Rice University

Waveform Tomography and its Application at a Ground Water Contamination Site

By

Fuchun Gao

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

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

[Signatures]

Alan Levander, Carey Croneis Professor, Chair
Earth Science

[Signatures]

Colin Zelt, Associate Professor
Earth Science

[Signatures]

Fenglin Niu, Assistant Professor
Earth Science

[Signatures]

William Symes, Noah Harding Professor, Chair
Computational and Applied Mathematics

HOUSTON, TEXAS
MAY 2004
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.
ABSTRACT

Waveform Tomography and its Application at a Ground Water Contamination Site

by

Fuchun Gao

This thesis develops and applies seismic waveform tomography to solve the unique problem of imaging complicated shallow sub-structures with high resolution. Shallow sub-structures are commonly characterized by seismic reflection/refraction imaging, georadar and seismic travel time tomography (e.g., Steeples, 1998; Carcione et al., 2000 and Azaria, 2002). Their resolving power or applicability is often limited. In contrast, waveform tomography, a full wave field inversion technique, resolves sub-structures at a resolution that is a fraction of the illuminating wavelengths.

Forward modeling in waveform tomography is based on a finite difference solution to the acoustic wave equation in the space-frequency domain. During inversion for model parameters, the technique efficiently calculates the gradient of a misfit function with respect to model parameters by correlating back-propagated and forward modeled wave fields, avoiding the forbidding task of explicitly computing Frechet kernels. Part of this study compares travel time and waveform tomography in a synthetic cross-well test. The two tomographic approaches are found to be complementary if data contains no significant low frequency spectra.

I then apply waveform tomography to two datasets from a ground water contamination site at the Hill Air Force Base (HAFB) to sample formation heterogeneities and to map the 3D geometry of a buried paleo-channel where DNAPLs
(Dense Non-Aqueous Phase Liquids) were dumped. The first is a VSP-surface seismic experiment. The final velocity model from waveform tomography applied to the VSP dataset generally correlates well with lithology logs, depth migrated 2D/3D reflection data and a velocity model from 3D travel time tomography. Large velocity variations vertically and laterally (200m/s) occur in a distance as short as ~1m. The model is interpreted geologically and petrologically. Scale features down to ~1.5m were recovered.

I then apply waveform tomography to 45 2D seismic profiles extracted from a 3-D surface seismic experiment at HAFB, and recover the 3D geometry of a buried paleo-channel acting as a trap for DNAPLs. By combining the identified cross-sectional geometry, the 3D geometry of the channel is reconstructed. The subsurface map could be used to plan injection/extraction well placements with good precision and low cost in the on-going ground water remediation program.
ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Alan Levander for his support, valuable advices and constant encouragement through the years of my study at Rice and help with my thesis writing. I would also like to thank one of my committee members, Dr. Colin Zelt, for his advices, suggestions, critical comments and teaching. Special thanks go to my committee members Dr. William Symes for his insights and teaching, and Dr. Fenglin Niu for helpful discussions.

I would like to thank Dr. Gerhard Pratt for initially introducing me into the field of waveform tomography, his guidance in further developing the technique and numerous advices. Without his expertise in waveform tomography, the application of waveform tomography in this study is impossible.

I also would like to thank all field crew members in the Hill Air Force Base experiments for their hard working and great teamwork. Thanks for the HAFB administration, PASSCAL and Duke Engineering, Inc. for their co-operations.

I would like to thank Mary Cochran and Alex Hemsath for their computer expertise. Thank Diana Dana for preparing data for me, Peng Shen and Sangwon Ham for their co-operations, and my fellow graduate students for friendships.

Finally I would like to thank my wife Annie, my son Kevin and entire family members for their sincere support. I dedicate this thesis to my grandmother with whom I grew up. In my mind she planted hope which brought me through all the adventures.
TABLE OF CONTENTS

Abstract ................................................................................................................... ii
Acknowledgements ................................................................................................. iv
Table of Contents ..................................................................................................... v
List of Tables ........................................................................................................... viii
List of Figures .......................................................................................................... viii

Chapter 1: Introduction ............................................................................................ 1
   1.1 Non-seismic methods ...................................................................................... 2
   1.2 Seismic methods ........................................................................................... 3
       1.2.1 Seismic refraction method ..................................................................... 3
       1.2.2 Seismic reflection method .................................................................... 5
   1.3 The present study .......................................................................................... 9

Chapter 2: Methodology .......................................................................................... 21
   2.1 Waveform tomography ................................................................................ 21
       2.1.1 Forward modeling .............................................................................. 21
       2.1.2 Inversion .......................................................................................... 23
   2.2 Source signature determination .................................................................... 29

Chapter 3: Waveform and travel time tomography ............................................... 31
   3.1 Introduction .................................................................................................. 31
   3.2 Method ......................................................................................................... 32
   3.3 Data analysis and results .............................................................................. 34
   3.4 The computing aspects of travel time and waveform tomography .............. 43
3.5 Discussion and conclusion................................................. 47

Chapter 4: Waveform tomography applied to the VSP dataset............... 50
4.1 Introduction................................................................. 50
4.2 The geological setting.................................................. 51
4.3 The experiment......................................................... 52
4.4 The VSP dataset and pre-processing.................................. 55
4.5 Travel time tomography applied to the VSP dataset................. 58
4.6 Waveform tomography.................................................. 61
4.7 Q value determination.................................................. 65
4.8 Waveform inversion..................................................... 68
4.9 Geological interpretation of the waveform tomography velocity model... 71
4.10 Resolution analysis.................................................... 74
4.11 Petrological interpretation............................................. 76
4.12 Comparison with other images....................................... 82
4.13 Conclusion............................................................... 85

Chapter 5: Waveform tomography applied to the surface reflection data set..... 87
5.1 Introduction............................................................... 87
5.2 The experiment.......................................................... 88
5.3 The dataset............................................................... 88
5.4 Travel time tomography................................................ 94
5.5 Waveform and travel time tomography.................................. 100
5.6 Waveform tomography.................................................. 102
5.6.1 Correlation with the VSP velocity model.......................... 119
5.6.2 Waveform inversion results ........................................... 121
5.6.3 Comparison with depth migrated images .......................... 133
5.7 3D geometry of the paleo-channel .................................... 134
5.8 Discussions and conclusions ........................................... 135
Chapter 6: Discussion and conclusion .................................. 137

References ........................................................................... 144
Appendix I ........................................................................... 150
LIST OF TABLES

Chapter 3

Table 1. Waveform and travel time tomography: computing aspects.................. 43

Chapter 5

Table 1. Travel time tomography for 45 2D seismic profiles......................... 95

LIST OF FIGURES

Chapter 3

Figure 1. The synthetic velocity model................................................. 33
Figure 2. A synthetic shot-gather.......................................................... 34
Figure 3. The synthetic model and travel time velocity model...................... 35
Figure 4. The velocity model from wideband waveform tomography.............. 36
Figure 5. The velocity model from narrowband waveform tomography........... 39
Figure 6. The cycle skip problem......................................................... 40
Figure 7. The velocity model by travel time and narrowband waveform tomography. 42
Figure 8. Salt-dome imaging by waveform tomography............................ 45

Chapter 4

Figure 1. Aerial photo of the Hill Air Force Base (HAFB).......................... 51
Figure 2. Map of clay top depths............................................................ 53
Figure 3. Source-receiver geometry of the VSP experiment........................ 54
Figure 4. A shot gather from the VSP dataset.......................................... 55
Figure 5. Statistics of picked arrival time differences............................... 58
Figure 6. Final velocity model and its associated ray paths ......................... 60
Figure 7. Frequency domain data ......................................................... 63
Figure 8. The relationship of amplitude versus offset .............................. 66
Figure 9. Determination of Q values ...................................................... 67
Figure 10. Estimated source signatures .................................................. 69
Figure 11. Observed and predicted seismograms in waveform inversion ........... 70
Figure 12. Final velocity model from waveform tomography ........................ 71
Figure 13. Geological interpretation of the waveform tomography velocity model.. 72
Figure 14. A synthetic test for the VSP experiment ................................... 75
Figure 15. Angular grains contacting each other ...................................... 78
Figure 16. Petrological interpretation ...................................................... 79
Figure 17. Two P- and S-wave velocity profiles measured in-situ .................. 81
Figure 18. Correlation between waveform tomography model and depth migration. 82
Figure 19. Comparison between waveform and travel time velocity models ...... 84

Chapter 5
Figure 1. Source receiver geometry for 45 2D seismic profiles .................... 89
Figure 2. Representative shot gathers for lines ........................................... 90
Figure 3. Ray coverages for travel time tomography .................................... 97
Figure 4. A synthetic test for the surface reflection experiment ...................... 101
Figure 5. Estimated source signatures for 45 profiles ................................. 103
Figure 6. Velocity models from travel time and waveform tomography .......... 112
Figure 7. Correlation of the VSP velocity model and surface reflection models..... 120
Figure 8. Observed and predicted seismograms for 45 seismic profiles .......... 122
Chapter 6

Figure 1. Wavefronts in 2D heterogeneous and averaged 1D models.................. 142
Chapter 1

Introduction

This thesis develops and applies seismic waveform tomography to solve the unique problem of imaging shallow-substructures with high resolution. Applications to datasets from a ground water contamination site at the Hill Air Force Base (HAFB), show that the waveform tomographic approach is a cost-effective and efficient seismic imaging tools for ground water remediation programs. Shallow sub-structure characterization has been the research topic of many investigators in the past several decades. It is associated with different problems, such as fracture detection (e.g., Tura et al., 1992; Zinni, 1995), buried waste detection (King et al., 1989), archaeological probing (Witten et al., 1992), earthquake hazard evaluation (Woolery et al., 1997), ground water mapping/remediation and coal mining (Gochioco, 1992) as well as gas/oil exploration.

As mentioned above, one of the shallow geo-technical problems is contamination site characterization and remediation. Contamination sites have been the targets of shallow sub-structure characterization to delineate the geometry of polluted aquifer systems or to characterize the distribution range of contaminants using different geophysical approaches, which are categorized as non-seismic and seismic techniques. Non-seismic techniques mainly include electrical resistivity log/resistivity spectrum, electromagnetic induction and ground-penetrating radar (GPR). Seismic techniques mainly include reflection, refraction imaging and seismic travel time tomography.
1.1 Non-seismic methods

The wide range (over seven orders of magnitude) of resistivity variation in rock has long made it an attractive physical parameter for subsurface exploration and description (Dobecki and Romig, 1985). In ground water investigations, where the desired targets are thick aquifers saturated with fresh water overlying a high-resistivity base, resistivity surveys are ideal for determining the thickness of the aquifer system and even a measure of water quality (Urish, 1983). In such investigations, resistivity data obtained in the field are interpreted qualitatively and/or quantitatively by inverting the observed data using theoretical physical models. For example, Benson et al. (1997) carried out electrical resistivity surveys to locate the gasoline that contaminates ground water in a parking lot in Utah county, Utah. Similarly, resistivity spectra were measured by Vanhala et al. (1992) to detect organic chemical contaminants such as toluene in clay minerals. More recently, in their efforts to delineate the 3D geometry of sludge disposal ponds at Lernacken, Sweden, 3D resistivity data were acquired by Dahlin et al. (2002) and inverted by 3D resistivity modeling for quantitative interpretation.

Electromagnetic techniques such as GPR and electromagnetic induction are less widely applied than resistivity methods because the latter are more sensitive to resistivity/conductivity than are the former methods, and also because electromagnetic data are difficult to interpret, according to Keller (1971). Furthermore, conductivity and saturation of soil or rock mass limit the penetration depth of GPR, thereby limiting its application. However, often with the aid of seismic reflection profiling, GPR has found substantial applications in ground-water remediation. For example, Carcione et al. (2000) mapped fuel spilled into ground using GPR at the Krzywa Air Base, Poland.
1.2 Seismic method

The technique developed and applied in this thesis study is a seismic approach to characterize shallow sub-structures to aid in ground water remediation. Seismic imaging techniques originally grew out of military purposes to locate enemy artillery in World War I, according to Dobecki and Romig (1985). It has been realized since the earliest days that seismic methods offer inherently greater resolution than potential field methods such as the resistivity approach discussed above. It is well known that high frequency content in propagating waves is essential for high resolution images in any wave-based method, but high frequencies attenuate faster than lower frequencies in most media. Seismic waves propagating in the earth have wave lengths short enough to “see” scale features we are interested in, yet are subject to attenuations low enough such that they propagate far enough into the earth for us to see broadband images. Therefore, seismic approaches often have the advantage of imaging larger areas with greater resolution than other geophysical techniques. This section presents a review of two most commonly applied seismic approaches: refraction, reflection method.

1.2.1 Seismic refraction method

Seismic refraction works well if velocity increases with depth and the substrate is composed of generally flat layers. The first step in this method is to pick first arrivals as a function of geophone positions which are usually linearly arranged. The second step is to fit straight lines to the plotted data to define different velocity layers. Velocities thus determined are apparent velocities and equal to the true velocities of different layers only if layer boundaries are horizontal. In the case of dipping layer boundaries, the line should
be shot at both ends to determine true velocity and dip angles of the layers using apparent velocities from two shots. The advantage of this method is that first arrivals are easy to identify and interpretation is straightforward. However, there are three pitfalls in the application of this approach (Dobecki and Romig, 1985): (1) Low velocity layers cannot be detected by refraction surveys because no refracted waves are generated from these layers. (2) A thin layer cannot be observed either because of triplication, i.e., head waves generated in the layer lag behind first arrivals from other layers. (3) The first arrival wavelet has finite band width, different from the assumption in interpretation that it is a delta function. Therefore uncertainty is introduced in refraction interpretation. Fine structures in layer boundaries are not observed partly due to this uncertainty, partly due to waveform healing. Because of these pitfalls and the fact that velocity structure in the target area is subject to large lateral variations, more sophisticated geophysical inversion methods such as refraction tomography are introduced and widely applied (e.g., Zelt and Barton, 1998).

In refraction tomography (or first arrival travel time tomography), first arrival times are picked instead of being fit by straight lines. The first arrival picks are theoretically predicted using an initial velocity model. The misfit which is the difference between the predicted and picked travel times is back-propagated into model space iteratively. The initial model is thus updated iteratively. The loop stops when the misfit is small enough or other criteria are satisfied. Sub-structure is characterized by the distribution of seismic velocity in the final model. This quantitative approach can characterize much more complicated structures than the conventional method mentioned above. The important assumption in this method is that the velocity structure in the subsurface varies smoothly
compared to the frequency of propagating waves such that travel times can be accurately computed by solving the Eikonal equation (e.g., Bleistein et al., 1995). If carefully implemented, this approach could characterize sub-structures in a resolution scaled with Fresnel zone and could provide valuable background velocity information for many applications. Lanz et al. (1998) applied refraction tomography to delineate the depth and geometry of buried wastes up to 40.0m depth. Azaria (2002) performed 3D refraction tomography so image of a paleo-channel where contaminating solvents were dumped was successfully mapped. However, low velocity layers cannot be imaged if only first arrivals are used even though isolated low velocity zone can. In that case, reflection tomography could play a role.

1.2.2 Seismic reflection method

The basic process in the reflection method is (1) to identify reflection events from layer boundaries, (2) to enhance these events by signal processing techniques, (3) to analyze seismic velocity based on these events and (4) to interpret or migrate them in the time or depth domain using the velocity information obtained by velocity analysis. By migration, sub-structures are delineated by their impedance contrasts. Migrated images are then interpreted geologically. The advantage of the reflection method over the refraction method is that low velocity layers and thin beds are not missed and resolution as high as a quarter of wavelength is possible. Another advantage is that with the same source-receiver offset, the reflection method can generally image deeper than the refraction method. It has been applied in ground water monitoring as well as petroleum exploration. For example, Birkelo et al. (1988) show that the top of a water table as shallow as 2.5m
deep can be mapped with proper equipment and field parameters, and even the drop of water table can be detected during a pumping test. Similarly Bachrach and Nur (1998) report that the water table variation in a California beach associated with tiding can be monitored, but the reflections do not correspond to the water table as defined as the phreatic surface; instead they are influenced by partial saturation. Dana (2003) mapped the 3D geometry of a polluted aquifer system at the Hill Air Force Base (HAFB), Utah, using 3D reflection data.

However, there are a number of factors limiting reflection image quality and therefore limit its applicability. The first factor is that the top most material in the earth attenuates high frequency energy more than low frequency, which tends to make the transmitted wavelengths too large for the scale features we are interested in. In fact, frequency contents of reflection waves are observed to never go beyond 600Hz if both sources and geophones are at the surface (Steeples, 1998). On the other hand, material in this part of the earth is very heterogeneous in terms of density, seismic velocity, mineral composition, and fluid content. Imaging such target features often require a wide frequency range. This factor is less limiting now than before because of technical improvements in source generating equipments, geophone and recording systems (Dobekci and Romig, 1985).

The second factor, which is more important, is the signal-to-noise ratio of reflections. Noises for reflections include background and natural cultural noise and signal-generated noise, such as Rayleigh and Love waves. In fact, the level of background cultural noise at our seismic experiment site at HAFB is unusually high. According to Levander et al. (2003), the principal noise sources included pumps operating as part of the remediation
effort, the local power grid, movement of vehicles and personnel associated with the remediation effort, and low-flying jet aircraft stationed at HAFB. Although we took what measures we could to reduce the noise levels (for example, not recording during takeoff and landing of jets), much of the seismic data have unavoidably high levels of background noise. The most effective noise reduction measure, shooting at night when wind and cultural noise sources were lower, was prohibited by HAFB as being unsafe.

Another significant type of noise is signal-generated; i.e., reflections may be mixed with other waves, such as refracted waves, diffracted/scattered waves and ground roll. This type of noise poses a bigger problem for the reflection method in shallow seismic characterization than deep seismic probing where reflections from deep structures could be naturally separated from refracted waves and ground roll. In shallow seismic characterization, shallow reflections are inevitably contaminated by signal-generated noises. Most commonly they are mixed with ground roll, which in many cases dominate the wave fields recorded at surface. Ground roll, which is surface waves formed by seismic energy trapped in surface layers, spread according to the first power of horizontal distances while reflection waves spread according to the square of the two-way slant distances. As a result, at typical recording distances, the amplitudes of surface waves are many times larger than that of reflection waves. This is especially true if low velocity layers exist because more energy is trapped in this layer. How to retrieve a relatively weaker signal (reflection) reliably from an overwhelming ground roll wave fields is a challenging problem. A few algorithms proposed so far such as Karhunen Loeve transform (Liu, 1999), wavelet transform (Deighan and Watts, 1997) and median filter
(Duncan and Beresford, 1995) attacked the problem with mixed successes, usually at the price of deformed reflection signals.

