and FDTT. c) IFTT model. d) FDTT model. Both IFTT and FDTT models provide an RMS misfit of 1 ms to the manually-picked traveltimes. Models e) to h) are from the SAGEEP 2011 blind test. They are named after their index numbers in Zelt et al. (2013): e) V8, f) V14, g) V12, h) V10. Among the 14 models presented in Zelt et al. (2013), V8 produces the smallest mean of the absolute difference from the true model, V14 produces the smallest mean of the relative difference, V12 produces the smallest standard deviation of the absolute difference, and V10 produces the smallest standard deviation of the relative difference. These four models provide RMS misfits of 1.4 ms, 1.6 ms, 1.2 ms, and 1.1 ms, respectively, to the synthetic traveltimes (Zelt et al., 2013). In each model, the black contours correspond to 500 m/s and 2000 m/s. The white contours are from the true model for comparison.

Figs. 3(c) and 3(d) show the IFTT and the FDTT models, both providing an RMS
misfit of 1 ms to the picked data. A velocity contour of 2000 m/s is used to interpret the boundary separating the sediments and bedrock (Zelt et al., 2013). The shape of the boundary defines the bedrock offset and the dipping fault. Both the IFTT and the FDTT models give a smooth expression of the bedrock offset (centered at 95 m) and the shallow part of the fault (< 30 m depth). For comparison, Figs. 3(e)-3(h) present the best traveltime tomography models from Zelt et al. (2013) according to four different statistical measures.

![Figure 3.4. Comparison of a horizontal slice at Z=25 m through the true model (Figure 3a), IFTT model (Fig. 3(c)), FDTT model (Fig. 3(d)), and models from Zelt et al. (2013) (Figs. 3(e)-(h)).](image)

The horizontal model slice at 25 m depth shows that the FDTT model overall matches the true model better than the IFTT model and the models from Zelt et al. (2013) in terms of a more accurate recovery of the absolute velocity values and a sharper velocity contrast defining the targeted geologic structures (Fig. 4). This horizontal model slice captures the bedrock offset, a sharp velocity change from ~1000 m/s to ~3000 m/s between ~90-100 m lateral position, and the shallow part of the dipping fault, an ~20-m wide low-velocity interval between ~180-200 m lateral position. All the estimated
velocity models show these two velocity features smoothly. The FDTT, IFTT and V8 models present these two features more sharply with larger velocity contrasts. The FDTT and V8 models are better than the IFTT model with more accurate high velocities between ~100-170 m lateral position, that directly contributes to the definition of two key features in the model, the bedrock offset and the dipping fault. Elsewhere the V8 model is inferior to the IFTT and FDTT models with generally less accurate velocity estimation, particularly an incorrect high-velocity feature at the left side of the model. But the comparisons are not ideal in that the models from Zelt et al. (2013) were obtained from inverting calculated 100 Hz traveltimes from 101 shots, whereas the IFTT and FDTT models were obtained from picked traveltimes of the waveforms from 25 shots. Therefore, the differences between the TT models in Zelt et al. (2013) and the IFTT and FDTT models presented here are not merely the result of using different algorithms.

Figure 3.5. Full waveform inversion (FWI) estimated models. a) and c) are produced by taking IFTT model (Fig. 3(c)) as the starting model, and by starting the inversion at 6 Hz and 12 Hz, respectively. b) and d) are produced by taking FDTT model (Fig. 3(d)) as the starting model, and by starting the inversion at 6 Hz and 12 Hz, respectively. All inversions end at 64 Hz. In the text they are named a) IFTT-FWI-6Hz-64Hz model, b) FDTT-FWI-6Hz-64Hz model, c) IFTT-FWI-12Hz-64Hz model, and d) FDTT-FWI-12Hz-64Hz model. In each model, the black contours correspond
to 500 m/s and 2000 m/s. The white contours are from the true model for comparison.

Taking the IFTT model (Fig. 3(c)) as the starting model, the IFTT-FWI-6Hz-64Hz model (Fig. 5a) is produced by starting FWI at 6 Hz and ending at 64 Hz. This model adds wavelength-scale structures that are weakly expressed or not visible in the starting model, including the low-velocity zone at 5 m depth defined by the 500 m/s contour and the deeper part of the dipping fault below 30 m depth as seen in the 45 m depth slice in Fig. 6(b). The FWI model also improves the estimation of other features, e.g., a steeper expression of the bedrock offset centered at ~95 m and a narrower expression of the shallow part of the dipping fault (Fig. 6(a)). The same observations apply to the comparison of the FDTT model (Fig. 3(d)) and the FDTT-FWI-6Hz-64Hz (Fig. 5(b)) model. Comparisons of the seismograms produced by the true model and the inverted models (Fig. 7) show a significantly better match of the seismic phases achieved by the FDTT-FWI-6Hz-64Hz model over the FDTT model. Compared with the FDTT model, the FDTT-FWI-6Hz-64Hz model achieves a 32% reduction in the waveform misfit function (Eq. 3). Both the qualitative and quantitative comparisons indicate a successful application of FWI.
Figure 3.6. Comparison of horizontal slices through the true model (Fig. 3(a)), traveltime tomography estimated models (Figs. 3(c)-(d)), and full waveform inversion estimated models (Figs. 5(a)-(b)) at a) $Z=25$ m, and b) $Z=45$ m. The black boxes highlight key features in the true model.
Figure 3.7. Representative synthetic acoustic seismograms from two end shots. a) and b) calculated from the true model (Fig. 3(a)), c) and d) calculated from the FDTT model (Fig. 3(d)), and e) and f) calculated from the FDTT-FWI-6Hz-64Hz model (Fig. 5(b)). A reducing velocity (Vr) of 2000 m/s is used to amplify subtle changes in the apparent velocities of both the near- and far-offset traces within the smallest possible time window. The seismic waveforms in the shaded areas were not used for waveform inversion as a result of the time window excluding later arrivals. Areas highlighted by the dashed gray ellipses are for detailed comparisons of the true and estimated model waveforms. Data are displayed with low-pass filter of 70 Hz.

Comparing the IFTT-FWI-6Hz-64Hz model (Fig. 5(a)) with the FDTT-FWI-6Hz-
64Hz model (Fig. 5(b)), the latter shows a more continuous low-velocity zone at ~5m depth, and a more accurate image of the deep part of the dipping fault (Fig. 6(b)). A realistic challenge for an application of FWI to real data is the lack of usable low frequency data due to noise, leading the inversion to converge to a local minimum (Bunks et al., 1995). We simulate this situation by producing the IFTT-FWI-12Hz-64Hz model (Fig. 5(c)) and the FDTT-FWI-12Hz-64Hz model (Fig. 5(d)), which are the equivalents of the IFTT/FDTT-FWI-6Hz-64Hz models with a substitute of 12 Hz as the starting frequency. We observe degradation in image quality from starting the inversion at a higher frequency (compare Fig. 5(a) to 5(c), and 5(b) to 5(d)), for example, the low-velocity zone is less continuous, and the image of the deep part of the dipping fault is less accurate. However, the FDTT-FWI-12Hz-64Hz model (Fig. 5(d)) is noticeably better than the IFTT-FWI-12Hz-64Hz model (Fig. 5(c)) and even slightly better than the IFTT-FWI-6Hz-64Hz model (Fig. 5(a)) in its continuous expression of the low-velocity zone, and a more accurate image of the deep part of the dipping fault.

Comparing the FWI models with the TT models, even the IFTT-FWI-12Hz-64Hz model (Fig. 5(c)), the least accurate of the FWI results among the four, is still significantly better than any of the TT models (Figs. 3(c)-(h)).