The signal-generated noise could pose an even worse problem for the reflection method if both lateral and vertical variations in sub-structures are severe. Reflection boundaries are less well defined, and diffracted/scattered waves resulting from small inhomogeneities can be additional noises for reflection events. The Common-Mid-Point (CMP) stacking and Normal Moveout (NMO) correction in standard reflection processing is based on the hyperbolic moveout assumption. When the sub-structure is complicated, the assumption is not valid. Velocity analysis based on the velocity spectrum method (e.g., semblance, see Yilmaz, 2001) in standard reflection processing could be biased. 3-D acquisition instead of 2-D acquisition, Pre-stack migration instead of after-stack migration may help to ease the problem, and velocity information for migration may be obtained from other approaches such as travel time tomography. Ray-based approaches such as travel time tomography provide very reliably velocity information for migration; however, provide little information about small inhomogeneities, i.e., diffractors and scatterers.

The challenge of shallow seismic characterization by the reflection method has long been realized in the geophysical community. In his comprehensive review of various approaches in engineering and environment geophysics, Steeples (2001) found that the seismic reflection method was not adopted for imaging sub-structures shallower than 30m until mid-1980s because of the difficulties. It has been applied occasionally for shallow imaging with mixed success since then. He further suggested that one means that seismic methods involving target depths smaller than 30m can be extended is to
analyze the wave types generally discarded by conventional seismic reflection investigators. This thesis study coincidentally made an effort in this direction by applying full wave field inversion which is totally not in the reflection seismology domain in terms of methodology.

1.3 The present study

Operable Unit-2 (OU-2) at the Hill Air Force Base is a ground water contamination site, where relatively dense ($\rho > \rho_{\text{H}_2\text{O}}$) chlorinated solvents known as Dense Non-aqueous Phase Liquids (DNAPLs) have been introduced into ground. The polluted aquifer system is limited to the uppermost 15-20m because of a clay layer which is an aquiclude. The natural ground water flow is down slope to the Weber river in the nearby valley through a natural drainage path that formed a distinct paleo-channel about 15m wide and 15m deep. The lateral heterogeneity of the infilling sediments in the paleo-channel has been increased by nearby land slumping and other mass waste events. To prevent the flow of contaminants the Air Force surrounded the OU-2 site with a subsurface containment wall made of bentonite clay.

This thesis study characterizes the shallow sub-structure by estimating the distribution of seismic P-wave velocity. It has been long recognized that the earth’s top most 20m features large vertical and lateral variations in seismic velocity. The P-wave velocity could go from 50m/s (Bachrach et al., 1998), which is much smaller than P-wave velocity in air (i.e., sound speed, 340m/s) to 1500m/s (acoustic velocity in water) within 20 m depth. The shallow seismic velocity structure is complicated because it is affected by many factors such as mineral composition, degree of cementation and consolidation,
porosity, pore fluid saturation, permeability and confining pressure, etc, and is also the 
locus of weathering processes and rapid fluid transport. The complicated velocity 
structure can be measured directly by ultra-sonic logs with high resolution (Liu et al., 
1997); however, it is expensive to drill a borehole and to make such measurements. 
Furthermore, such measurements can be made only on a point basis; i.e., lateral 
variations can only be interpolated from point measurements. That is why in-situ 
measurements are very few. It has been realized that even closely spaced borehole 
information could never duplicate the detail of the subsurface images using seismic data 
(Francese et al., 2002).

Seismic velocity structures can also be reconstructed by travel time tomography. 
However, typical travel time tomography has a vertical resolution from 5-10m depending 
on source-receiver geometry. The resolution is not adequate to resolve the complicated 
structure within 20m, although the method is very useful in characterizing large 
geological features and providing background velocity information.

Knowledge about shallow seismic velocity structures in the petro-physical sense is 
relatively limited so far, partly because of the complexity and partly because laboratory 
measurements are very hard to conduct at low pressure (<0.1MPa) due to transducer-
material (e.g., sand) coupling problems (Bachrach et al., 2000). Most research efforts to 
understand acoustic properties of shallow unconsolidated material are conducted by 
analyzing field seismic data (e.g., Baker et al., 1999; Bachrach and Nur, 1998).

This thesis study attempts to solve the problem of how shallow seismic velocity 
structure can be estimated reliably with high resolution in two-dimensions. This problem 
is significant for (1) the methodology of seismic imaging itself; (2) petro-physical study
of the shallowest part of the earth; (3) real applications of shallow seismic characterization, in this case at the OU-2 ground water contaminations site. This methodology also is potentially useful in waste detection, archaeological probing and applications in gas/oil exploration. The key to solving this problem lies in seismograms themselves. Seismic waveforms contain much information that is ignored by many existing seismic imaging approaches. This study shows through synthetic as well as real data applications at the OU-2 ground water remediation site that the problem can be solved by waveform tomography.

Waveform tomography is a geophysical inversion technique where waveforms instead of travel times are inverted to characterize sub-structures in terms of seismic properties such as seismic velocity, density and attenuation. It has been developed in the context of diffraction tomography (DT), where diffracted/scattered waves are inverted to map inhomogeneities smaller in size than the wavelengths of incident waves. In a comparison with holography, Wu and Toksoz (1987) tested a DT algorithm linearized by the Born approximation in different source receiver geometries and concluded that diffractors could be mapped with a resolution up to half a wavelength. Devaney (1982, 1984, 1989) discussed differences between geophysical diffraction tomography and X-ray tomography and difficulties regarding DT in non-uniform backgrounds. We (Gao et al., 2001) proposed a new algorithm by which the forbidding task of filter computation with respect to non-uniform backgrounds are avoided by an iterative method. While diffraction tomography has the advantage of linearized inversion, it is generally difficult to isolate diffraction/scattering waves from total wave fields.
On the other hand, full waveform inversion has been given more and more attention in the last decade or so because of the progresses in seismic forward modeling, increased speed in computer hardware and breakthroughs in inversion techniques. Full waveform inversion was not thought of as being practical by some researchers because it is too expensive to forward model seismic waves and to compute explicitly the Frechet kernel; the Frechet kernel is defined as the derivative of the seismic wave field with respect to model parameters. Lailly (1983) and Tarantola (1984) took an important step by realizing that models could be updated iteratively by back-propagating the data residuals and correlating the result with forward-propagated wave fields. The advantage of the correlation method is that only the gradient vector of misfit function with respect to model parameters is computed at each forward modeling step, avoiding direct computation of Frechet kernel which requires many forward solutions if it is computed explicitly. This is a breakthrough toward a practically useful algorithm for waveform inversion. This step was followed by many efforts to forward model seismic waves in an efficient way because performance of inverse procedure depends critically on the technique used for solving forward problems.

Gauthier et al. (1986) and Mora (1987) formulated their algorithms of seismic forward modeling in time and space domain, and Pratt and Worthington (1990), Pratt (1990), Geller and Hara (1993) formulated it in the frequency and space domain. Gauthier et al. (1986) pointed out that ray tracing is not adequate to characterize wave propagation in a complicated media, because of the high frequency approximation. On the other hand, numerical methods based on finite difference (or finite element) solutions to wave equation are able to synthesize full wave fields even in complicated media with errors
adaptively controlled. The authors forward modeled acoustic waves by finite difference methods in time and space domain, followed by waveform inversion. Mora (1987) did elastic waveform inversion in a similar methodology as in Gauthier et al. (1986).

While forward modeling in the time-space and the frequency-space domains are analytically equivalent, the advantage of forward modeling in frequency-space domain attracted a number of researchers' attention due to computational advantages. Marfurt (1984) first pointed out that forward modeling in frequency and space domain is the method of choice if larger number of source locations are involved. Pratt (1990) later pointed out that data from large aperture seismic surveys could be inverted effectively using only a limited number of frequency components, thus reducing the number of frequency-domain forward solutions in a way not possible with time domain methods. These advantages are added to other computational advantages such as the ease with which viscous attenuation and dispersion are incorporated into frequency domain methods and the ease with which inverse methods can be implemented to use the lowest data frequencies first, thus mitigating severe non-linearities in the waveform inversion problem (Pratt, 1999).

This study improves upon and applies the waveform inversion technique developed by Pratt (1999) to the dataset from the HAFB seismic experiment to obtain a high resolution images from the ground water contamination site. Although the waveform inversion technique in the frequency domain has many advantages, it also suffers from a shortcoming—cycle skipping. It can be understood conceptually in the context of phase ambiguity problem. In frequency domain, digital seismograms are expressed as complex numbers which consist of amplitudes and phases. It is well known that if phases are
added by multiples of $2\pi$, the resulting complex numbers remain the same, so-called phase ambiguity. But in time domain, cycles are skipped. The cycle skip problem manifests in waveform inversion techniques such as waveform tomography through the requirement that the initial velocity models be well known. How initial velocity models which meet the requirement of waveform inversion can be obtained is a problem. As we know, there are more approaches by which seismic velocity models are obtained. One of them is seismic travel time tomography which has been widely applied in exploration as well as earthquake seismology. What are the relative performances of travel time tomography and waveform tomography? Are velocity models from travel time tomography good enough for waveform tomography? What are the different roles the two methods play in a real seismic exploration? These questions and the cycle skip problem are answered and discussed in Chapter 3 of this thesis.

The next problem to solve is that the data we acquired from the HAFB seismic experiment are elastic waveforms whereas only acoustic waveforms are modeled by the waveform tomography technique of Pratt (1999). Mora (1987) developed and synthetically tested an elastic waveform inversion technique which is powerful because density, P-wave velocity and S-wave velocity structures can be determined simultaneously. Mora (1987) tested the technique using a relatively simple diffraction model and Mora (1988) further tested the technique using both reflected and transmitted waves. According to Mora (1987), the shortcomings of elastic waveform inversion are (1) large cost, and (2) requiring initial P-/ S-wave velocity models and density models because of the highly non-linear nature of waveform inversion. The cost for solving the two coupled P-Sv equations at least doubles that for the acoustic equation. Furthermore,
most seismic surveys only record single component data while three-component data are required to implement the elastic waveform inversion with full capacity. However, the elastic waveform inversion could be applied to single-component data by assuming \textit{a priori} relationships between P-wave velocity and density, and between P- and S-wave velocities themselves (Shipp and Singh, 2002). What is really determined is actually the P-wave velocity model. So far the technique hasn’t been applied in shallow sub-structure characterization.

This study adopts the idea of treating recorded elastic P-waves as acoustic waves which are then forward modeled by solving the acoustic equation. This approximation makes the waveform tomography much less expensive so that it is practically more useful and attractive, while it retains the resolving capacity of waveform inversion. This is a valid approximation because (1) the active seismic sources in most exploration experiments are P-wave sources, i.e., except for surface waves, the shear waves generated by such sources are minimum. Such sources include dynamite explosion, sledge hammer and rifles. In our case, the seismic source is a 0.223 caliber rifle. (2) The dominating wave phenomenon is P-type waves following first breaks in seismograms. How this approximation is applied will be discussed in full length in Chapter 4.

However, without reliable source signatures being estimated first, the observed waveforms still cannot be inverted because the shapes of forward modeled waveforms highly depend on source signatures. A seismic signature (or seismic source wavelet) is important not only in waveform inversion, but also in processing data to, for example, increase the resolution of seismic data by signature deconvolution, deghosting and
dereverberation. Therefore, how to estimate reliable source signatures has been the research topic of many researchers in the last several decades (e.g., Wold, 1938).

Methodologically two categories of approaches have been developed to estimate the source signatures: the statistical methods and the deterministic methods. The former category, based on the original stationary time series model of Wold (1938) was initiated and developed by Robinson (1954), Treitel and Robinson (1967) and Peacock and Treitel (1969). By these methods, source signatures are estimated from the seismic data. The underlying assumption is that the data in a given time window can be expressed as the convolution of a wavelet with the impulse response of the earth structure. Most existing methods of source estimation are in this category and have been applied successfully in seismic explorations. However, the underlying assumption is valid to different extents from case to case. Ziolkowski (1991) pointed out that the signatures are better estimated in a deterministic way; i.e., signatures can be measured directly near a source instead of being estimated by making assumptions on the statistical properties of seismograms. So far, methods to measure Vibroseis and air gun sources were proposed and tested (Ziolkowski, 1991) and the method for dynamite sources was proposed by the same author but not tested. The seismic source in our experiment is a .223 caliber rifle, and all seismograms were recorded within 40.0m offset. These relatively near-offset data are all ideal for source estimation if noise level is low. In this study, we adopt the non-statistical approach of Pratt (1999) to invert source signatures by minimizing a misfit function. The source signatures are updated in each inversion loop after the velocity model is updated. The advantage of the approach is that the source signatures determined in this way are not sensitive to errors in the velocity model, as is shown in Pratt (1999). Details of how
source signatures are determined and the resulting source signatures by a shot rifle are in Chapter 4.

The main purpose of waveform inversion is to obtain a high resolution velocity model. However, seismic velocity is not the only physical parameter by which materials can be characterized. One of the other physical parameters is attenuation, which is quantified by Q values. The larger the Q value is, the less seismic waves are attenuated. Intrinsic attenuation is a function of compaction, density, seismic velocity, pressure and temperature. Therefore map of Q values, together with other geological information, can help us to identify different rock types and even thermal structures. Furthermore, it is important to account for seismic attenuation in waveform inversion for two reasons: (1) amplitude decay. In waveform inversion, both the shape and amplitude of waveforms are modeled. The amplitude is controlled by geometrical spreading and attenuation. Therefore, forward modeling has to incorporate Q values to match the observed waveform better. In Chapter 2, we can see how Q values are incorporated into the forward modeling. (2) Seismic waves travel in different velocity in media with different attenuation property.

The next question is how to determine Q values. Q values for different rock types under different confining pressure, pore-fluid saturation and temperature can be measured in the laboratory (e.g., Winkler and Murphy, 1995). In the field, Q values have been determined for different areas in the world using different methods (e.g., Der et al., 1984; Snieder, 1987; Xie and Nuttli, 1988). Most of these methods determine Q values by modeling spectral decay in the frequency domain as a function of distance from the seismic source. In our case, since Q values can be incorporated into the finite difference
forward modeling, Q values are determined by matching the relationship of amplitude versus offsets for synthetic seismograms to that for the observed ones. That is basically a trial-and-error approach. Details of Q value determination will be discussed in Chapters 4 and 5.

With high resolution in the velocity model, how reliable the detailed structures in the model are imaged will be next question to be addressed. The question is equivalent to asking how well can we believe the details in the model? As a cautious check, this study compares the velocity model from waveform tomography to other images of substructures including (1) depth migrated 2D surface data from the VSP experiment, (2) the depth migrated 3D dataset from 3D reflection experiment (Dana, 2003), and (3) velocity model from 3D travel time tomography (Azaria, 2002). These comparison and/or correlation will help to bring the geological structure to light and indirectly answer the question of how reliable the waveform tomography images are. All the comparisons are made in Chapter 4.

After the shallow seismic velocity structure is imaged with high resolution, it is still at least one step away from obtaining a geological sub-structure map which can be used directly for the ground water remediation purpose. Realizing that what we have is a distribution of one physical parameter—seismic velocity and that the velocity structure and geological structure are well correlated but correlation is not unique because velocity can be also affected by factors such as fluid saturation in addition to rock type, we incorporate as much other information as available to make a geologically sensible interpretation of the velocity model. In our case, we know the geological setting of the target area and lithology logs from the two boreholes from our VSP experiment. It is
natural for us to correlate the lithology logs with the velocity model. The correlation serves two purposes: (1) to see how good the velocity model correlates with borehole lithology logs; If the velocity model is well resolved, the correlation should be good. (2) if the correlation is reasonably good, it can be used as a calibration for geological interpretation for the areas within the two boreholes. Geological interpretation is covered in full in Chapter 4.

The velocity model can also be interpreted from a petro-physical perspective because bulk and shear modulus of unconsolidated material can be evaluated by the Hertz-Mindlin theory and effective P-wave velocity can be computed using the Gassman equation (e.g., Bachrach et al., 2000). Since we know some of the parameters used in the Hertz-Mindlin theory such as mineral composition of rocks in the target area, it is possible to estimate a few parameters by theoretically predicting the observed velocity profile. Since seismic velocity for unconsolidated material under low confining pressure is difficult, if not impossible, to measure directly in the laboratory, this study may pave the way for new research direction in rock physics. Details of the interpretation can be seen in Chapter 4.

As in many tomographic studies, resolution analysis is necessary to see how truly the image reveals sub-structures. Generally, the resolution of waveform inversion scales with wavelength, but resolution as a function of space coordinates varies with source-receiver geometry, source and receiver intervals. This study evaluates the resolution through a synthetic test. In the test, a synthetic model is set up with exactly the same experiment configuration as that in a real experiment. The synthetic model is designed such that different scale features are mixed in the same model. Then synthetic seismograms are
computed and travel time tomography is first performed. Then the synthetic seismograms are treated as the observed waveforms which are inverted to see how the true model is reconstructed by waveform tomography.

In Chapter 5, this thesis shows that the 3D geometry of a buried channel is reconstructed with relatively high resolution by waveform tomography. The 3D geometry is reconstructed by combining 45 parallel slices of cross-sectional geometry of the channel. Each cross-sectional geometry is identified in a velocity model from waveform tomography applied to a 2D seismic profile, in the same strategy as it is applied to the VSP dataset. Some of the velocity models are correlated with depth migrated images from Dana (2003).

In summary, this thesis applies waveform tomography to solve the unique problem of shallow sub-structure imaging with high resolution. The problem is unique because of the high degree of heterogeneity in terms of geological structure, low confining pressure and high degree of un-consolidation. Different aspects of how the problem is solved are discussed. Chapter 2 summarizes the theoretical background of the method used in this thesis. Chapter 3 shows a comprehensive comparison between travel time and waveform tomography. Chapter 4 presents details of the seismic experiment in a ground water contamination site at HAFB and how the technique of waveform tomography is applied to real data. Geological and petrological interpretation of the resulting velocity model, its correlation with other images and resolutions analysis are also in Chapter 4. Chapter 5 shows the application of waveform tomography to reconstruct the 3D geometry of a buried paleo-channel where DNAPLs were dumped, followed by discussions and conclusions in Chapter 6.
Chapter 2

Methodology

This chapter presents the theoretical background of the methods developed and applied in this thesis. Following Pratt et al. (1998), section 1 introduces the method for waveform tomography, and discusses various technical aspects of it. Section 2 discusses the method for source determination (Pratt et al., 1998).