3.6. Discussion

This study demonstrates the use of frequency-dependent traveltime tomography (FDTT) and full waveform inversion (FWI) for near-surface seismic velocity estimation. FDTT uses first-arrival times and takes the seismic data’s frequency content into consideration to provide a starting model for FWI. FWI uses the waveforms of the early
arrivals to improve the model resolution by adding wavelength-scale structure that is not in the traveltime model. In this study, the low-velocity zone, the bedrock offset, and the dipping fault are well imaged in all versions of the FWI models (Fig. 5) that would allow for direct geologic interpretation from the velocity model itself. These results confirm the suitability of using TT, and FDTT in particular, to provide the starting model for FWI. The importance of a good starting model for FWI has been widely discussed and confirmed by several researchers using FWI for different scales of geophysical studies (e.g., Pratt et al., 2002; Shah et al., 2012; Wang et al., 2014). The unique contribution of this chapter is in proposing to use a new frequency-dependent traveltime tomography to produce a better starting model than that produced by conventional ray-theory-based IFTT.

In this study, FDTT is superior to IFTT in two respects. First, the FDTT model (Fig. 3(d)) more accurately estimates the magnitude of features, though it provides a similar scale of resolution as the IFTT model (Fig. 3(c)). This observation also applies to the model slice comparison (Fig. 4) of the FDTT model with other TT models in Zelt et al. (2013), albeit this comparison is not ideal in that Zelt et al. (2013) used 100 Hz WDVS traveltimes from 101 sources, while this chapter uses picked traveltimes from only 25 sources. Secondly, while both the IFTT and FDTT models (Figs. 3(c) and 3(d)) are appropriate starting models for FWI, FDTT provides a more suitable starting model in that the FDTT-FWI models reveal better velocity estimation compared with equivalent IFTT-FWI models that use the same frequency band-width of data (Fig. 5(b) compared with 5(a), 5(d) compared with 5(c)). More importantly, in a realistic situation where the low frequency data, in this case from 6 to 10 Hz, is not usable, the FDTT-FWI-12Hz-
64Hz model (Fig. 5(d)) still recovers the targeted features well, e.g., the deep part of the
dipping fault, but the IFTT-FWI-12Hz-64Hz model (Fig. 5(c)) represents a significant
degradation compared to the IFTT-FWI-6Hz-64Hz model (Fig. 5(a)).

Despite the potential for high resolution velocity estimation by FWI as shown in
this study, application of FWI to real field data is challenging because of FWI’s extreme
nonlinearity. It also requires heuristic preprocessing steps including pre-inversion
amplitude matching to account for assumptions made in the wave-field modeling
(Virieux and Operto, 2009), e.g., 2-D acoustic modeling as in this study. FWI seeks the
model with the best fit to the preprocessed data, and as such, there is always a risk of
introducing unrealistic fine-scale structure. FWI uses heuristic criteria (e.g. Jaiswal et al.,
2009) to stop the inversion iterations including a visual comparison of the modeled and
observed data (Fig. 7). As opposed to the difficulties of applying FWI to real data, no
obstacle prevents the use of FDTT as long as first arrivals can be picked. In addition,
traveltimes have a more linear relationship with seismic velocity than amplitudes. The
FDTT model is obtained in a more objective manner using a misfit of the predicted and
observed data quantified by a chi-square value taking the pick uncertainties into
consideration (Zelt, 1999). It also honors Occam’s principle that states that a minimum-
structure solution containing only the model features required by the data is the best
(Constable et al., 1987). Although the development of a more objective workflow for the
application of FWI is needed, we consider FDTT and other frequency-dependent/phase
inversion techniques (e.g., Ellefsen, 2009) to be a practical step forward over ray-theory-
based TT for the non-expert practitioners dealing with near-surface seismic data.

Although we applied a realistic time window to honor the fact that FWI typically
only makes use of the early arrivals when dealing with real data (Pratt, 1999), this study presents FDTT and FWI under nearly ideal circumstances since the data have no noise and the waveform data were calculated by the same algorithm used in FWI. Nevertheless, this chapter shows what is possible with high-quality, realistic, near-surface seismic data using a combined strategy of FDTT and FWI.

The usable high frequency limit of FWI is not necessarily the high end of the spectrum, or even the dominant frequency of the seismic data (e.g., Jaiswal et al., 2008), as is the case for this study. One factor that may influence the usable high frequency limit is the signal-to-noise ratio of each inverted frequency group. However, in this study with noise-free synthetic data, the frequency groups between the 20 dB points at 2 and 250 Hz, including those containing the dominant frequency and beyond, have effectively the same signal-to-noise ratio, contaminated only by a very small amount of numerical noise. Note, the relative energy differences between frequency groups, i.e., being highest near the dominant frequency, are not important because the frequencies are not inverted simultaneously. Therefore, in this study, as with real data, the following three factors play a role in determining the high frequency limit. First, the larger of the source and receiver spacing (the source spacing in this study) limits the highest possible frequency we can use while avoiding spatial aliasing that can introduce spurious artifacts into the inverted model (Brenders and Pratt, 2007b). Second, a poor starting model can limit the highest possible frequency we can use to avoid FWI converging to a local minimum that may cause spurious artifacts in the inverted model (Bunks et al., 1995; Pratt et al., 2002). Finally, the time windowing operation excludes later parts of the waveforms that may reduce the consistency of the high frequency data, e.g., a later-arriving seismic event may
be included within the time window of some shot gathers but outside the window of others. The effects of these factors are taken into consideration through empirical assessments during the inversions according to the criteria stated in the Methods section and we determined that 64 Hz is the usable high frequency limit, lower than the dominant frequency of 80 Hz.

3.7. Conclusions

The workflow of applying FDTT followed by FWI to near-surface seismic refraction data shows the ability to achieve a velocity image with wavelength-scale features for direct interpretation of targeted structures. The velocity contours can be used to define the boundary between layers and zones of different rock properties for geologic interpretation. We used a laptop with a single processor to process the data from a realistically feasible seismic refraction survey to produce the FDTT and FWI models. The modest field acquisition effort and computational methods are accessible for solving near-surface problems in the environmental and engineering industries.

The results show that FDTT better estimates velocity than IFTT and the FDTT velocity model serves as a more suitable starting model for FWI. FDTT uses frequency information in both the forward and inversion modeling steps (Zelt and Chen, 2013). Its improvements stem from the calculation of frequency-dependent traveltimes, naturally resulting in frequency-dependent sensitivity kernels. This is fundamentally different from traveltime tomography methods that adopt a frequency-dependent kernel for the inversion without forward calculating a frequency-dependent traveltime (e.g., Watanabe et al., 1999; Liu et al., 2009).
3.8. Acknowledgements

This research was funded by National Science Foundation grant EAR-1056073 and Department of Energy grant DE-FG07-97ER14827. The authors would like to thank Gerhard Pratt for providing the FWI code. The authors are also grateful to Priyank Jaiswal, Fuchun Gao, and Geoff Chambers for help with applying FWI.
Chapter 4

Detecting a known near-surface target through application of frequency-dependent traveltime tomography and full waveform inversion to P- and SH-wave seismic refraction data

4.1. Abstract

We have applied a combined workflow of frequency-dependent traveltime tomography (FDTT) and full waveform inversion (FWI) to 2D near-surface P- and SH-wave seismic data to image a known target consisting of a buried tunnel with concrete walls and a void space inside. FDTT inverted the P- and SH-wave picked traveltimes at 250 Hz to provide long-wavelength background velocity models as the starting models for FWI. FWI inverted 18-54 Hz P-wave data and 16-50 Hz SH-wave data to produce velocity models with sub-wavelength- and wavelength-scale features allowing for direct
interpretation of the velocity models as is usually carried out in conventional imaging using seismic reflection data. The P- and SH-wave models image the top part of the tunnel at the correct location at a depth of 1.6 m as a high-velocity anomaly. The P-wave models also image the air in the void space of the tunnel as a low-velocity anomaly. The inverted models were assessed by synthetic tests, the consistency of the inverted sources, and the fits between the predicted and observed data. As a comparison, conventional ray-theory infinite-frequency traveltime tomography (IFTT) was also applied in a combined workflow with FWI. The comparisons of the inverted models favor the use of FDTT over IFTT because: 1) The FDTT models better recover the magnitude of the velocity anomalies, and 2) The FDTT model serves as a better starting model for FWI, which results in a more accurate FWI velocity estimation with better recovery of the magnitude and location of the key features.

4.2. Introduction

Near-surface geophysics investigates the shallow crust to about 100 m depth to address environmental, engineering, archeological, neotectonic and resource exploration interests (Steeples, 2001; Butler, 2005; Doll et al., 2012). Ray-theory-based seismic traveltime tomography methods are commonly used for estimating the near-surface velocity structure, contributing to the understanding of lithology, geological structure, and geological processes (e.g., Zhao and Xu, 2010; Ramachandran et al., 2011; Baumann-Wilke et al., 2012). However, compared with global-scale or crustal-scale studies, where the size of anomalies is typically much larger than that of the seismic wavelength, in the near surface the size of anomalies is often comparable with or smaller
than the seismic wavelength (e.g., Gao et al., 2007), contrary to the infinite-frequency assumption of ray theory. This motivates the use of full wavefield methods such as full waveform inversion (FWI) for improved geophysical properties estimation with better resolution (e.g., Gao et al., 2006, 2007; Smithyman et al., 2009; Adamczyk et al., 2014).

The main challenge faced by FWI for real data applications is the degree of nonlinearity of the inverse problem, e.g., the well-known problem of the presence of local minimums in the objective function and the convergence to a local minimum using local optimization methods (Bunks et al., 1995). Virieux and Operto (2009) present an excellent overview to illuminate the state of the art of FWI. The practical challenges and difficulties for the applications of FWI to real data include: 1) the development of a sufficiently accurate starting model, 2) the lack of low frequencies, 3) the presence of noise, and 4) the approximate modeling of the wave-physics complexity (Virieux and Operto, 2009). Besides the summary of the suggested solutions in Virieux and Operto (2009), recent progress that aims to tackle these difficulties includes: 1) new formulations of the objective function and modified inversion workflows, e.g., minimization of the data residuals using the $L_1$ norm to reduce the effects of noise (Brossier et al., 2010), and $L_2$ norm minimization in Laplace-Fourier-domain waveform inversion that enhances the use of phase information and suppresses amplitude information (Kamei et al., 2014; Jun et al., 2014); 2) new preconditioning operators to help the inversion converge to a more geologically meaningful model, e.g., preconditioning with non-stationary directional Laplacian filters (Guitton et al., 2012); and 3) building a more accurate starting model for FWI using newly developed traveltime methods, e.g., stereotomography (Prieux et al., 2012) and frequency-dependent traveltime tomography (FDTT) (Zelt and Chen, 2016).
In this chapter we use FDTT (Zelt and Chen, 2016) to provide a starting model for FWI as a way of mitigating the nonlinearity of FWI. FDTT consists of a small modification to ray theory that takes the frequency content of the data into consideration in both the forward and inverse modeling steps. Theoretically, its advantage over ray-theory-based infinite-frequency traveltime tomography (IFTT) is more significant for near-surface data because of the relatively longer seismic wavelengths and smaller-scale, larger-magnitude heterogeneities. Applications of FDTT to realistic synthetic and real data reveal better results than IFTT, especially in terms of recovery of the true magnitude of velocity anomalies (Zelt and Chen, 2016). As a result, the model obtained from FDTT also provides more information for geologic interpretation. Moreover, Chen and Zelt (2016) show that FDTT can provide a more accurate starting model for FWI than IFTT using a realistic synthetic test, and the benefits are more significant in the absence of usable low frequency data.

The objective of this study is to apply FDTT and FWI to 2-D P- and SH-wave data (vertical source-receiver and horizontal transverse source-receiver components, respectively) to image a known buried concrete tunnel through estimation of high-resolution velocity models. The focus is to test the ability of a workflow combining FDTT and FWI to image very shallow, near-surface, sub-wavelength structures through estimated velocity models. As a comparison, the IFTT and IFTT-FWI results are also presented.

Previous studies show confidence in detecting sub-wavelength objects with seismic waves. Bachrach and Reshef (2010) show that a buried 0.15 m diameter pipe at 1.5 m depth can be detected with dense 3-D receiver arrays using ten times larger wavelengths
through seismic migration imaging methods. As a comparison, our target is buried at a similar depth with a dimension of about one fifth of the wavelength.

Seismic tomography with S-waves has inherent benefits compared with P-waves, specifically smaller wavelengths due to lower velocities, which can theoretically contribute to better image resolution. However, S-wave FWI in near-surface studies has not been as popular as P-wave FWI, mostly because of the difficulty in generating good signal-to-noise ratio data. In this chapter, we adopt Chambers’ (2009) modified version of Pratt’s (1999) acoustic frequency-domain FWI algorithm to invert horizontal transverse component (SH-wave) data.

In the following sections we describe the seismic experiment and data, followed by a review of the methods: FDTT (Zelt and Chen, 2016) and FWI (Pratt et al., 1998; Pratt 1999; Chambers 2009). We then present the results from a realistic synthetic P-wave test designed according to our knowledge of the target, followed by the real data results from the P- and SH-wave data.

### 4.3. Site and Experiment

The seismic data used in this chapter were collected in 2011 over a target consisting of a utility tunnel buried in shallow sediments at Rice University campus (Figure 1). The buried depth, lateral location and structure of the tunnel are known. The concrete walls of the tunnel are 0.6-m-thick on top, and 0.3-m-thick on the other sides. The void space within the tunnel is mostly empty and for modeling purposes is assumed to be filled with air in the following synthetic test.
Figure 4.1. A schematic view of the 2-D survey showing the source-receiver geometry and the known location and structure of the target: the concrete tunnel walls are 0.6 m thick on top, and 0.3 m thick on the other sides. The average 1-D linear-gradient background velocity models for P- and SH-waves between 0 and 10 m depth were determined from the P- and SH-wave traveltime data.
Figure 4.2. Two component seismic receiver. The receiver is vertically inserted into the ground using a single spike and records two orthogonal components, $a$ and $b$, that are symmetric with respect to the vertical axis.

The 24-meter-long seismic survey line was perpendicular to the buried tunnel. There are 25 shots and 72 receivers, yielding a total of 1800 traces. The shots have a 1-m interval between 0 and 24 m. The receivers start at 0.167 m with 0.333-m interval and end at 23.833 m. The surface elevation variation is 0.26 m. The P- and SH-wave data were collected separately. For the P-wave source, each shot was a stack of 10 vertical hammer blows on a trailer hitch ball vertically mounted on the ground. For the SH-waves, each shot was a stack of 10 horizontal hammer strikes on a trailer hitch ball mounted on the end of a wooden fence post lying on the ground weighted down by four
people and fixed by three vertical steel rods to avoid horizontal displacement caused by the hammer blows.