2.1. Waveform tomography

Following Pratt et al. (1998), the method for waveform tomography is described as the following:

2.1.1 Forward modeling

Forward modeling is the most important step in any inversion techniques. In this technique, seismic wave fields are forward modeled by solving the acoustic wave equation which can be expressed in discrete matrix format using finite difference as

\[ M\ddot{u}(t) + C\dot{u}(t) + K\dot{K}(t) = f(t) \] (1),

where dot denotes derivative with respect to time \( t \), i.e., double dots represent second derivatives and a single dot denotes first order derivative. \( C \) is the damping matrix, \( M \) is the mass matrix and \( K \) is the stiffness matrix. The mass, stiffness and damping matrices are computed by forming a discrete representation of the underlying (spatial) partial differential equations and the physical parameters (e.g., the seismic velocities, the bulk density and the attenuation parameters). \( u(t) \) and \( f(t) \) are vectors representing the wave
field to be determined and the source signature. The convention that bold low case characters denote vectors and capitals denote matrices is followed here.

Equation (1) is expressed in time-space domain. We now choose to solve the equation in frequency-space domain. This has distinct computational advantages for multi-source problems (Marfurt, 1984; Pratt, 1990). We will also see the advantage of a frequency-space domain solution in the inversion formulation. Fourier transforming with respect to time \( t \) to both sides of Equation (1) yields,

\[
Ku(\omega) + i\omega Cu(\omega) - \omega^2 Mu(\omega) = f(\omega)
\]

(2),

where \( u(\omega), f(\omega) \) are the wave field and source signature in frequency domain,

\[
u(\omega) = \int_{-\infty}^{\infty} \tilde{u}(t)e^{-i\omega t} dt \quad \text{and} \quad f(\omega) = \int_{-\infty}^{\infty} \tilde{f}(t)e^{-i\omega t} dt.
\]

For simplicity Equation (2) can be rewritten in matrix format as

\[
Su = f
\]

(3),

where \( S = K - \omega^2 M + i\omega C \) is the complex impedance matrix. Equation (3) can be solved as

\[
u = S^{-1}f
\]

(4).

In practice, it is not possible or desirable to actually invert the very large matrix \( S \). What is commonly done to solve Equation (3) is to use matrix factorization methods, such as LU decomposition (Golub and Van Loan, 1983). Once the matrix \( S \) is decomposed as \( S = LU \), where \( L \) is a lower triangle matrix and \( U \) is an upper triangle matrix, \( L \) and \( U \) can be repeatedly used for any other sources \( f \). This is very important for any iterative solution to inverse problems, in which many forward solutions are needed for real sources as well as virtual sources.
If the source signature is a Kronecker delta function $\delta_y$, it is easy to see that column vectors in $S^{-1}$ contain the discrete approximations to the Green's functions for impulses at different nodes. Approximate Green's functions are constructed in different ways for different seismic imaging techniques. For example, Green's functions can be constructed approximately by dynamic ray-tracing under the high frequency assumption. Green's functions here are computed implicitly by solving the acoustic wave equation in frequency-space domain, and are therefore band-limited to low-frequencies as determined by the spatial sampling employed.

2.1.2 Inversion

In this section, the methods for solving the non-linear seismic waveform inversion problem using local methods, such as the gradient method are summarized. The problem can also be solved by a Gauss-Newton or full Newton methods. Details of the two approaches can be found in Pratt et al. (1998). By posing the inverse problem in the frequency domain, we can obtain specific algorithms for Frequency Domain Inversion (FDI).

Suppose we have $n$ experimental observations, $d$, at a subset of nodal points corresponding to receiver locations. Assume we have a total $l$ nodal points and $n \leq l$. The object of the inverse algorithm is to infer a set of model parameters, $p$, that predicts the observations using forward modeling algorithm described above. Assume that we have a suitable initial velocity model, $p^{(0)}$, that is close enough to the global solution to allow successive relinearization. Given the initial model, we can calculate the response by $u = S^{-1}f$. 
The residual error at the \( n \) receiver nodal points, \( \delta \mathbf{d} \), is defined as the difference between the initial model response and the observed data at the receiver locations. Thus

\[
\delta d_i = u_i - d_i, \quad i = 1, 2, \ldots, n
\]  

(5),

where the subscripted quantities are the individual components of the corresponding vectors \( \delta \mathbf{d} \), \( \mathbf{u} \) and \( \mathbf{d} \). The subscript \( i \) is the receiver index.

The inverse algorithm is to minimize the misfit function which is defined as the \( l_2 \) norm of the data residual

\[
E(p) = \frac{1}{2} \delta \mathbf{d}^T \delta \mathbf{d}^*,
\]

(6),

where the superscript \( ^T \) denotes matrix transpose and \( ^* \) denotes complex conjugate, which is introduced to make the misfit function real valued.

By using gradient method, the misfit function is minimized by

\[
p^{(k+1)} = p^{(k)} - \alpha^{(k)} \nabla_p E^{(k)}
\]

(7),

where \( k \) is an iteration number, and \( \alpha \) is a step length chosen to minimize the misfit function in the direction given by the gradient of \( E(p) \) which represents the direction in which the misfit function is increasing most rapidly. The misfit function can always be reduced by pursuing the negative of that direction. The iteration in Equation (7) is performed until some criteria is reached, for example, the inversion will stop if the misfit function is smaller than a small positive number chosen beforehand.

The gradient direction is evaluated by taking partial derivative of Equation (6) with respect to the model parameters, \( p \),

\[
\nabla_p E = \frac{\partial E}{\partial p} = \Re \left[ \frac{\partial \mathbf{u}}{\partial p} \right]^T \delta \mathbf{d}^* = \Re \{ \mathbf{J}' \delta \mathbf{d}^* \}
\]

(8),
where \( J_q = \frac{\partial u_i}{\partial p_j} \), \( i = (1,2,\ldots,n) \); \( j = (1,2,\ldots,m) \), is the so-called Frechet kernel, an \( m \times n \) derivative matrix. Here \( m \) is the total number of model parameters and \( \mathbf{p} \) and \( \nabla_p E \) are column vectors of length \( m \). All vectors are in bold characters here and are assumed to be column vectors unless otherwise stated. Since the gradient of the real-valued misfit function with respect to real-valued model parameter is also real-valued, only the real part of \( \mathbf{J}^T \delta \mathbf{d}^* \) is taken.

For linear forward problems, the step length \( \alpha \) is computed as

\[
\alpha^{(k)} = \frac{|\nabla_p E|^2}{|\nabla_p E|^2} \tag{9}
\]

where \( | \cdot | \) represents the geometrical length of the vectors. For non-linear forward problems (such as the seismic problem), the step length must be found using line search techniques along the direction opposite to the gradient. Basically Equation (9) can be understood by noting that the step length is taken as the geometrical length of the gradient vector scaled by sensitivity which is measured in the Frechet kernel. If perturbation in the wave field is very sensitive to the perturbation in the velocity model, a smaller step length is taken than otherwise to avoid any dramatic and potentially erroneous effects in inversion.

From Equation (8), we can see that the gradient direction can be computed easily if the Frechet kernel is known. However, it is a computationally forbidding task to compute the Frechet kernel directly. For example, \( m \times n \) forward solutions are needed to compute the \( m \times n \) elements of \( \mathbf{J} \), which is too expensive to pursue. Lailly (1983) and Tarantola (1984) made the important realization that the steepest descent direction (negative
direction of $\nabla_p E$) could be computed without computing the Frechet kernel explicitly. They showed how the gradient of the misfit function could be computed by backpropagating the data residuals and correlating the results with forward-propagated wave fields. The derivation is summarized as following.

In order to explicitly link the computation of the gradient vector to the forward problem given in Equation (5), the $m \times n$ Frechet kernel matrix is augmented into $m \times l$, where $l$ is the total number of model parameters. Then we can write Equation (8) in a new format as

$$\nabla_p E = \text{Re}\{\hat{J}' \delta \hat{d}'\}$$

(10),

where $\hat{J}$ is the augmented $m \times l$ matrix and $\delta \hat{d}$ is the data residual vector of length $l$, obtained by augmenting $\delta d$ with $l - n$ zeros. Explicitly Equation (10) can be written as

$$\begin{bmatrix}
\frac{\partial E}{\partial p_1} \\
\frac{\partial E}{\partial p_2} \\
\vdots \\
\frac{\partial E}{\partial p_l}
\end{bmatrix} = \text{Re}\begin{bmatrix}
\frac{\partial u_1}{\partial p_1} & \frac{\partial u_1}{\partial p_2} & \cdots & \frac{\partial u_1}{\partial p_m} \\
\frac{\partial u_2}{\partial p_1} & \frac{\partial u_2}{\partial p_2} & \cdots & \frac{\partial u_2}{\partial p_m} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial u_l}{\partial p_1} & \frac{\partial u_l}{\partial p_2} & \cdots & \frac{\partial u_l}{\partial p_m}
\end{bmatrix} \begin{bmatrix}
\delta d_1^* \\
\delta d_2^* \\
\vdots \\
\delta d_n^*
\end{bmatrix} = \text{Re}\begin{bmatrix}
\frac{\partial u}{\partial p_1} \\
\frac{\partial u}{\partial p_2} \\
\vdots \\
\frac{\partial u}{\partial p_m}
\end{bmatrix} \begin{bmatrix}
\delta d_1^* \\
\delta d_2^* \\
\vdots \\
\delta d_n^*
\end{bmatrix}$$

(11).

An expression of partial derivatives in Equation (11) in terms of the forward modeling matrix in Equation (4) can be obtained by taking partial derivatives of both sides of Equation (4) with respect to $i$-th model parameter $p_i$:

$$S \frac{\partial u}{\partial p_i} = -\frac{\partial S}{\partial p_i} u \quad \text{or} \quad \frac{\partial u}{\partial p_i} = S^{-1} f^{(i)}$$

(12),

where the $i$-th "virtual" source term
\[ f^{(i)} = -\frac{\partial S}{\partial p_i} u \]  

(13),

is a vector of length \( I \). The partial derivative expressed in Equation (12) is a new forward modeling problem driven by the virtual source term shown in Equation (13). The virtual source is at the location of \( i \)-th model parameter. The virtual source represents the interaction of the background wave field, \( u \) with the model parameter \( p_i \). In this case, the computation is directly comparable to the Born method for computing wave field perturbations due to first order scattering.

The computation of the partial derivatives of the impedance matrix, \( \frac{\partial S}{\partial p_i} \), required in the computation of the virtual source in Equation (13), depends on specific details of the numeric set-up in the finite difference algorithm. It is usually trivial to compute.

Now the augmented Frechet kernel matrix can be expressed as

\[
\hat{J} = \begin{bmatrix}
\frac{\partial u}{\partial p_1} & \frac{\partial u}{\partial p_2} & \cdots & \frac{\partial u}{\partial p_m}
\end{bmatrix} = S^{-1}[f^{(1)} f^{(2)} \cdots f^{(m)}] \quad \text{or} \quad \hat{J} = S^{-1}F
\]  

(14),

where \( F \) is an \( I \times m \) matrix, the columns of which are the virtual source terms for each of the \( m \) physical parameters. Equation (14) gives an explicit formula for the Frechet kernel matrix. Computation of the elements of \( \hat{J} \) using Equation (14) requires \( m \) forward modeling problems to be solved. However, our ultimate goal is to compute the gradient vector and the Frechet kernel matrix need not be computed explicitly to obtain the gradient vector. Substituting Equation (14) into Equation (10), we can obtain

\[
\nabla_p E = \text{Re}\{\delta \hat{d}^*\} = \text{Re}\left\{F^t S^{-1} \delta \hat{d}^*\right\} = \text{Re}\left\{F^t v\right\}
\]  

(15),
where \( \mathbf{v} = [S^{-1}]^t \mathbf{d}^* \) is the back-propagated wave field. Expressed in Equation (15) is the so-called back-propagation method to compute the gradient vector. We can take a look at the \( i \)-th component in Equation (15) to see what it implies,

\[
(\nabla_p \mathbf{E})_i = \text{Re} \left\{ \mathbf{f}^t (i) \mathbf{v} \right\} = \text{Re} \left\{ \mathbf{u}^t \left[ \frac{\partial S}{\partial p_i} \right] \mathbf{v} \right\}
\]

(16).

From Equation (16), it is evident that the gradient can be computed as the scaled multiplication of the forward modeled wave field \( \mathbf{u} \) and the back-propagated residual wave field \( \mathbf{v} \). The scale factor \( \frac{\partial S}{\partial p_i} \) consists of only highly local non-zero values at, or near, the diagonal of the \( i \)-th row. The multiplication of the two wave fields in the frequency domain is equivalent to correlation in time domain. This is closely related to pre-stack reverse time migration and to the imaging condition of Claerbout (1976), except that it is done in frequency domain. Computing the gradient vector by the approach expressed in Equation (16) has a great advantage that only two forward solutions are needed to obtain the vector, instead of \( m \) forward solutions if the Frechet kernel is computed explicitly. Furthermore, the multiplication of two wave fields in frequency domain is trivial and efficient.

In summary, three forward modeling problems need to be solved in one iteration of an inversion using the gradient method. By solving the first forward modeling problem, the background wave field \( \mathbf{u} \) is obtained. The back-propagated wave field \( \mathbf{v} \) is obtained by solving the second forward modeling problem using residual waveforms as virtual sources. After the two forward problems have been solved, the gradient vector is evaluated. Finally step length \( \alpha \) is computed by solving the third forward modeling
problem using the computed gradient vector as a virtual source. Once the first forward solution \( \mathbf{u} \) is ready and hence the LU factors of the impedance matrix \( S \), the other two solutions can be rapidly computed using the computed LU factors. This is the advantage of forward modeling in frequency domain. Furthermore, it is trivial to correlate the forward modeled wave field \( \mathbf{u} \) and the back-propagated wave field \( \mathbf{v} \) in frequency domain, which is simply multiplication. This is another advantage to formulate the inversion in frequency domain.

### 2.2 Source signature determination

Following Pratt (1999), the algorithm for determining source signatures is summarized here. Given an initial velocity model and an initial source wavelet \( f \), we can compute the wave field \( \mathbf{u} \) by placing the initial source wavelet at the same location of a real source. Now we can modify the source signature by a complex-value scalar \( \lambda \) so

\[
\mathbf{u}' = \lambda S^{-1} f = \lambda \mathbf{u}
\]

(17),

where \( S \) is the impedance matrix. Now we want to find a suitable complex scalar \( \lambda \) which minimizes the misfit function defined in Equation (6). In order to minimize \( E(\mathbf{p}) \), set

\[
\frac{\partial E(\mathbf{p})}{\partial \lambda} = \frac{1}{2} \frac{\partial (\lambda \mathbf{u} - \mathbf{d})^t (\lambda \mathbf{u} - \mathbf{d})^*}{\partial \lambda} = \mathbf{u}'^t (\lambda \mathbf{u} - \mathbf{d})^* = 0
\]

(18).

It is obvious that

\[
\lambda^* = \frac{\mathbf{u}'^t \mathbf{d}^*}{\mathbf{u}'^t \mathbf{u}^*}
\]

(19),

minimizes the misfit function defined in Equation (6). Note that this complex-value scalar modifies both the amplitude and phase of the source signature for certain
frequencies. To evaluate this expression, we need one forward solution per source and per frequency. This algorithm is independent of the starting model for the source wavelet and it converges in one iteration. In deriving the solution above, it is assumed that the source signature was the only unknown, i.e., the velocity model is correct. Pratt (1999) showed that for the cross-hole problem, a good representation of the source signature can be obtained when the velocity model is only approximate.
Chapter 3

Waveform and Travel Time Tomography

3.1 Introduction

Seismic ray-based tomography has been widely applied in both exploration and earthquake seismology. It has been proven to be an effective seismic imaging tool in characterizing sub-structures in global (e.g., Grand et al., 1997), crustal (e.g., Zelt and Barton, 1998) as well as shallow/near surface (e.g., Azaria, 2002; Lanz et al., 1998) settings. More recently, waveform tomography (Pratt et al., 1998) has emerged, a computational considerably greater burden than ray based travel time tomography methods. In waveform tomography, wave-theoretical methods and direct arrival waveforms are used. So far, waveform tomography has not been applied as widely as ray-based tomography. In his study about the resolution limits in ray tomography, Williamson (1991) concluded that reconstructions based on ray tracing have a reduced resolution that scales with the Fresnel zone, (i.e., resolution \( \propto [\lambda L]^{0.5} \), where \( L \) is path length and \( \lambda \) is wavelength), while the superior resolution of waveform tomography scales with wavelength, (i.e., resolution \( \propto \lambda \)), and is theoretically independent of path length. What is common to both techniques is that sub-structures are imaged in terms of seismic velocity. Therefore it is natural for us to ask the questions such as (1) what their relative advantages and performances are; (2) what their different roles are in seismic imaging efforts; (3) what hampers waveform tomography from being widely applied so far. We (Pratt et al., 2002) addressed the first two questions previously using synthetic checkerboard models. This chapter addresses the first two questions by synthetic tests
using a more realistic model which is randomly generated with certain scale features. The model is called the random model hereafter. One of the hampering factors of the waveform tomography technique by Pratt et al. (1998) is that the original Fortran77 code can only handle relatively small models because static arrays in F77 can only be allocated in limited size while survey areas could be very large. This study modifies the Fortran77 code into a Fortran90 version which allocates arrays dynamically. The new version can handle relatively large models.