For both the P- and SH-wave sources, each receiver group consists of two 40 Hz geophones arranged to receive two orthogonal components, roughly like a two-component version of a Galperin geophone (Figure 2). The receiver was vertically planted into the soil beneath a grass field so that the plane of the orthogonal components is perpendicular to the 2-D survey line. Each geophone pair can produce vertical and horizontal (transverse) components. For the P-wave source, the two received components are summed to isolate vertical motion by canceling out the horizontal motion. For the SH-wave source, the two received components are subtracted to isolate transverse horizontal motion by canceling out the vertical motion.
Figure 4.3. P-wave data examples. (a) Shot gather from $x=0$ m (upper panel), and the average amplitude spectrum of the 25-ms window starting from the first-arrival pick of each trace (lower panel). (b) Shot gather from $x=24$ m (upper panel), and the average amplitude spectrum of the 25-ms window starting from the first-arrival pick of each trace (lower panel). The data in the upper panel are shown trace normalized to a common maximum amplitude. The red dots and lines indicate the picked first arrival times. The traveltime and early waveform advances due to the tunnel walls are indicated by the dashed gray ellipses.
Figure 4.4. SH-wave data examples. (a) Shot gather from x=0 m (upper panel), and
the average amplitude spectrum of the 25-ms window starting from the first-arrival
pick of each trace (lower panel). (b) Shot gather from x=24 m (upper panel), and the
average amplitude spectrum of the 25-ms window starting from the first-arrival
pick of each trace (lower panel). The data in the upper panel are shown trace
normalized to a common maximum amplitude. The red dots and lines indicate the
picked first arrival times. The traveltime advances and waveform anomalies
probably due to the tunnel walls are indicated by the dashed gray ellipses.

As data examples, Figures 3 and 4 show the raw shot gathers of P- and SH-waves,
respectively, from the two end shots. Overall, the data have good signal-to-noise ratio,
though the P-wave data have better signal-to-noise ratio in that the first arrival times are
clearer, and the early seismic waveforms are more continuous and consistent. There are
indications of traveltime advances and early waveform advances (Figures 3 and 4), that are interpreted to be associated with the high-velocity concrete tunnel walls.

4.4. Methods

This chapter uses frequency-dependent traveltime tomography (FDTT) (Zelt and Chen, 2016) and full waveform inversion (FWI) (Pratt et al., 1998; Pratt, 1999; Chambers, 2009). FDTT and FWI are complementary in the sense that FDTT can provide a starting model for FWI. Both methods solve a nonlinear inverse problem through a local decent method starting from an initial velocity model and iteratively updating it to reduce the differences between the predicted and observed data. FDTT inverts first-arrival times by calculating frequency-dependent traveltimes. FWI inverts the waveforms of early arrivals by solving the acoustic wave or SH-wave equation. In this section, we review the important components of the FDTT and FWI methodologies for completeness and we refer the readers to Zelt and Chen (2016), Pratt et al. (1998), Pratt (1999), and Chambers (2009) for more details. The P- and SH-waves are inverted separately. The same FDTT algorithm is used for the P- and SH-wave data because both of their first-arrival times follow the eikonal equation. We use Pratt’s (1999) frequency-domain acoustic FWI algorithm for the P-wave data, and Chambers’ (2009) modified version of it for the SH-wave data.

4.4.1. Frequency-Dependent Traveltime Tomography

The forward modeling component of FDTT consists of a wavefront-tracking algorithm that solves the 2-D eikonal equation on a square grid of velocity nodes using
finite-difference operators (Vidale, 1988), with modifications to allow for large velocity gradients (Hole and Zelt, 1995) and to take frequency into account (Zelt and Chen, 2016). The modification to calculate frequency-dependent traveltimes involves pre-smoothing the velocity model using a local wavelength-dependent operator determined by the chosen frequency corresponding to the center or maximum frequency content of the first-arrival waveforms depending on how the data are picked (Zelt and Chen, 2016). Since both the P- and SH-wave first-arrival-times follow the eikonal equation, they can both be modeled and inverted by the same algorithm. We will first introduce FDTT with the chosen modeling and inversion parameters for the P-wave data, and then explain the differences in the chosen parameters for the SH-wave data.

In this study, the first arrival picks, representing the observed data, correspond to the onset of energy of each seismic trace (e.g., Figure 3 for the P-waves). They were obtained using a semi-automated scheme whereby a few picks were made manually by eye, and the intervening picks were determined automatically using a cross-correlation scheme (Zelt, 1999), resulting in a total of 1800 picks. Analyzing the waveform data to determine the frequency for the traveltimes is the only extra step needed for preparing the input data for FDTT compared with that of IFTT. With good signal-to-noise ratio P-wave data, as in this study, the onsets can be picked accurately. As a result, the picks are assumed to be dependent on the high-frequency limit of the first-arrival waveforms (Zelt and Chen, 2016). We chose 250 Hz to model and invert the P-wave traveltimes after examining the spectrum of the first-arrival waveforms (Figure 3). At 250 Hz, the seismic amplitude is about ~5-10% of the normalized peak amplitude. Each P-wave pick is assigned a 1 ms uncertainty based on the data frequency, ambient noise and an
examination of the reciprocal time differences of the P-wave picks (Figure 5a). A 1-ms uncertainty is also consistent with our use of 250 Hz for modeling and inverting the P-wave traveltimes in that it corresponds to one quarter of a period. The uncertainty is used in the stopping criteria in the inverse step based on an appropriate data misfit between the calculated first arrival times (predicted data) and the picked arrival times (observed data) corresponding to a normalized chi-squared misfit of one (Zelt, 1999).

A best-fit 1-D model (Figure 6b) was determined from the P-wave traveltimes using the Zelt and Smith (1992) algorithm to serve as the starting model for P-wave IFTT and FDTT. A regularized inversion algorithm is used to minimize a combination of the data misfit and model roughness to provide the smoothest model with a proper data misfit (Zelt and Barton, 1998). In the inverse step, for IFTT each traveltime is sensitive to the velocities along the corresponding ray path as in conventional traveltime tomography, whereas FDTT uses sensitivity kernels that extend one wavelength on either side of the travel path (Zelt and Chen, 2016). A trade-off free parameter, lambda, weighs the relative importance of data misfit and model smoothness. Lambda is decreased automatically from its initial free-parameter value to introduce large-scale structure in early iterations and fine-scale structure in later iterations. At the end, lambda is adjusted to yield a final model that provides a normalized chi-squared misfit of one.
Figure 4.5. (a) P-wave, and (b) SH-wave histogram of number of shot-receiver pairs versus reciprocal first-arrival time difference in 0.5-ms bins.

The first-arrival waveforms of the SH-waves (Figure 4) have a slightly lower central frequency and lower high-frequency limit than that of the P-waves (Figure 3), and have a lower signal-to-noise ratio, although similar reciprocal time differences (Figure 5). As a result, we used a larger pick uncertainty of 1.5 ms. However, we used the same frequency of 250 Hz for the modeling and inversion of the SH-wave picks because FDTT can yield model artifacts when using frequencies lower than some threshold caused by the noise in real data (Zelt and Chen, 2016).

4.4.2. Frequency Domain Full Waveform Inversion

Pratt and his collaborators (Pratt et al., 1998; Pratt 1999) developed the theoretical background of the frequency-domain waveform inversion and implemented it for the isotropic visco-acoustic case. It has been widely used for crustal scale (e.g., Brenders and Pratt, 2007a) and shallow scale (e.g., Smithyman et al., 2009; Gao et al., 2006, 2007)

In the frequency domain, the 2D acoustic and SH-wave equations at frequency $\omega$ can be expressed as (Marfurt, 1984; Chambers 2009)

$$\omega^2 a u(\omega) + \nabla \cdot (b \nabla u(\omega)) = f(\omega).$$  \hspace{1cm} (1a)

For the acoustic case

$$a = \frac{1}{\lambda}; \quad b = \frac{1}{\rho}$$ \hspace{1cm} (1b)

and for the SH-wave case

$$a = \rho; \quad b = \mu$$ \hspace{1cm} (1c)

where the vector $f$ is the source signature, the wavefield vector $u$ is pressure and transverse displacement in the acoustic and SH-wave cases, respectively, $\rho$ is density, and $\lambda$ and $\mu$, are Lame parameters. Taking a finite-difference approximation to equation 1a results in a matrix equation that can be expressed as

$$Su = f$$  \hspace{1cm} (2)

where $S$ is a complex impedance matrix depending on the frequency and Lame parameters, $f$ represents the source signature vector and $u$ represents the wavefield vector to be calculated. Equation 1 can be expressed as $u = S^{-1}f$, and numerically solved by LU decomposition (Pratt et al., 1998; Pratt 1999). The Lame parameters in the impedance matrix $S$ are calculated by the given velocity and density models, and for simplicity, the
density model is assumed to be proportional to the fourth root of velocity (Gardner et al., 1974). In addition, in this chapter an attenuation model is not used or inverted (e.g., Jaiswal et al., 2008). The differences between solving the acoustic and SH-wave equations are in the implementation of the complex impedance matrix $S$, i.e., the substitution of $a$ and $b$ in equations 1b and 1c, but the processes are the same (Chambers 2009).

Except solving the wavefields, the pre-processing workflow and inversion framework for both the P- and SH-waves are the same. We first explain the processing and inversion steps for the P-waves and then point out the differences in the chosen parameters for the SH-waves.

Pre-processing of the raw observed data is crucial to improve the signal-to-noise ratio and to exclude the wavefield that is not included in acoustic forward modeling. The key pre-processing steps include: band-pass filtering, time and offset windowing, and pre-inversion amplitude matching. A zero-phase Ormsby filter with a 5-10-60-120 Hz bandwidth is applied to the P-wave data and the low and high cut frequency is determined in a trial and error manner according to the frequency band that is used for the inversions as discussed later. The time window, with a 10-ms cosine taper added before and after, starts at the manually picked first arrival times and the ending time is determined visually to exclude the seismic events that are not modeled by acoustic modeling (e.g., surface waves, converted shear waves). During the transformation of the data from the time domain to frequency domain, a single-value exponential damping factor was applied to each trace starting from its first arrival (Pratt, 1999). As a result, the energy of each trace within 50 ms of the first arrival is dominant in the inversions. The
near-offset (< 4-5 m) traces are muted out given the dominant presence of surface waves, along with some noisy individual traces beyond that offset according to a visual check of signal-to-noise ratio and trace consistency. The amplitude of the processed observed data is scaled through an empirical approach to match the amplitude-versus-offset (AVO) behavior of the predicted data (Brenders and Pratt, 2007a). The predicted data uses a two-excursion Keuper wavelet with a 40 Hz dominant frequency and the final velocity models from IFTT or FDTT as the starting model. Both the observed data and the predicted data are sorted into 8 bins by offset from 4 to 24 m with a 2.5-m interval. The root-mean-square (RMS) amplitude of each bin is calculated, and a least-squared fit of the logarithmic RMS amplitude versus offset is found for each set of observed and predicted data. The scaling factor is determined by the difference of the least-squared fits of the two sets of data [equation (15) of Brenders and Pratt, 2007a]. This amplitude correction approximately accounts for some of the effects on the amplitudes not included in the forward modeling, such as anelastic attenuation, source radiation pattern, and 3-D geometric spreading (e.g., Priyank et al., 2008; Smithyman et al., 2011). All of the processing steps are applied without shifting the traces in time.

The outermost loop of the inversion is over frequency from low to high to mitigate the non-linearity of the inverse problem. This implementation strategy is known as a multi-scale approach (Bunks et al., 1995; Sirgue and Pratt, 2004). Effectively, low-wavenumber, large-scale structure will first be introduced to the model updates, and then high-wavenumber, fine-scale structure will improve the resolution in later iterations. The choice of the starting low frequency and stopping high frequency follows a trial and error procedure according to a visual assessment of (1) the spectra and signal-to-noise ratio of
the real data, (2) the updated model for its overall geological sensibility and artifacts, and (3) the predicted data for their similarity with the observed data (Gao et al., 2006; Jaiswal et al., 2009). We determined that the usable frequency range for P-wave FWI is 18 to 54 Hz. The data are grouped with increasing frequency content. The first group is 18 and 20 Hz, the last group is 52 and 54 Hz, and the intermediate groups each contain 3 frequency components without skipping or overlapping from 22 to 50 Hz, resulting in 7 groups in total. For the inversion of each group of frequencies, the source signature is first treated as an unknown and is inverted from the observed data assuming a known velocity model provided by traveltime tomography or the output from the previous frequency group. The initial source signature is a Keuper wavelet with a 40 Hz dominant frequency covering the bandwidth of the preprocessed data. The source inversion is linear (requires only one iteration) and produces a good match for both the amplitude and phase of the true source wavelet [equation (17) of Pratt, 1999]. After this, the velocity model is updated multiple times according to the stopping criteria assuming a known source signature. The source signature and the velocity model are updated as the frequency of each group increases.

The non-linear seismic waveform inversion problem is solved by a local decent method. The data misfit is defined by the $L_2$ norm of the data residual in the frequency domain. The velocity model is iteratively updated according to the gradient of the data residual to minimize the misfit function. The data residual gradient gives the direction in which the misfit function decreases most for a given amount of change in the model. It is an iterative process and model updates stop when the misfit function diverges or when the largest model change among the entire grid is relatively small, less than 1% of the lowest velocity value in the model.
The regularization parameters are not defined in the objective function but applied to the data residual gradient directly (Brenders and Pratt, 2007a). A low-pass wavenumber filter was applied to the gradient to suppress the unstable, high-wavenumber components in the model. For data of a given frequency, $f$, in a model with a minimum velocity $c_{min}$, the upper limit of this filter is $2f/c_{min}$ (Brenders and Pratt, 2007a; Wu and Toksoz, 1987). A taper zone of one fifth of the upper limit is used to avoid artifacts generated from the filter. A cosine taper is used along the depth axis to suppress high gradients near the sources due to the singularity of the Green’s functions at their origins (Brenders and Pratt, 2007a). This taper has a value of 0 at 0.2 m depth (the depth of the deepest source due to elevation) and above, and a value of 1 at 1.2 m depth and below.

The SH-waves go through the same pre-processing workflow of band-pass filtering, time and offset windowing, and pre-inversion amplitude matching, with a difference in that the bandwidth of the zero-phase Ormsby filter is 5-10-50-100 Hz. The usable frequency range for SH-wave FWI is determined to be 16 to 50 Hz, resulting in 6 groups each contain 3 frequency components without skipping or overlapping.
4.5. Synthetic P-Wave Results

Figure 4.6. P-wave traveltime tomography models from synthetic data. (a) True model. (b) 1-D starting model. (c) Infinite-frequency traveltime tomography (IFTT) model. (d) IFTT model displayed as perturbations with respect to the 1-D model. (e) 250Hz frequency-dependent traveltime tomography (FDTT) model. (f) FDTT model displayed as perturbations with respect to the 1-D model. The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 100 m/s in both model and perturbation plots.