3.2 Method

We simulate wave propagation in the random model by solving the acoustic wave equation in frequency and space domain (see Chapter 2). The source-receiver geometry in the synthetic test is cross-well. In the 300m-deep-borehole on the left of the model are 101 sources with a uniform 3.0m interval. In the 300m-deep-borehole on the right are 101 receivers with the same uniform 3.0m interval. The two boreholes are separated by 150.0m (See Figure 1). The velocity model is numerically represented by seismic velocity values on a net of grids with a homogeneous 0.75m grid spacing, i.e., the model has 400 by 200 grids. Velocity varies from 1950.0m/s to 3150.0m/s with average at 2620.9m/s (see Figure 1). The velocity perturbations are strong enough to create discernable travel time anomalies and to create significant non-linear effects. Spectral analysis shows that along the horizontal direction, peak wavenumbers are at 37.5m and 10.0m wavelengths while along vertical direction, peak frequencies are at 100.0m, 21.4m, 20.0m, 14.3m, 12.0m, 10.3m, 9.1m and 8.1m wavelengths. It means that velocity structure along horizontal direction is simpler than that along vertical direction. This
study uses a synthetic Gaussian derivative source with frequencies between 50Hz and 950Hz and a dominant frequency at 300Hz. So the average wavelength is 8.5m by which most scale features in the synthetic model can be "seen" with enough resolution. It is helpful to know that the average Fresnel zone width is
\[
\sqrt[2]{\frac{\lambda}{3 \lambda}} = 35.7 m,
\]
where \( l = 150 m \) is the wave propagation length which is taken approximately the same as the width of the model.

Figure 1. The cross-well source-receiver geometry (left) in this study and the synthetic velocity model (right) which is randomly generated.
101 synthetic shot gathers are generated using the source signatures and velocity model. The shot gather from shot 50 which is located near the middle of the borehole is shown in Figure 2. The red dots show the first arrival picks. Picking uncertainties are believed to be less than half cycle of dominant frequency (~2.0 ms) since this is a synthetic noise-free data. Both travel time tomography using the code First Arrival Seismic Tomography (FAST) (Zelt and Barton, 1998) and waveform tomography are carried out.

![Figure 2. The synthetic shot gather for shot 50 which consists of 101 traces. Red dots show the first arrival picks.](image)

### 3.3 Data analysis and results

First step is travel time tomography. $101 \times 101 = 10201$ first arrival picks are used. The initial velocity model has a constant velocity at 2550.0 m/s (which is the middle point of the velocity variation range, not the mean of the true velocity model). RMS misfit
reduces from 3.2ms to 0.25ms in 10 iterations. Figure 3 shows the velocity model from travel time tomography and the true model for comparison. Compared to the true model, the travel time model reconstructs the big scale features in the true model except those at the top and bottom. This is because the ray density and angular variation at the top and bottom part is lower than in the middle. However, we can see the bottom part is better recovered than the top part because ray-based travel time tomography generally recovers high velocity anomalies better than low velocity anomalies due to Fermat’s principle:

Figure 3. The comparison of true velocity model and the model reconstructed by travel time tomography. A total of 10201 first arrival picks are used and RMS misfit is reduced from 3.2ms to 0.25ms beginning with a homogeneous model.
seismic rays choose high velocity features rather than low velocity features. This is true even for the better reconstructed part of the model. The net effect is that the velocity values in the model are generally over-estimated. The mean velocity of the travel time velocity model is 2731.2 m/s which is larger than that (2620.9 m/s) of the true model.

We then perform waveform tomography also using a constant velocity model as the initial model. 11 frequency components 52Hz, 100Hz, 148Hz, 196Hz, 244Hz, 292Hz, 340Hz, 388Hz, 436Hz, 468Hz and 500Hz are inverted sequentially from low to high and each is inverted for 10 loops. That is, the frequency component 52Hz is inverted first for 10 iterations and velocity model is updated. The updated velocity is then used as initial velocity for the 100Hz inversion and so forth. For simplicity, source signatures are assumed to be known and are the same as the synthetic source signatures. For real data application, this assumption certainly does not hold true.

![Figure 4. The velocity model from wideband (52-500Hz) waveform tomography (left), the true velocity model (middle) and the difference (wideband model – true model) between the two models (right).]
Figure 4 shows that wideband waveform tomography reconstructs the true velocity model much better than the travel time tomography does. Although the top most and bottom most parts of the model are not as well imaged as the middle part, however, they are also much better imaged than by travel time tomography. What is more interesting is that unlike the travel time tomography, waveform tomography reveals the low velocity features (see top part of the left plot in Figure 4) in a similar resolution as it does the high velocity features (see middle part of the model). The mean velocity of this model is 2619.3m/s which is very close to that of the true model (2620.9m/s). As we know, many geological features interesting to human being happen to have lower seismic velocities than surroundings. The most famous examples are oil, gas-hydrates, some waste deposits and sediments below basalt. Therefore waveform tomography is expected to have more applications in seismic exploration in the oil industry and environmental geophysics as its potential is made full use of. Visual comparison of the two models (left and middle plots in Figure 4) indicates that velocity distribution patterns down to ~3m scale features in the two models are similar to each other.

Visual comparison of color images can be deceptive, however. We therefore illustrate the difference more objectively by quantifying it as shown in the right plot of Figure 4. This plot shows the difference has a linear pattern almost parallel to the source and receiver lines. The pattern is related to the source-receiver geometry of cross-well experiments and other non-linear features are due to the non-uniform resolution as a function of space coordinates in the experiment. We can see the top most and bottom most parts are indeed less well imaged than other parts. Most of the differences have absolute values no larger than 100m/s.
It is mentioned above that waveform tomography can reconstruct the true image with high resolution by only inverting a sub-domain of all frequency components. This is one of the advantages of waveform inversion in the frequency domain rather than the time domain. The lowest frequency to be inverted depends on how close the initial velocity model is to the true model. This point will be illustrated in full length later in this chapter.

The highest frequency component to be inverted depends on (1) what the highest usable spectrum is in the data; (2) what the smallest scale feature we are interested in. Usable spectrum is defined here as those frequency components that are consistent throughout a whole dataset. In a real data application, the highest frequency component to be inverted usually depends on noise levels because the phase of high frequency components are more easily altered with large relative uncertainties than low frequency components, i.e., high frequency components are more easily to be inconsistent in the presence of noise. It is usually true that the highest usable frequency is usually lower than the highest frequency component that is still a significant component of the source spectrum, yet may be contaminated by random noise. However, noise is not the sole reason for that the data become inconsistent at higher frequencies. Other possible reasons include (1) higher attenuation for higher frequencies so high frequency components are better recorded by near offset receivers than far offset receivers; (2) higher frequency components are easier to be spatially-aliased; (3) the repeatability of the high frequency end of the source spectrum. For these reasons, it is important to first determine what the usable spectrum is for a real dataset. This issue will be discussed in detail in Chapter 4 where waveform tomography is applied to HAFB datasets.
Choice of the frequency components to be used is determined by a few factors. If computing cost is not an issue, waveform tomography could invert all usable frequency components to assure the highest quality image. It seems that the interval is a matter of choice. However, we certainly do not have to invert all frequency components to obtain a decent image by waveform tomography. To obtain a large velocity model by waveform tomography which requires intensive computing, it is wise to keep in mind the trade-off between computing cost and image quality to obtain a decent image at relatively low computing costs. The computing aspects of waveform tomography and travel time

![Figure 5](image)

Figure 5. The velocity model by narrow band (252Hz-500Hz) waveform tomography (right) is compared to the true velocity model (left). Obviously the tomography fails to reconstruct the true model.
tomography will be compared in detail later in this chapter.

In summary, wide-band waveform tomography is capable of reconstructing velocity at a resolution much higher than travel time tomography. However, real seismic experiments can not be conducted at very low frequencies and the resulting seismograms frequently do not contain much significant low frequency spectra. A more realistic starting frequency would be much higher than 52Hz. Figure 5 shows the velocity model by waveform tomography applied in the same way as shown above except that the frequency band is from 252Hz to 500Hz. Obviously the velocity model fails to reconstruct the true model.

![Graph showing observed and synthetic seismograms](image)

Figure 6. Schematic plot showing the cycle skip problem. The red curve is the observed seismic signal. The blue solid and dashed curves are the seismograms initially and finally predicted in the first application of waveform tomography. The black solid and dashed curves are the seismograms initially and finally predicted in the second application. Note a full cycle is skipped in the second application.

The problem is called cycle skipping. Figure 6 conceptually shows how cycle skipping happens. In Figure 6, the red solid curve is the observed seismogram corresponding to
one single frequency. If the initial velocity model is close enough to the true model and predicts the mono-frequency seismogram like the blue solid curve which is within half cycle of the observed seismogram, then waveform inversion of that frequency will end up fitting the observed seismogram perfectly like the blue dashed curve. However, if the initial velocity model is not that close to the true model and predicts the mono-frequency seismogram like the black solid curve, then inversion may very likely end up fitting the observed seismogram like the black dashed curve. In the frequency domain the black dashed curve fits the observed as perfectly as the blue dashed one; however, the black dashed curve skips exactly one cycle — a cycle skip. It is easy to see that the higher the frequency component is, the more likely cycle skips occur.

The cycle skip problem shows that if the initial velocity model in waveform tomography is not close enough to the true model, then inversion procedure will lead to a convergence not meaningful. This is especially true if gradient method is used in inversion because search direction is localized. In this case, the homogeneous initial velocity is not adequate to predict initial seismograms within half-cycle of the observed ones for the 252.0Hz component; however, it is for the 52.0Hz component as in wide band waveform tomography. So waveform tomography works only if the initial velocity model is adequately close to the true model. It is difficult to define a close model quantitatively, however, a close model depends relatively on the lowest frequency component used in the waveform inversion.

Then is a velocity model from travel time tomography a close model? If it is, what starting frequency components is it good for? The narrow band waveform tomography is performed again in exactly the same way as above except the initial velocity is the
velocity model from travel time tomography; i.e., the model shown in the left plot of Figure 3. Figure 7 shows the resulting velocity model reconstructs the true velocity model reasonably well, although not as well as the wide band velocity model. The difference between the two models (See right plot of Figure 7) is much larger than the difference between the wide band model and the true model (right plot of Figure 4) especially at top and bottom parts of the model. The difference is large at top and bottom parts are due to two reasons: (1) the initial velocity model (i.e., the travel time tomography model) is far away from the true one at these parts, i.e., severe artifacts at these locations; (2) lower resolution at these parts because they are not well illuminated. Nevertheless, the model by travel time tomography followed by narrow band waveform tomography successfully reconstructs most of the features in the true model, indicating the travel time velocity model is adequate to prevent the cycle skip problem. Since 252.0Hz is a relatively high

Figure 7. The velocity model reconstructed by narrow band waveform tomography using the velocity mode from travel time tomography as the initial model (left), the true model (middle) and the difference between the two models (right). Note the difference is larger than that in Figure 4.
starting frequency, a velocity model from travel time tomography is probably a good starting velocity model for most practical cases. This indicates that the two tomographic approaches could be complementary to each other.

For real data application, we have to take many other factors such as noises, source signatures into account. As pointed out by Pratt et al. (2002), even the combination of travel time tomography and waveform tomography is prone to potential difficulties. The message from this study is that waveform tomography could image the sub-structure with high resolution; however, it is less robust than travel time tomography, i.e., the predicted travel time does not change dramatically when the model varies.

3.4 The computing aspects of travel time and waveform tomography

This section briefly compares the computing aspects of both tomographic algorithms such as computing time required, memory usage. The code package FAST (Zelt and Barton, 1998) for travel time tomography and the code package FULLVW (Pratt et al., 1998) for waveform tomography are both written in Fortran 77. All tests in this study are carried out in a SUN Ultra-Enterprise workstation.

Table 1. Computing aspects of travel time and waveform tomography (model dimension: 221 × 421).

<table>
<thead>
<tr>
<th></th>
<th>Travel time tomography</th>
<th>Wide Band Waveform tomography</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU time (Hr)</td>
<td>1.5</td>
<td>9.4</td>
</tr>
<tr>
<td>Memory usage (Mb)</td>
<td>62</td>
<td>192.0</td>
</tr>
</tbody>
</table>

Table 1 shows that the wide band waveform tomography takes about ~6 times as long as the travel time tomography does, and occupies a memory ~3 times as large. It is
obvious that waveform tomography is much more expensive than travel time tomography. Furthermore, the cost of waveform tomography increases faster than the model dimension does because if the model dimension is doubled, the matrix in forward modeling would have 4 times as many elements, therefore the resulting computing time is at least 4 times as much.

The most time-consuming part of waveform tomography is forward modeling. The waveform tomography code FULLWV has incorporated most recent developments in computing algorithms to solve the forward modeling efficiently. For example, nested dissection is used to solve $Ax=b$ problem by LU decomposition. However, the code cannot be applied to large models because most operation systems such as UNIX put a limit on the size of a single array statically allocated and the limit which depends on operational systems can be easily exceeded when models get larger. For example, under IRIX operation system in SGI workstation, the code can only handle models with 300 by 300 grids because the code allocates arrays with the size of $16 \cdot m$, where $m$ is the total number of grids in the model. The sizes of these arrays can go beyond the limit easily. On the other hand, some surveys especially marine surveys could deploy receiver arrays 12.0km long. How can the code possibly be applied to these data?

The solution to the question is to use the dynamic array allocation feature in Fortran90. Fortran 77 allocates only static arrays which are subject to the limitation. Dynamic arrays can be allocated and de-allocated at the programmer’s will in a code without limitations imposed by the operation system. The code FULLWV has been re-written into Fortran90 version. The new version can handle models with 1000 by 1000 grids and can be applied
to a survey area 50km wide and 10km deep if frequency components up to 10Hz are inverted. The technical detail of the modification is summarized in Appendix I. One synthetic example using the Fortran90 version is shown in Figure 8.

![Figure 8](image.png)

Figure 8. A synthetic example using the Fortran90 version of FULLWV. The top plot shows the true velocity model. The middle is the starting velocity. The bottom plot shows the velocity model reconstructed by waveform tomography. 9 frequency components from 1 to 9Hz are inverted.

The synthetic example shows an application of waveform tomography to a typical 2D marine profile. The model has 598 by 281 grids with a uniform grid spacing 35.0m, representing a model 20.9km wide and 9.8km deep. The velocity variation in the model
reflects a typical marine environment: from 1.5km for seawater at the top, 2.0-3.5 km/s for sediments and 3.5km/s for a salt dome. Also source-receiver geometry in the example is that of a typical 2D marine survey: source is located 150.0m away from the near end of a 7.2km-long streamer which is moving along a straight line. 288 hydrophones distribute along the streamer with a 25.0m interval. 91 sources are detonated along the profile with a 150.0m interval.

The goals of this synthetic test are (1) to test the capacity of the new version FULLWV in handling large models; (2) to test how a salt dome can be imaged. The true model mentioned above is reasonably large for the first goal. For the second goal, we setup an initial velocity model (see Figure 8) which is basically is smoothed version of the true velocity model. The boundary of the salt dome is blurred and velocity within it is incorrect. In real data application, the initial velocity model can be obtained from NMO analysis and depth migration to know the layer boundaries in sediments and the overall shape of the salt dome. The initial model is not obtained from travel time tomography because the 7.2km-long streamer allows the deepest seismic rays to touch only the top of the salt dome.

The next step is to see how the image of the salt dome can be sharpened by waveform tomography. The bottom plot of Figure 8 shows the velocity model from waveform tomography inverting 9 frequency components from 1 to 9Hz. The plot shows that the salt dome can be imaged with sharpened boundary and a velocity closer to the true model. By inverting higher frequency components, the image could be improved, however, with larger cost.
3.5 Discussion and Conclusion

This chapter compares comprehensively the advantages and disadvantages of travel time tomography and waveform tomography. Since travel time tomography is much more widely applied than waveform tomography, this study also serves the purpose of illustrating various technical aspects of waveform tomography, and issues in its potential applications through comparison to travel time tomography. This comparison is also meaningful in another perspective that the distinctions between ray-based tomography, e.g., travel time tomography, and wave equation based tomography, e.g., waveform tomography are analogous to distinctions between ray-based and wave equation based other approaches, such as migration (Pratt et al., 2002). Therefore, the conclusions in this study may have implications for other seismic imaging approaches involving ray tracing and wave equations, such as migration.

The advantages of waveform tomography are (1) high resolving power associated with waveform inversion, and (2) the same high resolving power for low and high velocity features. Theoretically the resolution is scaled with wavelengths, i.e., the higher frequency the seismograms contain, the higher the resolution is. However, the high frequency components are easy to be inconsistent when subject to noises in real data application. So in real data application, the resolution depends on whatever the highest consistent frequency component is. The second advantage is also interesting because many well-known geological targets happen to have lower velocity than surroundings.

The disadvantages of waveform tomography are (1) highly non-linear inversion, and (2) high computing cost. Highly non-linear inversion implies extreme sensitivity of predicted seismograms to velocity models, which makes waveform tomography not so
robust as travel time tomography. This also implies that the initial velocity model has larger impact on the performance of waveform inversion. Because of the cycle skip problem and gradient method which assumes the initially predicted seismograms are close to the observed ones, waveform tomography is demanding on initial velocity model, requiring it to be close to the true one. Because of this weakness, waveform tomography is far more likely to fail than travel time tomography. The second disadvantage is obvious through comparison above. This disadvantage may appear less obvious as computers get more and more powerful.

The advantages of travel time tomography are (1) robustness, and (2) lower computing cost. Although being a non-linear inversion also, travel time tomography is more robust than waveform tomography because perturbations of travel times are much less sensitive to perturbations in velocity models than perturbations of waveforms are. Both advantages are the main reasons why travel time tomography is widely applied.

The disadvantages of travel time tomography are (1) lower resolving power, and (2) un-balanced resolving power for high and low velocity features. The resolution by travel time tomography is lower because travel time perturbations are insensitive to small inhomogeneities due to waveform healing. The longer rays travel, the more waveforms heal. So the resolution decreases with ray travel lengths. The larger wavelengths are, the more insensitive travel time perturbations are to small inhomogeneities. So the resolution also decreases with wavelengths. The resolution by waveform tomography is higher because the technique makes uses of diffracted/scattered waves which are generated by small inhomogeneities. Theory of diffraction tomography (e.g., Wu and Toksoz, 1987) tells that inhomogeneities as small as points could be resolved with a resolution up to half
a wavelength. Diffracted/scattered waves which contain information about small inhomogeneities cannot be described by seismic ray tracing, which gives rise to the need of developing image tools other than ray-based ones. The second disadvantage results from the fact that rays generally tend to sample high velocity features better than low velocity ones, which can be understood by Fermat’s principle. This disadvantage seems especially true when ray coverage is poor. However, low velocity features could be well resolved provided favorable source-receiver geometry and dense source-receiver coverage.