We first performed P-wave synthetic tests honoring the real acquisition geometry and frequency band, using a velocity model (Figure 6a) with a tunnel structure whose
location and dimension is based on the known information (See section 2). The concrete velocity is assumed to be 4000 m/s (Malhotra and Carino, 2003) and the air velocity is 330 m/s; the background model is the 1-D starting model used for IFFT and FDTT (Figure 6b). The synthetic wavefields were produced by the frequency domain method as mentioned in section 3.2 and then transformed to the time-domain for first-arrival-picking. Gaussianly-distributed random noise with a standard deviation of 1 ms was added to the manual picks. The 1-D background model (Figure 6b) was used as the starting model for IFFT and 250Hz FDTT. Both the IFFT and FDTT models (Figure 6c and e) provide a chi-squared misfit of 1 ms. Both the IFFT and FDTT models image the top of the concrete with a high velocity anomaly at the accurate location as indicated by the known outline of the tunnel, centered at about 1.6 m depth and 8 m lateral position, best seen in the model perturbation plots (Figure 6d and f). The FDTT model contains a stronger magnitude of the high velocity feature compared to that in the IFFT model. Also, the FDTT model contains a low-velocity anomaly below the top of the concrete, presumably due to the void space, that is not seen in the IFFT model plots (Figure 6c and d).
Figure 4.7. P-wave full waveform inversion (FWI) models from synthetic data. (a) IFTT-FWI model that uses the IFTT model (Figure 6c) as the starting model. (b) IFTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 6b). (c) FDTT-FWI model that uses the FDTT model (Figure 6e) as the starting model. (d) FDTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 6b). The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 100 m/s in both model and perturbation plots.

FWI took the final IFTT and FDTT models as starting models and produced the final IFTT-FWI (Figure 7a and b) and FDTT-FWI (Figure 7c and d). Overall, the FWI models appear to image the velocity anomalies associated with the tunnel more accurately than the TT models. Comparing the FDTT and FDTT-FWI perturbation plots (Figure 6f and 7d), the central part of the low-velocity anomaly under the top of the concrete is relatively weak and slightly off to the right in the FDTT model, but is stronger in the FDTT-FWI model centered within the concrete walls. Both the FDTT-FWI and
IFTT-FWI models image the high-velocity top of the concrete and low-velocity void space at the correct locations, but the FDTT-FWI model better recovers the magnitude of these velocity anomalies.

The synthetic tests set expectations for how well the tunnel can be imaged using TT and FWI with the realistic influencing factors that were also used for the real data applications with TT and FWI: acquisition setup, useable frequency band, and the chosen free parameters. The models, particularly the FDTT-FWI model, successfully image the high velocity top concrete wall and low velocity air, but the two sides and the bottom concrete walls are not clearly or consistently shown in all the models, not surprisingly since they are thinner than the top, significantly less than one seismic wavelength in the background model at this depth (about ~5 to 10 m wavelength assuming a 50 Hz wave as used in FWI). Also, the sides and bottom of the tunnel represent structure beneath a larger high-velocity zone and as such first-arrival waveforms are not expected to be very sensitive to structure immediately underneath.
4.6. Real P-Wave Results

Figure 4.8. P-wave traveltome tomography models from field data. (a) Infinite-frequency traveltome tomography (IFTT) model. (b) IFTT model displayed as perturbations with respect to the 1-D model (Figure 6b). (c) 250Hz frequency-dependent traveltome tomography (FDTT) model. (d) FDTT model displayed as perturbations with respect to the 1-D model (Figure 6b). The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 100 m/s in both model and perturbation plots.
Figure 4.9. Wavepaths for 250 Hz for P-wave FDTT model (Figure 8c); raypaths for P-wave IFTT model are very similar. Gray dots show source locations. For clarity, every third wavepath is plotted. The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom).

Figure 8 shows the IFTT and FDTT models derived from the real P-wave data. Both the IFTT and FDTT models provide a chi-squared misfit of 1 ms. They both contain the high-velocity top of the concrete tunnel centered at the correct lateral and vertical locations. The FDTT model shows a stronger magnitude of the top of the concrete as high as \( \sim 750 \text{ m/s} \) compared to \( \sim 550 \text{ m/s} \) in the IFTT model. Also, the FDTT model perturbation plot more clearly shows a low-velocity anomaly below the top of the concrete corresponding to the void space. The constraint on the final TT models can be qualitatively assessed looking at the ray/wavepath coverage (Figure 9). The wavepath coverage is good to \( \sim 5 \text{ m} \) depth, deep enough to image the tunnel.
Figure 4.10. P-wave full waveform inversion (FWI) models from field data. (a) IFTT-FWI model that uses the IFTT model (Figure 8a) as the starting model. (b) IFTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 6b). (c) FDTT-FWI model that uses the FDTT model (Figure 8c) as the starting model. (d) FDTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 6b). The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 100 m/s in both model and perturbation plots.

FWI took the IFTT and FDTT final models as starting models and produced the IFTT-FWI and FDTT-FWI models (Figure 10). The FWI models contain more small-scale structure than that in the TT models. The top of the concrete tunnel is the largest magnitude velocity anomaly in both models. Comparisons of the FWI models with the TT models (Figure 8) show that the FWI models increase the magnitude of the high-velocity feature representing the top of the concrete tunnel, e.g., from a maximum of \( \sim 750 \text{ m/s} \) in the FDTT model (Figure 8c) to a maximum of \( \sim 950 \text{ m/s} \) in the FDTT-FWI
model (Figure 10c). The FDTT-FWI model (Figure 10c and d) most clearly presents the low-velocity feature representing the void space inside the tunnel with stronger magnitude and a more accurate location than in the IFTT-FWI model.

The FDTT-FWI model better recovers the tunnel features than the IFTT-FWI model with a stronger magnitude, e.g., a maximum of ~950 m/s for the top of the concrete tunnel in the FDTT-FWI model (Figure 10c) compared to a maximum ~750 m/s in the IFTT-FWI model (Figure 10a), and a more accurate position, e.g., the low-velocity feature within the tunnel is more accurately located in the FDTT-FWI model (Figure 10c and d) than that in the IFTT-FWI model (Figure 10a and b) where it is ~1 m to the left of the known location. As in the synthetic tests, the two sides and the bottom walls of the tunnel are not imaged in the inverted FWI models.
Figure 4.11. One-dimensional velocity profiles created by averaging in the X direction between 7.1 and 8.6 m, corresponding to the assumed interior extent of the tunnel. (a) Averaged velocity profiles for the P-wave inverted models (Figures 8a and c, 10a and c). (b) Averaged velocity profiles for the SH-wave inverted models (Figures 12b and d, 13a and c). The vertical dashed line in (a) represents the seismic velocity of the air, 330 m/s. The horizontal dashed lines in both (a) and (b) indicate the positions of the top and bottom of the concrete tunnel.

Figure 11a shows 1-D velocity-depth profiles to summarize the relative performance of the methods in terms of how well they recover the high-velocity top of the concrete tunnel and the low-velocity air below. Each profile is a lateral average of the corresponding 2-D model over the known interior extent of the tunnel (X=7.1 to 8.6 m). The FWI models clearly outperform the TT models, and within each group, the FDTT and FDTT-FWI models are better than IFTT and IFTT-FWI models, respectively. Also,
in this case the FDTT model is comparable with IFTT-FWI model in recovering the high-velocity top of the concrete, although the IFTT-FWI model better recovers the low-velocity void space. The depth profiles also shows that the bottom of the concrete tunnel, thinner and deeper, is beyond the resolution of the smoothly inverted velocity models.

### 4.7. Real Sh-Wave Results

![Figure 4.12. SH-wave traveltime tomography models from field data. (a) 1-D starting model. (b) Infinite-frequency traveltime tomography (IFTT) model. (c) IFTT model displayed as perturbations with respect to the 1-D model. (d) 250Hz frequency-dependent traveltime tomography (FDTT) model. (e) FDTT model displayed as perturbations with respect to the 1-D model. The known position of the concrete tunnel is shown by the black rectangle.](image)
concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 50 m/s in both model and perturbation plots.