Travel time tomography and waveform tomography could be complementary to each other. Waveform tomography reconstructs sub-structures in high resolution when the observed data contain noise-free complete spectrum. However, noises are inevitable in field data. Spectrum of these data are significant only within some frequency bands, depending on seismic sources and receivers. Experiments are conducted in a way such that high frequency components are preserved for resolution purpose. It is shown in this study that travel time tomography can provide a reasonably good starting velocity model for waveform tomography even if low frequency spectrum is non-significant. The resulting model resolves sub-structures in a resolution comparable to that of the wide band waveform tomography.

The Fortran90 version of the waveform tomography code eliminates the limitation on array dimensions using the dynamic allocation of memory, thus improves the applicability of the waveform tomography technique. The future work would parallelize the algorithms in the forward modeling of the code in machines with shared or distributed memory.
Chapter 4

Waveform Tomography Applied to the VSP Dataset

4.1 Introduction

This study applies the 2D waveform tomography technique of Pratt et al. (1998) to a Vertical Seismic Profile (VSP) and a surface dataset from a groundwater contamination site at the Hill Air Force Base (HAFB), Utah. The datasets were acquired along with a 3D surface reflection and a 3D surface tomography dataset. The objective of the whole HAFB seismic experiment is to map the 3D geometry of the bottom of the aquifer system which is polluted by relative dense chlorinated solvents known as Dense Non-aqueous Phase Liquids (DNAPLs). Over a half century chlorinated solvents have been widely used and the unchecked release of these compounds into the subsurface produced wide spread ground water contamination. Only in recent decades have the severity and scale of the contamination been made apparent (Younger et al., 2002). The HAFB seismic experiment is to map the distribution of the contaminated water in the target area, a part of the ground water remediation efforts initiated by Department of Energy. The objective of the VSP experiment is to sample the lateral formation heterogeneities within a buried paleo-channel which is the structural host for the contaminants.

The VSP dataset is first undergone a series of pre-processing procedures include (1) adding geometry, (2) stacking and (3) first arrival picking. Second, travel time tomography is performed to obtain an initial velocity model for waveform tomography. Third, the study applies wave equation datum to get rid of ground rolls, followed by consistency checking for every frequency component. Finally waveform tomography is
applied to the first arrival waveforms to obtain a high resolution velocity model and source signatures.

4.2 The geological setting

The ground water contamination site (therefore the experiment site) is located at the Operable Unit 2 (OU-2) of Hill Air Force Base (Figure 1). The site geology consists of a heterogeneous mix of Recent-Quaternary sands and gravels of the Provo formation overlying a silty clay layer of the Quaternary Alpine formation, both formations part of the Lake Bonneville group. The alluvium is 3–17 m in thickness, and is variably permeable. The clay layer is at least 50m in thickness, and is largely impermeable,
causing it to act locally as an aquiclude. The top of the clay layer was incised in the Pleistocene by a drainage to the Weber River that formed a distinct paleochannel about 15m wide and up to 15m deep. Subsequent to burial, the area has experienced land slides and other mass wasting events, that have increased the lateral heterogeneity of the surficial sediments and have likely complicated the shape of the channel.

The paleochannel in the clay layer is a natural ground water channel and also acts as the structural trap for the DNAPLs, which have ponded in its deepest points. Originally trichloroethylene (TCE) and other DNAPLs used as degreasing agents were dumped in two pits at the OU-2 site. These trenches leaked the DNAPLs into the shallowest water table, which is perched on the Alpine formation clay. In an effort to contain the spread of the DNAPLs in the shallowest ground water table, a trench was dug around the entire site, which was then filled with bentonite clay. This trench is referred to as a containment wall.

4.3 The experiment

A series of high resolution seismic experiments were carried out in the OU-2 site (Figure 1). The experiments included a 3D surface seismic reflection experiment, a 3D surface tomography experiment, 6 borehole check-shot surveys and a survey combining 2 borehole vertical seismic profiles and a surface reflection profile, referred to as VSP experiment hereafter. The 3D surface seismic reflection experiment and the 3D surface tomography experiment are described in detail in Dana et al. (2002) and Azaria et al. (2002), respectively. The VSP experiment was designed to estimate material property
variations by lithological sampling from two boreholes and to image the subsurface structure. The VSP experiment is described in detail as following.

The two boreholes in the VSP experiment are separated by 21.35m horizontally. They are located at both ends of the light yellow line in Figure 2. The VSP source receiver geometry is further delineated in Figure 3. In Borehole 1 which is the southern borehole, 28 channels of data were recorded with geophone interval of 0.5m. In Borehole 2 which is the northern borehole, 25 channels of data were recorded with geophone interval 0.5m.

Figure 2. Map showing depths to clay top in the target area of the Hill Air Force Base (HAFB) seismic experiment conducted in Ogden, Utah in August, 2000. Green stars show the location of geophones and red stars the location of sources in the 3D seismic experiment. The light yellow line shows the location of surface geophone spread in the Vertical Seismic Profile (VSP) experiment described here. One borehole was drilled at each end of this line. The black line shows the geophone spread in the 2D seismic experiment conducted in 1998.
Another 60 channels of data were recorded at surface with geophone interval of 0.35m. The borehole data were recorded by a BHG-3 14-Hz 3-component geophone clamped to the borehole casing at different depths. The surface data were recorded by the Geometrics cabled seismograph system with 40Hz geophones. 31 sources are located at surface with 0.7m interval. The seismic source is a .223 caliber rifle. The rifle was fired 28 times at one source location and the borehole geophone was raised 0.5m and clamped again to the well casings after each shot until it reached the borehole top. Each source location is 35cm away laterally from every other geophone. A hole about 10cm deep was dug at each source location. A piece of plastic film was stuffed into the hole to prevent the hole from collapsing and to maintain a clean environment for rifle barrel. Before the rifle was fired, the hole in the ground was covered by a square of pad with a hole on it. The pad is to prevent dust from spreading out when shooting.

Figure 3. Source receiver geometry of the VSP experiment. The black and blue dots show the locations of receivers and sources, respectively.
The rifle barrel was pushed into the hole on the pad and then into the hole on the ground. The rifle operator stepped both feet on the pad when shooting.

4.4 The VSP dataset and pre-processing

The VSP dataset consists of data recorded at surface and at two boreholes. For surface data, since the rifle was fired repeatedly at one source location for 28 times, 28 shot gathers were recorded by the Geometrics recording system at one source location. 28 shot gathers associated with one source location is hereafter called a shot group. The first 13 shot groups were sampled at 0.5ms and each trace has 4096 samples. The rest 18 shot groups

![Seismogram recorded from shot 16. The 60 channels surface-recorded data are sandwiched between data recorded at Boreholes 1 and 2. Red dots show the first arrival picks. The surface data and some part of the borehole data are dominated by ground roll.](image)

Figure 4. Seismogram recorded from shot 16. The 60 channels surface-recorded data are sandwiched between data recorded at Boreholes 1 and 2. Red dots show the first arrival picks. The surface data and some part of the borehole data are dominated by ground roll.
groups were sampled at 1.0ms and each trace has 1024 samples. The first shot group is near Borehole 1 and the last near Borehole 2. Significant spectrum of the dataset is from 30 to 350Hz.

Source-receiver geometry was added into the whole dataset using Promax as a first step of pre-processing. The dataset originally downloaded from recording instruments were in SEGY format and each shot gather is associated with a Field File ID (FFID) number. All SEGY format datasets were first converted into Disk Image format by using the seismic processing software package Promax. By using the Promax module 2D Land Geometry Spreadsheet, a table where each FFID is associated with (X,Y) coordinates of a source and the (X,Y) coordinates of a group of geophones corresponding to that shot gather was edited. Finally using the Promax module Geometry Installation, all parameters in the table were written into headers of each shot gather associated with one FFID number.

28 shot gathers within one shot group recorded at surface were then visually checked for alignment in time domain. The purpose of this step is to identify shot gathers with good alignment so that they can be stacked to obtain a representative shot gather for that shot group. The representative shot gather presumably has higher signal-to-noise ratio than any of the single shot gathers within that shot group because random noises are more or less suppressed through stacking. To check trace alignment visually in time domain, three common receiver gathers for three geophones are sorted out from shot gathers in a same shot group. The three geophones included two located nearest to two boreholes and the one in the middle, i.e., 30-th geophone counted from Borehole 1. Traces in each common receiver gather should be well aligned in time domain if successive recordings
were done under the same condition. The most important reason for misalignment of traces is that when bullets were fired at a same hole repeatedly, the hole became deeper and wider. Therefore the response of the hole to the shooting evolves a little bit. After being visually checked, traces in each common receiver gather are found to be in groups with well-aligned traces in each group. Traces in the group with largest number of well-aligned traces are stacked and the resulting shot gather is chosen as the representative shot gather for that shot group. 60 traces in each of the 31 representative shot gathers were then combined with 53 traces recorded at two boreholes as one big shot gather. The resulting dataset has 31 shot gathers with 113 traces in each.

The third step in pre-processing is first arrival picking and signal conditioning. First arrival time picks are necessary in this study because they will be used in travel time tomography and will be used as the starting points of time windows for waveform tomography. Since arrival time picking can be subjective, the most important issue in absolute arrival time picking is consistency, i.e., picking by following a somehow objective criterion. The criterion followed in our case is that coherent arrivals are visually picked with amplitude amplified by 5 times and traces filtered if necessary. Under this criterion, 2400 first arrival picks were made. Some traces were not picked due to several reasons. One reason is that they are near offset traces with offsets <=1.5m. Signal amplitudes in near offset traces can be anomalously large due to strong ground motion and relative pick errors can be very large because of the short arrival times. Other traces are not picked because they are either overwhelmed by noises or bad traces. At the same time with picking, bad traces are killed. For example, the 2-nd and 8-th geophone
on the surface counted from Borehole 1 have never recorded any significant signals through the entire experiment and traces recorded by them were killed.

4.5 Travel time tomography applied to the VSP dataset

The first arrival waves in the data recorded at surface are either direct waves or refracted waves associated with shallow substructures. To control the quality of travel time picks, reciprocity of travel time picks from the data recorded at surface is checked since most source locations are near receiver locations. Ideally the time differences between reciprocal pairs of picks should be small. Figure 5 shows the statistics of time differences between each pair of reciprocal picks. 226 of the total 250 (i.e., 90.4%) time differences are within 2.5ms, indicating the overall picks are very much consistent. Picking uncertainty is therefore taken as 2.5ms.

Figure 5. Statistics of arrival time differences between each pair of reciprocal source-receiver. 226 of the total 250 (90.4%) time differences are within 2.5ms.
By checking picking profile, i.e., relationship of arrival times versus horizontal offset distance, it can be seen there is a thin layer on the top with velocity between 450.0 and 550.0m/s underlain by a layer with velocity around 350.0m/s, which may be one of reasons why ground rolls dominate the data recorded at surface.

An initial velocity model is set up for the travel time tomography. The initial model is derived from a 1D velocity model where velocity linearly increases from 300.0m/s at surface to 1500.0m/s at 24.0m depth. The initial model is 25.0m wide by 25.0m deep, numerically represented by 251×251 grids with homogeneous grid spacing 0.1m. The model size is slightly larger than experimental target size (21.4m × 15.0m) to accommodate all possible ray paths.

A total of 2400 first arrival picks are inverted in the travel time tomography for 30 iterations. The travel time RMS misfit is reduced from 5.08ms to 2.55ms, which is close to the picking uncertainty (2.5ms). Figure 6 shows the final velocity model from the travel time tomography. The model reveals that above 10m, which is the water table, the seismic velocity varies from ~300.0m/s to ~800.0m/s, and below water table, the velocity increases very fast. Approximately we can see a low velocity layer immediately below the top most thin high velocity layer. This agrees with the preliminary velocity analysis mentioned above. Scale features down to ~7.0m can be identified in a target dimension 21.4m by 15.0m. Furthermore resolution is not uniform in the entire target area, which can be seen in the ray coverage plot (see Figure 6).

Figure 6 shows that ray paths highly concentrate on areas with relative higher velocity. Holes in Figure 6 (poorly sampled areas) are due to two reasons: (1) lower velocity in the holes and (2) source-receiver geometry. The big hole at middle-bottom part of the model
Figure 6. Velocity model from travel time tomography (top) and its associated ray paths (bottom). Every fifth ray path is plotted for clarity. Blue dots show the locations of sources and black dots the locations of receivers. The locations of two boreholes are also shown.
is due to the source-receiver geometry. Also we can see that most deep-penetrating ray paths are provided by the borehole data instead of the surface data. In fact, without the borehole data, the limited offsets in surface data alone and the low velocity layer would confine the ray paths to the top most 3.0m only. The transmitted waves in borehole data provide key constraints on the deep structures in this tomography.

4.6 Waveform tomography

Before waveform tomography can be applied to the dataset, this study has to take care of a few problems. The first problem is to mute noises before first arrival and noises mixed with signals. For all traces to have the same time resolution (therefore same frequency resolution), the mute gate should be the same for every trace in the entire dataset. That is, the time window length for each trace is a constant which is to be determined. The time window should begin at the first arrival time and end at a point such that the window contains as much signal as possible. However, because of the ground rolls which are not modeled by solving the acoustic wave equation, the time window cannot be very long due to the trade-off that if the time window is very long, only far offset traces can be used (see Figure 3). One way to get out of this trade off is to get rid of the ground rolls.

Ground rolls can be suppressed by several approaches such as f-k filters and wave equation datuming (e.g., Duncan and Beresford, 1995). They can be suppressed by an f-k filter because dispersive surface waves such as ground rolls can be represented by a fan-shaped area in the f-k domain, therefore they can be suppressed by a so-called f-k fan filter. When being carefully implemented, the filter can recover signals from ground rolls.
reasonably well. However, signal waveforms are usually deformed to some extent by this filter, which is not ideal for waveform inversion. Ground rolls can also be eliminated using wave equation datuming by taking advantage of the property that surface waves do not penetrate deep. If re-datum the wave fields recorded at surface to a certain depth, then zero all the remaining wave fields above the depth and datum the re-datumed wave fields back to the original recording level, the surface waves will be presumably eliminated. The shortcoming of wave equation datuming is that near offset data will be highly deformed. Traces with 3.5m offset are not used in this study. These traces cannot be used even if not deformed by datuming because of non-elastic phenomena in wave fields very near sources.

Borehole data are contaminated by ground rolls not as extensively as surface data are due to the fact that some of the borehole data are recorded at depths larger than 3.0m. However, a type of noise unique to borehole data is tube waves which are seismic waves trapped in the borehole casing. This type of noise in our data travels at ~2000.0m/s which is much higher than the seismic velocity in surrounding media, thus signals can be totally lost in the tube waves especially when sources at surface are close to boreholes (see Figure 9). For this reason, borehole data recorded when sources are within 3.5m from the borehole are not used, in addition to the data recorded at depths shallower than 3.0m. The remaining borehole data cannot be fully used either because the first arrival waveforms are immediately followed by P-to-S mode conversions. For all these reasons, a 30-ms-long time window is chosen to include as much useful signal as possible in a single trace and to include as many traces as possible. The muted data are Fourier transformed into frequency domain.
The next problem to solve is to decide what frequency components will be inverted in the waveform tomography. The problem is solved by consistency check. Following Pratt (1999), we plot the real part of the frequency domain data as a function of source index and receiver index shown in Figure 7. See the caption of it for details of how the figure is plotted. Ideally if all surface data are consistent, the real part should be the largest along the diagonal line of the figure because the diagonal line corresponds to zero source-receiver offset, and should fluctuate symmetrically off-diagonal with respect to the diagonal. For borehole data, the pattern is more complicated because both up-going and down-going waves are recorded at a certain depth. However, if the vertical velocity structure is not very complicated, we should be able to see more or less a stripe pattern. The manner of displaying data is analogous to the interference diagram in optical holography.

Figure 7. Distribution of real part of frequency domain data for 40.0Hz as a function of source index and receiver index. The first 60 channels are for surface data, the next 28 for Borehole 1 and last 25 for Borehole 2.
Consistency for the frequency component 40.0Hz can be checked in Figure 7. Since near offset data are not used, we can see blanks in both surface and borehole data. The blanks in the surface data are mostly due to that near offset traces are all killed. Those in borehole data correspond to traces killed because they are either near the sources or contaminated by ground rolls. We can see the surface data display a stripe pattern parallel to the diagonal line and borehole data shows a concentrated band pattern, indicating that these data are largely consistent. Similarly, other frequency components are also checked. Finally, 9 frequencies components 15, 25, 30, 40, 50, 60, 80, 90 and 115Hz are identified as being consistent and will be inverted.

In order to assure a successful waveform inversion, frequency components which are checked for consistency from low to high are inverted successively, using the travel time velocity model as starting model. Frequency component as low as 15Hz is inverted first in this study for reasons: (1) the component is consistent also; (2) cycle skips occur less possibly if inverting a low frequency component. The travel time tomography velocity model is presumably close enough to the true model, however, uncertainties and artifacts in the model introduced by uncertainties in picks and ray tracing could possibly cause the cycle skip problem if high frequency components are inverted first. For the clean synthetic data in Chapter 3, the inversion could begin with a frequency component as high as 252.0Hz using a travel time velocity model as the initial model. For real data, because of the uncertainties and noises in data, it is difficult for even the combination of travel time and waveform tomography to get a high resolution velocity model (Pratt et al., 2002); (3) a successful inversion using low frequency component first paves the way for
successful inversions using higher frequencies later, therefore a successful waveform inversion.

Source signatures are estimated using the initial velocity model before the waveform tomography is applied. The source signature for each source or the source signature for a group of sources can be estimated. This study estimates a source signature for each source first. Usually some of the estimated source signatures are poorly constrained because data in some of the shot gathers are very noisy or useful data are very limited. In this case, the poorly estimated are grouped into adjacent sources whose signatures are relatively better estimated, and one source signature is estimated for one group of sources. Technically 3 steps are involved in estimating a source signature: (1) initially guess a signature, say, a Ricker wavelet with a similar dominant frequency and bandwidth as that in real data; (2) Forward model using the initial source signatures for each frequency component involved and obtain a wave field sampled at receivers; (3) cross-correlate the forward modeled wave fields and the observed wave fields and determine a correction parameter $\lambda$ for each frequency component. Then multiply the frequency components of the initial source signatures with their respective correction parameters which are complex numbers, and obtain a new source signature in time domain by inverse Fourier transform. This study estimates 49 frequency components of a source signature from 10Hz to 250Hz with frequency interval of 5Hz.

4.7 Q value determination

The amplitudes of seismic waves in the observed data decreases rapidly with offsets, indicating significant seismic attenuation in the target area. Since waveform tomography
fits both the phase and amplitudes of the observed data, seismic attenuation should be incorporated in the waveform inversion. The Q values are determined by trial-and-error method, i.e., synthetic seismograms are computed with different Q values and the relationships of amplitudes versus offsets are investigated for different cases. Q value is determined when one of the relationships matches the observed. In this study, the amplitude-versus-offset relationship is measured using the data recorded at surface and Q values are constant for the whole model as a first-order approximation.

![Graphs showing amplitude vs offset](image)

**Figure 8.** The amplitude versus source-receiver offset relationship for real data (top) and synthetic data with Q=20 (below). The relationship is measured over 4 shot gathers and averaged as red curve (top) and blue curve (below).

The amplitude-versus-offset relationships for real data and synthetic data for Q=20 are shown in Figure 8. The amplitudes for a same offset measured over 4 shot gathers are
averaged to obtain representative amplitude-versus-offset relations for both the real and synthetic data. After normalization and correction for 3D propagation for the synthetic relations from 2D waveform modeling, the relations can be compared (see Figure 9). We can see that the amplitude-versus-offset relation using Q=20 matches the observed one very well. The message from this study is that Q=20 is a good approximation for the real distribution of seismic attenuation in the target area.

Figure 9. The amplitude versus offset relationship. The red star and blue diamond curves are the same as in Figure 8 except the blue diamond curve is normalized against the red curve by matching the first three points. The blue triangles are corrected for 3D propagation and match the observed ones. The green curve is derived in the same way as the blue diamond curve except using Q=100.
4.8 Waveform inversion

9 frequency components are thereby inverted successively. By gradient method, each component is inverted for 10 loops. The model in waveform tomography is 24.4m wide and 22.0m deep which is padded larger than the actual target dimension (21.4m by 15.0m) to facilitate the simulation of boundary conditions. To have better constraints on deep part of the model, the gradient vector estimated in each loop is weighted such that the deeper part has larger weights than the shallower parts. The weights balance the tendency that the data have larger resolving power for shallow part of the model than the deep part. They are vertically cosine-tapered from 12.0m toward surface and from 18.0m toward the bottom, i.e., 22.0m.

The flow chart of the inversion is as the following: 9 frequency components from low to high are inverted individually and the velocity model is updated successively. Source signatures for individual sources are updated for amplitude only during each individual frequency component inversion. After one stage of frequency domain inversions, source signatures are inverted using the updated velocity model. Both amplitudes and phases are updated in the source signature inversion. The newly updated source signatures are then used in the next stage of frequency domain waveform inversion and so forth.

After 3 stages of waveform inversions, 69.4% reduction in the accumulated misfit function is achieved and the final velocity model predicts the observed waveform fairly well. Figure 10 shows the source signatures for all 31 shots. These source signatures are very much consistent except some have more details than the others. The dominant frequency of these source signatures are at ~80.0Hz. Figures 11 shows the waveforms observed, datumed and predicted using the final velocity model shown in Figure 12. Red
dots in Figures 11 are first arrival picks. Those traces without red dots are not used in waveform tomography. We can see that after the surface data are datumed, the near offset traces are deformed so they are not used in waveform inversion either. Figures 11 also shows that the waveforms recorded at surface and boreholes are different in shape, i.e., the waveform recorded at boreholes are more elongated in time domain and have less high frequency component than those recorded at surface because surface data are recorded by 40Hz geophones and borehole data by 14Hz geophones. Source signatures in Figure 13 are mainly determined using surface data. So we can see that the predicted waveforms fit the surface data better than borehole data.

![Source index](image)

*Figure 10. Estimated source signatures for 31 shots. Zero time of all source signatures is at 20ms.*

The velocity model from waveform tomography has several interesting features. (1) Compared to the travel time velocity model shown in Figure 6, the velocity model from waveform tomography reveals more details in sub-structure. Scale features down to ~1.5m are recovered in the waveform tomography while scale features down to ~6.0m are recovered in travel time tomography. (2) A low velocity zone occurs at the depths from 1.0m to 4.0m. (3) Velocity increases from ~300.0m/s at top to 1600.0m at 15.0m depth, an average vertical gradient of ~80.0m/s/m. (4) The velocity distribution near
Figure 11. Observed (top), datumed (middle) and synthetic (bottom) seismograms of shots 9 [(a)] and 18 [(b)]. Traces which are not used in waveform tomography are plotted in grey in the top two panels. They are either near-offset traces or contaminated by ground roll. Red dots in the panels show the observed first arrival picks.
southern borehole is more heterogeneous than that near northern borehole. (5) Velocity contrast as large as ~200 m/s could occur over a range as short as ~1.0 m.

![Velocity model](image)

**Figure 12.** The final velocity model from waveform tomography. Nine frequency components from 15 to 115 Hz (15, 25, 30, 40, 50, 60, 80, 90, 115 Hz) are inverted and misfit in the final model is reduced by 69.6%.

### 4.9 Geological interpretation of the waveform tomography velocity model

The velocity model from waveform tomography is geologically interpreted as a thin layer of desert hardpan on the top underlain by an unconsolidated mixture of sand, clay and gravel. Below water table which is at 10.0 m, the velocity increases rapidly. This interpretation is calibrated by and/or correlated with two lithological logs from the two boreholes on both sides of the model (Figure 12). Lithology log at Borehole 1 shows that
above 6.1m, silty clay mixes with mottles or sand partings at different depths. From 6.1m to 8.8m, gravelly sand dominates and further below sticky clays dominates again. This matches the velocity model in Figure 12 where small velocity variation can be seen above 6.0m and velocity between 6.0 and 8.5m are very different from above. The formation at Borehole 2 is very much homogeneous except for the top ~3.0 meters. Velocity variation shown in Figure 12 near this borehole correlates with this feature very well, i.e., a low velocity feature above ~3.0m depth is underlain by a very much homogeneous layer down to 12.5m. Below 12.5m, the velocity increases rapidly. The lithology log from Borehole 2 shows that a mixture of clay, sand and gravel above ~3.0m is underlain by a thick layer of clay down to 12.5m. Below 12.5m is the mixture of sand and gravel. Generally we can see that the correlation between the lithology logs and the velocity model is good.

Figure 13. Geological interpretation of the final velocity model in Figure 16 and its correlation with two lithology logs from two boreholes. The velocity structure generally correlates well with two formation logs. Velocity structure near Borehole 1 is more heterogeneous than that near Borehole 2 because of more heterogeneous formation near Borehole 1.
The good point-wise correlation between two lithology logs and the velocity model can be used as a calibration for geological interpretation of the velocity model between the two boreholes. The interpretation is shown in Figure 13. Based on the correlation, the low velocity layer beneath desert hardpan is most probably an unconsolidated mixture of clay, sand and gravel. The desert hardpan has higher velocity than the mixture because it’s more highly cemented, therefore more consolidated than the latter. The low velocity layer is above water table and a small amount of air introduced into the pores of this mixture would greatly reduce the effective bulk modulus, and therefore seismic velocities. Also from the correlation, we can see that a channel of low velocity at the depths between 6.1m and 8.8m extends horizontally from the southern borehole to x=8.0m. The composition of this channel is gravelly sand near the southern borehole and gradually evolves into a mixture of sand, clay and gravel in the middle of the model. This channel is sandwiched between a clay mixture above and a layer of brown clay below, lying immediately above the water table. Gravelly sand is porous and less consolidated than clay, resulting in velocity decrease.

The top of clay can also be identified based on the correlation between the lithologic logs and the velocity model. It can be seen that some of the clay is under the water table, and some of the clay is above it. Compared to the clay top mapping shown in Figure 2, the clay top identified in Figure 13 appears shallower, especially the clay top near the northern borehole. Figure 2 shows that the clay top is ~15.0m while the northern lithologic log shows that clay presents as shallow as ~3.0m. The difference is because the clay top shown in Figure 2 is interpolated using sparsely distributed well logs, showing broad features of the paleo-channel. The difference also suggests that the
geological structure in the target area can be more heterogeneous than that appears in Figure 2.

4.10 Resolution analysis

This section analyze the resolution of the velocity model shown in Figure 12 by a synthetic test which has exactly the same source receiver geometry and target area as those in the real VSP dataset from the HAFB seismic experiment. Three square-shaped diffractors are designed into the three-layer true model to test the relative performance of travel time and waveform tomography. The squares in the deepest layer are larger because resolution is expected lower there. To simulate the real data case better, 31 shot gathers generated synthetically are added with 31 gathers of random noises with power spectra at 3% of the average dominant power spectra of the data.

The travel time velocity model represents the large wave length features of the true model in the sense that velocity increases with depth (see Figure 14). However, the small scale features are not reconstructed because resolution in travel time tomography scales with Fresnel zone which is about ~6.0m for this test. The scale features of the diffractors are at ~1.5m which is much smaller than the Fresnel zone. That is why diffractors are not recovered in travel time tomography while they are recovered better in waveform tomography where resolution scales with wavelength. The diffractors in the top most two layers are recovered with decreasing resolution with depth and those in the deepest layer are not recovered at all because of poor illumination. This test shows that up to ~10m, the scale features down to ~1.5m are recovered with decreasing resolution.
Figure 14. The true model (top) and models from travel time tomography (middle) and waveform tomography (bottom) for the synthetic test.
4.11 Petrological interpretation

With the high resolution in it, the velocity model can even be interpreted in a petrophysical sense. Since the un-consolidated materials at the earth's top most a few meters are difficult to be characterized petro-physically in lab because of the coupling problem between transducers and the materials at very low pressure, this section shows a new approach by which petro-physical parameters such as porosity and degree of consolidation can be inferred from the seismic velocity information based on certain assumptions.

The approach can be summarized briefly in the following: the combination of the two models by Walton (1987) and Hertz-Mindlin (Mavko et al., 1998) shows that for a dry, dense, random pack of identical elastic spheres, the effective bulk and shear modulus are

$$K_{hm} = \left[ \frac{n^2(1-\phi)^2G^2}{18\pi^2(1-\nu)^2} \frac{R_c}{R_g} \rho_b g Z \right]^{1/3} \quad (1),$$

$$G_{hm} = \frac{5-4\nu}{5(2-\nu)} \left[ \frac{3n^2(1-\phi)^2G^2}{2\pi^2(1-\nu)^2} \frac{R_c}{R_g} \rho_b g Z \right]^{1/3} \quad (2),$$

where $K_{hm}$ and $G_{hm}$ are the bulk and shear modulus. $G$ is the shear modulus of the mineral grain in the material. The mineral composition in the VSP target area is quartz with clay and $G=39.0\text{GPa}$ (Mavko et al., 1998). $\phi$ is the porosity of the material. $\nu=0.17$ is the Possion's ratio for quartz with clay. $n$ is the average number of contacts per mineral grain, and is usually between $3<n<8$. Following Bachran et al. (2000), we take $n=6$ as an average number for the entire area. $\rho_b$ is the bulk density, which is $2.7 \text{ g/cm}^3$ above
water table and 2.8 g/km³ below water table. \( g = 9.8 m/s^2 \) is the gravitational acceleration. \( Z \) is the depth in meter. \( R_g \) is the mean radius of the mineral grain. \( R_c \) is determined by

\[
\frac{1}{R_c} = \frac{1}{2} \left( \frac{1}{R_1} + \frac{1}{R_2} \right)
\]

(3),

where \( R_1 \) and \( R_2 \) are the local radii of two contacting grains (Figure 15). It is easy to see that \( R_c \) is affected more by the smaller one between \( R_1 \) and \( R_2 \). In another word, \( R_c \) is very small if one of the local radii is very small. The ratio \( R_c/R_g \) is an indicator of material integrity. The ratio has positive impact on seismic velocity.

The underlying assumption in the Hertz-Mindlin model is that no slip exists between grains. The real material is usually a mixture of (1) a material with zero tangential stiffness and (2) a material that obeys the Hertz-Mindlin model. The no-slip assumption could result in a small error if only acoustic wave propagation is concerned and can be safely used in estimating the effective moduli (Mavko et al., 1998).

Walton’s model assumes that the random pack is dry. If fluids such as air are introduced in the pores, Gassman’s equation can be used to determine the effective moduli as

\[
\frac{K_{\text{eff}}}{K - K_{\text{eff}}} = \frac{K_{\text{km}}}{K - K_{\text{km}}} + \frac{K_{\text{air}} \text{ (or } K_{\text{water}})}{\phi(K - K_{\text{air}})}
\]

(4),

\[
G_{\text{eff}} = G_{\text{km}}
\]

(5),
Where $K_{\text{eff}}, K_{\text{air}}(K_{\text{water}})$ and $G_{\text{eff}}$ are the effective bulk modulus for the whole material, bulk modulus for the air (water) and effective shear modulus for the whole.

Figure 15. Conceptual plot showing angular grains contacting each other.

Once the effective moduli are determined, the effective velocity can be determined by

$$V_{\text{eff}} = \sqrt{\frac{K_{\text{eff}} + 4G_{\text{eff}}}{3\rho_b}}$$

(6).

The idea in this section is that by matching the theoretically calculated $V_{\text{eff}}$ to the velocity determined by waveform tomography, some of the petro-physics parameters such as porosity and $R_c/R_g$ ratio can be hopefully characterized. Since we didn’t measure in-situ the other parameters required by Equations (1) and (2) such as number of contacts per grain and Poisson’s ratio, it is impossible to determine the petro-physical parameters uniquely for the entire target area because of the heterogeneity in both structure and material. However, it is possible to characterize the unknown parameters generally based on constraints on the other parameters required by Equations (1) and (2). This section treats $n$, $\nu$, $g$, $\rho_b$, $Z$, $K$ as the known, and treats $\phi$ and $R_c/R_g$ as the
unknowns. Among the known parameters, the number of contacts $n=6$ is taken from the study of Bachran et al. (2000) and the others are well constrained. The 2D velocity model is averaged into a 1D velocity profile as the "observed" data in this section. By fitting the 1D velocity profile, a 1D profile of porosity and $R_c/R_g$ can be estimated (see Figure 16).

![Figure 16. Petrological interpretation of the final velocity model from waveform tomography. The 2D model is averaged into a 1D model (red curve) which is forward modeled using Eqs. (1) through (5). The known parameters are (1) mineral composition--quartz with clay ($K=39\text{GPa}$, $G=33\text{GPa}$, $v=0.17$). (2) $\rho_b = 2.7$ (above water table) and 2.8 $\text{g/cm}^3$ (below water table). (3) $g=9.8 \text{m/s}^2$. (4) $n$ is assumed to be 6. 1D profiles of porosity (middle) and $R_c/R_g$ ratio (right) which is a measure of consolidation can be determined by forward modeling the 1D velocity profile (left). Note that Gassman equation with water phase only fails to predict velocity change across water table.](image-url)
The 1D velocity profile has a pattern in it: it has two low velocity layer above water table and velocity increases from \(~800.0\,\text{m/s}\) at water table to \(~1500\,\text{m/s}\) at 15.0m. From the porosity and \(\frac{R_c}{R_g}\) profiles, we can see that the first low velocity layers are due to low porosity in desert hardpan above the layer and low \(\frac{R_c}{R_g}\) ratio in the layer. So the first low velocity layer between 1.0–3.0 depths are interpreted petrologically as a layer of more porous and less consolidated mixture of sand, gravel and clay beneath a layer of less porous and more consolidated desert hardpan. Similarly the low velocity layer between 6.0m and 9.0m is interpreted as a layer of less consolidated gravelly sand sandwiched between clay mixture above and clay below.

Figure 16 also shows that the Gassman equation with water only phase fails to predict the 1D velocity profile below water table. The Gassman equation predicts a sharp velocity increase across the water table while the “observed” velocity increases gradually toward depth. Figure 17 shows two 1D P- and S-wave velocity profiles measured in-situ by Liu et al. (1997) at the Memphis embayment, Tennessee. At the first location, the measured P-wave velocity (Figure 17, top) increases sharply beneath water table and at the other, it increases gradually (Figure 17, bottom). The first profile could be very well predicted by Gassman equation with only water phase while the second could not. The difference is due to that the velocity variation associated with water table is influenced not only by material variations, but also by partial saturation, therefore it is sensitive to the history of flow (Bachrach and Nur, 1998). The velocity variation across the water table in the averaged 1D model is similar to that seen Figure 17 (bottom), suggesting (1) non-uniform water saturation and (2) complicated flow history possibly due to pumping nearby for water purification in the ground water contamination site at HAFB.
Figure 17. Two P- and S-wave velocity profiles measured in-situ at the Memphis area (after Liu et al., 1997).
4.12 Comparison with other study

This section compares the velocity model from waveform tomography with images from (1) depth migration of surface data in the VSP experiment; (2) a velocity model from the 3D travel time tomography by Azaria (2002) using the 3D travel time tomography dataset.

![Diagram showing velocity model and depth migration](image)

**Figure 18.** The velocity model from waveform tomography is overlapped with the depth migrated VSP surface dataset. The reflection events approximately correlate with water boundaries around ~10m and clay top below water table. Velocity used in the migration is a step-wise constant 1D velocity model.

Figure 18 shows the comparison between the velocity model from waveform tomography and the depth migrated image using the data recorded at surface in the VSP
experiment. The velocity used in the depth migration is a step-wise constant 1D velocity model instead of the velocity model by waveform tomography to avoid any direct correlation between the waveform tomography velocity model and the final depth migrated image. Since the reflection events are retrieved from the dominating ground rolls using wave equation datuming, near offset data are highly deformed and are not used. So horizontally stack section only between ~3.0m to 19.0m are obtained, so is the depth migrated image. We can see (1) one reflection event associated with the velocity change from sand to clay at ~8.0m depth; (2) another reflection event dipping from left to the middle bottom and turning up. The second event is correlated very well with the velocity low in the middle lower part of the model.

Azaria (2002) performed a 3D travel time tomography using the 3D tomography data from the same HAFB experiment and projected the model into the VSP target area (see Figure 19 (a)) such that it can be compared with the waveform tomography model (see Figure 19 (b)). We can see the waveform tomography model reveals much more structure than the travel time velocity model. Generally this agrees with the rule of thumb that the resolution in travel time tomography scales with Fresnel zone while that in waveform tomography scales with wavelength. The difference in the two models is also due to the fact that the sampling rate in the VSP data is much higher than that in the 3D tomography dataset. The source interval in VSP dataset is 0.7m compared with 2.1-2.8m in 3D tomography dataset, and the geophone interval in the former is 0.35-0.5m compared with 2.1-2.8m in the latter.
Figure 19. Velocity model for the VSP target area from the 3D travel time tomography of Azaria et al. (2002) to compare with the model from waveform tomography.
4.13 Conclusion

This chapter applies waveform tomography to solve the problem of imaging the shallow substructure with high resolution. Chapter 3 of this thesis concludes that using the velocity model from travel time tomography, waveform tomography can reconstructs sub-structure with a resolution scaling with wavelength. To achieve the high resolution, first arrival travel time tomography is performed to provide initial velocity model for the subsequent waveform tomography. Although not all waveform can be used in the waveform tomography because of random and signal-generated noises, wave equation datuming is applied such that slightly more waveform data can be utilized.