A best-fit 1-D model (Figure 12a) was determined from the SH-wave traveltimes using the Zelt and Smith (1992) algorithm to serve as the starting model for SH-wave IFTT and FDTT. Figure 12b and c show the IFTT model inverted from the real SH-wave data in absolute and perturbation plots, respectively, and Figure 12d and e are the corresponding FDTT model plots. Both the IFTT and FDTT models provide a chi-squared misfit of 1.5 ms. They both show the high-velocity top of the concrete tunnel centered at the correct vertical location but laterally slightly to the left by ~1 m. The FDTT model better recovers the top of the concrete than the IFTT model with a stronger magnitude; the perturbation plots show that the high-velocity top of the concrete in the FDTT model is ~ 250 m/s faster than the background, and it is ~ 100 m/s faster than the background in the IFTT model.
Figure 4.13. SH-wave full waveform inversion (FWI) models from field data. (a) IFTT-FWI model that uses the IFTT model (Figure 12b) as the starting model. (b) IFTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 12a). (c) FDTT-FWI model that uses the FDTT model (Figure 12d) as the starting model. (d) FDTT-FWI model displayed as perturbations with respect to the 1-D model (Figure 12a). The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom). Contour interval is 50 m/s in both model and perturbation plots.

FWI took the final IFTT and FDTT models as starting models and produced the IFTT-FWI (Figure 13a and b) and FDTT-FWI (Figure 13c and d) models. The high-velocity top of the concrete tunnel is not well located in the IFTT-FWI model, probably due to the higher noise level in the SH-waveform data. Unlike the P-wave FWI models, the SH-wave FWI models do not show a low-velocity anomaly associated with the void space inside the tunnel, presumably because shear waves do not propagate through air. Both of the FWI models appear to image the thinner right side of the tunnel as a high-
velocity anomaly centered at 8.8 m lateral position and ~3 m depth, which does not exist in the TT models (Figure 12). By comparison, the left side of the tunnel is not imaged, probably because of the lack of wavepath coverage as compared to the right side (Figure 14). Notwithstanding what appear to be significant small-scale artifacts in the upper ~4 m of both SH-wave FWI models, presumably due to noise, the FDTT-FWI model images the tunnel structure (top and right side) with the best resolution of all the FWI models (P-wave and SH-wave). This is probably because of the combined benefits of the frequency effects in FDTT not found in IFTT and the shorter wavelengths of SH-waves compared to P-waves, e.g., about ~2.5 to 5 m for SH-waves compared to ~5 to 10 m for P-waves in the background velocity model at the depth range of the tunnel assuming a 50 Hz wave as used in FWI.

Figure 4.14. Wavepaths for 250Hz for SH-wave FDTT model (Figure 12d); raypaths for SH-wave IFTT model are very similar. Gray dots show source locations. For clarity every third wavepath is plotted. The known position of the concrete walls of the tunnel are indicated (0.6 m thick on top, 0.3 m thick on sides and bottom).
The SH-wave 1-D depth profiles (Figure 11b) show that both of the FDTT models (TT and FWI) are better than both of the IFTT models in recovering the magnitude of the top of the concrete tunnel. The SH-wave depth profiles do not account for the fact that inverted tunnel features are centered to the left of the known tunnel position (Figures 12 and 13), partly out of the range of the averaging between X=7.1 to 8.6 m (the interior extent of the tunnel) that produces the profiles, probably due to noise in the observed SH-wave data.

4.8. Discussion

For P- and SH-wave real data over a known target, this study demonstrates the use of frequency-dependent traveltime tomography (FDTT) and full waveform inversion (FWI) for near-surface seismic velocity estimation to detect sub-wavelength structure. FDTT uses first-arrival times and takes the seismic data’s frequency content into consideration to provide a starting model for FWI. FWI uses the waveforms of the early arrivals to improve the model resolution to wavelength-scale and detection to sub-wavelength scale. All of the inverted models from both the P- and SH-wave traveltime and waveform data image the high-velocity top of the concrete tunnel that is 2.1-meter-wide and 0.6-meter-thick, less than a wavelength of either the P- or SH-waves in the background velocity model at that depth (~5 m for P-waves and ~2.5 m for SH-waves assuming a 50 Hz wave as used in FWI). The P-wave FDTT-FWI model (Figure 10c and d) images the low-velocity void space inside the tunnel and the SH-wave FDTT-FWI model (Figure 13c and d) weakly images the right side of the tunnel. The sub-wavelength-scale inverted velocity anomalies allow for accurate geologic interpretation
from the velocity model itself, i.e., without the need for more conventional imaging techniques using seismic reflection data. The results also confirm the suitability of using TT, and FDTT in particular, to provide suitable starting models for FWI.

For the real data in this study we have the advantage of knowing the location and the structure of the target to ground truth the TT and FWI inverted models. In general studies without such prior information, the results can be assessed with synthetic tests such as those in Figures 6 and 7, and within the inversion workflows, e.g., the reasonableness of intermediate output, and the similarities between the predicted and observed data.

The P- and SH-wave TT models provide a 1 ms and 1.5 ms chi-squared misfit, respectively, determined by the picking uncertainties, to manually picked traveltimes. Fitting the data at the level of its uncertainty combined with regularization to minimize model roughness are consistent with Occam’s principle that states that a minimum-structure solution containing only the model features required by the data is the best (Constable et al., 1987).

Using realistic near-surface synthetic data, Chen and Zelt (2016) demonstrate the adequacy of using TT models as the starting model for FWI and particularly the advantages of using FDTT over IFTT, especially in the case of the lack of usable low frequency data. In this chapter, we also show the adequacy of using TT models, and especially FDTT models, as the starting model for FWI with three sets of data containing different levels of noise: the synthetic P-wave data with no noise, the real P-wave data with moderate noise, and the real SH-wave data with strong noise. To different extents,
these three cases show the advantages of FDTT over IFTT in that: 1) FDTT better estimates velocity anomalies with more accurate magnitude than that IFTT (Figures 6, 8, 11, and 12), and 2) FDTT provides a better starting model for FWI, resulting in the FDTT-FWI model containing velocity anomalies with more accurate magnitude and position than the corresponding IFTT-FWI model (Figures 7, 10, 11 and 13).
Figure 4.15. P- and SH-wave source signatures (trace normalized) at each shot location estimated from the processed data and inverted models. (a) Source signatures estimated from the processed P-wave data and the P-wave FDTT model (Figure 8c), (b) Source signatures estimated from the processed P-wave data and the P-wave FDTT-FWI model (Figure 10c); (c) Source signatures estimated from the processed SH-wave data and the SH-wave FDTT model (Figure 12c) (d) Source signatures estimated from the processed SH-wave data and the SH-wave FDTT-FWI model (Figure 13c). Note that the source signatures start before time 0 to account for the phase delay generated in the 2-D simulation (Pratt 1999; Morse and Feschbach, 1953).