Here are the conclusions: (1) The heterogeneous shallow seismic structure can be imaged with high resolution by waveform tomography. The heterogeneity at the shallowest part of the earth could be severe such that other seismic approaches may have difficulties in imaging the complicated structures. (2) When the ratio between the target dimension and the wavelength is small, waveform inversion techniques such as waveform tomography is necessary to achieve the objective of sampling lateral formation heterogeneities with high resolution. The VSP experiment shown in this chapter is the case. (3) The source signatures can be determined for individual shots and could be used by other applications; (4) The velocity model from waveform tomography reconstructs more details than the travel time tomography; (5) The waveform tomography reveals that (a) a low velocity layer at 1-3m depths; (b) large lateral variations. Velocity contrast as large as ~200.0m/s can occur in 1.0m range; (c) a large vertical gradient at ~80.0m/s/m; (d) small scale features down to ~1.5m. (6) The velocity model correlates well with two lithology logs and can be interpreted geologically; (7) It is possible to interprete the high-
resolution model from waveform tomography in a petrological sense. A few petrological parameters can be somehow constrained; (8) A constant $Q=20$ is determined as a first order approximation for the real attenuation distribution in the target area.

Resolution analysis shows that above $\sim 10\text{m}$, scale features down to $\sim 1.5\text{m}$ are recovered with decreasing resolution with depth even when data are slightly noise-contaminated. Large-scale layer boundaries further down in the target area can be still detected, however, small-scale features are not imaged due to poor illumination.
Chapter 5

Waveform Tomography Applied to the Surface Reflection Dataset

5.1 Introduction

Chapter 4 shows that waveform tomography can be applied to a Vertical Seismic Profile dataset to sample the lateral and vertical formation heterogeneities along the strike of a paleo-channel where DNAPLs were dumped. The transmitted waves recorded at depths in the VSP data constrain deep structures better than the surface data. However, surface data with large offsets that are not contaminated by ground rolls provide good depth-constraint seismic phases such as the refraction and reflection wave fields. This chapter will show that waveform tomography can be applied to 45 2D seismic profiles sorted out from the 3D surface reflection experiment to shed light on the 3D geometry of the paleo-channel, the structural host for the polluted ground water. Precise knowledge of the 3D geometry of structural hosts for polluted ground water (e.g., a channel in this case) is key to any ground water remediation programs in the world. This study shows how waveform tomography is applied to achieve the goal of determining the 3D geometry accurately.

The 45 seismic profiles are oriented approximately in the East-West direction and almost perpendicular to the surface spread of the VSP experiment and the strike of the paleo-channel. Each velocity model from waveform tomography therefore provides a cross-section view of the channel. The combination of these cross-sectional views gives a 3D structure of the channel, following exactly the philosophy behind any tomographic techniques, i.e., imaging slice by slice.
5.2 The experiment

The 3D surface reflection experiment was carried out using 619 Texan seismic recording systems and 40Hz single component geophones. The geophones were almost equally distributed into 6 parallel lines oriented roughly East-West (see Figure 2 in Chapter 4). Geophone inline interval is 35cm and cross-line interval is 2.1m. Seismic source is a .223 caliber shot rifle which was shot diagonally in a rotated brick pattern between two lines. The shot interval is 35cm. When shooting is finished between two lines, the ~103 geophones at the northern most line were moved to the southern end and seismic sources were fired again between the next adjacent 2 lines. This pattern was repeated and the six lines were rolling over to the south. Totally there are 45 lines. It took a crew of over 20 16 days to finish the experiment.

5.3 The dataset

Since we are doing 2D waveform tomography, 45 2D profiles were sorted out from the 3D dataset for this study. Shot gathers with sources immediately adjacent to the geophone lines were selected. These sources are 20cm away from the geophone line and distribute on both sides of a line except the southern and northern most lines (see Figure 1). Each geophone line is about ~36.0m long and the 20cm distance between the sources and the lines are neglected such that we regard them as strict 2D profiles. The number of usable shot gathers varies from line to line (see Figure 1). Overally the quality of dataset is good (Figure 2). The shot gathers for each line were sorted according to their locations at the line and merged. First arrival times are picked. Since we are doing travel time tomography to provide starting velocity models for waveform tomography, picks are
made on those trace with fairly clear signals to yield less RMS misfit in travel time tomography and to select better quality waveforms for waveform tomography. Wave-equation datuming is not applied to this dataset because (1) the offset in this dataset is

![Graphical representation of seismic profiles]

Figure 1. 45 2D seismic profiles sorted out from the 3D surface reflection dataset. Source locations (red dots) are 20cm away from geophone lines. The line numbers referred to in the text are counted from the north to south.
Figure 2. Representative shot gatherings from Lines 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40 and 44. Red dots show the first arrival picks.
Continuation of Figure 2.
Continuation of Figure 2.
long enough to have far-offset traces un-contaminated by ground rolls (see picks in Figure 2), and (2) previous application of wave-equation datuming shows that it only slightly enhance the window length of usable data because the waveform of some traces are deformed after datuming is applied.

Figure 2 shows representative shot gathers from one out of every 4 lines. Red dots show the first arrival picks used in travel time tomography. Refracted waves are evident in these plots.

5.4 Travel time tomography

Using the first arrival time picks, travel time tomography is performed using FAST for each of the 45 lines. The dimension of the model is [-1.5m, 37.5m] in X direction and [-0.5m, 32.5m] in Z direction. The model is represented numerically by finite difference grids with 0.1m interval in both dimensions. The model is padded larger than the actual survey scope to accommodate all possible ray paths. The initial velocity model is derived from a 1-dimensional velocity model where velocity increases linearly from 400.0m/s at Z=0.0m to 1500.0m/s at Z=15.0m, and to 1800.0m/s at 32.5m. This initial model is the same for all 45 profiles.

For each line, Table 1 shows the number of shot gathers, number of picks used, initial misfit and final misfit. The uncertainty in picks is assumed to be the same as that in the VSP picks, which is 2.5ms, a quarter of the period at 100Hz. However, RMS misfits for a few lines never goes below 2.5ms, indicating either picking uncertainty is larger than 2.5ms or data are not consistent due to (1) 3D effect in the source-receiver geometry; (2) out-of-plane wave propagation. For example, the misfit for line 7 never goes below
3.75ms. Reciprocity check for this line shows that the pick differences in most of the reciprocal pairs are real, indicating data are not very consistent due to the two possible reasons stated above. In this case, the model with the smallest misfit is chosen as the final

Table 1. Travel time tomography for 45 2D seismic profiles.

<table>
<thead>
<tr>
<th>Line No.</th>
<th>No. of sources</th>
<th>No. of picks</th>
<th>Initial RMS (ms)</th>
<th>Final RMS (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>790</td>
<td>5.46</td>
<td>2.56</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>614</td>
<td>3.57</td>
<td>2.77</td>
</tr>
<tr>
<td>3</td>
<td>22</td>
<td>1624</td>
<td>4.55</td>
<td>2.64</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>1738</td>
<td>4.42</td>
<td>2.46</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>1702</td>
<td>4.62</td>
<td>2.44</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>1245</td>
<td>5.91</td>
<td>2.42</td>
</tr>
<tr>
<td>7</td>
<td>24</td>
<td>789</td>
<td>6.78</td>
<td>3.75</td>
</tr>
<tr>
<td>8</td>
<td>20</td>
<td>1796</td>
<td>4.70</td>
<td>2.41</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
<td>1098</td>
<td>6.19</td>
<td>3.09</td>
</tr>
<tr>
<td>10</td>
<td>22</td>
<td>1061</td>
<td>5.44</td>
<td>3.06</td>
</tr>
<tr>
<td>11</td>
<td>29</td>
<td>1827</td>
<td>5.99</td>
<td>2.59</td>
</tr>
<tr>
<td>12</td>
<td>24</td>
<td>893</td>
<td>5.66</td>
<td>2.65</td>
</tr>
<tr>
<td>13</td>
<td>27</td>
<td>1331</td>
<td>5.82</td>
<td>3.10</td>
</tr>
<tr>
<td>14</td>
<td>29</td>
<td>1745</td>
<td>4.40</td>
<td>2.35</td>
</tr>
<tr>
<td>15</td>
<td>30</td>
<td>1796</td>
<td>4.70</td>
<td>2.41</td>
</tr>
<tr>
<td>16</td>
<td>29</td>
<td>1790</td>
<td>6.93</td>
<td>2.39</td>
</tr>
<tr>
<td>17</td>
<td>22</td>
<td>1476</td>
<td>6.92</td>
<td>2.58</td>
</tr>
<tr>
<td>18</td>
<td>23</td>
<td>1375</td>
<td>7.04</td>
<td>2.54</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
<td>1904</td>
<td>5.89</td>
<td>2.34</td>
</tr>
<tr>
<td>20</td>
<td>22</td>
<td>1405</td>
<td>6.78</td>
<td>2.50</td>
</tr>
<tr>
<td>21</td>
<td>24</td>
<td>1821</td>
<td>7.74</td>
<td>2.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>22</td>
<td>33</td>
<td>2301</td>
<td>9.41</td>
<td>2.61</td>
</tr>
<tr>
<td>23</td>
<td>27</td>
<td>1982</td>
<td>8.56</td>
<td>2.68</td>
</tr>
<tr>
<td>24</td>
<td>32</td>
<td>2304</td>
<td>4.88</td>
<td>2.58</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>1822</td>
<td>4.86</td>
<td>2.60</td>
</tr>
<tr>
<td>26</td>
<td>27</td>
<td>1708</td>
<td>6.10</td>
<td>2.97</td>
</tr>
<tr>
<td>27</td>
<td>30</td>
<td>1725</td>
<td>5.26</td>
<td>2.51</td>
</tr>
<tr>
<td>28</td>
<td>33</td>
<td>1956</td>
<td>4.98</td>
<td>2.66</td>
</tr>
<tr>
<td>29</td>
<td>32</td>
<td>2094</td>
<td>6.59</td>
<td>2.49</td>
</tr>
<tr>
<td>30</td>
<td>27</td>
<td>1628</td>
<td>7.75</td>
<td>2.56</td>
</tr>
<tr>
<td>31</td>
<td>25</td>
<td>1419</td>
<td>8.44</td>
<td>3.27</td>
</tr>
<tr>
<td>32</td>
<td>28</td>
<td>1472</td>
<td>8.61</td>
<td>2.68</td>
</tr>
<tr>
<td>33</td>
<td>29</td>
<td>1548</td>
<td>7.72</td>
<td>2.73</td>
</tr>
<tr>
<td>34</td>
<td>9</td>
<td>531</td>
<td>7.43</td>
<td>2.48</td>
</tr>
<tr>
<td>35</td>
<td>8</td>
<td>454</td>
<td>9.44</td>
<td>2.53</td>
</tr>
<tr>
<td>36</td>
<td>16</td>
<td>637</td>
<td>7.31</td>
<td>2.57</td>
</tr>
<tr>
<td>37</td>
<td>26</td>
<td>1085</td>
<td>7.87</td>
<td>2.48</td>
</tr>
<tr>
<td>38</td>
<td>22</td>
<td>698</td>
<td>7.17</td>
<td>2.51</td>
</tr>
<tr>
<td>39</td>
<td>24</td>
<td>816</td>
<td>8.02</td>
<td>2.62</td>
</tr>
<tr>
<td>40</td>
<td>14</td>
<td>542</td>
<td>10.27</td>
<td>2.77</td>
</tr>
<tr>
<td>41</td>
<td>17</td>
<td>682</td>
<td>8.45</td>
<td>2.49</td>
</tr>
<tr>
<td>42</td>
<td>18</td>
<td>653</td>
<td>7.05</td>
<td>2.97</td>
</tr>
<tr>
<td>43</td>
<td>21</td>
<td>852</td>
<td>6.76</td>
<td>2.69</td>
</tr>
<tr>
<td>44</td>
<td>20</td>
<td>896</td>
<td>11.35</td>
<td>2.80</td>
</tr>
<tr>
<td>45</td>
<td>9</td>
<td>363</td>
<td>9.66</td>
<td>3.09</td>
</tr>
</tbody>
</table>

(Continuation of Table 1)
Figure 3. Ray coverage for Lines 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40 and 44, i.e., ray coverage for one out of every 4 lines. Blue dot shows the locations of sources and black ones the locations of receivers. For clarity, one out of every three rays is plotted and each ray path is plotted by connecting one out of three points defining the ray path.
model, otherwise the model which fits the picks at the 2.5ms uncertainty level is chosen as the final model. From Table 1, it can be seen that travel time residuals for all lines except for Line 2 are reduced by >40.0% and the residuals for 32 out of the 45 lines are reduced by >50.0%. Figure 3 shows the ray coverage in the travel time tomography for Lines 1, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44 (i.e., one of every 4 lines).

The ray coverage in Figure 3 shows that (1) within each line, ray coverage for shallow part of the model (<3.0m) is better than deeper part (>8.0m); (2) ray coverage is different for different lines because of different size of useful data; (3) the largest depth ray samples in each line varies between ~9.0m to ~18.0m. Given the a priori knowledge that the paleo-channel is up to ~15.0m deep, this depth range of ray coverage indicates that the bottom of the channel can be imaged in the travel time tomography in some lines, but not as well imaged as the flanks of it. For some lines where rays only sample up to ~9.0m deep, the deepest part of the channel is not imaged unless the channel is equally shallow. However, the seismic waves actually sample areas centered at the ray paths instead of only the geometrical ray paths themselves, meaning that the part of model below the deepest ray paths can be possibly imaged in waveform tomography. (4) We can see clearly turning ray paths corresponding to refracted waves at both flanks of the channel in ray coverage plots of Lines 24, 28 and 40.

The final velocity models from travel time tomography are described below in the context of their comparisons with final velocity models from waveform tomography.
5.5 Waveform and travel time tomography

To show the relative performances of the two tomographic approaches and why it is necessary to apply waveform tomography, this study performs a synthetic test which is shown in Figure 4. 30 synthetic shot gathers with homogeneous shot interval 1.25m involves in this test. Each shot gather has 104 traces with geophone interval 0.35m. Both shots and geophones are on the surface. The source-receiver geometry is the same as a typical seismic profile sorted out from the 3D dataset. Since real data are usually contaminated by noises, random time series are generated and added to the synthetic seismograms. The random time series are scaled such that the power spectra of them are at 3% of the peak power spectrum of the synthetic data. Figure 4 shows the synthetic velocity model as well as models from travel time and waveform tomography. The velocity model from travel time tomography is taken as the initial velocity model for waveform tomography.

Figure 4 shows that the channel feature designed into the synthetic model on purpose is reconstructed by the travel time tomography in a broad sense. The velocity model from waveform tomography reconstructs the channel much better, i.e., the geometry of the paleo-channel can be better identified. Furthermore, even the small scale features are also reconstructed to some extent. Artifacts exist in the reconstructed image by waveform tomography due to (1) only frequency components up to 90Hz are inverted; (2) noises are added in the data. The idea behind this study is that the geometry of the buried-channel could be hopefully better reconstructed by taking advantages of the superior resolving power of waveform tomography.
Figure 4. A synthetic velocity model (top), a velocity model from travel time tomography (middle) and waveform tomography (bottom).
5.6 Waveform tomography

(1) Time windowing. As stated in the VSP data application, the main issue in windowing time domain data is to balance the trade-off between the number of traces included and the number of samples (i.e., time duration) in the time window. A 20ms time window following first arrivals is used in this study as a result of this balance. This time window was preceded and followed by a 5ms ramp to reduce window artifacts.

(2) Source signatures for individual shot or shot groups. Source signatures are initially determined by using the travel time velocity model for each line. The frequency components used in the source signature determination are from 5.0Hz to 180.0Hz with 5.0Hz interval. If some of the shot gathers are very noisy, then shot groups are formed and one source signature is generated for each source group. Once the source signatures are determined initially, they are used in subsequent velocity inversions and only amplitude of these source signatures are updated during the inversion. The updated velocity models then are used to determine the final source signatures which may be used again for the next stage of velocity inversion. The process can be repeated more than once if necessary until good matches are achieved between synthetic and observed seismograms.

Figure 5 shows all the final source signatures for each line. We can see that the source signatures are very much consistent. However, some of the source signatures have more details than others possibly due to different noise levels in different shot gathers and different local conditions.
Figure 5. Source signatures for sources or source groups in each of the 45 lines.
Continuation of Figure 5.
Continuation of Figure 5.
Source index

Continuation of Figure 5.
Continuation of Figure 5.
Continuation of Figure 5.
Continuation of Figure 5.
(3) Waveform inversion. The first step in waveform inversion is to select which frequency components to be inverted. So frequency components are checked for consistency in the same way shown in the VSP data application. For the surface reflection dataset, it is found that the selected waveforms have very much consistent spectrum between 30.0Hz to 90.0Hz. Therefore frequency components 30.0, 40.0, 50.0, 60.0, 75.0 and 90.0Hz are chosen to be inverted successively. A frequency component as low as 30.0Hz is inverted first to ensure no cycle skip problems. The inversion process may go through a few stages. In each stage, the frequency components from 30.0 to
90.0Hz are inverted individually for 10 loops. After each stage, the source signatures are inverted using the final velocity model in the previous stage and this pattern can be repeated more than once if necessary. Usually after the first stage, synthetic seismograms are computed and compared with the observed seismograms to see if there is cycle skip problems and if the source signatures have a proper shapes that the predicted waveforms are similar to the observed ones.

The final velocity models from waveform tomography are described in their comparisons with those from travel time tomography. Figure 6 shows all the velocity models from travel time tomography and waveform tomography for each of the 45 seismic profiles. The difference between the two models for each line lies in (1) that the model from waveform tomography has more details than the travel time model because of the difference in resolutions in the two techniques; (2) the model from waveform tomography has better constraints on the image of the paleo-channel, especially in the deepest part of it. For example, the travel time velocity model for Line 45 shows that the paleo-channel extends to ~7.0m whereas we can see clearly the valley extends to 9.0m in the waveform tomography model and the boundary of the channel is also better defined by waveform tomography even in the shallow part, i.e., the flanks of the channel. This is mainly because travel time tomography assumes that the waves only sample the geometry of the seismic ray paths which is reasonable under high-frequency approximation, whereas the waves sample actually areas centered at the ray paths (say, with radii at a wavelength at least) which is modeled in the waveform tomography. This is the reason why waveform tomography is applied to the dataset.
Figure 6. Travel time (left) and waveform tomography (right) models for each of the 45 2D seismic lines.
Continuation of Figure 6.
Continuation of Figure 6.
Continuation of Figure 6.
Continuation of Figure 6.
Continuation of Figure 6.
Continuation of Figure 6.
Continuation of Figure 6.

5.6.1 Correlation of velocity models from the VSP and surface reflection datasets

Since 11 2D seismic profiles intersect with the VSP target area, velocity models from the waveform tomography applied to the two datasets can be compared. Comparison between two models from two independent inversions using totally different datasets could serve as a blind test to see what structures in the models are true features if good correlation is found. Figure 7 shows the comparison. We can see that Lines 20, 21, 22, 23, 25, 26 and 27 correlate reasonably well with the VSP velocity model. Especially the good correlation between Lines 21, 27 and the VSP model indicates that the low velocity
Figure 7. 10 velocity models from the waveform tomography applied to Lines 19 through 28 are correlated with the VSP velocity model.
channel near the south borehole and the thrusting low velocity feature near the northern borehole in the VSP velocity model are true. The low velocity feature in the middle bottom part of the VSP model is not clearly reflected in Line 24. Among other reasons, this discrepancy may be due to (1) depth resolution is not enough in the surface reflection data for a small low velocity feature at 12.0m depth; (2) this feature is also located at the part of lowest resolution in the VSP model because of the source-receiver geometry. However, the low velocity feature is correlated well with the depth migrated image (Figure 18 in Chapter 4), indicating the feature is most likely real and the reason in (1) is responsible for the discrepancy.

5.6.2 Waveform inversion results

Figure 8 shows all the waveform inversion results for 45 seismic profiles. For each profile, a shot gather from field data and a synthetic shot gather calculated using the final velocity model and the final source signature are shown. Red dots in each panel show the first arrival times hand-picked on the field data. Traces which are not used in waveform tomography are shown in grey. Those traces which are not used are either near-offset traces or contaminated by noises such as ground rolls or pumping noises. Generally we can see most of the first arrival wave phenomena are well predicted by synthetic seismograms. However, some local phenomena due to local receiver and source conditions in the observed wave fields are not well modeled by waveform tomography because those local conditions are not taken into account in this application. The conditions have second-order effects on the inversion and are neglected here.
Figure 8. Field observed (upper panel) and synthetic seismograms (lower panel) from waveform tomography for each of the 45 2D seismic lines. Red dots are the first arrival times made on the field seismograms.
Continuation of Figure 8.
Continuation of Figure 8.
Continuation of Figure 8.
5.6.3 Comparison with depth migrated images

This section shows the correlation between the velocity models from waveform tomography and depth migrated images by Dana (2003). Figure 9 shows the velocity models from Lines 18 and 26, superimposed by depth migrated images. Although there are noises in the depth migrated images, the main events in the migrated images correlate well with the velocity structures reconstructed in the waveform tomography velocity model. The migrated events are associated with the velocity increase across the channel boundaries. The geometry of the channel can be identified both in the velocity models and the depth migrated images. Given the fact that the velocity used in the depth migration of Dana (2003) is absolutely independent from the velocity models obtained from waveform tomography, this good correlation indicates that the identified geometry of the channel is reliable.
5.7 3D geometry of the Paleo-channel

In the velocity model of each 2D seismic profiles, the cross-sectional geometry of the paleo-channel can be identified, following the 800.0m/s velocity contours which is the seismic velocity for clay above water table. By combining all 45 cross-sectional geometries identified, the 3D geometry of the paleo-channel can be reconstructed slice by slice.
slice, following the similar philosophy behind any tomographic imaging. Figure 10 shows the reconstructed 3D geometry of the channel.

The reconstructed 3D geometry shows it varies both along strike and cross-strike. Its west-side is generally steeper than the east side. Along strike, the channel is wider in the northern end than its southern end. These two features agree with the 3D depth migration results of Dana (2003). In terms of velocity, the average velocity of the material filling the channel is about 200–300.0 m/s lower than that below the channel. Velocity in the shallowest a few meters in the southern most 20m along the strike is lower than most other parts seen in the channel, corresponding to soft and unconsolidated soils in the southern part of the experiment site.

![3D geometry of a paleo-channel](image)

Figure 10. 3D geometry of a paleo-channel reconstructed by combining 45 parallel cross-sectional geometry of it identified from 45 slices of waveform tomography velocity model.

5.8 Discussions and Conclusions

This study applies waveform tomography to 45 2D seismic profiles to reconstruct the 3D geometry of a paleo-channel at a ground water contamination site in the Hill Air
Force Base (HAFB). Since DNAPLs are heavier than ground water, they presumably reside on the deepest points of the channel. It is important to determine the 3D geometry of the channel accurately in a ground water remediation program because it is the structural host for the polluted ground water and DNAPLs themselves.

This study for the first time applies waveform tomography to solve an environmental seismic imaging problem. It shows that waveform tomography reconstructed the structural host for ground water with relatively high resolution, using the travel time velocity models as starting models. The derived 3D geometry of the channel is generally consistent with that derived from 3D depth migration (Dana, 2003). The 3D geometry derived this way is cheaper than that derived by well logs. While well logs could only determine the depths to the clay tops point-wisely, waveform tomography could determine the depths in profiles.

Travel time and waveform tomography could be complementary to each other, which is further testified in this study. Travel time tomography reconstructs an image which represents the large wavelength features of the target area. Waveform tomography refines the image and features such the cross-sectional geometries of the channel are more focused and easier to be identified.

This study solves the problem of how to image complicated shallow structures in relative high resolution by waveform tomography. Knowledge accumulated in this study paves the way for applications of waveform tomography to the shallow seismic imaging problems associated with waste detection, ground water contamination/monitoring, etc.
Chapter 6

Discussion and Conclusion

This thesis addresses and solves the unique problem of how to image complicated shallow substructures with high resolution. This problem is significant wherever shallow seismic characterization is applied. This thesis further develops the form of waveform tomography of Pratt et al. (1998) and proposes its application to solve the problem. To justify the proposal, it makes a comprehensive comparison between the conventional approach of travel time tomography and waveform tomography. Then it applies waveform tomography to two real datasets from a ground water contamination site in the Hill Air Force Base (HAFB) to sample lateral formation heterogeneities and reconstruct the 3D geometry of the structural host of polluted ground water. In summary, this thesis touches the following aspects of waveform tomography: theoretical background, computer engineering issues, synthetic as well as real data applications. If resolution is the main goal in any seismic imaging studies, waveform inversion related techniques such as waveform tomography could be the solution because they provide ultimate high resolution in seismic imaging.

Chapter 2 shows the theoretical background of waveform tomography, following Pratt et al. (1998). It shows the methodological advantage of being formulated in frequency-space domain over those techniques formulated in time-space domain. Forward modeling in the former is more efficient than the latter if large number of sources involve. The inversion is more convenient in frequency domain since the correlation of the back-propagated residual wave fields and the forward modeled wave fields needed to compute
the gradient vector is trivial in frequency domain. Furthermore, frequency domain formulation brings the flexibility (1) to invert selected frequency components successively from low to high to mitigate the non-linearity of waveform inversion; (2) to invert selected frequency components with higher signal-to-noise ratio.

Chapter 3 shows the technical features and performances of waveform tomography in the context of its comparison with travel time tomography. Although the synthetic study utilizes the cross-well source-receiver geometry to better reconstruct images using transmitted waves, the conclusions of the study are not specific to that geometry and hold true for other source-receiver geometries as well. The synthetic model is randomly generated with scale features of different sizes. It is designed so because (1) the real crust and upper mantle are believed to have a random yet self-similar pattern in velocity structure, although the real scale features may not have the same sizes as shown in the synthetic model. (2) The model is ideal to test the different resolving powers of the two tomographic approaches. It is found in Chapter 3 that the waveform tomography resolves in a resolution higher than the travel time tomography does at the price of higher computing cost and being less robust. The computing cost for waveform tomography is higher because to forward model waveforms requires longer CPU time and more computing resources such as memory than to compute travel times. The algorithm in waveform tomography is less robust than that in travel time tomography because of the extreme sensitivity of waveform perturbation to velocity perturbations while travel times do not change dramatically due to same velocity perturbations. In another word, waveform tomography is far more likely to fail than travel time tomography. As a consequence, waveform tomography requires the initial velocity model close enough to
the true model in order to achieve meaningful convergence, especially when the gradient method is used in inversion.

Chapter 3 further shows that the two tomographic approaches could be complimentary to each other. Real seismograms usually do not contain much significant low spectra. Chapter 3 shows that a velocity model from travel time tomography provides a good starting velocity model essential for a successful waveform tomography. Real data usually contain noises which make the whole data inconsistent to some extent. As concluded by Pratt et al. (2001), even the combination of travel time and waveform tomography is prone to many potential difficulties, such as source estimations.

Chapter 4 shows how waveform tomography can be applied to solve a seismic imaging problem at a ground water contamination site in the Hill Air Force Base (HAFB). The object of the VSP experiment is to sample the lateral formation heterogeneities in a buried paleo-channel where DNAPLs were dumped. The survey scope is 21.0m wide and 15.0m deep, which is only ~5.0 times of the wavelength. The problem lies in that the media in the VSP target area is very heterogeneous while the resolution of commonly applied ray-based imaging tools such as travel time tomography is not adequate to resolve the details. The resolution of travel time tomography is around ~7.0m which is nearly one-third of the horizontal target dimension. Furthermore, heterogeneous media give rise to complicated wave phenomena such as diffracted/scattered waves. All these make the wave equation based approaches such as waveform tomography advantageous in shallow seismic imaging, though it is computing-intensive. The real data application in Chapter 4 proves that waveform tomography can solve the problem and play a unique role in high resolution shallow seismic imaging.
The high resolution velocity model from waveform tomography reveals details that could not be seen in the travel time tomography. It shows (1) a low velocity layer at 1-3 m depths. Lower velocity in the layer results from a thin layer of more consolidated desert hardpan above. Because of this low velocity layer, ground rolls in the data become dominating in terms of seismic energy. (2) Large lateral variations. Velocity contrast as large as $\sim 200.0 \text{m/s}$ can occur in $\sim 1.0 \text{m}$ range. The implications of large lateral variation will be investigated and discussed below. (3) A large vertical gradient at $\sim 80.0 \text{m/s/m}$ on average. (4) Small scale features down to $\sim 1.5 \text{m}$. Furthermore, the velocity model correlates well with two lithology logs available and can be interpreted geologically. Chapter 4 also shows it could be somehow interpreted petrologically.

The significance of high resolution velocity estimation by waveform tomography also lies in the fact that seismic velocity distribution in the shallowest part of the earth can not be easily measured in a rock physics lab because of the coupling problem between a transducer and unconsolidated materials under low pressure. Waveform tomography provides an alternative to estimate accurately seismic velocity of materials in shallow environment.

To show the remarkable velocity variation both vertically and laterally and its possible implications for other imaging approaches assuming small lateral velocity variations, the differences in travel times and wave fronts between the velocity model by waveform tomography and a 1D velocity model averaged from it are investigated (Figure 1). Figure 1 shows travel times and wave fronts in the two models from source locations 1 and 16, respectively. We can see that wave fronts in the waveform tomography velocity model
are much more complicated than those seen in the 1D average model. The differences in travel times ranges from -5.5ms to 6.5ms, a variation of 12.0ms which is almost equal to one dominant period (12.5ms). For ray-based imaging approaches such as Kirchoff migration, this could possibly lead to severe uncertainties.

By waveform tomography, source signatures can be determined for an individual shot (or shot groups). The determined source signatures are band-limited representation of the true source signatures. These source signatures can be used for other applications such as deconvolution. These source signatures are very much consistent, reflecting that the shot rifle made repeatable source signatures.

A constant Q=20 is determined as a first order approximation for the real attenuation distribution in the target area. This low Q value indicates that the target area attenuates seismic waves very significantly. The Q value determined results from all attenuating effects such as poroelastic wave propagation, anelastic absorption and scattering. It is impossible to distinguish these effects by our method. In their attenuation study by visco-acoustic modeling and lab experiments, Keers et al. (2001) shows that the Q values for water saturated sands mixing with non-aqueous phase liquid (NAPL, similar to DNAPLs in this thesis) range from 10~20, depending on different concentration of NAPL. The determined Q=20 is within that range. However, Q values are not constant in the whole target area, as is reflected by the misfit in the AVO inversion and different waveform modulation at different source-receiver offsets.

If it is challenging to image the heterogeneities in the buried channel, it is even so to image the 3D geometry of the channel itself lying beneath all the heterogeneities. Since the channel is the structural host for the polluted ground water, it is important to delineate
Figure 1. Comparison of wavefronts (upper 4 plots) in the VSP velocity model and a 1D velocity model averaged from it, and relative travel time differences between the two models. All plots on the left column are from Shot 1 and those on the right are from Shot 16.
its extension in space accurately because it is the key to a ground water remediation program. Chapter 5 shows that waveform tomography could re-focus the image of the paleo-channel initially reconstructed by travel time tomography.

By applying waveform tomography to 45 2D seismic profiles, 45 high resolution velocity models are reconstructed. By combining the 2D cross-sectional geometry identified at each velocity model, the 3D geometry of the channel is reconstructed. However, the underlying assumption is that out-of-plane seismic energy is minimum. Since mostly refraction waveforms are used in the waveform tomography, the assumption is approximately satisfied as pointed by Lanz et al. (1998).

The complementary nature of waveform and travel time tomography is further manifested by the applications of waveform tomography in this thesis. Starting from travel time tomography is one of the steps taken to harness the highly non-linear inversions into a meaningful convergence. This is important in real data application because the inversion is also sensitive to noises mixing with data.

The reconstructed 3D geometry by waveform tomography is cheaper than otherwise from well logs. The subsurface map can be used for engineering companies to decide where to place extraction/injection wells. For the first time, waveform tomography is applied at a ground water contamination site. Knowledge gained in this novel effort may pave the way for the wide applications of waveform tomography to other environmental problems and other seismic imaging problems from local to regional scopes.
References


Appendix I

Waveform Tomography : Fortran90 Version

This appendix summarizes the computer engineering aspects of the modification of the waveform tomography package of Pratt et al. (1998) from the Fortran77 into Fortran90 version. The object of the modification is to optimize the code such that it can handle regional size models in greater running efficiency.

Introduction

The original tomographic package of Pratt et al. (1998) is written in Fortran77 and has been tested extensively in academic environment. All arrays in the code are declared statically and are subject to stacking limit. The biggest problem to solve is that static arrays can not be declared large enough for model sizes in regional scale. For example, it is common for a 2-D marine profile to cover a target area 20.0km wide and 10.0km deep. Depending on what frequency components are modeled, such a survey scope may require a model with 1 million model parameters, i.e., 1 million grids in 2-D finite difference modeling. Since some of the arrays are allocated 16~50 times as large as the total number of grids, such a model results easily in memory shortage in most conventional machines.

The next problem is how to optimize usage of memory. Memory usage optimization was never an issue when the code was applied academically to relatively small models, but it may be a big issue for several reasons when the code is applied to large size models. First of all hundreds of arrays are declared in the whole package. Many of the arrays have sizes 16 to 50 times as large as the number of model parameters. When models are large, the code requires huge memory space, yet these arrays do not have to be allocated at the same time and some of them do not necessarily have to be allocated at all. Secondly, one of the arrays is declared to hold all the digital samples from all shot gathers in a survey. When the code is applied to a dataset several GBytes in size, it is easy for the array to have size larger than $2^{31}$, which is the maximum array size possibly allocated for a 32-bit code, such as this tomographic package. Without the optimization, sometimes only part of a large dataset can be used and application scope of the code is
therefore limited. How to maintain an acceptable running speed when it is applied to large datasets is another problem.

**Code Modification (Method)**

Before the problems mentioned above were attacked, the code was extensively tested and debugged in SGI machines because it was never before in SGI/IRIX system. Three compiler-dependent bugs were found and fixed. This debugging made the package more versatile and robust.

To solve the stacking limit problem, all arrays with sizes no smaller than the number of model parameters are modified into pointers or allocatable arrays, which are dynamically allocated. The advantage of arrays dynamically allocated is that (1) they are not subject to stacking limit; (2) the memory they hold can be released by de-allocation when they are not in usage any more. To use pointers as dummy variables, interfaces were added in code to facilitate the calling between main and subroutines or between subroutines themselves. The utilization of pointers, allocatable arrays and interfaces upgrades the code into Fortran90 standard. This upgrading involves 188 arrays in 32 of the total 125 Fortran programs and six header files. 64 interfaces were added in 27 Fortran programs. The largest model that is tested so far by using the Fortran90 version has 1 million model parameters (grids). If only up to 10Hz of marine data is inverted, a one-million-grid model can cover an area 70km wide and 17km deep. However, with a model of such size, running speed and memory allocation are big issues. One way to optimize memory usage is that some of the arrays can be allocated conditionally. Arrays are allocated only when it is necessary, taking advantage of the dynamic allocation features in Fortran90. Otherwise they are declared but not associated with any physical memory space. For this tomography package, conditional allocation implies that some of the arrays are allocated only when certain features of the package are used. For example, if there is no time domain output during forward modeling or inversion, as is a common practice, one of the largest arrays will not be allocated at all. Most of the local arrays in the main program FULLWVND.F are allocated conditionally after the modification. But more work can be done in this respect in the future.
To improve running efficiency, the package is modified such that for each source, all
do-loops over the number of total receivers are only executed over the receiver range
corresponding to that source. This modification greatly improves running efficiency
especially when both sources and receiver arrays are moving in a survey. For a marine
survey of total 91 sources and 285 receivers for each source, it takes ~6 hours and 45min
to finish one inversion loop before and after the modification. This modification is the
first step toward a version of the package that is tailored specifically for marine survey
geometry.

References

methods in frequency domain seismic waveform inversion. Geophysical Journal