The inverted source wavelets (Figure 15) are intermediate output in the FWI inversion workflows. The consistency of the inverted source wavelets among different shots can serve as a quality control tool to monitor the inversions (e.g. Gao et al., 2006; Jaiswal et al., 2009). Since several factors are coupled in the inversion workflow, the consistency among different shots can imply: 1) the effectiveness of the preprocessing steps to exclude noise and events that are not modeled by the acoustic solver, 2) the adequacy of using the TT model as the starting model for FWI, and 3) the reasonableness of the final FWI inverted model. The P-wave FDTT starting model and the FWI model inverted source signatures (Figure 15a and b) show good consistency in their phases among different shots, although the two on the left are slightly off. The SH-wave inverted source signatures (Figure 15c and d) are not as consistent especially after their first peaks, indicating artifacts in the inverted models and/or noise in the processed SH-wave waveforms.
Figure 4.16 P-wave data examples. (a), (c), (e) Shot gather for x=0 m. (b), (d), (f) Shot gather for x=24 m. (a), (b) Real processed data. (c), (d) Predicted data from the P-wave FDTT model (Figure 8c) and its corresponding inverted source (Figure 15a). (e), (f) Predicted data from the P-wave FDTT-FWI model (Figure 9c) and its corresponding inverted source (Figure 15b). The data are trace normalized. The early waveform advances due to the tunnel walls are indicated by the dashed gray ellipses, within which the FDTT-FWI-predicted data provide a better fit to the processed data.
Figure 4.17 SH-wave data examples. (a), (c), (e) Shot gather for x=0 m. (b), (d), (f) Shot gather for x=24 m. (a), (b) Real processed data. (c), (d) Predicted data from the SH-wave FDTT model (Figure 12d) and its corresponding inverted source (Figure 15c). (e), (f) Predicted data from the SH-wave FDTT-FWI model (Figure 13c) and its corresponding inverted source (Figure 15d). The data are trace normalized. Waveform anomalies probably due to the tunnel walls are indicated by the dashed gray ellipses, within which the FDTT-FWI-predicted data provide a better fit to the processed data.
Figure 4.18. P-wave frequency-domain data showing the real part of the Fourier components of the data at 40 Hz, with source number increasing from top to bottom and receiver number increasing from left to right. (a) Real processed data. (b) Predicted data from the P-wave FDTT model (Figure 8c) and its corresponding inverted source (Figure 15a). (c) Predicted data from the P-wave FDTT-FWI model (Figure 10c) and its corresponding inverted source (Figure 15b). Black areas
represent data not used in the inversion. The dashed gray ellipses highlight data fit improvements of the FDTT-FWI model over the FDTT model.

Figure 4.19. SH-wave frequency-domain data showing the real part of the Fourier components of the data at 30 Hz, with source number increasing from top to bottom and receiver number increasing from left to right. (a) Real processed data. (b)
Predicted data from the SH-wave FDTT model (Figure 12d) and its corresponding inverted source (Figure 15c). (c) Predicted data from the SH-wave FDTT-FWI model (Figure 13c) and its corresponding inverted source (Figure 15d). Black areas represent data not used in the inversion. The dashed gray ellipses highlight data fit improvements of the FDTT-FWI model over the FDTT model.

The similarity between the observed and predicted waveform data assesses the reasonableness of the FWI inverted models. In both time- and frequency-domain, the P- and SH-wave FWI model predicted data overall better matches the real processed data than the corresponding TT model predicted data (Figures 16-19). For example, the time-domain P-waveform advances (Figure 16), that are interpreted to be due to the high-velocity top of the concrete tunnel, are significantly better matched by the FWI model predicted data. As an example in the frequency domain, the better fits by the FWI model predicted data in the far-offset SH-waves (lower left and upper right corners in the grey ellipses, Figure 19) indicate a more accurate velocity model estimation, particularly in the deeper parts of the model.

Although the SH-wave data is noisier than the P-wave data, as is usually the case, this study shows the benefit of collecting SH-wave data and processing them with TT and FWI. If the P- and SH-wave data are inverted independently, the velocity anomalies common in both sets of inverted models, e.g., the high-velocity top of the concrete tunnel, help to validate each other and demonstrate the robustness of the overall modeling approach. In addition, SH-waves with lower velocity but similar usable frequency range as that of the P-waves, may detect smaller-scale features, e.g., the thinner side of the concrete tunnel on the right is suggested in the SH-wave FDTT-FWI model (Figures 13c
and d) though not in the P-wave FDTT-FWI model (Figures 10c and d).

**4.9. Conclusion**

The workflow of applying FDTT followed by FWI to near-surface P- and SH-wave seismic refraction data shows the ability to achieve a velocity image with sub-wavelength-scale features for direct interpretation of shallow structure. As the best inverted models from each set of data, both the P- (Figure 10c and d) and SH-wave (Figure 13c and d) FDTT-FWI models image the top wall of the concrete tunnel with a high-velocity anomaly. Additionally, the P-wave FDTT-FWI model includes a low-velocity anomaly corresponding to the void space in the tunnel, and the SH-wave FDTT-FWI model includes a high-velocity anomaly corresponding to one of the thinner sides of the tunnel, the locations of which are consistent with the prior knowledge of the tunnel. Besides the prior knowledge, the inverted models are validated by 1) the results of a realistic synthetic test; 2) The reasonableness of the intermediate output; and 3) similarities between the predicted data and the observed data.

The seismic survey was done in two days by three-five people and a laptop was used with a single processor to process and invert the data to produce the FDTT and FWI models. The modest field acquisition effort and computational methods are accessible to the environmental and engineering industries for near-surface studies.

**4.10. Acknowledgements**

This research was funded by National Science Foundation grant EAR-1056073 and
Department of Energy grant DE-FG07-97ER14827. The seismic acquisition instruments were provided by the IRIS PASSCAL Instrument Center. We thank the volunteers in the Rice earth science department for their help in acquiring the data. We thank Gerhard Pratt and Geoff Chambers for providing the original and modified FWI codes.
Conclusions

For near-surface engineering and environmental studies, this thesis promotes a combined use of frequency-dependent traveltime tomography (FDTT) and full waveform inversion (FWI) to process the near-surface seismic refraction data by presenting high-resolution inverted velocity models that accurately image the wavelength- and sub-wavelength-scale structures. These fine-scale structures allow for direct geologic interpretations for near-surface targets, e.g., the bedrock structure (Chapter 3) or a concrete tunnel (Chapter 4), as is usually carried out in conventional imaging using seismic reflection data.

Frequency-dependent traveltime tomography (FDTT) is a newly developed method that takes the frequency content of the seismic waves into consideration in both its forward modeling and inversion steps. One unique contribution of this thesis is to demonstrate that FDTT, compared to ray-theory-based infinite-frequency traveltime
tomography (IFTT) can serve as a significantly better starting model for FWI, particularly in the lack of high signal-to-noise ratio low-frequency data. The examples in Chapter 2 justify the use of wavelength-dependent velocity smoothing (WDVS) algorithm in FDTT for calculating a frequency-dependent traveltime. These examples reveal the frequency-dependent behaviors of WDVS that are consistent with the finite-frequency wave propagation, and compare the frequency-dependent traveltimes calculated using WDVS with that from synthetic seismographs, which shows a good match.

Though not presented in this thesis, there can be situations in real data applications where a single use of FDTT is preferable to a combined use with FWI. Compared to FDTT, FWI demands high signal-to-noise ratio data, and involves an extreme non-linear inversion problem. FDTT is applicable as long as the first arrival times can be picked, without placing greater demands on the data quality, as demanded by conventional ray-theory-based traveltime tomography.

The workflow of a combined use of FDTT and FWI is only demonstrated in 2D in this thesis. Zelt and Chen (2016) present both 2D and 3D examples for FDTT that achieves better velocity estimation compared to IFTT, as I show in the 2D cases in this thesis. Partially because of the computational cost, 3D waveform tomography methods have not been widely used in the academia. Future work as directed by this thesis can be applications of this combined use of FDTT and FWI to 3D refraction data.

Biondi (1997) attempts a finite-difference solution of the frequency-dependent eikonal equation derived from wave equation without high frequency approximation but
doubts its stability. An accurate and stable numerical solution of the frequency-dependent eikonal equation directly derived from wave theory, if achieved in the future, can be a substitute of WDVS to theoretically better model a frequency-dependent traveltime.
References


Biondi, B., 1997, Solving the frequency-dependent eikonal equation: Stanford Exploration Project, Report 73, November 18, P 325-338


Chen, J. and C.A. Zelt, 2016, Application of frequency-dependent traveltime tomography and full waveform inversion to realistic near-surface seismic refraction data: Journal of Environmental and Engineering Geophysics, in press.


